Demo: Sword Fight With Smartphones

Zengbin Zhang†
Computer Science, UC Santa Barbara
Santa Barbara, CA
xiangying.meng@pku.edu.cn

Jian Qiu†
China Mobile Research Institute
Beijing, China
qiujian@chinamobile.com

David Chu
Microsoft Research
Redmond, WA
davidchu@microsoft.com

Thomas Moscibroda
Microsoft Research Asia
Beijing, China
moscitho@microsoft.com

Abstract

We present a demonstration of a phone-to-phone *Sword Fight!* game. It utilizes our solution for achieving high speed 3D continuous localization described in the accompanying conference paper [1]. The approach uses acoustic cues based on time-difference of arrival and power level. It assumes at least two microphones and one speaker per phone, which is common on new smartphones. Accelerometers and digital compasses assist in resolving ambiguous acoustic-only localization. Continuous localization is achieved with the aid of a loose time synchronization protocol and a Kalman filter. Lastly, practical gameplay issues are addressed.

Categories and Subject Descriptors

C.5.3 [**Computer System Implementation**]: Microcomputers—*Portable devices*

General Terms

Algorithms, Design, Performance

Keywords

acoustic localization, mobile gaming, smartphones

1 Introduction

Consider a real-time, face-to-face multiplayer *Sword Fight!* game, in which two players Alice and Bob are located in the same room, and use their phones as if they were lightsaber swords. Alice and Bob wave their phones; the orientation and location of one phone relative to the other simulates sword striking and blocking.

Such *high-speed, locational, phone-to-phone (HLPP) games* do not exist on commodity phones and enabling them seems particularly challenging. Existing solutions that provide real-time, locational functionality in commercial game consoles such as the Kinect or the Wii rely heavily on the existence of a fixed, external infrastructure such as a microphone array or cameras. In contrast, enabling HLPP games such as sword-fight (as well as similar non-gaming HLPP applications) would require implementing a similar functionality entirely on the two participating phones, without the help of any such external infrastructure.

The underlying technical challenge that has to be solved in order to enable these novel HLPP-type applications on smartphones is a new real-time 3D mobile device localization problem. While numerous variants of localization and positioning problems have been studied in the context of mobile phones, the one required for our purpose is particularly challenging, because the two devices need to establish a relative coordinate system without the use of additional infrastructure nor hardware modifications. Moreover, the coordinate system must be maintained in real-time, i.e., as the relative positions of the phones change, position estimates should be continuously revised with delay that is ideally imperceptible to human observers.

Our accompanying paper [1] shows that in principle, such fine-grained, real-time 3D localization is feasible on commodity smart phones. Our basic insight is to leverage the different sensors already commonly found on mobile devices. The core algorithm uses each device's multiple microphones to perform *acoustic 3D triangulation* and derive position estimates. Some smartphones such as Apple's iPhone 4 and Google's Nexus One already ship with two mics for video conferencing purposes, and more devices are continuing this trend. To accurately perform triangulation, we develop a new method that combines time-of-arrival (TOA) and signal power cues. In addition, each device uses its standard accelerometer and digital compass to assist triangulation by resolving ambiguous positions, and identifying *alignment regions* when two phones might be well-positioned for triangulation. Outside alignment regions, inertial displacement is used to estimate position. In order to enable each phone to track the other phone's movement, position estimates are continuously collected, and our algorithm employs a Kalman filter to smooth point samples and decrease measurement

†Work performed while at Microsoft Research

SenSys'11, November 1–4, 2011, Seattle, WA, USA.
ACM 978-1-4503-0718-5/11/11

variance. The filter accounts for the various forms of measurement errors inherent in the phone-based triangulation.

This demo shows two things. First, it shows an implementation of our new high-speed, phone-to-phone 3D localization algorithms on commodity devices. Second, this demo prototypes the game play of the 3D localization primitive as embodied by the *Sword Fight!* game. We explore issues related to localization latency, user perceived playability, error tolerance of acoustic ranging with in-the-field noise, and user tolerance of continuous acoustic localization.

2 3D Localization Design Overview

Our basic algorithm consists of *primary mode* supplemented with *fallback mode*. Primary mode consists of acoustic localization using the two mics already available on the phone. Fallback mode consists of IMU-based displacement tracking and is used when mics and speakers are turned away from each other. Our protocol in primary mode starts off by executing the *Initialization Stage*. The protocol then switches to continuously iterating through *Tone Exchange Stage*, *Distance and Angle Measurement Stage*, and *Position Estimation Stage*.

1. *Initialization Stage:* Two phones establish loose time synchronization by calculating WiFi round trip time and exchanging local clock values.
2. *Tone Exchange Stage:* Both phones turn on their microphones, and emit audio tones. Each phone records both its own audio tone and that of the other phone. Both phones also sample their accelerometer and geomagnetic sensors to obtain a rotation matrix.
3. *Distance and Angle Measurement Stage:* The recorded tones are first used to compute raw TOA cues. Then the two phones exchange their TOA cues to calculate distance values and local angles. The local angle measurement together and the rotation matrix obtained in the previous stage are transmitted to the other phone.
4. *Position Estimation Stage:* The distance measurement, two angle values and rotation matrices are fed to a Kalman filter to estimate the phone's relative 3D location coordinate.

When localization detects that the phones' relative positions are near the alignment region, we switch from primary mode to fallback mode. Fallback mode has two objectives: (1) to provide location estimates while the two phones are not in the alignment region, and (2) to detect when the two phones have re-entered each other's alignment region. Both of these objectives are accomplished by continuously calculating the displacement of each phone with accelerometer and compass readings. Fallback mode operates as follows.

1. *Enter Fallback:* Initialization establishes initial coordinates based on the last primary mode coordinates.
2. *Data Exchange:* Phones continuously exchange accelerometer and compass readings.
3. *Position Update:* Based on local and remote accelerometer and compass updates, each phone calculates its relative displacement, and updates its position estimate.
4. *Exit Fallback:* If it is detected that the phones are back within the alignment region, primary mode is restarted.

3 Gaming Considerations

We have designed Sword Fight to operate as follows. Each user holds his phone as if it were a sword grip. Possible game actions are striking, blocking and standing neutral. Striking consists of jabbing the phone toward the opponent with the length of the phone parallel to the ground plane. Blocking consists of holding the phone still with its length perpendicular to the ground plane. Standing neutral is the default position otherwise. Swinging and swiping attacks are currently not registered.

When an attacker makes a strike, it must be directed toward the other party in order to have a chance of registering points. However, the defending party may block the strike as long as the block is registered to be in the same plane as the attack. For example, if the attacker strikes the torso, then the defender ought to block at the torso level. The two parties take turns attempting strikes. A successful strike scores a point, whereas a blocked strike opens up an attacker for an unblockable retaliatory counterattack which the defender must make in a small time window. In addition, more powerful strikes as measured by accelerometer magnitude register higher points.

We plan to address several practical issues as we build our demo. First, while we have found our localization latency and accuracy to be acceptable in controlled measurement settings, it is not clear how well it will hold up to the demands of Sword Fight gameplay.

Second, players need a means to receive feedback on successful strikes and blocks. This is particularly challenging in our case because the phone screen is not readily viewable and the acoustic channel is already used for localization. We plan to investigate whether the acoustic channel has additional bandwidth for user feedback, potentially in frequencies not used by localization.

Third, the error tolerance of acoustic localization when humans are actively manipulating the phones is unclear. A player could grip the phone and thereby obscure the mic. A player could turn his body away from the opponent, again obscuring the acoustic signal's straight-line path. It may be possible to detect such conditions and switch to fallback mode if this does occur, or otherwise build gameplay mechanics to discourage such error modes.

Fourth, it is unclear how users will react to continuous acoustic localization, which is rather audible and unobscured. One option is to investigate ways of embedding our localization acoustic tones within game music.

Overall, our results indicate that by exploiting acoustic signaling mechanisms, two microphones on each phone, and various sensor information, it is possible to achieve real-time phone-to-phone 3D localization with an accuracy sufficient for many applications. Our gaming demo will attempt to be the first such application to take advantage of the 3D localization primitive.

4 References

[1] J. Qiu, D. Chu, X. Meng, and T. Moscibroda. On the feasibility of real-time phone-to-phone 3d localization. In *SenSys*, 2011.