
Collaborative Capturing of Significant Life Memories

Norbert Gyorbiro

University of Aizu
ngyorbiro@acm.org

Henry Larkin

University of Aizu
research@logic.nu

Michael Cohen

University of Aizu
mcohen@u-aizu.ac.jp

Abstract

The continual advancement in technology and form-factor of personal devices have made possible “memory archiving” and “life logging,” the recording of a person’s life through wearable and ubiquitous computing technologies. The purpose of this research is to minimize the total data recorded while maximizing the number of significant or user-preferred “captured moments.” This media is captured automatically based on the subject's arousal level and collaboratively employing nearby camera phones. The recording groupware application analyzes collective emotion for more accurate capture.

Keywords

Collaborative life logging, significant memories, mobile groupware.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Algorithms, Design

Copyright is held by the author/owner(s).

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.

ACM 978-1-60558-930-5/10/04.

Introduction

Traditional life logging has two inherent limitations for many real world applications. One is the need for storing and transferring available information, especially if multiple devices are used to record activity. The second drawback is in making use of the recorded data, with an entire life recording being too large to review or present to others. In both instances, saving just the “highlights” of one's life is a possible solution.

Significant events are those that are unusual or different in a person's life. If a person were snowboarding, for example, then these times might be when they performed radical, dramatic stunts, or raced friends down trails. Being able to detect and record such moments could provide highlights of memories as a collation of recordings. Such applications extend themselves throughout a person's day: with friends at any event and recording times when there was laughter or when a surprise occurred, or riding a bike and meeting someone with whom we'd like to remember our first conversation.

Related works include Personal Audio Loop (PAL) [1], a continuous audio recording device designed to recover conversations of the recent past. The device uses a cyclic audio buffer that is overwritten with new recordings unless the user presses a button to keep the recently recorded data. StartleCam [2] is a wearable video camera recording system with an additional automated mode. It can automatically capture events that are likely to be interesting for the user, based upon physiological arousal level metered by continuously monitoring the wearer's skin conductivity. Recently captured images are stored or optionally transmitted to a web server.

The proposed personal informatics system can be seen as an evolutionary descendant of StartleCam. During the last ten years, the necessary equipment has miniaturized to fit in one's pocket in the form of a smartphone that communicates with a wireless physiological sensor that can be worn comfortably underneath clothing. Such developments allow possible productization but also open up new areas for research. A group of people in a mutual proximity can cooperate to create shared memories by automatically capturing important moments triggered by their own or other group members' networked devices. The capture can include various modalities: audio, images, and video based on the capability of the device and user preference. Furthermore, the system allows assessing the collective or emergent arousal level within the group.

Picard [6] considers the need of emotion theory for ambulatory people: emotional data collection and analysis outside artificial laboratory conditions. The proposed wearable system could provide a history of one's arousal level as well as collected arousal level of peers in one's proximity throughout daily activities.

Concept

The system records significant memories using automatic and collaborative media capture. The detection algorithm is based on physiological arousal level [7].

The system uses ad hoc peer-to-peer networking among a group of people. A significant event triggers a broadcast to neighboring nodes to capture and record media if they are facing the target. This media is uploaded to a shared drop box at later times.

Realtime sharing of arousal level within a group provides an auxiliary communication channel to extend the perception of the person's action. As a person experiences a significant event, the system sends a "look over here" signal to draw the attention of nearby observers. Observers may inadvertently face the person with their cameras, enhancing the distributed capture of media. Audio effects and haptic feedback can be used to convey the feeling of excitement.

Group emotion is also employed to capture an event where the person in focus may be experienced enough to not feel an increase in arousal level but spectators do. By employing a majority vote in the group's arousal level, such event can qualify as a significant event and is captured by the system. When replaying the recorded memories, aggregated arousal levels can be used as an unconscious collaborative rating of the highlights of the event ("the biggest snowboard jump") or when the group shared a similar excitement ("a shared surprise").

Recorded media might be scaled and composed together as salient stills [9] reflecting how significant a moment is or how many agree that is. Images of multi-angle capture can be automatically composed into 3D models with novel viewpoint selections by using a service such as Microsoft Photosynth [3], and enhanced by spatial playback of the recorded audio.

Usability

The recording system consists of two mobile devices: a smartphone and a heart rate monitor. The smartphone is held in one's breast pocket or hung around one's neck, with camera facing outward. Besides a decent camera, requirements are support of Bluetooth protocol

for wireless communication and a Global Positioning System (GPS) unit for determining location of devices participating in the collaborative capture, common among modern smartphone models.

Recent heart rate monitors come with sensors embedded in a fabric strap worn on the body with a small detachable unit that processes the signals and sends data over a wireless Bluetooth or low-power radio link to a mobile phone.

The application utilizes Bluetooth to receive data from the heart-rate sensor and also to coordinate the network of capture devices. Bluetooth works also outside of mobile data coverage, thus extending usability of the system. For day-long use, camera begins capture when heart rate picks up. This way battery is conserved, though at a cost of missing the first few seconds of sudden surprises (e.g. reading a book at a coffee shop when an earthquake occurs). Once active mode is entered (manually selected or automatically with heart rate), the camera records in a cyclic buffer. The buffer allows data from a few seconds previous to be saved (as there will be a small lag between the actual significant event and a user's physiological reaction) of that event. There is also lag between wireless communication to nearby devices.

Arousal Level Detection Algorithm

The heart rate monitor provides two valuable metrics. Heart rate (HR) measured in beats per minute (bpm), and normal beat-to-beat (NN) intervals measured in milliseconds (ms) used for Heart Rate Variability (HRV) analysis. The detection algorithm uses these two measures, and its construction here is based the

detailed review of methodologies in [8] with the applicability in a resource-constrained environment.

Encountering a surprising situation, the heart reacts with tachycardia, an increase in its pumping rate. This can be detected by recognition of the difference from the baseline HR. The baseline is dependent on the subject and also on the activity (an intense physical activity like snowboarding would increase the HR). Thus the baseline needs to be measured on a per person per activity basis. By establishing a platform for HR (a common formula with good approximation is $HR_{max}=220 - age$), several bands of arousal can be determined and detected from the baseline measure. A body-worn accelerometer or pedometer can provide information about the activity level of the person, an additional input that can be used in establishing HR baseline bpm level.

HRV correlates inversely with HR, but has been shown [5] to be more sensitive to arousal. Possible methods for analyzing HRV include time domain, frequency domain, and nonlinear methods. The time-domain based rMSSD measure is defined as the square root of the mean squared differences between successive NN intervals. It provides an estimation of short-term components of the HRV and was found [4] to become lower during a stressing situation and rises during relaxation periods.

The combination of the HR and rMSSD measurements can provide a good index for the arousal level in the current activity context. These calculations are relatively simple and mainly linear thus calculated quickly on smart phones. Besides speed, simplicity also

helps to conserve the battery, a highly valuable resource in mobile environments.

Network Algorithm

The network architecture is based on an ad hoc neighbor assistance design. Each device is aware of neighboring nodes within a proximity, creating an ad hoc wireless peer-to-peer network. A device broadcasts a Capture Request with its current GPS location to all the network addresses it knows of whenever a significant event is triggered. Where devices are equipped with electronic compasses (e.g. Apple iPhone 3GS), a significant simplification for nearby devices can be made. A nearby device simply records when it is facing the coordinates of the target device, within +/- 90 degrees.

Recorded media can be uploaded either manually or automatically on certain conditions (e.g. an idle time, a stationary time in one spot for an extended period, or when the device is plugged into a power source). An FTP server's details can be provided as the location for the drop box.

Conclusion

Utilizing the affective state of a group to detect significant memories for recording and later viewing represents an important step forward in life logging research. The system being developed employs collaborative capture, a mirror that reflects our significant moments as perceived by our peers. Further work will investigate various arousal detection algorithms, applications of emergent group emotions, and quintessential memory playback methods of recorded media.

Citations

- [1] Hayes, G. R.; Patel, S. N.; Truong, K. N.; Iachello, G.; Kientz, J. A.; Farmer, R. & Abowd, G. D. The Personal Audio Loop: Designing a Ubiquitous Audio-Based Memory Aid. Springer Verlag, pp. 168–179. (2004)
- [2] Healey, J. and Picard, R. W. StartleCam: A Cybernetic Wearable Camera. In Proc. of the 2nd IEEE Int. Symp. on Wearable Comp. ISWC. IEEE Computer Society, Washington, DC, 42. (1998)
- [3] Microsoft Photosynth. <http://photosynth.net/> (2009)
- [4] Lackschewitz, H., Huther, G., Kroner-Herwig, B. Physiological and psychological stress responses in adults with attention-deficit/hyperactivity disorder (adhd). Psychoneuroendocrinology. (2008)
- [5] Montebugnoli, L., Servidio, D., Miaton, R. A. & Prati, C. Heart rate variability: A sensitive parameter for detecting abnormal cardiocirculatory changes during a stressful dental procedure. J. Am. Dent. Assoc. 135(12), 1718–1723. (2008)
- [6] Picard, R.W. Affective Computing. MIT Press. (1997)
- [7] Russell, J. A circumplex model of affect. J. Pers. Soc. Psychol. 39(6), 1161–1178 (1980)
- [8] Sanches P. Supporting Self-Reflection in Everyday Life. Master's thesis, Swedish Inst. of Comp. Sci. (2008)
- [9] Teodosio, L., Bender, W. Salient stills. *ACM Trans. Mult. Comp. Comm. Appl.* 1, 1, 16-36. (2005)