Evaluating the Lifelog: a Serious Game for Reminiscence

William Burns¹, Paul McCullagh², Chris Nugent² and Huiru Zheng²

¹ School of Engineering, Ulster University, United Kingdom
² Computer Science Research Institute, Ulster University, United Kingdom

{wp.burns, pj.mccullagh, cd.nugent, h.zheng}@ulster.ac.uk

Abstract. Body worn cameras record lifelogs as a sequence of images and one therapeutic use is to promote reminiscence in older people. We investigate if the emotional response of the viewer can be used to identify images of interest and whether this can become a serious game for shared interaction with a family member or carer. To evaluate whether this is technically feasible we report on a small evaluation of five healthy participants (Male=3, Female=2) aged between 24-46 years of age. Participants reviewed lifelog images six months after the initial data collection. Galvanic skin response readings were recorded and matched to the image stimuli. By monitoring such responses it is possible to organise the lifelog into events, potentially highlighting activities of daily living and social interaction for subsequent reminiscence. Initial results indicate emotional responses can be quantified and detected but no clear classification of emotional trends emerged. We suggest improvements in methodology to make the approach viable and discuss the need for data reduction. As wearable technology improves, the approach can add to the quantified-self paradigm, allowing wider application to learning and training.

Keywords: Reminiscence, Serious Game, Galvanic Skin Response, Personalisation, Pervasive Computing.

1 Introduction

Reminiscence Therapy (RT) involves the “discussion of past activities, events and experiences with another person, usually with the aid of tangible prompts such as photographs” [1]. There is evidence to suggest that it is effective in improving mood in older people. As such, it may be the basis of a ‘serious game’ for older people and carers/family members to support shared interaction and inclusion. As a side effect, it may stimulate memories and hence support cognitive wellbeing, although this process is less well understood. According to Collins dictionary, a lifelog is “the practice of digitally recording all of one's daily activities by for example wearing a camera that takes photographs every few seconds”. The resulting images may be uploaded to a computer for later retrieval, thereby building up a digital record of daily existence.
activities. A lifelog could be used to monitor social interaction and to monitor food intake and other activities of daily living (ADL). Gordon Webb is a pioneer of the silver surfers [2] who have adopted the quantified-self paradigm [3], utilizing body worn cameras (e.g., Microsoft SenseCam) for lifelogging. He stated: “I have built up about 100 sequences, consisting of 60,000 to 80,000 images. That content varies from pleasant walks to conferences, so I can see who I was talking to. The camera marks the time that all the pictures were taken, and then I upload the images into a single folder, labelling the sequence so that I can retrieve it later.” [4].

In addition to this extension of the pervasive computing paradigm, human computer interaction continues to evolve by enriching the computer with artificial intelligence attributes. Affective computing couples the user with the technology in a close symbiotic relationship; it relates to or arises from human emotions and encompasses computer science, psychology and cognitive science [5]. The use of physiological recordings (e.g., heart rate, respiratory rate) is a critical input for quantifying affective computing. Speech metrics, facial and gesture recognition can be used to interpret a user’s emotions in this paradigm.

1.1 Galvanic Skin Response

The Galvanic Skin Response (GSR) measures the electrical conductance of the skin; it can vary depending on the amount of moisture, specifically sweat, on the skin. As the production of sweat is controlled by the sympathetic nervous system [6][7], the measurement of GSR has been used to indicate psychological or physiological arousal [8][9]. If the sympathetic nervous system is highly aroused, then the body activates sweat glands, which in turn increases skin conductance.

A GSR sensor measures the electrical conductance between two points on the body, typically at the extremities (fingers and palms), usually positioned about an inch apart and measured by injecting a harmless small current. Conductance varies depending on the emotional state of the subject. The most common use of GSR to measure arousal is as a component of a polygraph test (lie detector) [10], in conjunction with other parameters such as heart rate and respiration rate. Biofeedback therapy also uses GSR to measure and display a subject’s stress levels with the view to help them identify and control anxiety [11].

The environment in which the GSR data collection occurs can have an impact on the consistency of results. Factors such as temperature and humidity naturally affect the skin conductance of the subject [12][13]. In addition to this, GSR responses are typically delayed between 1-3 seconds after the presented stimuli [13] so temporal adjustment is needed to match responses to stimuli.

1.2 Serious Games and Reminiscence

A serious game has a primary purpose other than entertainment, in this case to promote social interaction and wellbeing. Review of lifelogs with a family member can be considered as a serious game, providing context for shared interaction. McCallum and Boletsis have provided a taxonomy of serious games for dementia.
The taxonomy categorises games as {cognitive, physical, social-emotional}, game types as {preventative, rehabilitative, educative, assessing} and player types as {potential patient, patient, public, professional}. Cognitive games trigger the cognitive abilities of the player; social-emotional games encourage players to link with their friends, providing shared experiences and discussion opportunities. Rehabilitative games have therapeutic functionality; assessing games to provide data to the player about his/her health status. Hence the work described in this paper may be categorised as: game category {cognitive, social-emotional}, game type {assessing} and player type {professional}. Of course, our intention is to investigate potential patients in the future, with a view to provide rehabilitation of cognitive function.

However, the jury is still out on the efficacy of such games for older people. A review by McCallum and Boletsis [15] concludes that: "many games developed for entertainment purposes are used for health reasons, acquiring the characteristics of serious games..... dementia games do have an effect on cognitive impaired people. If that effect is longlasting and/or transferable to the daily activities is a matter of further scientific investigation."

Our hypothesis is that context such as location, people involved and changes in environment within the lifelog will have an effect on the viewer’s arousal. The goal of this evaluation is to determine if GSR signals can be used to identify ‘images of interest’ during reminiscence of lifelog data. The rationale is that arousal can potentially be used to personalise the organisation of the collected data. This could result in streamlined lifelogs with more meaningful events that stimulate the user and hence promote shared interaction with a carer. By providing further context for the interaction, such as assessing for an event of interest: {what happened?, who is involved?, where did it take place?, when did it take place?, why did that happen?}, we can build this reminiscence interaction into a serious game.

2 Methods

Participants were asked to use software to review previously recorded lifelog data. A reminiscence software package was developed to support this. They used a wearable sensor around their wrist that connected, via two small wires/electrodes to two of their fingers, to collect GSR data. Following a brief calibration phase, participants were shown a series of images, in order to achieve a baseline measure. Participants navigated a previously recorded lifelog at their own pace while GSR data were subsequently recorded. The data collected were (i) GSR data sampled at 10.6 Hz (Comma Separated Variable, CSV format), (ii) Screen recording of both GSR waveforms and Reminiscence package screens, i.e. the evoking images.

The data were collected using a Shimmer sensor [16] with an additional GSR board with two electrodes attached to the index and middle fingers of the user’s left hand. In order to collect the GSR data from the sensor, the Shimmer Connect software [17] application was used to stream data from the sensor to the recording computer.

Participants were recruited to test the hypothesis that 1) reminiscing of lifelog data has a physiological effect on the reviewer and 2) that this change in physiological
arousal can be used to personalise lifelog and hence reminiscence. Five participants from a previous evaluation of a wearable lifelog system [18][19] were recruited to review their lifelog data whilst wearing a GSR sensor, see Table 3 for user demographics. The Faculty of Computing and Engineering Research Ethics Filter Committee at Ulster University granted ethical approval for this research (study number: 2014.030614.22).

As all of the participants were healthy volunteers (without cognitive decline); the aim was to test efficacy of the methodology. The recording took place six months following the collection of the initial lifelog data in order to minimise the possibility of simple memory recall and hence simulate reminiscence. Participants were asked to sit quietly while eight standard images were presented (as slides using Microsoft PowerPoint) to establish a personalised baseline, as GSR responses are highly individual and direct comparisons between subjects can be misleading [20]. Each slide was shown for 5 seconds and contained images of varied valence. The images aimed to promote both positive (kitten & puppies) and negative (snake & spiders) emotional responses. In addition, two of the eight slides contained coloured words that invoke the Stroop Effect [21]. The Stroop effect is known to evoke stress levels by affecting an identification task (e.g., the word “Red” written in “Blue”; identify colour or meaning). The baseline allowed for settling of the GSR personalised to each individual, for the ambient experimental conditions.

Participants were then presented with the lifelog images they had recorded using the miLifeCam sensor component [18]. Approximately 300 images had been recorded for each participant. These images were played in sequence at a rate of one image every two seconds. During this time, participants sat in a silent room with the researcher passively viewing the lifelog; they did not actively interact with the technology or researcher. Participants were asked to remain silent throughout the evaluation. The researcher recorded the GSR data as a CSV file for later analysis. In addition to the capture of the GSR data, stimulus and waveform screens were recorded, so that GSR readings could be matched with stimuli. The GSR data were subsequently manually aligned with the screen recording of the lifelog review to identify associations between the images displayed and the participant’s emotional arousal.

The researcher reviewed the screen recordings and the data were segmented into six categories (Table 1). The GSR data were visualised into charts and annotated with the categories of lifelog images, factoring in the delay in GSR response [13]. When interesting GSR events, such as sharp peaks and troughs, were visually identified, the corresponding visual stimulus was retrieved and added to the annotated GSR charts. A deviation of more than 5% of the range of the data over 5 successive time samples was used to determine this. In Section 3 an annotated GSR chart and summary explanation is presented for one of the five participants along with a meta-analysis of the remaining four.
3 Results

GSR responses together with an overlay of the visual stimuli were developed for each participant. From the collected lifelog data from all participants, four event types were identified along with two non-specific event types, presented in Table 1. This allows the identification of the type of activities that were viewed during the reminiscence period that produce potentially interesting peaks and troughs in the GSR data, i.e. changes in skin conductance.

Table 1. Event types used to segment the GSR data during the reminiscence period, with typical activities that occurred during these events.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Typical Activities Within This Event</th>
<th>Total number if GSR events Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>PowerPoint Viewing (baseline event)</td>
<td>Viewing the images described above.</td>
<td>1</td>
</tr>
<tr>
<td>Bathroom (ethical consideration)</td>
<td>Camera deliberately obstructed or removed for the participant to use the bathroom. Typically displayed black/dark images.</td>
<td>0</td>
</tr>
<tr>
<td>Stationary</td>
<td>Sitting at their desk. Not moving between environments with no significant interactions.</td>
<td>14</td>
</tr>
<tr>
<td>Travel</td>
<td>Moving between rooms. Walking along corridors and walking up/down stairs. Sometimes results in blurry images.</td>
<td>4</td>
</tr>
<tr>
<td>Social Interaction</td>
<td>A social interaction with a colleague, or researcher. Clear visual identification of another person in close proximity.</td>
<td>7</td>
</tr>
<tr>
<td>Eating</td>
<td>In the process of eating food, either alone or with another person. Clear image of food being consumed.</td>
<td>0</td>
</tr>
</tbody>
</table>

3.1 Sample Participant (P5)

P5 was a 46 year-old male with a miLifeCam dataset of 319 images. From the GSR dataset collected, seven GSR events were identified, see Figure 1. The first event was captured when the participant was looking at the PowerPoint presentation. This is the only participant of this evaluation who showed any significant response during the PowerPoint phase. Nevertheless, as this evaluation was investigating the effect of lifelog data on physiological signs, this event is not deemed important from a reminiscence perspective, yet it illustrates the possibility of GSR for quantification. The second event occurred just after a colleague joined the participant for lunch, see Figure 2. The interaction that seems to trigger the response is the image of the new person/interaction sitting down at the table.
Fig. 1. P5’s GSR data (conductance specified in kOhms versus time epoch) recorded using the Shimmer sensor during the reminiscence phase. Troughs in the data were identified and the corresponding image the participant was viewing at the time is overlaid upon the GSR.

Fig. 2. Screenshot of the Reminiscence Tool interface, presenting a sample social interaction image presented to P5 during reminiscence GSR event detailed in Figure 1.
On some of the images, P5 was looking at the food he was consuming. During the social interaction, it may be deduced that P5 was using his smartphone and it is reasonable to assume that this may have been the catalyst for the new GSR event. The fourth event occurs during a series of images clearly showing the two social interactions the participant was having. No other pieces of technology or persons were in the series of images. The next identified event was when the participant had returned to their office and began to work on an academic paper. After a series of images of the document on the screen, P5 reviews a paper copy of a document. At this time P5 lays back in his chair offering a view out of their office window. This triggers the sixth event. The seventh GSR event identified occurs as P5, again lying back in the chair, views the monitor with a document displayed.

Figure 3 summarises the GSR data of all five participants in this study. P3 and P5 have similar levels of GSR, as do P2 and P4. It should be noted that during P1’s evaluation phase, the temperature in the room during recording was 28 degrees Celsius, which resulted in much higher skin conductance levels by comparison with the other four participants.

![Graph showing GSR data for all five participants.](image)

**Fig. 3.** Variability of GSR over time during reminiscence.

### 3.2 All Participants

Following the analysis of the collected GSR data of all the participants and the visual identification of GSR changes, Table 2 was compiled to categorise the event types and occurrences. Of all the interesting events identified from the GSR data, the event type that had the most GSR events identified was paradoxically “Stationary” with 14 events identified. “Social Interaction” event types were the second most common event type to generate a GSR response with seven occurrences. “Travel” initiated four GSR responses. Somewhat counter-intuitively, eating was not reflected in GSR events.
Table 2. Participant demographics and associated events evoked from lifelog images (P1 - 5).

<table>
<thead>
<tr>
<th>Participant Demographics</th>
<th>Life Log images</th>
<th>Events</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 (Female, 45)</td>
<td>327</td>
<td>E1.1 Baseline (Stroop)</td>
<td>Due to high temperature, the GSR looked to be saturated (see Figure 3 GSR P1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E1.2 Stationary</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E1.3 Social Interaction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.1 Stationary</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.2 Social Interaction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.3 – 2.6 Stationary</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.7 – 2.8 Travel</td>
<td></td>
</tr>
<tr>
<td>P2 (Male, 34)</td>
<td>310</td>
<td>E2.1 Stationary</td>
<td>P2 had the most peaks and troughs in their GSR data. Nevertheless, these GSR features were not as defined as those of P5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.2 Social Interaction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.3 – 2.6 Stationary</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2.7 – 2.8 Travel</td>
<td></td>
</tr>
<tr>
<td>P3 (Female, 24)</td>
<td>325</td>
<td>E3.1 – 3.4 Stationary</td>
<td>Despite all of the GSR events occurring during a Stationary event type, the lifelog data shows that P3 carried out five of the six event types in Table 2</td>
</tr>
<tr>
<td>P4 (Male, 32)</td>
<td>335</td>
<td>E4.1 Travel</td>
<td>Both travel and stationary event identified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E4.2 – 4.3 Social Inter-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>action</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E4.4 Travel</td>
<td></td>
</tr>
<tr>
<td>P5 (Male, 46)</td>
<td>319</td>
<td>E5.1 Baseline (Puppies)</td>
<td>More detail can be inspected in Fig 1, with evoking stimuli in Fig 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E5.2 – 5.4 Social Inter-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>action</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E5.5 – 5.8 Stationary</td>
<td></td>
</tr>
</tbody>
</table>

4 Discussion and Recommendations

This experimental approach attempts to correlate visual stimuli with subjectively assessed GSR arousal. If we can measure relevant events with GSR then it be a means of prompting the recording of the lifelog itself. The rationale for this is the need to reduce the overall lifelog data set for pragmatic deployment. Lifelogs generate so much data of a similar nature that they become less useful, suffering from the ‘big data’ anomaly. Contextual information can potentially facilitate a reduction in data collection and storage.

The images used in the PowerPoint presentation (baseline) were chosen in order to evoke positive and negative emotions. The use of IAPS (International Affective Picture System) [23] could have been used as it provides images that have a standardised emotion evoking score [24]. GSR may change according to both exogenous (external) and endogenous (internal) events. Certain aspects of the collected lifelog data trigger some form of emotional arousal as identified by the GSR recording. Nevertheless, they may not be consistent enough for us to conclude that changes in environment, as captured visually, prompt an emotional response. While there may be a relationship between the images viewed by the participants and their GSR arousal, addi-
tional modes of stimulation could produce a more representative arousal level. If additional information, such as a sound captured, was played back during reminiscence, it may have a more consistent effect on the participant’s arousal levels. Other relevant sensory stimuli are not capable of being captured, as part of the lifelog (e.g. the ‘smell’ of an environment that may prompt an emotional response, a technique well known to supermarkets). Research is being conducted into the transmission of smells using technology systems [22].

The detection of specific events within GSR data is challenging due to the sensor measurement being ambiguous and dependent on hidden contexts such as physical activity and room and body temperature [26]. Additional physiological metrics could be included in order to obtain a more robust measure of emotional stimulation in reminiscence. Using GSR alone as a metric for assessing emotional response when reviewing large datasets appears promising but may not be specific. Metrics such as heart rate, body temperature and respiration rate would also result in more quantitative and robust results. Nevertheless, intrusiveness would be increased.

It is also unclear if the responses are as a direct result of the stimuli presented to the participants, as opposed to environmental factors (sound, temperature and physical comfort during the review process etc.) or even boredom. To address this issue the addition of eye-tracking technology would help us to establish task engagement. By combining the GSR data with eye tracking it may be possible to establish what aspects of a particular image may have influenced the physiological change. A less sophisticated alternative to eye tracking, would be asking participants to talk ‘out loud’ about what they felt when reviewing the images. For example, if they remembered anything about the people or environments that were shown.

The identification of a GSR ‘event’ should also be empirically verified in an automated way. The use of a derivative and threshold approach would allow the automated detection of specific change points. Another approach is the use of CUSUM (Cumulative Sum) as a means to detect changes in the recorded physiological data [25]. A more rigorous approach would harness data analytic algorithms for the identification of significant events. Advances in wearable technologies mean that the collection of physiological data, such as GSR, heart rate and body temperature are becoming more ubiquitous (and less costly). Nevertheless, these commercially available devices are typically tied to their own software ecosystem and do not allow third party developers access to the raw data in order to classify specific emotions. If the raw data were available, the scope of emotional response recognition has many applications outside the realm of healthcare technologies.

One limitation of this investigation is that all the participants were younger, <=46 years of age, and as such would typically have better memory recall than older participant for whom reminiscence is designed. A further evaluation should be undertaken with older participants and their lifelog data, as age and medication consumption have been shown to have an effect on the GSR results [12]. The use of IAPS images would establish a verified baseline. This study was carried out with only images and no additional stimuli such as sound. The introduction of sounds would have some effect on the participant’s arousal level. The use of sounds in addition to visual stimuli would be suited to serious gaming.
4.1 A serious game for reminiscence

In this work we have evaluated the use of GSR with healthy volunteers. With a lifelog dataset, collected by a person with Dementia (PwD), a serious game can be developed to match the taxonomy of [14], specifically the Cognitive/Social-emotional serious games category, thus providing an Assessing serious game, possible a Rehabilitative game. By modifying the Reminiscence Tool used to review the life-logged data, (Who, What, When, Where, Why) questions based on the collected images could be asked, Fig 4. For example: Who is this, What are they/you doing, When and Where was this and Why were you here/doing that? Using the GSR recording hardware, the user’s emotional response and thus agitation could be monitored resulting in onscreen hints and prompts. Coupled with eye-tracking, the user’s engagement could be monitored and multimedia prompts could be provided. Metrics such as the number of correct and incorrect answers, the duration of task completions, total engagement and physiological responses could be used by professionals and carers to provide an assessment over time.

![Image of a serious game interface](https://example.com/image)

Fig. 4. Serious game for assessment of reminiscence (proposed).

4.2 More general applications

In the longer term it is indeed possible that the emotional response, and thus attention, of students could be monitored in real time by educators that could i) specifically target a student who is falling behind in a class or, ii) modify their teaching techniques to accommodate students preferred learning methods. By responding to changes in emotional response while looking at visual stimuli further possibilities for use in serious gaming can be identified. Difficulty, duration and intensity can be adjusted to match the user’s affective state. One example would be how a pilot would respond emotionally to various scenarios during a flight simulation. In addition to testing the emotional levels of a user in a serious game, mindfulness could be explored in real-time. During a stressful simulation, an affective serious game could not only monitor the stress levels of the user, but assess how they overcome the increased stress levels.
to solve the problem posed. The proposed methodology can also be applied to the design of the architecture of virtual environments. By placing the user in a virtual environment, for example a newly proposed museum, the layout of the building in addition to the placement of exhibits could be evaluated for changes in emotion. The placement of artwork in a dimly lit room would potentially evoke a different response than if the same artwork were placed in a well-lit and cluttered area. The physical layout of the building could also be evaluated for ‘pain points’ in navigating to various exhibits.

5 Conclusion

This paper presented a feasibility study into the use of GSR to ascertain if there is a physiological response from participants reviewing lifelog data, collected six months previously. We addressed whether such physiological data could be used to personalise the lifelog. Results show that the reminiscence of lifelog images affects four of the five participants evaluated. Incidentally approximately 10% of participants in GSR studies are non-responders (hypo-responsive) with regards to their electro dermal activity [13]. The small sample size and data variability mean that no specific event type can be identified, nor was any specific emotion identified. Additional stimuli during the life-log recording, such as sound, could trigger more GSR responses than reminiscing with images alone. With the computation power of smartphone technologies, and the wearable nature of the GSR sensor, it would be possible to modify the miLifeCam sensor component to integrate the GSR sensor input to trigger the capture of images as and when the user’s arousal levels change above a certain threshold (or indeed based upon some other criterion). This affective miLifeCam component would in turn reduce the number of images captured and mean that the data presented during reminiscence could be more emotionally relevant to users. As wearable technology improves, the approach can add to the quantified-self paradigm, allowing wider application to learning and training.

6 References