Poster: Creation of a Co-Located Mobile-phone Users Group Using Voice

Satoshi Moriyama
Aoyama Gakuin University
5-10-1 Fuchinobe, Chuo-ku
Sagamihara, 252-5258, Japan
+81-42-759-6318
satoshi@rcl-aoyama.jp

Yuta Nishii
Aoyama Gakuin University
5-10-1 Fuchinobe, Chuo-ku
Sagamihara, 252-5258, Japan
+81-42-759-6318
nissy@rcl-aoyama.jp

Junji Takahashi
Aoyama Gakuin University
5-10-1 Fuchinobe, Chuo-ku
Sagamihara, 252-5258, Japan
+81-42-759-6318
takahashi@it.aoyama.ac.jp

Yoshito Tobe
Aoyama Gakuin University
5-10-1 Fuchinobe, Chuo-ku
Sagamihara, 252-5258, Japan
+81-42-759-6318
yoshito-tobe@rcl-aoyama.jp

ABSTRACT
We often share some files or information among co-located people using mobile phones in our daily lives. Conventionally, we explicitly need to choose users to be connected in the same space. Therefore, defining co-location is useful for many applications. Based on the background, we propose XOLOC, a system to create a co-located mobile-phone users using voice. Specifically, a person in the group utters some voice and if another device can capture the voice correctly, the device is determined to be in the same group.

Keywords
Co-location; Voice recognition; Mobile phone

1. INTRODUCTION
When we define "co-location" as present within a certain distance from the device of speaker, it might be wrongly determined that it is not in the same area even if it is actually in the same room, or that it is in the same area even if it is not actually in the same area. In addition, since people can connect from nearby rooms, it is difficult to define the same area using wireless signal. For these reasons, we defined "co-location" as the voice data, which the speaker's smartphone obtained, are the same as the voice data, which smartphones of the audience obtained.

Some researchers have investigated Japanese verbal acoustics. Miki et al. discriminated each verbal with the ratio of F1 to F2 and F2 to F3, and the relative value of the groove between F2 and F3 [1]. Urakami et al. recognized a speaker using formant of vowels [2]. In the work, they discriminated and matched speakers within a designed voice database of sole vowels. Our work utilizes the knowledge obtained from these previous works and applies it to define a group of co-located users using determining formants of voice.

MobiSys'16 Companion, June 25-30, 2016, Singapore, Singapore
ACM 978-1-4503-4416-6/16/06.
http://dx.doi.org/10.1145/2938559.2948772

2. Design of XOLOC
XOLOC can analyze WAVE files and output the results of matching lists. XOLOC has three functions: i) obtaining formant frequencies of the speech waveform, ii) matching each formant frequency, and iii) generating a list of audiences in the same area as the speaker. In ii), there are two steps to determine if each feature is approximate or not. The first step is focusing on differences of speaking among individuals, and the second step is the difference of vowels. Since there are some formant frequencies, such as F1, F2, or F3, we compare them. The formant frequencies of vowels are distributed over the contrast value each, and discriminate if these voices are the same with their features.

3. Poster and Video Clip
In our poster, the participants can see more details of the design. Additionally, the video of our application can be found at http://rcl.it.aoyama.ac.jp/XOLOC-demo.mp4

4. REFERENCES
[1] Yoshio Miki, Takuya Niikawa, Kanji Nohara, and Kentarou Okuno. 2006. Acoustic Identification of Japanese Vowels and its Application to Pathological Articulation. The Japan Journal of Logopedics and Phoniatrics, Vol. 47 (2006) No.2 P 155-165.

[2] Urakami Naoyuki, Shoji Yuta, Shiraishi Jun, Yamauchi Hironon, Fukumizu Yohei, and Izumi Tomonon. 2012. Speaker Recognition Using Formant of Vowels. IEICE Technical Report, SP2012-101, 2013.