Deutscher, M., Jeffrey, P., and Siu, N. "Information Capture Devices for Social Environments". To appear in Proceedings of *Second European Symposium on Ambient Intelligence (EUSAI 2004)*, November 8-10, Eindhoven, The Netherlands. Short Paper, in press.

Information Capture Devices for Social Environments

Meghan Deutscher[1], Phillip Jeffrey[2], Nelson Siu[1]

[1]HCT Laboratory, University of British Columbia
Vancouver, British Columbia, Canada V6T 1Z4
{deutschm, nsiu}@ece.ubc.ca
[2]MAGIC, University of British Columbia
Vancouver, British Columbia, Canada V6T 1Z4
phillipj@ece.ubc.ca

Abstract. We put forth a proposal for a system that creates an augmented social environment where devices allow for an integrated capture of information about a social gathering (e.g. pictures, audio recordings, written thoughts and expression). Users can subsequently navigate around a collection of images, while listening to sounds and conversations recorded with the photo at the time of capture. We describe the devices used to capture this information and the user experience of the integrated *Memory Collage* application. The paper will also discuss the implications found for ambient intelligence within social environments.

1 Introduction

Our project's primary aim is the exploration of audiophotography and other methods of capturing information in a social environment. Audiophotography [1] is a domain that studies the value and practice of recording sound with still photographs. Motivated by the research of Frohlich and Tallyn we are trying to explore the full affordances of the audio capabilities found on many digital cameras today.

Although audio provides an additional dimension to memory recall on top of viewing photos alone [1], the ways in which sounds most relevant to the viewer are recorded has not been explored. Video recording can be argued as a method that incorporates all relevant sound and image together, but they are not selective in the information they capture, making indexed retrieval difficult at a later date. To ensure a high probability of capturing sound and images of importance, a system needs to determine the relevant events in its environment.

We are developing applications for the capture of events in one's social environment, and then presenting the audio enriched photos to the user in ways that evoke memories and emotions. This paper will detail our progress to date in augmenting a social gathering using microphones, cameras and sensors to obtain a more omnipresent view for sharing and personal viewing. We will then discuss the issues revealed throughout the project pertaining to ambient intelligence in social settings.

2 Related Work

The last few years have seen the development of a number of computational systems that capture information within their environment. Sumi, Matsuguchi, Ito, Fels, and Mase [2] developed a system, which integrates wearable and embedded sensors, video, and microphones, to observe interactions amongst multiple users. The resulting information is captured on an individual basis and does not allow for an omnipresent view of events. The Eavesdropper camera [3] automatically captures spontaneous moments in pictures when triggered by sounds such as laughter, voices, or noises. Frohlich and Tallyn [1] suggest that ambient "sounds-of-the-moment" convey a richer memory when paired with their corresponding photographs. We see an opportunity in developing a system that actively seeks to provide a novel means for recording important moments and multiple perspectives of a party.

3 Memory Capture and Display

An informal user study was conducted at a party organised by the authors to observe and record social interaction by participants within an instrumented environment. The location was a lounge on the campus of the University of British Columbia. The twenty-one participants consisted of professors, researchers, and graduate students. Questionnaires were distributed for initial feedback regarding their thoughts on the devices in the environment. We asked specifically about privacy issues and the obtrusiveness of the devices. Various sound recorders were placed in strategic locations to record ambient noise and coherent conversations near high traffic areas (e.g. foosball table, food table). Digital cameras carrying digital voice recorders (Fig. 1) captured snapshots of the event with the idea that recorded audio would likely correlate well with a photo. Touch sensors (Fig. 2) provided participants an opportunity to note significant moments through a time-stamping mechanism. These time-stamped events and photos served as a temporal index into the audio recordings so that meaningful audio clips could be played during photo viewing.

Fig.1 Digital camera with digital voice recorder mount.

Fig 2. Decorative event capturing touch sensor.

Using the information collected at the testing party, an application which we call a "Memory Collage" (Fig. 3) was created to elicit responses as to the usefulness of the

sound and photo combination in evoking emotional memories of an event. The Memory Collage application enables users to spatially manipulate images and arrange them in a personalised manner, unlike the structured organisation common to photo-sharing applications and browsers [4]. Users were able to create meaningful arrangements of their choosing, as well as play back a clip of sound recorded when a photo was taken.

Preliminary feedback indicates that this method of event capture and viewing allowed participants to recall situations they ordinarily would have had difficulty remembering. This phenomenon was particular apparent when an audio clip captured spontaneous remarks or comments that are normally quickly forgotten after an event. The Memory Collage demonstrated that there is a significantly stronger memory recall of events from audio-photographs.

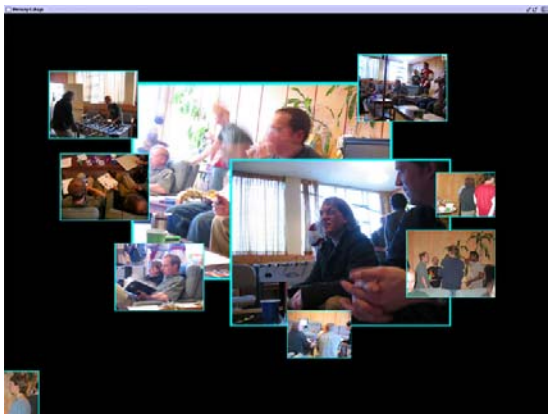

Fig. 3. Photos placed in personalised arrangement

4 Conclusions and Future Directions

Our project provides a first step into the research of information capturing devices in social environments, which we view as a new and relevant area of interest for ambient intelligence. The greatest challenge may involve using devices that rely on assumptions or fixed variables in a social environment. For example, we were not able to place recording devices in ideal locations, and we were unsure whether participants were using touch sensors for the intended purpose. These problems were encountered because of the inherent unpredictability of the environment. A system designed to be used in a shared, social space should be able to account for this unpredictability.

Our future directions are two-fold. First we wish to investigate more salient methods of capturing sound. Some participants found that sounds were not always correlated to the photos they interacted with, producing a degree of confusion. This may be because the affordance of the system suggests that sounds should be very closely correlated to the focus of the photo. A possible solution could be the use of a

Local Positioning System such as RemoteEyes [5] and equipping each participant with a personal microphone. Sound clips could then be chosen by locating who is in each photo. Secondly, we wish to develop a framework for the use of intelligent devices in a social environment. This framework should detail how to build a flexible system that can be fitted to a specific environment, depending on its variable and fixed attributes, such as the event space, the number of people, or the type of event.

We present these ideas to provide a catalyst to the research of ambient intelligence in social environments. By increasing these efforts, we hope to provide further insight into the relationship between people and their environment when interacting with intelligent systems.

Acknowledgements. We would like to thank Jussi Ängeslevä for his useful feedback during an earlier draft. We are grateful to the participants of our study for giving us their time and to the anonymous reviewers for their comments. We would also like to thank the Advanced Systems Institute (ASI) of BC for funding our project through an ASI communications award.

References

1. Frohlich, D., Tallyn, E.: Audiophotography: Practice and prospects. In: CHI 99 Extended Abstracts on Human Factors in Computing Systems, ACM, Pittsburgh, PA (1999) 296-297

2. Sumi, Y., Tetsuya, M., Ita, S., Fels, S., Mase, K.: Collaborative capturing of interactions by multiple sensors. In: Proceedings of the Fifth International Conference of Ubiquitous Computing (Ubicomp 2003) Interactive Poster, ACM, Seattle, WA (2003) 193-194

3. Martin, H., Gaver, B.: Beyond the snapshot: From speculation to prototypes in audiophotography. In: Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, ACM, New York, NY (2000) 55-65

4. Kuchinisky, A., Pering, C., Creech, M. L., Freeze, D., Serra, B., Gwizdka, J..: FotoFile: A Consumer Multimedia Organization and Retrieval System. In: Proceedings of Conference on Human Factors in Computing Systems (CHI 99), ACM, Pittsburgh, PA (1999) 496-503

5. Changsong, S., Wang, B., Vogt, F., Oldridge, S., Fels, S.: RemoteEyes: A remote low-cost position sensing infrastructure for ubiquitous computing. In: Proceedings of the First International Workshop on Networked Sensing Systems (INSS 2004). University of Tokyo, Japan (2004) 31-35