Polite Ringer II: A Ringtone Interaction System Using Sensor Fusion

Ming-Chang Tsai
National Taiwan University
R98944021@ntu.edu.tw

Fu-Chiang Chou, Yih-Feng Kao, Kai-Cheng Yang
HTC Corporation
Fuchiang_Chou@htc.com, Calvin_Kao@.htc.com, abraham_yang@htc.com

Mike Chen
National Taiwan University
mikechen@csie.ntu.edu.tw

ABSTRACT
We present a system which automatically reduces the ringtone volume as soon as users start picking up their mobile devices to answer incoming phone calls. Our system uses multiple sensors already in mobile phones (e.g. accelerometer, gyroscope, and proximity sensor) to extract 47 features, and uses supported vector machine (SVM) as the classifier to identify the act of picking up a phone. We have collected data under several typical conditions: including walking vs stationary and picking up the phone from inside a bag vs pockets. Our results show that the system can correctly identify users picking up their mobile phones 95% of the time on average.

Author Keywords mobile devices, activity detection, sensor fusion, SVM.

ACM Classification Keywords H5.2. [User Interfaces]: User-centered design

General Terms Design, Experimental

INTRODUCTION
Ringtone, vibration and visual feedback have all been used to indicate incoming phone calls on mobile devices. Ringtone is useful in situations where users can not feel the vibration or does not have line of sight to the devices; however, its main drawback is that it disturbs people nearby. Users may feel embarrassed when they forget to put their phones into silent mode in meetings or in quiet places such as libraries. To silence a phone, users typically have to first pick up the devices, and silence it by pressing the volume switch.

As shown in **Figure1**, our system is designed to automatically reduce the ringtone volume when users start picking up their mobile devices to answer phone calls.

RELATED WORK
We had previously presented a similar system called *Polite Ringer* [1]. Since *Polite Ringer* depended solely on accelerometers for recognizing users' activities, it could work only for case when the mobile phone was picked up from a table or a nearly horizontal surface.

UbiComp'11, September 17–21, 2011, Beijing, China.
ACM 978-1-4503-0630-0/11/09.

Different to the previous work, our new system combines the fusion of multiple sensors with the supported vector machine for activities recognition. It is expected that the new system can work in more general cases in comparison to *Polite Ringer.*

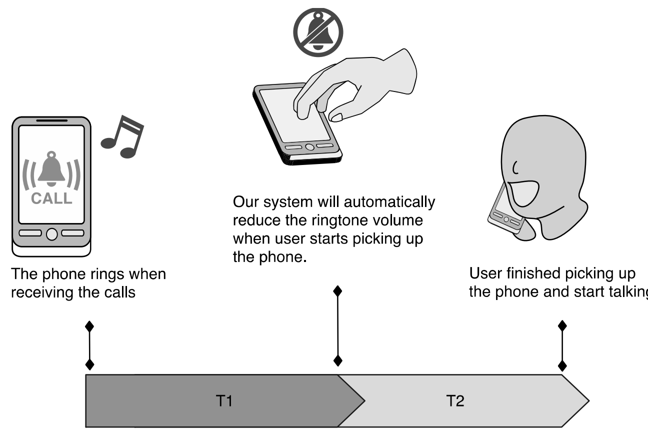

Figure 1. Our system will turn down or turn off the ringtone volume during the time interval T2

SYSTEM OVERVIEW
The workflow of the system, as shown in **Figure 2,** can be separated into three major steps: data acquisition, feature computation and SVM classification.

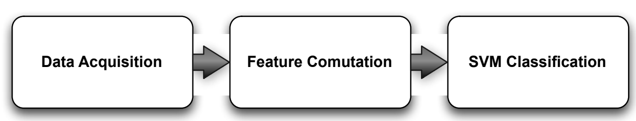

Figure 2. The workflow of our system

Data Acquisition
Once the phone receives the incoming phone call and starts ringing, the sensors will be triggered to continuously collect data from the environment. The sensor we select are listed as follow: accelerometer, gyroscope, compass, proximity sensor and the light sensor**.**

Feature Computation
Features will be computed on 1000ms windows with 500ms overlapping between consecutive ones from the original data. Means, standard deviations, and entropies on the frequency domain are used as the features for the classification. Our system uses FFT to retain information on

the frequency domain. Details about the features we select are also listed in **Table 1.**

SVM Classification

After the feature computation, each time window will be classified by the supported vector machine into the three different states, namely BEFORE, DURING and AFTER the phone picking activities. The transitions between the first and the second states will be identified as the time when user start picking up the phone, as shown in **Figure 3.**

Sensor Data	Type of Features
Accelerometer on the x, y, and z-direction	Mean, STD, power, entropy on frequency domain
Norm value of the accelerometer	Mean, STD, power, entropy on frequency domain
Accelerometer on the vertical direction to earth	Mean, STD
Gyroscope on the x, y, and z-direction	Mean, STD, power, entropy on frequency domain
Norm value of the Gyroscope	Mean, STD, power, entropy on frequency domain
Compass sensor on the x, y, and z-direction	Mean, STD
Norm value of the Compass sensor	Mean, STD
Proximity Sensor	Number of changes in the window
Light Sensor	Mean, Number of changes in the window

Table 1. The type of sensors and features

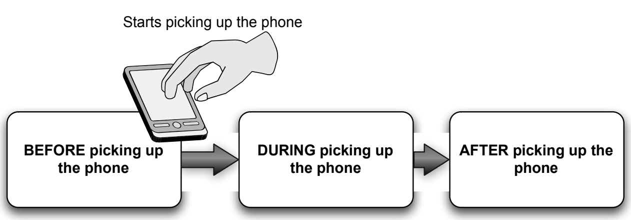

Figure 3. The SVM classifies the uses' activities into three different states. The time when users start picking up the phone will be identified as the transition between the first and the second states.

PRELIMINARY EXPERIMENTS

We had done a quick and dirty experiment to evaluate the performance. We had collected data from 6 participants, 5 male and 1 female, under 5 typical conditions. And we had evaluated our system based on this data set. The details of the testing conditions are listed in **Table 2**.

Data Collection

We used HTC's Sensation[2] smartphone as the device to collect data. During the data collection, we made phone calls to the device, and the participants were asked to pick up the phone about 10 seconds after hearing the ringtone. Each participant would perform the same procedure for 3 times under each testing condition. Thus, we collected 18 samples for each conditions, 90 samples in total. Volunteers were video-recorded throughout the whole session for later evaluation.

Evaluation and Results

Data under different conditions were evaluated separately as following.

For each turn, we selected samples of one participant as the testing set, while the rest would be used to train the SVM classifier of our system. Then we evaluated performance based on whether the system could correctly identify the participants picking up the phone from testing samples. We considered the system to be incorrect if it mistakenly recognized the event before the participants actually picked up the device.

The above procedure would be repeated until all samples had been considered as the testing data. The average accuracy of each condition is listed in **Table 2**.

User	The Device	Accuracy
Sitting	On the table	100%
Sitting	In the pocket	100%
Walking	In the pocket	83%
Walking	In the bag (the bag was carried by the user)	94%
Not specified	In the bag (the bag was on the table)	100%

Table 2. The details of the conditions and the evaluation results.

REFERENCES

1. An Empirical Investigation of the Effect of Understandability on Acceptance of a Sensor-Based Ringtone Interaction System. *Proceedings for the workshop: Using Audio and Haptics for Delivering Spatial Information via Mobile Devices, MobileHCI 2010* (2010), 6-9
2. HTC Sensation. http://www.htc.com/us/products/sensation4g-tmobile