Lifelog Access Modelling using MemoryMesh

Lijuan Zhou

Master of Science in Computer Science (Hons.)

A dissertation submitted in fulfilment of the

requirements for the award of

Doctor of Philosophy (Ph.D.)

to the

Dublin City University

Faculty of Engineering and Computing, School of Computing

Supervisor: Dr. Cathal Gurrin

January, 2016

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _______________ (Lijuan Zhou)

Student ID: 11212053

Date: _______________

Table of Contents

List of Tables

List of Figures

Nomenclature

ADL Activities of Daily Living, page 86

AmI Ambient Intelligence, page 6

ANNs Artificial Neural Networks, page 90

BLPF Butterworth Low-Pass Filter, page 94

BoF Bag of Features, page 119

CH Colour Histogram, page 120

CI Confidence Interval, page 78

ES Event Segmentation, page 101

EST Event Segmentation Theory, page 116

GPS Global Positioning System, page 12

GSR Galvanic Skin Response, page 91

HITS Hyper-link-Induced Topic Search, page 145

ICMR ACM International Conference on Multimedia Retrieval, page 86

ICR Image Category Recognition, page 121

IDF Inverse Document Frequency, page 145

IR Information Retrieval, page 2

JSON JavaScript Object Notation, page 52

LES Lifelog Event Segmentation, page 12

ME Metabolic Equivalent, page 100

MET Metabolic Equivalent of Task, page 100

ML Machine Learning, page 97

NDCG Normalized Discounted Cumulative Gain, page 77

OD Object Detection, page 122

OpenCV Open Source Computer Vision Library, page 125

PAR Physical Activity Recognition, page 11

PAR Physical Activity Recognition, page 87

PC Pervasive Computing, page 5

PDA Physical Daily Activity, page 11

PIR Passive Infra-red Light, page 48

PL Personal Lifestyle, page 99

QS Quantified Self, page 3

RANSAC RANdom SAmple Consensus, page 122

RF Random Forest, page 109

SIFT Scale-Invariant Feature Transform, page 120

SURF Speeded Up Robust Features, page 121

TF Term Frequency, page 145

UC Ubiquitous Computing, page 6

WWW World Wide Web, page 5

SVM Support Vector Model, page 95

Abstract

Lijuan Zhou

Lifelog Access Modelling using MemoryMesh

As of very recently, we have observed a convergence of technologies that have led to the emergence of lifelogging as a technology for personal data application. Lifelogging will become ubiquitous in the near future, not just for memory enhancement and health management, but also in various other domains. While there are many devices available for gathering massive lifelogging data, there are still challenges to modelling large volume of multi-modal lifelog data. In the thesis, we explore and address the problem of how to model lifelog in order to make personal lifelogs more accessible to users from the perspective of collection, organization and visualization. In order to subdivide our research targets, we designed and followed the following steps to solve the problem:

1. Lifelog activity recognition. We use multiple sensor data to analyse various daily life activities. Data ranges from accelerometer data collected by mobile phones to images captured by wearable cameras. We propose a semantic, density-based algorithm to cope with concept selection issues for lifelogging sensory data.

2. Visual discovery of lifelog images. Most of the lifelog information we take everyday is in a form of images, so images contain significant information about our lives. Here we conduct some experiments on visual content analysis of lifelog images, which includes both image contents and image meta data.

3. Linkage analysis of lifelogs. By exploring linkage analysis of lifelog data, we can connect all lifelog images using linkage models into a concept called the MemoryMesh.

The thesis includes experimental evaluations using real-life data collected from multiple users and shows the performance of our algorithms in detecting semantics of daily-life concepts and their effectiveness in activity recognition and lifelog retrieval.

Acknowledgements

I here to thank my supervisor Dr. Cathal Gurrin for his valuable guidance not only on research topics but also on research methodologies. His profound knowledge and innovative thoughts about lifelogging is nothing but impressive and enlightening the PhD life from 2011-2015.

I would also thank Prof. Alan Smeaton, Prof. Noel O'Connor and Margaret Malone who helped a lot in various aspects in my daily research.

There is no words that I can say to express my love and appreciation to my parents for bringing me up and giving me the best education ever that I could have by being paradigmatic parents.

I also appreciate the help for my experiments from the supportive and insightful group of our research lab. Time can pass, but good memory of working together and having fun on Friday nights never fades away.

I would like to thank my external and internal examiner for their valuable feedback, Prof. Joe Carthy and Dr. Claus Pahl.

Thanks to the main financial support from SFI for the research in this thesis. The work of this PhD research has emanated from research conducted with the financial support of Science Foundation Ireland under grant number ”SFI 11/RFP.1/CMS/3283”. Special thanks to IBM Brazil and University of Oxford for funding my research trips during my PhD. Their expertise and practical skills in lifelogging and public health have been of great help for carrying out my research. Also, great appreciation goes to Active Living Research Center and Dr. Jacqueline Kerr for supporting me conducting research over in UCSD.

Thank you everyone for steering me in the right direction.

Chapter 1

Introduction

1.1 Motivation

As sensor technology is becoming ubiquitous, more people are beginning to utilise inexpensive sensors for recording aspects of their lives and activities. Sensor technology has become prevalent in recent years as the cost of wireless communication decreases and a new generation of digital devices is becoming available to everyone. Meanwhile sensors as data collection tools are also attracting ever-increasing research interest in both academia and industry. Sensors can be applied to many domains such as personal health monitoring [118, 21, 126], military applications [106], home activity detection and security surveillance [120], and social group interaction [63].

Such increasing ubiquity of low-cost sensors allows for the creation of digital lifelogs [74], or detailed digital footprints of our on-line and real-world activities. A lifelog is a digital archive of people's life experience in a format of images, biometric readings, computer interaction records, emails etc. Since lifelogging is a method by which people may chronicle their existence digitally, a lifelog can be said to “represent a comprehensive ‘black box’ of human life activities and may offer the potential to mine or infer knowledge about how we live our lives” [74]. Indeed, for many individuals, a lifelog already exists in the form of documents, photos, videos, location logs, purchase histories, electronic health records and so on. In this research, the focus is on the type of lifelogs that are made possible by the new generation of personal sensing devices. This large volume of data gathered by the sensor devices brings big potential for new user centric applications while at the same time, bringing challenges in modelling large amounts of multi-modal lifelog data from the sensors.

In this work, the focus is on how to organize large-volume, rich-format lifelog data. For some types of lifelog data management is not a big challenge due to data homogeneity. For example, some people log their footsteps using a pedometer, while others may log their movements using FitBit or similar activity loggers. However, recent advances in customer technologies (e.g. Narrative Clip, the Autographer and even Google Glass) are bringing the challenges of managing large volumes of lifelog data to the fore, and these challenges provide the motivation of this research: to explore the potential for new lifelog application technologies that make the lifelog manageable, useful and more than just a big-data archive. These new developments of logging and communication technologies heralds a new era for lifelogging, by expanding this research from human memory reminiscence to human health management, self surveillance, entertainment and more. In recent years we have become used to fast and effective Information Retrieval, but lifelogging poses new challenges. This thesis explores the first retrieval models for accessing lifelog data.

According to Gurrin et. al. “A lifelog is the actual data gathered. It could reside on a personal hard drive, in the cloud or in some portable storage device. The lifelog could be as simple as a collection of photos, or could become as large and complex as a lifetime of wearable sensory output (for example, GPS location logs or accelerometer activity traces)" [74] and "Lifelogging is the process of passively gathering, processing, and reflecting on life experience data collected by a variety of sensors, and is carried out by an individual, the lifelogger. The life experience data is mostly based on wearable sensors which directly sense activities of the person, though sometimes data from environmental sensors or other informational sensors can be incorporated into the process" [74]. In this work, a lifelog is defined as a long-term archive of consecutively collected passive sensor data and image data that can be chronologically organized using digital data processing techniques and applicable in people's life in different scenarios including searching and browsing. It is within this understanding of the concept of a lifelog that this research is framed.

Lifelogs can be captured in many ways with a wide range of sensors, most specifically wearable sensors. Wearable sensors include FitBit, BodyMedia, GPS locator, Looxie, SenseCam, mobile phones and biomedical wearable textile etc. All these devices act as sensors that can record one or more aspects of our life experience both individually and socially. Since there exists a variety of such sensors, we can consider that the task of gathering lifelog data is sufficiently supported by existing technologies. Nevertheless, the challenging task for enhancing human memories using lifelogging is not to simply gather data, but to mine it, organise it and generally to support the potential for positive captology [131].

There are many international academic and industrial conferences and activities that are concerned with this new research area, including the SenseCam Conference, Pervasive Health, UbiComp, Quantified Self and so on. All these conferences and meet-ups are attracting increasing levels of attention from the research community and society as a whole. Aside from academics, *Quantified Self — Self Knowledge*

Through Numbers is a collaboration between users and tool makers who share an interest in self knowledge through self-tracking [149]. QSers exchange information about their personal projects, the tools they use, tips they have gleaned, and lessons they have learned. QSers blog, meet face to face, and collaborate on-line. This conference has also branched in many cities around the world like Dublin, London, Sydney, Beijing, Silicon Valley, Portland, Berlin, Athens, Guangzhou, Seattle, Davis, Milan, Louisville, Rotterdam and so on. Most of the research team here at Dublin City University that supported me during this research are active participants in the Dublin QS group and we attend the regular meet-ups, where we get to understand the real-world lifelogging needs of other participants. There are also many companies active in the space. Most recently, Narrative [1], a Swedish start-up has had major success on kickstarter[2] with their lifelogging camera up to the date when this thesis is being written. Arduino, which provides open-source electronic prototyping platforms, allow to create interactive electronic objects. One award-winning development recently has been a lifelogging camera built on the Arduino platform.

Although lifelogging is beginning to 'take off', the potential is not yet completely realised. Sellen and Whittaker have criticised a perceived lack of direction in lifelog research and have gone so far as defining an initial five *R*s of memory access [142] as potential benefits for human memory of engaging in lifelogging. The five *R*s are *R*ecollecting, *R*eminiscing, *R*etrieving, *R*eflecting and *R*emembering intentions. Each of the five *R*s defines a different reason why people access their memories or their personal life archives. Specifically in this research, we are trying to develop new approaches to lifelog data organisation that support technology mediated *R*ecollecting and *R*eminiscing of past life experience. *R*ecollecting is con-

[1] http://memoto.com/

[2] https://www.kickstarter.com/

cerned with reliving past experiences for various reasons. For example, we may want to recall who was at an event, or where we parked the car. *R*eminiscing, which is a form of recollecting, is about reliving past experiences for emotional or sentimental reasons. In this research, we are applying Internet web information retrieval technologies into the domain of lifelogging to explore the potential of exploiting links between lifelog data to enhance the effectiveness of lifelog data access.

In addition, Bell and Gemmel stated in their book *Total Recall* [15], that the era of lifelogs have "the potential to revolutionise our healthcare, productivity and social lives, it will change what it means to be human." There are many different potential application areas from personalised healthcare to enhanced learning; however, we have only begun to address the research challenges. This is like the early days of WWW search; back then there were many millions of WWW pages, but no suitable search engine. The "Google" of lifelogging is still non-existent. Indeed, for this research, the inspiration is drawn from the early days of the WWW; when novel algorithms catapulted a small digital library project in Stanford from 'backrub' built on Lego hardware literally to become the search giant Google that we know today. However this thesis is not trying to be the "Google" of lifelogging, but focuses on some pioneering research on effective means of lifelog retrieval. The research contained in this thesis is therefore to be viewed as a first-step in a long process of developing lifelog data organisation and retrieval systems and not as a solved problem. This thesis therefore investigates the application of novel search and managing technologies, based on WWW linkage algorithms, and explores the possibilities to apply them to the new domain of lifelogging. We divide the process of research based on the state-of-the-art lifelog and information retrieval technologies and research outputs.

Pervasive computing researchers focus on lifelog research using pervasive technology to collect data to enhance lifelog data modelling and data visualization. While pervasive computing is a new paradigm, it is also described as *Ubiquitous Computing*, *Ambient Intelligence*, or *'everyware'*. Although these terms are all slightly different, they all indicate the trend of how the supporting technologies and knowledge required to supporting lifelogging is receiving increasing attention.

As with any research involving user data, one of the big concerns with lifelogging lies in the privacy issues. Privacy issues like many other issues within the research scope of lifelogging, has triggered a significant amount of research discussion. O'Hara et. al. proposed that lifelog data acquisition, storage and retrieval can be undiscriminating and therefore possibly raises issues about privacy, identity and empowerment-related issues[123]. Other researchers also proposes different solutions to privacy issues. Ye et. al. proposed to use Face Blurring to filter out face information in all lifelog images [171]. Gedik et. al. proposed some architecture and algorithms to prevent lifelog location privacy from leakage [69]. One of the thorough research on lifelog privacy issues is by Jacquemard in his Ph.D. thesis about ethics of lifelog technology [82]. While in this thesis, we do not intent to elucidate whether lifelog privacy issues are positive or negative. This research concentrates on exploration of the potential of searching through lifelogs and leaves others to consider the privacy, social and ethical issues inherent in lifelogging.

Before moving onto the main body of research presented in this thesis, we go through the significance of this research and propose our research questions based on our hypotheses.

1.2 Significance of This Research

Internet of Things (IoT) [163, 169] is a scenario in which objects, animals or people are provided with sensors and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction [169]. IoT is a promising technology due to the benefits it offers in a modern world of complexity. It has evolved from the convergence of wireless technologies, micro-electromechanical systems and the Internet [169]. IoT technology is getting trendy in solving problems like in-home healthcare services [125], collaborative warehousing environment [136], product delivery service [86]. Also how to implement IoT has received significant research attention of late [110, 156]. We notice a trend in which ubiquitous computing is becoming a part of everyday life.

Lifelogging is one branch of pervasive computing that is accelerating in use and application recently. With the recent availability of wearable sensing technologies and an acceptance of personal data gathering and on-line social sharing (e.g, on Facebook timeline), lifelogging has become a mainstream research topic. We now have the ability to gather and store large volumes of personal data using an inexpensive wearable devices and smart phones. Despite massive availability of lifelogging tools, how to collect, organize and represent lifelog data is still under much discussion [31, 174]. New technologies especially various sensors in data capture, storage and computing will undoubtedly bring about a revolution in the way how humans interact with technology in the forthcoming years. One will be able to record as much of life experience as people wish, in previously unimaginable detail. Everything we see, hear, food we eat, fun we have, health of our body, all can be captured digitally. This is likely to revolutionise many aspects of our lives, for example, our healthcare, our learning and our productivity. As Bell and Gemmell state:“the coming world of Total Recall will be as dramatic a change ... as the digital age ... It will change the way we work and learn. It will unleash our creativity and improve our health. It will change our intimate relationships of loved ones both living and dead. It will, I believe, change what it means to be human ... I can look back over my activity logs and notice where I have spent too much time on low-priority projects, or took too little time at a key place, or burned up a surprising number of hours reading Internet news ... lifelog (if properly organised) will reveal the meaning of your life."[15]. This is the focus of this research and forms the basis for the significance of this research. We define new access models for lifelog applications based on applying techniques from WWW search. We also proved as shown in the conclusion of this research that the proposed new retrieval models are applicable and feasible in lifelog research and show promise as an avenue for further research and optimisation.

Three converging technology streams have brought us to the point at which we can consider a world of ubiquitous lifelogging, they are data capture, storage and organization technologies [15]. Firstly, data capture technologies, such as smart phones are already ubiquitous and we can already see a new generation of lifelogging devices coming to market, such as location tracking, personal and environmental sensing devices and more specifically, life activity specific recording devices such as SenseCam, Narrative or Google Glass.

Secondly, data storage technologies have progressed to such a point that it is now possible to capture 5,000 photos per day (using a Vicon Revue wearable camera - a refinement of the original SenseCam) and store 30 years worth on a 1 TB hard drive (at the low resolution of some wearable lifelogging devices, such as the Microsoft SenseCam, that will be discussed later). Looking into the future, one will be able to store all of the life experience in a continual stream of digital video while data storage technologies further improve in the coming years. The advent of cloud-based data storage technology also provides excellent opportunities for storing vast life archives of video, audio, photos and various sensor streams to build a rich lifelog.

The third and final technology stream is the data organisation and retrieval software, which is the focus of this thesis and is an enabling technology in this area. Initial organization and access tools for lifelog have been based on a browsing methodology. A browsing tool for lifelog [56] is only the very beginning of what is possible and when faced with large lifelog archives, a browsing tool has been shown to be inefficient to allow a user to either locate or gain benefit from their content [175]. It is our conjecture, and the basis of this research, that lifelogging needs a Googlisation phase that moves the lifelog access state-of-the-art away from browsing and towards search. This Googlisation phase should include more event recognition and association from a psychological perspective, and search life experience through an associative module, like associative human memory. The application of search for lifelogging provides a first approach to retrieve and recollect from lifelogs.

Given the multimedia nature of lifelog content, the natural next technological step is the generation of search engines that operate on manual textual annotations of lifelog content; it would allow for text based searching through the archive. However manual annotation is unlikely to be scalable. A lifelog will form a multimedia-rich, densely linked and constantly changing hypermedia archive[25], where the access methods are not simply text queries in a desktop environment. The queries can often be non-textual (e.g. context queries based on interactions with people or locations, or visible objects) statements of an immediate information need, or the user will access the content by following associative links (as one does with one's own content-addressed or associative memory). This research addresses these issues by integrating information retrieval knowledge with inputs from cognitive psychology to develop appropriate knowledge representation technologies based on the 5*R*s if lifelog access[142].

5 *R*s can be originated from Cognitive Science. In the Stanford Encyclopaedia of Philosophy: cognitive science is the interdisciplinary study of mind and intelligence, embracing philosophy, psychology, artificial intelligence, neuroscience, linguistics, and anthropology. Its intellectual origins are in the mid 1950s when researchers in several fields began to develop theories of mind based on complex representations and computational procedures. Its organizational origins are in the mid 1970s when the Cognitive Science Society was formed and the journal *Cognitive Science* began. Since then, more than seventy universities in North America, Europe, Asia, and Australia have established cognitive science programs, and many others have instituted courses in cognitive science.[3]

Even we have theory bases like 5 *R*s to support that lifelog research needs to combine computing science and psychological cognitive science, we have enough evidence to see the possibility of applying cognitive science to our lifelog research, which leads us to the main contribution of this research, the*MemoryMesh*, a rich multimedia archive of life experience data that is interlinked as human episodic memory is considered to be interlinked. The potential benefit in terms of enhanced access and retrieval of applying the WWW linkage algorithms over the linked MemoryMesh therefore is the basis of this research.

[3]http://plato.stanford.edu/entries/cognitive-science/

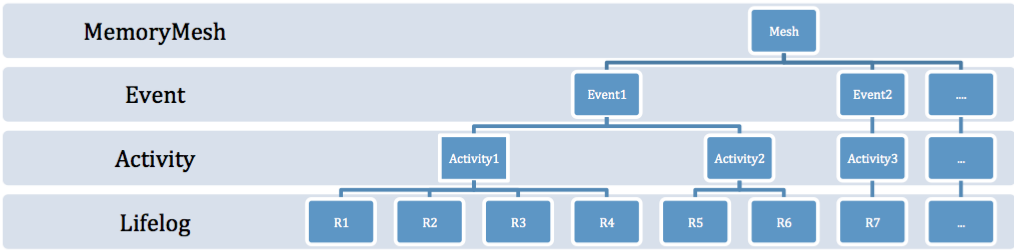


Figure 1.1: The vertical hierarchy of MemoryMesh structure

1.3 Hypothesis

In order to solve the problem of organising lifelog and making it retrievable for personal usage, we focus on three key challenges: *physical activity recognition*, *lifelog event segmentation*, and *lifelog retrieval*. Figure 1.1 shows the hierarchy of how these are related together. The bottom of this hierarchy is the raw and original lifelog data that is gathered from multiple devices. The second to bottom layer is the activities layer. This layer explains the process of extracting human activities from the raw original lifelog sensor data. The third layer is the event layer. This layer contains more contextual information and explains the process of making lifelog data more meaningful by segmentation into events and enhancement with contextual information. The top layer is the integration layer. With all events composing the MemoryMesh of life experience, a lifelogger would be able to access to lifelog via the MemoryMesh to provide enhanced search and retrieval. This suggests one initial solution to the problem of accessibility of lifelog for real-world users.

As stated, there are three key challenges in this research:

- Physical Activity Recognition

  Accelerometer data has been widely used for physical activity recognition using user-annotation through supervised machine learning techniques [11]. Ellis et. al. evaluated all modern popular supervised machine learning methods and proposed that random forest is the best for accelerometer and GPS data based physical activity recognition method [64]. However, only 5 of 15 activities are effectively recognized due to lack of sufficient features, with 34 features used in study. In this study, we present a supervised machine learning method of transportation mode prediction from GPS and accelerometer data using 64 features which were also applied in the previous work of Ellis [64]. It is computationally efficient when compared with Casale's 20 features [35] and Ellis's 34 features [64].

- Lifelog Event Segmentation

  For text based event detection or segmentation, Naughton and Carthy et. al proposed a sentence level event segmentation approach to classify events from unstructured text [119]. Event extraction from video has also been widely explored [116]. But there has been limited work that has been done in the area of lifelog event segmentation. The reasons for challenges or limitations lie in ethical privacy issues, data collection and data sharing and forthcoming data processing technologies. Doherty et. al. has developed an initial approach to conduct event segmentation in lifelog research [58, 52]. This approach is applied in the SenseCam browser [56] as shown in Figure 1.2. These provide the basic approaches to event segmentation that are enhanced in this work and described later in this thesis, and the construction of linkage based MemoryMesh.

- Lifelog Retrieval

  Information retrieval has been playing an essential role for over 50 years since

Figure 1.2: Touch screen of the DCU SenseCam browser by Doherty et. al.

information management research began to be explored in the 1950s [139]. In the defined research pathway of this research, as shown in Figure 1.1, a lifelog is composed of retrievable units called events. Like WWW, these units can be linked or associated as event episodes that can be linked as a model of associated memory analogically. Therefore, how we retrieve lifelog data is one of the main research topics in this thesis. There is plenty of research about concept based data indexing for information retrieval [77]. However, there is still limited work on lifelog retrieval. Aizawa et. al. proposed to use lifelog content and context for efficient retrieval [4]. But this method heavily relies on human input of context information. Jones et. al. proposed that lifelog access can be ad hoc search that allows ad hoc queries which users partially remember, either to inform themselves of some forgotten details or to share with others [87]. While in this thesis, we build linkage graphs for lifelog retrieval based on retrievable episodes like events and apply tailored algorithms to enhance retrieval performance.

The available suite of low-cost sensors today makes it possible to automatically capture life experience in rich detail and subsequently to identify events or moments from real-world activities and allocate importance factor to each event. While the prior work on event segmentation and life experience contributes to life event browsing [58] by stating that lifelog access can be divided into a three-layer structure of data organization, this thesis proposes that lifelog browsing model can be defined as a four-layer structure of lifelog browsing, as shown in Figure 1.1. The top of the structure is a linked graph based MemoryMesh. Figure 1.1 demonstrates the primitive representation of data hierarchy of the MemoryMesh system. The key of the methods of constructing MemoryMesh lies on activity recognition, event segmentation, event linkage analysis, MemoryMesh construction and retrieval.

So we make the following hypotheses for this research based on this research pathway:

1. Lifelogs can be captured by multiple sensors and segmented into different events. Event segmentation techniques in use today can be enhanced by the introduction of life activity detection into the event segmentation process (layer 2 and 3 in Figure 1.1). Activities are units of events while events are the units of personal lifelog, which means multi-layer hierarchical event segmentation is preferable to single layer ones, which were the focus of all previous lifelog event segmentation efforts.

2. Lifelogs can be constructed as a MemoryMesh (the top layer in Figure 1.1), of linked rich hypermedia, and algorithms that derive from web search, can help to search, organize and recommend events from lifelog archives. Events (as

in hypothesis 1) can be taken as units of lifelog for organization and linkage analysis based access.

1.4 Research Questions

One of the major differences between lifelog data and traditional multimedia data is that lifelog can have additional metadata collected by additional sensors. This makes lifelog data more explorable than traditional multimedia data. Based on the listed two hypotheses above and its attribute of meta data, we propose the following research questions:

1. Can we extract daily life activities from raw lifelog data?
2. If we can recognize activities from the raw lifelog data, can these activities facilitate human life event segmentation from lifelogs? If so, how efficient and accurate can this be?
3. If human lifelogs can be constructed as different units of daily events, in what way these daily events can be effectively linked together and constructed as human memory mesh? We define it as MemoryMesh in this research.
4. How can we support information access using the MemoryMesh?

In the thesis, we explore and address the problems of how to model lifelogs in order to make personal lifelogs more accessible to users from the perspective of collection, organization and visualization. In order to subdivide our research targets, we designed the following 3 steps:

1. *Lifelog activity recognition.* We use multiple sensor data to analyse various daily life activities. This includes data from accelerometer collected by smart

phones and images captured by wearable cameras. We propose a semantic, density-based algorithm to cope with concept selection issues. This will be described in Chapter 4 as lifelog sensor discovery.

2. *Visual discovery of lifelog images*. Most of the input information we take everyday is present to us as visual images. Images contain significant information about our lives. Here we conduct some experiments on visual content analysis of lifelog images, which includes both image content and image meta data. Image data provide rich information to lifelog event segmentation. This will be shown in Chapter 5 as lifelog visual discovery.

3. *Linkage analysis of lifelog*. By exploring linkage analysis of lifelog data, we can connect all lifelog images using linkage models into a concept called *MemoryMesh*. This will be presented in Chapter **??** as lifelog linkage analysis.

The thesis includes experimental evaluations after the experimental settings and results in each chapter. All experiments use real-life data from multiple users. The performance of our approaches to detecting semantics of daily-life concepts and their efficacy in activity recognition and lifelog retrieval is also included in each chapter.

1.5 Terminology

In order to describe the methods that we use to solve the research questions presented in the previous section, we elucidate some terminology here.

- **Lifelogging**

Lifelogging is the process of capturing, processing, and organizing and discovering lifelog data that are gathered by physical sensors and software both passively and subjectively, including self-reported information, indoors and outdoors [51, 177]. Although Bush's Memex refers to what is now lifelogging in 1940s, the first commonly accepted work in lifelogging was by Steve Mann in late 1980s and early 1990s when he started to wear an apparatus (wearable camera) over head. The device has been evolving ever since towards the appearance of ordinary eyeglasses.

- **Lifelogger**

  A lifelogger is an individual that conducts lifelogging using wearable sensors. In this thesis, we use lifeloggers to generally refer to people who collect data for lifelog.

- **Lifelog**

  Lifelog is the actual data that lifeloggers collect for lifelogging. It is actually existent data. The lifelog data can include sensory data like 3 dimensional accelerometer reading, orientation, GPS traces, bluetooth, heart rate capture etc. Multimedia data can be images, video and audio. All these provide lifelog hypermedia data analysis supports.

- **Lifelogging System**

  Different from lifelog software that focus on one component of lifelogging, a lifelogging system is a holistic software that covers all aspects of lifelogging, including data collection modules, data storage modules, data analysis and interactive user interfaces. These systems may be designed for different purposes. SenseCam, Autographer and their associated data processing systems

[56] are designed to assist human memory. FitBit is an example of health information lifelog that tracks human activities. It helps illustrate life activities using charts and tables to facilitate user to record and review their moving details[4]. SenseSeer is a lifelogging system using a mobile phone as a data collection tool and a set-up server for data storage and processing [133]. The digital library of lifelogging data and its supportive system is called *surrogate memory* [74]. We also developed a number of prototype lifelogging systems for our research purposes [175, 176, 178]. The details about implementation of these systems are mainly covered in the Chapter 3.

- **Lifelog Concepts**

  According to *The New Oxford Dictionary of English*, *Concept* is "an idea or thought what corresponds to some distinct entities or class of entities, or to its essential features, or determines the application of a term, and thus plays a part in the use of reason or language". While this definition of regular concept focuses on idea and thought of entities, *Lifelog Concepts* are more limited to concepts that are relevant to personal live experience. These concepts can be names of people, descriptive content of happenings and objects in lifeloggers' life experiences and so forth. Different lifeloggers have different lifelog concepts. These lifelog concepts construct patterns of people's life and the weights of lifelog concepts can imply the importance of that concept in the life of the lifelogger. In this thesis, lifelog concepts are very diverse and range from categories such as people (e.g. male, female), nature (e.g. lake, beach), weather (e.g. rainbow, fog) to even sentiments (e.g. unpleasant, euphoric), see Table 1.1. Lifelog concepts are the concepts that appear or occur in the life of a lifelogger. These concepts include not only physical objects but also

[4] http://fitbit.com/

natural scene and emotional states. These concepts are essential to constructing lifelog linkages and personal archive or human digital memories [72]. Therefore, lifelog concepts are one of the elements that assist in building up relations between lifelog events and they can also be employed to generate lifelog search query sets.

Lifelog concepts compose parts of the ontology of lifelog MemoryMesh. These concepts inherently indicate the relationships between different activities and events of lifeloggers. Therefore these concepts can be applied in formalisation of digital life representation. However, in this research, we focus more on the representation of lifelog concepts and their attributes in linking lifelog events. The formalisation of concept representation and their mutual relations is one of the future work out-of-the-scope of the research presented here.

- **Activity**

  Lifelog activity is a vague concept in many lifelogging research scenarios, especially when it is mentioned at the same time as event, they are often used interchangeable. But in this thesis, we define lifelog activities to be physical activities that individuals behave in their lives like walking, standing, stepping-down, driving etc. While events reflect a logical sequential combination of activities and do not have one word or two words of naming, rather events can have detailed descriptions with time information, contextual information etc. The naming of activities is also included in the set of concepts in lifelogging, therefore the activity set is a subset of the concept set.

- **Lifelog Event**

  There is still no final definition of event in lifelog research. Zacks and Tver-

natural elements	
time of day	day, night, sunrise/sunset
celestial bodies	sun, moon, stars
weather	clear sky, overcast sky, cloudy sky, rainbow, lightning, fog/mist, snow/ice
combustion	fire, smoke, fireworks
lighting effects	shadow, reflection, silhouette, lens effects
Environment	
scenery	mountain/hill, desert, coast, landscape, cityscape, forest/park, graffiti
water	underwater, sea/ocean, lake, river/stream, other
flora	tree, plant, flower, grass
fauna	cat, dog, horse, fish, bird, insect, spider, amphibian/reptile, rodent
People	
age	baby, child, teenager, adult, elderly
gender	male, female
quantity	none, zero, one, two, three, small group, large group
relationship	family/friends, co-workers, strangers
Image elements	
quality	in focus, selective focus, out of focus, motion blur, noisy/blocky
style	picture-in-picture, circular warp, gray-color, overlay
view	portrait, close-up/macro, indoor, outdoor
type	city life, party life, home life, sports/recreation, food/drink
impression	active, euphoric, happy, funny, unpleasant, inactive, melancholic, scary, calm
Human elements	
transportation	bicycle/motorcycle, car/van/pick-up, truck/bus, rail vehicle, water vehicle, air vehicle

Table 1.1: ImageCLEF[5] - Image Retrieval in CLEF ImageCLEF, refer to the concepts for the visual concept annotation

sky defined events as "a segment of time at a given location that is conceived by an observer to have a beginning and an end" [173]. While by Hard et. al.'s definition [76], events should have a hierarchical structure, and should be composed of sub-events. According to Qiu et. al., an event is a natural unit for human memory [133] and there should be no major context change within one event. We consider lifelog events as life episodes that are chronologically consecutive and psychologically memorable that can be consistently segmented by lifelog owners.

- **MemoryMesh**

  MemoryMesh is a dynamically organized and hyper-linked lifelog that can retrieve elements (lifelog nodes) according to relations between them. Different from conventional lifelog archives, a MemoryMesh attaches more meaning concepts in lifelog data by extracting and analysing the data itself and forming a linked hypermedia mesh. It is a digital analogous system of human memory using associative human memory mechanism modelling. The contribution of this research is to build a lifelog model for effectively and efficiently utilizing a a large volume of lifelog data to enhance personal digital memory, which is named *MemoryMesh* in this thesis.

- **Lifelog Linkage**

  In MemoryMesh, all information is connected by some association or relations which represent the cue of linking one lifelog node (event) to another. The linkage model is one of the main research tasks in this research and it is covered in Chapter **??**.

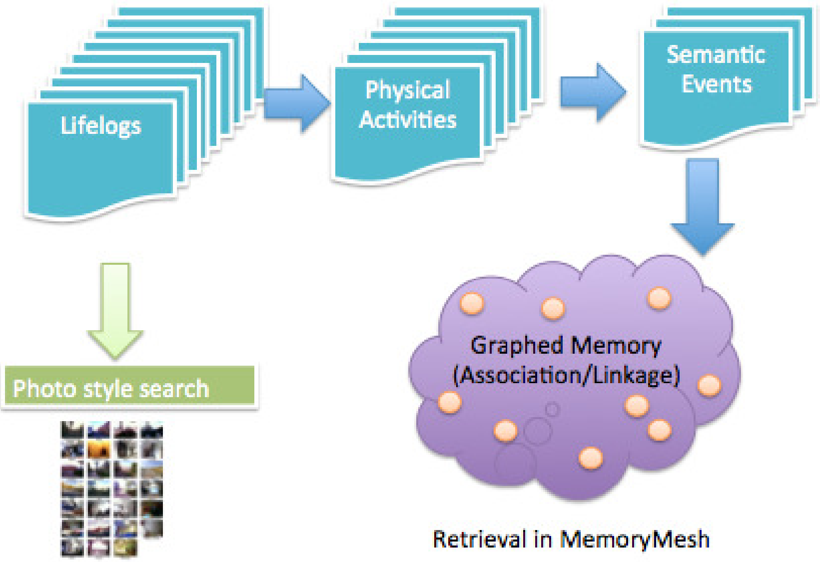


Figure 1.3: Overview of data flow in lifelogging systems, which shows the difference between traditional lifelog research systems and the MemoryMesh.

1.6 Contribution and Novelty

The goal of the research is to explore human activities and events in people's lifelogs. It also includes the potential influence of linkage models to support human access to their lifelog by applying linkage analysis to enhance retrieval effectiveness. These components are all main aspects of building future lifelogging systems as shown in Figure 1.3. Figure 1.3 illustrates the data flow of MemoryMesh lifelog modelling systems. The contribution of the research in this thesis can be summarised as:

1. *New event model*

   This thesis firstly proposes to apply activity recognition to human lifelog event segmentation to create a new approach to event segmentation;

2. *Linked MemoryMesh as a hypermedia or a linked graph*

This thesis firstly proposes the concept of MemoryMesh for lifelog research. It also implies application of the new model of event segmentation to facilitate linkage analysis and development of MemoryMesh;

3. *Retrieval model with linkage for MemoryMesh search*

   While the most previous work on lifelog retrieval is mainly database search or text ranking without event analysis, this work is an exploration of combining text analysis and WWW search methods in lifelog research;

1.7 Thesis Organisation

The thesis is focused on exploring personal lifelogging and provides initial results of a linkage analysis-based approach to making a lifelog more accessible in terms of data collection, organisation and visualisation. In this chapter, we introduced the motivation and the significance of this research. We also proposed our research questions and summarised our research plan. The remainder of this thesis is organised as follows:

Chapter 2 overviews the research on lifelogging from both academic and industry perspectives.

Chapter 3 introduces in detail all research methodologies that are applied over this research, from text mining techniques to image processing approaches, from information retrieval modelling to evaluation matrices.

Chapter 4 presents an overview of contextual discovery of lifelog using multiple sensory data. These data include accelerometer readings, GPS, WiFi, bluetooth that are collected via ubiquitous sensors. The long-term user study allows us to use devices and technology to record data passively instead of generating content objectively.

Chapter 5 describes visual content analysis for event segmentation using image processing techniques, which is the pre-process for lifelog linkage analysis in Chapter **??**.

Chapter **??** provides an overview of linkage analysis for lifelog about extracting semantic relations between event entities in lifelog. We extract relations between lifelog episodes, or events.

Chapter 7 summarises this thesis. Firstly, it reviews the research questions and concludes all solutions proposed in this research for the research questions. The contribution and future work of this study is also discussed. We also mention any limitations of this research in Chapter 7.

Chapter 8 lists all the publications the author has published in related areas.

Chapter 2

Overview of Lifelogging: Trends and Methods

2.1 Introduction

Personal lifelogging has been researched at the same time as the emergence and increasing ubiquity of digital technologies especially sensor technology. However, for the most part, information retrieval research has focused on data generated *by* humans, as opposed to data that is generated *about* humans. Although concepts like digital life, lifelogging, digital wellness, modern biometrics etc. are being discussed more frequently these days, the effective usage of sensors in human experience retrospection and reminiscence is still under discussion and is primarily laboratory bound. Few researchers are actually taking a holistic view of lifelog research. In this chapter, we represent our work on exploration and exploitation of WWW retrieval models in personal lifelogging, especially applying web linkage models in lifelog research. We introduce related work in the most relevant aspects: wearable sensors for lifelogging; contextual sensing for lifelogging; event segmentation and retrieval

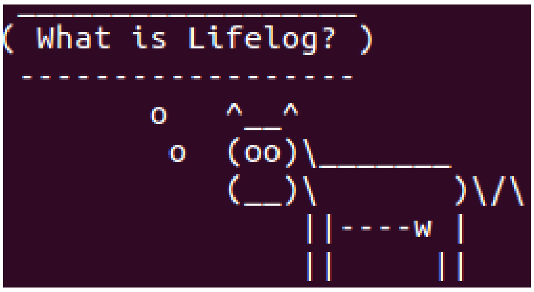


Figure 2.1: What is Lifelog

based on lifelog linkage analysis; life activity recognition and recording; use cases and evaluation of lifelog systems.

Although sensor technology has been researched for a long period of time, lifelog is a relatively new concept and consequently, there has not yet been a significant volume of related research. However, the following sections overview the state-of-the-art research in the area of lifelogging.

2.2 Five *W*s of Lifelog

Lifelogging is not just a trend in collaborative research with inputs from computer science, health science, psychology, physics, chemistry etc, but also affects many aspects of our life by changing the interaction approach to our living environment and cognition. We explore lifelogging through five main aspects: What, Why, hoW, Who and When.

2.2.1 What is Lifelogging?

From a perspective of either scientific research or industry application, there is still no agreed definition of lifelogging in terms of consideration of usage or meaning. Even so, in all definitions of lifelogging, it includes the ongoing sensing of totality of waking life experience and combines as much information as we can gather about our daily lives.

Recent practice of the MyLifeBits project of Bell & Gemmell combined active and passive logging by using wearable cameras and capturing real-world information accesses [15, 70]. Gurrin et. al. asserted that lifelog is the actual data that people gather either passively or subjectively for recording their lives. It could reside in a personal hard drive, in the cloud or in some portable storage device. The lifelog could be as simple as a collection of photos, or could become as large and complex as a lifetime of wearable sensory data (for example, GPS location logs or accelerometer activity readings) [74]. Lifelogging, for the purposes of this work, is considered to be the passive capture of daily activities using environmental and wearable sensors. The data that comprises a lifelog can take various forms, such as body sensor readings like body temperature, heart rate, galvanic skin response etc., image/pictorial recordings, which is also called visual lifelogging [161]. The potential to capture such a large data collection presents many challenges, including data analysis, visualisation and motivating users of different ages, backgrounds and technology experience to lifelog that are represented as large archives of rich multiple media and sensor content.

2.2.2 Why Lifelog?

When mentioning the purposes of why we lifelog, there could be countless reasons and unsurprisingly, many of which would not be understood yet. Wang et. al. found that lifelog could be applied to remembrance or re-finding previous events from the past [161]. Recent results from a pilot study have examined both sedentary and movement behaviour of a population of users. Early results from this research have shown considerable potential in the field of movement behaviour research where the duration of journeys to/from work or school, as a specific targeted activity, can be accurately estimated just from SenseCam images [92]. Also individually speaking, lifelog gives people a valuable insight into the life of lifeloggers. They can put their life experience into print as one would with photo albums of important events, such as birthday parties or Christmas family gatherings etc. It is the author's belief that the coming era of lifelogging technologies may also help people to replay their life entirely or become digitally consoling for lost of family members.

Corbin et. al. concluded that lifelog could be beneficial to lifestyle and behaviour recognition [45]. A sedentary lifestyle, coupled with overeating, can create an energy imbalance that causes abnormalities in the body. In terms of diagnosing diseases, it is currently undergone in a manner of question and answer or self-reporting, but its accuracy and precision is debated as for some diseases, patients can not describe their problem properly [121]. Lifestyle meta data can be utilized to assess the quality of human lives and provide doctors with a reference to diagnose diseases.

Given that most lifelogging work is based on examining the data from wearable sensors, we note that wearable sensors can be applied in many areas: logging autonomic activity [67], fall detection [40], location and activities recognition [107, 109, 147, 102], users spatial context recognition [147, 41], rehabilitation [168, 146, 128], natural disaster rescue [117], home healthcare [164, 152], martial art games [78], Gait analysis [150, 20], etc. However, all of these technologies are very domain dependent. Most researchers naturally tend towards the domain of lifelogging for personal health monitoring and awareness, especially for mental health. Sports is another applicable case of collecting data like movement, sweat, heart rate etc. In this work, we take a more holistic overview and explore data gathering and organisation for the generic lifelog domain, through our currently planned evaluations and cooperative linkages.

Zhou et. al. present a report over this based on a survey of various approaches to capturing lifelog data, which includes SenseCam/Vicon Revue, wearable smart phones, wearable video cameras, location loggers using GPS, bluetooth device loggers, human body biological state monitors (temperature/heart rate etc.)[174]. Potential lifelogging devices are evaluated and the advantages and disadvantages of different capture methods are presented, including consistency and integrity of capture, 'life coverage' of the captured data, as well as people's experience and attitude to these data capture devices. This has been done with user studies and surveys. This work also suggested the most suitable model of data capture for personal life logging in a variety of domains of use cases [174], which provides some support to our data collection that is used throughout this research.

2.2.3 HoW We Can Log Life?

Wearable sensors are pervasively utilised for lifelogging with recent availability of wearable sensing technologies [174], as shown in Figure 2.2. With the increasing acceptance of personal data gathering and on-line social sharing (e.g, on Facebook time-line, Twitter), lifelogging has become a mainstream research topic. SenseCam is a wearable device designed to capture aspects of wearer's daily lives [80, 79, 142] to support human memory reminiscence. The sensecam has spawned numerous related devices that are discussed below.

Given the low-cost and ubiquity of sensors, inexpensive smart phones come equipped with an array of sensors. Consequently we now have an ability to gather and store large volumes of personal data by using an inexpensive and ubiquitous smartphone. Qiu et. al. proposed a real-time lifelog recording and transferring system developed for Android users [132]. Researchers in University of Southampton

Figure 2.2: Sensor devices, a small collection

also developed distributed and pervasive healthcare Deja View system for supporting memory [48]. Most of these researchers focused on data capture. However, with many available lifelogging tools, how to collect, organize and represent lifelog data is still under much discussion [31, 32, 174].

Lifelogs can also be collected passively using ambient sensors. During the last decade, we have witnessed tremendous progress of this field and the development of numerous communication standards for low-power wireless communication. ZigBee is a specification for a suite of high level communication protocols using small, low-power digital radio devices based on an IEEE 802 standard for personal area networks [13, 10].

Utilising similar sensors, there has been a recent burst of research activity also in the area of smart homes, where ambient assistance intelligence technology can be applied to houses designed for older people or people with special needs [37, 99, 146, 127]. The Centre for Affective Solutions for Ambient Living Awareness

(CASALA) in Ireland has built 16 apartments for ambient living research [1]. Doyle et. al. proposed the requirements for gathering healthcare data in aware homes [60], and their results reveal that the old residents of these smart homes have a desire to play an active role in managing their health. And they also proposed potential concerns surrounding the delivery of such information through technology.

Lifelog technologies evolve in different directions, which include:

1. Web 2.0 blog

   Figure 2.3 is an example for Web 2.0 blog. As we can view from Web 2.0 technologies, people can write down their life diaries digitally and publish them into their weblogs in order to make a record of their experiences or thoughts. The limitation of this format of "lifelog" is that it requires human input with the associated cognitive effort and it is just a digital version of the conventional diary that people have kept for centuries. What makes lifelogging different is the fact that it should be an automated process that requires minimal or no human input to gather and organise [74].

2. Body Sensor Network

   Body sensor networks were first proposed by Lo et. al. as a wireless sensor platform for pervasive healthcare monitoring [112]. There are quite a few available sensors for constructing a body sensor network to collect data about body movement and body condition. Although they can utilise similar wearable sensors, the focus of this research effort is different to the focus of our work on lifelogging, so we do not consider them any further in this thesis.

- Microsoft SenseCam

[1] www.casala.com

My blog

LucidLead Blog

How to be a Big Data Scientist

Sharing is Caring
Collecting

What is Big Data?

here is no question that 'big data' has hit the business , government and scientific sectors . Indeed, the term 'big data' has acquired the trapping so f religion! However, there are a lot of examples of companies that were into 'big data' before it was called 'big data' — ater m coined in 1998 by two researchers. • But, what exactly is 'big data'? In short, the term 'big data' applies to information that cannot be processed or handled using traditional processes or tools . Big data is an architecture and most related challenges are IT focused.

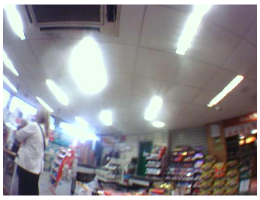
One standard definition of Big data is 4-Vs.

- Volume
- Variety
- Voracity
- Velocity

Figure 2.3: Web 2.0 blog example

(a) SenseCam

(b) SesenCam sample 1



(c) SesenCam sample 2

Figure 2.4: SenseCam and sample pictures it takes

The Microsoft SenseCam (see Figure 2.4a) is a small lightweight wearable camera used to passively capture photos and other sensor readings from a user's day-to-day activities [80]. It is designed to be worn around the neck taking user oriented photos of what is in front of the user. It captures on average 3,000 images of a typical day, equating to almost 1 million images per year. The Microsoft SenseCam was the first widely deployed lifelogging device and utilised by researchers in various fields of research, most notably in terms of human healthcare and memory studies. Hodges et. al. detail the potential memory benefits of a personal visual lifelog such as life records that is generated by devices such as the SenseCam or the OMG Autographer [79].

- Autographer

  The OMG Autographer (see Figure 2.4a) is a small wearable device which incorporates a digital camera and multiple sensors including a 3-axis accelerometer to detect motion, a thermometer to detect ambient temperature, a passive infra red sensor to detect presence of a person, and a light sensor to detect light. It is used to record a detailed visual record of daily activities. The camera itself can take 5-mega pixel-shots, with dedicated glass hybrid wide-angle lens which capture a fixed-focus 136-degree view. Other technical specifications include 8GB of storage, which can hold up to 28,000 images (which works out at over 12 days of capture on the highest frequency setting). There's also a medium and low frequency option, which users can switch between using the two-button set-up on the right edge of the device. This is also how users can enable Bluetooth syncing with iPhone, which springs up a notification window on the smartphone, propelling people into the dedicated application. The device charges through micro-USB, but we had no issues

Figure 2.5: Autographer

with it running for several days before the battery indicator dipped below half capacity[2]. The Autographer is the main lifelogging device that is utilised in this research and has been used in the data gathering experiments described in the remainder of this dissertation. The Autographer was chosen for this research because of it's quality of photo capture, reliability, all-day battery and ready availability.

- Bodymedia[3]

  BodyMedia (see Figure 2.6) is a wearable healthcare device developed by a medical and consumer technology company called BodyMedia. It records the cardiac activities of an individual and BodyMedia released a human physiology data set for testing at the 2004 International Conference on Machine Learning. The device has been used in hundreds of clinical studies [166].

- Shimmer[4]

  Shimmer (see Figure 2.7) is a wearable tool to collect accelerometer data

[2]http://www.engadget.com/2013/07/29/autographer-wearable-camera-hands-on-price/
[3]http://www.bodymedia.com/
[4]http://www.shimmersensing.com/

Figure 2.6: BodyMedia

Figure 2.7: Shimmer

[26] from individuals. It has a very sensitive accelerometer sensor with 100 Hz collection. Shimmer is lifelog recorder for GPS data and high frequent accelerometer which is very crucial for physical activity recognition.

- SenseSeer [133, 6]

  SenseSeer is an Android mobile phone application developed by Qiu et. al. [133]. It can collect all sensor data types (including camera) accessible in an Android phone and send them to a server for storage and processing. SenseSeer is an example of a software lifelogging tool for gathering lifelog data.

2.2.4 Who Can Benefit From Lifelogs

Lifelogs can be beneficial as a support for a new kind of social networking between family and friends. Aharonya et.al. looked into the connection between individual's social behaviour and their financial status, network effects to decision making by employing a ubiquitous computing approach and found that the value of social factors for choice, motivation, and adherence enables quantifying the contribution of different incentive mechanisms between family and friends [2].

Lifelogging is also beneficial to intervention in increasing physical activity [143] participation, which potentially supports health interventions by reducing any diseases that are caused by lack of physical activity. Chronic disease sufferers with ailments such as dementia can used lifelog tools to record their daily activities using fitness monitoring wristbands or cameras like the Autographer. The data can be passively or objectively transferred to doctor-patient communication tools to help doctors to diagnose correctly and find solution precisely [73]. So, all in all, there are three beneficiaries of lifelogging: lifeloggers, society and organizations and corporates. In this work we focus on supporting applications of lifelogging and explore how novel organisation technologies can be applied to assist in retrieval from lifelogs. As such, this research can benefit individuals primarily, but also by inference, society.

2.2.5 When to Do Lifelog

Widely speaking, we start to log lives in many ways even from the very beginning of our lives. Before we were even born our parents could have our ultrasound screen of foetus. Also when we were in childhood we could have a large number of family photos. These family photos can be one part of our lifelog repository. Unpredictable and creative new technologies are helping people to achieve goals of recording life easier. By applying ever-evolving new technologies, it could be very beneficial if we can take lifelogging as an all-life activity.

The point is that lifelog is for all-life recording or "lifelong-lifelogging". Especially the human generation that grows up with digitalized life, it is inevitable to have digital records of health history, on-line messages, family and friends photos. Such data can be massive, non-hierarchical archives consisting of interlinked life experiences. Such data can compose of a person's "*lifelong-lifelogging*". This is also a motivation of this research on lifelogging as a next generation of personal data, gathered throughout life. Such data needs to be automatically organised and made searchable. This is the focus of the research carried out in this thesis.

2.3 Previous Uses of Lifelog

New technologies can always be controversial when being applied in real-life use-cases, and it is the same with lifelogging. CardioTrainer[5] is an activity measuring mobile application, using mobile phone's accelerometer, combined with visualization and other feedback to help users increase their physical activity levels. Individual portable or wearable gadgets, no matter it is SmartBand or Narrative, these lifelog applications or props, can be used as complete activity tracker. These are mainly built for fitness and entertainment like monitoring work-out goals, or remembering the most valuable events of lives. Mostly these are for commercialization reasons. These devices provide a connection between mind/body and the digital world. And the data can be uploaded and synchronized easily with the consent of application or gadget users.

[5] http://www.worksmartlabs.com/cardiotrainer/about.php

2.3.1 Lifelog for Activity Recognition

Activity recognition fits into a bigger framework of context awareness. Activity for human lives can have different definitions according to different scales and perspectives and even goals. Some assert that life should be divided to be normal human activities and disordered ones in the view of pathology and medical science [12]; while some sociologists propose human life activities can be categorized as individual activities and social activities [65]. Some neurologists classify life activities into smaller and detailed ones like eating/walking/talking etc. for the convenience of behaviour analysis [88, 92].

Despite that, computer scientists explore life activity recognition using external devices like accelerometers, mobile phones, GPS locator, home sensors etc. in a more statistical and mathematical manner. By 3-axis accelerometer data, Bao and Intille proposed to use accelerometer for activity recognition [11] with users' annotation, and Bieber and Kwapisz's perspective in activity recognition is in agreement with this, proposing that activity recognition can be achieved by using phone accelerometer data [19, 102]. The techniques behind these activity recognition can be based on a machine learning classification.

2.3.2 Lifelog for Social Behaviour Analysis

Lifelogging can be utilized in investigating social mechanisms of people in the real world [2]. The evolution of lifelog research for physical measurement analysis liberates the lifelog research from the outer ring, making it more popular with public, not just people who have memory impairment. Caprani et. al. applied lifelog research specially in sharing and explored sharing as a motivation for family reminiscence and socialness [33, 32, 34].

2.3.3 Lifelog for Memory Enhancement

Lifelog for memory research had maintained the dominant focus of lifelogging research, as a means to help an individual review and remember past activities. Berry et. al. explored the neural basis of effective memory with SenseCam and its effectiveness in memory therapy [80, 16, 18, 17]. In our previous papers, we present a new generation of lifelog system to support reminiscence through incorporating event segmentation and group sharing [175].

One of the most important aspects of reviewing lifelogs for memory enhancement is the concept of a lifelog event, which is analogous to an episode or event in human episodic memory. Without detecting events, a user may need to slowly look through more than 3,000 images for every day to find a certain memory cue. The idea of segmenting the large lifelog data stream into events means that it is faster and easier to locate a desired sequence of images via a browsing mechanism. Initial event segmentation research was conducted in two main directions: 1) personal photo management using context and content information with the combination of user annotation [122]; 2) event detection from videos for specific aims like sports [137].

Meanwhile, a lifelog is more meaningful for a user and easier to browse if it can be segmented into different conceptual events. Face detection and novelty recognition is also applied in previous work on event segmentation [57]. Visual concept detection also informs lifelog event segmentation in object detection from lifelogs [28]. One approach for automatic segmentation of lifelog into events by different methods of improvements in the selection of normalization, fusion, and vector distance techniques based on location change, photo scene change etc. with manually groundtruthed events [58]. As far as we know, this is the only real prior work on this area for lifelog event segmentation. Events are broad and inflexible units of retrieval, we will add this with the prior approach as described in Chapter 5. It is our conjecture that we can develop improved event segmentation models by integrating additional sensors and a new technique for event segmentation that will be described in Chapter 5.

2.4 Challenges in Lifelog Research

2.4.1 Privacy Consideration and Subject Protection in Lifelog

Privacy is the most frequently mentioned ethical issue in conjunction with lifelog technology in the academic debates recently [75], especially so since most lifelogging research involves wearable cameras that capture other individuals in the course of their daily activities. Increased personal information and emerging lifelog management technologies based on privacy-by-design help to alleviate the issues alleviate some of the issues. However it is still a challenge to do any research in the lifelogging area, given these privacy considerations and a subsequent lack of willingness of users to give their private lifelog data for research purposes. Other privacy concerns include lifelogs being subject to enforced access and accidental disclosure of private lifelog data.

"Digital society did not turn out to be what people hoped", said MIT IDEASLAB's Alex (Sandy) Pentland's Sustainable Digital Ecology[6]. *"People should have the same rights and obligations in people's digital selves. You should control the information about you, who gets it and what they want to do with it."* Privacy consideration is the first issue to be considered during the preparation for experimental data for this research, since the Autographer can record very detailed life under proper set-up. Ethical approval was obtained from Research Ethics Committee in our uni-

[6] http://www.youtube.com/watch?v=O22acc48lKo

versity for the use of participants SenseCam images for a wide variety of lifelogging experimentation. For the particular research in this dissertation, we utilised a number of datasets and for our main research tasks, the participants were wearing a SenseCam or Autographer consecutively for over a period of one month. This guarantees a variety of activities (for example, 'Eating' at home or outside, 'Walking' in the street or countryside, etc.) needed in our experiments and was carried out in accordance with the ethical principals under which my research team operates.

In this research, we tried to comply with all the requests from our data collection participants, who may be concerned about any data that they may not wish to be included in the experimental dataset. In this case, such data could have been removed from the collection prior to the individual donating the data for use during the experiments in this thesis. In the appendix there is an example of one such ethics form that we have employed during this research. This form covers our data collection of privacy polices about collecting and using the data, as shown in Appendix A.

2.4.2 Big Data Storage and Processing

Big data generally comes with four elements of data attributes: volume, variety, velocity and veracity [74]. Lifelog data, due to the fact that it conforms these four characteristics, is considered to be a scenario that big data techniques can be applied to. Challenges to lifelogs [75] are various as in processing large volume of data. How to effectively and efficiently organize and access lifelog data is a big research question in this research area [74].

It is our conjecture, and the focus of this research, that WWW or Internet search and organisation technologies can be applied to lifelogging research with the inclusion of concerns about data protection. Data organisation techniques, such as inverted indexes can be applied to storing lifelog data. The primary contribution of this work, however, is the hypothesis that WWW linkage algorithms, such as PageRank algorithm [24] and other information retrieval basics can be applied to solving organisation and search challenges in large personal lifelog archives.

2.5 Lifelog Retrieval Systems

Initial research into lifelogging focusses on individual capturing devices (e.g. Steve Mann [114], Aizawa [4], Tano [151], SenseCam [80], deja-view [48]). There has been some work on holistic systems for lifelogging. Lee et. al. presented personal lifelog management system by summarizing photos [103]. In order to provide benefits to end users, the lifelog system must be an end-to-end system, incorporating data gathering, data storage and some forms of data retrieval. While early lifeloggers including Steve Mann [114] made inroads into capturing lifelog content, the most famous attempt to address the challenges to retrieval and value extraction from lifelog is the MyLifeBits project at Microsoft Research, which was concerned with gathering and making a searchable and long-term personal life archive for one individual [70]. Other notable work in the area has produced systems as a real-time life experience tool by Qiu et. al. [132] which utilises a conventional smartphone as the capture device with real-time analysis of the life experience, and the work of De Jager et. al. [48], who have developed a hardware device that operates in conjunction with a cell-phone to enable real-time capture and feedback. Despite prior work in lifelogging systems, there are few holistic systems for lifelogging. Within this research plan, we will develop lifelog retrieval systems to support the experimentation outlined later. One prototype system that was already built and presented in the 19th Multimedia Modelling conference (MMM 2013) [175] is described in Chapter 5 in more detail.

So far, there is little research about lifelog retrieval system evaluation. One of the main research in this area is conducted by Jones et. al [87]. They stated that in order to evaluate information access applications for personal lifelog, experimental collections are required which are sufficiently large and diverse to represent the expected features of real user personal lifelog [87]. The main contribution in this work is applying the evaluation techniques of information retrieval to personal lifelog retrieval and this is shown to be efficient in multiple works [95, 91]. So most of lifelog retrieval evaluation is conducted in the similar way as conventional retrieval systems using a test collection methodology to evaluate the individual components and user studies to evaluate the holistic solutions to the event segmentation and MemoryMesh.

2.6 Linked Archives

Since a lifelog can be a huge archive of multimedia data, it is our conjecture that organisation and retrieval performance can be enhanced by incorporating linkage-analysis algorithms and concepts. However, there have been no prior considerations of this hypothesis. The prior work in organizing lifelogs was based on either video-style playback or event segmentation with browsing or basic search [74]. In this work, we propose that lifelog can be represented as a densely linked hypermedia archive, called MemoryMesh. We introduce how this can be constructed and the potential to improve retrieval performance.

With the ever-increasing universality of sensor industry, many sensors can be used to sense our environment, in a manner similar to how humans sense their surroundings. We believe that digital sensors make it possible to automatically identify important events, just like a human can. In our experiments, we try to discover important events using the context extracted from many physical and virtual sensors. Since our lives are regular over a given period of time, any notable happening can impact our memory. These special moments become the important ones because they have a notable impact on our daily life. By analysing our context history, our life pattern can be discovered using pattern mining algorithms. By comparing the context of that moment with the life pattern, users' important moments can be discovered.

2.6.1 Lifelog MemoryMesh and Node Modelling

The Dewey decimal system for categorizing items in a library collection is a classic example of a hierarchical categorization scheme. People's family tree has a structure, so does lifelog data if it is taken as linked mesh of each lifelog events. Nodes are kingpins of the MemoryMesh system. As we will describe later, a node is akin to an event in prior lifelog work. Since an event is rich in data and associated meta data, it is shown that these nodes can be linked together into a mesh or web of human lifelog experience. This is called the MemoryMesh and is the core data construct used in this research. The research presented in this thesis shows that the MemoryMesh provides a viable option for managing lifelog data and making it accessible to the end user. While it is not proposed to solve the challenge entirely, this work is a pioneering endeavour to chart a pathway for future research efforts by taking the first steps in this exciting and challenging research area.

2.7 Conclusion

In this chapter, we introduce the background of lifelogging and its collection through various sensors. These sensors are categorized into two main classes: wearable sensors (entertainment supportive sensors, health monitoring sensors, sport supportive sensors) and fixed/unwearable sensors (home installable PIR, remote control window sensors, sleep sensors). We also present the concept and limited prior work on event segmentation/annotation, life activity recognition and lifelog retrieval system and system evaluation. In the very end the motivation and expectation for this research is discussed.

Chapter 3

Research Methodology

In this chapter, we overview the research methods employed in this research. These research methods range from approaches to lifelog data collection, data annotation, analysis methods that include machine learning, t-tests, image-object detection algorithms and evaluation matrix. We explain why, how and when these methods are employed in the thesis. This chapter also describes the experimental set-up from system configuration to evaluation measurements and user studies.

3.1 Introduction

As we described earlier, lifelogging is all-of-life logging [51, 15]. However, the majority of previous research in lifelogging has been focused on physical activity recognition and lifelog event segmentation for lifelog linkage analysis using wearable mobile devices to collect multiple sensor data, including Autographer and SenseCam images. While the data collection is getting more widely acceptable and the ever-decreasing cost of sensors and mobile devices makes it more feasible, there are still a few challenges in lifelog research that need to be addressed in order to create more valuable systems for real-life applications. These questions are:

1. Lifelog data collection. It includes multiple sensing platforms for construction and application.
2. Lifelog data storage. It describes how the data is stored in a cross-platform cloud based service.
3. Lifelog data analysis. Lifelog data analysis needs the semantic enrichment to support organization and access.
4. Lifelog data representation and visualization.

In the remainder of this chapter, we firstly introduce data gathering and storage process applied in this thesis. In the section 3.10, the prevailing approaches to solving the problem of lifelogging management and access are presented. In the last sections, we describe the approaches taken in solving these problems and describe the evaluation methods that are used in chapters 4, 5 and **??**.

3.2 Study Data Collection

As we presented our work on a survey on sensors for lifelogging that was presented at SenseCam 2012, we found there were many existing devices that can be used for lifelog data gathering [174], including SenseCam, Autographer, wearable sensors (e.g. BodyMedia Fit Armband [130], FitBit [46]), programmable devices (e.g. Arduino, Tyndall, Shimmer), Heart Rate Monitors, smartphone apps and various ambient sensors. In this research, we focus on applying wearable devices for data collection with the aim to enhancing individual user-centred applications of lifelogging. Wearable sensors can be applied to collect numerous kinds of data about the

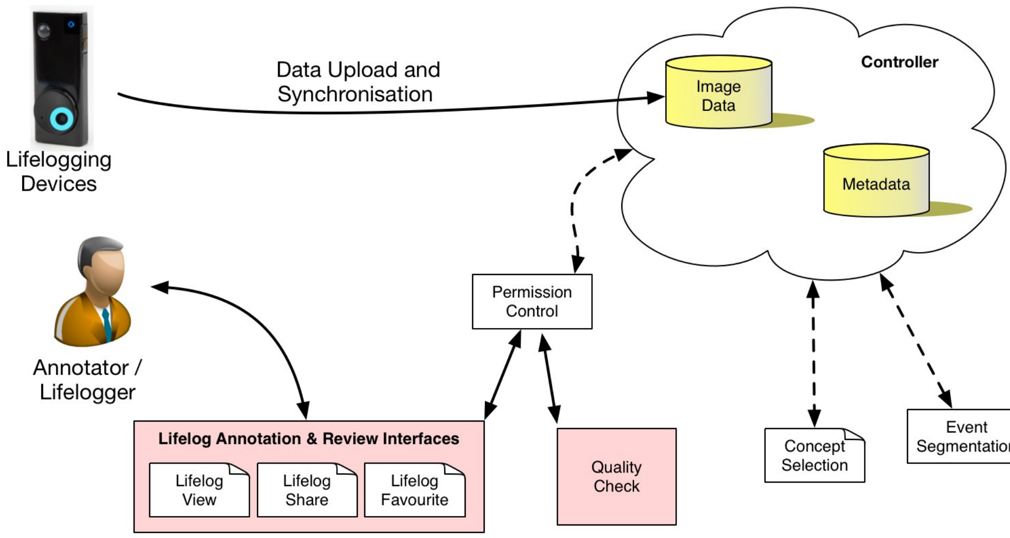


Figure 3.1: The MemLog system architecture

individuals including user locations, biometrics, activities, somatometric indices, visual content, aural content. We utilise lifelogging software developed for Android Smart phones, both in-house [132] and off-the-shelf (MOVES app), SenseCam and Autographer for visual capture, smartphone sensors and applications, Heart Rate Monitors to gather a media rich lifelog for individuals. The types of data include: passive capture visuals, explicit capture visuals, audio capture, location, acceleration, compass, user physical activity (walking, sitting, driving, flying, etc.), temperature, light levels, PIR output, bluetooth devices nearby and other context sources. It is our conjecture and understanding that a richer set of source data provides a deeper and more semantically meaningful event segmentations, data annotations and linkage graphs, which better support our planned experimentation on MemoryMesh evaluation. The metrics of data collected by these devices are shown in Table 3.1. In all, we aim to gather different sensor data for this research.

Figure 3.1 is the diagram of the MemLog lifelogging system that was developed to support most of the research proposed in this thesis. This architecture is represen-

Device	Sensor Type	No. Data Entries
SenseCam	image	2412325
Autographer	image	725773
	GPS	725773
	bluetooth	725773

Table 3.1: Sensor data Types for collection. Autographer collects one image with all sensor readings.

	GPS	Blue-tooth	Capture Frequency	Last Hours
1	on	on	High	5.5h
2	off	on	High	5.5h
3	on	off	High	5.5h
4	off	off	High	5.5h
5	on	on	Medium	15h
6	off	on	Medium	18h
7	on	off	Medium	18h
8	off	off	Medium	19h
9	on	on	Low	21h
10	off	on	Low	25h
11	on	off	Low	25h
12	off	off	Low	28h

Table 3.2: Autographer comparison over data capture type, frequency and maximum lasting time

tative of the software architectures of this research. In all, three different lifelogging applications were developed during the course of this work: *ZhiWo*, *ShareDay* and *MemLog*. *MemLog*, which is the most important of these.

The collected datasets include continuous collection of more than 25 types of phone-based signal, including location, accelerometer, Blue-tooth based device proximity, communication activities, installed applications, currently running applications, multimedia and file system information, and additional data generated by our experimental applications. This design is implemented as a Django based web service writing in Python, HTML, CSS and JavaScript. All the data is processed on the server side as back-end services.

3.3 Collection Devices

We used wearable devices with multiple sensors for data collection in this thesis. There are a large number of devices to choose from, but our initial exploration of the devices [174] suggested that a reduced set of broad sensing devices was sufficient to encourage data gathering and reduce the overhead on the individuals gathering data. Consequently, we centred our data gathering around two types of device; sensor-rich smart phones and wearable cameras such as the Autographer and SenseCam. These wearable mobile devices all can be worn for at least 5 hours after one full charge according to Table 3.2 when running at full-frequency capture. For the Autographer data collection, we asked each wearer to recharge their device after lunch resting time regardless of whether the device has battery or not to ensure a longer lasting data collection over the day. For SenseCam, one-day charge can last over 18 hours data collection, so the device only needs to be charged during night sleep time. In this case, there was a protocol designed to ensure close to full-day high-frequency data capture. Each sensing modality is now described, beginning with the two core modalities, but also including peripheral sensing modalities using additional sensors that we have utilised for some of the experimentation described in this thesis.

- Mobile phone sensing platform

  The SenseSeer platform [133] is a real-time lifelogging software that can work independently or in conjunction with a SenseCam. The software runs on Android smart phones and was developed at DCU. It utilizes energy conservation software on the smartphone to support all-day-long sensor capture and build a semantically rich life narrative. All available sensors on the phone (including camera, accelerometer, GPS, Blue-tooth, etc.) are employed to

Sensor	Description
Acceleration	Physical movement of the user
GPS	Geographic location
Blue-tooth	Social context
WiFi	Indoor location, location cache
Camera	Automatic photo/video capture
Speaker	Environmental sound/noise level

Table 3.3: Available sensors equipped on a smartphone

capture the current user context. Data is analysed on the phone and uploaded dynamically as a life-stream for further analysis when the WiFi network is available.

- SenseCam and Autographer

  SenseCam is a wearable camera with a wide-angle lens and Accelerometer, PIR, thermometer that takes periodically photos passively and automatically while users only need to wear it around neck using a lanyard. SenseCam has been applied in multiple research projects has has been proved to be helpful in facilitating human reminiscence and retrospective memory [80, 104, 56, 54]. Autographer is an evolution of the SenseCam that also hangs on a lanyard around neck and as such is orientated towards the activities that the user is engaged in. Similar to the SenseCam, it collects multiple lifelog data including images, GPS traces, bluetooth. Given that the Autographer is an evolution of the SenseCam, for much of this work, we employ the Autographer as opposed to the SenseCam

- Funf on Android Phone

  Most of modern smart phones already carry accelerometers, gyroscopes and other sensors and so can be used to log user daily life information. The problem with the Autographer is that the most frequent capture rate is 10 seconds per one capture, so it can not provide enough accelerometer data for sensory lifelog analysis. Funf Journal is an open source software to be installed in Android phones [2]. The orientation of worn phone considered for each of the accelerometer axis was: the x-axis, y-axis and z-axis were respectively aligned parallel to the coronal/frontal plane with positive direction pointing towards tail, the sagittal/median plane with positive direction pointing towards anterior; and the transverse plane with positive direction towards right.

- Basis B1 Band

  Basis B1 Band is a sleep and fitness tracker, and it can be worn like a watch. The data collected from this device (shown in Table 3.4) includes heart-rate that we need for activity recognition in our some of the work described in later chapters.

Different devices have different capture lifetimes owing to their batteries and levels of battery consumption. Battery lifetime depends on the type of data being captured and the frequency, as shown in Table 3.5

3.4 Data Collection Rules and Protocols

In our research process, we keep pre-defined data format to maintain data for experiments and follow the rules to collect our lifelog data:

- keep data in easily processed form, like csv, JSON , database etc.;
- keep data collected with time stamp, as well as data modification and all day gathering;
- keep data that is relevant to our research need;

date	calories	gsr	heart-rate	skin-temp	steps
2015-02-16 10:08Z	1.6	5.45083e-05	71	82.4	0
2015-02-16 10:09Z	1.2	5.51151e-05	71	82.4	0
2015-02-16 10:10Z	1.2	5.5421e-05	74	82.4	0
2015-02-16 10:11Z	1.2	5.48109e-05	71	83.3	0
2015-02-16 10:12Z	1.3	5.48109e-05	72	83.3	0
2015-02-16 10:13Z	1.2	5.5421e-05	73	83.3	0
2015-02-16 10:14Z	1.2	5.51151e-05	74	83.3	0
2015-02-16 10:15Z	1.4	5.51151e-05	75	83.3	0
2015-02-16 10:16Z	1.5	5.51151e-05	73	83.3	0
2015-02-16 10:17Z	1.5	5.57286e-05	74	83.3	0
2015-02-16 10:18Z	1.5	5.57286e-05	75	83.3	0
2015-02-16 10:19Z	1.3	5.6349e-05	72	83.3	0
2015-02-16 10:20Z	1.5	5.60379e-05	72	83.3	0
2015-02-16 10:21Z	1.6	5.6349e-05	74	83.3	0
2015-02-16 10:22Z	1.4	5.60379e-05	75	83.3	0
2015-02-16 10:23Z	1.5	5.6349e-05	73	83.3	0
2015-02-16 10:24Z	1.4	5.66617e-05	75	84.2	9
2015-02-16 10:25Z	2.1	5.5421e-05	76	84.2	2
2015-02-16 10:26Z	1.5	5.72925e-05	72	85.1	0
2015-02-16 10:27Z	1.6	5.72925e-05	73	85.1	0
2015-02-16 10:28Z	1.4	5.60379e-05	74	85.1	0
2015-02-16 10:29Z	1.5	5.66617e-05	76	85.1	0
2015-02-16 10:30Z	1.7	5.79302e-05	76	85.1	0
2015-02-16 10:31Z	1.8	5.72925e-05	74	85.1	0
2015-02-16 10:32Z	4.6	5.85751e-05	113	85.1	55
2015-02-16 10:33Z	4.5	6.02188e-05	113	86	46
2015-02-16 10:34Z	9.3	5.85751e-05	104	86	103
2015-02-16 10:35Z	8.4	5.89002e-05	117	86	86
2015-02-16 10:36Z	1.6	5.69762e-05	111	86	0
2015-02-16 10:37Z	1.8	5.98864e-05	85	86	0
2015-02-16 10:38Z	3.8	6.0553e-05	94	86	33

Table 3.4: Basis B1 Band Collected Data

Devices	On Time	Charge Time
Autographer	+5.5 hours	2pm, 12am
Mobile Phone	1 day	12am
Heart Rate Monitor	7 days	12am

Table 3.5: Battery life-time for different devices. Autographer is set to be high capture with all sensors on, which generates up to 3,000 photos per day

- keep data private to protect lifeloggers privacy.
- all data should be manually cleaned (to remove any potentially private data) before uploading, like any data that collector would not like to share can be deleted from the system.

We ask all participants to sign the *Informed Consent Form* [42] and seek to gain ethics approval prior to commencing this work.

3.4.1 Surveys

Survey is the research method that is used to assess thoughts, opinions, and experience of research participants [8]. It is often applied by psychologists and sociologists to analyse human behaviours. Especially in mobile CHI research area, survey is a useful research method to get an idea of real-world user needs and provide potential directions for design and research [96]. In this research, we claim that lifelogging is a research field that deals mostly with human needs, so in the early phrase of this research, we conducted a survey about lifelog data collection and the feeling of participants (who were aware of the field) about lifelogging. The survey includes telling users about what wearable sensors are, what data these sensors are collecting and asking how they are feeling about the concept of lifelogging. Only those subjects who are positive to wearable sensors are asked whether they would like to join the data collection process. The results of the survey can be found in the

3.4.2 Collected Data Formats

We focus on using data collectable via lifelog wearable cameras (SenseCam, Autographer) and smartphone software sensing, to achieve a easy-to-go data gathering process. The data collected using the above mentioned devices are:

- Lifelog Images

  The Autographer can record data with three type resolutions: 2592 pixels $*$ 1936 pixels, 640 pixels $*$ 480 pixels and 256 pixels $*$ 192 pixels. Average RGBSift extraction time for these three type of images are 1m30.30s, 4.64s and 0.92s. Thus, under the consideration of processing performance and image quality, we mainly use the medium-size Autographer images.

- Accelerometer

  Funf Journal [2] the maximum frequency for accelerometer capture is 100 Hz which is a good supplement when the maximum frequency for Autographer accelerometer capture is 0.1 Hz.

- GPS

  GPS data is sensitive to show people's privacy as it records locations of people's home, work places, routine of life. While in this research, GPS data is not public to share for the security and privacy of participants. GPS data is collected as a format of 3 dimensional matrix for a day. Each row of matrix records one reading of a time-labelled moment of the lifelogger's life. Each row is represented in latitude, longitude and altitude. But we only use lati-

participant.	social role	no of images
1	professor	11258
2	final year PhD student	20284
3	first year PhD student	15105
4	finance department worker	3523
5	assistant officer	5105
6	part-time first year PhD student	3921

Table 3.6: Social activeness detection result

tude and longitude for locating positions. For example, *53.385380, -6.257100* represents the location of School of Computing at Dublin City University.

3.4.3 Datasets

During this research, we have amassed a lifelog log dataset that may be open-sourced in the near future, which, also to our best knowledge, is an unprecedented lifelog dataset, which we refer to as the *DCULifelogdataset*, see Table 3.7.

- DCULifelogdataset

  DCULifelogdataset includes data collected from 6 participants over a period of a month from recruited by DCU under consent forms. These participants are trained carefully how to use Autographer and mobile phone and how to upload it into MemLog system for annotation and further experimental purposes. This dataset will be used by multiple experiments in this thesis including activity based activity recognition and event based hyper linked lifelog analysis.

- DIAL

  DIAL dataset is collected by 46 participants for one week each using SenseCam and embedded sensors. This dataset is provided by UCSD as a collaboration between two universities when the author was provided an opportunity

	DCULifelogdataset	UCSD_DIAL
Sizes	$123G$	$20G$
No. of Images	725773	35335
No. of Uploads	9711	46
Temporal Durations	$2515h$	$351h$

Table 3.7: Attribution of two datasets used in this thesis. Uploads are about how many person-days of lifelogged data. One person one day is one upload

to visit UCSD in 2013. In this research, as for the limitation of data sharing protocol, we are only using SenseCam data that was worn around neck, not any of the related hip-worn sensors. Each participant wore SenseCam for a week and were asked to upload their data after the one-week data collection period.

3.5 Data Analytic For Semantic Enrichment

The currently prevailing theory for detecting unknown objects or concepts is still using annotated data for training a predictive model. The basic lifelog gathered is annotated with a semantic enrichment step which integrates existing Insight visual processing tools (e.g. visual features from photos and semantic enrichment tools (e.g. weather, environmental lighting, visual concept detection). This task produces an enhanced media rich annotation for the data gathered, constructing the basis of the MemoryMesh.

In this thesis, the author has built an new annotation tool that is applied in lifelog data annotation. The annotation feeds the need of real-world lifelog and other wearable computing application and research. It includes local and remote data uploading, service infrastructure set-up and easy-to-use annotation interface [178]. The annotation applied in this thesis are activity annotation and event segmentation.

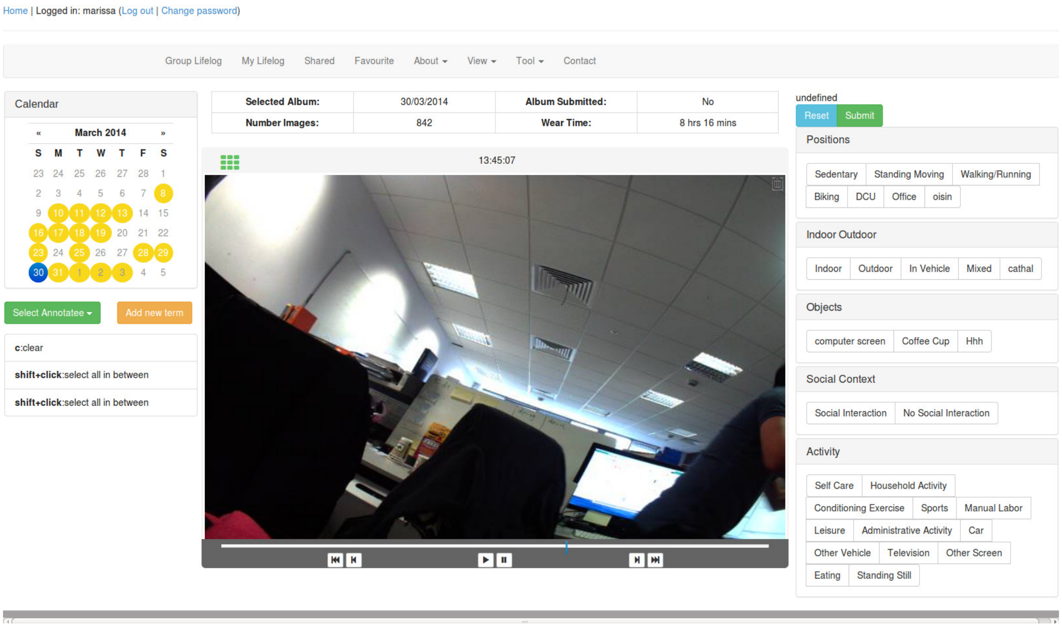


Figure 3.2: The MemLog system interface for the main annotation page

The two annotation phrases are both embedded into our MemLog system (Figure 3.1 and 3.2) [178].

3.5.1 MemLog Overview

It is our belief that concept detection is a basic technology for event segmentation and linkage analysis. To the best of our knowledge, there is no such a system available that allows us to annotate concepts from an organized list in lifelog data. Therefore, we designed and developed the MemLog system to meet this research need. The MemLog system is an end-to-end lifelog management and annotation system and a screen-shot of MemLog is shown in Figure 3.1, in which an annotator can be seen annotating a piece of content. On the left of the screen is the calendar to access the lifelog data, on the right is the annotation panel and the main part of the screen is the data panel, which displays (in this case) a full-sized playback view of the event being annotated. MemLog has been designed with multiple attributes that meet practical research purposes, including both local and remote data uploading, manual and automatic data annotation, and quality checking etc. These attributes have been gathered after extensive consideration of the needs of large-scale lifelog annotation efforts. The most relevant and important attributes of MemLog can be summarised thus:

- *User Authorization.*

  User authorization is necessary to enhance privacy and security of individuals lifelog data. User authorization is implemented through a user-group-permission map data table design scheme. The users can be data owners, annotators, researchers reviewing the data and managers who oversee the flexible process and allocate annotation tasks.

- *Uploading and Appropriate Storage.*

  Lifelogs can exist in different formats and sizes. MemLog allows for the uploading of lifelog data from the most widely used lifelogging tools, such as the Microsoft SenseCam, the Vicon Revue and the OMG Autographer. In addition, MemLog is designed to be usable in both a data centre environment or on a personal laptop or desktop computer. Therefore, in the uploader module, it supports both local transfer and remote upload.

- *Extensible Categories and Concept Labels.*

  Following discussions with collaborators who engage in extensive levels of lifelog annotation, we have identified the need for both core concepts as well as user (or task)-specific concepts. Hence we have implemented a core set of concepts to facilitate general concept annotations. These concepts include over 1,000 conceptual terms. For personalised concept annotation, users can add their own personal label sets to the predefined concepts. Each concept can belong to one category while each category contains one or more concept labels. For example, we have a category called "activity", it includes some labels including sitting, walking, standing still, cycling, driving etc.

- *Automatic Label Recommendation*.

  A lifelog for an individual user can surpass 4,000 images per day, hence automatic label suggestion is a valuable tool in the annotation process. The automatic label recommendation process in MemLog segments lifelog images into episodes or events and generates multiple most likely potential labels for each event (see Section 3.1). Annotators can check and modify these recommended labels, reviewing them in a dedicated review screen for validation and update.

- *Cross-Annotator Quality Check*.

  To avoid discrepancies between humans manually annotating visual data, the system is designed to have a quality check to re-access a configurable percentage of past annotations and correct when necessary. In this way, any problem annotators can be identified. In addition, the quality check is also useful for label recommendations.

- *Event Reviewing*.

  Event reviewing is to show users daily events in a format of segmented lifelog images. The algorithm behind event segmentation is supervised machine learning based on user annotated data and holistic trained model. After event segmentation, all events are presented to users chronologically with key-frames of each event. Each event can be demonstrated either in a conventional thumbnail view or in movie view (sequential playback of lifelog images).

- *Favourite and Sharing*.

  It has been shown that people have desire to share their lifelog data with their family members and friends [30]. MemLog is designed to have functionalities to facilitate people to manipulate their lifelog in terms of favourite and share with designated group members.

3.5.2 MemLog Deployment Architecture

Currently we have deployed our system in a university super computer data center and locally to support multiple research tasks. Figure 3.4 is the model view controller design of the MemLog system. The controller acts as the data storage centre, storing the images and the database of meta-data (including annotations). The lifelogging devices upload directly to the Controller and the users can access pages (the annotation, review and quality check pages) through the HTTP interface by means of a Permission Controller that regulates the views of the data that different users see. The event segmentation and concept selection processes automatically executes over the data in the Controller.

- Automatic Labelling of Lifelog Content

  Human effort in manual lifelog image annotation is a huge task when the lifelog dataset becomes large [34, 75], which every lifelog will, by its very nature. One of many potential characteristics of lifelog data is the recurrence of data that demonstrates people's daily life. These recurred lifelog subsets of life episodes is called life events. These events represents people's life patterns in the way of demonstrating repetitive life episodes. We use this as the

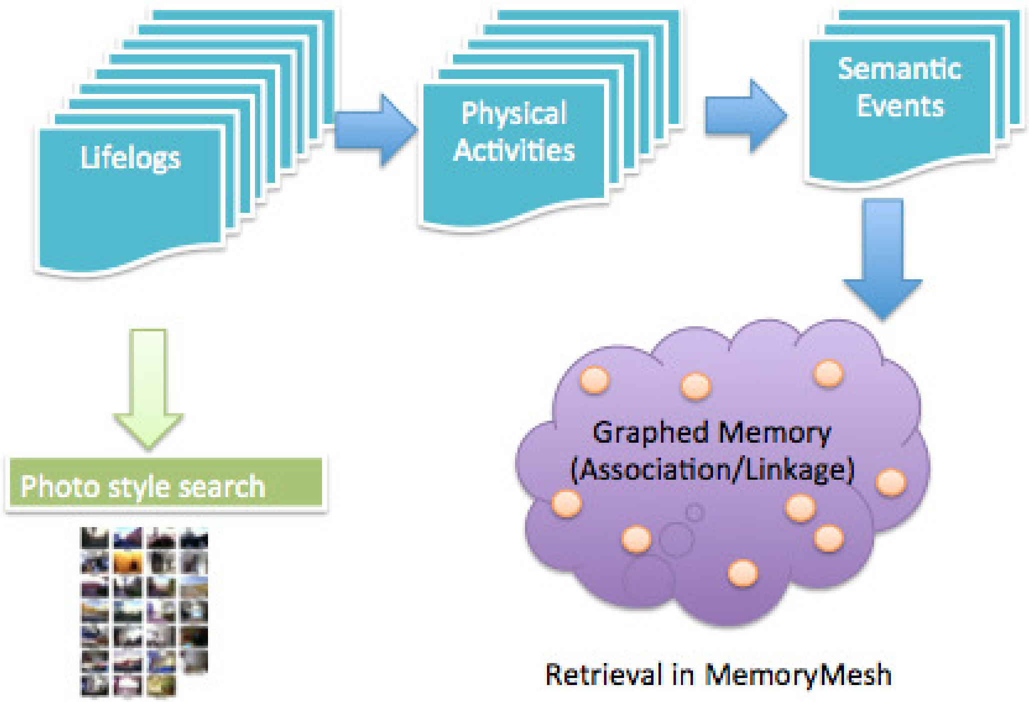


Figure 3.3: Overview of the data flow in the MemLog system

basis for label recommendation in our automatic annotation system process. Automated label recommendation is a crucial step towards efficient and effective management of this increasingly high volume of content [140], while such data can be assigned with category and content information using general semantics [100]. In this work, we compare the image similarity of lifelog events and recommend labels for current lifelog events using previously annotated labels to adaptively enhance automatic annotations.

- Physical Activity Annotation

Physical activities are annotated using *ZhiWo* system, as shown in Figure 3.7. *ZhiWo* was developed for this research purpose, to annotate physical activities in real-world deployment. Physical activities used in this research include walking, jumping, falling, driving, running, cycling, step-down, step-

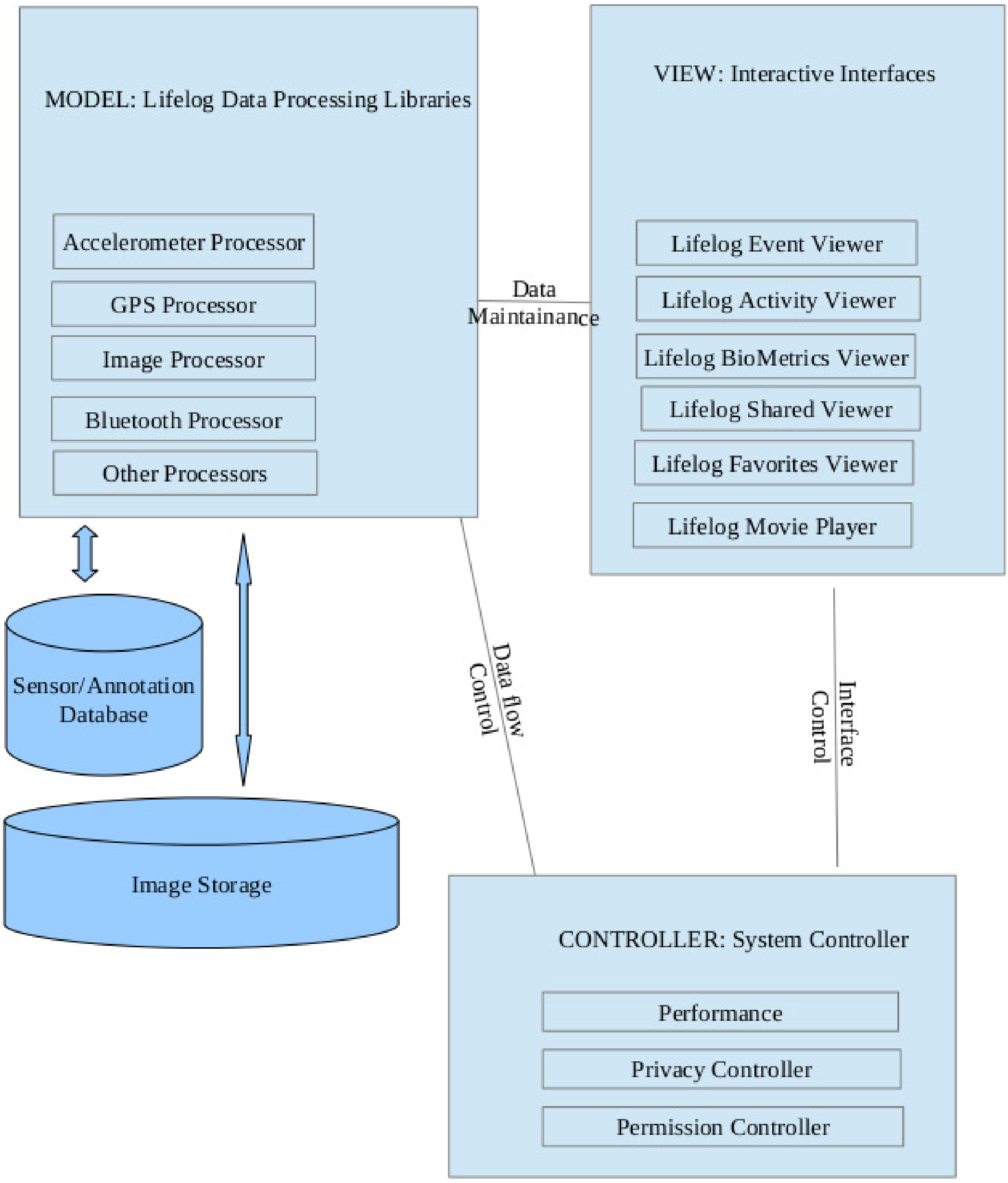


Figure 3.4: The Model-View-Controller design of the MemLog system

up, standing-still.

Subjects are asked to use cell phones to record their activity after finishing an activity, therefore, each activity label is assigned to a series of data between last activity and the current activity. Each time user subject a labelled activity, all sensory data associated is also sent to the cloud server. By using this approach, we successfully collected 7,205 annotated activities. More details about ZhiWo is described in Section 3.7

- Lifelog Event Annotation

  Events annotation are used to mark the boundary for event segmentation and as a source of metadata for the MemoryMesh. Events are different from activities in that, activities are objects and subject irrelevant, which events are objects and subjects relevant. Let's say "Having dinner" is an activity, while "having dinner with Ryan at student dinning hall at 12pm on 13th of March 2013" is an event. Also, activities are more general than events. An event may contain multiple physical activities, as during a period of having dinner, a person could be also sitting or standing.

3.5.3 Image Semantic Concept Annotation

In the image semantic concept annotation task, our goal is to detect the presence of the various concepts in lifelog images and provide us with the annotations on both per-image and event basis. To support efficient (i.e. non-browsing) access to lifelog, semantic analysis tools are needed. These act as software sensors to enrich the raw sensor streams with semantically meaningful annotations, that support the basis of retrieval and linkage techniques. For example, raw accelerometer values on a smartphone can identify the physical activities of a user [11], bluetooth and GPS sensors allow us to determine where and with whom people are with [29], while using automatic detection of concepts is possible from lifelog images [84, 145].

3.6 *ShareDay*: Inter People Event Linkage Annotation

People have always collected mementos over lifetime. With the digitization of mementos (photos and videos etc.), researchers have begun to realize the benefit of this to support reminiscence [129]. Sharing digital information is already commonplace, through emails, mobile phones and social networks. However, sharing lifelog data between family members, to our knowledge, has not yet been explored. Shared reminiscence between family members can serve many functions such as maintaining memories of past relatives, creating bonds and teaching younger family members from the elders' experiences. We believe that sharing lifelog within a family would enrich reminiscence and story-telling. In this thesis, we describe a novel software system to support sharing lifelog data.

According to Icek Ajzen's Theory of Planned Behaviour [5], human behaviour is a result of conscious intention instead of unconscious developments. According to Fogg, persuasive technology uses seven strategies to influence behaviour: reduction, tunnelling, tailoring, suggestion, self monitoring, surveillance, and conditioning.

- *Reduction* simplifies a task that the user is trying to do.
- *Tunnelling* guides the user through a sequence of activities, step by step.
- *Tailoring* provides custom information and feedback to the user based on their actions.

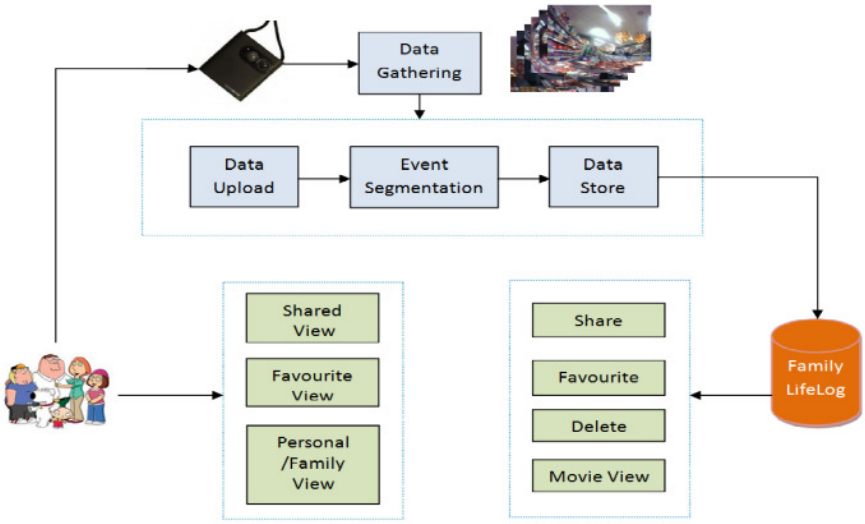


Figure 3.5: Overview of the ShareDay system for group memory enhancement

- *Suggestion* gives suggestions to the user at the right moment and in the right context.
- *Self-monitoring* enables the user to track his own behaviour to change his behaviour to achieve a predetermined outcome.
- *Surveillance* observes the user overtly in order to increase a target behaviour.
- *Conditioning* relies on providing reinforcement (or punishments) to the user in order to increase a target behaviour.

A previous study on intergenerational sharing [32] has shown that both older and younger people were more likely to wear a lifelogging device for the purpose of sharing images rather than simply wearing a lifelog device for private browsing or reminiscence. ShareDay was designed to support browsing and sharing through lifelogs on cross-platform implementations. To support family reminiscence we have designed the system to be used on a touch screen device displayed in a communal area at home so that all family members can upload, view and share their

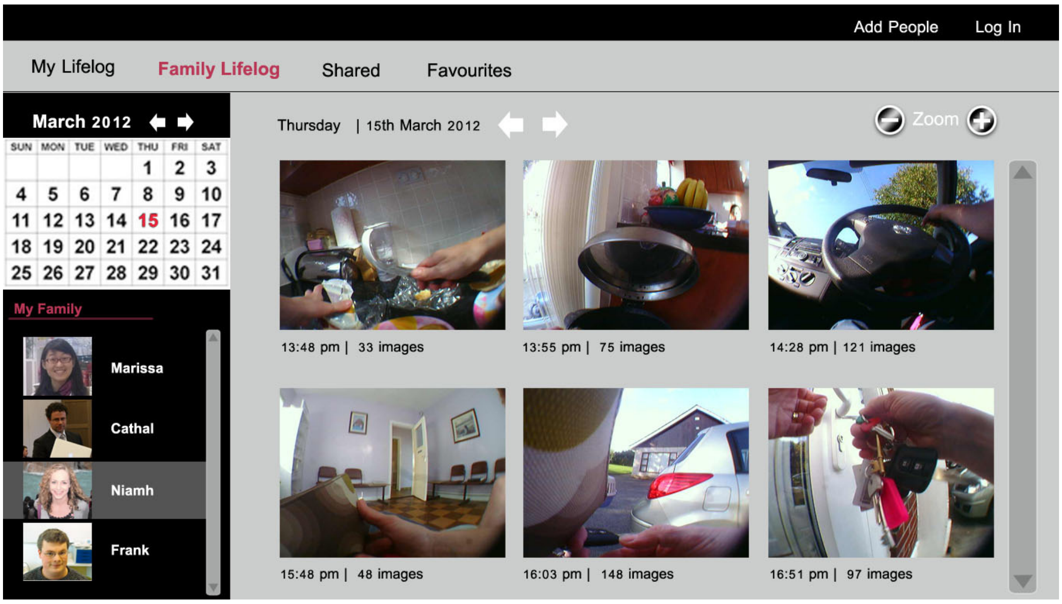


Figure 3.6: ShareDay system overview: family view

lifelog. Users can also view specific person's shared data by clicking on their profile (see Figure 3.6). The lifelog data can be viewed in two modes: personal and family. In family view, all group members can see all daily events and images in an overall, shared or favourite view. In personal view, only the logged-user can upload data and view their data. If logged, users can manage (share, mark as favourite or delete) and browse their visual lifelog. There are three main functions incorporated into the ShareDay system: managing personal data, shared data and favourite data. Figure 3.6 is the main view of the system, which shows all events gathered by all family members in that day.

The continuous lifelog data streams require segmentation into logical units (experiences). In prior work, the lifelog data has been segmented into events which are statically defined. The challenge that needs to be fixed here is how to build a static MemoryMesh over dynamic events. The solution here is to initialize the MemoryMesh according to a fixed period of data and add up new events referring to multi-level similarity based association. We examine and quantitatively evaluate how effective video shot and scene boundary detection, digital photo event detection and existing lifelog event segmentation techniques are, and develop/evaluate a new experience segmentation model which is dynamic and of variable length.

To collect a personal lifelog (in our view of lifelogs for this thesis) the user must initially capture images and other sensor data, and upload them to the system. Visual lifelogs present a challenge for developers as they need to represent the users day accurately and in a user-friendly manner, without requiring the user to browse through up to 3,000 images per day. For ShareDay, we integrated event segmentation model based on the work of Doherty et. al., [53], which organizes a sequence of SenseCam images into a set of events.

Events represent daily activities such as walking, eating, shopping, talking, etc. Key-frame images representing events are selected and displayed for each event, with six large key-frame images being selected for each event. When a user is logged in they can manage (share, favourite or delete) and browse through their visual lifelog. With regard to sharing, a user can share/mark favourite their daily events/images so that other family members can see shared data, which is designed under the concern of privacy.

Sharing lifelog information has a wide range of advantages for both user and communities [135, 160, 134]. The initial screen of the system displays shared lifelog of each family member. The user can touch on the name of their family member to view lifelog. A user can also browse through shared lifelog organized chronologically by touching the Shared tab. Group sharing to support family reminiscence is the primary aim of the proposed lifelog management system. However, we also wanted to ensure that users had control over their own lifelog as the content can be extremely personal. To accommodate for this users can select to share im-

(a) User Login (b) Activity Annotation (c) Activity Timeline

Figure 3.7: Snapshots of the interfaces of *ZhiWo* personal daily activity tagging and recognition system

ages/events when they are logged into their accounts. These images automatically be transferred to communal lifelog data set which all members of the family can view.

In our previous studies the participants reported that when they wanted to share images they had difficulty finding the images due to the vast lifelog datasets that are so easily accumulated [174, 32]. An easy fix for this was to provide users with a Favourites button. These selected images/events are automatically added into a favourites folder which users can find on the main menu bar. This work has been published in Publication 6 and 8.

3.7 *ZhiWo*: Physical Activity Annotation

We present an intuitive lifelog activity recording and management system called *ZhiWo* as shown in Figure 3.7. To support our planned research, we needed a large archive of annotated real-world activity data to develop models of user activity using machine learning techniques. Based on prior experience, it was not considered sufficient to ask users to 'go back' and annotate past data, nor were there any off-the-shelf tools available for us to use that maps raw sensor data to real-life activities.

Hence, we developed the third of three software tools, the purpose of which was to support the sensor data reading and annotation in a free-living, real-world environment.

ZhiWo system is an Android mobile phone application built on Android Version 2.1. It embraces all sensors in the mobile phone as protocols to gather sensor readings. By using a supervised machine learning approach, sensed data collected by mobile devices are automatically classified into different types of daily human activities and these activities are interpreted as life activity retrieval units for personal archives. It solves the problem in layer 2 in Figure 1.1.

This work has been published in Publication 4.

3.8 Development Technologies Employed

All of the three software tools had different user and operation requirements. Hence, each was designed and built independently. *ZhiWo*, for example, needed to run on a mobile device, hence it was developed in the Java language for application on the Android platform. ShareDay is a cross- platform application built in Flex language, because of the focus on end-user ease-of-use and integration with touch-screen computers in the home. MemLog, on the other hand, was developed to be used at DCU (for this experimentation) and also in partner institutions, hence it was developed as a web application and built using Django 1.7.1 and Python language, with all the data being processed in the server side.

3.9 Data Analytics Approaches

In this section, we describe the methods we chose to employ for data analytic and lifelog semantic enrichment. Further details of the application of these approaches are given in the following chapters.

3.9.1 Feature Extraction

Feature extraction is a research topic that focuses on finding the best features to describe the specified dataset in order to present all data in that dataset more precisely and in a manner that can be more easily processed by a machine. In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. In this research, there are two main steps that need to extract features from raw data: sensory data feature extraction and visual data extraction. For sensory data feature extraction, the raw accelerometer data and GPS data are transformed into a 46 dimensional space by using mathematical calculation and numeric attributions like mean, standard deviation, variance etc. For visual data feature extraction, visual images are transformed into a 4,000 dimensional space to represent each image using color histogram method.

3.9.2 Feature Representation and Fusion

Different feature representation for lifelog data can massively change results of research evaluations. While single feature representation can fairly well describe lifelog data, it is common trend to use feature fusion techniques to achieve a better performing data feature representations. These include the following methods:

- Bayesian Network

Bayesian Network is a probabilistic graphical model (a type of statistical model) that encodes a set of random variables of interest and their conditional dependencies via a Directed Acyclic Graph (DAG). Bayesian Network has been testified to be powerful in extracting and encoding knowledge from data when used in conjunction with statistical techniques. For example, a Bayesian network could display the probabilistic relations between people's choices to some restaurants, if a person felt better about a restaurant A comparing with the restaurant B, it is more possible this person will go to restaurant A instead of B.

- Supervised Learning

  Supervised Learning is a machine learning process that needs pre-defined learning materials, like categorized and annotated data. In supervised learning process, inputs are objects with an pair of values, one is the labelled desired output, the other one is normally a vector of multiple dimensions. In ideal situation, the inputs should include all possible objects, while in reality, with the limitation of enrolling in all data, we can only conclude limited possible pairs. SVM based classification is a widely applied supervised learning methods, which is used mainly in this thesis for both sensory data classification and visual content analysis.

- Unsupervised Learning

  Different from Supervised Learning processes that need pre-labelled objects, unsupervised learning does not stake a claim to annotated datasets, but it tries to find hidden structure in unlabelled data.

  A Hidden Markov Model is a generative statistical Markov Model in which the system being modelled is assumed to be a Markov process with unob-

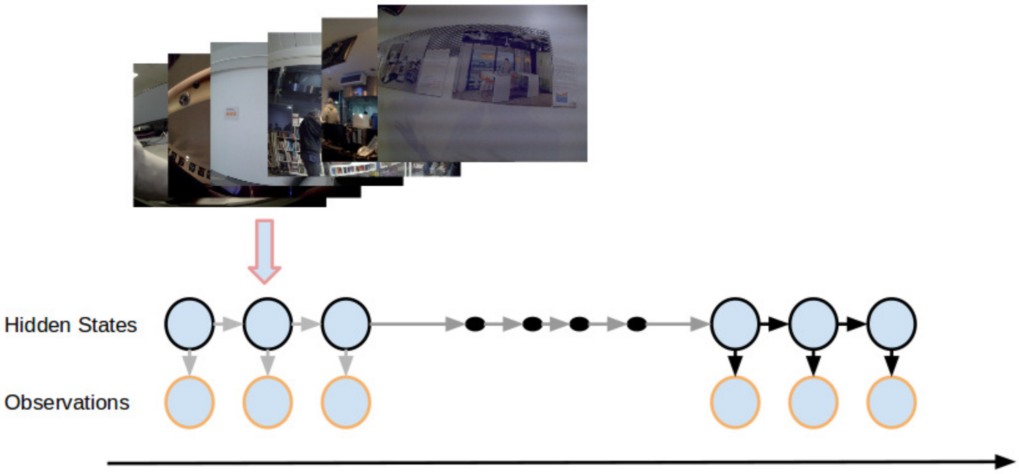


Figure 3.8: HMM structure for event modelling

served or hidden states. It is a probabilistic model of joint probability of a collection of random variables $\{O_1, O_2, ..., O_T, Q_1, Q_2, ..., Q_T\}$. The O_t variables are discrete observations and O_T variables are hidden and discrete states. Under an HMM, two conditional independence assumptions are made about random variables that make associated algorithms tractable [161].

1. $P(Q_t|Q_{t-1}, O_{t-1}, ..., Q_1, O_1) = P(Q_t|Q_{t-1})$ It is that given the $(t-1)_{st}$ hidden variable, the i_{th} hidden variable, is independent of previous variables.

2. $P(O_t|Q_t, O_t, ..., Q_1, O_1) = P(O_t|Q_t)$ It is that the t_{th} observation depends only on the t_{th} state.

Expectation Maximization(EM) is an iterative method for finding maximum likelihood estimate of the parameters of a hidden Markov model for a given set of observed feature vectors. One of most commonly used EM algorithm is Baum-Welch algorithm [165]. Baum-Welch algorithm optimally estimates the probability of the HMM model by iteratively re-estimating model parameters.

Clustering is another applicable unsupervised learning method for lifelog data analysis. Clustering generally iterate over a set of data with a goal to achieve for ceasing the process of iteration. Here we use k-means (e.g., k-means, mixture models, hierarchical clustering).

3.9.3 Feature Scaling

Feature Scaling is a method used to standardize the range of independent variables or features of data[1]. Different features have different scales in terms of enumerating representation. In order to encode all features into the same dimensional space, feature scaling can be applied in contributing to this process. Feature scaling helps to eliminate the difference between real-world data and modelling features.

3.10 Evaluation Procedure

In this section, we describe our evaluation methods employed in the research presented in this thesis. There were a number of evaluations required of the components of the MemoryMesh, hence we needed to consider many approaches to evaluation, all borrowed from the fields of information retrieval and multimedia analytic. These evaluation approaches are utilized in this research for image processing, physical activity analysis and linkage analysis.

3.10.1 Precision and Recall

Prevision and Recall are standard approaches to evaluation in Information Retrieval (binary-rank-order), Machine Learning and Pattern Recognition. Precision, which

[1] http://en.wikipedia.org/wiki/Feature_scaling

can also be called positively predicted rate, is the fraction of corrected retrieved instances 3.1, while recall (also known as sensitivity, see Equation 3.2) is the fraction of relevant instances that are retrieved. Therefore, both precision and recall are based on an understanding and measure of relevance of programs. Generally speaking, precision is the probability that a retrieved item is relevant, while recall is the probability that a relevant item is retrieved in a search.

In Equation 3.1 and 3.2, I_{rel} means items that are relevant, while I_{ret} means items that are retrieved. Suppose we have a retrieval program for finding the lifelog images that represents a lifelogger is using computer, in which, totally we get 1,000 images, 200 images with computers and 800 images without computers. After one run of our algorithm, we get 250 images that are recognized by our program are images with computer. But actually, there are only 170 of 250 images that are genuinely computer images. Then, our precision in this run is $170/250 = 0.68$, recall is $170/200 = 0.85$.

$$Precision = \frac{|I_{rel} \cap I_{ret}|}{|I_{ret}|} \tag{3.1}$$

$$Recall = \frac{|I_{rel} \cap I_{ret}|}{|I_{rel}|} \tag{3.2}$$

The relation between precision and recall is shown in Figure 3.9.

3.10.2 F Measure

In order to determine if our tests are significant, we conduct some test statistics. In our case, we use the F statistic or called F-measure. $F - value$ is a single measure that represents the harmonic mean of precision and recall, see Equation 3.3. In this case, recall and precision are evenly weighted, so it is also called F_1 value. It

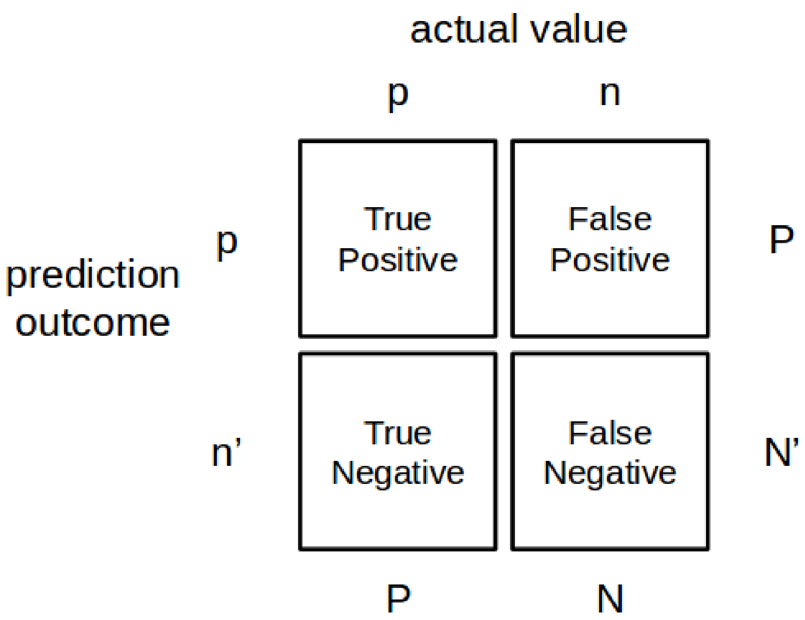


Figure 3.9: Relation between precision and recall

conveys the significance of tests or evaluations.

$$F = 2 \cdot \frac{Precision * Recall}{Precision + Recall} \tag{3.3}$$

3.10.3 AP and MAP

As stated earlier, precision and recall are single-value metrics based on the holistic list of documents returned by the program or system without considering the ranked order in which the returned documents are presented. In order to include the consideration of ranked order of retrieved items, we introduce average precision, which computes the the average value of $p(r)$ in precision-recall curve over the interval from $r = 0$ to $r = 1$. The precision-recall curve here is a curve that represents every precision-recall dots for top N retrieved items. This curve plots precision $p(r)$ as a function of recall r. Average precision is the average value of of precision at all positions. In Equation 3.4, $P(k)$ is the precision of the I_k in the retrieved item list, and $relk$ is 0 if I_k is relevant, otherwise, it is 0.

Most standard among the TREC retrieval tasks community is Mean Average Precision (MAP), which provides a single-figure measure of quality across recall levels, see Equation 3.6. Among all evaluation measures, MAP has been proved to be especially good in terms of discrimination and stability.

$$AP = \frac{1}{|I_{rel}|} \cdot \sum_{k=1}^{n} P(k) * rel(k) \tag{3.4}$$

$$rel(k) = \begin{cases} 1 & \text{if } I_k \; is \; relevant \\ 0 & \text{if } I_k \; is \; not \; relevant \end{cases} \tag{3.5}$$

$$MAP(Q) = \frac{1}{Q} \cdot \sum_{q}^{Q} AP(q) \tag{3.6}$$

3.10.4 NDCG

NDCG stands for Normalized Discounted Cumulative Gain [162]. NDCG originates from an earlier, more primitive, measure called Cumulative Gain [83]. It is a measure of ranking quality that is widely used in Information Retrieval. It measures the quality of ranking based on two assumptions:

1. Highly relevant documents are more useful when appearing earlier in a search engine result list (have higher ranks);

2. Highly relevant documents are more useful than marginally relevant documents, which are in turn more useful than irrelevant documents.

The two assumptions also meet the requirements of our lifelog user accessibility. We use this as one of measurements in lifelog retrieval experiments in Chapter **??**.

3.10.5 Significance Test

Significance tests are used in this research to determine whether an experimental result is statistically significant.

1. Student's t-test

   A t-test is any statistical hypothesis test in which the test statistic follows a Student's t distribution if the null hypothesis is supported. We use python statistical module for student t-test for different user groups.

2. Confidence Interval

   Confidence Interval is interval of numbers containing the most plausible values for our statistical results. In all research evaluation, we use the confidence interval value of 0.95.

3.10.6 Cross-fold Validation

Cross-validation, is also called rotation estimation [50, 159, 159], is a predictive model assessment technique for evaluating how the results of a statistical analysis generalizes to an independent dataset. Its main goal of Cross-fold Validation is to estimate how precisely a predictive model generates based on known data performed in the real-world settings on an unknown dataset, for instance from a real problem. This approach can limit problems like over-fitting, which describes random error or noise instead of the underlying relation [9].

In practise, we apply 10-fold cross validation, which means that we separate the test dataset into 10 sub-folders. We do 10 times of experiments, in each one of these experiments, we choose one folder as the training dataset (known data) and other 9 folders as the test dataset (unknown data).

To avoid any bias in the experimentations and software development, k-fold cross validation is employed. This is a technique where the training set is split into k disjoint subsets of equal sizes and a model is trained for each subset with the overall performance of the software being calculated as the mean accuracy of each subset. In this way, the optimization of the software for a single set of training and test data can be avoided. In this work, k is set to be 5, which has been empirically determined to provide a reasonable balance between processing time required to train each model and the accuracy of the validation.

3.11 Data Representation and Visualization

There is a need to represent lifelog activities in a descriptive format like a graphic form to allow users to have a better visualization experience. Visualization is a combination of analysis and illustration. This is because a picture speaks more than words. We used multiple visualization tool to represent our research results, these tools include Dojo graphics and charting [2] and Google charts [3].

1. Daily activity

   We visualize daily activity using pie chart [4] to give a lifelog information about what is the activity composition of the users day, as in Figure 3.10.

[2] http://dojotoolkit.org/features/graphics-and-charting
[3] https://developers.google.com/chart/interactive/docs/gallery/piechart
[4] http://demos.dojotoolkit.org/demos/pieChart/demo.html

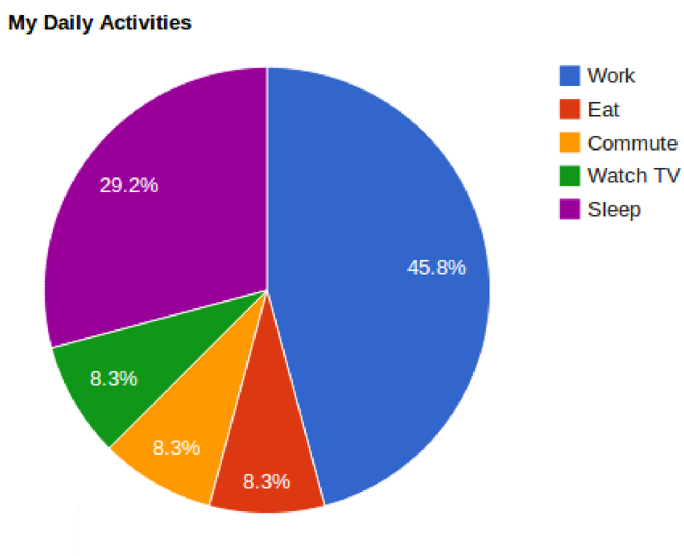


Figure 3.10: Personal daily activity pie-chart visualization

2. Events

   We developed an on-line events viewing system to help lifeloggers to review their daily lifelog. See Figure in Appendix B.4 for an overview of the system interface. All events are revealed as nodes in the lifelog graph.

3. Lifelog Linkage

   In this research, linkage analysis is based on double directed networked graphs of lifelog nodes. That means that we use node-link graph in Computer Science to analyse lifelog linkage attributes. See Figure in Appendix B.6. We also experiment using HTML5 Charts.js[5] and Infovis tool-kit [6].

[5]http://www.chartjs.org/
[6]http://philogb.github.io/jit/demos.html

3.12 Conclusion

In this chapter, we discussed research methodologies used in this research. These methodologies scale from data collection techniques to data mining algorithms, from data analysis evaluation matrices to systems for representation. These methodologies is supportive for all processes of this research. Data collection provides dataset for experiments; data mining algorithms and image processing approaches are mainly applied in Chapter 5 for lifelog visual discovery; evaluation matrices include evaluation methods for information retrieval, image content analysis and linkage analysis; system models are designed and implemented to visualize research results and provide a platform for user studies in Chapter 4, 5 and **??**.

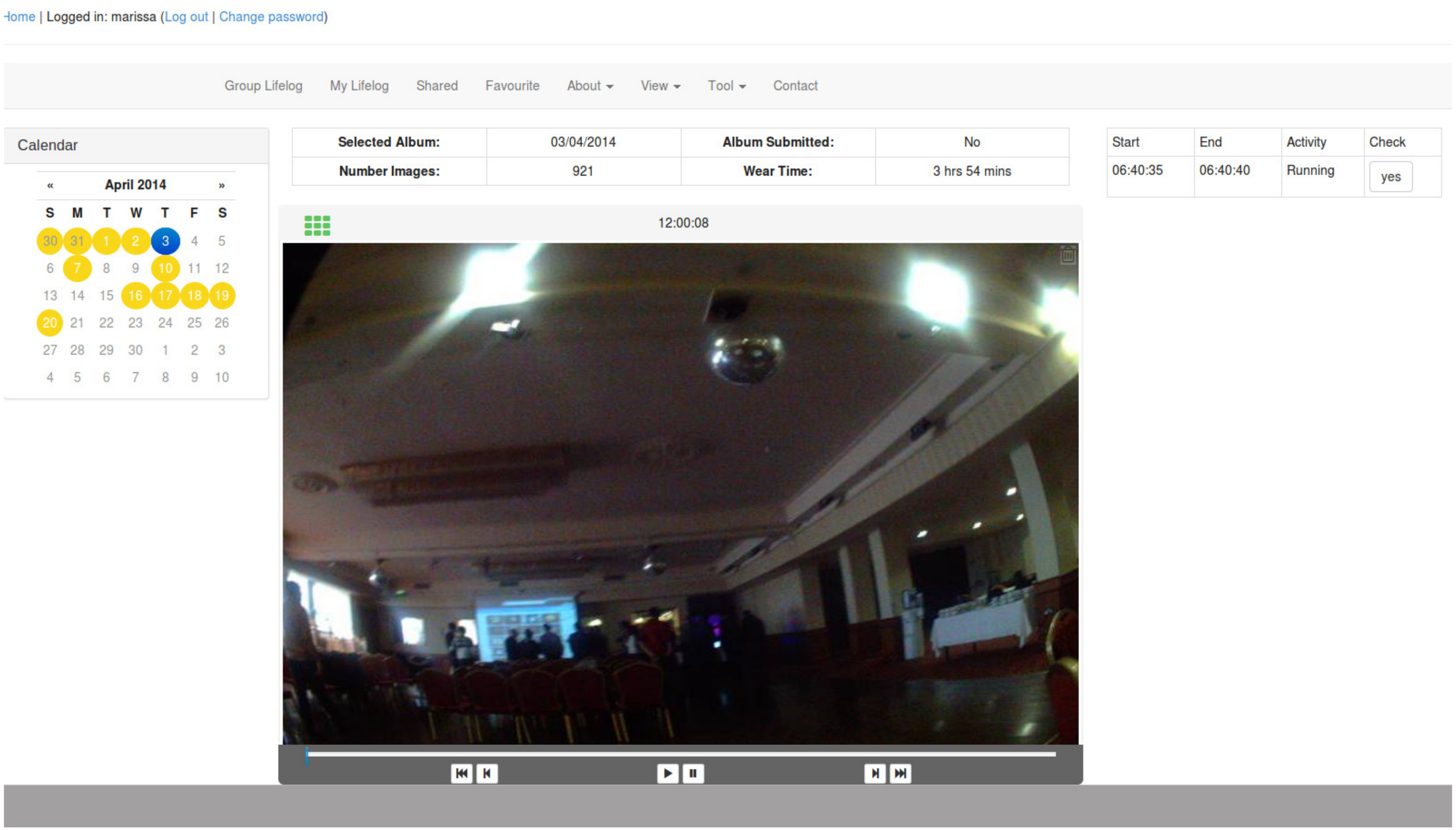


Figure 3.11: Event viewing in the MemLog system for the subject A

Chapter 4

Contextual Discovery of Lifelog with Sensor Data

Considering how people can benefit from lifelogging, it is likely that people will use imagery to log good memories, or use GPS to log their locations, or use Accelerometer to log their physical activities. We know that physical activities have highly relation to physical wellness [45, 44]. Physical activities have been testified to be significant in maintaining good body condition and preventing multiple diseases, including type two diabetes and some cardiovascular fitness [93]. In modern digitalized world, sensor technologies, in almost every aspect of daily business and personal life, simplify both common and complex processes and chores. In houses, technology breeds convenience, resulting in time and personal energy savings. Using sensors for managing personal lifestyle and physical activities has been researched in various scenarios. In this chapter, we provide an overview of contextual discovery from lifelogs using multi-modal sensor data. This data includes accelerometer, GPS, WiFi, bluetooth that are collected by multiple sensors. The

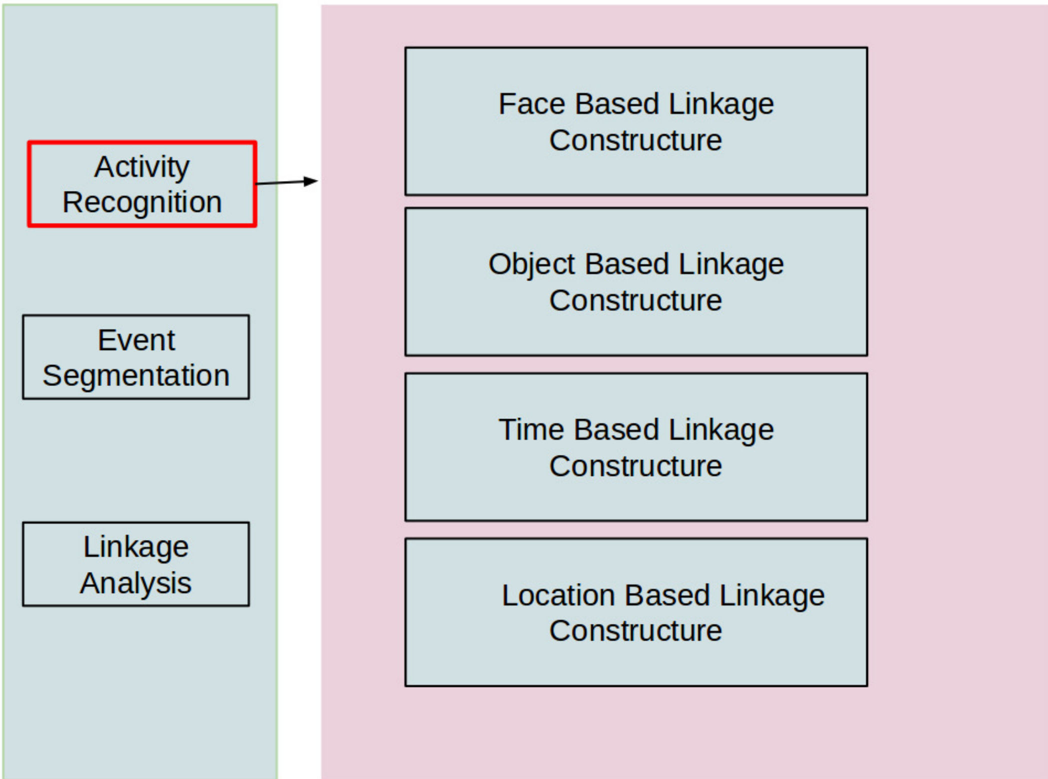


Figure 4.1: Overview of Chapter 4

long-term user study allows us to use devices/technology to record lives for later usage passively instead of generating content objectively. This chapter also introduced a self-reporting research method to provide us with ground truth data to conduct multiple research experiments.

4.1 Introduction

Lifelogging as in previous research was mostly used in memory enhancement, which in most cases were defined as predominantly visual capture of lives [56, 80, 79]. While recently, more and more research and applications are utilising sensory information of lifelog to explore positive use cases like daily physical activity assessment [154, 22]. Moreover, recent discoveries have found that using sensors to understand users' behaviour is a non-visual manner but contextual information can provide positive feedback in users' behaviour prediction [101]. For sensory contextual discovery from lifelogs, this thesis specifically focuses on physical activity recognition by introducing input of heart rate from already available sensor source, which, to our knowledge, is currently not facilitated by any tool or highly matched technique yet.

Over the past two decades, the accelerometer has been established and refined as a tool for tracking physical activities [154, 22]. Smart phones can be used as a platform for lifestyle monitoring [61, 62]. All sensory data will be taken as the context information on lifelog data analysis.

According to WHO, physical inactivity (lack of physical activity) has been identified to be the fourth leading risk factor of global mortality (6% of deaths globally). In the book *Physical activity and health: a report of the Surgeon General* published by National Center for Chronic Disease Prevention [121], it is addressed that Americans (or more generally, people) substantially improve their wellness and quality of life by embracing moderate amounts of physical activities in their daily lives. The study of human motion using sensor devices including accelerometers, GPS etc. has become increasingly extensive in various research areas. The exploitation of lifelog access includes extracting and employing contextual raw data for user access and further research. Thus, in this chapter, we focus on automatically recognising Activities of Daily Living (ADL) using portable, non-costly sensors and cell phones. Valuable information regards an individuals' degree of functional ability and lifestyle [22, 94]. More specifically in this research, physical activities are important semantics that facilitates to build linkages between lifelog events for retrieval purposes. The labels assigned to each activity are taken as attributes to associated event.

We have introduced that there has been little research effort into physical activity based event detection from lifelog, with the fact that prior state-of-the-art systems focus on analysing a small number of wearable sensors. For example, the previous work on event segmentation by Doherty et. al. incorporates visual processing of raw lifelog data [57] from SenseCam to identify visually significant changes in the composition of SenseCam images as a trigger for an event change. Since event segmentation is the core basis of this research, rather than simply accepting the current approaches, we propose to develop and evaluate a new technique for lifelog event segmentation that integrates Activity Recognition into the Event Segmentation, based on our conjecture (and related hypothesis) that visual processing is not sufficient to segment life activities into events. Initial results of our research into Activity Recognition have been gathered and the relevant paper was presented at ICMR 2013. In this chapter, we introduce activity recognition methods that we have used in our previous work and its application in personal lifestyle detection

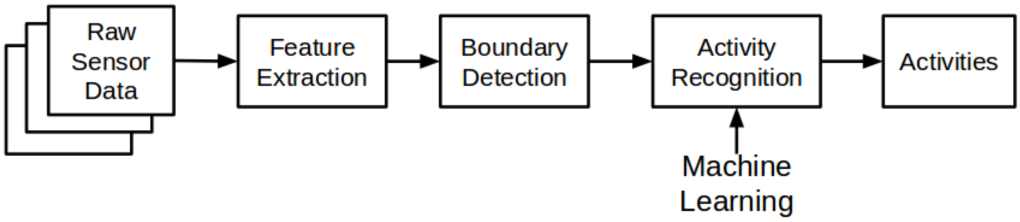


Figure 4.2: The process of physical activity recognition

and event segmentation for the next chapter. It includes why we are doing physical activities, how we are conducting physical activities recognition and how it is used in facilitating event segmentation and linkage analysis. Figure 4.2 illustrates the process of physical activity recognition, which we describe now.

4.2 Synopsis of Physical Activity

Activity is an umbrella term for lifelog researchers because it is often confusingly used as a synonym for *event* and life-activity, both physical and semantic. Physical activity recognition using sensors has been a topic of research for about 10 years. Bao et. al. proposed to use user-annotated acceleration data for activity recognition [11, 102]. Physical activity is a complex behaviour while here we narrow down categorization of physical activities to be a limited and digitally recognizable set of physical behaviours.

Physical Activity, "exercise" and "physical fitness" are terms that are often used interchangeably in epidemiology and especially behavioural epidemiology research while their definitions are distinguished [36]. In the definition of Casperson et. al., physical activity is any kind of bodily movement produced by skeletal muscles that brings about energy expenditure.

Human activity has different definition in different areas. Jennifer et. al. stated that human daily activities include walking, jogging, climbing stairs, sitting, and standing [102]; activities which are particularly relevant to exercise and healthcare. In our research, we will extend this scale to consider most of the 16 activities according to Kahneman's work in Science Magazine [88], see Table 4.2.

The United States Department of Agriculture writes on their website mentioning that "Physical activity simply means movement of the body that uses energy."[1] According to their definition, physical activities should be categorized as moderate or vigorous as per the exertion required in their performance.

These activities form the target for our new activity recognition technique, which in turn both supports semantic annotation of lifelog events and the actual effectiveness of the event segmentation model itself. However in this research, we only address 15 of the 16 activities (ignoring Intimate Relations).

In contrast, the definition of physical activity we have proposed in this thesis is more centred around lifelog access applications than those in common physical measurement uses and thus is considerably more restrictive. Physical activities are divided into two main categories: elementary physical activities and combined physical activities.

Elementary physical activities are physical activities that can not be split into smaller activities like standing, walking, sitting etc. Six types of elementary activities that are detected automatically are "WALKING", "SITTING", "STANDING", "LAYING", "WALKING_UPSTAIRS", "WALKING_DOWNSTAIRS", using Supportive Vector Machine [133] as shown in Table 4.1. We use abbreviations in later description of how the categorization of elementary activities are applied.

Combined physical activities are activities that can be divided into multiple

[1] http://www.choosemyplate.gov/physical-activity/what.html

1	W	WALKING
2	SI	SITTING
3	ST	STANDING
4	L	LAYING
5	WU	WALKING_UPSTAIRS
6	WD	WALKING_DOWNSTAIRS
7	C	CYCLING

Table 4.1: Elementary physical activity abbreviation and labels

1	Intimate relations (not explored)	9	Preparing food
2	Socializing	10	On the phone
3	Relaxing	11	Taking Care of my Children
4	Pray/worship/meditate	12	Computer/Internet/Email
5	Eating	13	Housework
6	Exercising	14	Working
7	Watching TV	15	Commuting
8	Shopping	16	Napping

Table 4.2: Combined target activities in this lifelogging research

different or same elementary activities as shown in Table 4.2.

Physical activities play significant role in daily lives as every human is conducting meek (walk/shaking hand) or vigorous (boxing/running) kinds of physical activities, just as stated by Caspersen et. al. “No member has no activity just as no person has no fitness - all are active or fit to greater or lesser degrees” [36]. Therefore, physical activity recognition is an important means of recording body movements of humans. Additionally, since physical activity recognition can facilitate digital calorie expenditure reckoning, people would like to use modern devices to record their life data about physical activities, which is one of the underlying motivations of this research to perform. To the best of our knowledge, the exact correlation between accelerometer output and energy expenditure has not yet been established, and given that accelerometers only calculate base don physical movement, the energy expenditure based on cognitive processing and other baseline body activities will not be captured by accelerometers.

Equation 4.1 can be used to express the caloric contribution of each category to the total energy expenditure due to physical activity according to [36].

$$kcal_{asleep} + kcal_{occupation} + kcal_{leisure} = kcal_{totaldailyphysicalactivity} \tag{4.1}$$

4.3 Adaptive Hierarchical Physical Activity Recognition

In this section, we are introducing our proposed adaptive hierarchical physical activity recognition methods using sensor data collected by mobile phone and Autographer.

We used Artificial Neural Networks as a classifier. ANNs have advantages of being adaptive and capable of providing correct answers even if the previously unseen features are entered as inputs [105]. Our wearable system is based on a new set of 20 computationally efficient features and the Random Forest classifier [35]. The following section will describe how the date is collected and how the features are extracted and applied in physical activity recognition.

4.3.1 Data Collection

The data for physical activity recognition was collected with both Autographer and Android phones with our in-house lifelogging software, as previously mentioned). The Autographer is set at high capture rate to ensure that we gather enough data (about 3,000 images per day) and the mobile phone data capture rate is set at 10Hz.

For all subjects, each activity was manually identified and coded with predefined activity labels. Autographer sensory data collection frequency of Autographer sensory capture and Android phone sensory capture are respectively 0.1Hz and 50Hz, thus, a period of T = 1/f = 100ms and with a resolution of 50 bits for acceleration and one picture in 10 seconds. The data collection was passive, hence it had little impact on the wearer's daily life. To increase accuracy, we asked wearers to wear the phone in the back-hip pocket, while Autographer was worn on the typical lanyard around neck. More details of this dataset is introduced in Chapter 3. Data collected by BASIS band include calorie, galvanic skin response , hear-rate, skin temperature and steps as shown in Figure 4.3.

The public Human Activity Recognition Using Smart phones Data Set [7] shared by University of California, Irvine can be downloaded here [2] [7]. This dataset was collected with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, SITTING, STANDING, LAYING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS) wearing a smartphone (Samsung Galaxy S II) on the waist. This dataset provides us with the basic training set for physical activity recognition for training our machine learning model for predicting.

4.3.2 Concept of Adaptiveness

The activity recognition work is significant because the model permits us to gain useful knowledge about the habits of millions of users passively just by having them carry cell phones in their pockets. In addition, activity is a natural segmentation enhancement element for our work on event segmentation (next chapter). In our research, we use a number of light-weight and power-efficient methods for ac-

[2] http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones

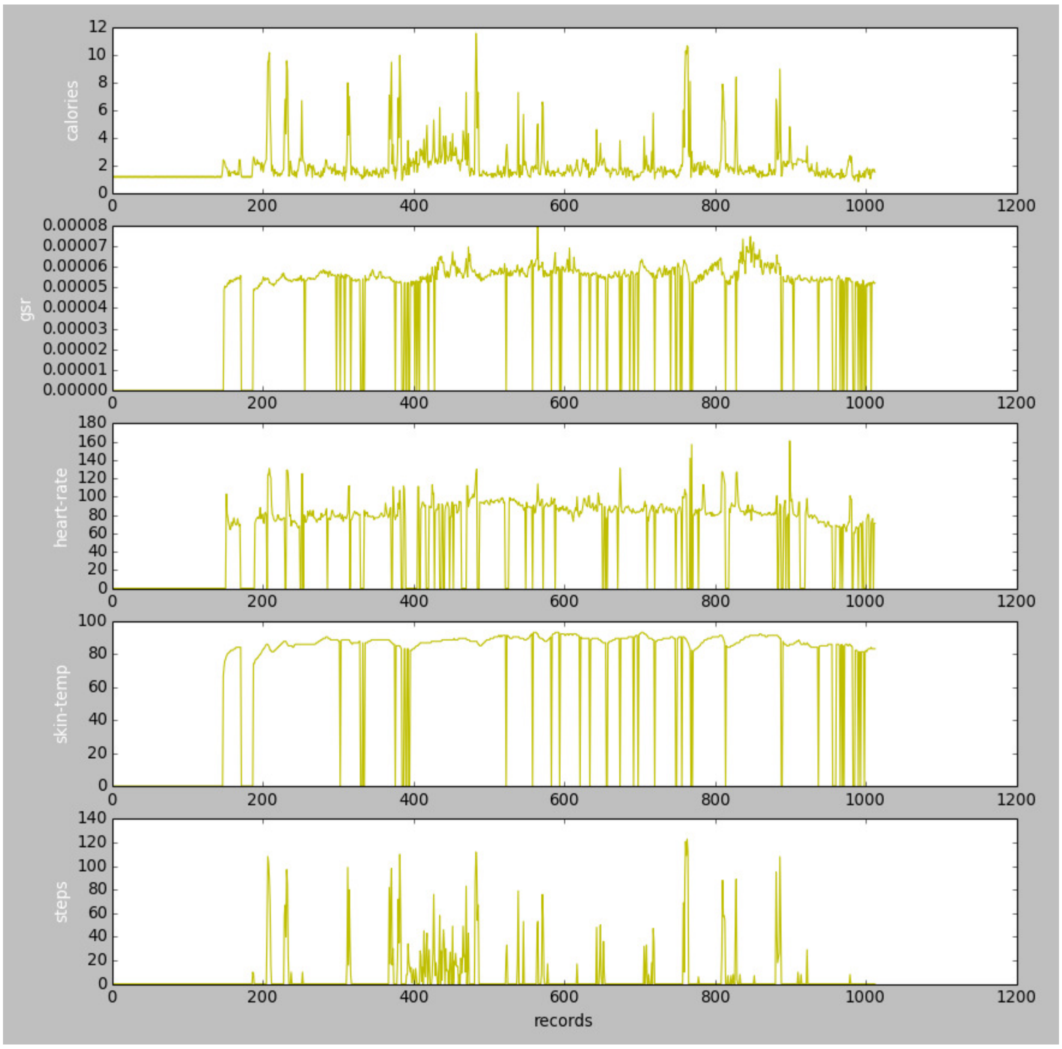


Figure 4.3: Basis Band data distribution of one day's data of one subject

```
CAM,2012/10/23 20:14:14,00172978.JPG,P
ACC,2012/10/23 20:14:25,-0.020,00.847,00.465
TMP,2012/10/23 20:14:25,0022.5
CLR,2012/10/23 20:14:25,00027
PIR,2012/10/23 20:14:25,1
BAT,2012/10/23 20:14:25,41675
MAG,2012/10/23 20:14:25,577,685,-4096
```

Figure 4.4: Raw SenseCam sensor data for activity recognition

tivity recognition: accelerometer-based and GPS location-based. It may be deemed necessary to enhance these methods with additional sensor sources, but this would be the subject of experimentation. In this section we outline our work for the development of activity recognition model. All data for this activity recognition had been collected in prior research in which the author was personally involved.

Each individual is unique not just in appearance but also in life behaviours that might appear similar to a human but actually massively different to sensitive sensors due to bias and sensitiveness of digital devices collection. When designing human computer interactive systems, adaptiveness is a vital element to consider [81]. Especially in our physical activity recognition process, we follow the research clues stated in [153], that the recognition process should adopt the rule of adaptiveness to allow the system to learn to adapt to individuals. Back to applying adaptiveness in the physical activity recognition process, instead of using one single model for physical activity recognition for all users, adaptive physical activity recognition generates personalized predictive models for each user as an optimization necessity in physical activity recognition.

4.3.3 Feature Extraction for Physical Activity Recognition

Different researchers use different quantitative data onto the features of physical activity recognition. Some use mathematical and statistical approaches. Lee et. al applied two levels of feature extraction [107]: four features including mean, standard deviation, spectral entropy and correlation, which are extracted from the first level for state classification, while at the next level (i.e., the activity recognition), three features including autoregressive coefficients, signal magnitude area, and tilt angle are used for activity classification.

Feature extraction should be the most important part for physical activity recognition process using 3-axial linear acceleration (a_x, a_y, a_z) and 3-axial angular velocity (v_x, v_y, v_z). Sensor-based activity recognition is a challenging task due to the inherent noisy nature of the input from digital devices. Thereby, in the raw data pre-precessing phrase, noise filtering is employed to get clean raw signal data and then it is sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window), similar to the approach used in *DCULifelogdataset* in Section 3.4.3. Considering the sensor acceleration signal could have been affected by gravitational and body motion, we used a Blutterworth Low-Pass Filter [3] to separate the acceleration signal into body acceleration and gravity. All features are normalized and bounded within [-1,1]. The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cut-off frequency is used. For each window, a vector of features is calculated using variables from the time and frequency domain. We base our feature extraction on his research but also extend to a wider spectrum by introducing more mathematical calculation, like described in Chapter 3, Section 3.9.1.

The acceleration signals with N samples during a period of time T can be represented as:

$$A(T) = \begin{bmatrix} a_{x,1} & a_{x,2} & \dots & a_{x,N} \\ a_{y,1} & a_{y,2} & \dots & a_{y,N} \\ a_{z,1} & a_{z,2} & \dots & a_{z,N} \end{bmatrix} \tag{4.2}$$

where $a_{x,t}$, $a_{y,t}$, and $a_{z,t}$ are the acceleration signals from x, y, and z-axis, respectively $(t = 1, 2, ..., N)$.

For each instant t_i of acceleration record, the three acceleration values $(a^x_{t_i}, a^y_{t_i}, a^z_{t_i})$ represents the three axis of the accelerometer x, y and z at the time-stamp of t_i. One sequence 1 second raw measures is subsequently used to obtain one instance vector. The instance vector $v = (v_1, v_2, ..., v_{39})$ includes 10 x-axis values, 10 y-axis values, 10 z-axis values, three means and three standard deviations and three variance for each axis. In the equation 4.5, j stands for the start point of x, y or z axis values.

Probability models and Support Vector Machines need a prerequisite of data input that is how to combine heterogeneous input data in SVM training and testing model.

$$mean = \frac{1}{N}\sum_{1}^{10} v_i \tag{4.3}$$

$$sd = \frac{1}{N}\sqrt{\sum_{1}^{N}(x_i - \mu)^2} \tag{4.4}$$

$$variance = \frac{1}{N-1}\sum_{j}^{j+10}(x_i - x_j)^2 \tag{4.5}$$

Correlation among these 3 axes represents the relation among three axes and helps to find out the strength and direction of a linear relation among three axes. The correlation among three axes can be expressed as:

$$R(T) = \begin{bmatrix} r_{x,x} & a_{x,y} & a_{x,z} \\ a_{y,x} & a_{y,y} & a_{y,z} \\ a_{z,x} & a_{z,y} & a_{z,z} \end{bmatrix} \tag{4.6}$$

where, $r_{i,j} = \frac{1}{N-1}\sum_{t=1}^{N}(\frac{a_{i,t}-\bar{a_i}}{s_i})(\frac{a_{j,t}-\bar{a_j}}{s_j})$, and $i, j \in x, y, z$, here where $r_{i,j}$ represents the correlation between two axes of the sensor, $a_{i,t}$ and $a_{j,t}$ the value of the t_{th} sample, $\bar{a_i}$ and $\bar{a_j}$ the mean, while s_i and s_j are the standard deviation for both axes, respectively.

All features included in this extraction process are included as in Table 4.3.

Analogously, the model produced by SVM for combined physical activity recog-

Feature	Explanation
mean()	Mean value
std()	Standard deviation
mad()	Median absolute deviation
max()	Largest value in array
min()	Smallest value in array
sma()	Signal magnitude area
energy()	Energy measure. Sum of the squares divided by the number of values.
iqr()	Interquartile range
entropy()	Signal entropy
arCoeff()	Autorregresion coefficients with Burg order equal to 4
correlation()	correlation coefficient between two signals
maxInds()	index of the frequency component with largest magnitude
meanFreq()	Weighted average of the frequency components to obtain a mean frequency
skewness()	skewness of the frequency domain signal
kurtosis()	kurtosis of the frequency domain signal
bandsEnergy()	Energy of a frequency interval within the 64 bins of the FFT of each window.
angle()	Angle between to vectors.

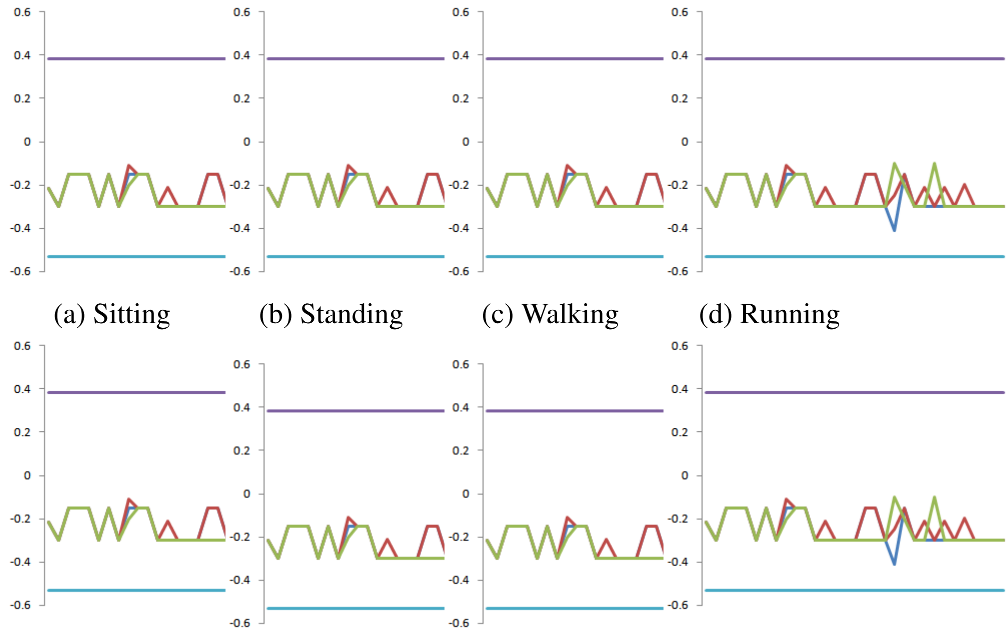
Table 4.3: Activity feature extraction methods and their brief explanation

nition shares the same features as for elementary physical activity recognition.

4.3.4 Accelerometer Enhanced Physical Activity Recognition

Physical accelerometer feature extraction is required for physical activity recognition. In activity recognition, there is a fair amount of uncertainty and risk involved with estimating the future value of next possible activities due to the wide variety of potential outcomes.

Accelerometer data records human body movement in three dimensions. In our research, the devices will be set to a high-collection rate (5Hz). To implement our algorithm we firstly annotate accelerometer data that is collected from users as they performed daily activities such as driving, walking, jogging, climbing stairs, sitting, cooking, and standing, and then aggregated this time series data into learning examples that summarize the user activity over a time interval. We then use machine



(a) Sitting (b) Standing (c) Walking (d) Running

(e) Stepping Up/Down (f) Falling (g) Driving (h) Eating/Drinking

Figure 4.5: Activity recognition results for each activity.

This is our first experiment that aims to identify physical activities. Y-axis stands for the fluctuation rate, and X-axis is the time. The time interval for this data is 5 seconds. All samples are randomly selected from our dataset. All dimensions are scaled to fit in [-1,1].

learning (ML) based on annotated training data to induce a predictive model for activity recognition. For this we are using an off-the-shelf SVM (Support Vector Machine) tool called LIBSVM [38].

WiFi demonstrates the location information of users by the strength of the signal. Take WiFi "School of Computing" for example, if this signal is detected, it shows that the user is in the School of Computing building. And in many cases, this user could be conducting an "Computer/Internet/Email" activity and it is less likely to be "Shopping". We use WiFi to detect user activities by applying the look-up mechanism.

GPS is efficient for speedy movement with location change on a vast scale, here

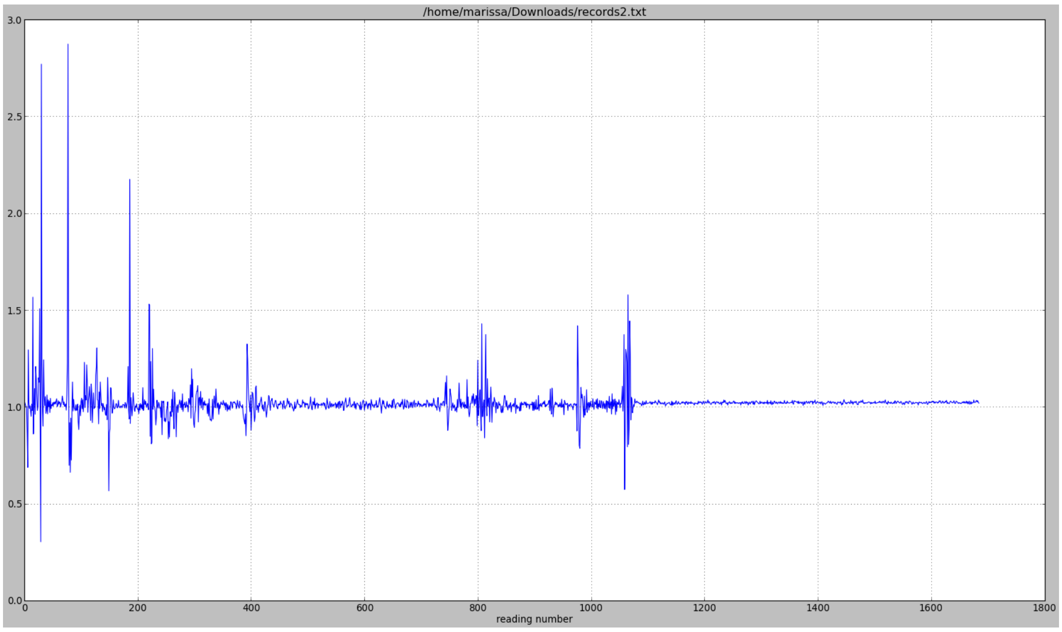


Figure 4.6: Accelerometer recorded signal for activity recognition
Accelerometer recorded signal for activity recognition

we can extract very rich context information by combining location information with relevant geographic knowledge. We use GPS location and map positioning and altitude numbers into human-readable locations and store these locations in the database. A new record, if it is within the scope of an existing location, should be mapped to that location. The approach has been successfully used in my prior work in GPS locating in late 2011.

4.3.5 Physical Activity Recognition Based on Feature Fusion

Using the combination of multiple features (GPS location, WiFi, Accelerometer) and time, we integrate the features from different perspectives. We conduct this comparison to examine whether the combination of these features outperforms any individual feature. These features are taken as different dimensions in the machine learning model training.

In physics and other sciences, a non-linear system, in contrast to a linear system, is a system which does not satisfy the superposition principle - meaning that the output of a non-linear system is not directly proportional to the input.[3]

Future fusion combines multiple feature sets in a linear way to expand feature sets. In this research, all features are combined in a linear space to mitigate the complexity of the feature space.

4.4 Personal Lifestyle Recognition

Our previous work on public health research provides some basis of our work on how to use lifelogs for public health research as a means of lifestyle detection [141]. This research demonstrates some of our thoughts about how lifelog can be applied in public and individual health prevention and monitoring. One additional source of evidence we will gather is concerned with Personal Lifestyle . We define PL to be either active or sedentary according to tri-axial accelerator data (a_x, a_y, a_z) and GPS (latitude g_{la}, longitude g_{lo}) data. We compute the coefficient of activity degree by:

$$A = a * \sqrt{a_x^2 + a_y^2 + a_z^2} + b * \sqrt{g_{la}^2 + g_{lo}^2} \tag{4.7}$$

Equation 4.7 is the lifestyle detection method firstly defined by this thesis. In this equation, the parameters a and b are empirically set as 0.4 and 0.6. If $A > threshold$, then the current movement is defined as active, otherwise it is defined as inactive. The $threshold$ here will be determined by specific experiments that will be conducted over large group of users. After we get assignment of all active or inactive features of all movements of the users, we can get the activity degree

[3] http://en.wikipedia.org/wiki/Nonlinear_system

1	Physical activity	MET
2	Light intensity activities	<3
3	sleeping	0.9

Table 4.4: MET levels for different activities

of the users at any point in time. We empirically think that users should be active during travel or moving time like commute/cooking/doing housework. If not, the user will be considered as an inactive person. The Personal Lifestyle (PL) value of people in specific time period feedbacks into the activity recognition in the way that active people has a shorter time episode of activities. The result of this recognition are evaluated by a user annotation-based experiment.

More careful descriptions and considerations of the components of daily living will facilitate cross-study comparisons. Physical activities monitoring using digital devices has been proved to be helpful for personal physical wellness maintaining [68, 45, 44].

The Metabolic Equivalent of Task (MET), or simply Metabolic Equivalent (ME) , is a physiological measure expressing the energy cost of physical activities and is defined as the ratio of metabolic rate (and therefore the rate of energy consumption) during a specific physical activity to a reference metabolic rate, set by convention to 3.5 ml o_2*/kg/min or equivalently as in Equation 4.8:

$$1MET = 1\frac{kcal}{kg*h} = 4.184\frac{kJ}{kg*h} \tag{4.8}$$

MET is the ratio of the work metabolic rate to the resting metabolic rate. It is used as a means of expressing the intensity and energy expenditure of activities in a way comparable among persons of different weight. Actual energy expenditure during an activity depends on the person's body mass index; therefore, the energy cost of the same activity will be different for people of different weight and height.

As described earlier, Event Segmentation (ES) has received little attention in research thus far, with the only comparable prior work on lifelog being done by Doherty et. al. [59, 58]. In this work, we use multi-model sensor data for ES. This multi-model sensor data includes images, GPS location, accelerometer, WiFi, bluetooth. With accelerometer and location, we need to find activities like *Socializing* or *Computer/Internet*; with bluetooth data, we can explore the social life of the users. The input of the algorithm is the activities and the output are events or episodes of life activity.

Instead of taking original raw sensor data as input for event segmentation in previous work, we take recognised activities as the input for event segmentation units, as stated in Chapter 1, which decreases the dimension of the unprocessed source data and has a potential to increase the segmentation accuracy.

4.5 Experimental Set-up and Variables

We have implemented an event segmentation algorithm based on time series and accelerometer data that is collected by a SenseCam/Autographer camera and instantiated as the ShareDay system [175]. ShareDay is a lifelog system to support reminiscence through incorporating event segmentation and event sharing. This was specifically trailed in a family scenario where ten families were employed to engage in lifelogging for a period of two weeks. This is the first system developed by the author and first user trial for this work. This system will be used as a framework upon which I build the next systems that underlie the planned research. This allowed us to trial the initial event segmentation approach with real-world users. The figure in Appendix B.3 is the snapshot of the main interface of our implemented system.

We evaluated our system through a user study into system usability, usefulness of system, and potential for event sharing between users. Table 4.5 demonstrates the participant distribution in the ShareDay evaluation system. In total, we got 4 participants wearing wearable sensors for a month. We found in this study that all users thought the system easy to use, easy to learn, an efficient way to access information, useful, fun and interesting. Most users would like to look at shared images only when they want to find some things of interest or for social reasons.

Participant	Gender	Age	NumberOfImages
FK	male	56	4618
CK	female	58	11071
DK	male	28	3249
SK	male	23	2443

Table 4.5: Initial user study evaluation of ShareDay

Although the ShareDay system provided a framework for my future experiments and a baseline event segmentation model, from the perspective of this research, the most important result from this experiment is actually the understanding gained that lifestyle activities could potentially form a valuable input into the event segmentation process, and this has formed one of the hypotheses of this research (hypothesis 2) in Chapter 1, that we can develop a better event segmentation approach by utilising activity recognition. This is also part of the initial study covered by this research.

4.6 Evaluations

In this section we describe how we evaluated Activity Recognition, which is the input of the Event Segmentation model and a source of annotation . To evaluate the potential of the activity recognition and its potential as a new data source for event segmentation, we are following the test collection approach in which we employ users to gather a suitable quantity of representative real-world source data and then generate a ground-truth of correct answers by engaging the same users to annotate their data with the correct annotations [56]. It is then possible to develop algorithms and techniques to map between the source data and the annotated source from the users. This allows us to employ machine learning techniques with K-Fold ($k = 10$) cross validation to develop non-biased software tools and to refine and re-evaluate these tools using a single, reusable gathered dataset.

In order to evaluate the potential of activity recognition, we have access to all participants' data over a period of one week. It was crucial that the participants engage in as natural a lifestyle as possible during the data gathering phrase. Since data gathering for lifelogging is a very challenging activity, this data gathering exercise is being done in conjunction with other researchers. These users were chosen from a broad spectrum of user groups, from students to retired individuals, with varying incomes, age ranges and education levels. After gathering data for a prescribed period of time, the users engage in the manual annotation process to build the ground-truth for the test collection. All annotators are provided images of the sequences to prompt important activity information. If the recognised activity was in agreement with the real one in the time aligned wearable camera images, it was marked as 1, otherwise it will be marked as 0.

In addition, the data is sub-sampled to ensure that the same number of positive and negative examples are used when training and evaluating the models, so as to further eliminate bias from the process. In total, we get 654,585 annotated activity records.

$$A_{pa} = \frac{Num_{correctlyrecognisedactivities}}{Num_{allactivitiesonecategory}} \tag{4.9}$$

Activity	**MM-AR**	**ACC-AR**	**WiFi-AR**	**LOC-AR**
WALKING	.5652	.4672	.1245	.1652
WALKING_UPSTAIRS	.6943	.6652	.0124	.0025
WALKING_DOWNSTAIRS	.9587	.9922	.2145	.0145
SITTING	.8962	.8521	.2351	.8566
STANDING	.7587	.7542	.0024	.0041
LAYING	.7754	.7752	.0085	.0057
RUNNING	.4524	.4523	.0014	.0056
DRIVING	.9869	.9122	.4120	.1245
CYCLING	.7568	.1254	.7547	.5652

Table 4.6: Activity recognition accuracy of four methods (MM-AR, ACC-AR, WiFi-AR and LOC-AR), the rate of automatic recognition result numbers divided by numbers of respective manually marked result. (In this daily activity recognition, we set average number of daily activity to be 30 according to our experimental and empirical observation.)

This is the approach to training and evaluation that will underpin all of the planned semantic software development in this work. It is beneficial in that it allows refinement and reworking of techniques long after the data gathering and annotation phase is over. Another benefit is that it allows for the comparative evaluation against existing techniques by simply implementing them over the new test collection and observing the performance. In Table 4.6, as we use binary category classification, all values are the rate that the activity is recognised as 1(correctly) and that is also manually annotated as 1 over all categories.

For our first experiment into activity recognition, implemented a multi-modal human daily activity recognition method for personal lifelog data analysis. Different from prior work which used single modal data for activity recognition, our approach combined three models of data (accelerometer, WiFi and GPS). Aside from this, we extracted features from the multi-modal data by calculating mean/standard deviation/median number/correlation of all three axial accelerometer data. Also, for GPS data, the centre point nearest neighbour is applied; for WiFi, strength of WiFi signal gives evidential hints of lingering in a place. The experimental result shows that the proposed human activity recognition approach combining GPS and WiFi data as auxiliary to conventional accelerometer based method can outperform any single-modal method, suggesting that the likelihood of certain activities occurring are correlated with the GPS and WIFI data, which are location-aware data sources.

As shown in Figure 4.5, feature samples for ten activities (Sitting, Standing, Walking, Running, Stepping Down, Falling, Driving, Eating/Drinking, Sleeping/Napping) under different features (GPS location (purple line), WiFi strength to centre WiFi device (blue line), the mean value of accelerometer data ACC-X, ACC-Y, ACC-Z in other three lines). In accelerometer data, all standard deviation/correlation are normalized to be within -1 to 1. WiFi strength is normalized to be -1 to 0, in which 0 means strongest and -1 means weakest. GPS location is calculated to be earth distance between the position of the subject and centre point.

The reason for this first experiment identifying these 10 activities instead of the 15 lifestyle activities mentioned at the beginning of this chapter is that as the lifestyle recognition experiment is still in progress and images had not been taken into consideration in this first experiment, some of activities like *Computer* and *Relaxing* are hard to be distinguished using the current approach. Thus we only extract physical activities for this first experiment, but for our later experiments, we identify the 15 lifestyle activities as a source of event segmentation input and experience annotation data. In the table, for "sleeping" activity, only WiFi signal can be detected well, that is mainly because we take time into consideration. In the algorithm, if time is correct and the signal strength is not changing, then is likely to be "sleeping" activity.

We explore the performance of our proposed multi-model recognition method of activity recognition by comparing with single-modal methods, see Table 4.6. In this table, ACC means accelerometer based AR method, WiFi represents wireless signal based approach, LOC shows the location based method and MM stands for the combination of three aspects. The data we used is the all participants data mentioned in the above context. We can easily see that different features work on different activities. Take activity *Stepping Up* and *Stepping Down* for example, ACC-Z works better for these activities as accelerometer fluctuates most strongly in the Z axis. The more detailed research output can be seen in Table 4.6.

One new finding from this research is a concept called Activity shifting. Activity shifting is a concept that we define in this thesis, because as we analysed the activity recognition results, we found that there is a possibility for some activities to be misclassified as a similar one. The phenomenon that the activities are recognised as another activity we call *Activity Shifting*. Activity Shifting (Table 4.7) demonstrates the possibility of different activities marked as another activity by MM-AR, in which, if the value is higher, the more possible that the activity in left of the table marked as the activity in header of the table. The values in the diagonal are the accuracy rate of the activities recognised right in the MM-AR approach. In this case, the possibility demonstrated in the Table 4.7 is the Activity Shifting potential. The activity shifting rate is higher if two activities are physically continuable. Take *Standing* and *Walking* for example, AS rate is 0.2586 in *Standing-Walking*, and 0.1253 in *Walking-Standing*. Also Activity shifting rate is 0.0053 in *Running-Driving*, and 0.0019 in *Walking-Standing*, which is lower than the counter part as *Running* and *Driving* differ in GPS changing feature (higher fluctuation). But the Activity Shifting rate is not diagonal, as the possibility of activity A being recognised as activity B is not equal to the possibility of activity B being recognised as activity A.

Different activities have different accelerometer fluctuation patterns, which are effective in recognising activities, but as shown in Figure 4.5, that means there are a few activities they are highly similar to each other. See activity *Standing* and *Sleeping/Napping*, both have low fluctuation rate although for the *Sleeping* activity, ACC-X meets minor regular up-down fluctuation as the body slowly moves when breathing. All these features construct unique activity fluctuation patters referring to multi-modal features. We reduce the potential impact of Activity Shifting. We also suggest that Activity Shifting is a valuable target for future research activities, but consider it to be outside of the scope of this research presented in this thesis.

4.6.1 User Annotation of Personal Activities

ZhiWo, as previously mentioned, is a prototype human activity recognition system based on a concept of "Knowing Me". The data source for *ZhiWo* is a raw sensor stream sampled from a smartphone worn on the body. *ZhiWo* semantically processes the sensor stream to identify distinct user activities (e.g. walking, resting, commuting, eating, shopping, etc.) based on a machine learning model trained by human annotation of activities on the mobile device. It provides browsing techniques based on an activity time-line view and can support locating specific activities of interest from within large lifelog archives.

To achieve the greatest accessibility of data gathering for various users, an application was developed for smartphones (Android OS) for collecting lifelog, see Figure 2.2. Users collect data by wearing the smartphone on a lanyard about the neck, or otherwise attached to the clothing in a manner so the smartphone camera is orientated towards the activities the user is engaged in. The data collected in-

Activity	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
WALKING	.5652	.1235	.0185	.0954	.1523	.0892	.0569	.1293	.1475	.7524
WALKING_UPSTAIRS	.0012	.6943	.2586	.0563	.2356	.3214	.2512	.0122	.1256	.1286
WALKING_DOWNSTAIRS	.1023	.1253	.9587	.2013	.4213	.1003	.0908	.0823	.0512	.4213
SITTING	.0021	.2013	.0805	.8962	.2106	.2030	.0456	.0053	.0101	.2121
STANDING	.1023	.0056	.1059	.0825	.7587	.2315	.2510	.1238	.1546	.1564
LAYING	.2013	.1203	.0101	.2404	.0170	.7554	.0589	.2665	.1565	.0127
RUNNING	.0012	.1235	.2451	.2131	.1123	.1452	.4524	.2441	.2211	.1124
DRIVING	.1231	.1254	.1212	.1598	.0987	.0984	.0587	.1245	.0869	.1547
CYCLING	.0178	.0123	.1231	.0019	.0085	.0089	.0065	.7568	.0256	.0215

Table 4.7: Potential error rate, the cross number means the percentage of activities that are mislabelled by the proposed method. In the table, A1:Sitting, A2:Standing, A3:Walking, A4:Running, A5:Stepping Up, A6:Stepping Down, A7:Falling, A8:Driving, A9:Eating/Drinking, A10:Sleeping/Napping. Note that since we are using binary classification, these figures do not need sum to 1 either horizontally and vertically. A higher figure indicated a greater potential for mislabelling. These figures are important as a guide to our future work.

cludes photos, tri-axial accelerometer readings, GPS, WiFi, bluetooth signals, and ambient environmental measurements such as temperature. This data can be either transferred to the server in real-time, uploaded on demand, or uploaded in bulk upon charging to save the battery power of the smartphone.

Our experience shows that personal activity recognition models with supervised machine learning model outperform naive approaches of setting thresholds. Therefore, we applied a supervised machine learning method to identify different activities automatically. The machine learning technique we employed was Support Vector Machine with a linear kernel. We have identified a set of the sixteen most enjoyable life activities as defined in previous work of Kahneman et. al. [88] and our software automatically identifies appropriate activities, chosen from this sixteen. The source data for training the SVM is the sensor stream data, segmented into events, and the user annotation of life activities, which they can do by identifying their current activities through the interface of the smartphone application.

4.6.2 Activity Recognition

For activity recognition evaluation, we used the DIAL dataset and DCULifelog-Dataset. We set one window to be 60 seconds long and 30 steps within one window for an adult normal walk. The frequency for one window is 30Hz,

Among 17 most commonly used classification methods, Random Forest and SVM are evaluated to be the most effective classifiers [90]. In this research, we explore the effectiveness of using these three menthols.

As shown in Table 4.8, most of elementary activities can be recognised at accuracy of over 90%, while some of them like laying has a very high accuracy as the time feature is also added into the feature set for training and predicting.

As shown in Table 4.9, most of combined activities can be recognised at a lower

Classifier	Activity	Accuracy(%)	Time taken (s)
RF	WALKING	95.12	5.23
	SITTING	95.85	3.6
	STANDING	88.56	3.54
	LAYING	99.56	2.76
	WALKING_UPSTAIRS	92.25	4.36
	WALKING_DOWNSTAIRS	89.58	3.25
	CYCLING	90.54	4.54
SVM	WALKING	95.12	5.09
	SITTING	93.34	5.06
	STANDING	98.56	3.25
	LAYING	99.56	2.33
	WALKING_UPSTAIRS	92.25	3.25
	WALKING_DOWNSTAIRS	89.58	3.05
	CYCLING	91.33	3.91
ANN	WALKING	94.25	5.12
	SITTING	92.51	5.11
	STANDING	93.25	3.21
	LAYING	99.44	2.56
	WALKING_UPSTAIRS	95.09	4.17
	WALKING_DOWNSTAIRS	88.22	3.91
	CYCLING	90.96	4.21

Table 4.8: Evaluation of activity recognition.

no.	**Activity**	**Accuracy(%)**		
		RF	SVM	ANN
1	Socializing	20.83	12.35	15.85
2	Relaxing	35.2	39.58	54.23
3	Pray/worship/meditate	46.25	33.65	40.21
4	Eating	59.8	62.23	45.26
5	Exercising	33.65	23.25	29.46
6	Working	85.66	86.58	88.16
7	Watching TV	25.33	25.65	33.21
8	Shopping	88.54	**86.59**	**95.33**
9	Preparing food	33.21	25.51	24.95
10	On the phone	13.25	18.68	29.56
11	Taking Care of my Children	23.25	12.35	18.5
12	Computer/Internet/Email	53.55	54.25	36.5
13	Housework	25.21	15.94	17.38
14	Working	68.56	78.54	74.46
15	Commuting	**94.25**	89.99	95.21
16	Napping	12.2	14.24	9.58

Table 4.9: Combined Activity Recognition

accuracy than elementary activity recognition as shown in Table 4.8. "Shopping" and "Commuting" are the two most detectable activities with the highest detection accuracy. The reason for this is that for "Shopping", GPS location for shopping centres are added into feature set; for "Commuting", driving feature is almost (81%) showing that it is commuting, also GPS feature also works for this detection.

4.6.3 Lifestyle of Activeness

One research topic we include in this chapter is the level-of-activity detection of personal lifestyle. Level-of-activity is also called *activeness* in this thesis. We use Equation 4.7 to detect the lifestyle of the subjects and use the factor as a index to present the activeness for personal health care research. Although this is not the main topic of the research of this thesis, but it is one of the research directions that ubiquitous computing can be applied in support of daily life, which is the main

Participant	Social Role	still(%)	Moderately Active(%)	Highly Active(%)
1	professor	66.25	14.24	19.51
2	final year student	80.33	11.08	8.59
3	first year student	40.98	35.25	23.77
4	finance department worker	39.51	33.09	27.4
5	assistant officer	36.25	54.24	9.51
6	part-time first year student	38.41	45.11	16.48

Table 4.10: Evaluation of activeness recognition

purpose and aim of this research.

The output of Equation 4.7 is a score. This score was calculated by giving one point for every second, threshold to three states (1) "still" (2) "moderately active" (3) "highly active". The result of activeness for 6 subjects are shown in the Table 4.10.

As shown in the Table 4.10, different people with different professions have significantly different lifestyle activeness. First year student has more active lifestyle than the final year student who had deadlines for theses. Also the age of group also affects the lifestyle activeness. As shown in the table, younger age group has more active lifestyle than the older group.

4.7 Conclusion

In this chapter, we first give our definition of human daily activities and later describe the activity recognition methods for sensor data. For activity recognition, we have done experiments on physical activity recognition (10 activities) and more work is to be done over 15 activity recognition. We also point out that we can also do lifestyle detection for optimizing activity recognition and for healthcare research in the future. We have also outlined the approach we are taking to development and evaluation, that will be employed many times throughout this research. In the following chapter, we will introduce how to apply the activity recognition in event segmentation.

Chapter 5

Visual Discovery from Lifelogs

5.1 Introduction

In our view of lifelogging, the data includes not only sensory data that provides contextual information about the activities of the individual, but also visual content that has previously proved to be efficient in human memory reminiscence [32, 56, 129]. These visual contents make up the vast majority of the lifelog data that we generate in our lifelog (up to 86% are images). However, the potential of lifelogging as an assistive technology is hampered by the complexity of visual content analysis [28]. From prior research, we know that people perceive daily activities in terms of discrete events [56] and in terms of lifelogging, this requires the creation of event segmentation algorithms that segment a continual stream of lifelog data into a set of discrete events. Events provide both a unit of access and a unit of retrieval for lifelog. In this chapter we will review past efforts at event segmentation as well as propose our approach to segment lifelog events by introducing features from both contextual and conceptual data from lifelog that is discussed in the previous chapter 4. This chapter also provides input for the linkage analysis in the Chapter **??**.

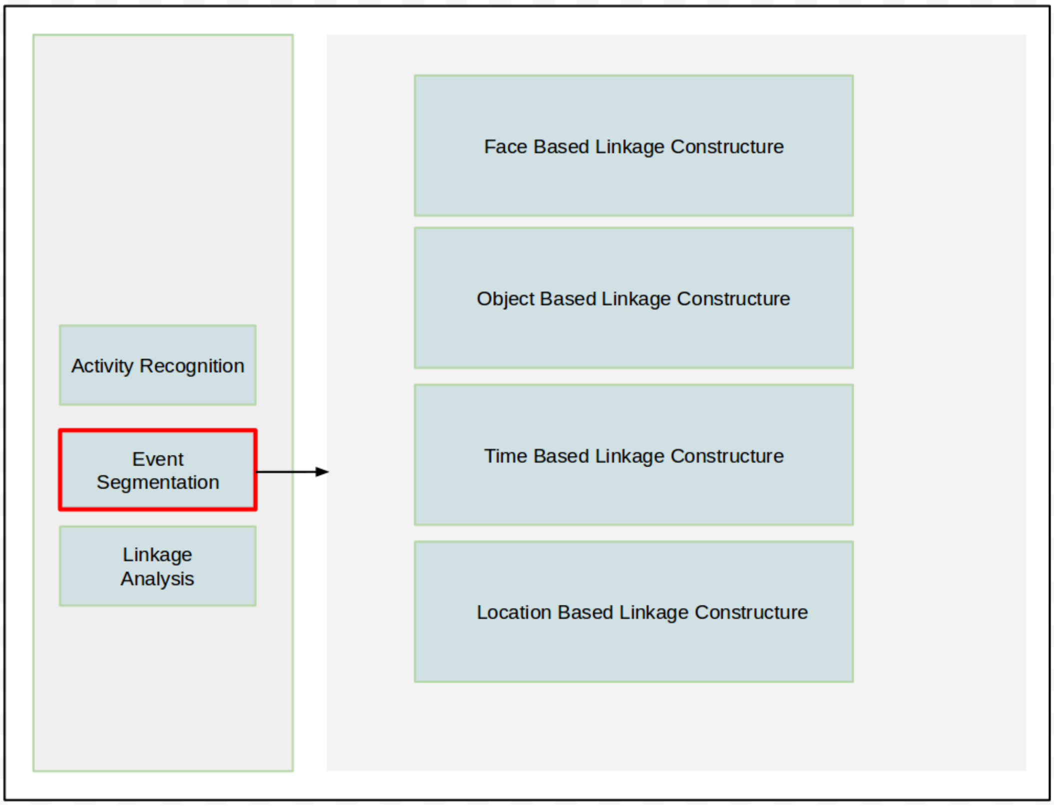


Figure 5.1: Overview of work in chapter 5

Event segmentation, in lifelogging research, means the automatic process of segmenting chronologically consecutive images and body/ambient sensors into daily life events. Researchers in the lifelog area have employed a number of methods for lifelogs event segmentation . Event Segmentation Theory proposes that event boundaries tend to occur when features of the situation being observed are changing the most, because these tend to be the times at which prediction errors increase [172].[1] Byrne et. al. validated everyday concept detection in visual lifelog and applied it to daily event segmentation [28]. Wang et. al. proposed the concept of visual lifelogging and brought up using Hidden Markov Model to analyse time series of lifelog images [161]. Doherty et. al. also has done a recap of continuing work in this trend [52].

The motivation for re-examining the event segmentation models in use today is that there is limited research undertaken on lifelog event segmentation and association for lifelog, yet the quality of the event segmentation has an impact on any holistic solution given that the event segmentation model typically generates the documents, the basic data types to be used in developing search and access methods for lifelog data. In the lifelogging community, it has been recognised by Lin et. al. that a visual lifelog should be segmented into shots/activities/events to make it manageable [111]. Events in lifelogs [52] are life episodes that composes people's life, like *working in the office*, *shopping with mum* etc. The concept *Event* differs from the concept *Activity* in the way that events are more scenario related while activities are more physical behaviour based. Event segmentation is important because it provides a basic unit for lifelog retrieval or linkage for any lifelog application, which is the research topic covered in the next chapter. Without event segmentation, a SenseCam wearer (or any other lifelogger) would have to examine

[1] http://www.apa.org/science/about/psa/2010/04/sci-brief.aspx

up to 5,000 photos per day to review past experiences. Indeed, the initial Microsoft SenseCam viewing software simply provided a frame-by-frame animated playback mechanism for accessing up to thousands of images per day. In this thesis, we provide a solution to the problem promised in the former part of this thesis by introducing lifelog event segmentation methods, events linkage mesh construction and retrieval models. Finally we depict our experimental system configuration and illustrate our experimental prototypes.

5.2 Visual Features for Lifelog Visual Content Analysis

Events, or lifelog data in general, can be made semantically enriched by detecting who is in the event (in the images), how many friends are in the images when and where the image was taken, etc. Visual features can provide the information extracted from lifelog photos. According to Cathal Gurrin, the lifelogger with first person experience of lifelogging, averagely, there will be around $1T$ of data collected by a lifelogger within a year. It would be headache if these data is not user-friendly organized and accessible to users. With the aim of providing all lifeloggers with a user-friendly lifelog service, there are various problems we should solve in this area:

1. How to segment events from a continuous lifelog stream?
2. What are the key factors that we can extract from visual images to represent to users?
3. With visual images and extracted content, in what format this content can be applied into lifelog linkage analysis and search?

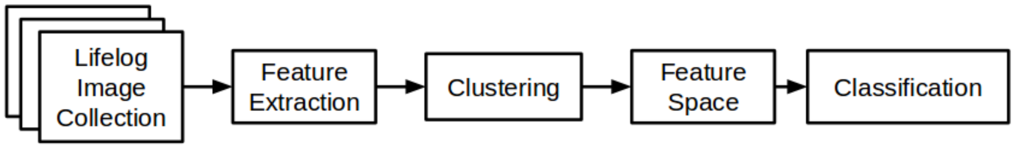


Figure 5.2: Process of lifelog visual discovery

This is also in accordance with the hypothesis 1.3 we propose in Chapter 1.

In order to address these three questions, there are two sources of data we can employ: images and sensor records. For images, an example is if it reminds us of some visual memory of past experience, we should be able to associate a flower with a garden or a friend who sent the flower. This means objects in the photos can trigger memory of past experience. Therefore, our work emphasis can be established onto an object extraction process for lifelog linkage analysis. This is the fundamental idea for the next chapter and the lifelog linkage analysis. In order to do so, we extract visual information such as objects in images as a starting point for visual content based event segmentation described in this chapter.

Automatic detection of concepts is already a very active area in multimedia research in general and is applied to video from broadcast TV, movie, TV news etc. as well as to still images. The Figure 5.2 shows the process of how we conduct our visual discovery for lifelog image data. Later in this section, we describe different visual content discovery algorithms for lifelog analysis and in the section of evaluation, we provide the results of comparing the different visual content discovery methods for lifelog data.

5.2.1 Bag of Visual Words

The *Bag of Key points* or *Bag of Visual Words* method is based on vector quantization of affine invariant descriptors of image patches. BoF is one of the popular visual descriptors used for visual data classification. Bag of Features (BoF) is inspired by a concept called Bag of Words that is used in document classification. A bag of words is a sparse vector of occurrence counts of words; that is, a sparse histogram over the vocabulary. In computer vision, a bag of visual words of features is a sparse vector of occurrence counts of a vocabulary of local image features.

BoF typically involves in two main steps. First step is obtaining the set of bags of features. This step is actually off-line process. We can obtain set of bags for particular features and then use them for creating BoF descriptor. The second step is we cluster the set of given features into the set of bags that we created in first step and then create the histogram taking the bags as the bins. This histogram can be used to classify the image or video frame.

In this research, we first extract opponentSIFT points for 279,230 images, each image gets $N(i)$ sift points; each opponent sift has 384 features. For different images, n is different. Heuristically speaking, the more complicated the image is, the more sift points can be extracted from images, n is larger. All images averagely contain 1,000 sift points, then 4 random sift points are taken out from each image in our 279,230 image datasets to compose a 1,116,920 sift points repository. This repository is used to cluster to get cluster center to construct visual vocabulary. In this research, we chose 1,000 cluster center as the visual vocabulary. Each word in this vocabulary contains 384 features in the same space as raw original images. Then every image is mapped into these 1,000 centres constructed space using Bag of Visual Word approach. With this method, all images with different number of sift points are reconstructed into 1,000 dimensional space. The value of each dimension for a image represents the visual occurrence of that center in the image. With the visual vocabulary, a key-frame for each image can be represented by a 1,000 dimensional feature vector, analogous to the bag-of-words representation of text documents. Our weighting schemes for bag of visual words image retrieval is *term frequency with normalization* [170] as shown in Equation 5.1.

$$weight = \frac{tf_i}{\sum_i tf_i} \tag{5.1}$$

5.2.2 Colour Histograms

Colour Histogram is widely applied to different visual features presentation. A colour histogram is a representation of the distribution of colours in an image. For digital images, a colour histogram represents the number of pixels that have colours in each of a fixed list of colour ranges, that span the image's colour space, the set of all possible colours. So each feature of one image is actually the closest point to cluster centres. This is a easy algorithm to implement although the process of matching to the closest cluster center can cause the problem of losing information. There are a number of approaches to colour extraction from images that can be considered.

5.2.3 Colour SIFT

Scale-Invariant Feature Transform (SIFT) bundles a feature detector and a feature descriptor. The detector extracts from an image a number of frames (attributed regions) in a way which is consistent with (some) variations of the illumination, viewpoint and other viewing conditions. The descriptor associates to the regions a signature which identifies their appearance compactly and robustly. In-depth description of the algorithm can be referred to [113] and tutorials [2]. We choose to use Colour SIFT because it is proved to be effective in detecting and describing local features in images.

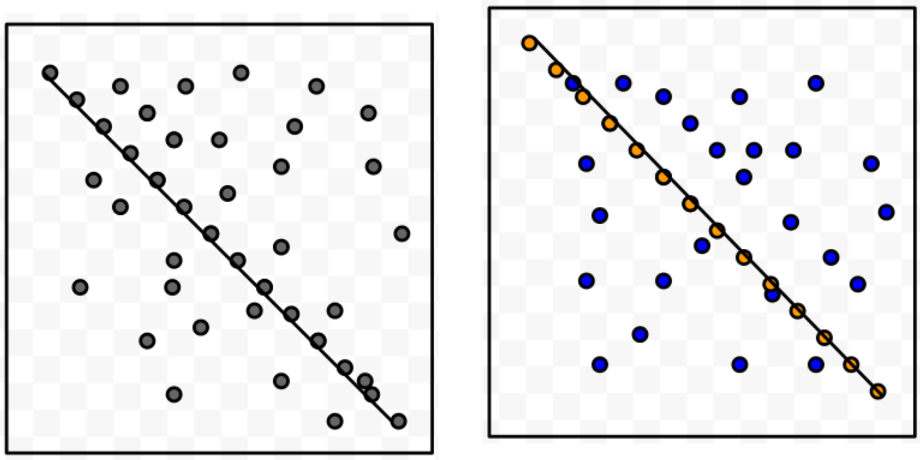
Event segmentation can be taken as a Image Category Recognition problem by analysing visual information on the level of objects and scene types. Colordescriptor [158, 157] is a open-source tool for extracting sift points for images[3] Each image will be described as a $128 * N$ matrix, where N is the number of SIFT feature points for that image. After getting all SIFTs for all images, we get a subset 10,000 images of all images to describe as the base space for decomposing and normalizing all images. Here we set subspace dimensions to be 1,000 dimensions to keep enough information. These 1,000 points are taken as cluster centres for all images. The reason why we choose 1,000 is due to the concern of speed and preciseness. As some of our experiments show, if dimension space grows by 5 times, the speed slows down by 25 times. And 1,000 dimension is precise enough for our experiments on concept detection.

5.2.4 SURF

Speeded Up Robust Features (SURF) is a robust local feature detector, first presented by Herbert Bay et. al. in 2006, that can be used in computer vision tasks like object recognition or 3D reconstruction [14]. SURF was inspired by SIFT but is claimed to be more robust than SIFT [14]. In this thesis, we detect SURF features using MATLAB detectSURFFeatures of MathWorks.

[2] http://www.vlfeat.org/overview/tut.html

[3] http://koen.me/research/colordescriptors/download



(a) Lifelog dataset with many outliers for which a line has to be fitted.

(b) Fitted line with RANSAC; outliers have no influence on the lifelog detection result.

Figure 5.3: RANSAC Approach Illustration

5.2.5 RANSAC

RANdom SAmple Consensus (RANSAC) algorithm proposed by Fischler and Bolles [66] is a general parameter estimation approach designed to cope with a large proportion of outliers in the input data. RANSAC is a re-sampling technique. It uses the least amount of data points to generate candidate solutions to estimate the underlying model parameters [49]. Imagine we have thousands of lifelog images and we want to find similar events that share with the same visual information like meeting with the same people or having dinner in the same restaurant. RANSAC approach is used in this thesis to find out outliers in the lifelog data. As shown in Figure 5.3, RANSAC can find out outliers without influencing the result.

5.3 Object Detection

One of precedent idea of discovering visual lifelog is based on object detection [159] . We assume that if two lifelog events share sufficient objects, then these two events should be higher similar than the events that does not share common objects in the visual lifelog. Here we train our machine learning model for different objects for lifelog image visual discovery, in our research, the objects that we detect for lifelog visual discovery are introduced in the following sub sections. These are the initial list for event segmentation. Later in this research, more objects are used.

5.3.1 Face Detection using Haar Cascades

Face detection is essential in lifelog research not only for a potential application purpose of finding people the lifelogger interacts with in lifelog, but also for ethical purposes of deleting all images with some specific people's faces if these people require for a deletion for a concern of privacy [123, 30]. It is mentioned that there is a dichotomy between lifelogger's ideal and actual behaviour based on current societal expectations and technology restrictions [30]. Face detection for lifelogging research is one part of technology exploration that narrows down that gap[4]. Our face detection techniques is based on object recognition work of Viola et. al. [159].

5.3.2 Social Activeness

Although there is some prior research on physical activeness discovery from lifelog [55, 74], there is very little discussion over social activeness of lifelog data analysis. We define *social activeness* as a measurement of how active a lifelogger is at a level of joining in social life with other individuals. Social activeness indicates to what extent the lifelogger is socially active. Intuitively, we define a new measure for social activities of a lifelogger over a period of time T to be the number of faces detected by lifelog device divided by the length of the period, as in Equation 5.2:

[4] http://docs.opencv.org/trunk/doc/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html#face-detection

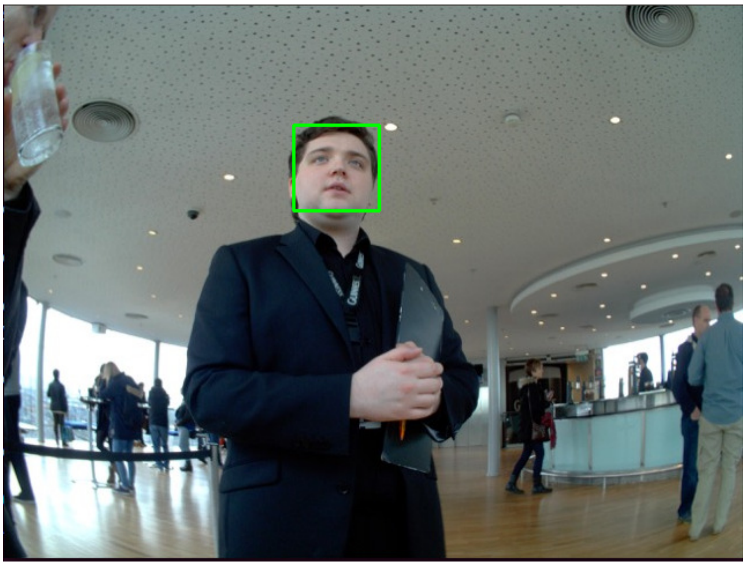


Figure 5.4: Face Detection Result

subject.	social role	no. of images	no. of faces	social activeness
1	professor	11258	482	4.28
2	final year PhD student	20284	626	3.09
3	first year PhD student	15105	238	1.58
4	finance department worker	3523	49	0.56
5	assistant officer	5105	65	0.77
6	part-time first year PhD student	3921	207	5.28

Table 5.1: Social activeness detection result

$$SA_T = \frac{\sum_0^T n_f}{|T|} \tag{5.2}$$

5.3.3 Object Detection in Lifelog Data

Object detection uses features and machine learning algorithms instead of human eyes and brain to find object in visionary scene, based on idea of replicating the process of the human brain. Object detection is the prerequisite for object based lifelog linkage constructing and retrieval.

We use Open Source Computer Vision Library (OpenCV) [23] for object detection task. OpenCV is the most used libraries in robotics for detection and to understand the objects captured by image sensors. We choose two key concepts that we detect from our lifelog data: “car” and “computer” to show our process of concept detection for lifelog visual data discovery. These two key concepts are chosen due to their ubiquity in our dataset. The concept ”Computer” is the most common concept in the dataset. ”Car” is the most common non-computer related concept.

1. Car detection

   In order to detect cars in our lifelog, we use UIUC Image Database for Car Detection [1] dataset for training our car detection model. This dataset has 1,050 training images (550 car and 500 non-car images), we also extract some car training images from our lifelog dataset, and this gives us 2,045 more car images and 513,459 non-car images.

   We employ boosted cascade of simple features based rapid object detection proposed by Viola et. al. [159]. Here, we used 1,050 car images for training to get a classifier, which a cascade of boosted classifiers working with haar-like features [167]. All training images are scaled to the same size, 48px width and 24px height.

   Here are steps how we get the training set and generate the cascade for car detection.

   (a) Get the positive training set

      find TrainImages/ -type f -name ’*pos*’ -printf “TrainImages/%f 1 0 0 100 40*

      n” >cars.info

(b) Get the negative training set

find TrainImages/ -type f -name '*neg*' -printf "TrainImages/%f n" >bg.txt*

(c) Get the vectors

opencv_createsamples -info cars.info -num 550 -w 48 -h 24 -vec cars.vec

(d) Check the vectors

opencv_createsamples -vec cars.vec -w 48 -h 24

(e) Train the cascade

opencv_traincascade -data data -vec cars.vec -bg bg.txt -numPos 500 -numNeg 500 -numStages 2 -w 48 -h 24 -featureType LBP

(f) Test a car using generated cascade

./cardetect –cascade=data/cascade.xml

By using car information, we would know whether a lifelogger is "walking along the road" or "walking towards his/her car". This is important when we are using this information for event and activeness detection. "walking towards his/her car" can mean the following event is highly likely to be driving.

2. Computer detection

   Computer relevant lifelog contents is a significant symbol of the individual in a work environment. We use SURF Homography detector as shown in Figure 5.5 to recognize computers in lifelog images. SURF Homography detector algorithm is widely applied in finding known objects in a complex image. As in lifelog visual dataset, we find that, most of images about lifeloggers life are repetitive as a matter of circulated life activities and events.

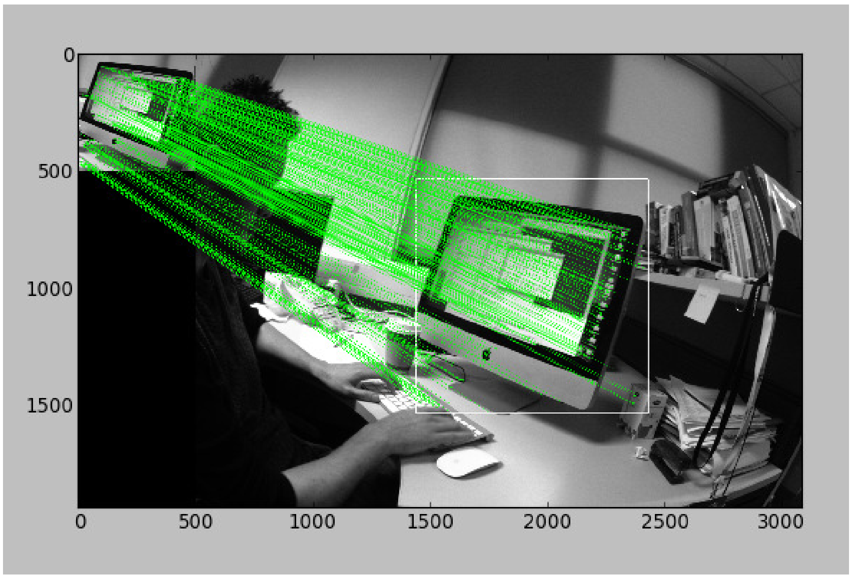


Figure 5.5: Computer detection from lifelog data using SURF Homography detector. A lifelog image when a lifelogger was talking with a colleague in the office in front of a computer.

5.4 Multi-modal Event Segmentation

In previous work, Doherty et. al. initially broke up sequences of SenseCam images into a series of chunks, where the boundary between these chunks corresponds to periods when the device has been turned off for at least 2 hours [52]. In this work, each image is then represented by MPEG-7 descriptor values and values from SenseCam sensors described earlier. The MPEG-7 descriptors we select for event segmentation are: colour layout, colour structure, scalable colour, and edge histogram [58, 59]. All these features are applied in three different types of event segmentation: visual event segmentation, contextual event segmentation and conceptual event segmentation.

5.4.1 Visual Event Segmentation

Our observation is that most consecutively similar images have length of shorter than 35, So we choose a window size 35 for comparing image's similarity.

1. Compare adjacent images (or blocks of images) against each other to determine how dissimilar they are.
2. Determine a threshold value whereby higher dissimilarity values indicate areas that are likely to be event boundaries. e.g. a boundary is more likely to occur at a time of significant visual or sensory change as opposed to when little change occurs.
3. Remove successive event boundaries that occur too close to each other.

The similarity between two images is calculated by Equation 5.3. Here, A_i means image A, and B_i means image B. n is the dimension of features, here we use 1,000.

$$similarity = \frac{\sum_{i=1}^{n} A_i * B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} * \sqrt{\sum_{i=1}^{n} B_i^2}} \tag{5.3}$$

5.4.2 Context Features for Event Segmentation

Context means information that is gathered by additional sensors, like accelerometer, WiFi and location etc. In addition to visual features that are introduced by the last section, the event segmentation algorithm in this section includes contextual information to facilitate event segmentation.

Contextual information includes location, infra-red signals, time of the day etc. In the experiments in the following sections, we combine the features from visual content (concepts) and contextual content by feature fusion [148] that is mentioned in Chapter 3.

In the experiment on evaluating activity classification, we carried out an assessment of our algorithm on data sets using both clean (correct) concept annotation and on concept annotation with errors.

5.4.3 Conceptual Features for Event Segmentation

Concepts that are discussed to be detected from images in previous sections in this chapter are also used to segment events. The idea is that if the same objects appear in consecutive images even their similarity is lower than the threshold, the images should also segmented into one event instead of being split into different events. This is based on the observation from our dataset that some images should be segmented into the same event as due to the scalability and orientation of images, their similarity might be massively different.

Concept features for lifelog analysis has not yet been established. In this research, we employ the same concepts in Wang's work [161], including *band1 basin basket bicycle book bottle bowl building bus car cashier cat chair child cloth clothes cooker cup cutlery cycle_lane dark deli drink face finger finger_touch food fridge fruit glass glove group1 hand hand_gesture hand_washing handle_bar hanger hanging_clothes indoor inside_bus inside_car keyboard kitchen laptop microwave milk mobile_phone monitor newspaper notebook office outdoor page_turning paper people pet phone_screen plastic_bag plate pram_buggy presentation projection projector remote_control road_path road_sign screen shelf shop sink sky soap steering_wheel table1 taking_notes toy traffic_light tree trolley TV vegetable water window yellow_pole* . This is the full list of all concepts.

Therefore, besides the space of visual features and context features, concept features are also included into the subspace of visual content, subspace of contextual information and subspace of concepts as shown in Equation 5.4.

$$image\_set = \begin{bmatrix} image_1 : v_1 & ... & v_n & c_1 & ... & c_n & cpt_1 & ... & cpt_n \\ image_. : . & . & . & . & . & . & . & . & . \\ image_n : v_1 & ... & v_n & c_1 & ... & c_n & cpt_1 & ... & cpt_n \end{bmatrix} \quad (5.4)$$

Here we normalize features from different sources (MPEG-7, accelerometer, light level, ambient temperature, and passive infra-red and concept of event) into a same metric space [5].

5.5 Experimental Set-up and Variables

5.5.1 System Design and Implementation

Visual concepts in lifelogging are visible facts of lives that demonstrate location, time and objects of life events. In this work I build on top of pre-existing event annotation tools as well as develop my own tools (specifically activity recognition) to support the panned research. In this process, images are annotated with activity attributions and all the annotated images with time stamps and other sensor data are taken into training session in the following experiments.

In the event segmentation, we evaluate the performance of our daily activity based event segmentation by comparing with prior work [59, 58]. This prior work is also the ground truth of our work on event segmentation.

In this part, we use the same test collection and evaluation approach as in the last chapter, annotated for activities and events. This will address hypothesis 1 in

[5] http://en.wikipedia.org/wiki/Feature_scaling

participant	events	detected events	correct	**precision(%)**		
				v	cont.	conc.
1	556	687	325	32.54	33.25	58.45
2	1025	1456	857	46.54	75.32	83.6
3	786	845	348	41.42	46.31	44.27
4	159	201	82	25.48	42.56	51.57
5	214	321	98	49.55	42.54	45.79
6	145	278	86	15.26	31.23	35.1

Table 5.2: Conceptual event segmentation result

our research plan.

We use the user annotation method for activity based event segmentation algorithm, comparing with images based event segmentation algorithm accuracy 29.2% [58].

In this section, we describe how we evaluate our approaches for lifelog visual discovery and analyse our results from a real-life application.

This section now discusses the evaluation we carried out on augmenting lifelog events with images/videos from other sources of information. We extensively evaluate the results using collected datasets.

5.5.2 Result of Event Segmentation Approaches

Table 5.2 shows the precision of different event segmentation approaches (visual, contextual and conceptual) over the data from 7 different participants. As shown in the table, the highest detection rate is for lifelog 2, who is a PhD student having a regular lab/home life for this detection period of time. The lowest rate is for lifelog 6, which is partially due to lack of sufficient training data.

The results shown in Figure 5.6 present the results from different approaches for different participants' data. As shown in this figure, the conceptual event segmentation approach outperforms the other two approaches of visual and contextual

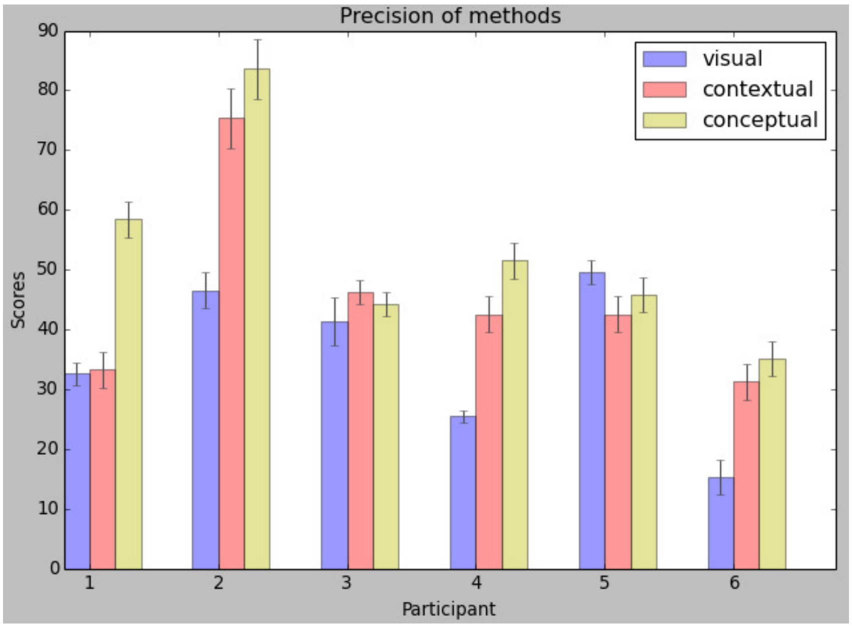


Figure 5.6: Event segmentation results of different approaches for different participants. Confidence interval value is set to be 0.95.

information based event segmentation. The higher the value is, the better the feature set for event segmentations are. This means using concepts for event segmentation is the best practice for lifelog event segmentation. There is still scope for extending this work in future.

Figure 5.6 also shows that for the participant 3's data, the contextual event segmentation approach outperform the other two approaches. The reason for this should be that the data from the participant 3 himself as as we know shown in our forms of participants, the lifestyle of participant 3 highly lies in cycling, driving and sitting in the most of cases (65%). The images for this participant has a high relation with the activities/contextual information.

For the participant 5, due to lack of diversity of different events, as participant had a big concern of privacy, most of data collected by the participant 5 are about work. Therefore, the work environment is quite dominant (35.65%) in this sub-dataset. The concept space for this participant's data is sparse due to lack of diversity. There is why even just with image visual features, the performance of event segmentation is also good enough comparing fusing more features like concepts and contextual information.

5.6 Conclusion

In this chapter, we introduce the visual discovery of lifelog using visual image analysis methods like SIFT, SURF, color histogram, face detection etc. By using feature fusion methods, we combine 3 different type of features, including visual content, contextual information and concepts into the experiment of evaluating different approaches for lifelog event segmentation. The performance checking evaluation shows that the overall performance of the feature fusion of three different spacial information is the best for lifelog event segmentation. We explored how we use the activity features that are introduced in the previous chapter for event segmentation to support lifelog retrieval in MemoryMesh.

Chapter 6

Lifelog Linkage Analysis with Event Modelling and Retrieval

Data in lifelogs are heterogeneous, which eliminates available approaches to efficiently organize lifelog. Linkage provides a perspective of looking over lifelog data as an event-based data graph. The process of establishing lifelog connections is also called relation extraction. Relation extraction from lifelogs is about extracting semantic relations between entities in lifelogs, like activities and events that are discussed in the previous chapters. In this chapter, we provide an overview of linkage analysis for lifelogs and explore possible linkage analysis methods of lifelog research including extracting relation between lifelog data (specifically lifelog events) based on the attributes of these events.

The limitation of *ShareDay* and *ZhiWo*, along with all other lifelog retrieval systems developed to date, lies in the fact that there is no linkage analysis between activities and events, also these systems are temporary not permanent in terms of organizing lifelogs in a searchable manner. This motivates us to think deeply over

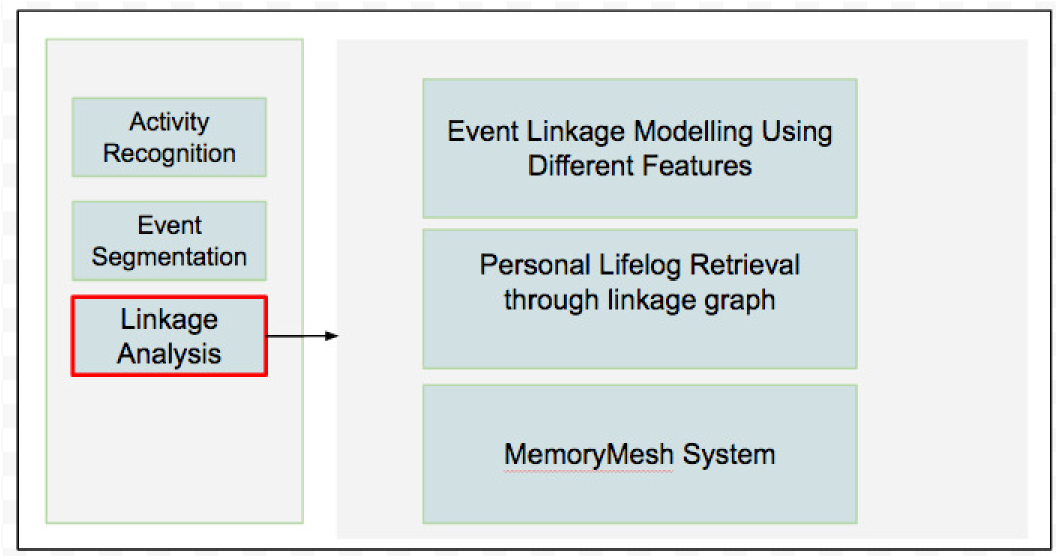


Figure 6.1: Hierarchy of work in chapter **??**

how to link lifelog data dynamically, which we call the MemoryMesh in this thesis.

In this chapter, we also present the MemoryMesh and identify how it can be used to support lifelog search.

6.1 Introduction

The idea of MemoryMesh comes from human associative memory. The potential of associative memory can be expanded through Memory Maps. Memory Maps are like a Mind Map as described by Buzans and Harrison [27]. This chapter uses WWW retrieval models to mimic the associative memory in cognitive science to build up a MemoryMesh for lifelog linkage analysis and retrieval.

The SenseCam (as mentioned in Chapter 2) introduced the research community to the potential of wearable cameras as lifelogging tools to gather a media rich lifelog for individuals and groups. We refer to lifelogging as "a form of pervasive computing, consisting of a unified digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive", as used by Kitcher and Dodge [51]. With considering this definition, we focus our work in this thesis on developing data structures to support lifelog archives by exploring linked-data potentials for the multimedia content that exists in the archive.

As stated previously, a lifelog can consist of more than 5,000 photos daily, along with hundreds of times more sensor readings. Very quickly, such a lifelog becomes too large to browse. Therefore, it becomes necessary to organize this data and to support search and retrieval for the purposes of easy accessibly and usability of lifelog data for lifeloggers. Doherty et. al. identified the 'event' as a suitable atomic unit of retrieval and proposed automatic segmentation of lifelog data into events, made accessible through a browsing interface [58]. However it was found that 75% of browsing effort fails to find a known event from a large lifelog [54]. Adding a search facility over automatic (sensor-based) event annotations reduced the failure-to-find error rate to 25% and the search-time by a factor of ten. In the MyLifeBits project [71], a database search mechanism is provided, which was shown to be effective at locating nuggets of information from Bell's extensive archives [176].

Both Doherty and Bell & Gemmel's work show the potential of search interfaces to lifelog for supporting a user with an information need. However, they assume that the information need is simply that, and information need, without giving much consideration to the types of information need. Given that searching a life experience archive is a new activity, to simply assume one type of information need is not the best way to proceed, in our experience. In order to understand the potential information needs, Sellen & Whittaker [89], as a guide to future development of lifelogging technologies, have identified the five reasons why people would access their memories, and by association, their lifelog. The five R's of memory access are

Recollecting, Reminiscence, Retrieving, Reflecting and Remembering Intentions.

While a flat database or text-index based representation of lifelog events can support the five R's retrieval [142], it is our conjecture that a better organization structure and data access methodology will support more efficient and effective lifelog retrieval and will better serve to provide retrieval facilities that address some (or all) of the 5 Rs, which we see as use cases of lifelogs.

Hence, we propose the MemoryMesh, which draws knowledge from WWW search, coupled with cognitive psychology, to develop a novel lifelog index structure that models the lifelog as a densely linked hypermedia archive. This allows for the application of new types of information retrieval concepts such as the WWW-inspired PageRank algorithm and supports multi-faceted browsing through the lifelog, supporting many use-cases. We now discuss the MemoryMesh, in terms of its construction and potential for enhanced interaction with lifelog. At the same time, we review what we mean by linkage in lifelogs and take it as our basis for linkage evaluation.

6.1.1 Fundamental Principles for Building Memory Mapping

Every happening in life can be memorized and these happenings generally link together through memory associations. Coggle [1] is a software about building information sharing through information associations. It has been evidenced that images captured by wearable devices can be utilized to facilitate people's ability to connect to their past, and these images do this in various ways [142]. Memory mapping is the representation of how we conceive of, and make claims about lifelog linkage analysis. A mind mapping generated using Coggle is shown in Figure 6.2. Episodic memories or autobiographical memories [43, 155] record a trace of life experience

[1] https://coggle.it/

(often in the form of visual imagery) with conceptual context. Episodic elements or properties of past events or happenings: times, places, associated people and emotions, and other contextual information can be stored in our memories. They allow people to figuratively travel back in time to remember all the context of the event. In order to utilize lifelogs to enhance people's episodic memory, we stick to the following fundamentals to build up memory mapping for MemoryMesh in this research thesis:

1. *Time and Space Accuracy*

   The time and location of events which are elements of the MemoryMesh should be similar to the real-life happenings, which ensures the good match of human memory and contents in the MemoryMesh.

2. *Contextual Information are Assisting to Events*

   Contextual information can assist human memory of events and experiences. Some people can listen to music at the time of finishing some tasks. So later, when the same music is playing, the task could be vivid in memory due to the musical contextual information. In our lifelog, due to privacy issues, we did not collect sound, but other contextual information like time, location, people around can be triggers of human episodic memory.

3. *Visual Content Matching*

   Visual content is a ground-truth match for linking similar items or events as per their location and environment. Visual content provides information like indoor/outdoor, people are we are seeing, event we are doing etc. In this thesis, we combine the similarity of images with other associated attributes to build up mapping for lifelog linkage.

Figure 6.2: Mind mapping of linked events in one's daily life

We conceive of fundamental principles for building memory mapping as a guideline for us to construct a lifelog linked graph, as in Figure 6.2. When considering constructing lifelog linkage for lifelog retrieval and analysis, these principles guide us on what needs to be considered into lifelog linkage modelling.

6.1.2 Guidance Consideration for Linkage Analysis

When we think of establishing linkage analysis for lifelog data, we need to consider two main factors: nodes and connections (or links). What are nodes? The first thing to understand about lifelog linkage analysis is what should be linked in lifelog. Nodes in lifelog analysis in this chapter are lifelog events that we segment using the approaches introduced and evaluated in Chapter 5. What is linkage? This is what we try to solve in this chapter. Before we introduce the linkage model we build for lifelog, we list a few items we consider when we are trying to build up the linkage graph for lifelog data.

1. **Picture**

   When we recall memories, some memories are easier to recall than the others. Pictures give a trigger that sparks old memories; they act as memory cues. Non-changed pictures can record what really happened in our lives and what we experienced in our past life. One way of automatically utilizing images as a source of building up our lifelog linkage for digital memory is to try to detect "content" of images by employing image processing technologies like concept detection, as has been discussed in the previous chapter.

2. **Events**

   Events are the key factor for linkage analysis because they act as the unit of linkage and retrieval. The main problem here is how to model events into a linkage space, which means what should be used to represent an event. As introduced in the previous chapter, we used visual, conceptual and contextual information of activities to represent activities for event segmentation, in this chapter, we explore using more event attributes to annotate events automatically. This is introduced in the following sections.

3. **Connections**

   A lot of the brain's work is based on association and it automatically links different subjects together to create a system. This is like Mind Map shown in Figure 6.2. Mind Map has been successfully implemented in a few products like MindJet[2], FreeMind[3] and XMind[4]. It has been shown that it can accelerate your learning capacity by helping people instantly see connections and links between different subjects [47]. Mind Mapping is a guidance for

[2]http://www.mindjet.com/
[3]http://freemind.sourceforge.net/wiki/index.php/Main_Page
[4]http://www.xmind.net/

us to build up linkage for lifelog that is mainly aiming enhance human digital memory.

6.2 Event Linkage Modelling

In a perspective of life of individuals, we can model various experiences (events) as nodes in a connected graph with potential links between different experiences as shown in Figure 1.1. Different colours of nodes represent different life events, and the connection line between nodes indicates the strength of connections between different events. The strength of different nodes shows to what extent these two events are connected. Some events can be strongly and massively connected with a lot of nodes, these events can be called core events of a specific episodic memory. Some nodes can be alienated, and it means that this event can hardly be associated to other events, which may be very novel.

Following the development and evaluation of the new approaches to event segmentation, this thesis turns attention to the creation of the event linkage model that is addressing the Hypothesis 2 outlined in the introduction of this thesis. Our own human memories are heavily dependent on a form of event segmentation (episodic memory) and event linkage (associative memory). Hence our linkage model will form a connected graph, see Figure 1.1. This is also the construction of the MemoryMesh. As new events will be added into the MemoryMesh, the MemoryMesh is designed to be a self-updating system with adaptiveness to the updated data.

We refer to this as the MemoryMesh because it is modelled as a linkage graph (like the WWW or in Figure 1.1) with nodes representing events and edges representing associative links between the events. It is this modelling as a graph that allows for the application of WWW linkage analysis models to lifelog linkage analysis.

There is no similar prior work in this area. All prior work has considered the lifelog to be composed of a discrete set of events and relies on database and search access to locate or rank desired events in response to a humans information need. This work is different because, as we mentioned in the introduction, we are modelling the lifelog as a linked graph, a linked hypermedia and this opens up new opportunities for exploration of new models of lifelog retrieval, as we are about to describe.

In order to build this MemoryMesh of linked events in lifelog, we explore the features of different events and build relations between different events according to their shared features, such as time, location, people or objects (objects extracted automatically from images using existing tool sets, based on SVMs). The strength of the relationship is initially assumed to be determined by the number of features shared, which means the more features that two events share, the stronger relationship that two events have. This is in effect an implementation of the *TF* document similarity from text retrieval [138, 115]. The limitation with such a model is that it requires a threshold and a distribution of event links employed in this work and this is the subject of future work. For example, it may be appropriate to engage a learning to rank [39] process, in which a pre-training phase allows us through a user study to identify what makes a good MemoryMesh link, and then train the retrieval model that generates the ranked list of similar events based on the user input. There is a body of research to be done also to identify the required linkage type (i.e. one to one or more likely, one to many), but it is not something that we focused on in this research. Such issues have been observed and considered when we were developing and evaluating the MemoryMesh.

In naturally occurring phenomena, such as height of buildings, distribution of wealth, links among WWW pages, a distribution called a Pareto (or power-law) distribution is observed. It is our expectation that we also observe such a distribution on the developed MemoryMesh. In addition, we evaluate if a concept such as Inverse Occurrence Frequency aids in the creation of the MemoryMesh. Inverse Occurrence Frequency is a concept that we are considering that is based on the *IDF* ranking feature from information retrieval [115], that models the importance of a document (event) to a collection as a whole (the MemoryMesh) based on the uniqueness of the concept. Following MemoryMesh construction, we have developed the first purpose-built browsable lifelog archive. In order to evaluate the usefulness of the MemoryMesh as a source of data for searching over, we need to develop and evaluate retrieval models, as described below and detailed later.

6.2.1 Context Based Linkage Analysis

Context information, such as life location and environment etc. would be considered as important aspects of memory. In this thesis, contextual information about lifelog are extracted from sensor data like WiFi, GPS, PIR etc. This information provides a matrix of measuring how the event is wrapped up by a context that has potentials of associating this event with any other events that may share similar contextual information.

6.2.2 Face Based Linkage Generation

Face detection helps lifeloggers to search for specific digital memory about people that showed up in his/her life before. Hence, faces (detected using the approach previously described) is one of the sources of links between nodes in the MemoryMesh. If two nodes are shown to share faces, then the two nodes have a link generated between them.

Lifelogger	No. of Images	Average Faces per Event
1	67927	56
2	26806	58
3	24623	28
4	8038	23
5	24441	58
6	3921	19

Table 6.1: Face detection for lifeloggers and their evaluation on faces per event

6.2.3 Concept Based Linkage Generation

Objects can be triggers for human memory system. It is not uncommon that one person recalls of his or her memory of the moment when he or she got a birthday gift toy when he or she sees the toy again after many years. Based on this phenomenon, we have reasons to examine how objects that can be revealed or detected from one's lifelog image repositories influence human memory retrieval process.

6.2.4 Lifelog Retrieval Using WWW Retrieval Methods

Concepts provide a clue for lifelog exploration in terms of enriching lifelog presentation spaces with semantic concepts [161]. After the construction of the MemoryMesh, there exists a linkage graph of events. It now becomes necessary to further address the second hypothesis by applying semantic graph algorithms (inspired by WWW search algorithms) that show the benefit of modelling the lifelog as a MemoryMesh. Here we describe our linkage models as eigenvector methods those are used widely in Internet web data analysis. In our research, we explore the performance of these algorithms and models for life experience linkage modelling. We compare for search and browsing against a ground truth of a *TF-IDF* text search implementation over the MemoryMesh.

1. **TF-IDF**

   TF-IDF is a basic approach for content based information retrieval [139]. *TF* refers to *Term Frequency* while *IDF* represents *Inverse Document Frequency*. The value of *TF-IDF* reflects how important a word is to a document in a collection or corpus. It increases proportionally to the number of times a word appears in the document, but is offset by the frequency of the word in the corpus, which helps to control for the fact that some words are generally more common than others. In our MemoryMesh system, all events are viewed as documents, while attributes of these events are taken as terms that form a document that represents the event for a search engine (in a similar manner to how Google Image search would work). In this case, we can apply the *TF-IDF* method in calculating the similarity of events and finding the most relevant events. This method is the baseline of our MemoryMesh linkage graph model and provide, both a baseline, as well as a user text querying system..

2. **HITS**

   Hyper-link-Induced Topic Search (HITS) (also known as hubs and authorities) is a link analysis algorithm that rates web pages in terms of their authoritativeness, or link to authoritativeness , developed by Jon Kleinberg [97]. It was a precursor to PageRank [124]. The idea behind Hubs and Authorities stemmed from a particular insight into the creation of web pages when the Internet was originally forming [139]; that is, certain web pages, known as hubs, served as large directories that were not actually authoritative in the information that it held, but were used as compilations of a broad catalogue of information that led users directly to other authoritative pages. In our experience linkage model Event HITS, a good hub represents a event that is connected to other events, and a good authority event represents a event that is linked by many different hubs/event [115]. To begin the ranking, $\forall p$, $auth(p) = 1$ and $hub(p) = 1$. We consider two types of updates: Authority Update Rule and Hub Update Rule. In order to calculate the hub/authority scores of each node, repeated iterations of the Authority Update Rule and the Hub Update Rule are applied. A k-step application of the Hub-Authority algorithm entails applying for *k* times first the Authority Update Rule and then the Hub Update Rule. In Authority Update:

$$\forall p, auth(p) = \sum hub(i) \tag{6.1}$$

In Hub Update:

$$\forall p, hub(p) = \sum auth(i) \tag{6.2}$$

This can also be achieved by examining the eigenvectors associated with the top ranked positive and negative eigenvalues. Our conjecture is that applying a HITS style algorithm to the MemoryMesh allows the identification of the most densely linked events from the MemoryMesh; the most important clusters of events. From Cognitive Science, we would assume that these are the most likely events to be merged. One additional positive feature of the HITS algorithm is that it is ideally suited to retrieval time processing, hence it is naturally scalable as the MemoryMesh grows over time. The other two algorithms mentioned below are typically processed pre-retrieval over the entire linkage graph, hence they are not ideal for smooth scaling. In our linkage analysis, all events are taken as nodes of networked MemoryMesh. In Equation 6.1 and 6.2, all nodes are assigned with the value of authority and hub. Authority $auth$ stands for the importance/novelty of the events in MemoryMesh while hub hub represents the ability of association of events in MemoryMesh. In this algorithm, we not only can get the most novel events (highest authority value) but also most most associative memorable events (the highest hub value).

3. **PageRank**

   From the perspective of constructing linkage between items, elements or entity in a system or organization, linkage in MemoryMesh is similar to link structure in the Web, which consists tremendous web pages, as well as forward links and back links [124]. PageRank is a link analysis algorithm proposed by the two founders of Google, Page and Brin in 1999 [124]. It is considered that it was named after Larry Page and used by the Google Internet search engine, that assigns a numerical weighting to each element of a hyper-linked set of documents, such as WWW, with the purpose of "measuring" its relative importance within the set. The algorithm may be applied to any collection of entities with reciprocal quotations and references. In event PageRank, any event is assigned with a score of significance according to the importance of the moment to the life. And connection between events is valued according to the features shared by two events.

$$PR(p_i) = \frac{1-d}{N} + d \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)} \tag{6.3}$$

Applying a PageRank style algorithm over the MemoryMesh allows for the immediate identification of the most commonly reoccurring events and event types from the MemoryMesh, as the PageRank algorithm on WWW allows for the identification of the most important WWW pages independent of any information need. In MemoryMesh, PageRank also allows users to identify the most “authoritative” events in their life. This is due to the independence of PageRank results to the query set itself. However, we use personalized PageRank algorithm. We initialize the events weights by user specified value of life importance, or by a query into the lifelog system, or via a sample event that acts as a query into the system (for the purposes of more-like-this search or for evaluation), as in this work.

In any case, we can prime the initial state of linked graph for MemoryMesh based on an information need. In Equation 6.3, p_i and p_j stand for different events in the MemoryMesh. This iteration process calculates the importance of event nodes by the “in” linkages. $M(p_i)$ is the set of event nodes that links into the event node $p_(i)$, and it is also called outbound set, while the inbound set $L(p_j)$ includes all event nodes that are linked from the event node p_j. The PR value of nodes shows the popularity of event nodes, which is the ability of being associated from any other linked events. d is called damping factor in the MemoryMesh, as in our thoughts, the owner of an MemoryMesh structure can stop the association at some stage. N is the number of events in the MemoryMesh.

4. **SALSA**

SALSA algorithm was proposed by Lempel [108] in 2001. It is a stochastic approach for link-structure analysis, which examines random walks on graphs derived from the link-structure [108, 144]. Both SALSA and Kleinberg‘s Mutual Reinforcement approach employ the same meta-algorithm. SALSA is equivalent to a weighted in-degree analysis of the link-structure of WWW sub-graphs, making it computationally more efficient than the Mutual Reinforcement approach [108].

(1) The hub matrix H, defined as follows:

$$h_{(i,j)} = \sum_{k:k\in(i_h,k_a),(j_h,k_a)\varepsilon G} \frac{1}{deg(i_h)} * \frac{1}{deg(k_a)} \tag{6.4}$$

(2) The authority matrix A, defined as follows:

$$a_{(i,j)} = \sum_{k:k\in(k_h,i_a),(k_h,j_a)\varepsilon G} \frac{1}{deg(i_a)} * \frac{1}{deg(k_h)} \tag{6.5}$$

Applying SALSA allows us to evaluate an alternative approach to the two common WWW linkage analysis techniques just described. SALSA is similar in many ways to both HITS and PageRank, and as such allows for the identification of the most important, commonly occurring events from the MemoryMesh. The linkage between two event node in MemoryMesh can be viewed as an endorsement. SALSA is a query dependent ranking algorithm. That means, whenever there is a query, the computation of ranking should be calculated at query time. In our experiments, we initiate hub and authority scores with the similarity of any two events.

Similarly to HITS algorithm, this algorithm computes an authority and a hub value for each event nodes in the MemoryMesh graph, and these values can be viewed as the principal eigenvectors of two matrices. HITS uses straight adjacency matrix, SALSA gets the values according to in and out degrees of event nodes. In formula 6.4 and 6.5, $h_{(i,j)}$ is the transition probability of the event node i to the event node j in the hub graph. The event node k points to i and j in the implication, $deg(i_h)$ and $deg(k_h)$ are hub degree of nodes *i* and *k*, while $deg(i_a)$ and $deg(k_a)$ are authority degree of nodes *i* and *k*.

However, unlike internet web page search, in which the web page weight is computed according to its interrelation with other web pages, the significance initialization of the MemoryMesh relies on the importance of the events to specific users, based in their different life experience patterns, or information needs, which means that for different users and at different times, the graph is very different. This is a concept that we propose and evaluate n this research, and drives us to consider real-world evaluations of this research, beyond the test collection evaluations that have been considered thus far. There is significant opportunity for future work in this area.

6.3 Personal MemoryMesh and Lifelog Retrieval

In information retrieval, many researchers are dedicated to exploring methods for ad hoc query understanding or parsing. Same in lifelog retrieval, this is also a big issue as currently, many of lifeloggers like elderly people, ad hoc queries could be much more easier for use.

Here we extend our approach for MemoryMesh constructing to N users instead of just one user. In this thesis, N is equal to 7 which is subject to the data set collection policies over this thesis. We conduct the same qualitative analysis as one user over N users to check the adaptability of our approach to population.

6.3.1 Review of MemoryMesh Rationale

We propose that a densely linked hypermedia is a more suitable lifelog data structure than a flat database-based organisation. In effect we consider the lifelog to be more like a WWW structure, based on documents (events) and links. We know from prior research that exploiting the linkage structure of the WWW allowed the PageRank [124] algorithm to significantly enhance the effectiveness of large-scale information retrieval on the WWW. PageRank was deployed in the Google search engine and was considered an integral part of the ranking process. In the case of lifelogs, we propose that modelling the lifelog as a linked data archive (as is done in the human memory system) will bring similar benefits, which will not only support more efficient and effective retrieval, but also better support real-time user interaction with lifelog and the five R's of memory access. By applying PageRank-style algorithms, the MemoryMesh will know the importance, novelty of events and will pre-calculate the links between them, thus making retrieval more efficient and effective.

6.3.2 MemoryMesh Construction Through Lifelog Linkage Analysis

On the WWW, the links between web pages are pre-existing and created by web page authors as they create new websites and web pages. Consequently, the WWW grows organically and algorithm can mine the latent qualitative judgements inherent in each WWW link. This enhances the effectiveness of search algorithms.

Since we propose that a lifelog is a densely linked hypermedia archive, we can model it as a graph. From mathematics, we know that a graph is an ordered pair of $G = (V, E)$ comprising a set V of nodes together with a set E of edges, which are

2-element subsets of V (i.e. connect two nodes). In the MemoryMesh, the graph is a representation of a set V of events where some pairs of events are symmetrically connected by edges (the set E). An edge e is a link that is created between events that are considered sufficiently similar. Each event is represented by a semantic annotation (typically from wearable sensors as in [54]), which forms the content for both query/retrieval and for linkage generation within the MemoryMesh. Given a set of events V, the strength of edges between the events are calculated and the most appropriate N edges are inserted into the graph. There are many methods for selecting N edges. Here we present three example methods:

- Visual similarity between SenseCam images from one event and SenseCam images in other events. Either low-level (e.g. SIFT or regionalised colour) or high-level (visual objects co-occurrence) visual similarity can be calculated and can inform the strength of links, with links above a threshold bring selected,
- Multi-axes similarity sources from lifelog meta-data, such as events at the same location, time, noise level, actors involved, and many other sources of linkage evidence,
- External sources of similarity, by looking to semantic-web-style external sources of semantic data to identify real-world links between events.

In addition, the linkage model could be single or multi-layer. A single-layer mesh would allow for a single link between events whereas a multi-layer mesh would allow for multi-faceted browsing, which would provide additional flexibility to the MemoryMesh. Regardless of the layering, it is likely that a proposed linkage distribution would need to be adhered to when calculating a meaningful value for

N, across single or multiple layers. As new events are continually inserted into the lifelog, the MemoryMesh would need to re-calculate linkages either dynamically or periodically.

6.3.3 Lifelog Retrieval in MemoryMesh

Having the links between events in the MemoryMesh provides a number of benefits. Firstly, it supports real-time browsing for reminiscence and reflection (two of the 5Rs), without the need to continually, dynamically generate links to related events. Secondly, the rich linked hypermedia allows for a user browsing session to be targeted into the best region of the lifelog to begin a linked reminiscence session or provided guided reminiscence tours. In addition, the pre-calculation of multi-layer linkages would make the MemoryMesh more flexible to new and novel use-cases. Finally, links between events in the MemoryMesh provides an additional source of evidence when retrieving information from the lifelog, which would allow for the application of common WWW algorithms such as HITS and PageRank, which have shown useful on the WWW.

Applying a HITS [98] style algorithm to the MemoryMesh has potential to identify the most important user-context/query related events from the MemoryMesh. Indeed exploring the top non-principal eigenvectors from the MemoryMesh (applying HITs techniques) could help to identify clusters of similar experience from within the lifelog. Employing PageRank could support the efficient selection of both mundane and novel events in the lifelog. As lifelogs increase in size, identifying novel events becomes increasingly important.

Enhancing the MemoryMesh to incorporate a multi-level event linkage model, it becomes possible to retrieve from, and browse through, the lifelog using various criteria, such as exploring the lifelog via user activity links, behavioural similarity,

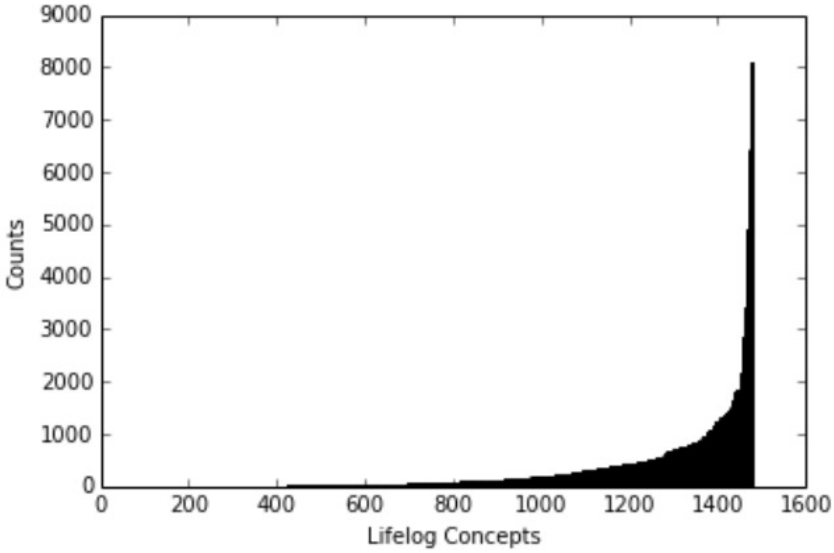


Figure 6.3: Concepts distribution for lifelogger 1. *The graph is generated using Python 2.7 and Matplotlib 1.4.3*

.

co-occurrence of similar objects, similar environmental context, or even the colour or textures of the SenseCam images grouped into events. It is our conjecture that moving from the flat collection of annotated events into a multi-dimensional linked hypermedia helps to generate a more useful lifelog.

6.4 Experimental Set-up and Variables

6.4.1 Deep Learning for Concepts

In order to enhance the concepts that we can employ for annotation, we use 1,701 concepts for image concept detection by enlisting help from one open-source deep learning tool called Caffe [85].

No.	Concept	Occur.	Concept	Occur.
1	laptop	8094	laptop computer	3036
2	notebook	7060	laptop	3036
3	notebook computer	7060	notebook	2426
4	computer mouse	6639	notebook computer	2426
5	mouse	6639	monitor	2320
6	monitor	6447	CRT screen	2223
7	printer	5852	screen	2223
8	CRT screen	4944	home	1581
9	screen	4944	home theatre	1581
10	desktop	4894	desktop	1488

Table 6.2: Top 10 lifelog concepts for lifelogger 1 (left) and lifelogger 3 (right)

As per Table 6.2, it is very appealing but unsurprising to see how much lifelog concepts of lifelogger 1 are related to office work as the real life of lifelogger almost spent more than 40 hours a week working in front of computer. Comparably, lifelogger 3 is also a person who works with computers/laptops much, but instead, lifelogger 3 spends more time at home than lifelogger 1, which also represents the reality. Therefore, we can conclude from the concept distribution that lifestyle of lifeloggers can be detected based on image concepts excluding traditional accelerometer data.

6.4.2 Significance Test for Different Lifeloggers

In order to test whether our approach for lifelog retrieval is applicable to all lifeloggers in our dataset, we need to test the significance over different data collected by different lifeloggers. Statistical significance indicates whether or not the difference between two lifelog groups that can most likely reflect difference between them in a real world. Here we take one set of data collected by one lifelogger as a group of sample. We also use t-test to see whether the mean of our concept samples differs in a statistically significant way from the theoretical expectation. We set *null*

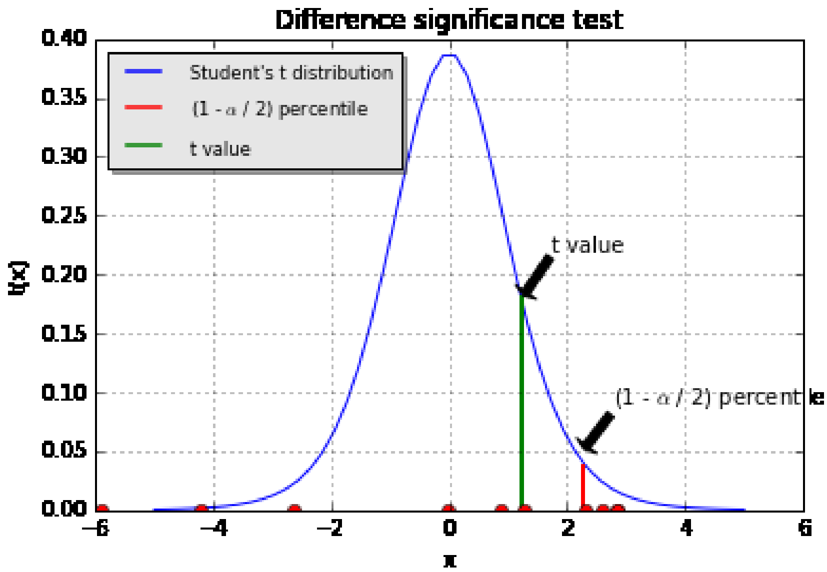


Figure 6.4: Student's t-test result. *The graph is generated using Python 2.7 and Matplotlib 1.4.3*

hypothesis for this case to be "*There is no significant difference between two lifeloggers' data*" and alternative hypothesis is " *There is significant difference between two lifeloggers' data*".

The significance level α in Figure 6.4 is a threshold, which means for a given hypothesis test, if α is larger than or equal to $P - value$, then it is considered statistically significant. The data we use here for t-test is the concepts distribution for two different lifeloggers. The α value is 0.05. Figure 6.4 is the t test result of our collected data for lifelog linkage analysis. As shown in Figure 6.4, the t-value is within the acceptance area, thus we can assert that there is no big difference between lifeloggers data and the proposed retrieval model should be applicable for both datasets. There are future work opportunities in examining this across diverse population.

6.4.3 User Study

This section provides overview and instruction with regard to the evaluation of interactive lifelog information retrieval system with users with reference to Kelly's work [90]. For retrieval research tasks, we ask 5 users to put in queries and rank the retrieved results by 1-5. In our case, 5 stands for most relevant, and 1 means least relevant. All these ranks are used to calculate precision, recall and F-measure. Appendix B shows the interface of our retrieval system for lifelog data.

6.5 Evaluations

To address hypothesis 2 and to actually evaluate the real-world usefulness of the MemoryMesh as the lifelog organisation tool, it is necessary to engage a new type of evaluation for the MemoryMesh. The first version of the event segmentation approach is used in our prototype system *ShareDay* (see Figure 3.5). The initial user study has been conducted and result can be seen in Table 4.5. For our ongoing work, rather than reusing the previous test-collection based evaluation, which is more suitable to a real-world user evaluation over a prolonged period of time by a number of users, we follow the approach of our prior work in MMM2013 (see Publication 6 and Table 4.5. It engages users to actively use a new lifelogging prototype that is built for this research. This prototype is based on the *ShareDay* system, but different from the browsing-focussed 2013 system, it is focused and targeted at evaluating the MemoryMesh and the linkage algorithms. Hence it incorporates an information retrieval module (implementing the baseline and 3 proposed linkage algorithms) that supports users in both searching for past events as well as browsing linked MemoryMesh in a manner similar to user browsing of associative memory and the multi-modal nature of the lifelog. The retrieval from the MemoryMesh in response to user queries supports users with an information seeking need such as an object that was seen before, a meeting in the past, or any previous experience that has been recorded in the MemoryMesh. This is a challenging task because the nature of the query is far more complex than in WWW search due to the complexity of associative memory. The semantic gap is significantly more noticeable in MemoryMesh retrieval. The retrieval model of lifelog data retrieval needs to ensure high accuracy of retrieved results by concerning latent semantically relevance ranking of associated events via the linkage graph.

We evaluate an early model of the MemoryMesh under the instruction described in Kelly's work [90]. This evaluation is based on identifying the quality of the links between the events that consist of the MemoryMesh; this naturally employs a precision-oriented retrieval measure. In this model, we use randomly selected events as query topics and evaluate based on the linkage likelihood. All threshold of linkage is 10%, that means, only top 10 percent strongest similarity are selected as effective similarity values for linking two events.

6.5.1 Evaluation Based on Reminiscence

In Sellen and Whittaker's 5*R* theory, retrieval, reminiscence, recollection, remembering intention and reflection are proved to be significant in lifelog research for human memory enhancement. In this research, we follow the tracks of lifelog researchers to evaluate our MemoryMesh to be effective in human memory reminiscence.

Here, we consider two types of user access case:

1. **Retrieval**.

   Retrieval here implies to find all relevant events in our MemoryMesh system.

For this we evaluate how the system retrieves events from lifelog dataset in response to a given event user. Users mark each retrieved result with different level of relevance which shows how the retrieved docs are ranked. And this feedback is used to calculate the precision, recall and other measurements for evaluation.

2. **Reminiscence**.

   Reminiscence is more about browsing through lifelog events. Browsing through all retrieved results gives users full accessibility to their lifelog data. Therefore, we evaluate reminiscence in lifelog by evaluating user browsing history and their feedback.

   Table 6.3 and 6.4 are the results of evaluation for lifelog reminiscence. This evaluation includes comparison among 4 different features sets and linkage models: contextual information, face recognition, conceptual information and feature fusion. These are all features mentioned in previous sections.

Feature Sets	Precision@10	Recall@10	F-measure@10
context	45.48	45.32	46.18
Faces	43.26	38.24	19.07
Concept	70.82	48.04	32.49
Feature Fusion	**83.28**	**58.62**	**55.43**

Table 6.3: Evaluation of different features for linkage analysis.

In this evaluation, only top 10 results are included. $Precision@10$ means precision for top 10 results, same for $Recall@10$ and $F-measure@10$. As shown in the Table 6.3, the feature fusion method outperforms (83.28% of accuracy) the other ones, which means the more features that are included into the lifelog nodes representation, the more accurate the retrieved results are. The faces based linkage model due to scarcity of the features as the faces are not included in all events. That means in this model, if two events both have no faces, their similarity could be high.

6.5.2 Evaluation Based on Event Retrieval

Evaluation based on event retrieval uses events as queries to retrieve similar event and users annotate the retrieved results with different scales (1-5) to mark the relevance of the retrieved items. Table 6.4 is the evaluation results for all 6 participants. All best performance is marked out with bold, and worst performance is marked out with italic. As we can see from the evaluation of lifelog retrieval Table 6.4 that simpler algorithm like *TF-IDF* outperforms other algorithms in most cases, while *SALSA* algorithm is also a good choice for lifelog information retrieval. Also, the more data a lifelogger has, the more accurate the retrieval becomes, which means more data contributes more to the solution of the sparsity problem in our dataset.

One interesting finding from Table 6.4 is that PageRank significantly outperforms other linkage graph based retrieval methods by averagely 2%. The PageRank applied here is *Personalized PageRank* as introduced in Section 3, which is query independent. That indicates that the output of the ranked list is particularly finalized by the importance of the life events that is annotated by users themselves. This means that with user input of importance as one dimension of the cosine similarity calculation affects the ranked list positively by almost 2%.

6.6 Conclusion

In this chapter, we talk about the event segmentation for lifelog data. Further this research, events segmented can be presented as experience nodes in our whole life,

participant	**Activity**	MAP	P@5	P@10	R@5	R@10	NDCG@5	NDCG@10
1	TF-IDF	75.12	**88.33**	70.82	**89.94**	**90.21**	67.08	69.14
	PageRank	77.43	83.21	74.09	85.21	87.33	60.99	63.77
	HITS	60.33	76.95	67.4	86.91	65.34	34.23	65.34
	SALSA	45.45	57.72	67.89	70.34	44.5	57.81	60.7
2	TF-IDF	89.22	83.12	75	88.93	85.43	**78.09**	**80.47**
	PageRank	67.55	69.04	76.78	65.43	70.02	56.04	60.43
	HITS	56.77	45.22	66.78	55.45	66.78	45.06	56.5
	SALSA	47.44	33.45	54.14	66.76	65.55	45.53	56.98
3	TF-IDF	33.45	30.56	40.56	45.67	65.78	56.95	60.32
	PageRank	67.11	67.44	77.89	62.21	76.54	55.43	64.55
	HITS	57.55	49.32	60.45	44.38	60.54	34.5	46.78
	SALSA	56.32	45.65	70.44	44.3	60	54.54	60.56
4	TF-IDF	60.54	59.04	70.3	45.34	49.8	55.67	61.11
	PageRank	**69.56**	67.88	**80.9**	45.54	56.43	56.09	66.5
	HITS	34.35	30.56	40.78	45.3	*30.23*	34.56	54.3
	SALSA	45.43	33.45	48.9	33.45	34.56	*26.04*	69.63
5	TF-IDF	65.65	49.04	66.78	45.43	50.34	46.7	50.69
	PageRank	63.44	59.04	77.67	44.5	76.53	32.35	39.45
	HITS	49.32	45.45	53.45	55.43	56.7	44.59	60.08
	SALSA	45.45	33.45	56.43	34.43	56.59	46.54	54.64
6	TF-IDF	43.40	39.56	56.54	37.54	45.43	43.42	45.45
	PageRank	45.43	34.43	59.98	43.40	45.62	33.56	*34.53*
	HITS	27.55	19.04	33.23	*15.43*	*16.02*	*26.44*	30.87
	SALSA	*19.34*	*12.34*	*22.34*	24.43	43.0	33.4	35.78

Table 6.4: Evaluation of lifelog retrieval. P: precision; R: recall; NDCG: Normalized Discounted Cumulative Gain. All are in percentage (%)

these experience nodes are linked together by sharing features like same location, same people circle, same objects encountered etc. Then we describe the three proposed linkage analysis approaches and how we intend to evaluate them.

We have proposed and presented the MemoryMesh, which organises a lifelog as a densely linked hypermedia. We suggest that applying a MemoryMesh organization model to a lifelog will increase the flexibility and usefulness of the lifelog and support new types of user interaction and better support real-time user search and retrieval. Future work that relies on this initial work could examine more types of information need and more of the 5*R*'s of lifelog access. What we present here is initial findings that suggest the value of a MemoryMesh as a lifelog organization methodology.

Chapter 7

Conclusion and Summary

In this thesis, we proposed a number of research questions and hypothesis all centred on the development of improved models of lifelog data organisation and management. We have described the construction of the lifelog utilizing various lifelog data collected by wearable sensors, especially by conducting event segmentation for experience retrieval and activity recognition. The aim of this research was to build a personal life log retrieval system based on activity recognition and its assistance in event segmentation, model this as an associative MemoryMesh and explore the potential of MemoryMesh.

Early in this thesis, we have discussed the potentials of lifelogging and gathering large amounts of physical and visual sensors data. In the past decades, much effort has been focused on miniaturising hardware and sensors to facilitate the capture of large amounts of such data, and it is only recently that the community has seriously started tackling the problem of effectively managing these data. To verify our research, users' annotation is collected, they are used as ground truth. In the remaining sections, we talked about using four approaches to detect users' important moments: machine learning , information retrieval and mining sequential pattern.

Overall, we have designed a model of doing lifelog information retrieval based on activity recognition and event detection from lifelog data. The lifelog events constructed the retrievable unites of the lifelog information retrieval system to support retrieval from the lifelog.

7.1 Research Objectives Re-visited

The major purpose of this thesis has been to propose a lifelog access model to encourage application of lifelog technologies through three main aspects: sensory discovery of lifelog data, visual exploration of lifelog images and linkage constructing and analysis to discover the patterns of individual's lifelog.

1. "*If we can recognize activities from lifelog, can these activities facilitate human life event segmentation from lifelog? If so, how efficient and accurate can this be?*"

   We have used experiments to prove that lifelog activities can be employed in lifelog event detection when the lifelog activity tags are added into event detection models as additional features.

2. "*If human lifelog can be constructed as different unities of daily events, in what way these daily events can be linked together well and constructed as humans memory mesh?*"

   The approaches introduced in this thesis about how to organize different lifelog events and apply associative memory theories in constructing MemoryMesh to support effective retrieval.

3. "*How can we support search and retrieval over this MemoryMesh of life experience using multiple types of information need?*"

We developed a model in the MemLog system to visualize lifelog nodes and to support lifelog information retrieval. We evaluated this through six participants' lifelog and discussed the results from different perspectives.

7.2 Contribution

This thesis follows a logical sequence - from explaining what is lifelog and what is cognitive science to how the brain works and how cognitive science and information retrieval theories can be applied in lifelog research. This thesis tries to accommodate all the requests that lifelog needs, especially for retrieving lifelog episodes and meaningful lifelog events to propagate lifelog application in real-life except from some discussion over privacy issues with lifelog.

In this thesis, we try to solve the three research questions that are proposed in Section 1.4, Chapter 1.

Firstly, we explored using machine learning algorithms to facilitate physical activity recognition in lifelog research. ZhiWo (literally means *Know Me*) was a system that was developed to allow system users to mark down their daily activities and share with other by uploading to a server throw http service. The data collected from voluntary upload of users are applied to train models for activity recognition. The data establish the basics for further lifelog exploration.

Secondly, we conducted some experiments on how visual content and contextual information can be applied in the lifelog event segmentation. This is also the base for lifelog information retrieval by providing the structure of retrievable nodes from lifelog information dataset.

Finally, we presented the concept of MemoryMesh and a system called MemLog (see Appendix 7 for more information) to demonstrate how a lifelog system can be used for lifelog retrieval. The retrievable unites are life events that are automatically detected. The graph structure of linked lifelog establishes MemoryMesh. We have shown through initial experimentation the potential of MemoryMesh.

7.3 Future Work

There are many opportunities for future work, such as a refined model of the event segmentation process using activity recognition and then the construction of the MemoryMesh. This could be evaluated against state-of-the-art construction of the MemoryMesh. We will also suggest evaluation of the linkage models in a browsing experiment. Both the event segmentation process and the MemoryMesh organisation and access algorithms will be evaluated using the appropriate methodologies as described.

We would also like to see new techniques that can enhance daily and real-time update for each memory event node. We also suggest exploring more new retrieval models, in addition to TF-IDF, PageRank or HITS based retrieval models. We also suggest more 5*R*s exploration over evaluation based on real-world use cases. Whether additional sensing and additional annotation would enhance a better experience or performance for lifelog users and researchers would be also a significant future research task.

One of the future work that was also mentioned earlier in this thesis was about ontology in lifelog. Ontology provides multiple dimensional representation of lifelog concepts. While in this research, we focused on using concepts to detect relationships between different lifelog events and activities, in the future, more exploration of concepts and lifelog ontology could be also a potential large research topic in this field.

In addition, we think it is likely that the applicable lifelog access models would have different determinants, may relate to very specific utility scenarios and therefore may require different intervention and promotion strategies.

It is our conjecture that the biggest contribution of this work is in the motivation for linked models of lifelog and we encourage continued exploration. We consider the work in this thesis to be initial exploratory work. The proposed MemoryMesh is a concept that warrants future study. It is not yet a solved problem.

7.4 Final Thoughts

Lifelog is undoubtedly one of the most valuable technology in the coming years. Before 19th century, nobody knew what photography was, and now everybody can be a photographer. Not only for recording personal lives, beautiful nature scene, but also for self-reflection of making people feel fulfilled by the concept of recording and been valued. In the new century, we can except lifelog to be as common as digital photography today.

Although we have strong concern about privacy protection, but a lifelog system has to be well designed to protect the privacy of data owners. Lifelog owners have full control of their data if it is store in a cloud based lifelog management system. Just like websites with SSL certificates to ensure the protected transaction between the website and the user, lifelog system designers and developers do also need a good network security sense to put the privacy issues to the front.

But also, life is humdrum interspersed with moments of different emotions like interest, joy and sorrow. Apart from recording images and events, how to record emotions at moments is still not within any feasible research scope yet, but it would be a very interesting and desirable topic to go forward to with in the future.

Overall, it is more beneficial in a long run with lifelog data stored. And the trend is getting indispensable when data can be applied to daily lives to bring enhancement. But still, how to utilize the data in a proper manner and how to apply the data to enhance people's lives is still a subject for research.

This research is meant to be unbound and to make initial progress, as well as to draw forth by abler people to share ideas over lifelog research. It also invites more people is to explore, to find something new and to connect with each other. We feel that we have made initial findings, and opened doors for future research.

We have shown potential from replacing the database methodology of all past lifelog research with a novel linked model. It is our hope that the concept of MemoryMesh will be further explored and enhanced by future research efforts.

Chapter 8

List of Publications

1. **Zhou, Lijuan Marissa**, Brian Moynagh, Liting Zhou, Tengqi Ye and Cathal Gurrin. *MemLog, an Enhanced Lifelog Annotation and Search Tool.* In: The 19th International Conference on Multimedia Modelling, 5-7 Jan 2015, Sydney, Australia

2. Peter Scarborough, Richard Harrington, Anja Mizdrak, **Lijuan Marissa Zhou**, Aiden Doherty. *The Preventable Risk Integrated ModEl (PRIME) and its use to estimate the health impact of public health policy scenarios.* In: Journal of Scientifica: Experts in Electrophysiology and Imaging, Volume 2014

3. **Lijuan Marissa Zhou**, Hongfei Lin, Jun Yan. *Tags Know You Better:A New Approach to Enhancing MIR System.* In: Journal of Computer Science, Iusse 2, Volume 2014

4. **Zhou, Lijuan Marissa**, Gurrin, Cathal and Zhengwei Qiu. *ZhiWo: Activity Tagging and Recognition System for Personal Lifelogs*, In:2013 ACM International Conference on Multimedia Retrieval, April 16 - 19, 2013, Dallas, Texas, USA.

5. Frank Hopfgartner, Cathal Gurrin, **Zhou, Lijuan Marissa** and Yang Yang. *User Interaction Patterns for the Design of Lifelogging Systems*. In: Springer book "Semantic Models for Adaptive Interactive Systems" (SEMAIS), 2013.

6. **Zhou, Lijuan Marissa** and Caprani, Niamh and Gurrin, Cathal. *ShareDay: A Multi-modal Lifelog System for Group Sharing*. In: The 19th International Conference on Multimedia Modeling, 7-9 Jan 2012, Huangshan, Anhui, China.

7. Hopfgartner, Frank, Scott, David, Wang, Hongyi, Yang, Yang, Zhang, Zhenxin, **Zhou, Lijuan Marissa**, and Gurrin, Cathal. *Helping the Helpers: How Video Retrieval Can Assist Special Interest Groups*. In: The 19th International Conference on Multimedia Modeling, 7-9 Jan 2012, Huangshan, Anhui, China.

8. **Zhou, Lijuan Marissa** and Caprani, Niamh and Gurrin, Cathal, *ShareDay:A Memory Enhancing Lifelogging System Based on Group Sharing*. In: Summer School on Social Media Modeling and Search SSMS 2012, 10-14 Sept 2012 , Santorini, Greece.

9. **Zhou, Lijuan Marissa** and Lu, Xueke and Gurrin, Cathal. *A Realtime Lifelogging Solution for iOS Devices*. In: Irish HCI, 20-21 June 2012, Galway, Ireland.

10. **Zhou, Lijuan Marissa** and Gurrin, Cathal. *A Survey on Life Logging Data Capturing*. In: SenseCam Symposium 2012, 3-4 Apr 2012, Oxford, UK.

11. **Zhou, Lijuan Marissa** and Lin, Hongfei and Gurrin, Cathal. *EMIR: A Novel Emotion-based Music Retrieval System*. Advances in Multimedia Modeling, 2012: 627-629.

12. **Zhou, Lijuan**, Lin, Hongfei and Liu, Wenfei. *Enriching Music Information Retrieval Using Emotion Detection*. In: SIGIR 2011 Workshop on Enriching Information Retrieval. Beijing, China.

13. Wu, Di, **Zhou, Lijuan**, and Lin, Hongfei. *The Design and Implementation of Graduate Occupation Recommendation System.* Journal of Guangxi Normal University(Natural Science Edition), 29(1) 2011: 179-185.

14. **Zhou, Lijuan**, Lin, Hongfei and Luo, Wenhua. *Criminal Network Recognition Mechanism Based on Entity Relationship*. Application Research of Computers, 28(3) 2010: 998-1002

15. Li, Jing, Lin, Hongfei, and **Zhou, Lijuan**. *Emotion Tag Based Music Retrieval Algorithm*. Information Retrieval Technology, 2010: 599-609.

Bibliography

[1] Shivani Agarwal, Aatif Awan, and Dan Roth. Learning to detect objects in images via a sparse, part-based representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(11):1475–1490, 2004.

[2] Nadav Aharony, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*, 7(6):643–659, 2011.

[3] Dal Ahn, Jun-Seok Park, Chul-Soo Kim, Juno Kim, Yongxi Qian, and Tatsuo Itoh. A design of the low-pass filter using the novel microstrip defected ground structure. *Microwave Theory and Techniques, IEEE Transactions on*, 49(1):86–93, 2001.

[4] K. Aizawa, D. Tancharoen, S. Kawasaki, and T. Yamasaki. Efficient retrieval of life log based on context and content. In *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pages 22–31. ACM, 2004.

[5] Icek Ajzen. Theory of planned behavior. *Handb Theor Soc Psychol Vol One*, 1:438, 2011.

[6] Rami Albatal, Cathal Gurrin, Jiang Zhou, Yang Yang, Denise Carthy, and Na Li. Senseseer mobile-cloud-based lifelogging framework. In *Technology and Society (ISTAS), 2013 IEEE International Symposium on*, pages 144–146. IEEE, 2013.

[7] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L Reyes-Ortiz. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *Ambient assisted living and home care*, pages 216–223. Springer, 2012.

[8] Earl Babbie and Earl Babbie. *Survey research methods*. Number 04; H62, B3 1990. Wadsworth Belmont, CA, 1990.

[9] Michael A Babyak. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*, 66(3):411–421, 2004.

[10] N. Baker. Zigbee and bluetooth strengths and weaknesses for industrial applications. *Computing & Control Engineering Journal*, 16(2):20–25, 2005.

[11] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. *Pervasive Computing*, pages 1–17, 2004.

[12] R.A. Barkley et al. Major life activity and health outcomes associated with attention-deficit/hyperactivity disorder. *Journal of clinical psychiatry*, 63:10–15, 2002.

[13] P. Baronti, P. Pillai, V.W.C. Chook, S. Chessa, A. Gotta, and Y.F. Hu. Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and zigbee standards. *Computer communications*, 30(7):1655–1695, 2007.

[14] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision–ECCV 2006*, pages 404–417. Springer, 2006.

[15] C Gordon Bell and Jim Gemmell. *Total recall: How the e-memory revolution will change everything*. Dutton, 2009.

[16] E Berry, M Conway, C Moulin, H Williams, S Hodges, L Williams, K Wood, and G Smith. Stimulating episodic memory: Initial explorations using sensecam. In *psychonomic society. 47th annual meeting*, volume 11, pages 56–57, 2006.

[17] Emma Berry, Adam Hampshire, James Rowe, Steve Hodges, Narinder Kapur, Peter Watson, Georgina Browne, Gavin Smyth, Ken Wood, and Adrian M Owen. The neural basis of effective memory therapy in a patient with limbic encephalitis. *Journal of Neurology, Neurosurgery & Psychiatry*, 80(11):1202–1205, 2009.

[18] Emma Berry, Narinder Kapur, Lyndsay Williams, Steve Hodges, Peter Watson, Gavin Smyth, James Srinivasan, Reg Smith, Barbara Wilson, and Ken Wood. The use of a wearable camera, sensecam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: A preliminary report. *Neuropsychological Rehabilitation*, 17(4-5):582–601, 2007.

[19] G. Bieber, J. Voskamp, and B. Urban. Activity recognition for everyday life on mobile phones. *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, pages 289–296, 2009.

[20] P. Bonato. Wearable sensors and systems. *Engineering in Medicine and Biology Magazine, IEEE*, 29(3):25–36, 2010.

[21] Stephan Bosch, Mihai Marin-Perianu, Raluca Marin-Perianu, Paul Havinga, and Hermie Hermens. Keep on moving! activity monitoring and stimulation using wireless sensor networks. In *Smart Sensing and Context*, pages 11–23. Springer, 2009.

[22] Carlijn VC Bouten, Karel TM Koekkoek, Maarten Verduin, Rens Kodde, and Jan D Janssen. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *Biomedical Engineering, IEEE Transactions on*, 44(3):136–147, 1997.

[23] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc., 2008.

[24] Sergey Brin and Lawrence Page. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks*, 56(18):3825–3833, 2012.

[25] Peter Brusilovsky and Christoph Peylo. Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education (IJAIED)*, 13:159–172, 2003.

[26] Adrian Burns, Barry R Greene, Michael J McGrath, Terrance J O'Shea, Benjamin Kuris, Steven M Ayer, Florin Stroiescu, and Victor Cionca. Shimmer–a wireless sensor platform for noninvasive biomedical research. *Sensors Journal, IEEE*, 10(9):1527–1534, 2010.

[27] Tony Buzan, Barry Buzan, and James Harrison. *The mind map book: Unlock your creativity, boost your memory, change your life*. Pearson BBC Active, 2010.

[28] Daragh Byrne, Aiden R Doherty, Cees GM Snoek, Gareth JF Jones, and Alan F Smeaton. Everyday concept detection in visual lifelogs: validation, relationships and trends. *Multimedia Tools and Applications*, 49(1):119–144, 2010.

[29] Daragh Byrne, Barry Lavelle, Aiden R Doherty, Gareth JF Jones, and Alan F Smeaton. Using bluetooth and gps metadata to measure event similarity in sensecam images. In *IMAI07-5th International Conference on Intelligent Multimedia and Ambient Intelligence*, pages 18–24. World Scientific, 2007.

[30] Niamh Caprani. *The design of an intergenerational lifelog browser to support sharing within family groups*. PhD thesis, Dublin City University, 2013.

[31] Niamh Caprani, Aiden R Doherty, Hyowon Lee, Alan F Smeaton, Noel E O'Connor, and Cathal Gurrin. Designing a touch-screen sensecam browser to support an aging population. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 4291–4296. ACM, 2010.

[32] Niamh Caprani, Cathal Gurrin, and Noel E. O'Connor. Sharing as a motivation for lifelogging. In *SenseCam 2012: 3rd Annual Symposium*, SenseCam 2012, 2012.

[33] Niamh Caprani, Noel E O'Connor, and Cathal Gurrin. Motivating lifelogging practices through shared family reminiscence. In *Computer Human Interaction 2011*, CHI, 2011.

[34] Niamh Caprani, Paulina Piasek, Cathal Gurrin, Noel E O'Connor, Kate Irving, and Alan F Smeaton. Life-long collections: Motivations and the implications for lifelogging with mobile devices. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 6(1):15–36, 2014.

[35] Pierluigi Casale, Oriol Pujol, and Petia Radeva. Human activity recognition from accelerometer data using a wearable device. In *Pattern Recognition and Image Analysis*, pages 289–296. Springer, 2011.

[36] Carl J Caspersen, Kenneth E Powell, and Gregory M Christenson. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports*, 100(2):126, 1985.

[37] M. Chan, C. Hariton, P. Ringeard, and E. Campo. Smart house automation system for the elderly and the disabled. In *Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on*, volume 2, pages 1586–1589. IEEE, 1995.

[38] C.C. Chang and C.J. Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[39] Olivier Chapelle, Yi Chang, and Tie-Yan Liu. Future directions in learning to rank. In *JMLR Workshop and Conference Proceedings*, volume 14, pages 91–100, 2011.

[40] J. Chen, K. Kwong, D. Chang, J. Luk, and R. Bajcsy. Wearable sensors for reliable fall detection. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 3551–3554. IEEE, 2006.

[41] B. Clarkson, K. Mase, and A. Pentland. Recognizing user context via wearable sensors. In *Wearable Computers, The Fourth International Symposium on*, pages 69–75. IEEE, 2000.

[42] Research Ethics Committee. Dcu research ethics forms. `http://www4.dcu.ie/research/research_ethics/rec_forms.shtml`. Accessed: 03/03/2013.

[43] Martin A Conway. Episodic memories. *Neuropsychologia*, 47(11):2305–2313, 2009.

[44] C.B. Corbin and R. Lindsey. *Fitness for life*. Human Kinetics Publishers, 2006.

[45] C.B. Corbin, R. Lindsey, and G. Welk. *Concepts of physical fitness: Active lifestyles for wellness*. McGraw-Hill Boston, 2006.

[46] Kathryn L Dannecker, Sean A Petro, Edward L Melanson, and Raymond C Browning. Accuracy of fitbit activity monitor to predict energy expenditure with and without classification of activities. *Medicine & Science in Sports & Exercise*, 43(5):62, 2011.

[47] Anthony V D'Antoni, Genevieve P Zipp, Valerie G Olson, and Terrence F Cahill. Does the mind map learning strategy facilitate information retrieval and critical thinking in medical students? *BMC medical education*, 10(1):61, 2010.

[48] D. De Jager, A.L. Wood, G.V. Merrett, B.M. Al-Hashimi, K. O'Hara, N.R. Shadbolt, and W. Hall. A low-power, distributed, pervasive healthcare system for supporting memory. In *Proceedings of the First ACM MobiHoc Workshop on Pervasive Wireless Healthcare*, page 5. ACM, 2011.

[49] Konstantinos G Derpanis. Overview of the ransac algorithm. *York University, Toronto, Canada*, 2010.

[50] Pierre A Devijver and Josef Kittler. *Pattern recognition: A statistical approach*, volume 761. Prentice-Hall London, 1982.

[51] Martin Dodge and Rob Kitchin. " outlines of a world coming into existence": Pervasive computing and the ethics of forgetting. *Environment and Planning B*, 34(3):431–445, 2007.

[52] Aiden R Doherty. *Providing effective memory retrieval cues through automatic structuring and augmentation of a lifelog of images*. PhD thesis, Dublin City University, 2008.

[53] Aiden R Doherty, Ciarán Ó Conaire, Michael Blighe, Alan F Smeaton, and Noel E O'Connor. Combining image descriptors to effectively retrieve events from visual lifelogs. In *Proceedings of the 1st ACM international conference on Multimedia information retrieval*, pages 10–17. ACM, 2008.

[54] Aiden R Doherty, Katalin Pauly-Takacs, Niamh Caprani, Cathal Gurrin, Chris JA Moulin, Noel E O'Connor, and Alan F Smeaton. Experiences of aiding autobiographical memory using the sensecam. *Human–Computer Interaction*, 27(1-2):151–174, 2012.

[55] A.R. Doherty, N. Caprani, C.Ó. Conaire, V. Kalnikaite, C. Gurrin, A.F. Smeaton, and N.E. OConnor. Passively recognising human activities through lifelogging. *Computers in Human Behavior*, 27(5):1948–1958, 2011.

[56] A.R. Doherty, C.J.A. Moulin, and A.F. Smeaton. Automatically assisting human memory: A sensecam browser. *Memory*, 19(7):785–795, 2011.

[57] A.R. Doherty and A.F. Smeaton. Combining face detection and novelty to identify important events in a visual lifelog. In *Computer and Information*

Technology Workshops, 2008. CIT Workshops 2008. IEEE 8th International Conference on, pages 348–353. IEEE, 2008.

[58] A.R. Doherty and A.F. Smeaton. segmenting lifelog data into events. In *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on*, pages 20–23. IEEE, 2008.

[59] A.R. Doherty, A.F. Smeaton, K. Lee, and D.P.W. Ellis. Multimodal segmentation of lifelog data. In *Large Scale Semantic Access to Content (Text, Image, Video, and Sound)*, pages 21–38, 2007.

[60] J. Doyle, B. O'Mullane, A. O'Hanlon, and R.B. Knapp. Requirements gathering for the delivery of healthcare data in aware homes. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, pages 254–257. IEEE, 2011.

[61] Genevieve F Dunton, Yue Liao, Stephen S Intille, Donna Spruijt-Metz, and Maryann Pentz. Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones. *Obesity*, 19(6):1205–1212, 2011.

[62] Genevieve Fridlund Dunton, Keito Kawabata, Stephen Intille, Jennifer Wolch, and Mary Ann Pentz. Assessing the social and physical contexts of children's leisure-time physical activity: an ecological momentary assessment study. *American Journal of Health Promotion*, 26(3):135–142, 2012.

[63] Nathan Eagle, Alex Sandy Pentland, and David Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274–15278, 2009.

[64] Katherine Ellis, Suneeta Godbole, Simon Marshall, Gert Lanckriet, John Staudenmayer, and Jacqueline Kerr. Identifying active travel behaviors in challenging environments using gps, accelerometers, and machine learning algorithms. *Frontiers in public health*, 2, 2014.

[65] Yrjö Engeström, Reijo Miettinen, and Raija-Leena Punamäki. *Perspectives on activity theory*. Cambridge University Press, 1999.

[66] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[67] R.R. Fletcher, K. Dobson, M.S. Goodwin, H. Eydgahi, O. Wilder-Smith, D. Fernholz, Y. Kuboyama, E.B. Hedman, M.Z. Poh, and R.W. Picard. icalm: Wearable sensor and network architecture for wirelessly communicating and logging autonomic activity. *Information Technology in Biomedicine, IEEE Transactions on*, 14(2):215–223, 2010.

[68] B.J. Fogg. Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002(December):5, 2002.

[69] B. Gedik and L. Liu. Protecting location privacy with personalized k-anonymity: Architecture and algorithms. *Mobile Computing, IEEE Transactions on*, 7(1):1–18, 2008.

[70] J. Gemmell, G. Bell, and R. Lueder. Mylifebits: a personal database for everything. *Communications of the ACM*, 49(1):88–95, 2006.

[71] J. Gemmell, G. Bell, R. Lueder, S. Drucker, and C. Wong. Mylifebits: fulfilling the memex vision. In *Proceedings of the tenth ACM international conference on Multimedia*, pages 235–238. ACM, 2002.

[72] Cathal Gurrin, Daragh Byrne, Noel O'Connor, Gareth JF Jones, and Alan F Smeaton. Architecture and challenges of maintaining a large-scale, context-aware human digital memory. 2008.

[73] Cathal Gurrin, Zhengwei Qiu, Mark Hughes, Niamh Caprani, Aiden R Doherty, Steve E Hodges, and Alan F Smeaton. The smartphone as a platform for wearable cameras in health research. *American journal of preventive medicine*, 44(3):308–313, 2013.

[74] Cathal Gurrin, Alan F Smeaton, and Aiden R Doherty. Lifelogging: Personal big data. *Foundations and Trends in Information Retrieval*, 8(1):1–125, 2014.

[75] Cathal Gurrin, Alan F Smeaton, Zhengwei Qiu, and Aiden Doherty. Exploring the technical challenges of large-scale lifelogging. In *Proceedings of the 4th International SenseCam & Pervasive Imaging Conference*, pages 68–75. ACM, 2013.

[76] Bridgette M Hard, Barbara Tversky, and David S Lang. Making sense of abstract events: Building event schemas. *Memory & cognition*, 34(6):1221–1235, 2006.

[77] Paula Hatch, Nicola Stokes, and Joe Carthy. Lexical chaining for web-based retrieval of breaking news. In *Adaptive Hypermedia and Adaptive Web-Based Systems*, pages 327–330. Springer, 2000.

[78] E.A. Heinz, K.S. Kunze, M. Gruber, D. Bannach, and P. Lukowicz. Using wearable sensors for real-time recognition tasks in games of martial arts-an initial experiment. In *Computational Intelligence and Games, 2006 IEEE Symposium on*, pages 98–102. IEEE, 2006.

[79] S. Hodges, E. Berry, and K. Wood. Sensecam: A wearable camera that stimulates and rehabilitates autobiographical memory. *Memory*, 19(7):685–696, 2011.

[80] Steve Hodges, Lyndsay Williams, Emma Berry, Shahram Izadi, James Srinivasan, Alex Butler, Gavin Smyth, Narinder Kapur, and Kenneth R. Wood. Sensecam: A retrospective memory aid. In *Ubicomp*, pages 177–193, 2006.

[81] Tim Hussein, Heiko Paulheim, Stephan Lukosch, Jürgen Ziegler, and Gaëlle Calvary. *Semantic Models for Adaptive Interactive Systems*. Springer Science & Business, 2013.

[82] Tim Jacquemard. *Ethics of lifelog technology*. PhD thesis, Dublin City University, 2014.

[83] Kalervo Järvelin and Jaana Kekäläinen. Cumulated gain-based evaluation of ir techniques. *ACM Transactions on Information Systems (TOIS)*, 20(4):422–446, 2002.

[84] J. Jeon, V. Lavrenko, and R. Manmatha. Automatic image annotation and retrieval using cross-media relevance models. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 119–126. ACM, 2003.

[85] Yangqing Jia. Caffe: An open source convolutional architecture for fast feature embedding. *h ttp://caffe. berkeleyvision. org*, 2013.

[86] YU Jie, Nachiappan Subramanian, Kun Ning, and David Edwards. Product delivery service provider selection and customer satisfaction in the era of internet of things: A chinese e-retailers perspective. *International Journal of Production Economics*, 159:104–116, 2015.

[87] Gareth JF Jones, Cathal Gurrin, Liadh Kelly, Daragh Byrne, and Yi Chen. Information access tasks and evaluation for personal lifelogs. 2008.

[88] D. Kahneman, A.B. Krueger, D.A. Schkade, N. Schwarz, and A.A. Stone. A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702):1776–1780, 2004.

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

[90] Diane Kelly. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(12):1–224, 2009.

[91] Liadh Kelly, Jinyoung Kim, and David Elsweiler. Workshop on evaluating personal search. In *ACM SIGIR Forum*, volume 45, pages 81–86. ACM, 2012.

[92] Paul Kelly, Aiden Doherty, Emma Berry, Steve Hodges, Alan M Batterham, and Charlie Foster. 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*, 8(1):1–9, 2011.

[93] Paul Kelly, Aiden R Doherty, Alex Hamilton, Anne Matthews, Alan M Batterham, Michael Nelson, Charlie Foster, and Gill Cowburn. Evaluating the feasibility of measuring travel to school using a wearable camera. *American journal of preventive medicine*, 43(5):546–550, 2012.

[94] Adil Mehmood Khan, Young-Koo Lee, Sungyoung Y Lee, and Tae-Seong Kim. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *Information Technology in Biomedicine, IEEE Transactions on*, 14(5):1166–1172, 2010.

[95] Basel Kikhia, Josef Hallberg, Johan E Bengtsson, and Stefan Savenstedt. Building digital life stories for memory support. *International journal of Computers in Healthcare*, 1(2):161–176, 2010.

[96] Jesper Kjeldskov and Connor Graham. A review of mobile hci research methods. In *Human-computer interaction with mobile devices and services*, pages 317–335. Springer, 2003.

[97] J.M. Kleinberg. Hubs, authorities, and communities. *ACM Computing Surveys (CSUR)*, 31(4es):5, 1999.

[98] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.

[99] T. Kleinberger, M. Becker, E. Ras, A. Holzinger, and P. Müller. Ambient intelligence in assisted living: enable elderly people to handle future interfaces. *Universal access in human-computer interaction. Ambient interaction*, pages 103–112, 2007.

[100] Alfred Korzybski. *Science and sanity: An introduction to non-Aristotelian systems and general semantics*. Institute of GS, 1958.

[101] Christopher A Kurby and Jeffrey M Zacks. Segmentation in the perception and memory of events. *Trends in cognitive sciences*, 12(2):72–79, 2008.

[102] J.R. Kwapisz, G.M. Weiss, and S.A. Moore. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 12(2):74–82, 2011.

[103] Hyowon Lee, Alan F Smeaton, Noel E O'Connor, and Gareth JF Jones. Adaptive visual summary of lifelog photos for personal information management. 2006.

[104] Hyowon Lee, Alan F Smeaton, Noel E OConnor, Gareth Jones, Michael Blighe, Daragh Byrne, Aiden Doherty, and Cathal Gurrin. Constructing a sensecam visual diary as a media process. *Multimedia Systems*, 14(6):341–349, 2008.

[105] M.W. Lee, A.M. Khan, and T.S. Kim. A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation. *Personal and Ubiquitous Computing*, 15(8):887–898, 2011.

[106] Sang Hyuk Lee, Soobin Lee, Heecheol Song, and Hwang Soo Lee. Wireless sensor network design for tactical military applications: remote large-scale environments. In *Military Communications Conference, 2009. MILCOM 2009. IEEE*, pages 1–7. IEEE, 2009.

[107] S.W. Lee and K. Mase. Activity and location recognition using wearable sensors. *Pervasive Computing, IEEE*, 1(3):24–32, 2002.

[108] R. Lempel and S. Moran. Salsa: the stochastic approach for link-structure analysis. *ACM Transactions on Information Systems (TOIS)*, 19(2):131–160, 2001.

[109] L. Liao. *Location-based activity recognition*. PhD thesis, University of Washington, 2006.

[110] Zhicheng Liao, Yun Xiong, and Yangyong Zhu. Daciot: A data cloud for the internet of things. In *Computer Science and Applications: Proceedings of the 2014 Asia-Pacific Conference on Computer Science and Applications (CSAC 2014), Shanghai, China, 27-28 December 2014*, page 181. CRC Press, 2015.

[111] Wei-Hao Lin and Alexander Hauptmann. Structuring continuous video recordings of everyday life using time-constrained clustering. SPIE, 2006.

[112] B. Lo, S. Thiemjarus, R. King, and G.Z. Yang. Body sensor network-a wireless sensor platform for pervasive healthcare monitoring. In *The 3rd International Conference on Pervasive Computing*, volume 13, 2005.

[113] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.

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

[115] C.D. Manning, P. Raghavan, and H. Schütze. *Introduction to information retrieval*, volume 1. Cambridge University Press Cambridge, 2008.

[116] Gérard Medioni, Isaac Cohen, François Brémond, Somboon Hongeng, and Ramakant Nevatia. Event detection and analysis from video streams. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(8):873–889, 2001.

[117] F. Michahelles, P. Matter, A. Schmidt, and B. Schiele. Applying wearable sensors to avalanche rescue. *Computers & Graphics*, 27(6):839–847, 2003.

[118] Aleksandar Milenković, Chris Otto, and Emil Jovanov. Wireless sensor networks for personal health monitoring: Issues and an implementation. *Computer communications*, 29(13):2521–2533, 2006.

[119] Martina Naughton, Nicola Stokes, and Joe Carthy. Sentence-level event classification in unstructured texts. *Information retrieval*, 13(2):132–156, 2010.

[120] Anton Nijholt. Google home: Experience, support and re-experience of social home activities. *Information Sciences*, 178(3):612–630, 2008.

[121] United States. Public Health Service. Office of the Surgeon General, Centers for Disease Control, National Center for Chronic Disease Prevention, Health Promotion (US), President's Council on Physical Fitness, and Sports (US). *Physical activity and health: a report of the Surgeon General*. Jones & Bartlett Learning, 1996.

[122] Neil Ohare, Hyowon Lee, Saman Cooray, Cathal Gurrin, Gareth Jones, Jovanka Malobabic, Noel OConnor, Alan Smeaton, and Bartlomiej Uscilowski. Mediassist: Using content-based analysis and context to manage personal photo collections. *Image and video retrieval*, pages 529–532, 2006.

[123] K. OHara, M.M. Tuffield, and N. Shadbolt. Lifelogging: Privacy and empowerment with memories for life. *Identity in the Information Society*, 1(1):155–172, 2008.

[124] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: bringing order to the web. 1999.

[125] Zhibo Pang, Lirong Zheng, Junzhe Tian, Sharon Kao-Walter, Elena Dubrova, and Qiang Chen. Design of a terminal solution for integration of in-home health care devices and services towards the internet-of-things. *Enterprise Information Systems*, 9(1):86–116, 2015.

[126] A. Pantelopoulos and N.G. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(1):1–12, 2010.

[127] K.H. Park, Z. Bien, J.J. Lee, B.K. Kim, J.T. Lim, J.O. Kim, H. Lee, D.H. Stefanov, D.J. Kim, J.W. Jung, et al. Robotic smart house to assist people with movement disabilities. *Autonomous Robots*, 22(2):183–198, 2007.

[128] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers. A review of wearable sensors and systems with application in rehabilitation. *J Neuroeng Rehabil*, 9:21, 2012.

[129] S Tejaswi Peesapati, Victoria Schwanda, Johnathon Schultz, Matt Lepage, So-yae Jeong, and Dan Cosley. Pensieve: supporting everyday reminiscence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2027–2036. ACM, 2010.

[130] C.A. Pellegrini, S.D. Verba, A.D. Otto, D.L. Helsel, K.K. Davis, and J.M. Jakicic. The comparison of a technology-based system and an in-person behavioral weight loss intervention. *Obesity*, 20(2):356–363, 2012.

[131] O.E. Pinzon and M.S. Iyengar. Persuasive technology and mobile health: A systematic review. *Persuasive Technology*, page 45, 2012.

[132] Z. Qiu, C. Gurrin, A. Doherty, and A. Smeaton. A real-time life experience logging tool. *Advances in Multimedia Modeling*, pages 636–638, 2012.

[133] Zhengwei Qiu. *A lifelogging system supporting multimodal access*. PhD thesis, Dublin City University, 2013.

[134] Reza Rawassizadeh. Towards sharing life-log information with society. *Behaviour & Information Technology*, 31(11):1057–1067, 2012.

[135] Reza Rawassizadeh and A Min Tjoa. Securing shareable life-logs. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 1105–1110. IEEE, 2010.

[136] Paul J Reaidy, Angappa Gunasekaran, and Alain Spalanzani. Bottom-up approach based on internet of things for order fulfillment in a collaborative warehousing environment. *International Journal of Production Economics*, 159:29–40, 2015.

[137] David A Sadlier and Noel E O'Connor. Event detection in field sports video using audio-visual features and a support vector machine. *Circuits and Systems for Video Technology, IEEE Transactions on*, 15(10):1225–1233, 2005.

[138] G. Salton and C. Buckley. Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5):513–523, 1988.

[139] Gerard Salton and Michael J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc., New York, NY, USA, 1986.

[140] Supheakmungkol Sarin, Michael Fahrmair, Matthias Wagner, and Wataru Kameyama. Holistic feature extraction for automatic image annotation. In

Multimedia and Ubiquitous Engineering (MUE), 2011 5th FTRA International Conference on, pages 59–66. IEEE, 2011.

[141] Peter Scarborough, Richard A Harrington, Anja Mizdrak, Lijuan Marissa Zhou, and Aiden Doherty. The preventable risk integrated model and its use to estimate the health impact of public health policy scenarios. *Scientifica*, 2014, 2014.

[142] A.J. Sellen, A. Fogg, M. Aitken, S. Hodges, C. Rother, and K. Wood. Do life-logging technologies support memory for the past?: an experimental study using sensecam. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 81–90. ACM, 2007.

[143] A.J. Sellen and S. Whittaker. Beyond total capture: a constructive critique of lifelogging. *Communications of the ACM*, 53(5):70–77, 2010.

[144] Alessio Signorini. A survey of ranking algorithms. *Department of Computer Science, University of Iowa*, 2005.

[145] A. Smeaton. Content vs. context for multimedia semantics: the case of sensecam image structuring. *Semantic Multimedia*, pages 1–10, 2006.

[146] D.H. Stefanov, Z. Bien, and W.C. Bang. The smart house for older persons and persons with physical disabilities: structure, technology arrangements, and perspectives. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 12(2):228–250, 2004.

[147] A. Subramanya, A. Raj, J.A. Bilmes, and D. Fox. Recognizing activities and spatial context using wearable sensors. *arXiv preprint arXiv:1206.6869*, 2012.

[148] Quan-Sen Sun, Sheng-Gen Zeng, Yan Liu, Pheng-Ann Heng, and De-Shen Xia. A new method of feature fusion and its application in image recognition. *Pattern Recognition*, 38(12):2437–2448, 2005.

[149] M. Swan. Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, 1(3):217–253, 2012.

[150] R. Takeda, S. Tadano, M. Todoh, M. Morikawa, M. Nakayasu, and S. Yoshinari. Gait analysis using gravitational acceleration measured by wearable sensors. *Journal of biomechanics*, 42(3):223–233, 2009.

[151] S. Tano, T. Takayama, M. Iwata, and T. Hashiyama. Multimedia informal communication by wearable computer based on real-world context and graffiti. In *Multimedia and Expo, 2006 IEEE International Conference on*, pages 649–652. IEEE, 2006.

[152] Eric J. Topol. *The creative destruction of medicine: how the digital revolution will create better health care*. New York, Basic Books, 2012.

[153] Pedro; Moreira Rui S.; Soares Christophe; Morla Ricardo Torres, Jos M.; Sobral. An adaptive embedded system for physical activity recognition. In *Proceedings of Iberian Conference on Information Systems & Technologies*, pages 186–191, 2012.

[154] Richard P Troiano, David Berrigan, Kevin W Dodd, Louise C Mâsse, Timothy Tilert, Margaret McDowell, et al. Physical activity in the united states measured by accelerometer. *Medicine and science in sports and exercise*, 40(1):181, 2008.

[155] E. Tulving. Precis of elements of episodic memory. *Behavioral and Brain Sciences*, 7(2):223–68, 1984.

[156] Priyanka Upadhyay, Gurpreet Matharu, and Naveen Garg. Modeling agility in internet of things (iot) architecture. In *Information Systems Design and Intelligent Applications*, pages 779–786. Springer, 2015.

[157] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2010.

[158] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Empowering visual categorization with the gpu. *IEEE Transactions on Multimedia*, 13(1):60–70, 2011.

[159] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.

[160] Wolfgang Wahlster, Alexander Kröner, and Dominik Heckmann. Sharedlife: towards selective sharing of augmented personal memories. In *Reasoning, Action and Interaction in AI Theories and Systems*, pages 327–342. Springer, 2006.

[161] Peng Wang and Alan F Smeaton. Using visual lifelogs to automatically characterize everyday activities. *Information Sciences*, 230:147–161, 2013.

[162] Yining Wang, Liwei Wang, Yuanzhi Li, Di He, Wei Chen, and Tie-Yan Liu. A theoretical analysis of ndcg ranking measures. In *Proceedings of the 26th Annual Conference on Learning Theory (COLT 2013)*, 2013.

[163] Roy Want, Bill N Schilit, and Scott Jenson. Enabling the internet of things. *Computer*, (1):28–35, 2015.

[164] S. Warren, J. Yao, and G.E. Barnes. Wearable sensors and component-based design for home health care. In *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, volume 3, pages 1871–1872. IEEE, 2002.

[165] Lloyd R Welch. Hidden markov models and the baum-welch algorithm. *IEEE Information Theory Society Newsletter*, 53(4):1–14, 2003.

[166] Gregory J Welk, James J McClain, Joey C Eisenmann, and Eric E Wickel. Field validation of the mti actigraph and bodymedia armband monitor using the ideea monitor. *Obesity*, 15(4):918–928, 2007.

[167] Phillip Ian Wilson and John Fernandez. Facial feature detection using haar classifiers. *Journal of Computing Sciences in Colleges*, 21(4):127–133, 2006.

[168] J.M. Winters and Y. Wang. Wearable sensors and telerehabilitation. *Engineering in Medicine and Biology Magazine, IEEE*, 22(3):56–65, 2003.

[169] Feng Xia, Laurence T Yang, Lizhe Wang, and Alexey Vinel. Internet of things. *International Journal of Communication Systems*, 25(9):1101, 2012.

[170] Jun Yang, Yu-Gang Jiang, Alexander G Hauptmann, and Chong-Wah Ngo. Evaluating bag-of-visual-words representations in scene classification. In *Proceedings of the international workshop on Workshop on multimedia information retrieval*, pages 197–206. ACM, 2007.

[171] TengQi Ye, Brian Moynagh, Rami Albatal, and Cathal Gurrin. Negative faceblurring: A privacy-by-design approach to visual lifelogging with google glass. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 2036–2038. ACM, 2014.

[172] Jeffrey M Zacks, Nicole K Speer, Khena M Swallow, Todd S Braver, and Jeremy R Reynolds. Event perception: a mind-brain perspective. *Psychological bulletin*, 133(2):273, 2007.

[173] Jeffrey M Zacks and Barbara Tversky. Event structure in perception and conception. *Psychological bulletin*, 127(1):3, 2001.

[174] Lijuan M. Zhou and Cathal Gurrin. A survey on life logging data capture. In *SenseCam 2012: 3rd Annual Symposium*, SenseCam 2012, 2012.

[175] Lijuan Marissa Zhou, Niamh Caprani, Cathal Gurrin, and Noel E. O'Connor. Shareday: A multi-modal lifelog system for group sharing. In *MMM*, 2013.

[176] Lijuan Marissa Zhou and Cathal Gurrin. Memorymesh: lifelogs as densely linked hypermedia. In *Proceedings of the 4th International SenseCam & Pervasive Imaging Conference*, pages 90–91. ACM, 2013.

[177] Lijuan Marissa Zhou, Cathal Gurrin, and Zhengwei Qiu. Zhiwo: activity tagging and recognition system for personal lifelogs. In *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*, pages 321–322. ACM, 2013.

[178] Lijuan Marissa Zhou, Brian Moynagh, Liting Zhou, Tengqi Ye, and Cathal Gurrin. Memlog, an enhanced lifelog annotation and search tool. In *Multimedia Modelling*, 2015.

Appendices

Appendix A

Ethical Forms

Ollscoil Chathair Bhaile Átha Cliath
Dublin City University

29th April 2015

Dr Rami Albatal
INSIGHT

REC Reference: **DCUREC/2015/126**

Proposal Title: NTCIR-Lifelog

Applicant(s): **Dr Rami Albatal, Dr. Cathal Gurrin**

Dear Rami,

Further to expedited review, the DCU Research Ethics Committee approves this research proposal.

Materials used to recruit participants should note that ethical approval for this project has been obtained from the Dublin City University Research Ethics Committee.

Should substantial modifications to the research protocol be required at a later stage, a further submission should be made to the REC.

Yours sincerely,

Dónal O'Mathúna

Dr Dónal O'Mathúna
Chairperson
DCU Research Ethics Committee

DCU Research & Innovation

Taighde & Nuálaíocht Tacaíocht
Ollscoil Chathair Bhaile Átha Cliath,
Baile Átha Cliath, Éire

Research & Innovation Support
Dublin City University,
Dublin 9, Ireland

T +353 1 700 8000
F +353 1 700 8002
E research@dcu.ie
www.dcu.ie

Appendix B

MemLog System Retrieval Interface and Implementation

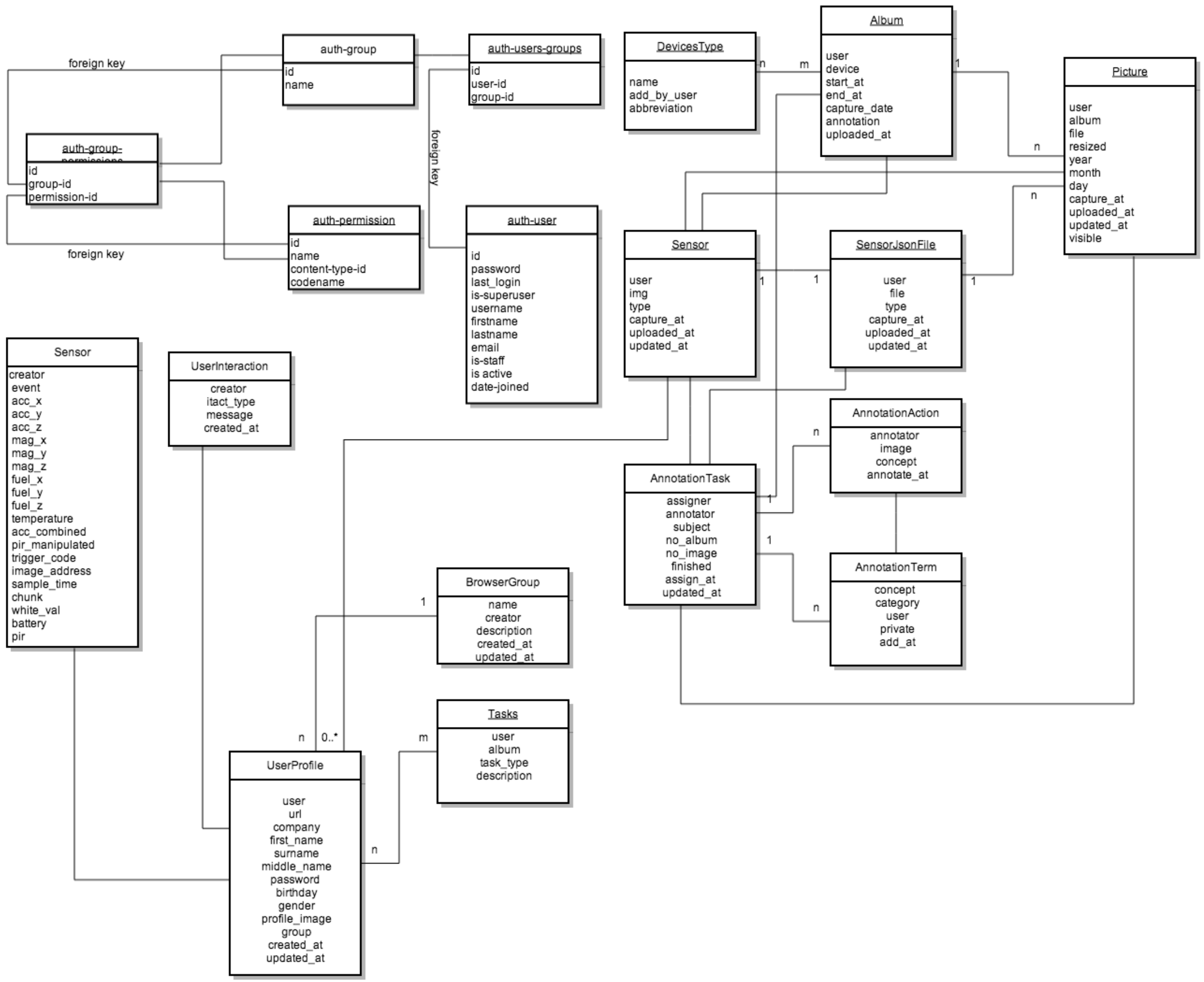


Figure B.1: Appendix figure MemLog ER diagram

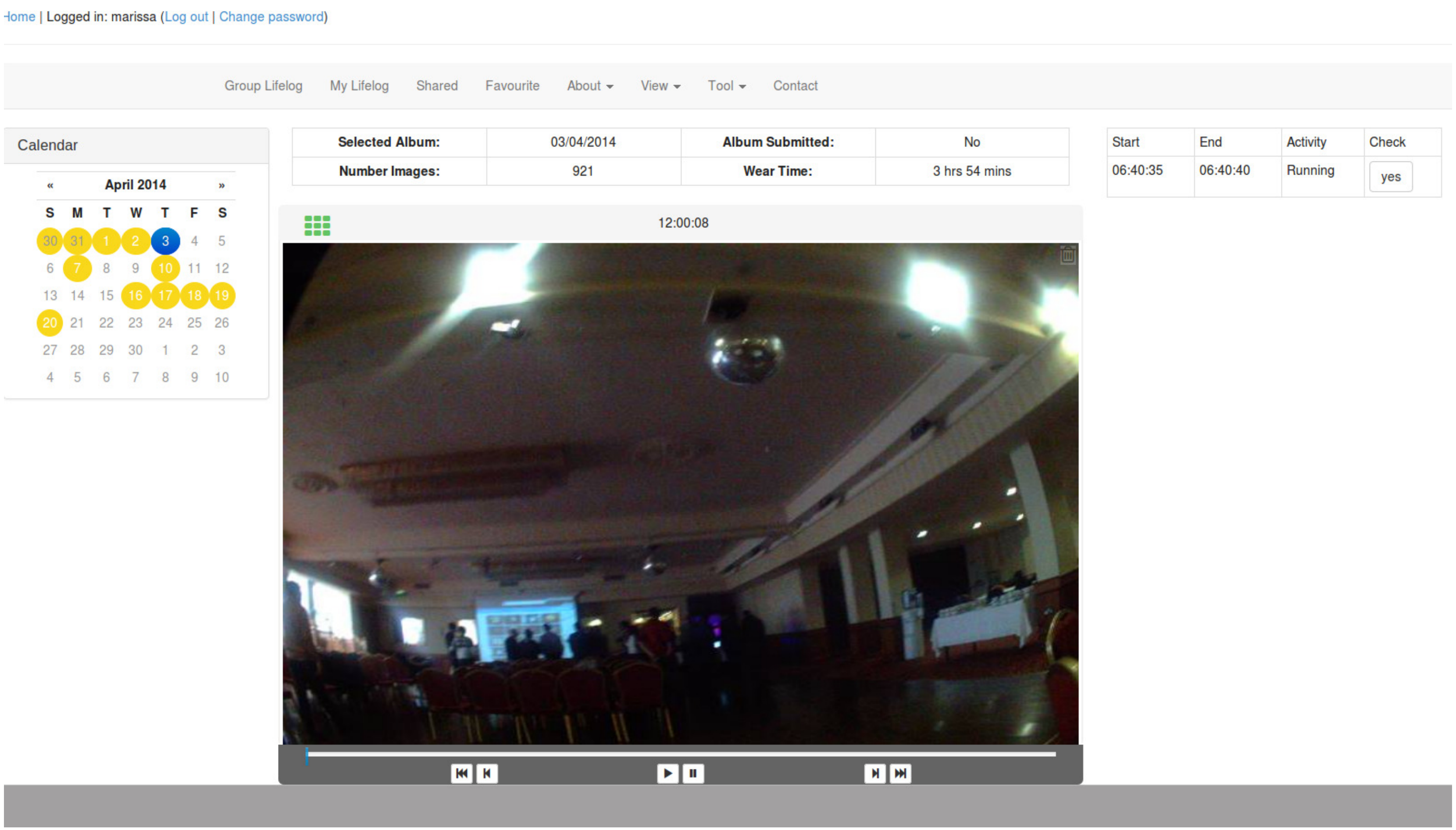


Figure B.2: MemLog screen-shot for the login view

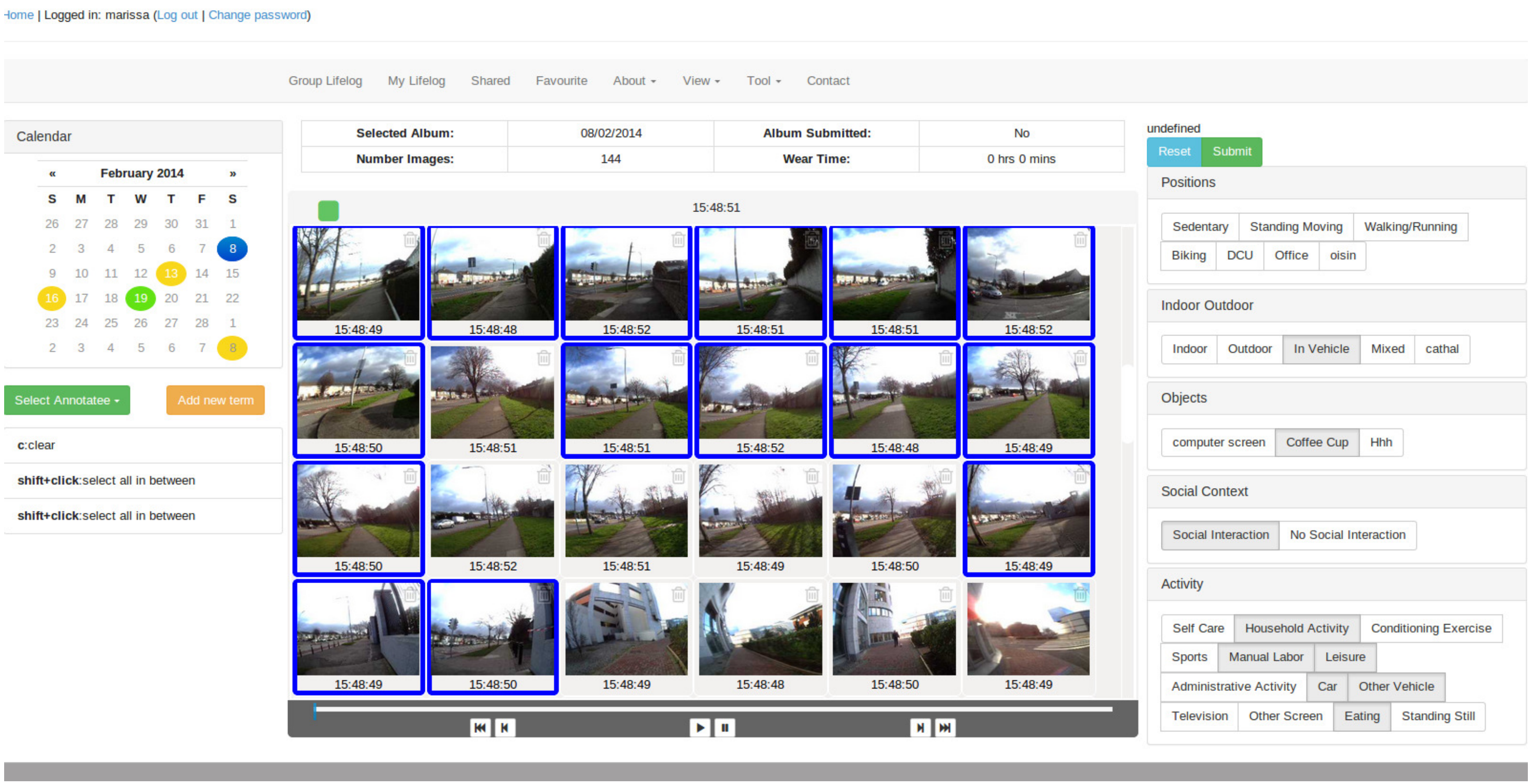


Figure B.3: MemLog Screen-shot for the thumbnail view

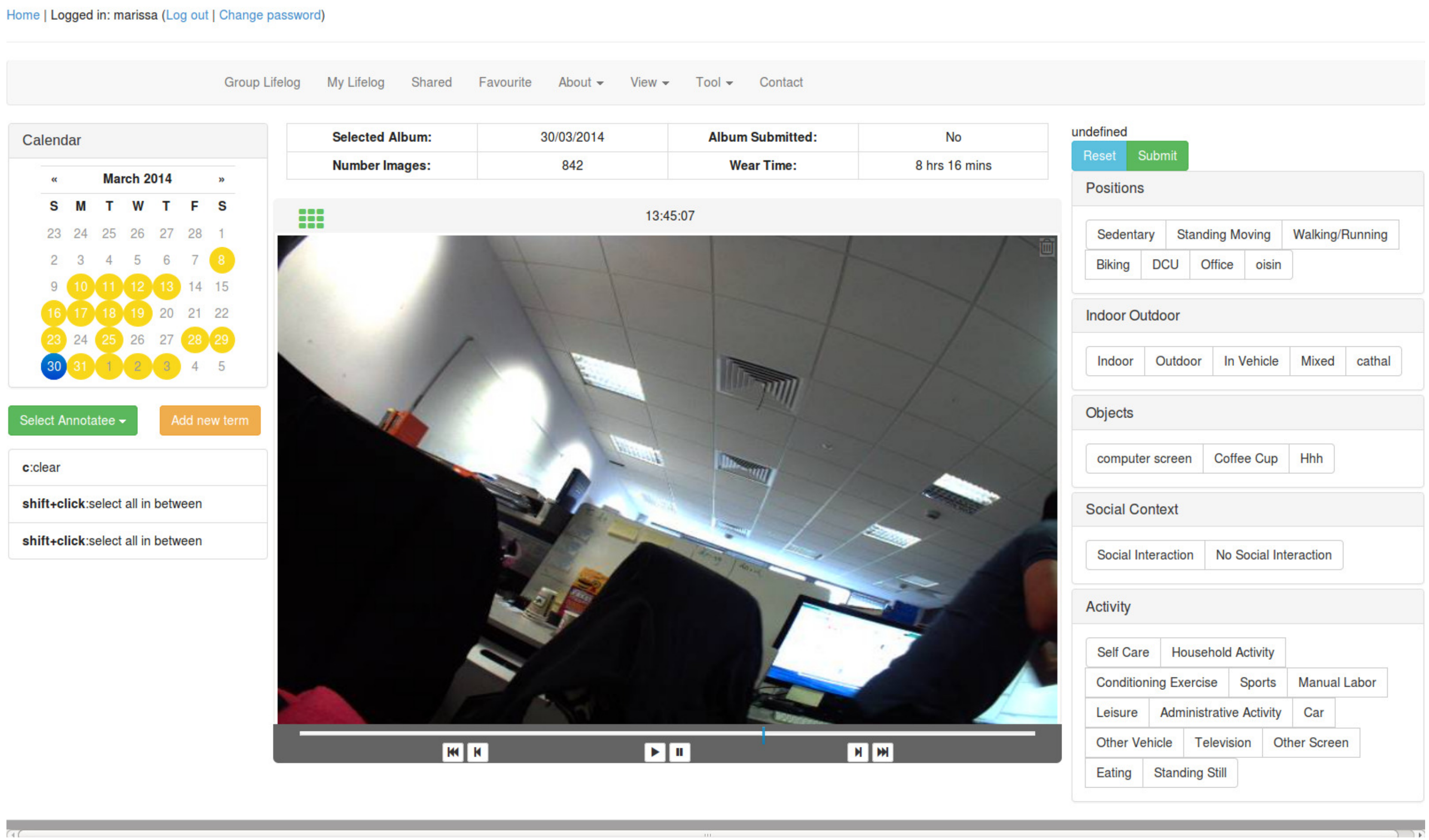


Figure B.4: MemLog Screen-shot for the annotation view

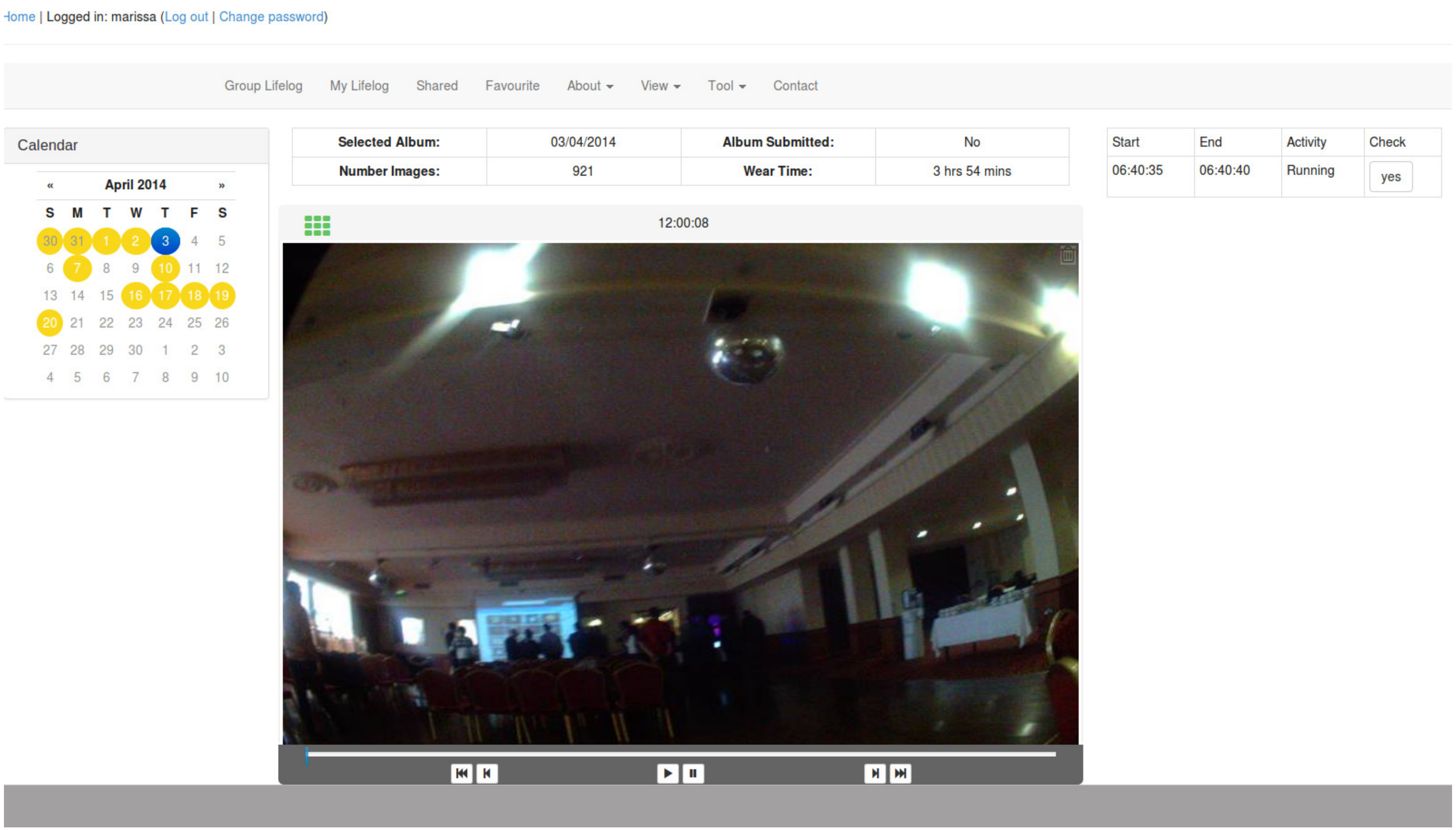


Figure B.5: MemLog Screen-shot for the activity view

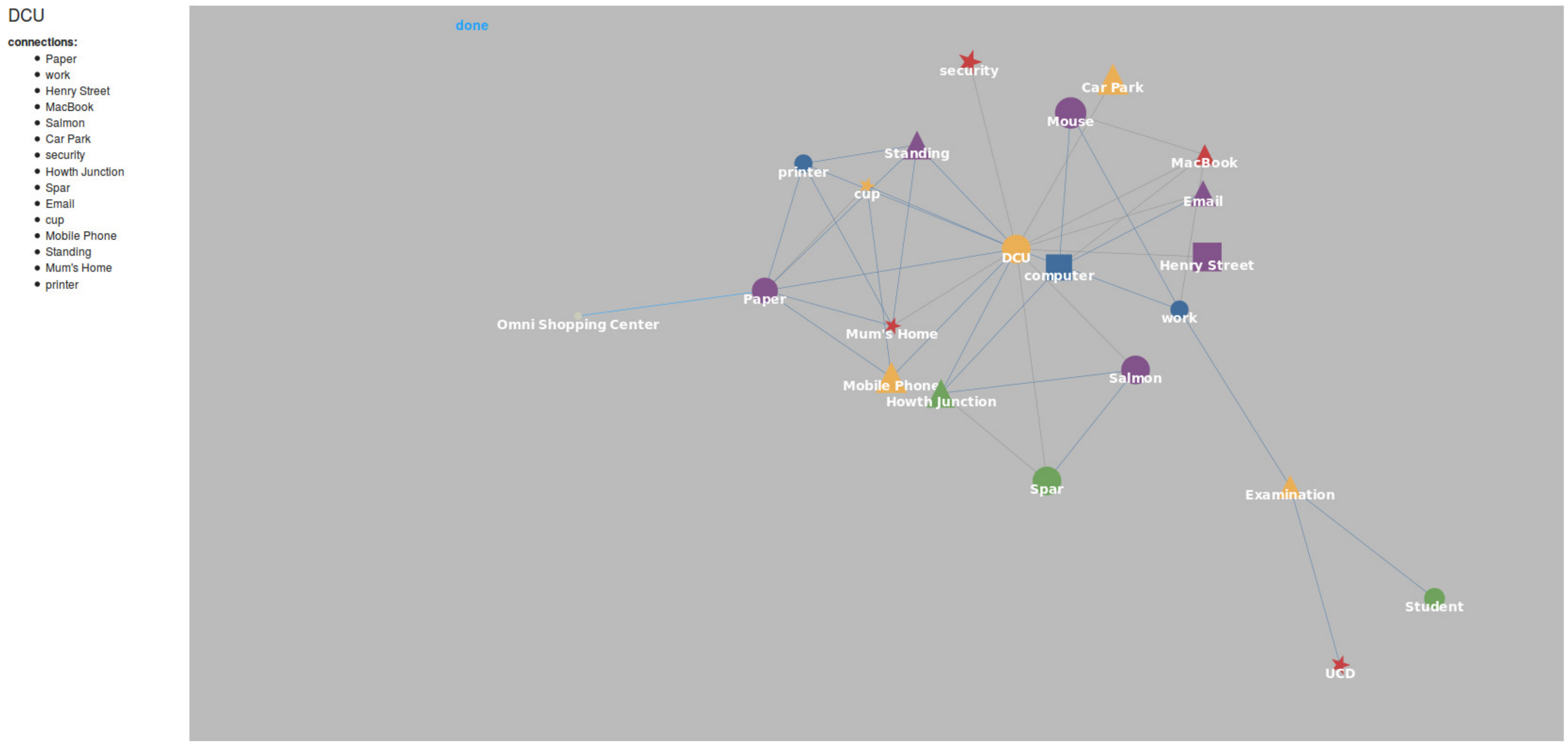


Figure B.6: Memory node linkage representation in the ZhiWo system

Appendix C

Survey on User Consent to Wearable Sensors

The source of this survey is under here[1].

[1] https://docs.google.com/spreadsheet/viewform?formkey=dFFJTWFQUG9uMnJObmpVcFN3Qkp6Nnc6MQ

Feeling to Wearable Sensors

This is a survey caring about people's feeling to wearable sensors, including wearable cameras, amateur fitness devices, professional biometric devices and location recorders. Thanks a million for doing this survey.

* Required

Part 1: Wearable Cameras

There are a variety of different wearable cameras available right now, such as SenseCam, Looxcie, Go-Pro, variable mobile phone applications and so on. These sensor devices can shot video or photos as your personal lifelog.

1. **What type of wearable cameras (if any) have you worn for a period of at least a few hours? ***

   Check all that apply.

   - [ ] SenseCam/Vicon Revue
   - [ ] Looxcie
   - [ ] Go-Pro
   - [ ] Wearable Camera Glasses
   - [ ] Photo Capture Mobile Phone Software (lifelapse, clarity mobile app etc.)
   - [ ] I have never worn a wearable camera
   - [ ] Other: ....

2. **If you have worn a wearable camera, for how long?**

   Mark only one oval.

   - ( ) Less than one day
   - ( ) More than one day but less than one week
   - ( ) More than one week but less than one month
   - ( ) More than one month but less than one year
   - ( ) More than one year

3. **If you have worn a wearable camera, how do you feel about wearing such a device?**
Mark only one oval.

- ( ) Very happy
- ( ) Happy
- ( ) Don't mind
- ( ) A little unhappy
- ( ) Very unhappy
- ( ) Other: ____________

4. **If you are unhappy with wearing a wearable camera in the last question, please specify the reasons.**

____
____
____
____
____

5. **If you have not worn a wearable camera, how do you think you would feel to wear such a device?**
Mark only one oval.

- ( ) Very happy
- ( ) Happy
- ( ) Don't mind
- ( ) A little unhappy
- ( ) Very unhappy
- ( ) Other: ____________

6. **In response to the previous question, why do you feel that way?**

____
____
____
____
____

7. **How do you feel if being captured by wearable cameras worn by others?** *
Mark only one oval.

- Very happy
- Happy
- Don't mind
- A little unhappy
- Very unhappy
- Other: ....

8. **In response to the previous question, why do you feel that way?** *

....
....
....
....
....

9. **Which of the following two devices you would feel more comfortable to be in the presence of ?** *
Mark only one oval.

- Wearable Photo Capture Devices
- Wearable Video (including audio) Capture Devices

10. **Why do you select that option in the previous question?** *

....
....
....
....
....

11. **Do you concern about the appearance of the wearable camera devices?** *
Mark only one oval.

- Yes
- No

12. **Wearable camera devices are believed to help offset some of the negative impacts of memory disability. Therefore, which of the following would best describe your opinion of when to begin using wearable cameras? ***

Mark only one oval.

- When I am healthy with no apparent memory disability
- At the onset of a memory disability
- Never

Part 2: Amateur Fitness Devices

Amateur Fitness Devices include Fit Bit, Nike + Pod etc that are able to store information such as the elapsed time of the workout, the distance traveled, pace, or calories burned by the individual wearing devices. Normally the amateur devices are small and portable.

13. **What type of amatuer fitness device (if any) have you worn for a period of at least a few hours?**

Check all that apply.

- [ ] Fit Bit
- [ ] Nike + Pod
- [ ] I have never worn any amateur fitness device.
- [ ] Other: ....

14. **If you have worn any Amateur Fitness Devices, for how long?**

Mark only one oval.

- Less than one day
- More than one day but less than one week
- More than one week but less than one month
- More than one month but less than one year
- More than one year

15. **If you have worn an amateur fitness device, how do you feel about wearing such a device?**

Mark only one oval.

- Very happy
- Happy
- Don't mind
- A little unhappy
- Very unhappy
- Other: ....

16. **In response to the previous question, why do you feel that way?**

....

....

....

....

....

17. **If you have not worn such a fitness device, how do you think you would feel to wear such a device?**
Mark only one oval.

- ( ) Very happy
- ( ) Happy
- ( ) Don't mind
- ( ) A little unhappy
- ( ) Very unhappy
- ( ) Other: ....

18. **In response to the previous question, why do you feel that way?**

....

....

....

....

....

19. **If you want to buy&wear an amateur device, which of the following will you consider?**
Check all that apply.

- [ ] Comfort to wear
- [ ] Easy to wear
- [ ] Good look
- [ ] Cheap
- [ ] Functional (can obtain your personal data that you want)
- [ ] Easy to upload data and to examine data analysis result
- [ ] Can easily share data with family and friends
- [ ] Other: ....

Part 3: Biometric Devices

It could be Heart Rate Monitor, BodyMedia SenseWear Pro, Equivital(Body sensor and GPS recorder), Readiband(Sleep tracking from your waist). These devices are mostly used to record biometric data, such as galvanic skin response (GSR) and skin temperature (ST) , physiological responses such as changes in heart rate or increased sweat production, sympathetic nervous activity. These data can be used for health analysis or professional sport training.

20. **What type of biometric devices (if any) have you worn for a period of at least a few hours? Please list one in one line.**

....

....

....

....

....

21. **If "yes" in the last question, how long did you wear this device?**
Mark only one oval.

- ( ) less than 1 day
- ( ) More than 1 day but less than 1 week
- ( ) More than 1 week but less than 1 month
- ( ) More than 1 month

22. **If you have worn an biometric device, how do you feel about wearing such a device?**
Mark only one oval.

- ( ) Very happy
- ( ) Happy
- ( ) Don't mind
- ( ) A little unhappy
- ( ) Very unhappy
- ( ) Other: ....

23. **In response to the previous question, why do you feel that way?**

....

....

....

....

....

24. **If you want to buy&wear a biometric device, which of the following will you consider?**
Check all that apply.

- [ ] Comfort to wear
- [ ] Easy to wear
- [ ] Good look
- [ ] Cheap
- [ ] Functional (can obtain your personal data that you want)
- [ ] Easy to upload data and to examine data analysis result
- [ ] Can easily share data with family and friends
- [ ] Other: ....

Part 4: Location Devices

GPS recorder

25. **What type of location recording devices (if any) have you worn for a period of at least a few hours? Please list one in one line.**

....

....

....

....

....

26. **If "yes" in the last question, how long did you wear this device?**
Mark only one oval.

- ( ) less than 1 day
- ( ) More than 1 day but less than 1 week
- ( ) More than 1 week but less than 1 month
- ( ) More than 1 month

27. **If you have worn an location recording device, how do you feel about wearing such a device?**
Mark only one oval.

- Very happy
- Happy
- Don't mind
- A little unhappy
- Very unhappy
- Other: ....

28. **In response to the previous question, why do you feel that way?**

....
....
....
....
....

29. **If you have worn an location recording device, why do you wear that gps device?**
Mark only one oval.

- Share location with family and friends
- Your phone has gps, which was turned on for better location based services
- For personal data collection, such how far have you travelled
- Other: ....

---

30. **Age** *

....

31. **Gender** *
Mark only one oval.

- Male
- Female

---

32. **We welcome any comment here!**

....

....

....

....

....