

Affective computing for wearable diary and lifelogging systems: An overview *

Jana Machajdik¹, Allan Hanbury², Angelika Garz¹, Robert Sablatnig¹

¹ Computer Vision Lab
Vienna University of Technology
{jana, garz, sab}@caa.tuwien.ac.at

² Information Retrieval Facility
Vienna, Austria
a.hanbury@ir-facility.org

Abstract

Diaries have transformed over the last decade. Originally in handwritten text format, photo albums and visual diaries became popular as photography became commonly available. Traditionally intended to remain private, the dimension of audience was added in blogs along with Internet communication. However, humans have a limited capacity to record their lives. Lifelogs overcome this limit and collect and store a person's personal information digitally. This can be done by recording all computer and cell phone activity and mobile context (e.g. GPS), but also adding multiple wearable sensors such as "always-on" cameras or bio-sensors. The enormous amounts of data created needs to be processed in order to be made useful to humans. In this paper, we review state-of-the-art literature on wearable diaries and lifelogging systems, and discuss the key issues and main challenges.

1 Diaries, Blogs and Lifelogs

Humans record their history, from global events that move the world to small everyday life events that they consider important. In this paper, we review the current trends of personal record keeping. We focus on approaches using wearable sensors to passively capture data.

First, forms of personal record keeping are discussed, especially how diaries have changed over time, becoming increasingly multimodal. Section 2 describes key issues and main challenges of building modern diary systems. In Section 3 we present selected projects which offer unique diary solutions. We conclude with thoughts on future challenges are offered.

1.1. Traditional diary

A diary is a sequence of entries arranged chronologically, created to report on what has happened over the course of a period of time. Personal diaries usually include the writer's thoughts and feelings. Originally handwritten in the form of books, the diary transformed from paper to electronic formats. Along with new media, new forms of diaries also developed.

1.1.1 Visual diary

Since photography became commonly available, photos were added to the text in the diary to illustrate events. In some cases the emphasis shifted towards photography altogether and it became common to create *photo albums* with short text comments. Photos simplified the record keeping and added a further dimension to the records that made the memories seem more real [26]. Sontag [26] goes even further and says that it is the photographic capture of reality that gives us the feeling of the realness of our lives, it helps us reconstruct our personal history. We don't believe our perception until a photo confirms it [26] as is also illustrated by the popular Internet-phrase "pics or it didn't happen"

*This work was supported by FFG, project no. 830043

demanding photographic proof of an unbelievable story. Artists use *visual diaries* to sketch drawings of their ideas and to collect images or other media. They serve them as inspiration and as a means to reflect on their artistic growth [1]. In *travel diaries* the author might even add small souvenirs, brochures, postcards or other nostalgic items to the diary.

1.1.2 Blogs

The most distinct and today probably also most popular form of diary are *blogs*. While the “paper version” was traditionally intended to remain private, this aspect widely changed as people adopted the new medium to chronicle their lives with the additional dimension of audience. Although the reasons given by bloggers such as *documenting my life*, *expressing opinions*, *letting off steam (shouts, feelings and thoughts)*, *inspiration (“to write is to think”)*, and *building communities*, the bloggers are conscious about their audience and censor the information they publish [21]. Admittedly, the entries serve rather for self-presentation and narcissism than for creating an extensive digital memory [21]. Still, the main point of a traditional diary (public or private) is that the content is actively generated by the user. However, documenting a life takes discipline and effort and, hence, the amount of information in the traditional diary is constrained either by the capacity of the medium or the writer. This limit forces the individual to filter the content he/she writes down.

1.2. LifeLog

This issue of necessity (and difficulty) to limit information that can be preserved in a diary by a single human is being addressed by new technologies. Technological advancements today make it possible to explore new ways and possibilities to capture, collect and store information. First envisioned by Vannevar Bush in 1945 [5], a LifeLog presents the notion to capture and store a whole lifetime of a person’s personal information digitally, so that it can be retrieved whenever needed. In 2001, this idea was revived by an experiment of Gordon Bell [2], who scanned all of his paperwork, photographs, medical records and other personal data. The initial focus for lifelogging was on desktop applications only, however, it has shifted towards mobile access and capture. Mobile devices are capable of more than just storing - they capture data passively, without the user’s conscious initiative. Lifelog systems collect a variety of signals - subsets of the following data categories (as described in [7]):

Passive visual capture: Wearable “always-on” cameras automatically take images or videos.

Biometrics: Wearable sensors measuring bio-signals, such as heart rate, galvanic skin response, skin temperature or body motion.

Mobile context: Cell phones can provide information such as location cues in the form of GPS data, wireless network presence and GSM location data. Co-present Bluetooth devices may indicate people present nearby.

Mobile activity: Call logs, SMSs, even email-logs, activity on the web and social networks sites can be gathered from mobile phones.

Desktop/laptop computer activity: Basically all PC/laptop activity of the user can be monitored, the time and duration of each task measured, documents saved, etc.

Active capture: Indirect (writing blogs is monitored as computer activity) or direct (add photos, write comments and annotate the lifelog’s content) capture.

Such technologies, however, are not limited to the obvious use of personal reminiscence only. Various applications featuring at least a subset of these possibilities have been used, e.g. in medical and therapeutic solutions, security enhancement, or to encourage self-reflection.

2 Challenges

The main challenges for building a lifelog system (as also described in [7]) can be divided into two parts: the *information capture* and the *post-processing*.

2.1. Information capture

Hardware, is the basis of the system. The main requirements on such devices are: *stability* (in terms of hardware and software), *reliability* of measurements (measured correctly and in time), a long *battery life*, sufficient *on-device storage*, unobtrusive *wearability*, and ideally being completely wireless. The *robustness* of hardware and software is especially crucial in passive sensors, since crashes are invisible to the user and it can take hours or days until the crash is observed and the user intervenes. Various sensors are available today, however few comply with the above requirements, e.g. Microsoft's SenseCam, BodyMedia SenseWear armband, Polar heart rate monitor, the Q sensor (Affectiva), Exmocare watch, various spycams, and smartphones equipped with GPS etc.

A *central storage* to collect and process the data from the various mobile devices is part of the system. Though an automatic and wirelessly data upload would be ideal it is not yet feasible, resulting in the need of manual data transfer which presents a noticeable inconvenience to the user.

The human issue presents a further considerable difficulty in conjunction with multiple mobile devices. *Forgetfulness* is the main issue, since forgetting to recharge batteries or to download data as well as leaving devices behind will result in missing or corrupt data. Concurrently, the user's *comfort* and with it his/her willingness to wear mobile sensors has to be kept in mind. Finally, the system and its developers need to consider the *privacy* of the users and their near surrounding, as well as applicable privacy laws. It should be possible to easily turn off sensors when it is appropriate (e.g. in bathrooms, dressing rooms, schools or when people object to being recorded).

2.2. The WORN effect

The reason diaries are (or can be) useful, is not solely to remember events – they can also serve as a means of self-reflection. Self-reflection is an important adult process leading to further self awareness and development [22]. For this to happen it helps when the writer re-visits old diary entries to contemplate on them. However, this function of a diary is little exploited since most diaries (and now lifelogs) are *WORN* - *write once, read never* [3]. In most cases, this is due to our limited capacity to create and read diaries. In traditional diaries, the user has to filter the data he/she writes down, resulting in too little information being recorded and the parts that they looking for in retrospect might be missing. Lifelogs, on the other hand, present the exact opposite problem. The captured information is too much for a human to sift through leading to the same result, namely not finding what is needed. In both cases the recorded information will be useless.

2.3. Processing

In order to make the captured multimodal and multimedia data searchable, and hence usable for the user, elaborate algorithms incorporating the steps enumerated below are needed [7]:

Synchronization of multiple devices. This is a pre-processing step where the input from the devices is time-aligned. Decisions to be taken are how to combine data from multiple devices having different sampling rates and how to handle corrupted, misaligned or lost time settings of a device. These occur e.g. due to crashes, traveling between time zones or time travel.

Data conversion and feature extraction. Multimodal raw data has to be converted into usable formats by extracting low-level features such as the heart rate from ECG or the step count from

accelerometer. Furthermore, the data is filtered and clustered based on similarity, rejecting low-quality and redundant data, such as over or underexposed images.

Transformation of the data into an interpretable format. This step is best explained mentioning a few examples such as transformation of GPS data into location names, of scanned documents into searchable text by applying OCR, or detecting semantic information in images through e.g. face recognition.

Data augmentation. The data is augmented by adding input from new external sources such as weather information or social context from social network platforms. The information from various sources is interrelated to gain new insights.

Finally, the system is made most useful when appropriate *visualizations* are used to present the extracted information and insights to the user.

2.4. Critique

Kalnikaitė [15] lists the main weaknesses of current prosthetic memory aids: (1) the main focus of research is mainly focused on technology, (2) there is little understanding on how these tools interact with the human organic memory, (3) little attention is given to the user needs, (4) there exist few tools for capturing autobiographical memory and (5) that there are too few long-term usage studies of current systems. An explicit description of potential value for users is necessary for lifelogging systems. Sellen and Whittaker [25] further argue that instead of trying to create one generic all-in-one system, designers should clearly define the specific purpose of their intended system. The goal of such systems should be synergy, not substitution of organic memory. It has been shown [15] that humans do not use memory aids as long as they are confident of their memory, even if the aid is accurate. One way to achieve useful systems is to strategically target the weaknesses of human memory, such as the *seven sins of memory* [24]. According to [25], following the current thesis to capture “as much as we can” has two weaknesses: we can never capture everything, and capturing too much is overwhelming for maintenance. Instead of *capturing the experience*, collections of digital data can serve as cues to trigger (organic) autobiographical memory about past events.

3 State of the Art

Current lifelogging applications differ in the goals or areas of usage, there is no all-in-one integrated approach yet. We identified two main roles that current lifelog-like applications are trying to fulfill:

- *Autobiographical memories* are captured for the purpose of re-experiencing the past by re-viewing old data. A possible expansion towards a more proactive role could be done by offering day summaries, statistics or other information that encourages the user to reflect on his/her life.
- *Personal memory assistant* type systems are concentrated on the present. They are meant to help in daily tasks, in situations where the user’s memory fails, e.g. “Where did I put my keys?”. Their purpose is to quickly execute very specific queries and produce an accurate answer to the search question. In the future, such systems should take a more proactive role and suggest relevant information to the user, even before they search for it.

However, clear definition of goals for such systems are too often omitted, thus the lines between the two roles are somewhat blurred. In the following we will mention a few such systems and present selected projects in detail:

Memory Glasses [9] derive information from context in a proactive fashion and display it on the vision field of the glasses. So far, it recognizes user-defined (tagged) locations and person names.

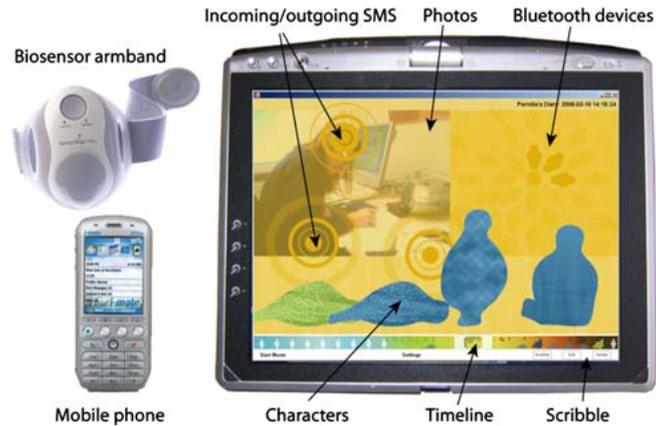


Figure 1. The Affective Diary system and user interface (from [27])

Ubiquitous Memories [16] is a wearable camera/glove system for remembering situations by touching objects. It is made for finding physical objects. *Virtual Augmented Memory* [12] implements a wearable camera for recognizing people based on face detection. *iRemember* [28] logs audio data, and implements speech recognition to enable phonetical or text searches within the data, as well as creation of summaries.

There are also other tools, which are rather *active* digital memory aids compared to the *passive* lifelogs, since they require active and conscious user input (such as manual calendar entries, specific *command* phrases, pushing of buttons, etc.) - we do not cover these in this paper.

In the following, we selected three projects to describe in detail. The emphasis is on systems that incorporate the fact that humans are emotional beings by using affective computing [23].

3.1. The Affective Diary

The Affective Diary [27] explores the emotional aspect to creating diaries and is designed to support self-reflection. It uses an armband equipped with bio-sensors (worn throughout the day) to automatically add information about how the user was feeling to the diary system. A user interface is provided where the physiological information is interpreted and visualized in the form of colorful blob as shown in Figure 1.

The sensors in the armband measure galvanic skin response (GSR), skin temperature, body motion, and steps. From these measurements it is possible to infer the emotional/physical state, which is visualized in the color of the blob, with blue indicating a calm state to red which indicates an excited state of the body. The body motion is encoded in the posture of the humanoid blob, ranging from a lying posture for situations where no motion occurred to a fully upright position during scenes with high physical activity. The user can review the data from the day, add her/his comments to any situation or include photos. Moreover, the software communicates with the user's mobile phone to additionally save SMS, call logs or the presence of other nearby Bluetooth devices logged by the phone throughout the day, indicating other people's presence in a ca. 10m proximity. Most users testing the system reported that it was a surprisingly revealing experience to use the Affective Diary.

3.2. SCOPE

Although the SCOPE [13] project is not designed as an explicit diary application, it presents a useful application of lifelog technology. SCOPE is a system for detecting childrens' potential emergency situations. It consists of wearable devices for pre-school children and a computer server in the school. The wearable device [19] includes a small camera, a triple-axis accelerometer with a dual-axis gyro-

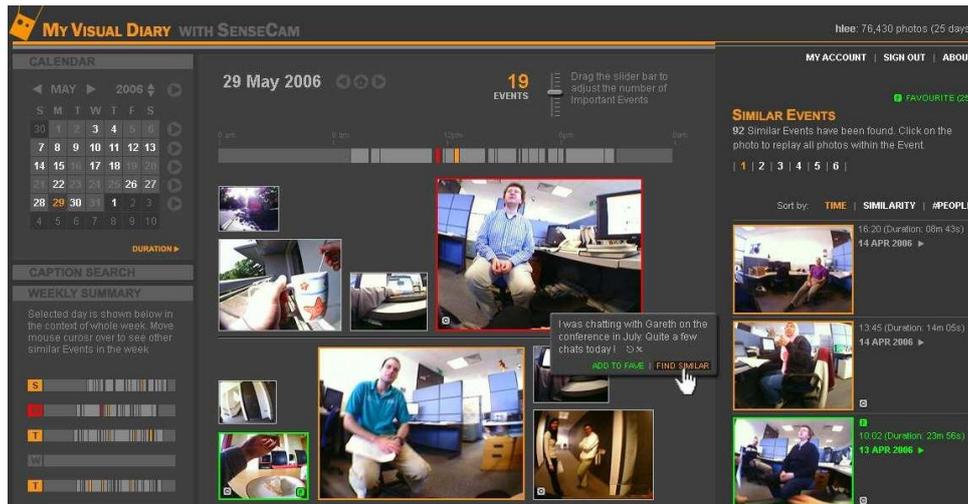


Figure 2. LifeLogging with SenseCam - user interfaced (from [18])

scope, a GPS receiver, and a heart rate monitor. An RF module communicates with the server and detects other devices (worn by the other children nearby). The measured signals are sent to the server, which then automatically recognizes each child's activity mode using an acceleration spectrogram, standard deviation of the heart rate, and the mutual spatial distances between children in the group. The activity modes of the children are classified into *walking*, *outdoor playing*, *eating* and *indoor playing*. When a potential emergency situation is detected, the device sends a photo that is uploaded online and initiates an "alarm". A photo is also taken every time there is an activity change. A summary of all activities during the day is created for parents. The system is currently being tested in Japan.

3.3. LifeLogging with the SenseCam

Microsoft's SenseCam is a wearable fisheye camera, typically worn on a string around the neck, designed to take photographs without user intervention. A passive infrared detector increases the probability that a picture is taken when another person stands in front of the wearer. A multiple-axis accelerometer is used to avoid taking blurry pictures that occur due to strong camera movement. The camera uses a timer to take a picture every few seconds, resulting in up to 3000 images per day. Worn daily, this adds up to approximately 1 million images per year.

To organize the data, *event segmentation* into distinct groups such as having lunch, driving etc. is performed [10], then for each event a *landmark photo* is selected which represents the event's content to the user [4], and the finally the *event novelty* is calculated based on how visually unique each event is in a given period of time [11]. It is assumed that a more unusual and unique moment will be more interesting to the user than commonly occurring daily rituals. For visualization, the landmark photos of events that have higher novelty value are presented in bigger sizes than events with lower novelty. The user interface is shown in Figure 2. The system is being expanded by also bringing in data from mobile phones to aid scene and location classification [8], techniques taken from the video domain, such as everyday concept detection [6] and first experiments with adding biometric sensors [17].

4 Conclusions and Future Work

In this paper we outlined the key challenges of lifelog solutions as well as three different state-of-the-art systems. The Affective Diary offers an aesthetic user interface that encourages users to reflect on their day, to evaluate their emotional reactions which are measured by wearable bio-sensors and

visualized as colorful blobs. The interpretation is left to the user. This system shows that even relatively simple, but well conceived setup can animate humans to reflect on their lives. The SCOPE system shows a unique security application of a lifelog-like system. Hereby, data from mobile context and the heart rate and motion of the user are processed to recognize emergency situations of children. The passive camera is only triggered to show the exact circumstances of such events. The image data is not processed, neither is user input possible — the system is completely automated and passive. LifeLog using SenseCam is a large-scale project aiming to incorporate all plausible sensors, mobile data and activity monitoring, theoretically suitable for a wide range of applications. However, to be made useful, the enormous amounts of data first have to be structured. So far, the main effort is concentrated on interpreting the visual data from the passive camera. However, the incorporation of the other data is relatively disconnected, the potential in large part untapped. Especially the inclusion of biometric data is missing, since emotions are one of the strongest indicators on which events are important to us, especially in the context of self-reflection [22, 27]. We propose a fusion of the Affective Diary system (including a bio-sensor to capture emotion) with a SenseCam-like wearable camera and image processing to create an affective visual diary/lifelog. We think that this combination makes sense, since emotional events are proven to be the most memorable ones and human memory is in large part based on visual information.

In conclusion, current lifelog solutions may answer the “when”, “where” and “who” — and in some cases even the “what”. However, thoughts and feelings — the most valuable aspect of traditional diary entries — are mostly missing in current lifelog systems. The “how (did I feel)?” is only partly answered by some (only the potential presence of a strong emotion is measured). The question “what did I think?”, however, still remains open. We don’t expect that to change in the near future. Documenting one’s thoughts will stay the responsibility of the user. However, encouraging (and offering the space for) (self-)reflection and documentation of thought should be an integral part of lifelogs. This is also the biggest challenge. And along the way to meet it, solutions have to be found to avoid the “write-once-read-never” effect. We can expect new creative solutions. After all, the current systems all have an explorative character and this is expected to stay the same for the near future.

References

- [1] D. K. Beattie. *Assessment in Art Education*. Art Education in Practice. Davis Pubns, 1 edition, March 1998.
- [2] G. Bell. A Personal Digital Store. *Communications of the ACM*, 44:86–91, January 2001.
- [3] G. Bell and J. Gemmell. *Total Recall: How the E-Memory Revolution Will Change Everything*. Dutton Adult, Sept. 2009.
- [4] M. Blighe, A. Doherty, A. F. Smeaton, and N. O’Connor. Keyframe Detection in Visual Lifelogs. In *PETRA*, 2008.
- [5] V. Bush. As We May Think. *The Atlantic Monthly*, 176(1):101–108, July 1945.
- [6] D. Byrne, A. R. Doherty, C. G. Snoek, G. J. Jones, and A. F. Smeaton. Everyday concept detection in visual lifelogs: validation, relationships and trends. *Multimedia Tools Appl.*, 49:119–144, August 2010.
- [7] D. Byrne, L. Kelly, and G. Jones. *Handbook of Research on Mobile Software Engineering*, chapter Multiple Multimodal Mobile Devices: Lessons Learned from Engineering Lifelog Solutions. IGI Publishing, 2010.
- [8] C. O. Connaire, M. Blighe, and N. O’Connor. SenseCam Image Localisation using Hierarchical SURF Trees. In *MMM*, 2009.

- [9] R. W. Devaul. *The memory glasses: wearable computing for just-in-time memory support*. PhD thesis, 2004. AAI0806327.
- [10] A. Doherty and A. F. Smeaton. Automatically Segmenting Lifelog Data Into Events. In *WIAMIS*, 2008.
- [11] A. Doherty and A. F. Smeaton. Combining Face Detection and Novelty to Identify Important Events in a Visual Lifelog. In *CIT*, 2008.
- [12] J. Farrington and V. Oni. Visual augmented memory (vam). In *ISWC*, pages 167–, Washington, DC, USA, 2000. IEEE Computer Society.
- [13] M. Hamanaka, Y. Murakami, A. Usami, Y. Miura, and S. Lee. System for Detecting Kindergartners’ Potential Emergency Situations. *WMSCI*, Vol. I:296–301, July 2010.
- [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 *UbiComp*, pages 177–193, 2006.
- [15] V. Kalnikaité. *Re-thinking Lifelogging: Designing Human-Centric Prosthetic Memory Devices*. PhD thesis, University of Sheffield, 2009.
- [16] T. Kawamura, T. Fukuhara, H. Takeda, Y. Kono, and M. Kidode. Ubiquitous memories: a memory externalization system using physical objects. *Personal Ubiquitous Comput.*, 11:287–298, April 2007.
- [17] L. Kelly and G. Jones. An Exploration of the Utility of GSR in Locating Events from Personal Lifelogs for Reflection. In *iHCI*, 2010.
- [18] H. Lee, A. Smeaton, N. O’Connor, G. Jones, M. Blighe, D. Byrne, A. Doherty, and C. Gurrin. Constructing a SenseCam Visual Diary as a Media Process. *Multimedia Systems*, 14:341–349, 2008.
- [19] S. Lee, J. Sohn, A. Usami, and M. Hamanaka. Development of Wearable Device by Kid’s friendly Design for Kid’s Safety. In *HCI*, 2010.
- [20] N. Muhlert, F. Milton, C. Butler, N. Kapur, and A. Zeman. Accelerated forgetting of real-life events in Transient Epileptic Amnesia. *Neuropsychologia*, 48(11):3235 – 3244, 2010.
- [21] B. A. Nardi, D. J. Schiano, M. Gumbrecht, and L. Swartz. Why We Blog. *Communications of the ACM*, 47:41–46, December 2004.
- [22] M. Pasupathi, T. Weeks, and C. Rice. Reflecting on Life. *Journal of Language and Social Psychology*, 25(3):244–263, 2006.
- [23] R. W. Picard. *Affective Computing*. MIT Press, Cambridge, MA, USA, 1997.
- [24] D. L. Schacter. The seven sins of memory. Insights from psychology and cognitive neuroscience. *Am Psychol*, 54(3):182–203, Mar. 1999.
- [25] A. J. Sellen and S. Whittaker. Beyond total capture: a constructive critique of lifelogging. *Commun. ACM*, 53:70–77, May 2010.
- [26] S. Sontag. The Image-World. In *On Photography*. Picador, 1977.
- [27] A. Ståhl, K. Höök, M. Svensson, A. S. Taylor, and M. Combetto. Experiencing the Affective Diary. *Personal Ubiquitous Comput.*, 13:365–378, June 2009.
- [28] S. Vemuri, C. Schmandt, and W. Bender. iremember: a personal, long-term memory prosthesis. In *CARPE*, pages 65–74, New York, NY, USA, 2006. ACM.