Designing Embodied Interfaces for Casual Sound Recording Devices

Ivan Poupyrev[†]

Interaction Lab, Sony CSL,

3-14-13 Higashigotanda,

Shinagawa, Tokyo

141-0022 Japan

ivan@csl.sony.co.jp

Haruo Oba,

Takuo Ikeda

Interaction Development Team,

Sony Creative Center,

1-7-1 Konan, Minato, Tokyo

108-0075 Japan

{oba, takuo}@dc.sony.co.jp

Eriko Iwabuchi

Ochanamizu University,

2-1-1 Otsuka,

Bunkyo, Tokyo

112-8610, Japan

[†] www.sonycsl.co.jp/person/poup/projects/

Abstract

In the *Special Moment* project we prototype and evaluate the design of interfaces for casual sound recording devices. These devices are envisioned to be used by a casual user to capture and store their everyday experiences in the form of "sound albums" – collections of recordings related to a certain situation. We formulate a number of design principles for such recording devices, as well as implement and evaluate two working prototypes. A *candle recorder* allows for capturing the general atmosphere at a party, and the *children's book recorder* records the interactions between parents and children while reading a book together.

Keywords

Sound recording, context-aware, tangible & ambient UI, embodied interaction, life-logging, aesthetic computing.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): User Interfaces: Input devices and strategies.

Introduction

Sound recording is one of the most important means for capturing, storing, and reproducing people's memories. There are many uses of sound recording: a student recording a lecture, a sound artist capturing envi-

Figure 1: The candle recorder used in the context of a party.

Figure 2: Children's book recorder is attached to a book to capture the interaction between mother and child.

ronmental sound for an art installation, a blogger recording a podcast and so on. However, all these examples are cases of either professional or enthusiastic hobbyist use. Sound recording devices are rarely used *casually* for "capturing a moment" just for the sake of creating a memento, as we often do when we take quick snapshots with a digital camera.

We present the *Special Moments* project that explores the design of recording devices for casual sound recording and playback. We envision small handheld sound recording devices that can be used by anyone to record and store sound snapshots. The device itself would form a *sound album*, providing an easy way for reproducing a recording without using a computer, home sound system, or any other external devices.

In this paper we propose interface design principles for such casual recording devices and describe implementation of two working prototypes: the *candle* and the *children's-book* sound recording devices (Figure 1 and Figure 2). We conclude the paper with informal evaluation that we carried out to assess the validity of our approach.

Related Work

The majority of sound recording devices, such as voice recorders, are general-purpose devices designed for use by professionals, i.e. news reporters, or sound enthusiasts. The value of recording sound in everyday lives is not yet clearly understood by casual users. Indeed, when Apple Computer researchers conducted a study on recording habits in 1992, they intentionally chose only those subjects that are comfortable with the use of audio data, so that researchers could avoid "getting caught in the question of why one might want to

use audio" [1]. Therefore, when designing recording devices for casual users the purpose and utility of such devices should be immediately evident to the user.

General-purpose recording devices also tend to make people uncomfortable and restrict their speaking behaviour [2]. The awareness of being recorded may change intonation and speech patterns similar to the way people's facial expressions and postures change when their picture is being taken. Therefore, we suggest that the design of casual recording devices should be unobtrusive, and it should fit in the particular context of where a recording is being made.

Several research projects investigated how sound recording can be integrated into everyday life. For example, *life-logging* devices continuously record audio-visual information [3, 4]: their goal is to capture and preserve every single event that we experience. Although our project is related, there are important differences. The underlying assumption of life-logging is that as much data must be captured as possible. However, browsing large volumes of data to find parts that are useful to the user can become difficult [5]. Browsing long audio recordings is especially time-consuming: special tools are usually required [6] which makes it less suitable for a casual user. To reduce this difficulty we designed *context-sensitive* devices so the recording takes place only when desired changes in the context are recognized. By using this approach we hope to reduce the amount of recorded data without sacrificing important information.

The context-dependent audio recording has been explored earlier in the context of note-taking [7, 8]. For example, Audio Notebook [8] links sound recording to

the notes written on a the computationally enhanced notebook. Our work was inspired by audio-enhanced note-taking, but we aimed to explore audio recording in a much wider range of situations. We also aimed at designing devices where recording does not have to be explicitly controlled by pressing a button [8] or selecting an item from a menu. Instead embodied and gestural interaction models [9, 10] are explored to seamlessly blend interaction into normal user practices.

Context-dependent Recording Devices

Special Moments is based on a number of principles:

C*ontext-dependent and context-aware*. In everyday life there are few situations where recorded sound is valuable to the end-user. We surveyed and analyzed a range of human activities at home, work, and outdoors to identify *contexts* where sound recording may present value for a casual user. The design of our devices is *limited* to these particular contexts. The recordings are triggered by the changes in these contexts.

Embodied interaction. The changes in the context can be identified by tracking the state of *physical objects* that users are operating. By embedding sensors into these objects we can track their state and trigger audio recording or index the audio data at the appropriate moment. The everyday objects then become interfaces to audio recording: instead of operating buttons and switches the user can simply use these everyday objects in the normal way. The recording therefore, can be controlled with little or no conscious effort, by using gestures and physical manipulation [9, 10] instead of dedicated interfaces. In a sense there is "no interface" in our devices.
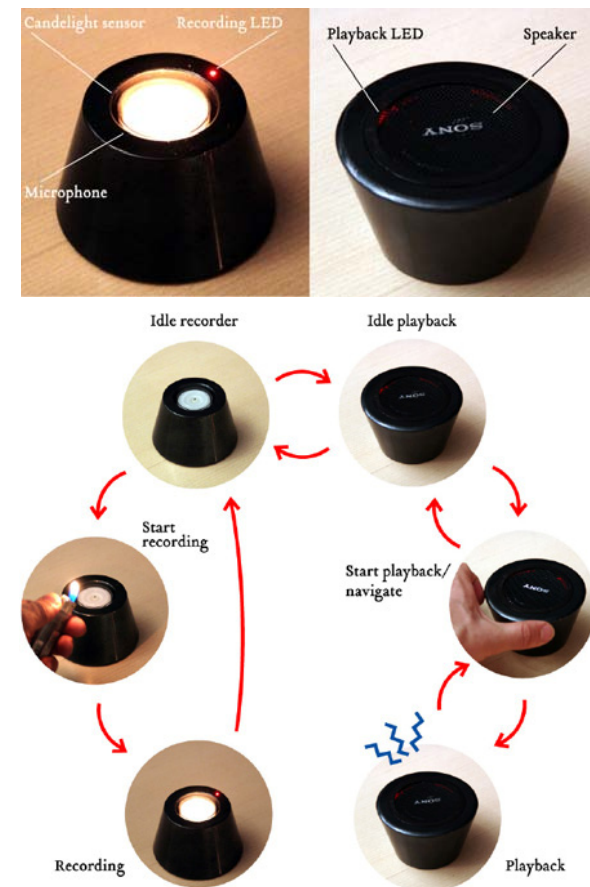

Figure 3: State transition diagram of the candle recorder

Self-contained sound albums. The device should have both recording and playback functionality to be enjoyed without using any other external devices. We envision users having multiple sound recording devices, and therefore, do not provide tools for deleting recordings.

We developed two recording devices based on these principles: the *candle* and the *children-book* recorders.

The candle recorder

The *candle recorder* can be used in the context of a party, dinner, or other similar occasion. Lighting a candle in such a context forms a natural trigger indicating a beginning of an event to start recording. Hence, the recorder is designed as a candle stand (Figure 1). The candlelight is sensed and tracked by an inexpensive infrared light sensor, with an LED lit up to indicate that a recording is taking place (Figure 3). When the user extinguishes the flame, the recording stops. To play back the recording, the user simply flips the candle stand over and the device becomes a small speaker (Figure 3). To start the play back the user twists the device, twisting it again will fast forward to the next recording and twisting it in the opposite direction will stop the play back.

The children's-book recorder

The second recording device designed was a *children's-book recorder* (Figure 2). It can be used by parents when they read a book together with their children. The device consists of two elements that are attached to a children's book and detect when the book is opened or closed.

Opening the book triggering the start of the recording; as parents interact with their child while reading the book, the device records both the parents and child's voices. Closing the book stops the recording (Figure 4). To listen to the recording, the user can plug headphones into the back of the device, which switches it into playback mode. In playback mode, opening the book starts playback of previously recorded data. To
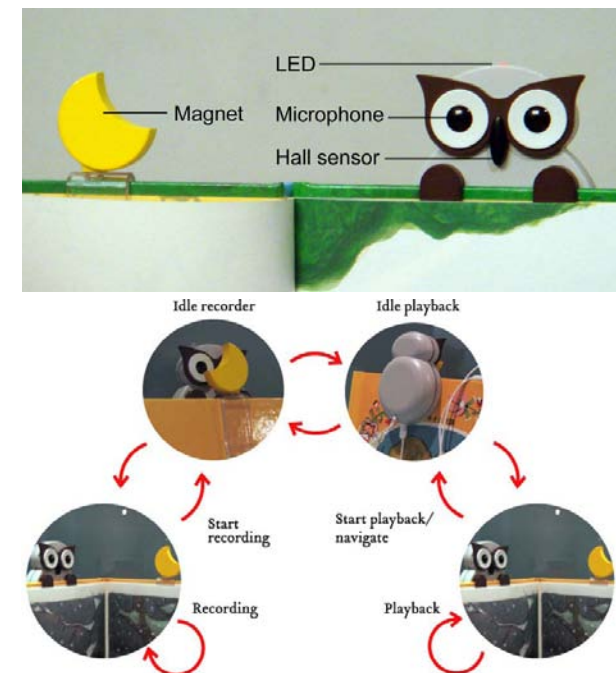

Figure 4: State-transition diagram of book recorder.

navigate through multiple recordings, the user can briefly close and open the book, this gesture is be interpreted as navigating to the next recording. Closing the book for a longer duration stops playback and resets the device to the first recording. We envisioned that the recorder will be used only with one book so that the child, once grown up, can re-experience reading the book with his or her parents.

Implementation notes

An MP3 player was disassembled and interfaced directly to the input/output ports of the 8bit RISC microprocessor Atmel Mega88. The sequences of commands were
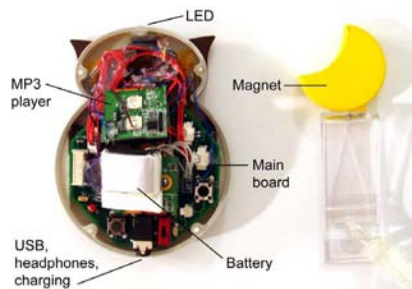
Figure 5: Implementation of book recording device

programmed to simulate MP3 player button presses so most functionality, such as playback and recording, could be controlled programmatically (Figure 5).

The light of a candle has a strong infrared component, allowing the candle recorder to use an inexpensive infrared photo transistor to detect if the candle is lit. An ADXL203 two-axis accelerometer is used to track the device orientation; an ADIS16100 gyroscope is used to track and interpret twisting gestures. We could not use a digital compass, a natural choice in such applications, due to magnetic field interference from the speaker.

The *book recorder* uses a Hall sensor for tracking the opening and closing of the book. The Hall sensor responds to the changes in magnetic field, therefore an external magnet is required. We mounted a Hall sensor inside the nose of the owl and the magnet is attached on the opposite side of the book. A soft metal plate was installed behind the Hall sensor as a magnetic field concentrator. It allowed us to increase sensor range from 10 to 30 millimeters, sufficient for most children's books. Figure 5 presents the configuration of the device.

Observations of device uses
The goal of our initial evaluation was to estimate the validity of our design assumptions and direct future development. We felt that it was too soon to conduct formal user studies and instead a series of field observations were carried out. The recording devices were taken to real dinner and home parties; and the book recorder was used to record conversations with children. In total 6 field studies were conducted, 21 participants between 25 and 45 years old were interviewed. Due to the specifics of the contexts (i.e. requests to fill a formal questionnaire at a dinner party were usually re-

jected), the interviews were informal but centered around two groups of questions:

1. Is the value of sound recording devices clearly understood? Does a strong relationship between the contexts and devices make their acceptance easier?
The strong context-dependence made the purpose and value of sound recording clear and unambiguous. In no situations were the participants wondering why anyone would use such devices. This is even more important because for the majority of participants the idea of recording sound as a memory snapshot was an unfamiliar idea. The concept of a book recorder was accepted with significantly more enthusiasm than the candle recorder. A typical reaction was "*... where can I buy it?*" One mother complimented the recording device and noted that children before five are changing extremely fast so there is a real feeling that everyday is special and everyday something is lost.

2. Was the embodied interface too simplistic? Did the user feel that they were in control? Was there any functionality missing?
In general no one had any difficulties with controls and most participants praised the interface's elegance and ease of use. In particular the combination of physical book and recorder was very warmly received: the interaction felt natural and unobtrusive. The participants did point out several problems. For one, sound browsing functionality was too limited with only skipping to the next recording provided. It seems that at least simple tools for browsing and time stamping are required.

When using the book recorder children would initially attempt to grab the device since it looked like a toy. During the field studies we realized that most of the

books for small children (i.e. under 5 years old) are short picture books and take just 10 to 20 min to complete. Initially, however, we incorrectly assumed that parents would read one book over a few days. Furthermore, small children often request the *same* book over and over again. Two of the participants took us off-guard with an unexpected suggestion: " *… can I record myself reading this book and then simply give it to my child so that I do not have to read the same book for the n-th time? …*"

A particular interest was a question of *privacy*. For the book recorder privacy was not a problem. However, for the candle recorder, many subjects noted that it might be seen as a "stealth" recording device and hence even more detrimental to privacy. We observed that *privacy* is intimately connected to *trust*: when the candle recorder was used strictly between the members of the family the problem of privacy was not raised and the candlelight recorder was accepted as an enjoyable addition to the event.

Discussion and Conclusions

The main contribution of this paper is exploration of context-dependent and embodied interfaces for media capture. As ubiquitous computing proliferates into the real world, more elements of our environment will be augmented with sensors. This will allow for the creation of new context-aware recording devices for capturing and preserving peoples' experiences.

Even though our evaluation was fairly informal, there are several conclusions we can make. First, designing context-dependent interfaces for media recording is a highly promising venue for further exploration. The study showed however that context has to be defined

more precisely, for example the age and relations between people have to be taken into account. Sound recording is still perceived as invasive; hence families, couples and close friends are perhaps a primary audience of such devices. An interesting venue for further exploration is to use recording devices to engage people in their activities. For example, can we make reading the same book by parents a less of a chore by providing a motivation: a memory to be treasured?

REFERENCES

1. Degen, L., R. Mander, and G. Salomon, *Working with Audio: Integrating Personal Tape Recorders and Desktop Computers*, in *CHI'92*. 1992. ACM. p. 413-418.
2. Campbell, N., *Towards Synthesizing Expressive speech: Designing and Collecting Expressive Speech Data*, in *EUROSPEECH* 2003. p. 1637-1640.
3. Mann, S., et al., *Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences.*, in *CHI'2005*. 2005. ACM.
4. Gemmell, J., et al., *Passive Capture and Ensuing Issues for a Personal Lifetime Store*, in *CARPE'2004*. 2004, ACM. p. 48-55.
5. Sellen, A., et al. *Do Life-Logging Technologies Support Memory for the Past? An Experimental Study Using SenseCam*. in *CHI'2007*. 2007: ACM. p. 81-90.
6. Arons, B. *SpeechSkimmer: Interactively Skimming Recorded Speech*. in *UIST '93*. 1993: ACM. p. 187-196
7. Whittaker, S., et al. *Filochat: handwritten notes provide access to recorded conversations*. in *CHI'94*. 1994: ACM. p. 271-277.
8. Stifelman, L., B. Arons, and C. Schmandt. *The Audio Notebook Paper and Pen Interaction with Structured Speech*. in *CHI'2001*. 2001: ACM. p. 182-189.
9. Ishii, H., A. Mazalek, and J. Lee, *Bottles as a Minimal Interface to Access Digital Information*, in *Extended Abstracts of CHI '01*. 2001, ACM. p. 187-188.
10. Harrison, B., et al. *Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces*. in *CHI'98*. 1998: ACM. p. 17-24.