Memento: An Emotion Driven Lifelogging System with Wearables

Shiqi Jiang[1], Pengfei Zhou[1], Zhenjiang Li[2], Mo Li[1]

[1] Nanyang Technological University, [2] City University of Hong Kong
[1]{sjiang004, pfzhou, limo}@ntu.edu.sg, [2]zhenjiang.li@cityu.edu.hk

***Abstract*—Due to the increasing popularity of mobile devices, the usage of lifelogging has been dramatically expanded. People collect their daily memorial moments and share with friends on the social network, which has been an emerging lifestyle. We see great potential of lifelogging applications along with rapid growth of recent wearable market, where more sensors are introduced to wearables, i.e., electroencephalogram (EEG) sensors, that can further sense the user's mental activities, e.g., emotions. In this paper, we present the design and implementation of Memento, an emotion driven lifelogging system on wearables. Memento integrates EEG sensors with smart glasses. Since memorable moments usually coincides with the user's emotional changes, Memento leverages the knowledge from the brain-computer-interface (BCI) domain to analyze the EEG signals to infer emotions and automatically launch lifelogging based on that. Towards building Memento on COTS wearable devices, we study EEG signals in mobility cases and propose a multiple sensor fusion based approach to estimate signal quality. We also present a customized two-phase emotion recognition architecture, considering both the affordability and efficiency of wearable-class devices. Our experimental evaluation shows that Memento is responsive, efficient and user-friendly on wearables.**

I. Introduction

Lifelogging digitizes human daily lives, which was widely adopted in the therapy for a series of neurodegenerative diseases using dedicated devices [13], [14], [23]. Due to the increasing number of smartphones, the use of lifelogging has been dramatically expanded. It has already became an emerging lifestyle for people to collect their memorial moments and share with friends. Lifelogging applications or services [3], [4] on smartphones are able to log users' lives in various forms such as texts, images, audio clips, videos, etc.

Recent wearable market has been rapidly growing in terms of both technology advances and penetration. Wearable devices, especially smart glasses are equipped with the first-person camera, microphone, and rich on-board sensors. They are always carried by the users and exposed to the ambient environment. These characteristics make us believe that wearables have the potential to provide better lifelogging services. It is, however non-trivial to build a lifelogging system on wearables. Rather than being directly migrated from the existing systems, lifelogging on wearable prompts new questions. In this paper, we make attempts to examine the questions we think important when designing an user-friendly, affordable, and efficient lifelogging system on wearables.

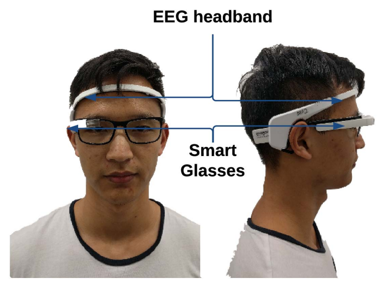


Fig. 1. Emotion-driven lifelogging system.

1) *The burden of human intervention*. Most existing lifelogging systems on mobile devices need manual operations [3], [4], which require excessive human interventions. Though various user interaction methods are introduced on wearables, including the gesture, voice and even wink controls, the intervention overhead causes non-negligible burdens and may impair users' willingness to conduct lifelogging. Due to the lag of manual operations, many brief but valuable moments could be easily missed as well. Therefore, an automatic lifelogging system without requiring user intervention is expected.
2) *The affordability of lifelogging on wearables*. Most passive lifelogging systems rely on dedicated devices [6], [7], which is, however, not appropriate for most COTS wearables devices. On the one hand, due to the limited size, the energy resources of wearables are bounded, which cannot afford continuous sensing. On the other hand, passive lifelogging usually leads to massive irrelevant lifelogs. Therefore the affordability of lifelogging systems on wearables should be carefully considered.
3) *Inside context awareness*. Recent years, the commercial Electroencephalography (EEG) [2] headband comes into the market. EEG is a measure of brain waves, which could be a reflection of the environment and generally results in physical and psychological changes of people, influencing behaviors. EEG offers the opportunities for us to understand the user mental states. In this paper, we attempt to introduce EEG sensors in order to dig more information from the inside world of users and improve the lifelogging accuracy based on that.

We investigate the questions above through the design and

978-1-5090-2991-4/17/$31.00 ©2017 IEEE

implementation of an emotion driven lifelogging system on wearables, Memento. The goal of Memento is to provide an automatic seamless lifelogging experience on COTS wearable devices, while considering the affordability, user-friendly and efficiency in practice. Shown in Fig. 1, Memento integrates EEG electrodes with smart glasses. By leveraging the Brain-Computer-Interface (BCI) domain knowledge, Memento derives emotions from EEG signals and launches the lifelogging process automatically based on that. The implementation of such a system, however, entails two crucial challenges:

First, signals collected from wearable EEG sensors are not always reliable — EEG signals are usually mixed with various external noises, from the electric appliances and the interfering signals from the muscle activities such as the blink, jaw or heart beat. On the other hand, the electrodes of wearable EEG are not fixed on the scalp (see Fig. 1 as a reference). The movements of the user could thus cause a drift of the electrodes, which will in turn lead to a unpredicted change of the harvested EEG signals. In this paper we carefully study the frequency features of major noises and apply an efficient filter to remove the signals due to non-brain activities. Meanwhile we propose a sensor fusion based approach with the help of IMU readings to detect electrodes drifting and assist estimating signal quality.

Second, the high computational complexity of existing emotion recognition algorithms prohibits them to be adopted directly on wearable devices. To address such an issue, we split the emotion recognition into two phases and install them on smart glasses and private clouds (personal PC or smartphone). Instead of extracting emotions in real time, we trigger lifelogging by analyzing the emotional changes. The exact emotion information are obtained offline and tagged on the lifelogs.

In summary, our contributions of this work include:

- The proposal of a new and natural way to automatically trigger lifelogging. We introduce the use of EEG to the lifelogging system on wearables.
- A series of techniques to integrate EEG electrodes with wearables. We present our signal processing and two phase emotion recognition design, in order to meet the computation and energy constraints of wearable-class hardware.
- The full design and implementation of Memento, an emotion driven lifelogging system on smart glasses. We show the capacities of Memento are efficient in support of emotion tagging and lifelogging workload. The experimental evaluation indicates that Memento is able to provide satisfactory battery life and user experience.

II. Lifelogging and Emotions

Lifelogging is generally considered to be the digitalization of personal experience for personal use. In the past decades, several dedicated passive lifelogging devices are proposed for various purposes, e.g. augmenting lives [28], memory aid [14] and health care [23]. Recent years, wearables gain their popularities and show the potential for lifelogging. The question therefore arises that how to architect a better lifelogging system on wearables.

We believe that lifelogging could be improved by capturing the internal experience of users. Rather than continuously logging or requiring manual operations, the system is able to automatically log memorial moments by understanding the user's mental state changes. For example, when the user gets very happy or sad, the happiness or sadness could be recorded automatically. Some unexpected situations or accidents could be logged as well when the user suffers strong emotional disturbances. We also believe that the lifelogs with emotion information could not only improve the lifelogging system itself but also enhance many other existing services such SNS, health care, memory aid, etc.

Research on emotions is not new. Generally, the bodily changes follow directly the perception of the exciting fact, and that our feeling of the changes as they occurs is the emotion [17]. The consensus is that emotions are physiological and measurable. Some basic emotions can be found universally across individuals when certain stimulations are given. Ekman's research findings [12] led to classify six emotions as basic: anger, disgust, fear, happiness, sadness and surprise. It is then improved to the bipolar model where arousal dimension (*how energized the experience feels*) and valence dimension (*how negative or positive the experience feels*) are considered.

The emotion changes can be observed via several physiological effects such as heart beats, facial expressions, voices and brain activities. EEG is a measure of brain activities via the brain electric signals. In order to harvest signals, EEG electrodes are attached on the scalps. The emotion changes lead to distinguished patterns in EEG signals from the certain positions. There is an increasing number of EEG based emotion recognition techniques [15], [26], [27].However, most of EEG based emotion recognition algorithms are not designed for the wearable-class platform. To adapt them, the problem we are faced with is the algorithm complexity. The influence of the high complexity is twofold: time constrains and energy overhead. On the one hand, most of the algorithms are proposed for the off-line emotion recognition, even on more powerful platforms. On the other hand, the energy issues are rarely considered in the previous designs.

III. Memento System Design

To enable lifelogging with wearables, we present our design of an emotion driven lifelogging system, namely Memento. Fig. 2 illustrates the system architecture. Memento contains three modules. The signal processing module takes EEG readings as input to detect users' emotional changes, e.g., emotion events. When an emotion event is detected, the lifelog collector starts to conduct the lifelogging. The camera on smart glasses serves as the major media to capture the daily life moments in the form of video. Finally the recognized emotions are tagged on the respective lifelogs.

The EEG electrodes loosely contact on user's scalp and the sensed EEG signals can be easily polluted by interfering

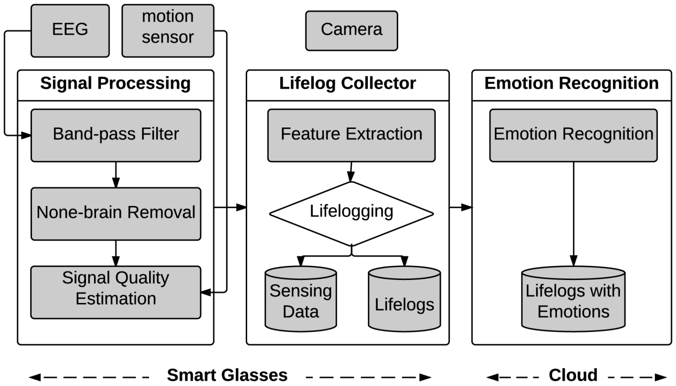


Fig. 2. System architecture of Memento

activities, e.g., eye blink, head movement, etc. To address this issue, in the signal processing module we propose an effective band-pass filter to remove ambient noises, utilize the kernel-based correlation to exclude non-brain activities, and further leverage motion sensors and the rigid feature of the wearable to score the quality of the perceived EEG signals (§III-A). With all high-score EEG signals, the lifelog collector module performs lifelogging according to user's emotions. However, recognizing emotions from EEG signals is computational intensive, not afforded by the wearable platform directly. To address this issue, we propose a two-phase emotion recognition solution and install them on the smart glasses and the cloud, respectively. On the smart glasses side, without specifying the exact emotion type, Memento merely detects emotion changes by leveraging the intermediate results from the feature extraction phase. Once a significant change is detected, the lifelogging is launched. The collected lifelogs as well as sensed data would be uploaded to the cloud. On the cloud a sophisticated recognition algorithm runs to recognize the exact emotion type, which serves as the emotion tag for the subsequent recorded lifelogs (§III-B).

A. Signal preprocessing

In this section, we describe the EEG signal preprocessing techniques applied in Memento. Preprocessing has been discussed in the previous BCI work. However, instead of collected in the controlled environment the signals of Memento are harvested in the uncontrolled and mobility cases, where the EEG electrodes are not fixed on the scalp and more internal and external noises are introduced. Specifically we process the signals as the following steps:

Band pass filter. EEG frequency normally ranges from 1 to 100Hz, which is further divided into several bands: Alpha (8-14Hz), Beta (12-30Hz), Gamma (>30Hz) and Delta (<4Hz), etc.. On the one hand, as shown in Fig. 3 on the whole EEG frequency band, several interferences are usually involved. The noise from electric appliances locates at 50Hz. The interference from the activities of the heart (ECG) generally overlap at the low frequency band (1-6 Hz). The muscle activities such as blinks also introduce the noise to EEG signals. On the other hand, Alpha (8-14 Hz) and Beta (12-30 Hz) contain the significant features to represent metal states, which are widely used in the motion recognition. Therefore we apply a 7-31 Hz band pass filter. Some research points out where the EEG frequency bands begin/end varies depending on the input sources and the EEG frequency boundary usually varies a few Herz, we set a fixed band pass filter in our design. We do not focus the specific signal characteristics on each band. The goal of the band pass filter is to get the most significant partial of signals, while filtering out the external and internal noises.

Non-brain activity removal. Non-brain activities including chews, blinks and other muscle activities, produce electric signals (EMG) and might be harvested by EEG electrodes. Frequent non-brain activities involve lots of interference to EEG signals. In most EEG study electrodes are usually attached on TP9, TP10, AF7 and AF8 points according to 10-20 system [1]. On this area, the most significant non-brain activities are blinks. In order to detect then remove the blink effects, we examine the following methods: we first notice that EEG amplitude changes on AF7 and AF8 channels when blinking, as shown in Fig. 4. We decompose signals into frames with the fixed length and apply a threshold based method to detect if a frame contains blinks. However due to the tail effect of the blinks and other brain activities, high false positive rate is obtained. The frequency characteristic might be distinguished between signals from the brain activities and blinks, but continuous time-frequency transformation could involve a so heavy workload that we do not consider it. In our design, we apply a kernel correlation based approach on time domain. We empirically set a frame with the size of 400 ms which equals to a typical blink duration. We train the kernel of blinks off-line. Fig. 4 illustrates the results when we apply the trained kernel on EEG signals directly. In each frame, correlation between the signals and the kernel is computed to detect if the blink effects are contained. After that, for every frames containing blink effects we apply a modified algorithm from [24] to remove the effect of blinks.

Signal quality estimation. To further understand the signal quality, we apply the signal quality estimation. We evaluate the signal quality in three aspects. First we make use of the reference electrode to estimate the contact status. The reference electrode emits the constant current. The energy harvested on the reference electrode indicates the contact status, if the contact status is good the power loss should be limited, otherwise only less power could be harvested on the reference electrode. We apply a binary filter to determine the contact status and use it as the trigger of the whole processing procedure. Second, we use the power correlation across electrodes to infer the electrode drift. When EEG electrodes drift we have the observation that since the wearable is a rigid-body, all the electrodes should be influenced at the same time and the drift impacts the signal power harvested in a short period. Fig. 5 illustrates the signal power on AF7 and AF8. When the drift occurs the correlation of the signal power on these two channels is observed. Third, we notice that the signal quality is vey likely to get worse when the user is in the

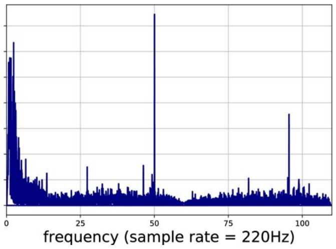


Fig. 3. Frequency distribution of raw EEG signals.

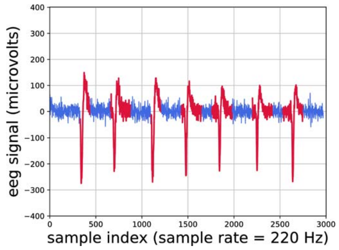


Fig. 4. Detecting blinks from raw EEG signals using kernel correlation

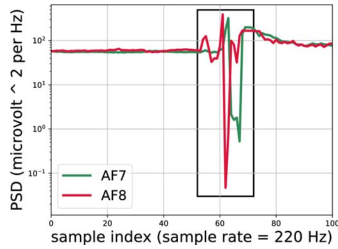


Fig. 5. Band power correlation across EEG electrodes when drifting.

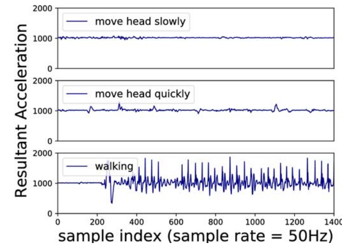


Fig. 6. Resultant accelerometer readings under different usage cases.

vigorous movements. Therefore the motion sensor is utilized. Fig. 6 shows the resultant acceleration in different movement states. We use the variance of the resultant acceleration to infer the movements and use it as another indicator of the signal quality. Finally we linearly combine the two indicators, the power correlation and the acceleration variance with the even weights as the signal quality score.

B. Two phase emotion recognition

Many emotion recognition algorithms are too heavy to adapt them on wearables directly, especially considering the near real-time processing constrain and the computation capacities. To best our knowledge, most of algorithms are designed for recognizing emotions off-line. Emotion recognition algorithms are based on different theories, but the procedures of these algorithms share the similar patterns: they can be decomposed into the feature extraction phase and the classification phase. We observe that the overall execution time is too heavy for wearable-class devices though, the classification phase is the major contributor and usually the feature extraction phase is light enough. It is because that most of feature extraction approaches leverage near linear algorithms, while tons of iterations are involved in the classification phase. For instance, in [27], the lifting based wavelet transformation is used to extract features, which is acceptable yet, but in the classification phase, the Fuzzy C-Means clustering is applied, introducing a vast number of computation. Based on this observation, we make attempt to separate the emotion recognition algorithm into the feature extraction phase and the classification phase, then install them on the wearable and the cloud, respectively.

EEG segmentation. Before introducing the feature extraction phase on Memento, we first describe our data segmentation strategy. Algorithm 1 shows the pseudo-code of out segment strategy. The design goal is to balance the data redundancy and the processing efficiency. We apply a sliding widow whose size is 10 seconds containing 2200 samples. We set 50% overlap initially. Leveraging the results of signal quality estimation (§III-A), we determine to keep the current segment or drop it. We use an adaptive threshold instead of the fixed threshold. A fixed threshold which can perform well in most possible circumstances is hardly determined. For instance, an absolute high threshold might lead to the data fragmentation while a low threshold might introduce more noisy segments. In this paper we average the signal quality score of all validated segments, then a certain tolerance score is subtracted to ensure the diversity (line 20-26). We use it as the adaptive threshold. The invalidated segments are dropped. If the consecutive segments are invalidated, it indicates the current situation gets worse. We reduce the overlap 10% for each consecutive dropped segment. We always reset the overlap size to 50% when a segment is accepted (line 11-18). Finally we store the validated segments in the circular buffer. The buffer follows the FIFO rule (line 4-6). If there is no space for the coming segment, the existing item would be replaced.

Feature extraction. Our system fetches the EEG segments from the buffer continuously and then extracts features. Since the exact emotions would not be recognized at this phase, we consider to use the intermediate results from features to trigger lifelogging. Some features that only describe the statistic characteristics are inappropriate. In Memento we use the fractal dimension (FD) based approach. FD is an index for characterizing fractal sets by quantifying their complexity. FD calculation has been used in the EEG analysis [26] and is considered to positively correlate with how energized the user is, hence the arousal level. Particularly we make use of Katz's FD calculation [18] and apply it directly on the waveforms. We obtain the FD of a curve using

$$D = \frac{\log n}{\log \dfrac{d}{L} + \log n} \tag{1}$$

where L is the total length of the curve and d is the diameter that can be estimated as the farthest distance with the beginning point. n is the number of the steps in the curve, which is adjusted by the granularity. Particularly we define $n = L/a$ where a is the average distance among the successive points.

Lager FD values are associated with higher arousal levels. To tolerate the absolute arousal differences among individuals, we do not use the absolute arousal level as the threshold to trigger the lifelogging procedure. We propose to use the trend of emotional changes, which is calculate as the derivation of the FD values. The FD trends represents the mental change direction, positive or negative and the intensity as well. We further use the sliding average to smooth the trend curve. Fig. 7 illustrates when the stimulation is applied the positive trend occurs, and the lifelogging process is triggered.

Algorithm 1: Segmentation Strategy

segmentation($data$);
Input : The EEG data after preprocessing $data$
Output: Circular buffer $Buffer$
$currentSegment \leftarrow$ capture data from $startLoc$ with $windowSize$;
$isValid \leftarrow valid(currentSegment)$;
if *isValid* **then**
  index = valid segment nums / bufferSize;
  Buffer$[index] \leftarrow$ currentSegment;
end
$startLoc \leftarrow advance(startLoc, isValid)$;
return *Buffer*;

advance($startLoc, isValid$);
Input : Last segment start index $startLoc$ and whether it is valid $isValid$
Output: The start index for the next segment
if $isValid$ **then**
  $consecutiveDrop \leftarrow 0$;
else
  $consecutiveDrop \leftarrow consecutiveDrop + 1$;
end
$startLoc = startLoc + windowSize * (1 - max(0, overlap - consecutiveDrop * 0.1))$;
return *startLoc*;

valid($segment$);
Input : The current segment $segment$
Output: If the current segment is accepted $isValid$
$qualityScore \leftarrow$ the signal quality score;
if $qualityScore >= threshold - tolerance$ **then**
  $threshold \leftarrow$ average score of all valid segment;
  return $True$;
else
  return $False$;
end

Two phase emotion classification. Finally if Memento determine to launch the lifelogging. A 120 seconds video is captured. Currently we set a fixed capturing duration. We find that adjusting the settings of lifelogging such as the lifelogging methods (video, audio, pictures), lifelogging configuration (FPS, resolution) and lifelogging duration could lead to different lifelogging qualities and energy consumptions. How to best balance them is one of our future work (§VII). The lifelogs and the features of EEG are compressed and saved in the storage of the smart glasses, as well as the related sensor readings including motion sensor readings, GPS and so on. We notice that although the computation and energy resources are so limited on wearables, the storage is satisfied with the lifelogs and sensing data. The data is uploaded to the private cloud (PC or smartphone) later when the smart glasses are in charged or upon the user requests. On the cloud, we run the sophisticated emotion recognition algorithm to recognize the emotions. Particularly we adapt the emotion recognition algorithms from [26].

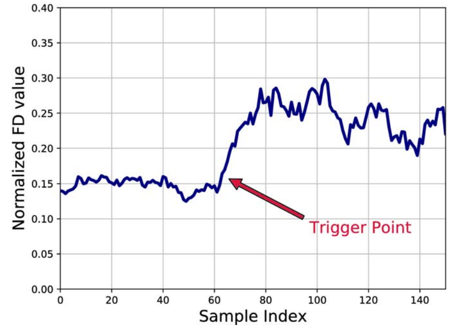
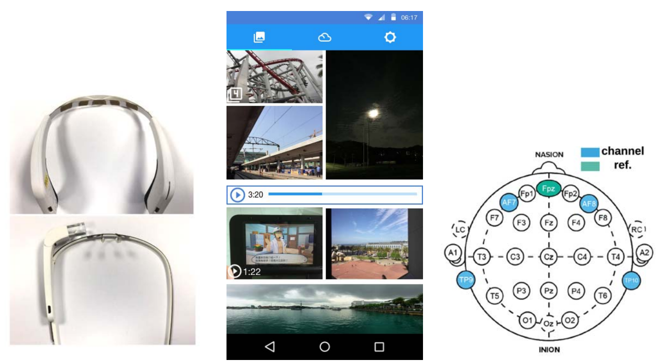


Fig. 7. Using the arousal trends as the trigger to perform lifelogging

IV. Implementation

We close the description of Memento by introducing the specific implementations of current prototype.



(a) hardwares used in the prototype. (b) UI of Memento on smartphones. (c) electrode locations in 10-20 system.

Fig. 8. The hardware and software implementation of Memento prototype.

Hardware. As illustrated in Fig. 8(a), the current prototype is comprised of two commercial devices: a Muse EEG headband [5] and a Google Glass Explorer Edition. Muse headband is designed as a personal meditation assistant, which helps user for meditation exercises. Four channels are supported on the Muse headband and they are illustrated by 10-20 system [1] in Fig. 8(c). Besides two pairs of channel electrodes and one reference electrode, the Muse headband is also equipped with an accelerometer. A PIC24 MCU is used to conduct simple processing operations e.g. low-pass filtering and band filtering. The headband has a separated battery with the size of 250mAh. Only Bluetooth can be used for the data communication on the Muse headband. We use a Google Glass as the smart glasses platform, since it is one of the most completed smart glasses we can find on the market. Android OS on the Google Glass also provides the openness and easy development to developers. The core of the Google Glass is the OMAP4430 system-on-chip (Soc), which owns a dual-core ARM Cortex-A9 CPU, 2BG RAM and 16GB flash storage. Its rich sensors satisfy the requirements of lifelogging and sensing, including microphone, camera, accelerometer, gyroscope, ambient light sensor and so on. The power supply comes from a 570mAh liPo battery. Both Wi-Fi and Bluetooth are supported. In our prototype, raw EEG reading are preprocessed by the band-pass filter on the Muse headband. Since the headband is not programmable, the rest of the preprocessing operations are done on the Google Glass. EEG data is pushed from the headband to the glasses via Bluetooth. We demonstrate the smartphone services on a LG Nexus 5 and a Samsung Galaxy. The back-end service runs on our experiment workstation with a Intel Xeon processor, 16GB RAM and 1T disk.

Software. We implement the Memento functionalities in three software packages and deploy them on the smart glasses, the smartphone and the back-end server, respectively. 1) We implement the preprocessing and lifelogging functins as an Android Wear application. Most of the code is written in Java. For the common utility algorithms used in the data cleaning and feature extracting e.g. Goertzel algorithm, Katz's FD calculation, etc., we implement them as the executed libraries in native code. 2) Memento service on the smartphone. We implement the application with GUI to browse the lifelogs in the private cloud and the Google Glass. The application is also used to share the lifelogs to the SNS easily. Fig. 8(b) shows the current user interface on Android. 3) Back-end server on the private cloud. We implement the full version of the emotion recognition on the back-end server in Java. We setup a web service on the server. Lifelogs along with related sensing data uploaded through the encrypted links. The back-end sever also provides the features of querying and browsing lifelogs.

Lifelog Store. Lifelogs are distributed both on wearables and the private cloud. On the Google Glass, media lifelogs are stored in the raw format. Before storing the EEG features, we compress them to reduce the size. We use SqLite database to maintain an index. An entry in the index represents a lifelogging event, which contains the corresponding media lifelogs and EEG data. Currently we have four attributions in each entry, which are the timestamp, the location obtained from GPS on the smartphone, the file pointers of the media lifelogs and the EEG features. The lifelogs on the Google Glass are uploaded to the private cloud when the glasses are in charge or upon the user's requests. On the private cloud, lifelogs are stored in the similar structure with on the glasses, but the emotion tags are added. Each entry is assigned one type of basic emotions, along with its arousal and valence levels.

V. Evaluation

In this section, we present our evaluation on Memento. We first introduce the dataset and the experiment settings. Then we evaluate Memento in terms of the system performance and the energy consumption.

A. Dataset

DEAP as the ground truth. We are given the access to DEAP [19], which is an open dataset for the analysis of human affective states. In DEAP EEG signals of 32 participants are recorded by the dedicated equipments with 32 electrode channels, as each watches 40 one-minute videos. The self-assessment tool [8] is used to let participants rate each video in terms of the level (1 to 9) of arousal, valence, dominance and so on. In total the size of DEAP is about 2.7GB and the video used can be assessed via Internet. The collected signals of DEAP are from the dedicated equipment and the results of the emotion assessments are obtained by the widely used self-assessment tool. Therefore we make use of DEAP as the ground truth.

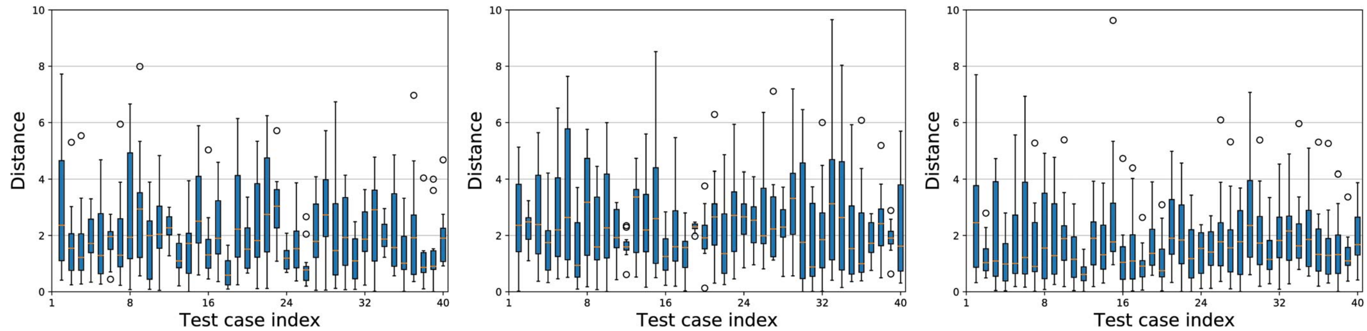
EEG data collected in the lab. Using one-minute videos in DEAP mentioned above as the stimulations, we collect the EEG readings by Memento. The raw EEG readings are recorded. This experiment is conducted in the lab, the participant is not allowed to move around. In total 40 pieces of raw EEG signals are recorded corresponding to each one-minute video in DEAP.

Lifelogs, EEG and sensor data collected by customized Memento. We collect the EEG readings, motion sensor readings and lifelogs by Memento under different scenarios, including the laboratory, the office, the street and the park. We build a customized Memento for the data collection. Instead of automatically launching the lifelogging process, the customized Memento continuously records lifelogs in the form of video. When the significant emotion changes occur and Memento is triggered to performance lifelogging, the event is marked. We add an external battery to supply power to the Google Glass in order to make it continuously capture videos at the best quality. Totally we get about 6 hours videos and sensor data traces.

Energy consumption trace. We use Monsoon power monitor to collect the real time energy consumption traces of the Google Glass and the smartphones.

B. Emotion Recognition Accuracy

To evaluate the emotion recognition accuracy, we use Manhattan distance on the Arousal-Valence model to measure the differences of emotion samples. First we run our full emotion recognition algorithm on DEAP EEG data and compare the results with the self-assessment value in DEAP. For each one-minute video, we have a sequence of EEG segments. Every segment is processed to extract the recognized emotion, which is represented by a pair of values on the Arousal-Valence model. We compare every point with the corresponding self-assessment value and calculate the distances. Fig. 9(a) illustrates the distance distribution of all 40 test cases. For each test case, we calculate Root Mean Square (RMS) between recognized emotions and self-assessment value. Overall the average RMS is 2.42 and the maximal and minimal RMS is 4.25 and 0.90, respectively. Among them, 60% of cases is less than 2.5 and the RMS of over 80% test cases are less than 3.29. In the similar way, we run the emotion recognition



(a) emotion recognition on DEAP data V.S. self-assessment in DEAP
(b) emotion recognition on the data collected by Memento V.S. the self-assessment in DEAP
(c) emotion recognition on the data of DEAP V.S. Memento

Fig. 9. Emotion recognition on the data of DEAP and collected by Memento.

algorithm on the data collected by Memento. We show the details in Fig. 9(b). Overall the average RMS is 2.76 and the maximal and minimal RMS is 4.22 and 1.5, respectively. In all the test cases, the RMS of more than 80% cases is less than 3.66. We further compare the recognition results between on the DEAP data and on the data collected by Memento. We align the recognized values in the time sequence and calculate the distance between each two values who are closed in time. Illustrated in 9(c), the average RMS 2.12 is obtain and RMS from more than 80% cases is less than 2.7.

Second we evaluate the performance of our preprocessing and arousal extracting algorithms. We use the performance obtained in the previous experiment as the baseline and compare against three modified processes that exclude either the non-brain removal, the signal quality estimation or both of them. In this experiment, we run tests on the data collected on Memento and calculate the average RMS with the self-assessment value on the whole 40 test cases. In Table I we list the results. From the table we observe that 1) the full preprocessing design provides more than 32% better performance comparing to the straightforward design (band-pass filter only). 2) Among the components of the preprocessing algorithm, the non-brain removal brings about 24% improvement. We think it is attributed to the fact that the non-brain noise is universal. 3) About 9% is brought by the signal quality estimation. The reason is that we collect the data in a controlled lab environment. The signal quality trends staying stable.

Process	**Averaged RMS**
Full design	2.76
Band pass filter (BPF)	3.63
BPF + Non-brain removal	2.92
BPF + Signal quality estimation	3.32

TABLE I
BREAKDOWN EMOTION RECOGNITION PERFORMANCE.

C. User Study

Next we evaluate Memento by a user study to answer the question: how the emotion-driven lifelogging approach meet the users' expectations.

Comparing to the previous continuous lifelogging or manual lifelogging, it is necessary to learn if the emotion driven lifelogs meet the users' expectations. We conduct a simple user study. From the continuous lifelogs collected in the experiment, we ask the volunteers to manually select the memorized moments. For each lifelog, the volunteer labels the corresponding start/end time. In order to completely derive the memorized moments, the volunteers are asked to select clips in multiple rounds. In every round, the clips selected in the previous rounds are skipped. We apply three selection rounds and use these user selected clips as the base line.

We compare the lifelogs marked by Memento with the base line. For each lifelogs of Memento, we particularly categorize the comparison result into the following buckets: for the video lifelog, we evaluate the overlap between the clips marked by Memento and the selected video clips. If the overlap rate is more than 50%, we label it as *fit*. If the overlap is ranged from 30%-50%, then we label it as *semi-fit*. We mark the video lifelog as *none-fit* if the overlap is less than 30%.

From the 6 hours lifeog traces, Memento marks 16 pieces of lifelogs in total. Fig. 10 shows the comparison results. We observe that more than 40% are labeled as *fit* and more than 75% are marked as *semi-fit* and upper.

We also evaluate the coverage of the selected videos against the lifelogs collected by Memento. We define the selected video is not covered when the overlap with the Memento lifelogs in any forms is less than 30%, covered otherwise. From the 6 hours lifelogging traces, the volunteers manually select 24 clips in total. Among them more than 70% lifelogs are covered. When we narrow down the test case space into the clips only from the selection first round, Memento achieves about 83% coverage.

D. Energy consumption

Next we evaluate the energy consumption of Memento. The EEG sensing overhead mainly comes from the EEG electrodes and the MCU on the EEG headband. Since the APIs provided by the Muse EEG headband are limited, we are not allowed to test in more comprehensive settings. The power consumption of the headband is stable during our experiments. The average

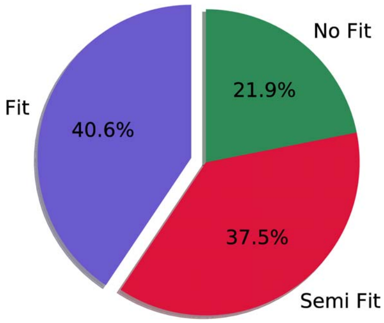


Fig. 10. How the emotion-driven lifelogs fit the users' expectations.

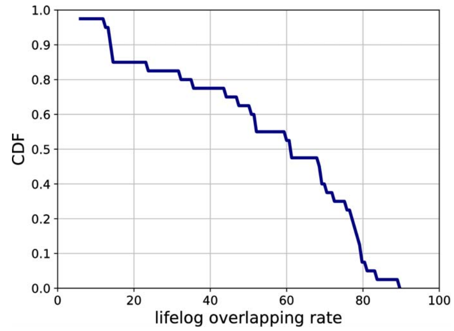


Fig. 11. The coverage ratio of the emotion-driven lifelogs against the clips selected by users.

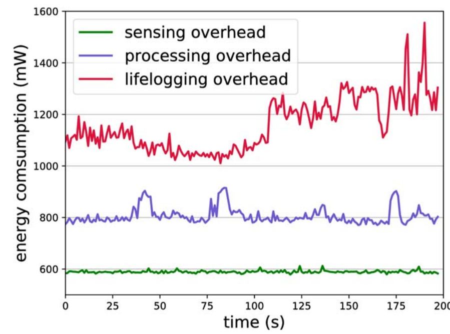


Fig. 12. Energy consumption of each component.

battery life of the Muse EEG headband is about 4.2 hours with the battery size of 250mAh.

On the smart glasses, Monsoon Power monitor is used to collect the real time energy consumption traces under the following circumstances: 1) Memento continuously receives the EEG and motion data. We regard it as the sensing overhead. 2) Memento cleans the data, extracts the FD values and computes the arousal trends. We label it as the processing overhead. 3) When the certain arousal changes are detected, the lifelog collector is triggered to perform lifelogging in the form of video. We regard it as the lifelogging overhead.

Fig. 12 illustrates the real time energy profiles of these three circumstances. We have the following observations: 1) Compared with the sensing overhead (589mW), the additional overhead brought by the processing is less than 200mW. 2) The energy consumption of the lifelogging, especially video capturing is a burden, which is about 1200mW. Currently we cannot avoid it due to the existing architecture of the camera and OS. However, since we apply an emotion driven approach, the lifelogging process is only triggered opportunistically. 3) Receiving data also brings the overhead in our implementation. Basically it is closed to the lowest energy to hold the CPU wake-lock. We think it could be largely reduced or even removed if EEG electrodes could be physically connected with the smart glasses and Low Power Unit (LPU) could be used. We regard it as our future work.

VI. Related Work

Memento touches several different research areas. In this section, we survey the prior work which is the most related to Memento mainly in the following two aspects:

Emotion recognition and its applications. Emotion is one of the human natures and can be inferred in several ways. Existing work exploits the recognition of emotions by leveraging the acoustic [21] or the visual signals [10], where the facial or spoken expressions are carefully analyzed. Comparing to the voice or exoressions, the physiological signals i.e. EEG, are more direct indicators of how we are feeling. There are a number of EEG based emotion recognition algorithms proposed in recent years. For example, to classify basic emotions, support vector machine is used in [15]. In [27], wavelet based methods are applied to extract features, fuzzy k-means and fuzzy c-means clustering are considered to do the classification. In [29], higher order crossing analysis(HOC) is adopted. However, these algorithms mainly focus the emotion recognition accuracy. The time constraints and the energy consumptions are rarely considered. Recent years some sensing techniques are also proposed to infer affective activities. In [22], the relationship between the phone usage pattern and the personalities is discussed. The authors of MoodScope [25] propose their findings that by analyzing the communication history and application usage patterns, users' daily mood can be inferred. Comparing to these systems, Memento leverages EEG directly and interprets the users' affective activities in the perspective of physiological properties.

Lifelogging system. In the past decades, many lifelogging systems are proposed. We categorized them into two groups: dedicated equipments for certain purposes and lifelogging systems on personal mobile devices, smartphones specifically. SenseCam [14] is a body mounted camera, which passively captures photos through a wide-angle lens. Extending it, Footprint Tracker [13] studies the effects of multiple memory cues. Recently wearables gain the popularities. Some wearable cameras [6], [7] can be used to log daily lives. Recent ZOE [20] leverages the advanced system-on-chip(Soc) techniques, continuously senses a number of user activities and provides the dialog based user interaction. However, most of these lifelogging devices are designed to passively log or request users' intervention. Memento aims to provide a automatic, affordable lifelogging solution on COTS wearable-class platforms. Leveraging the rich on-board sensors and sensing algorithms, some context awareness lifelogging approaches [11] on smartphones are proposed. Lifelogger [9] collects a diverse of sensor data and focuses on providing the robust lifelog query services. UbiqLoq [30] proposes a lightweight framework allowing developers easily create lifelogging application based on it. These lifelogging systems are developed to record the external environment. Meanwhile the internal experience of people is important and helpful for lifelogging as well [16]. Memento makes attempts to introduce the user emotions to improve the lifelogging.

VII. Discussion and Future Work

In this section, we briefly discuss the issues related to the current design and implementation of Memento.

Balance between lifelog quality and energy consumption. Memento senses the changes of users' emotions and launches the lifelogging process based on that. The smart glasses can record lifelogs in forms of audios, photos and video clips, etc.. It is obvious that the video clips contains more information than audios and photos do in the ideal cases. However, in practical, on the one hand, each lifelogging method with different configurations (e.g. sampling rate, resolution and FPS) has different energy profiles. On the other hand, the environment impacts the quality of lifelogs. For instance, the light conditions might influence the quality of videos and photos, however the audio clips are not impacted. The noise level might lead to the poor audio records but the photos and videos still holds the information of views. How to best balance the logging quality and the energy consumption is another further work.

LPU utilization. One of the problems Memento tackles is the energy bottlenecks. Heterogeneous computation could significantly reduce the energy consumption. In our implementation, the Muse headband uses a micro control unit to apply the preprocessing operations on the EEG signals such as smoothing and band-pass filtering. On the Google Glass, currently the co-processor or LPU is not utilized, since Android does not open the related API. In this paper, we focus on showing the proof-of-concept of the emotion-driven lifelogging system, and make it affordable on COTS wearables. For the future work, we plan to design the custom hardware to integrate EEG sensors physically, and utilize the LPU to improve the performance.

VIII. Conclusion

In this paper, we present the design and implementation of Memento, an emotion driven lifelogging system on wearables. Memento senses the emotional changes of users and automatically launches the lifelogging based on that. To the best of our knowledge, Memento is the first of its kind lifelogging system. Through a series of techniques, Memento integrates EEG and proposes the two phase emotion recognition that makes it efficient and affordable on wearables. Finally Memento outputs lifelogs tagged with emotion information which we believe could enhance many existing services.

IX. Ackowledgements

We acknowledge the support from MOE by its Tier 1 grant RG130/16 and Tier 2 grant MOE2016-T2-2-023, as well as the NAP grant and COE grant from NTU.

References

[1] "10-20 system," https://goo.gl/8Yzwv5.
[2] "Electroencephalography," https://goo.gl/okX9pC.
[3] "Everyday," http://everyday-app.com/.
[4] "Instant," http://instantapp.today/.
[5] "Muse eeg headband," http://www.choosemuse.com/.
[6] "Narrative clip," http://getnarrative.com/.
[7] "Vicon," https://www.vicon.com/.
[8] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, pp. 49 – 59, 1994.
[9] S. Chennuru, P.-W. Chen, J. Zhu, and J. Y. Zhang, "Mobile lifelogger–recording, indexing, and understanding a mobile users life," 2010, pp. 263–281.
[10] J. F. Cohn, "Facial expression and emotion," in *Proceedings of the 8th International Conference on Multimodal Interfaces*, 2006, pp. 233–238.
[11] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J. A. Landay, "Activity sensing in the wild: A field trial of ubifit garden," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2008, pp. 1797–1806.
[12] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.
[13] R. Gouveia and E. Karapanos, "Footprint tracker: Supporting diary studies with lifelogging," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2921–2930.
[14] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, "Sensecam: A retrospective memory aid," in *Proceedings of the 8th International Conference on Ubiquitous Computing*, 2006, pp. 177–193.
[15] R. Horlings, D. Datcu, and L. J. M. Rothkrantz, "Emotion recognition using brain activity," in *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*, 2008, pp. 6:II.1–6:1.
[16] L. Ivonin, H.-M. Chang, W. Chen, and M. Rauterberg, "Unconscious emotions: Quantifying and logging something we are not aware of," *Personal Ubiquitous Comput.*, vol. 17, no. 4, pp. 663–673, Apr. 2013.
[17] W. James, "What is an emotion?" *Mind*, vol. 9, no. 34, pp. 188–205, 1884.
[18] M. J. Katz, "Fractals and the analysis of waveforms," *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145 – 156, 1988.
[19] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis ;using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, Jan 2012.
[20] N. D. Lane, P. Georgiev, C. Mascolo, and Y. Gao, "Zoe: A cloud-less dialog-enabled continuous sensing wearable exploiting heterogeneous computation," in *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, 2015, pp. 273–286.
[21] C. M. Lee and S. S. Narayanan, "Toward detecting emotions in spoken dialogs," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 2, pp. 293–303, March 2005.
[22] H. Lee, Y. S. Choi, S. Lee, and I. P. Park, "Towards unobtrusive emotion recognition for affective social communication," Jan 2012, pp. 260–264.
[23] M. L. Lee and A. K. Dey, "Lifelogging memory appliance for people with episodic memory impairment," in *Proceedings of the 10th International Conference on Ubiquitous Computing*, 2008, pp. 44–53.
[24] Y. Li, Z. Ma, W. Lu, and Y. Li, "Automatic removal of the eye blink artifact from eeg using an ica-based template matching approach," *Physiological Measurement*, vol. 27, no. 4, p. 425.
[25] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong, "Moodscope: Building a mood sensor from smartphone usage patterns," in *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, 2013, pp. 465–466.
[26] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time eeg-based emotion recognition and its applications," 2011, pp. 256–277.
[27] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "Lifting scheme for human emotion recognition using eeg," in *2008 International Symposium on Information Technology*, Aug 2008, pp. 1–7.
[28] D. H. Nguyen, G. Marcu, G. R. Hayes, K. N. Truong, J. Scott, M. Langheinrich, and C. Roduner, "Encountering sensecam: Personal recording technologies in everyday life," in *Proceedings of the 11th International Conference on Ubiquitous Computing*, 2009, pp. 165–174.
[29] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from eeg using higher order crossings," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 186–197, March 2010.
[30] R. Rawassizadeh, M. Tomitsch, K. Wac, and A. M. Tjoa, "Ubiqlog: A generic mobile phone-based life-log framework," *Personal Ubiquitous Comput.*, vol. 17, no. 4, pp. 621–637, Apr. 2013.