6-2013

Demo: SocioPhone: Everyday face-to-face interaction monitoring platform using multi-phone sensor fusion

Youngki LEE
Singapore Management University, YOUNGKILEE@smu.edu.sg

Chulhong MIN

Chanyou HWANG

Jaeung LEE

Inseok HWANG

See next page for additional authors

DOI: https://doi.org/10.1145/2462456.2465702

Demo- SocioPhone: Everyday Face-To-Face Interaction Monitoring Platform Using Multi-Phone Sensor Fusion

Youngki Lee[1], Chulhong Min[2], Chanyou Hwang[2], Jaeung Lee[3], Inseok Hwang[2,4], Younghyun Ju[2], Chungkuk Yoo[2], Miri Moon[3], Uichin Lee[5], Junehwa Song[2]

[1]School of Information Systems, Singapore Management University
[2]Computer Science Department, KAIST, [3]Web Science and Technology Division, KAIST,
[4]Center for Mobile Software Platform, KAIST, [5]Knowledge Service Engineering Department, KAIST

[1]youngkilee@smu.edu.sg, {chulhong, chanyou, leejai, inseok, yhju, ckyoo, miri.moon, junesong}@nclab.kaist.ac.kr, [5]uclee@kaist.edu

Categories and Subject Descriptors

K.8 [**Personal Computing**]: General; C.3 [**Special-Purpose and Application-based Systems**]: Real-time and embedded systems

Keywords

Interaction; Conversation; Social; Platform; Volume Topography

1. INTRODUCTION

In this demo, we introduce SocioPhone, a novel initiative toward everyday face-to-face interaction monitoring platform. Among diverse verbal, aural, visual cues expressed during face-to-face interaction, SocioPhone captures diverse *meta-linguistic* information from conversations and provides *interaction-aware applications* on-the-fly. Undoubtedly, conversations are a key channel for face-to-face interaction. Specifically, monitoring conversational turns, i.e., alternation of different speakers (including none speaking), is the first crucial step to derive diverse interesting aspects of conversations, e.g., who is talking right now, how long and often one talks, how quickly one responds to another, and so on. In this demo, we will show the core technique of SocioPhone, *volume-topography*-based *turn monitoring,* which is highly accurate and energy-efficient under diverse real-life situations. In addition, we will demonstrate several example applications running on SocioPhone.

2. CONVERSATION TURN MONITORING

With SocioPhone, multiple smartphones work together to detect turn changes and associated speakers, along with a short in-situ training. Naturally placed phones belonging to a conversation group simultaneously sense a speaker's voice signals, but capture the signals with different strengths depending on their positions. Such relative sensory readings can be fused in realtime to form a *volume topography*, i.e., a signature vector of volume values sensed over different phones (See Figure 1). Our key observation is that such a topography is unique to each speaker, showing enough discrimination power to identify turns and associated speakers. With a short training period, e.g., 30-60 seconds at the beginning of a conversation, frequent turn-taking of speakers can be quickly and precisely traced through simple vector matching.

Our volume-topography-based technique has important advantages. First, volume parameters can be instantly and reliably estimated, even with a very short sensing window, e.g., 0.3 seconds; this enables to monitor dynamic turn-taking behavior in a highly agile way. Second, volume-topography is less susceptible to diverse environmental noises as it is built in-situ to reflect the current noise characteristics. Third, our approach is computationally much lighter than existing techniques for speaker recognition and diarization; it does not require complex signal processing such as MFCC extraction and GMM matching. Finally, we note that the method works well even at very low sampling rates (as low as 500 Hz), which has the potential to reduce users' privacy concerns.

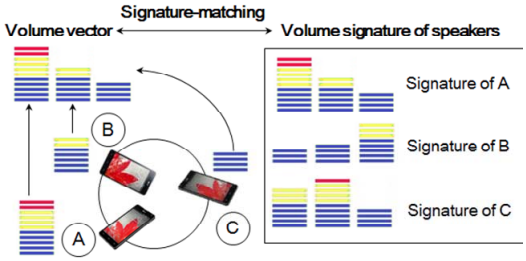
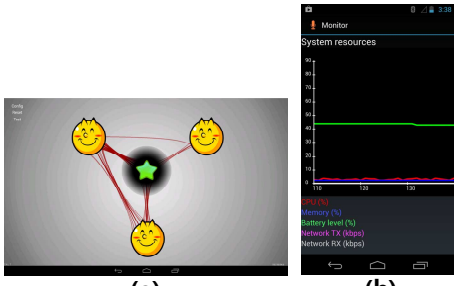


Figure 1 Illustration of in-situ turn monitoring



(a) (b)

Figure 2 Screenshots of applications

3. DEMONSTRATION

Turn Monitoring: We first demonstrate in-situ turn monitoring enabled by SocioPhone. We plan to setup a live visualization of the turn monitoring in the natural conversation between the guest and our staffs. Figure 2 (a) shows a screenshot of our conversation visualizer, denoting who are taking a turn from whom. In addition, we plan to demonstrate the resource feasibility of SocioPhone through our live resource monitor (see Figure 2 (b)).

Applications: We present a short video of everyday speech therapy for autistic children. This video delivers a concise digest of our field study at a local kindergarten with an autistic child. Following that, we show an initial deployment of our application designed to promote desirable conversation practices.

ACKNOWLEDGEMENT

This research was supported by the SW Computing R&D Program of KEIT(2012-10041313, UX-oriented Mobile SW Platform) funded by the Ministry of Knowledge Economy of Korea.

[1] This work was done while this author was at KAIST.

MobiSys'13, June 25–28, 2013, Taipei, Taiwan.
ACM 978-1-4503-1672-9/13/06.