

# In Search of Lost Sounds: Designing a Reminiscence Aid in Everyday Soundscape

Wenn-Chieh Tsai, Chiao-Yin Hsiao, Hung-Chi Lee, Ching-Hsiu Huang, Jane Yung-jen Hsu

National Taiwan University

Taipei, 10617 Taiwan

{d96944012, r99922012, d99922020, r99944023, yjhsu}@csie.ntu.edu.tw

## ABSTRACT

In this study, we re-introduced auditory sense in everyday practice of reminiscence. SoundTag, a technology probe, was used to explore the design space of using sounds as tags to annotate and retrieve event records in everyday life. Events with similar soundscapes could be matched by a *sound tag*. We selected timbre as the metric to determine the similarity between event soundscapes and implemented a proof-of-concept on smartphones. A one-week trial was conducted to evaluate the concept and profile the target user of using SoundTag in everyday reminiscence.

## Author Keywords

Soundscape, auditory sense, reminiscence.

## ACM Classification Keywords

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

## INTRODUCTION

Smartphones and their multimodal sensors play an important role in our everyday reminiscence. People can easily capture, store, access, and share digital mementos with their phones. However, an overwhelming amount of digital mementos could impose a heavy burden on the people to organize and retrieve specific mementos to recollect past experiences. Among all media, sound has lower information intensity than video and photos. According to [1], the lower density of information a colder medium contains, the more interpretation the viewer takes. Thus sound requires more participation from the viewers resulting a richer memory reconstruction. In addition, sound can facilitate richer perceptual recalls than text. It evokes strong emotions and emotions can enhance memory processes [2]. As photographs and diaries reflect from a third-person perspective, sounds, on the contrary, relive the experience in a first-person view [3].

In this work, we examined how our ‘secondary’ sensory system, the auditory sense, could make impact in our

personal daily practice of reminiscing. SoundTag is a technology probe that was deployed onto participants’ smartphones. It is a note-taking app with a catch to let users annotate events with any recorded sound clips and could retrieve similar events matched by the similarity of soundscape in these *sound tags*. A post-trial interview was conducted after each participant’s one-week field study. Thematic analysis was employed to derive insights into the design space of embodied human-sound interaction in the phases of autobiographical memory process.

## SOUNDTAG PROBE

SoundTag was designed to explore the varied uses of sound tag in participants’ daily life. Thus we installed the application on participants’ own mobile devices to reduce the intrusiveness of new equipment. We chose iPhone and iPod Touch 4G as our client-end devices. Users can collect and retrieve their digital mementos in two main modes, *My Memories* and *Recall by Sound*. A server is used to analyze sound tags by calculating their timbre similarity with other sound tags then sends the result back to the mobile devices. In the next few sections we will describe those two modes in more detail.

## My Memories

It is the place for storing memory cues. A list shows all event records sorted by timestamp. An event record could be composed of a photo, text descriptions, and sound tags. The length of a sound tag is limited up to 10 seconds. The limitation is arbitrary in the hope to capture the smallest self-contained particle of a soundscape [4]. We made the design so that event records could behave more like human memories which could be reconstructed while recollected. So users are encouraged to modify event records more often.

## Recall by Sound

In this mode, all sound tags are listed and grouped by event. Tapping on a sound tag will retrieve event records sorted by soundscape similarity in descending order. Users can then review the events by tapping on corresponding list items.

We also enable users to provide real time capture as a sound tag to do the search in situ. We call it *timbral listen*. In other words, users could record a sound clip to use it to retrieve past events if it happens to be something they came across before. They can redirect their phone to “listen.” The sound clip is formed either stopped by the users or reaching

the 10 seconds time limit, and then the smartphone will send the sound clip to the server and return the list back as soon as possible.

### SOUNDSCAPE SIMILARITY

To calculate the soundscape of a sound clip, we first started by dividing the whole audio signal into overlapped short-time frames based on the assumption that audio signals are more stationary in a short period of time. Time-domain features were extracted from the frames, and frequency-domain features were extracted from the spectrum generated by applying the short-time Fourier transform.

We chose timbre as the primary characteristic of a sound tag in our design for soundscape similarity. The intuition is that people can easily differentiate different sound sources by their distinct timbres. Timbre helps people conceptualize components in a soundscape and categorize the similar sound sources from different soundscapes. As a result, the *timbral features* defined by MARSYAS were picked as the primary factor in calculating the similarity [5]. The similarity is calculated from the Euclidean distance of two points of timbral features with each feature as one dimension in feature space of each sound sample.

### FIELD STUDY

We conducted a one-week field trial to explore the characteristics of our participants to understand how they would use the system. We aimed our target at younger adults who were regular smartphone users. Therefore, 10 users (4 male, 6 female) from different background were recruited. The users aged from 20 to 27 years old, with iPhone or iPod Touch 4G usage experiences ranged from 1 to 18 months.

Each participant was asked to collect 60 complete event records in their daily life and they are free to use both recall functions during the week. A complete event record was defined as an event record containing one photo, two text tags, and at least two sound tags. Users were suggested to upload their event records everyday to get the latest similarity ranking. For the purpose of exploring the design space, creative usage is encouraged.

### FINDINGS

#### Sound Tags on Memories

After a week of field trial, 578 sound tags were collected in total. The number was varied from person to person and ranged from the lowest of 9 to the highest number of 126. Each participant was creative in how they interact with the probe. Among all categories characterizing the clips participant collected, two groups were collected the most. They are the ambient sounds in background and human verbal sounds in foreground. Most sound tags were directly used to indicate real life objects in a soundscape.

We discovered the reason why participants captured human voice is not only it is the most evocative and memorable

type of sounds, but also the meaningful key feature they wanted to annotate on the events. The other major collection was the ambient sounds without a specific focus. Participants used this type of sounds to re-experience the event context. Furthermore, they intended to use these sound tags to represent where the event took place.

### Insights From Two Types of Users

The users who showed positive attitudes really enjoyed the field trial. In their daily life, sounds have already played an important role in reminiscence before using our design. The sounds they enjoyed help construct the whole atmosphere and enrich the memories. A common characteristic of this type of users was good storytelling skill. With even a simple question, they could share enriched stories with us.

On the other extreme end, another type of users did not get much surprise from the idea. Although some of them agreed the novelty of using sound to help reminiscence, they believed that sound tags are not convenient. For example, one user used only texts to record her daily life. She believed that all memories could be retrieved by simple texts. Other types of media can provide more details, but they are not necessary.

Despite the individual differences in user experiences, there were several common changes in the behavior among all users. They all mentioned that they became more sensitive to every slight sound, which had been omitted, in their daily life before.

### SUMMARY

We presented SoundTag, a reminiscence aid in daily soundscape. In this system, sound tags were used to annotate and retrieve event records based on their soundscape similarity. A proof-of-concept probe was implemented and deployed in a field trial. Users' usages and preferences were analyzed. Sound, one of our gradually forgotten sensory, has shown its potential to enrich memory tagging and influence on ambient soundscape awareness needs further exploration.

### ACKNOWLEDGEMENTS

This work was supported by the Taiwan National Science Council Grants # 99-2218-E-002-002.

### REFERENCES

1. McLuhan, M. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill (1964).
2. Jäncke, L. Music, memory and emotion. *Journal of Biology*, 7, 6 (2008), 21.
3. Bijsterveld, K., and Van Dijck, J. *Sound souvenirs: audio technologies, memory and cultural practices*. Amsterdam University Press (2009).
4. Schafer, R. M. *The Soundscape*. Destiny Books, 1993.
5. Tzanetakis, G. and Cook, P. MARSYAS: a framework for audio analysis. *Organised Sound*, 4 (2000), 169-175