

Whole-Home Gesture Recognition Using Wireless Signals (Demo)

Qifan Pu, Siyu Jiang, and Shyamnath Gollakota
University of Washington
{qp, jsy1218, gshyam}@uw.edu

ABSTRACT

This demo presents WiSee, a novel human-computer interaction system that leverages wireless networks (e.g., Wi-Fi), to enable sensing and recognition of human gestures and motion. Since wireless signals do not require line-of-sight and can traverse through walls, WiSee enables novel human-computer interfaces for remote device control and building automation. Further, it achieves this goal without requiring instrumentation of the human body with sensing devices. We integrate WiSee with applications and demonstrate how WiSee enables users to use gestures and control applications including music players and gaming systems. Specifically, our demo will allow SIGCOMM attendees to control a music player and a lighting control device using gestures.

CATEGORIES AND SUBJECT DESCRIPTORS

C.2.1 [Network Architecture and Design]: Wireless communication

KEYWORDS

Wireless; Gestures; User Interface

1. INTRODUCTION

As computing moves increasingly away from the desktop, there is a growing need for new ways to interact with computers. The Xbox Kinect is an example of a commercially available input sensor that enables gesture-based interaction using depth sensing and computer vision. The commercial success of these kinds of devices has spurred interest in developing new user interfaces that remove the need for a traditional keyboard and mouse. Gestures, for instance, enable a whole new set of interaction techniques for always-available computing embedded in the environment. For example, using a swipe hand motion in-air, a user could control the music volume while showering, or change the song playing on a music system installed in the living room while cooking, or turn up the thermostat while in bed. However, the burden of installation and cost make most vision-based sensing devices hard to deploy at scale. For example, throughout an entire home or building. Given these limitations, researchers have explored ways to move some of the sensing onto the body and reduce the need for environmental sensors [1, 4]. However, even on-body approaches are limited to

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
SIGCOMM'13, August 12-16, 2013, Hong Kong, China.
ACM 978-1-4503-2056-6/13/08.

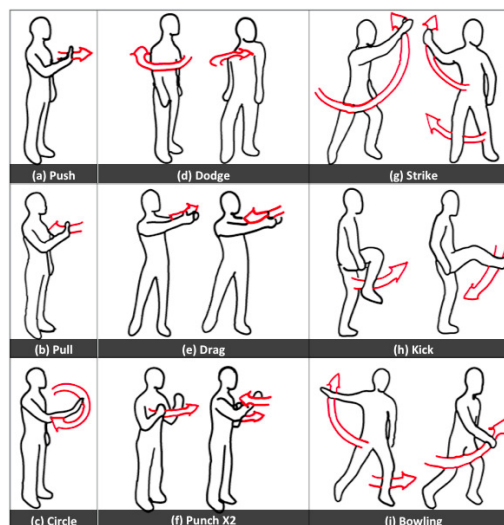


Figure 1—Gesture sketches: WiSee can detect and classify these nine gestures in line-of-sight, non-line-of-sight, and through-the-wall scenarios with an average accuracy of 94%.

what people are willing to constantly carry or wear, and may be infeasible in many scenarios (e.g., in a shower).

This demo presents NAME, the first whole-home human-computer interaction system that enables gestures and human motion tracking, without requiring either user instrumentation or an infrastructure of cameras/sensors. NAME achieves this by leveraging the Wi-Fi signals in an environment. Since these signals do not require line-of-sight and can traverse through walls, very few signal sources need to be present in the space (e.g., a Wi-Fi AP and a few mobile devices in the living room).

2. WISEE

To achieve this goal, we address the following challenges:

(a) *How do we extract human gesture and motion information from wireless signals?* WiSee leverages the property of Doppler shift, which is the frequency change of a wave as its source moves relative to the observer. The canonical example is the change in the pitch of a train's whistle as it approaches and departs from a listener. In the context of wireless signals, if we consider the multi-path reflections from the human body as waves from a source, then a human performing a gesture or moving, results in a pattern of Doppler shifts at the wireless receiver. Thus, a user moving her hand away from the receiver results in a negative Doppler shift, while moving the hand towards the receiver results in a positive Doppler shift.

However, human motion results in a very small Doppler shift that can be hard to detect from a typical wireless transmission (e.g.,

WiFi, WiMax, LTE, etc.). For instance, consider a user moving her hand towards the receiver at 0.5m/sec. This results in a Doppler shift of about 17 Hertz for a WiFi signal transmitted at 5 GHz. Since the bandwidth of WiFi's transmissions is at least 20 MHz, the resulting Doppler shift is orders of magnitude smaller than WiFi's bandwidth. Identifying such small Doppler shifts from these transmissions can be challenging.

WiSee presents a receiver design that can identify Doppler shifts at the resolution of a few Hertz from WiFi signals. The basic idea underlying WiSee is to transform the received WiFi signal into a narrowband pulse with the bandwidth of a few Hertz. The receiver then tracks the energy of this narrowband pulse to detect the small Doppler shifts. To do this, we design a *data-equalizing re-encoder* at the receiver that transforms each received symbol into the same symbol [3], by applying interference cancellation and then re-encoding the data corresponding to the first OFDM symbol. The solid line in Figure 2 shows the resulted narrow band signal, with most of its energy centered around its central pilot frequency. When a user moves her hand towards the receiver, part of this energy shift out and form a peak at the corresponding positive Doppler frequency, as shown by the dotted line.

(b) *How does WiSee map Doppler shifts to gestures?* WiSee leverages the fact that human gestures are typically continuous motions where different body parts move along paths that are constraint by the skeletal structure. For example, a punching gesture, requires the arms to accelerate in the forward direction, decelerate, and then finally stop momentarily before repeating the motion in the backward direction, which results in a unique Doppler pattern; WiSee classifies these patterns using a binary pattern matching algorithm. Specifically, WiSee can classify the nine gestures shown in Fig. 1 in line-of-sight, non-line-of-sight, and through-the-wall scenarios with an average accuracy of 94%. Further, since human motion such as walking results in higher Doppler energy than finer gesture motion, WiSee can also track human motion from behind a wall without requiring any devices on the other side of the wall.

(c) *How does WiSee work in the presence of other users?* A typical home may have multiple people who can affect the wireless signals at the same time. WiSee uses the MIMO capability that is inherent to 802.11n, to focus on gestures and motion from a particular user. MIMO provides throughput gains by enabling multiple transmitters to concurrently send packets to a MIMO receiver. If we consider the wireless reflections from each human as signals from a wireless transmitter, then they can be separated using a MIMO receiver [2].

Traditional MIMO decoding, however, relies on estimating the channel between the transmitter and receiver antennas. These channels are typically estimated by sending a distinct known preamble from each transmitter. Such a known signal structure is not available in our system since the human body reflects the same 802.11 transmitter's signals.

Our solution to this problem is inspired by the trigger approach taken by many multi-user games that use Xbox Kinect, in which a user gains control of the interface by performing a specific gesture pattern. In WiSee the target human performs a repetitive gesture, which we use as that person's preamble. A WiSee receiver leverages this preamble to estimate the MIMO channel that maximizes the energy of the reflections from the user. Once the receiver locks on to this channel, the user performs normal (non-repetitive) gestures that the receiver classifies using the Doppler shifts. A complete description of WiSee can be found in [5].

3. DEMO SETUP

Our demonstration setup is shown in Fig. 3. It includes a transmitting device and a WiSee receiver, both of which are imple-

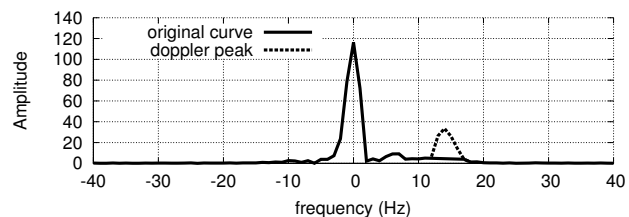


Figure 2—Extracting Doppler Shift: The figure plots the frequency profile of the narrowband signal centered at one of the sub-channels. In the presence of a gesture, a small portion of the energy shifts out of the central bulk and forms a Doppler peak.

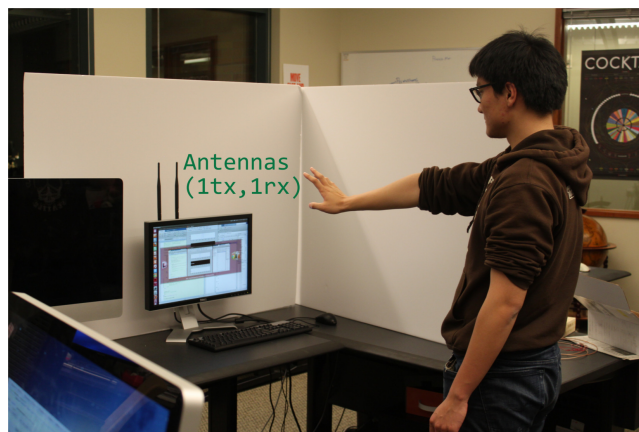


Figure 3—Demo Setup The system consists of an transmitting wireless device and a WiSee receiver. The user in this figure, controls the Linux window switcher application using WiSee.

mented in GnuRadio using the USRP N210 hardware. The transmitter periodically transmits packets and the WiSee receiver listens to these transmissions and detects human motions and classifies the gestures based on Doppler information. The gesture information will be sent to the higher layer applications to enable gaming and music control. Our demo focuses on demonstrating the ability of WiSee to detect and classify in-air gestures near our setup. Specifically, the proposed demo at SIGCOMM will integrate WiSee with a lighting control device and a music player. The users will be able to interact with these applications using in-air gestures.

4. REFERENCES

- [1] G. Cohn, D. Morris, S. Patel, and D. Tan. Humantenna: using the body as an antenna for real-time whole-body interaction. In CHI 2012.
- [2] S. Gollakota, F. Adib, D. Katabi, and S. Seshan. Clearing the RF Smog: Making 802.11 Robust to Cross-Technology Interference. In SIGCOMM, 2011.
- [3] S. Gollakota and D. Katabi. Zigzag decoding: combating hidden terminals in wireless networks. In ACM SIGCOMM, 2008.
- [4] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier. Digits: freehand 3d interactions anywhere using a wrist-worn gloveless sensor. In UIST 2012.
- [5] Q. Pu, S. Gupta, S. Gollakota, and S. Patel. Whole-home gesture recognition using wireless signals. In MOBICOM 2013.