Aiding episodic memory in lifelog system focusing on user status

44161507-3 YE XIN

Supervisor: Prof. Jiro TANAKA

Graduate School of Information, Production and Systems
Waseda University

A thesis submitted for the degree of
Master of Engineering

September 2018

Abstract

Episodic memory, the collection of past personal experiences, is vital to human for maintaining past and anticipating future. Lifelog is the detailed record of an individual's life, involving a process of capturing, processing and recalling life experience passively. Therefore, lifelog is widely used in aiding episodic memory. With appropriate cues, images taken by wearable cameras can help lifeloggers recalling their past experiences. Most existing cues are focusing on contextual information which can be captured from lifelog images or wearable sensors. However, important user-related memory triggers like lifeloggers' own inner psychological sentiment and external physiological movement are not involved in current research. In order to improve the situation, we creatively proposed two user-related cues to extract important memory for lifeloggers. These two cues are special sentiment cue and special movement cue.

Our system consists of three parts: a web-based lifelog viewer for lifeloggers to conveniently retrieve memory, a sentiment detection system to analyze lifeloggers' sentiment situation using face photo and a movement detection system to capture lifeloggers' movement status. In order to retrieve our goal, we integrate three hardware devices into our system: an Autographer is fixed to a hat on lifeloggers' head for capturing face photo, another Autographer hang on lifeloggers' neck for getting lifelog images and a smart phone in lifeloggers' pocket for detecting movement situation. By using our proposed system, lifeloggers can efficiently browse lifelog images taken during specific sentiment and specific movement. Moreover, we also combine our user-related cues with traditional contextual cues in our system.

We have invited some participants to test the usability and efficiency of our system. The result have shown the positive potential of aiding episodic memory by using our approaches.

Keywords: Episodic Memory, Lifelog, Special Sentiment, Special Movement

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Jiro TANAKA, for the continuously support on my study and research. He is the person who leads me into HCI field and helps me find out my own way in lifelog research. During 2-year study in Waseda University, he uses his knowledge, patience, motivation and enthusiasm guides me a lot. He let me know the importance of reading paper from top conferences. He mentors me how to express my research perspective in seminars and presentations. He guides me to find my interest and develop it into research topic. He also gives me countless advice on how to build my system and how to write a master thesis. It is my great honor to have the opportunity to do research in IPLab and to be a student of Prof. Jiro TANAKA..

Secondly, I would like to thank all members from IPLab. Thank you for your tireless efforts to help me sort out my research ideas and put forward constructive suggestions for improvement. Thank you for guiding me when I feel down and encouraging me to move on. You are not only my best partner in my research, but also my best friend in my life. Thanks for your companionship and support.

Thirdly, I would like to express my grateful to volunteers who joined the experiment of this research. Thank you for taking time out of your busy schedule to help us get valuable experimental data.

Last but not least, I want to thank my roommates and my parents. You will always be my strongest backup.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Introduction

1.1.1 Episodic Memory

Memory is the ability of our neural system to encode, store and retrieve information. It is the accumulation of one person's past activity, feeling and many other information [1]. Memory can be divided into many categories according to different classification methods. As shown in Fig.1.1, the most common method is to divide memory into long-term memory, short-term memory and sensory memory according to memory retention time. Episodic memory is a part of long-term memory, together with semantic memory, it forms the integrity of memory system.

The concept of episodic memory first appeared around 1970s. Unlike semantic memory which concentrates on the actual facts about this world, episodic memory is the memory of autobiographical events. These autobiographical events including contextual information(who, what, when, where, why) which can be explicitly elaborated and memorized. Generally speaking, semantic memory helps us gain objective knowledge while episodic memory helps us recall past experiences and shape our personality. Episodic memory allows an individual to vividly go back in time to recall the event which occurs at a particular time and space [2].

For example, you can recall the memory of what has happened at your first date and what you have eaten for your last meal. These are all the results of recalling episodic memory.

Episodic memory is vital to human being and Tulving [3] muse about it as “a true marvel of nature”. Moreover, according to the experiment done by Stanley B. Klein [4], the loss of episodic memory not only impact individual’s memory for the past, but also may extend the influence of people’s ability to anticipate the future. However, Cabeza R’s study [5] proved that the brain regions involved in episodic-memory recognition is much more active than that of episodic-memory recall. Meanwhile, because of the limitation of brain capacity and some severe memory illness, it is essential for us to find some external solutions for aid episodic memory retrieving.

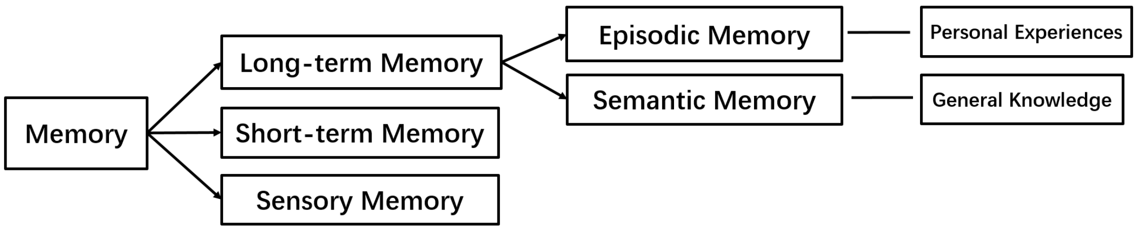


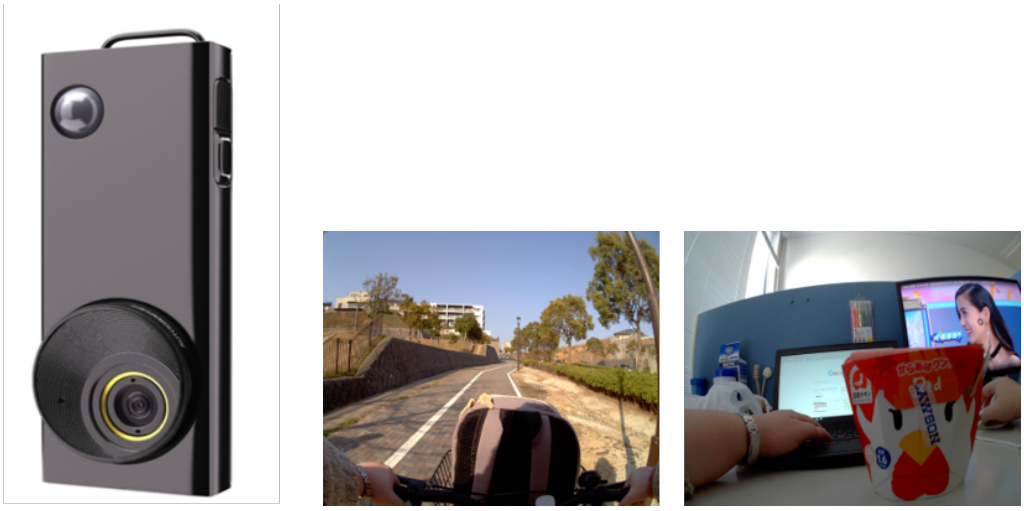
Fig. 1.1 Composition of memory

1.1.2 Lifelog

Lifelog is the detailed chronicle of personal life involving large number of data that can be captured automatically by wearable technology. It is a process of gathering, processing and recalling life experience passively. Individuals, defined as lifeloggers, will carry out multifarious sensors which can sense individual’s living environment and activities. [6]

Lifelog was first introduced to the world in 1980s by Steve Mann [7] who continuously captured physiological data together with live first-person video using a wearable camera. After that, lifelog has experienced many other experiment done by researches. With the significant development of wearable and storage technology, lifelog nowadays have become more and more popular and lifelog recording devices are easier to be accessed. The representatives are SenseCam, Vicon Revue, Narrative Clip and Autographer. Researches on lifelog mainly concentrate on three directions: (1) Using lifelog and other wearable technology to aid memory retrieve. (2) Using new interactions for lifelog sharing. (3) Discussion on lifelog privacy. Moreover, no matter what kind of research, lifelog always need viewer to review images which is taken using wearable camera. Typical lifelog viewer is mainly based on three platforms: web, desktop and smartphone. Among them, web stands out with its superior performance. Lifelog web viewer is graphical and easy to navigate. Due to its platform independent characteristic, lifeloggers can easily interact with lifelog viewer at any platform with a web browser.

Fig 1.2(a) shows Autographer, a widely used wearable cameras in lifelogging. Developed by OMG Life, Autographer is a wearable technology which is embedded with five different sensors to detect environment changing and will automatically take picture every 30 seconds [7]. Fig 1.2(b) shows some sample images taken by Autographer. We can clearly judge the environment condition of lifeloggers, so as to assist in memory retrieving.



(a) Autographer

(b) Sample Images

Fig. 1.2 Overview of Autographer

Visual information in personal lifelog can be very helpful for human memory aiding. Lee M L [8] proved that lifelogging technologies have the potential to provide memory cues for people who struggle with episodic memory impairment. And Aizawa K [9] have put forward five cues based on context and content for efficient retrieval of lifelog data. Moreover, Chen Y [10] introduced a work-in-process prototype lifelog searching system to augment human memory, which has obtained positive experimental results.

1.2 Organization of the Thesis

The rest of the thesis is organized as follows: In chapter 2, we describe the research purpose and the approaches we use. Chapter 3 will give the detailed introduction of how the target system is designed. Chapter 4 presents the implementation of how to utilize sentiment cue and movement cue in the system. Chapter 5 is about the evaluation which is used to prove the usability and efficiency of our system. Chapter 6 states 2 related work and some background on emotion detection and activity detection. Chapter 7 concludes the work and will discuss the future expectation.

Chapter 2

Goal and Approach

2.1 Problem

Episodic memory is vital to human for supporting life activities, including life management, social interactions and problem solving in daily life. Due to the limitation of brain function, aiding episodic memory using lifelog system has become more and more popular. This new approach not only can benefit those with significant memory impairment, but also has the potential of making normal individual's life more effective.

For aiding memory, we need to define memory triggers in lifelog system, which is defined as cue [11]. Most existing cues in lifelog research focus on contextual information like how, what, when, where and who. Representative is Lee ML [12] who use objects detected from lifelog image as memory cue and Chowdhury S [13] who involved GPS heat map as memory cue. We can find that these cues have nothing to do with users' own behavior and care little about users themselves.

However, according to the investigation by Sellen A J [14], psychology as design framework can help define the types of memory. Moreover, in our daily life, important memory always happens with the following two scenarios:

- Specific sentiment. When people feel sad, surprise, happy or other special emotion, it means some special event has happened.

- Sudden change of body shape. When people change their body movement from walking to still or from still to cycling, this always means the end of previous event and the start of a new event.

- Meet with others. When people meet with others, important memory always involved.

- Exposed to certain objects. When people is exposed to some certain objects, important memory might be reminded.

If we summarize the scenarios mentioned above, the latter two can be considered into contextual information which has already been implemented by researchers like Chowdhury [13]. The former two scenarios can be summarized as user-relative and research in aiding episodic memory using lifelog discard user's won status when considering memory cues. This is the problem we want to improve by using our proposed method.

2.2 Research Goal

In this research, we aim to involve user's own status into lifelog system, so as to help finding important memory for lifeloggers. Here, user's own status including inner psychological mood and external physiological activity. For a clearer explanation of our research perspective, we corresponds the user status to two cues:

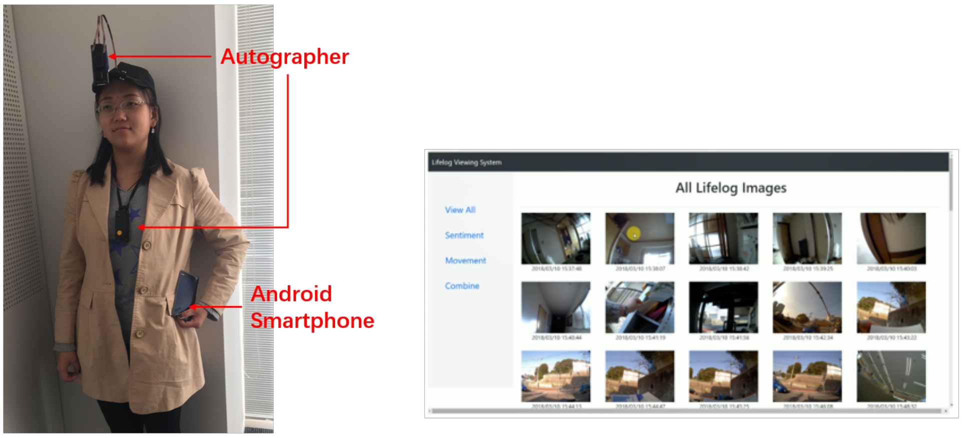
- Special sentiment cue.

- Special movement cue.

We innovatively use these two cues to propose a new user-related lifelog viewing prototype for aiding episodic memory more efficiently.

2.3 Research Approach

Our proposed system mainly consists of two parts: a wearable sensor system and a web-based lifelog viewer. Unlike traditional contextual lifelog cues, we introduce personal sentiment and movement condition as new cues to enhance the effectiveness of memory recall. To achieve this goal, we set up a wearable sensor system (see Fig 2.1(a)). We use Autographer attached on user's head to automatically capture face photo, which is used in emotion detection. Meanwhile, we use an Android smartphone embedded with motion sensors like accelerometer and gyroscope to continuously get data to acquire movement situation. Moreover, we use another Autographer hang on user's neck to capture lifelog image constantly.



(a) Hardware Overview (b) Lifelog Viewer

Fig. 2.1 Proposed System

By using our wearable sensor system, we can get pure sentiment data and movement data. After that, these two data will be input into the process system to judge sentiment and movement situation when each lifelog image is taken. Because of the cross-platform compatibility and usability of web, then we build a web-based lifelog viewer with the output. The appearance of the viewer is shown in Fig 2.1(b). Lifeloggers can easily access important memory by using two proposed new cues in the viewer.

2.4 Novelty and Potency

1. In order to enhance the efficiency of aiding memory, we novelty put forward two new user-relative cues for memory retrieve in lifelog system. They are special sentiment cue and special movement cue.

2. We use wearable camera (Autographer in our research) for observing emotion situation by constantly capturing face photo.

3. In our research, we use movement data not only for memory retrieving, but also for event segmentation purpose.

Chapter 3

System Design

3.1 System Overview

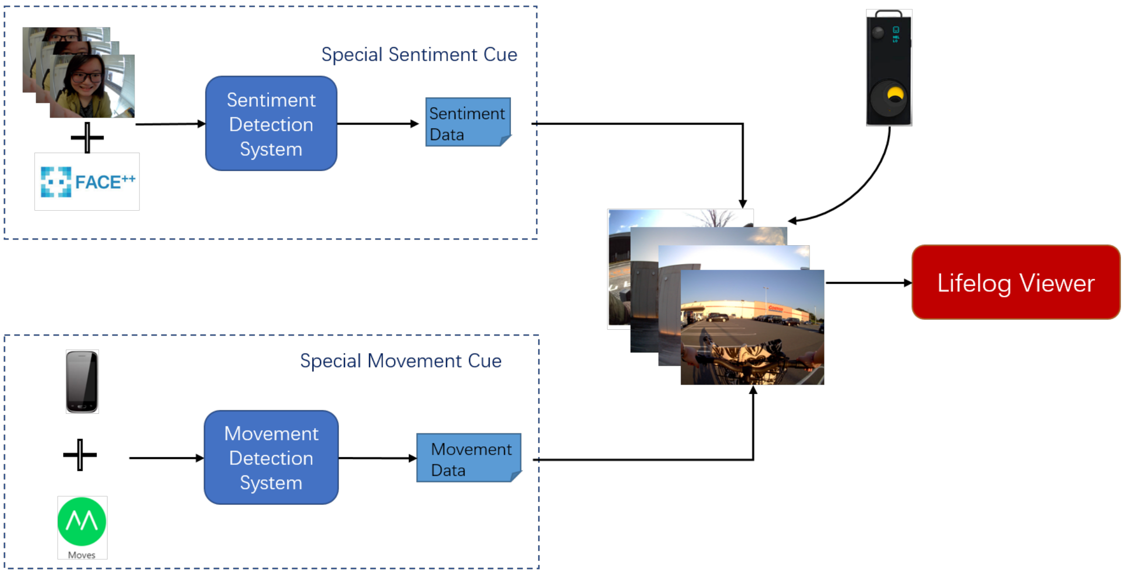


Fig. 3.1 System Overview

Our proposed method aimed to enhance the efficiency of aiding episodic memory. To achieve our goal, we come up with the idea of providing two user-related cues for lifeloggers. The system overview is shown in Fig. 3.1. The system is mainly consisting of 1 web-based lifelog viewer and two accessory sub systems: a sentiment detection system used to capture sentiment data and a movement detection system used to capture movement data. To get the sentiment data, we use face photo, together with Face++ API to form a sentiment detection system so as to get required sentiment data. For special movement cue, we use accelerometer in Android smartphone and Moves API to get movement data. After getting these two data, we combine them with lifelog image to form a lifelog viewer.

3.2 Special Sentiment Cue

3.2.1 Sentiment Process Procedure

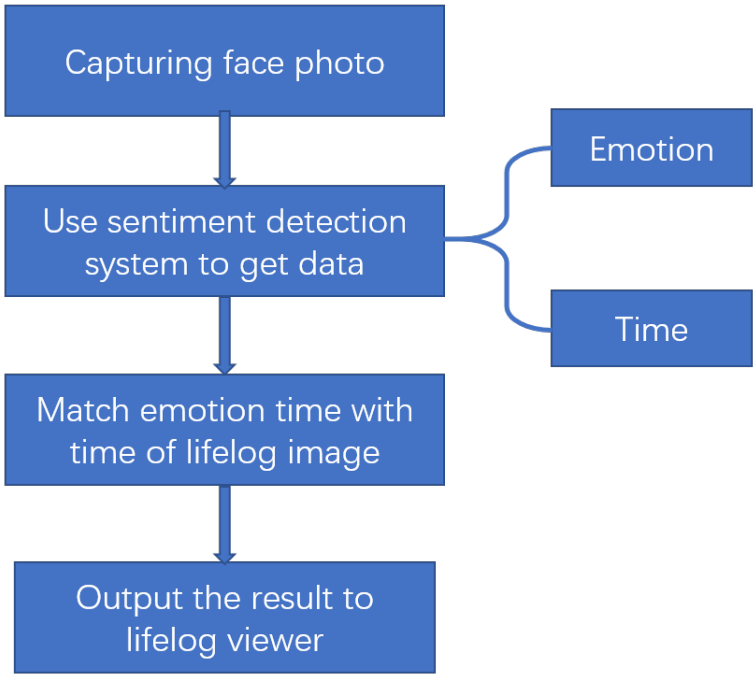


Fig. 3.2 Sentiment Process Procedures

Fig 3.2 shows the main procedure of processing sentiment data. Firstly, we need to capture face photo automatically using Autographer. Secondly, the face photo will be input into sentiment detection system to get sentiment data. Here, emotion data consist of two parts: the emotion status and the time information which shows when this emotion has been detected. Thirdly, we match the time of each emotion with the record time of each lifelog image. Finally, we output the result to lifelog viewer for lifeloggers. Fig 3.3 shows the procedure more vividly which shows an example. In the example, the face photo was detected as happiness and it was detected at 2018/03/18 15:12:38. Then, we can find the lifelog image taken at this time. Finally, we can judge that this lifelog image was detected as "happiness".

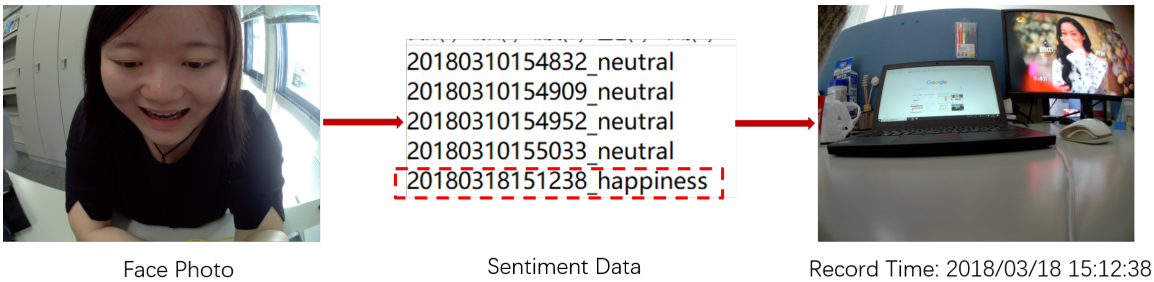


Fig. 3.3 An Example of Sentiment Process Procedures

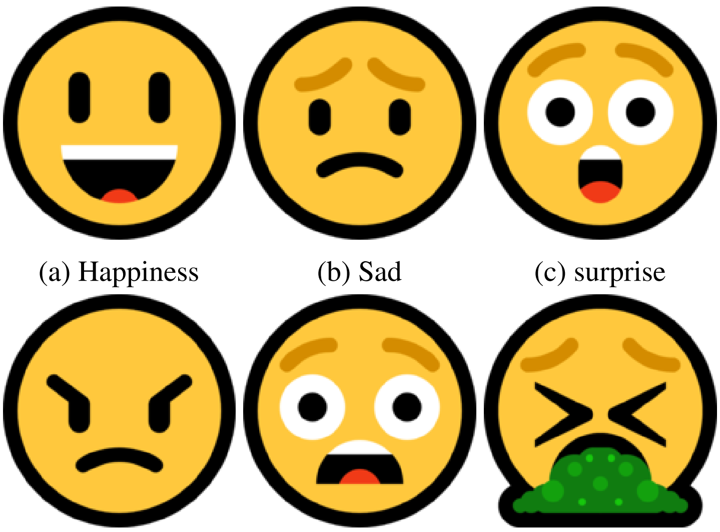
3.2.2 Sentiment Classification

Human's emotions are complex and diverse. They can be categorized from different perspectives. At the end of the 1860s, Paul Ekman conducted a cross-cultural study in the tribes of New Guinea, Papua, and then suggested that the facial expressions corresponding to the emotions were universal and had nothing to do with culture. Therefore, he suggests that facial expression is a biological instinct that can be classified safely and correctly. Since then, he formally put forward six basic emotions in 1972 [26]. Sentiment classification in our system uses six basic emotions put forward by Paul Ekman. Six emotions are:

- Happiness. Happiness is a mental or emotional state of well-being which can be defined by positive or pleasant emotions ranging from contentment to intense joy.
- Sad. Sad is an emotional pain associated with feelings of disadvantage, loss, despair, grief, helplessness, disappointment and sorrow.
- Surprise. Surprise is a brief emotional state experienced as the result of an unexpected significant event.

- Anger. Anger is an intense emotional response involving a strong uncomfortable and hostile response to a perceived provocation, hurt or threat.
- Fear. Fear is a feeling induced by perceived danger or threat that occurs in certain types of organisms
- Disgust. Disgust is an emotional response of revulsion to something considered offensive, distasteful, or unpleasant.

In order to demonstrate emotions more clearly and vividly, we create several avatar to represent each emotion. Fig 3.4 shows the emoji avatars. (a) – (f) represents happiness, sad, surprise, anger, fear and disgust.



(a) Happiness (b) Sad (c) surprise

(d) Anger (e) Fear (f) Disgust

Fig. 3.4 Emotion Avatar

3.3 Special Movement Cue

3.3.1 Movement Process Procedure

Fig 3.5 shows the procedure of processing movement data.

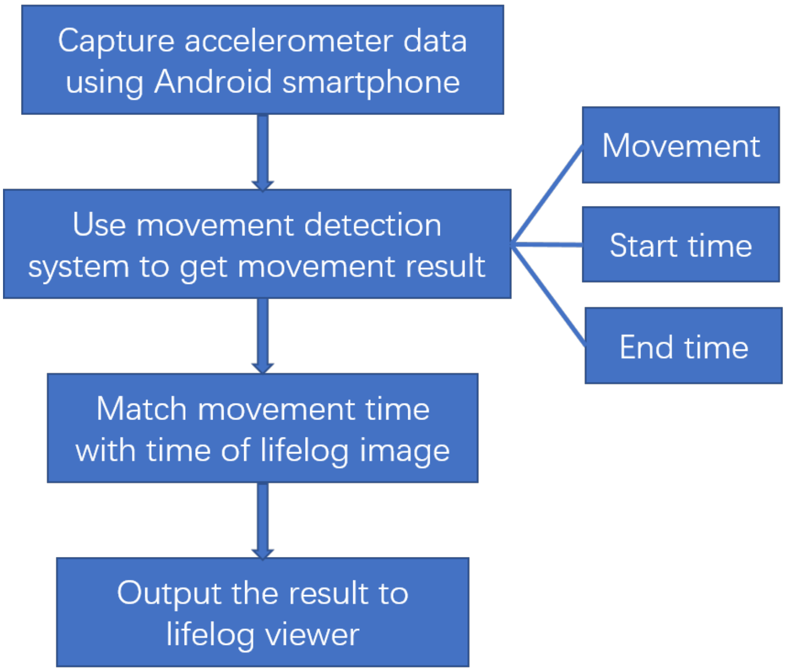


Fig. 3.5 Movement Process Procedures

Firstly, we will use accelerometer embedded in Android smartphone to capture the pure movement data of lifeloggers. Then, we import the raw accelerometer data into movement detection system. Together with Moves API, we can clearly get the result of movement situation. The result consists of variety information, including date, movement, start time, end time, etc. We mainly use movement, start time and end time in our research. After getting movement situation, we will match movement time with the time of lifelog image. By matching time, we can clearly find out the movement situation of each lifelog image. Finally, we will output the result to lifelog viewer. Fig 3.6 shows movement process procedure vividly by showing an example. In the example, the walking status is detected from 2018/3/11 15:33:03 to 2018/3/11 15:55:18. For the lifelog image showing in the right side, the record time is 2018/3/11 15:40:44. It is explicit that this lifelog image belongs to movement "walking". Therefore, the movement status of this lifelog image is judged as "walking".

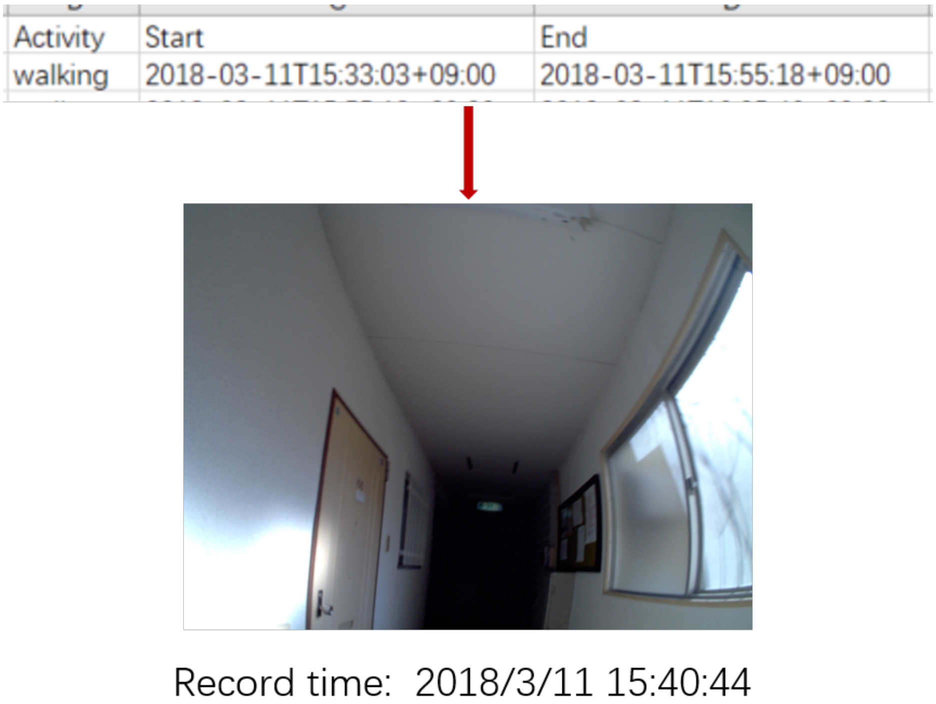


Fig. 3.6 An Example of Movement Process Procedures

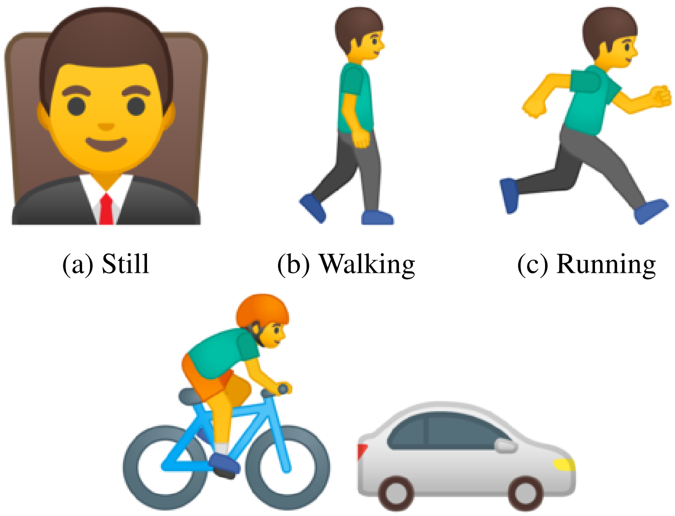
3.3.2 Movement Classification

Move is the change in position of an object over time in physics. Move is described in terms of displacement, distance, velocity, acceleration, time and speed. Human movement describes the way human moves. The classification of movement is quite complicated. To simplify the result, our proposed method involves mainly five kinds of movement:

- Still. Still is used to describe the situation when human has no significant movement. It includes sitting, standing, lying, etc.
- Walking. Walking is one of the main gaits of locomotion of human. It means using two feet for slowly movement.
- Running. Running is a method of terrestrial locomotion which allow humans to move rapidly on foot. The speed of running is faster then walking.
- On bike. Biking/on bike is the use of bicycles for transport, recreation, exercise or spot.

- On vehicle. On vehicle means human use all kinds of vehicle for transport, including truck, saloon, van. jeep, etc.

In order to demonstrate movments more clearly and vividly, we also create several avatars to represent each movement. Fig 3.7 shows the avatars. (a) – (e) represents still, walking, running, on bike and on vehicle.



(a) Still (b) Walking (c) Running

(d) On Bike (e) On Car

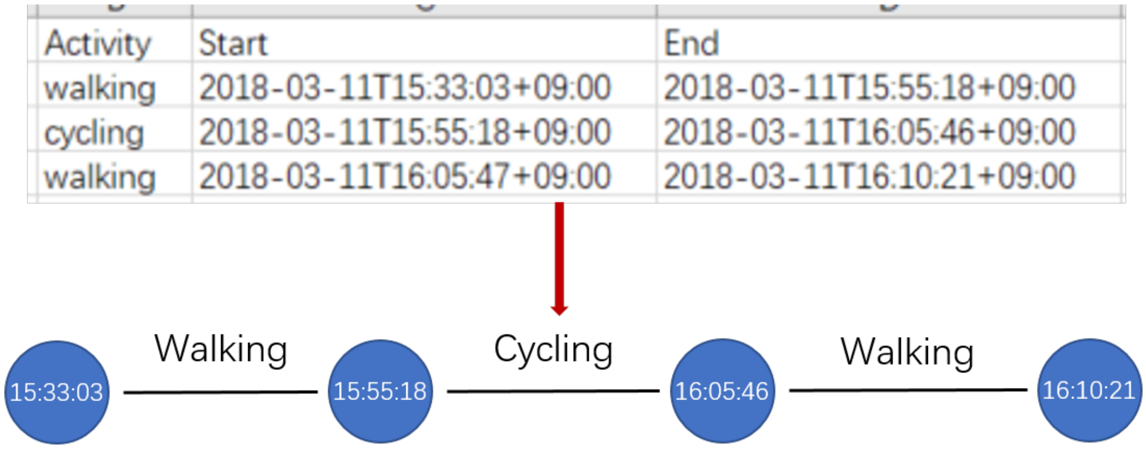
Fig. 3.7 Movement Avatar

3.3.3 Movement Timeline

Timeline refers to the connection of one or more aspects of events in series by Internet technology, forming a relatively complete recording system and then presenting it to the user in the form of graphic or text. Timeline can be used in various fields and the greatest role is to systematize, complete and exact the past. Timeline not only has an unparalleled superiority in the field of event segmentation, but also has the ability of making abstract data visible.

In our research, we mainly involve five movement status: still, walking, running, on bike and on vehicle. They can basically represent all movement status. Moreover, our monitoring on movement is continuous and uninterrupted. Therefore, based on the principle of data visualization, we use movement data to generate a behavior timeline in order to make movement cues more clearly and remarkable.

Fig 3.8 shows an example of how to generate timeline using movement data. We can see that the extracted movement data including three movement periods. The movement situation of each period is walking, cycling and walking. Because movement monitor is constant, we can figure out that the end time of first walking period is the start time of cycling period, and the end time of cycling period is the start time of second walking period. Therefore, we can generate an uninterrupted movement timeline according to the movement status, start time and end time information of each period. The generated timeline is shown in the figure. From the timeline, we can judge that from 15:33:03 to 15:55:18, user is walking. From 15:55:18 to 16:05:46, user is cycling. From 16:05:46 to 16:10:21, user is walking again. By using the timeline, user can fairly see his/her movement situation. Combined with lifelog image taken at each movement period, it is easy to segment events and use movement cue to recall episodic memory.



Activity	Start	End
walking	2018-03-11T15:33:03+09:00	2018-03-11T15:55:18+09:00
cycling	2018-03-11T15:55:18+09:00	2018-03-11T16:05:46+09:00
walking	2018-03-11T16:05:47+09:00	2018-03-11T16:10:21+09:00

Fig. 3.8 Movement Timeline Generation

3.4 Cue Combination

In order to help memory retrieve, memory cue in lifelog has involved all aspects, including variety of contextual information (where, when, who, what, why knowledge). At the same time, researchers are exploring what kind of cues are more useful and efficient among them. One of the representatives is Liang[15]. In their work, they set up an experiment to test the efficiency of using each contextual cue and build a lifelog viewer based on the result. In their work, they use mainly four cues as followed:

1. Object (what). They use computer vision API from Microsoft to analyze objects from lifelog images and use objects like sky, bike etc. as a memory cue.

2. Location (where). They use Google maps API and GPS information captured using Autographer to generate location information for each lifelog images. Lifeloggers can browse images from different location groups.

3. Face (who). By using Microsoft's face recognition API, their system can extract human's face from lifelog images. Different faces mean lifeloggers are meeting with different people. Therefore, face can be an important memory cue.

4. Time (when). Time information can be extracted from file name of lifelog images. An example of file name is "B00005534_21I7BF_20180310_160027E.jpg". The latter two parts show the date and time information of the image. Therefore, the time result is 2018/03/10 16:00:27 in our example.

Liang's work has achieved quite positive result and their way of interaction is user-friendly. As their work concentrated on contextual cues and our proposed approach concentrate on user-related cues, we try to combine these two different cues to form a stronger combination. Therefore, in cue combination part, we will demonstrate six kinds of cues: Object, Location, Face, Time, Sentiment and Movement. Meanwhile, in order to make our system easy to operate, we provide "multi-select" for users. In cue combination part, lifeloggers can select multiple cues (e,g, location+face+sentiment) and then shows the related lifelog images. By using this method, users can awake those memory with blurry cues and can better recall a specific event.

3.5 Usage Scenario

Fig 3.9 shows the usage scenario of our proposed system. Lifeloggers need to wear some wearable devices to enable the data acquisition of our system. The devices include an Autographer attached on hat, an Android smartphone and another Autographer on neck. After recording lifelog information, lifeloggers need to import the data in these 3 devices into specific file path. After that, a web-based lifelog viewer will be generated. Lifeloggers can use special sentiment cue and special movement cue to recalling their episodic memory efficiently by using the lifelog viewer.

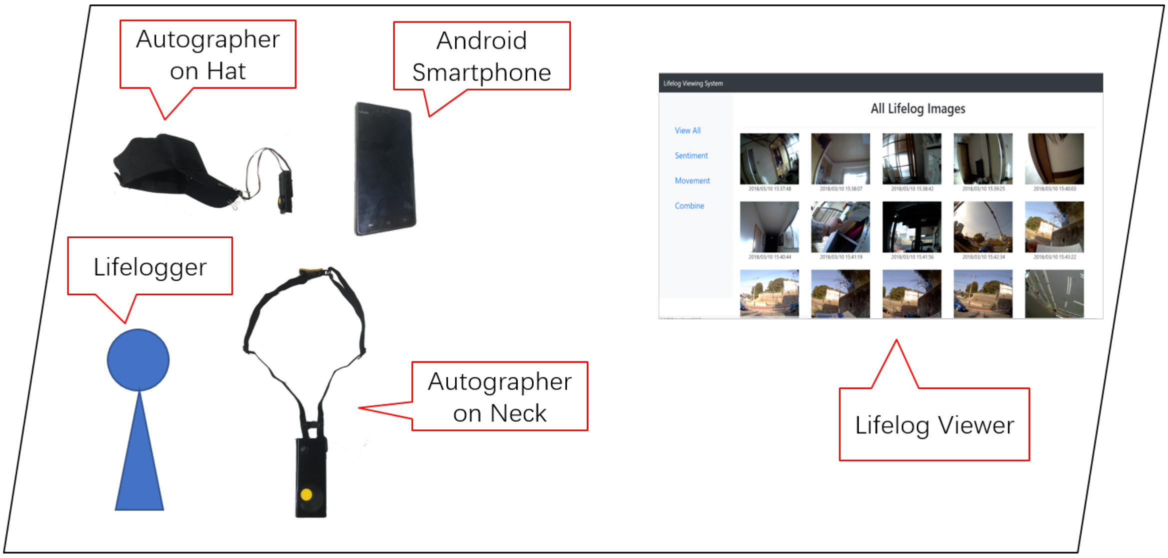


Fig. 3.9 Usage Scenario

Chapter 4

System Implementation

4.1 Hardware Implementation

4.1.1 Hardware Overview

Fig. 4.1 Hardware Overview

Fig 4.1 shows hardware overview. For using our system, users need to carry the following three wearable devices: an Autographer attached to a hat on lifelogger's head for face photo acquisition, another Autographer hang on lifelogger's neck for lifelog images capturing, an Android smartphone in lifelogger's pocket for movement data acquisition. To sum up, our system mainly involve two kinds of hardware: Autographer and Android smartphone. The usage of 2 hardware will be described in detail in section 4.1.2 and section 4.1.3.

4.1.2 Autographer

With the development of lifelog, there are more and more wearable devices that are developed for lifelogging. Wearable cameras are one of them. These cameras can passively and constantly capturing lifelog images. Autographer is a hands-free wearable camera that is used in current research. It is 90x36mm in size and weighs approximately 58 grams which makes it easy to clip on any clothes or hang on a neck. The camera is 5 megapixels and has a 136-degree wide-angle lens which imitates human eyes [16]. Without external interference, the capturing frequency of Autographer is 30 seconds per image. However, unlike other wearable cameras, the capturing frequency can change due to changes in external environment. Autographer is embedded with some sensors, including color sensor, magnetometer, PIR, temperature GPS, etc. When there is a sudden change of sensor value, Autographer will take an extra image which make lifelog images more valuable. Therefore, Autographer is favored by lifelog researchers for its superior performance and relatively intelligent shooting.

(a) Autographer

(b) Face Photo

Fig. 4.2 Autographer for Face Photo

Our system mainly involved two Autographers. One for capturing face photos and one for capturing lifelog images. In order to take face photo, we attach an Autographer to a hat as is shown in Fig 4.2 (a). Lifeloggers only need to wear this hat for emotion detection purpose. This device will record face photo constantly. The example of face photo is shown in Fig 4.2 (b). Another Autographer is hanging on lifelogger's neck for capturing traditional lifelog image. Fig 4.3(a) shows the Autographer and the neck sling. Fig 4.3 (b) shows an example of lifelog image.

(a) Autographer (b) Lifelog Image

Fig. 4.3 Autographer for Lifelog Image

4.1.3 Android Smartphone

Smartphone is a mobile phone that has a mobile operating system and can expand functions by installing applications, games and other programs. Android is a free and open source based operating system based on Linux, mainly used in mobile devices like smartphones. It is led and developed by Goggle and the open mobile alliance. As one of the most popular mobile operating systems, Android smartphone have covered more than half in the market according to the survey by Kantar. Nowadays, with the development of mobile technique, smartphones are always embedded with variety of sensors like accelerometer, gyroscope, gravity sensor, magnetic field sensor, etc. These sensors can be read by developer using Android SDK and is useful in detecting smartphone holder's movement condition.

Fig 4.4 (a) shows the Android smartphone that is used in our research. It is X5 ProD developed by Vivo. Developers can read sensors embedded in Android smartphone using Android SDK. Fig 4.4 (b) shows an example of reading accelerometer's value in x, y, z axis.

We can find that if we simply put our smartphone parallel to the ground, the value in x, y, z axis will be approximately equal to 0, 0, g (gravity unit).

(a) Smartphone

(b) Sensor Reader

Fig. 4.4 Android Smartphone for Movement Detection

4.2 Software Implementation

4.2.1 Development Environment

Hardware Environment

Personal computer from IPLab: Lenovo Thinkpad x260

Software Environment

In order to develop our system, we mainly use Java/HTML/CSS/JavaScript language and IDE MyEclipse. As we involve Android devices in our system, we also use Android SDK. For the web server part, we use Apache Tomcat 7.0 x. Meanwhile, we use MySQL for database. The running environment of our web viewer is Chrome and Firefox.

4.2.2 Structure and Framework

Browser/Server (B/S) Structure

Along with the rise of Internet technology, B/S structure is a change and improvement structure of Client/Server (C/S). Under this structure, user interface is implemented through a WWW browser. Very little transaction logic is implemented in the front end (browser), and the main transaction logic is implemented on the server side (server). The advantages of B/S structure are that it greatly simplifies the load scale of client computer, reduces the cost of system maintenance and upgrades and reduces the total cost of user. The B/S structure of our system is shown in Fig 4.5. The structure is consisting of three layers: presentation layer, application layer and data layer. There will be request and reply transaction between each layer. User's behavior is mainly happening at presentation layer. User can use our system at any platform with a web browser.

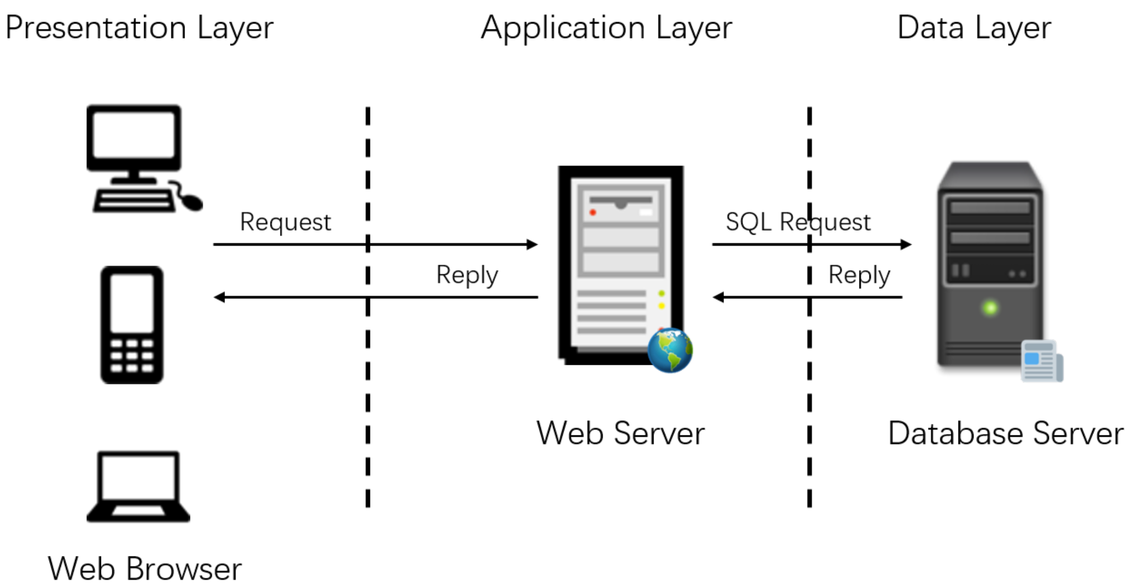


Fig. 4.5 B/S Structure

Struts+Spring+Hibernate (SSH) Framework

SSH is an integrated framework of Struts+Spring+Hibernate. It is a popular open source framework for web application. SSH framework is divided into four layers to help developers build web application with clear structure: presentation layer, business logic layer, data persistence layer and domain module layer. Therefore, SSH framework have the advantage of good reusability and convenient maintenance. In SSH framework, Struts is the overall infrastructure of the system and control business jump. Hibernate framework is used to support the persistent layer. Spring is in charge of managing the Struts and Hibernate. The overall framework is shown in Fig 4.6. Persistence layer uses hibernate framework and SQL to control operation to database. Business layer use Spring framework to control business code. DAO classes are used to communicate between business layer and persistence layer. Struts framework is used in Presentation Layer. This layer involve some .jsp files which show user interface and some .xml files in charge of jumping between each .jsp files.

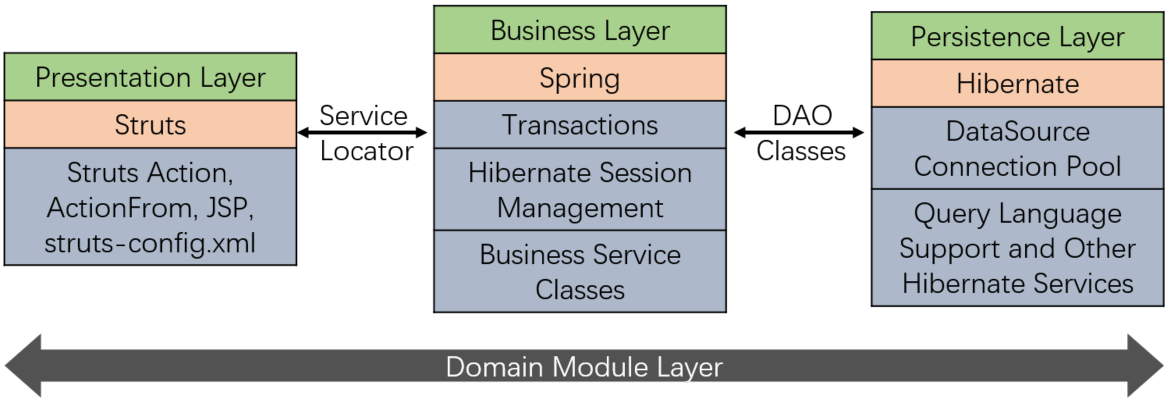


Fig. 4.6 SSH framework

4.3 API Introduction

4.3.1 Face++ API

Introduction:

Face++ can analyze the emotion of the detected face and return the confidence level of the face in different kinds of emotions. The higher the confidence level of a certain emotion is, the closer it is to the real emotion of that face. Currently, Face++ can identify anger, disgust, fear, happiness, neutral, sadness and surprise 7 kinds of most important emotions. The images used for emotion detection must be less than 2 MB and the form need to be JPG(JPEG) and PNG.

Request URL:

https://api-cn.faceplusplus.com/facepp/v3/detect

Call methods:

POST

Request Parameters:

Table 4.1 Request Parameters of Face++

Field	Type	Description
api_key	String	The API key to call this API
api_secret	String	The API secret to call this API
image_url	String	The URL of face photo
return_attributes	String	Attributes(Emotion in our research) that need to be returned.

Return Value:

Table 4.2 Return Value of Face++

Field	Type	Description
request_id	String	The unique string to specify each request
facest	Array	An array to represent detected face (including attributes)
image_id	String	The unique identification for face photo
time_used	Int	The total time required for the request in milliseconds
error_message	String	Returns only when request failed

Core Code:

```java
public String Face(String fileName) throws Exception {

    File file = new File(fileName);
    byte[] buff = getBytesFromFile(file);
    String url = "https://api-cn.faceplusplus.com/facepp/v3/detect";
    HashMap<String, String> map = new HashMap<String, String>();
    HashMap<String, byte[]> byteMap = new HashMap<String, byte[]>();
    map.put("api_key", "1tZeg15CPDu_MGA16f-Mw1_PWq0i4k97");
    map.put("api_secret", "AZGOFY3iDKcgq6iIECvPGGJyKwSPq_gI");
    map.put("return_attributes", "emotion");
    byteMap.put("image_file", buff);
    try {
        byte[] bacd = post(url, map, byteMap);
        String str = new String(bacd);
        return str;

    } catch (Exception e) {
        e.printStackTrace();
        return null;
    }
}
```

Fig. 4.7 Core Code of Invoke Face++

4.3.2 Moves API

Introduction

Moves is an APP developed by Protogeo which can automatically record smartphone holder's movement situation. The company provides both Android and iOS version. We mainly use the Android one in our proposed approach. Moves can automatically record user's movement information at the backstage. Therefore, user (lfelogger in our case) just need to keep their smartphone in their pocket or their bag for data capturing. Moves provide mainly four kinds of movement: walking, cycling, running and on vehicle. Considering the characteristics of human movement, we can also analyze the still status from data provided by Moves.

API Base URL

http://api.moves-app.com/api/1.1

Return Value

Moves API returns the value in the format of JSON. Fig 4.8 shows part result of returning value example. The return value consists of a date and a summary of that date's activity. In the example, the detected activity is walking. The important attribute of each activity is the name, the start time, the end time, duration, distance, steps and calories. We use name, start time and end time in our research.


```
"type": "move",
"startTime": "20180312T071430+0200",
"endTime": "20180312T074617+0200",
"activities": [
    {
        "activity": "walking",
        "group": "walking",
        "manual": false,
        "startTime": "20180312T071430+0200",
        "endTime": "20180312T072732+0200",
        "duration": 782,
        "distance": 1251,
        "steps": 1353,
        "calories": 99
    },
 ]
```


Fig. 4.8 Part of Return Value

4.4 Special Sentiment Cue

Default Page.

Fig 4.9 shows the default page of our lifelog viewer. The default page demonstrates all lifelog images which is ordered by time. By clicking the “sentiment” button in the sidebar, we can then go into the sentiment page.

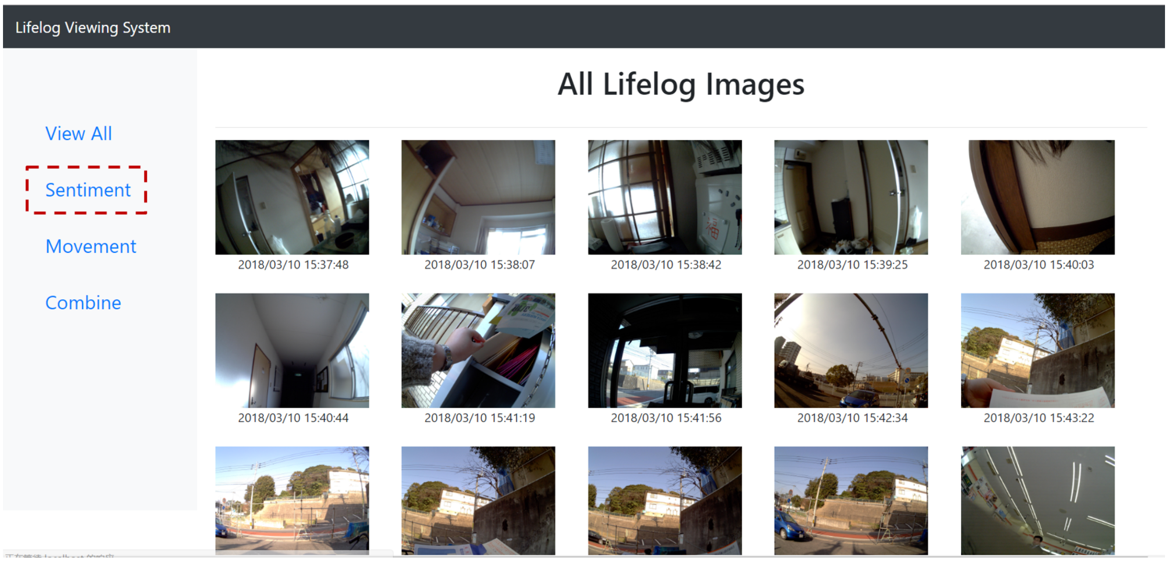


Fig. 4.9 Default Page

Sentiment Page.

Fig 4.10 shows the result of using special sentiment cue to classification all lifelog images. We can find that in the sentiment page, we mainly involve six kinds of sentiment, they are happiness, sad, surprise, anger, fear and disgust. In order to make emotion vividly, we use emoji to represent each emotion.

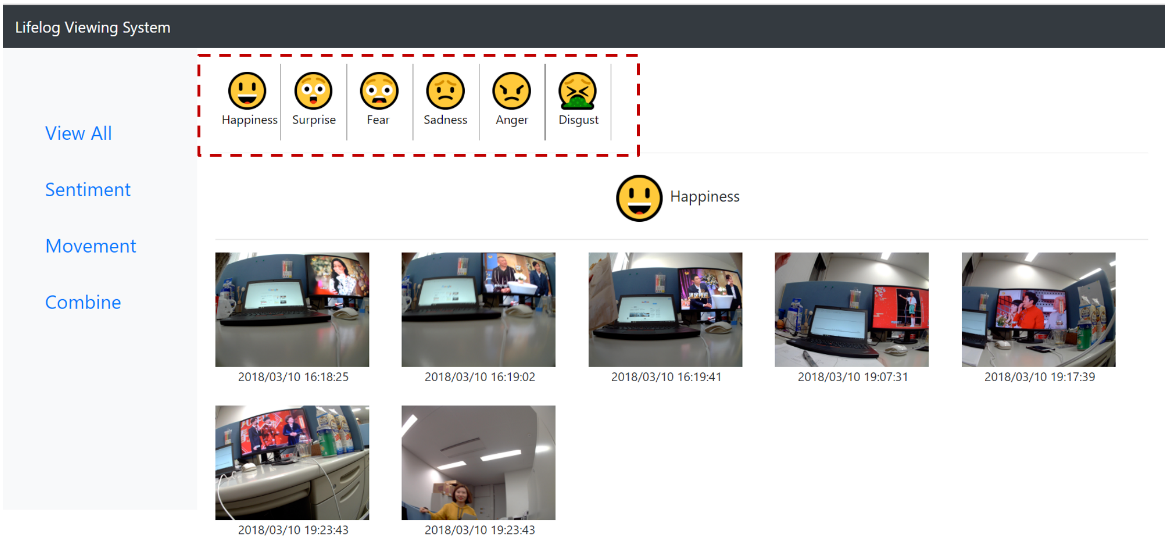


Fig. 4.10 Sentiment Page

Sentiment Result.

Fig 4.11 shows how we present sentiment result. The corresponding lifelog image of each emotion will appear under each sentiment's signal.

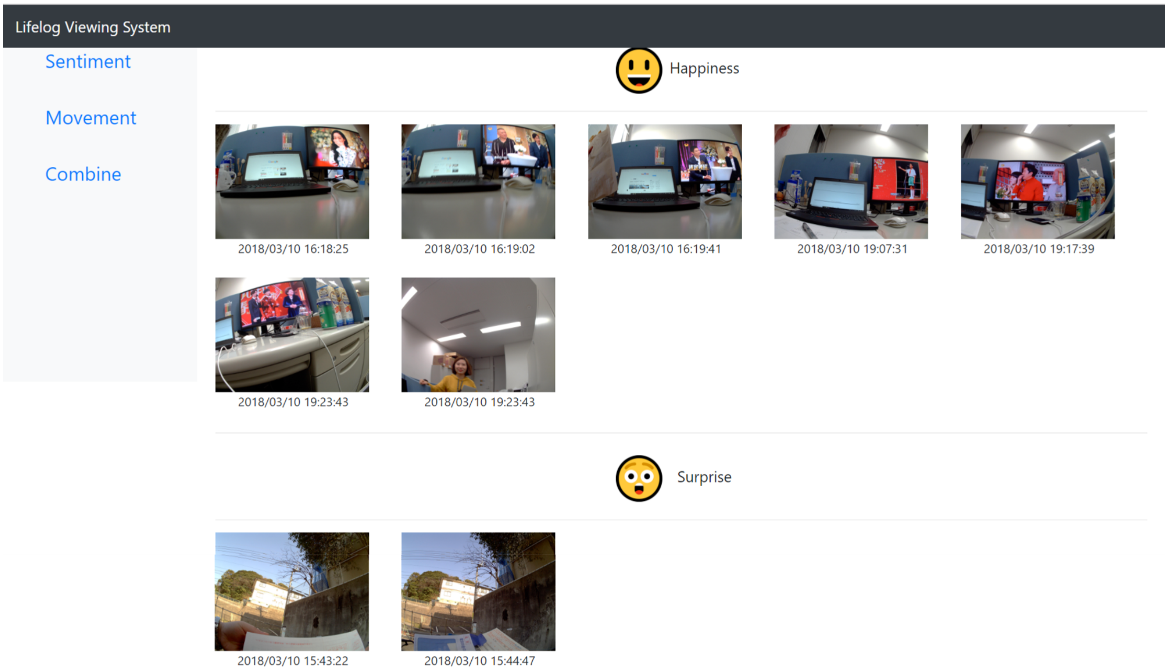


Fig. 4.11 Sentiment Result Page

4.5 Movement + Sentiment Cue

Default Page.

Fig 4.12 shows the default page of our lifelog viewer. By clicking "Movement" button on sidebar, we can enter movement section.

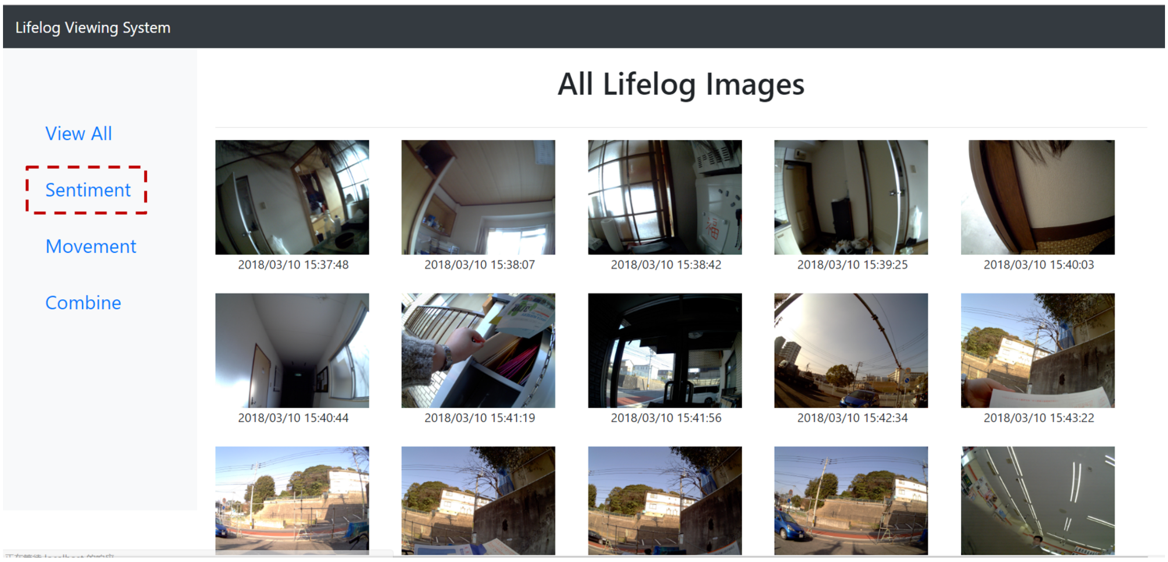


Fig. 4.12 Default Page

Movement Timeline.

Fig 4.13 shows movement timeline. we can see that we involve still, walking, running, on bike and on vehicle and there is a movement timeline showing in the middle of the page. By clicking the "view" button of each movement phase, we can enter to browse the detailed lifelog image of each segmented event.

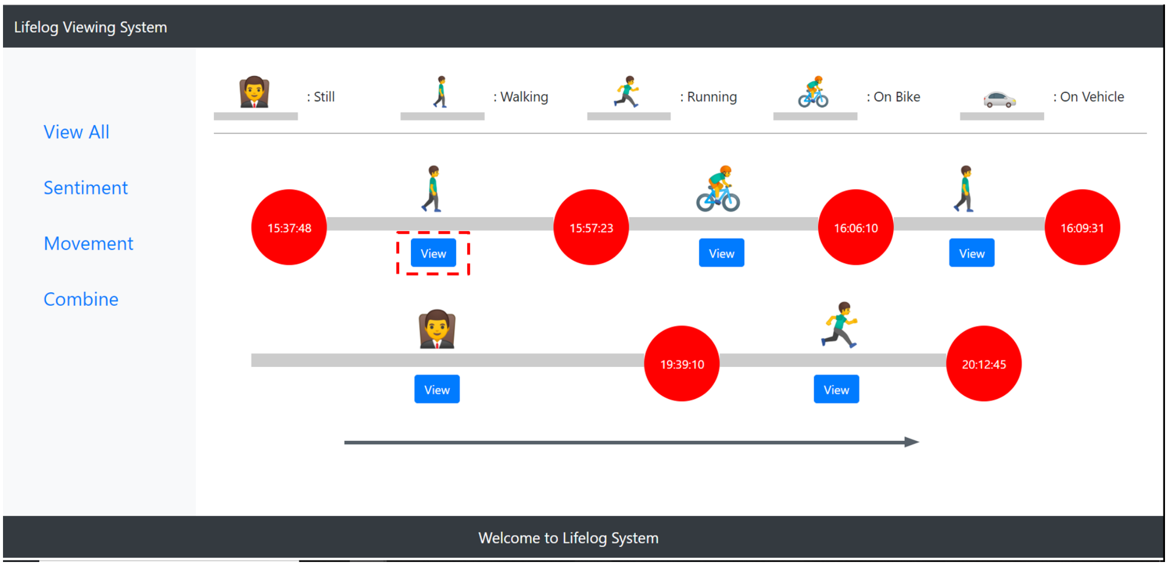


Fig. 4.13 Movement Timeline

Fig. 4.14 Movement Detailed Page

Detailed Page.

Fig 4.14 shows detailed page. The page consists of three parts: slides section(a), choose bar(b) and result section(c). In slide section, the images corresponded to each movement will be played automatically. This can reduce the time of browsing all images. As choose bar shows 6 sentiment emoji, lifeloggers can click the button to view the images satisfy both specific movement and specific sentiment. The result of images will be shown in result section.

4.6 Cue Combination

Default Page.

Fig 4.15 shows the default page of our lifelog viewer. By clicking "Combine" button on sidebar, we can enter cue combination section.

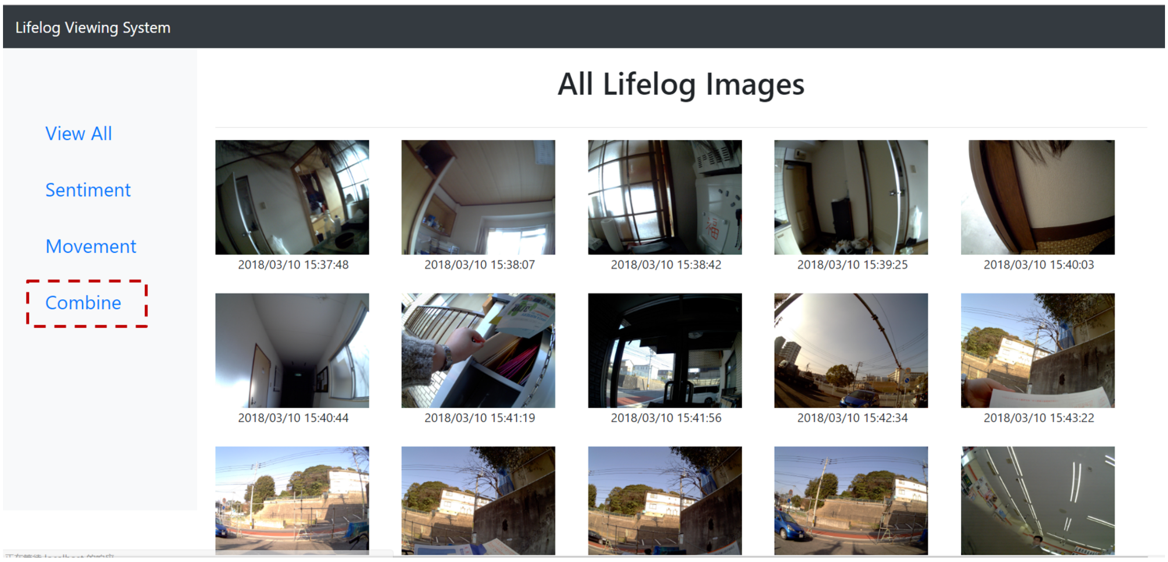


Fig. 4.15 Default Page

Cue Combination Page.

Fig 4.16 shows cue combination page. We can see that there are 6 kinds of cues: object, location, face, time, sentiment and movement. The design of this page is based on Liang's work. Users can click cues they want to combine, for example, Indoor, Specific Face and

Happiness in Fig 4.16. After that, users can click "Click to Confirm" button to view the combination result.

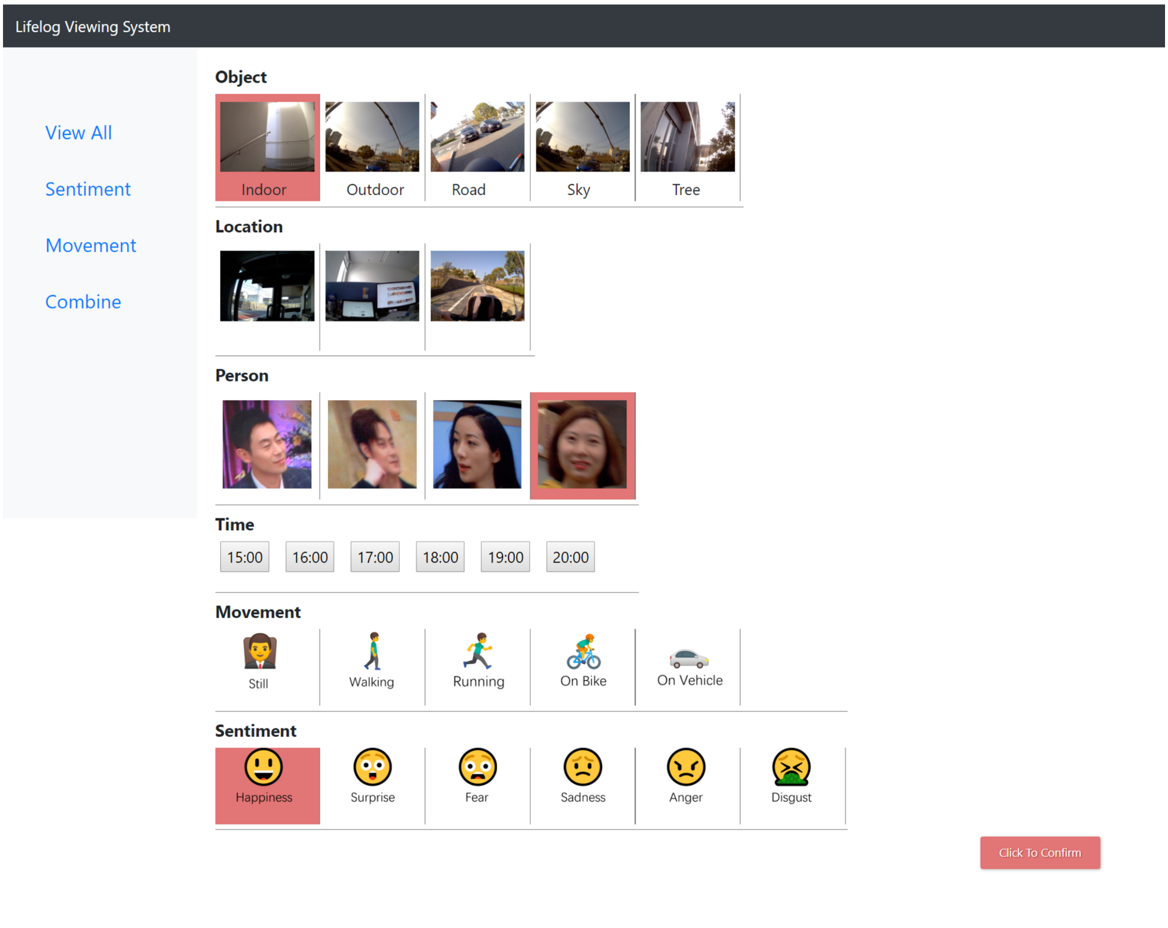


Fig. 4.16 Cue Combination Page

Combination Result Page.

Fig 4.17 shows combination result page. After choosing combination cues, the result will pop up.

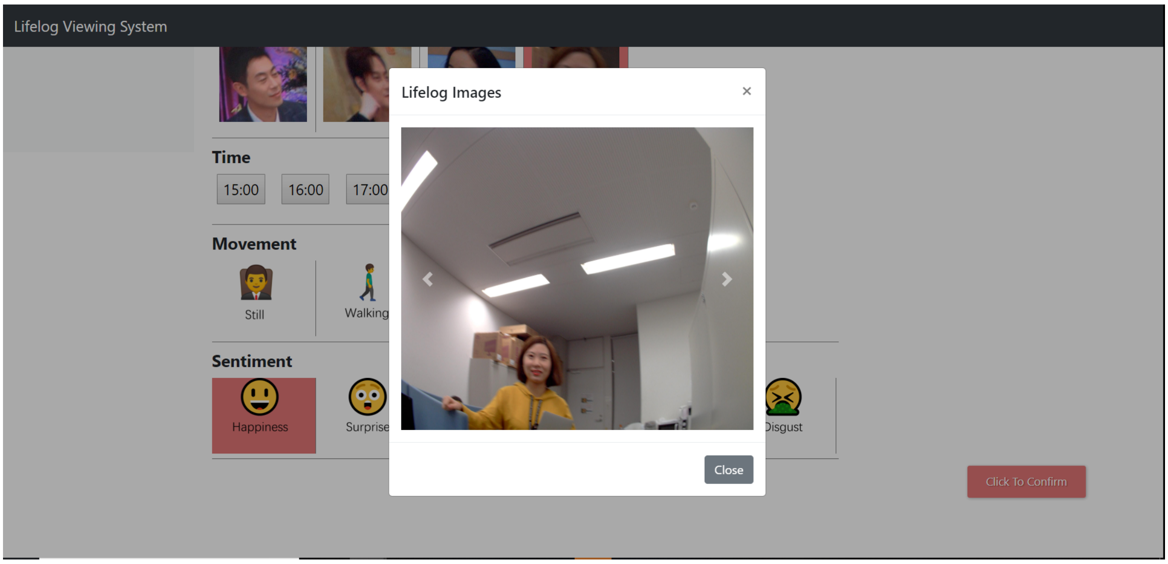


Fig. 4.17 Combination Result Page

Chapter 5

Evaluation

5.1 Participants

As Table 5.1 shows, in order to evaluate usability and efficiency of using our system, we recruited 6 participants, aging from 22 to 26. All participants are students who have general knowledge of computer and have the experience of using web browser. Before our experiment, there is a pre-description of how to use our devices and system.

Table 5.1 Subject Demographic Information

Participants	3 males. 3 females
Age	22-26; Mean: 22.3
Profession	Students

5.2 Method

The method we used for evaluation is described in the following 3 steps:

1. Each participant is asked to use our wearable devices for successive 5 hours in 1 day. There are no strict regulations on when to start. However, we strongly recommend choosing 10:00am to 3:00pm as recording period.

2. After recording, hardware devices can generate three data: sentiment data, movement data and lifelog images. These data are imported into our proposed lifelog system and a related lifelog viewer is generated.

3. Participant then tried to use our web viewer by their own. Our viewer contains three parts: movement cue viewer, sentiment cue viewer and a combination cue viewer. Movement cue viewer extracts important images and orders by using movement timeline. Sentiment cue viewer extracts vital memory using participant's sentiment situation. Combination cue viewer let users choose their own cue preference. Participants are required to record how many events they can reminisce by using each viewer (shown in Table 5.2).

Table 5.2 Questionnaire for Efficiency

Question	Answer
How many events you can remember before using our system?	
How many events you can recall by using special sentiment cue?	
How many events you can recall by using special movement cue?	
How many events you can recall by using combination cue?	

4. After using our system, a personal interview is given to each participant. The interview is mainly based on following questionnaire. The answer is grading from 1 to 5 (1 = very negative, 5 = very positive).

Table 5.3 Questionnaire for Usability

Question	1	2	3	4	5
Do you think it is comfortable to wear our devices?					
Do you think our viewer is easy to use?					
Do you feel extracted images useful?					
Do you think our system help in aiding memory intuitively?					

5.3 Result

In our evaluation, all participants have successfully completed the plan and have given effective feedback. In order to get the result of our evaluation, we collected 6 questionnaires from 6 participants (3 males and 3 females) and analyzed their feedback. Since our evaluation is divided into two aspects: efficiency and usability, our results analysis will also be carried out separately in these two aspects. Q3-Q6 reflects the efficiency of our proposed system. To get the result, we calculate the average amount of each question from 6 participants. As shown in Fig 5.1, the result of Q3, Q4, Q5 and Q6 is 6.50, 8.00, 9.83, 7.17. From the result, we can see that compare to recalling memory in default manner (not using any assistant), using any method in our proposed system can increase the number of recalling events. With the combination of movement cue in its interface, we can find the result of using movement cue shows better performance. It is surprise to find that combination cue do not have so positive feedback. In fact, during the interview after evaluation, participants mentioned that when using combination cue to recall events, at least one trigger need to be memorized. Therefore, it is not so easy to recall memory by using combination cue.

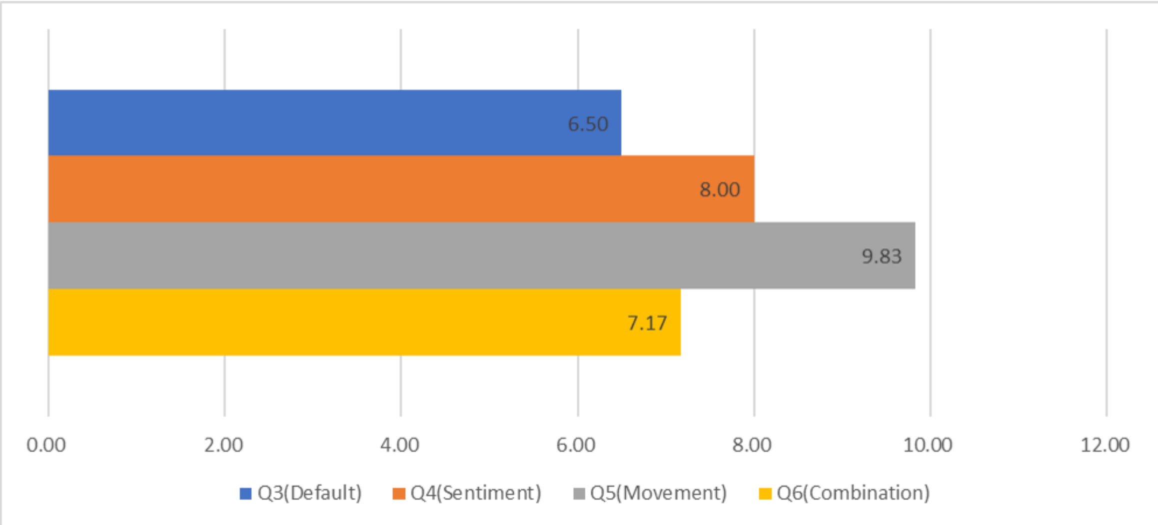


Fig. 5.1 Analyze Result of Q3-Q6

Another important aspect of our evaluation is the usability of proposed wearable device system and web-based viewing system. The answer of Q7-Q10 can show the result. Q7-Q10 are based on Likert Scale, which describe the answer in five-level: 1. Strongly disagree 2. Disagree 3. Neither agree nor disagree 4. Agree 5. Strongly agree. Therefore, the average score of each question can describe the agreement of each question. And the higher the score is, more positive the result will be. The average score result of Q7-Q10 is shown in Fig 5.2.

We can see that the score of all question is above average level, which shows positive result in our system's usability. Moreover, we can see that the usability of hardware system gets lowest score. This reminds us we need to improve our hardware devices to make it easier to be carried out.

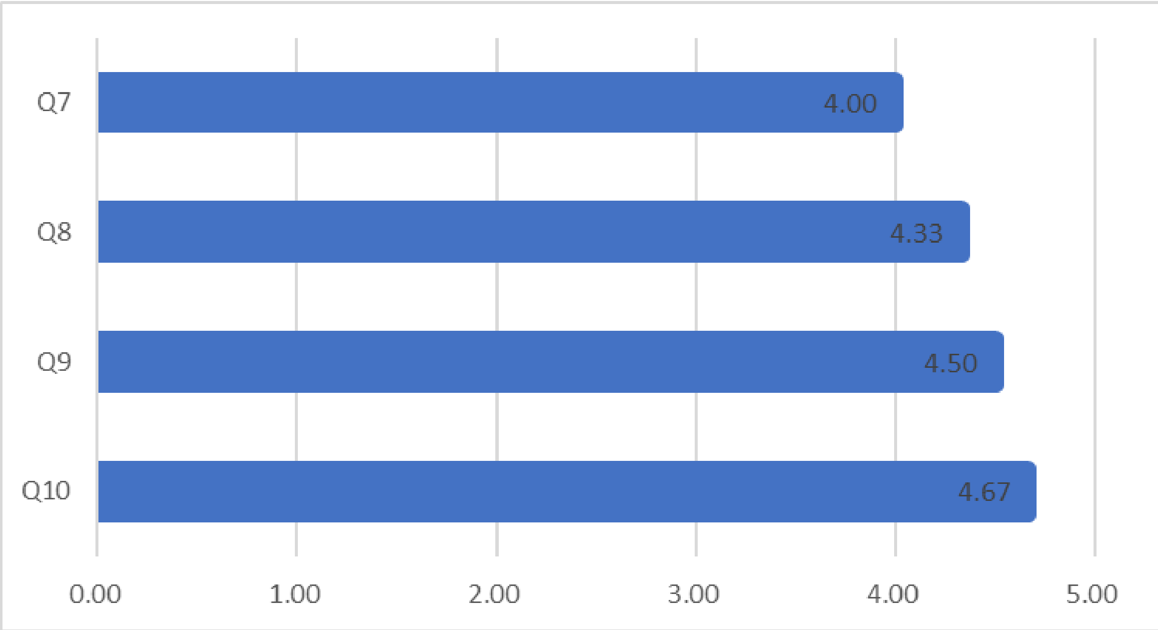


Fig. 5.2 Analyze Result of Q7-Q10

Chapter 6

Related Work

6.1 Related Works on Sentiment Detection

Emotion is a general term for a series of subjective cognitive experience. It is a psychological and physiological state produced by various feelings, thoughts and behaviors. The most common emotions are happiness, anger, sadness, surprise, fear, love, etc. Whether positive or negative, emotions can motivate people to act and influence individual's future somehow. Ever since human beings are aware of the importance of emotions to ourselves, with the increase of sensor accuracy and the rise of computer vision, we are becoming more and more concerned with the recognition and interpretation of emotions. Nowadays, there are mainly three methods to realize emotion detection: (1) Speech emotion detection [17, 18]. (2) Physiological signal detection [19]. (3) Facial emotion detection [20, 21].

In our proposed methods, we choose face photo to detect emotion changes among three methods which is mentioned above. To finally decide which method to choose, we have investigated the merits and demerits of three methods:

- Speech emotion detection: Detection accuracy is the highest one. However, it only detects emotional changes when people speak, making detection discontinuity.

- Physiological signal detection. Subjects need to wear all kinds of heavy and tedious sensors, which make daily life inconvenience and not natural enough.

- Facial emotion detection. The accuracy has become higher due to deep learning staff. Subjects only need to wear a hat hanged with a camera. Detection can be persistent [22].

6.2 Related Works on Movement Detection

Movement detection, or activity detection aims to recognize the actions of human from the observation on human's actions. It has captured the attention of computer science communities ever since 1980s. Up to now, mobile devices are becoming increasingly sophisticated, and the latest mobile phones usually embedded with all kinds of sensors. Among them, accelerometer is widely-used for movement detection [23]. Accelerometer measures the combination of gravitational acceleration and object's motion acceleration in the direction of x, y, z axis. Gravitational acceleration is always vertical to the earth, it can measure the angel change between objects and ground. Motion acceleration can detect the speed changing of the object [24].

6.3 General Related Work

Our work is closely related to the system developed by Harvey J A [25]. In their proposed system, they innovatively arose the idea of involving sitting behavior into the lifelog system. They use a sensor, named activePal, to catch subject's sitting event. Meanwhile, they used Vicon Revue, one kind of wearable camera, to capture lifelog image automatically. In their research, they extracted lifelog images when user is sitting for observation purpose. They useed their system to observe the living habit of elder adults. The idea of involving user's own status into the lifelog system and making behavior as a new cue inspired the birth of our proposed system.

Chapter 7

Conclusion

7.1 Summary

In general, our research explains the limitation of memory cue in lifelog system for aiding episodic memory. In order to improve the situation and enhance the efficiency of aiding episodic memory, we propose two user-related cues in this research and has implemented them into a web-based lifelog viewer.

In the proposed system, we mainly implement special sentiment cue and special movement cue as two new user-related cues for lifeloggers to retrieve memory. We use two Autographers and an Android smartphone to capture sentiment data, movement data and lifelog images. After processing the obtained data, we generate a web viewer for lifeloggers to use sentiment cue and movement cue to view lifelog images and retrieve memory. Moreover, we combine our user-related cues with traditional contextual cues (what, where, who, when) in "combine" part of our system.

We assume that lifeloggers can wear our proposed hardware devices in their daily life. After recording for the whole day, lifeloggers can upload their data onto our system and generate their own web viewer. They can retrieve their memory efficiently and conveniently by using our proposed system.

In order to test the usability and efficiency of our proposed system, we have included some participates in our evaluation. The feedback is quite positive.

7.2 Future Work

Although we have proposed a prototype of utilizing two new cues in lifelog system, there are still some limitations and future possibilities to improve the efficiency of aiding episodic memory. Because of the large scale of images and the restriction of APIs, processing speed in our system is a limitation. We hope this can be solved by optimizing image processing algorithm in the near future. Also, we are thinking of the possibility of involving more user-related cues. This can be done by monitoring other physiological signals like skin conductance.

References

[1] Lauralee Sherwood. Human physiology: From cells to systems. *Cengage Learning*, pages 157–162, 2015.

[2] Larry R. Squire and Stuart M. Zola. Episodic memory, semantic memory, and amnesia. *Psychology*, 8(3):205–211, 1998.

[3] Endel Tulving. Episodic memory: from mind to brain. *Annual Review of Psychology*, 53(1):1–25, 2002.

[4] Klein Stanley B, Judith Loftus, and Kihlstrom John F. Memory and temporal experience: the effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition*, 20(5):353–379, 2002.

[5] Roberto Cabeza, Shitij Kapur, Craik Fergus I, M, and McIntosh Anthony R. Functional neuroanatomy of recall and recognition: A pet study of episodic memory. *Journal of Cognitive Neuroscience*, 9(2):254–265, 1997.

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

[7] Autographer, March 2017. https://en.wikipedia.org/wiki/Autographer.

[8] Matthew L Lee and Anind K Dey. Lifelogging memory appliance for people with episodic memory impairment. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 44–53. ACM, 2008.

[9] Kiyoharu Aizawa, Datchakorn Tancharoen, Shinya Kawasaki, and Toshihiko 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.

[10] Yi Chen and Gareth JF Jones. Augmenting human memory using personal lifelogs. In *Proceedings of the 1st augmented human international conference*, page 24. ACM, 2010.

[11] Matthew L Lee and Anind K Dey. Providing good memory cues for people with episodic memory impairment. In *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, pages 131–138. ACM, 2007.

[12] Matthew L Lee and Anind K Dey. Using lifelogging to support recollection for people with episodic memory impairment and their caregivers. In *Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments*, page 14. ACM, 2008.

[13] Soumyadeb Chowdhury, Md Sadek Ferdous, and Joemon M Jose. A user-study examining visualization of lifelogs. In *Content-Based Multimedia Indexing (CBMI), 2016 14th International Workshop on*, pages 1–6. IEEE, 2016.

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

[15] Liang Jie. Memory recall in lifelog system focusing on memory cues for retrieving events. In *Master Thesis*. Graduate School of Information, Production and Systems, Waseda University, 2017.

[16] Katja C Thoring, Roland M Mueller, and Petra Badke-Schaub. Ethnographic design research with wearable cameras. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2049–2054. ACM, 2015.

[17] Feng Yu, Eric Chang, Ying-Qing Xu, and Heung-Yeung Shum. Emotion detection from speech to enrich multimedia content. In *Pacific-Rim Conference on Multimedia*, pages 550–557. Springer, 2001.

[18] Cynthia Breazeal and Lijin Aryananda. Recognition of affective communicative intent in robot-directed speech. *Autonomous robots*, 12(1):83–104, 2002.

[19] Foteini Agrafioti, Dimitris Hatzinakos, and Adam K Anderson. Ecg pattern analysis for emotion detection. *IEEE Transactions on Affective Computing*, 3(1):102–115, 2012.

[20] Paula M Niedenthal, Jamin B Halberstadt, Jonathan Margolin, and AAse H Innes-Ker. Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology*, 30(2):211–222, 2000.

[21] George Caridakis, Lori Malatesta, Loic Kessous, Noam Amir, Amaryllis Raouzaiou, and Kostas Karpouzis. Modeling naturalistic affective states via facial and vocal expressions recognition. In *Proceedings of the 8th international conference on Multimodal interfaces*, pages 146–154. ACM, 2006.

[22] Arman Savran, Koray Ciftci, Guillaume Chanel, Javier Mota, Luong Hong Viet, Blent Sankur, Lale Akarun, Alice Caplier, and Michele Rombaut. Emotion detection in the loop from brain signals and facial images. 2006.

[23] Jonghun Baek, Geehyuk Lee, Wonbae Park, and Byoung-Ju Yun. Accelerometer signal processing for user activity detection. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pages 610–617. Springer, 2004.

[24] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. Activity recognition from accelerometer data. In *Aaai*, volume 5, pages 1541–1546, 2005.

[25] Juliet A Harvey, Dawn A Skelton, and Sebastien FM Chastin. Acceptability of novel lifelogging technology to determine context of sedentary behaviour in older adults. *AIMS Public Health*, 3(1):158–71, 2016.

Appendix A

Evaluation Questionnaire

Questionnaire

Ladies/gentleman,

Thanks for helping us to complete this questionnaire. All the information will be only used in this research. Your information will be greatly valued and of course will be strictly confidential. Thanks again for your cooperation!

About you

1. How old are you? I am __23__ years old.

2. Male or female?

○ Male
● Female

Record

Question	Answer
3. How many events you can remember before using our system?	6
4. How many events you can recall by using special sentiment cue?	8
5. How many events you can recall by using special movement cue?	10
6. How many events you can recall by using combination cue?	7

Questions (Each question is graded from 1 to 5.)

7. Do you think it is comfortable to wear our devices?

totally uncomfortable ○—○—○—●—○ very comfortable

8. Do you think our viewer is easy to use?

totally hard ○—○—○—●—○ very easy

9. Do you feel extracted images useful?

totally useless ○—○—○—○—● very useful

10. Do you think our system help in aiding memory intuitively?

totally helpless ○—○—○—○—● very helpful

Fig. A.1 Questionnaire