WiFinger: Talk to Your Smart Devices with Finger-grained Gesture

Hong Li, Wei Yang*, **Jianxin Wang, Yang Xu, Liusheng Huang**
University of Science and Technology of China
qubit@ustc.edu.cn

ABSTRACT

In recent literatures, WiFi signals have been widely used to "sense" people's locations and activities. Researchers have exploited the characteristics of wireless signals to "hear" people's talk and "see" keystrokes by human users. Inspired by the excellent work of relevant scholars, we turn to explore the field of human-computer interaction using finger-grained gestures under WiFi environment. In this paper, we present Wi-Finger - the first solution using ubiquitous wireless signals to achieve number text input in WiFi devices. We implement a prototype of WiFinger on a commercial Wi-Fi infrastructure. Our scheme is based on the key intuition that while performing a certain gesture, the fingers of a user move in a unique formation and direction and thus generate a unique pattern in the time series of Channel State Information (CSI) values. WiFinger is deigned to recognize a set of finger-grained gestures, which are further used to realize continuous text input in off-the-shelf WiFi devices. As the results show, WiFinger achieves up to 90.4% average classification accuracy for recognizing 9 digits finger-grained gestures from American Sign Language (ASL), and its average accuracy for single individual number text input in desktop reaches 82.67% within 90 digits.

Author Keywords

Wireless; Micro-motion Recognition; Channel State Information.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces

INTRODUCTION

The pervasiveness of WiFi signals spurs a surge in relevant research, including motion detection, gesture recognition, localization. Multiple advances have been made both in terms of accuracy and granularity. In particular, recent literatures in keystroke recognition [1] and mouth motion [2] detection have demonstrated the viability of micro-movement detection with high accuracy using fine-grained radio reflections.

---

*Corresponding Author

The research community has studied various methods to classical gesture recognition, which can be classified into three categories: audio and radio based approaches, computer vision based approaches and wearable sensors based approaches. Audio and radio based approaches like SoundWave [3] and radio tomography technologies show a tendency of depending on large scale signal changes, they are not desirable for identifying micro-movement gestures. Vision based approaches have the fundamental limitations of requiring line of sight with enough lighting, and the infrared spectrum of sunlight often interferes with systems that rely on infrared, *e.g.*, LeapMotion [4] and Kinect [5]. Hence, such systems are not appropriate for outdoor use. Wearable sensors based approaches need wearable sensing devices that attach to the users' hand or body, which probably cause inconvenience in some occasions such as when a user is swimming or bathing, and it may incur extra cost.

Recently, studies "see" target movements by detecting and analyzing fine-grained radio reflection which overcomes the limitations of the above mentioned approaches. These Wi-Fi signal based human motion recognition systems, such as WiSee [6], Wi-Vi [7], WiKey [1] and WiHear [2], have been proposed based on the observation that different human motions cause different multi-path distortions in WiFi signals. WiSee uses Universal Software Radio Peripheral (USRP) to capture the Orthogonal Frequency Division Multiplexing (OFDM) signals and extract Doppler shift in signals reflected by movement of body or limbs to recognize nine gestures. Wi-Vi applies virtual multi-antenna technology to identify signal fluctuations caused by body movement. However, WiSee and WiVi focus on recognizing a set of coarse-grained gestures or motions, such as kicking, punching, stepping forward/backward. Additionally, they rely on sophisticated signals which are extracted from software radios, but not readily available in Commercial Off-The-Shelf (COTS) IEEE 802.11 devices. WiKey and WiHear employ COTS WiFi devices, one for continuously sending signals, and one for continuously receiving signals. They are both using in specific application scenarios such as hearing people's talk, detecting human's keystrokes. And both of them are not specialized in fine-grained gesture recognition. Melgarejo *et al.* proposed a fine-grained gestures recognition scheme which utilized sophisticated WARP v3 board equipped with two RE14P directional patch antennas, however, this scheme needs specialized devices to extract detailed RF signatures [8]. In our scheme, we focus on finer-grained micro-movement gestures recognition and achieve accurate text input on computer or mobile devices to realize human-computer interaction. Our scheme extracts Channel State Information (CSI) values from Intel

5300 (iwl5300) 802.11n Network Interface Card (NIC) by simply modifying the driver as in [9].

In this paper, we show the potential to identify finer-grained gestures by using feature information that is easily available on COTS wireless devices. The key insight is that while performing a certain finer-grained micro motion, the fingers of a user move in a unique pattern in the time series of CSI values which we call CSI waveform, for that gesture. In this paper, we target gestures of 9 digits finer-grained gestures from the standard ASL, as shown in Figure 1. Those gestures only involve finger movements and no hand or body movements. The unique finger movements of each gesture introduce relatively unique multipath distortions in WiFi signals and this uniqueness can be exploited to recognize gestures.

To translate the above high-level ideas into a working system entails a variety of technical challenges: (1) How does WiFinger detect and capture the subtle signal changes caused by finger movements of target user? (2) How does WiFinger analyze the tiny radio reflections and extract distinguishable features? (3) How does WiFinger compare shape features of any two finer-grained micro-movement gestures? (4) Will WiFinger be robust against unpredictable variances(*e.g.*, different users may perform the same gesture in slightly differently ways)?

This paper addresses the above challenges, and prototypes the WiFi signals based finer-grained gesture recognition system called WiFinger using two WiFi devices, a transmitter (*e.g.*, a router), and a receiver (*e.g.*, a laptop). The transmitter continuously emits signals to the receiver. The target user performs finger gestures between the transmitter and the receiver.

Our main contributions are summarized as follows:

- *Micro-movement analysis via physical layer CSI on off-the-shelf WLAN infrastructure in a non-invasive and device-free manner*: Our design of WiFinger provides the first solution using ubiquitous wireless signals to realize continuous text input in smart WiFi devices. WiFinger captures finger-grained gesture patterns from physical layer by modifying the driver of Intel 5300 802.11n NIC. Compared with Glove-Based input systems such as Sayre glove, Data glove, etc [10, 11]. WiFinger does not need users wear gloves or sensors, and meanwhile, it achieves continuous text input on smart WiFi devices.

- *Recognizing finger movements using radiometric characteristics of finger movements*: We address the challenge of identifying finger-grained gestures using subtle radiometric characteristics under WiFi environment. Our solutions are benefiting from the observation that the CSI series patterns vary among different finger-grained gestures, single CSI subcarrier is not enough to finger-grained gesture recognition. WiFinger adopts a novel method by averaging the extracted finger gesture profiles per 6 subcarriers, and concatenating them to form a distinguishable feature for gesture identification.

- *Accurate text input using WiFinger in WiFi devices*: To evaluate the effectiveness of WiFinger, 10 users are volunteered to provide training and testing data for our scheme.
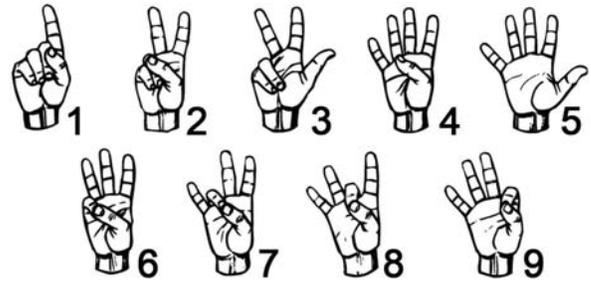

Figure 1. Part of finger gestures [12] that WiFinger can detect and recognize.

Users 2-10 perform 9 digits finger-grained gesture with 35 instances per gesture. To test the impact from the number of training set, we ask user 1 performs each finger-grained gesture 70 times. Finally, we get 3465 finger-grained gesture instances in our scheme. After that, we randomly generate 20 sequences of digits using matlab, there are 18 random numbers from 1 to 9 in each sequence. All users select 5 sequences to perform in WiFinger. As is demonstrated by the results, WiFinger achieves average recognition accuracy of 90.4% for each user, and its average accuracy of recognizing continuously number text input reaches 82.67%.

BACKGROUND

Prior work on gesture recognition can be generally categorized into two groups: device-based and device-free. Device-based gesture recognition techniques including audio or radio based [3, 13, 14, 15], vision based [5, 16, 17, 18] and sensors based [4, 19, 20, 21, 22]. They need extra hardware devices which might lead to high costs, and they are also constrained by the outside environment such as the ambient noise, light condition, etc. Device-free motion recognition techniques recognize human gestures using the variations of wireless signals. It does not need any additional hardware devices. Researchers have proposed various solutions which can be grouped into three categories:

RSS Based. Sigg *et al.* used USRPs which are specialized hardware devices to capture Received Signal Strength (RSS) values from WiFi signals [23, 24]. They utilized RSS values of WiFi signals to recognize four activities including lying down, crawling, standing and walking and achieved over 80% recognition accuracy for these four activities. Abdelnasser *et al.* proposed WiGest which leverages RSS values in WiFi signal strength to sense in-air hand gestures around the user's mobile device [25]. WiGest can achieve a classification accuracy of 96% for the application actions. According to the studies by former researchers, RSS values are not suitable for recognizing fine-grained motions such as gestures in standard ASL because RSS values only provide coarse-grained information about channel variations and do not contain fine-grained information about small scale fading and multi-path effects caused by micro-movements.

CSI Based. Recently CSI values obtained from COTS Wi-Fi NICs (such as Intel 5300 and Atheros 9390) are widely applied to recognize human activities [2, 26, 27, 28, 29, 30] and localize objects [31, 32, 33]. Han *et al.* proposed WiFall

that detects a single human activity of falling [26]. Zou *et al.* proposed Electronic Frog Eye that counts the number of people in a crowd using CSI values by treating the people reflecting the WiFi signals as "virtual antennas" [29]. Wang *et al.* proposed E-eyes that exploits CSI values for recognizing household activities such as washing dishes and taking a shower [27]. Zhou *et al.* proposed to use CSI to detect the presence of a person in an indoor environment [28]. Wang *et al.* proposed WiHear which employs specialized directional antennas to obtain CSI changes caused by lip movement for recognizing spoken words [2]. Ali *et al.* proposed WiKey that uses CSI values obtained from COTS to recognize keystrokes [1]. These splendid research specialized in specific application scenarios which do not contain continuous text input using CSI characteristics. Therefore, we propose a scheme which provides a solution to achieve text input using finger-grained gestures from standard ASL under WiFi environment.

Specialized Devices Based. Software Defined Radio (SDR) is a special purpose equipment that provides sophisticated signal features. Researchers have put forward various schemes that utilize SDRs to transmit and receive custom modulated signals for human activity recognition [6, 34, 35, 36]. Pu *et al.* proposed WiSee that uses a special purpose receiver design on USRPs to extract small Doppler shifts from OFDM WiFi transmissions to recognize human gestures [6]. Kellogg *et al.* presented AllSee that employs specialized analog envelop detector circuit for recognizing gestures within a distance of up to 2.5 feet using backscatter signals from RFID or TV transmissions [34]. Adib *et al.* designed WiTrack and WiTrack2.0 that apply specially designed frequency modulated carrier wave radio frontend to track human movements behind a wall [36]. Recently, Chen *et al.* built an SDR based custom receiver design to track keystrokes using wireless signals [37]. Melgarejo *et al.* proposed a fine-grained gesture recognition scheme using sophisticated WARP v3 board equipped with two RE14P directional patch antennas [8]. In contrast to all these schemes above, WiFinger does not need SDR or any special purpose devices, we utilize COTS WiFi NICs to recognize fine-grained finger gestures.

CHANNEL STATE INFORMATION

In wireless communications, channel state information refers to known channel properties of a communication link. It depicts how a signal propagates from the transmitter to the receiver and represents the combined effect of scattering, fading and power decay with distance. Modern WiFi devices that support IEEE 802.11n/ac standards typically consist of multiple transmitting (TX) and receiving (RX) antennas and thus support Multiple-Input Multiple-Output (MIMO). In OFDM system, the channel between each pair of TX and RX comprises of multiple subcarriers and the narrow band flat-fading channel with multiple TX-RX antennas is modeled as

$$\mathbf{y} = \mathbf{H} \times \mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is the received vector, $\mathbf{x}$ is the transmitted vector, $\mathbf{n}$ is the noise vector and $\mathbf{H}$ is the complex valued Channel Frequency Response (CFR). Let N_t and N_r denote the number of TX and RX antennas respectively, and N_c is the number
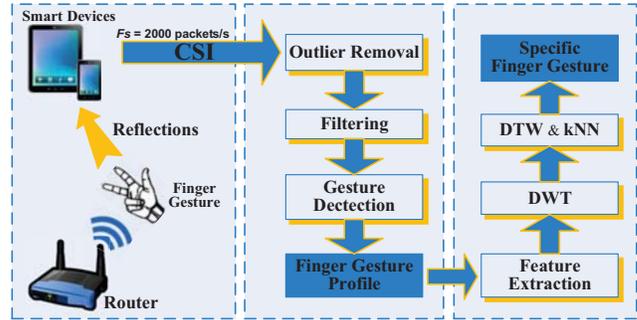

Figure 2. Framework of WiFinger

of subcarriers, then $\mathbf{H}$ is a $N_c \times N_t \times N_r$ dimensional matrix. Noise is often modeled as circular symmetric complex normal with $\mathbf{n} \sim \mathbf{c} \cdot \mathbf{N}(0, \mathbf{S})$. CSI is an estimation of $\mathbf{H}$. As the received signals are the resultant of constructive and destructive interference of several multipath signals reflected from walls or other surround objects. While performing a gesture near the receiver, hands or fingers can affect the propagation of multipath signals, which leads to the variations in amplitude and phase of CSI values among all subcarriers of every TX-RX antenna pair and these variations can be used to identify fine-grained gestures. The tool in [9] is employed to report CSI values of 30 OFDM subcarriers, thus the measured $\mathbf{H}$ is a matrix with $30 \times N_t \times N_r$ dimension. In addition, we consider the CSI values captured from each TX-RX pair as one *CSI stream*. Since CSI provides a finer-grained representation of the wireless link compared with RSS, recent researchers tend to utilize CSI rather than RSS.

WIFINGER OVERVIEW

WiFinger is a wireless system that utilizes commercial WiFi devices to achieve human-computer interaction by recognizing people's finger-grained gestures. Figure 2 illustrates the framework of WiFinger, which can be divided into 3 parts.

For the first part, a transmitter and a receiver comprise the CSI sampling unit. The transmitter is a wireless Access Point (AP), *e.g.*, a router or a laptop serving as a hot spot, and the receiver can be a smart device, *e.g.*, a laptop or a mobile phone. In WiFinger, the transmitter has 1 directional antenna and the receiver is equiped with 3 (other than 1) omni-directional antennas to achieve higher Signal Noise Ratio (SNR) as well as higher data rate [38].

While the system is working, the transmitter continuously sends wireless packets to the receiver. These packets were further sampled at the receiver side to extract CSI values. To ensure finger-grained motions recognition, we set the sampling rate to be $F_s = 2000 \ packets/s$. From each sampled packet, a CSI matrix $\mathbf{H}$ is extracted, which has a size of $30 \times 1 \times 3$, as mentioned in the previous section. Although we can obtain 1×3 CSI streams at the same time, only the CSI stream 1-1 is selected, thus $\mathbf{H}$ can be simplified to a 30×1 dimensional matrix. The impact of multiple streams will be covered in the Discussion and Limitation section. The selected stream $\mathbf{C}$ contains a series of CSI matrices and is represented as $\mathbf{C} = [\mathbf{H}_{t_1}, \mathbf{H}_{t_2}, \cdots, \mathbf{H}_{t_L}] = [\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_{30}]^T$

with size $30 \times L$, where $\mathbf{c}_i(i = 1, 2, ..., 30)$ represents different subcarriers.

In the second part, WiFinger minimizes the interference of environment noise by applying various de-noising methods to the original CSI stream. Then we use a novel gesture detection method to eliminate the "silent" part of CSI stream, and arrange gestures as separated finger gesture profiles.

The third part serves as a gesture recognizer. From each gesture profile, a special designed feature vector is extracted. The recognition of gestures is achieved by applying k-Nearest-Neighbours (kNN) algorithm, where the distance between two feature vectors is measured with Dynamic Time Warping (DTW). Besides, we compress the feature vectors with Discrete Wavelet Transformation (DWT) to reduce the computational complexity of DTW.

SIGNAL PREPROCESSING

The original CSI stream extracted from commodity WiFi NICs are inherently noisy, and its subcarriers usually have different noise levels, thus we process the CSI values of different subcarriers independently. To refine the CSI stream for finer-grained gestures recognition, WiFinger firstly removes the impracticable outliers of each subcarrier, then passes them through a low-pass filter to remove the high-frequency noise. Although strict low-pass filter can obviously remove noise, the noise within passband cannot be eliminated thoroughly. Thus, further process is necessary. We show that a weighed moving average method reduces these in-band noises efficiently.

Outlier Removal

Because of the internal state transition such as transmission power changes and transmission rate adaptation, there are some burst noises in the selected CSI stream. Figure 3(a) shows the CSI waveform of two consecutive finger gestures of No.1. It can be observed that there are some abrupt fluctuations in the waveform. Obviously, these outliers are not induced by finger movements. Given that outliers probably affect finger motion detection, WiFinger utilizes Hampel identifier [39] to eliminate these outliers. It declares all points out of the interval $[\mu - \gamma \times \sigma, \mu + \gamma \times \sigma]$ as outliers, where μ and σ are the median and median absolute deviation of the data sequence, γ varies in different situations and the most widely used value is 3. Figure 3(b) plots the CSI waveform after removing outliers.

Low-pass Filtering

In WiFinger, the speed of human finger gestures is low, and the signal changes caused by finger motions lie at the low end of the frequency spectrum while the noise induced by hardware imperfection, *e.g.*, carrier frequency offset, and channel propagation, *e.g.*, shadowing, has a relatively high frequency. Butterworth low-pass filter is a natural choice which has maximum flat frequency response in the pass band and roll off towards zero in the stop band, hence ensures the fidelity of signals in target frequency range and removing out-band noise greatly. Therefore, we apply the filter on the received samples to eliminate the out-of-band interference. We experimentally observe that the frequency of the variations in CSI
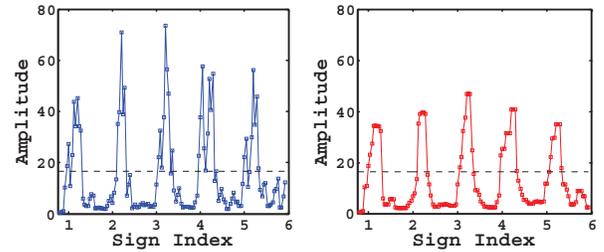
(a) Before 5-point median filtering (b) After 5-point median filtering

Figure 4. Finger gestures detected by sign indicator.

stream due to finger gestures is often within $1 \sim 60$ Hz. We set the output of the filter with a cut-off frequency of 60 Hz when applied to the CSI values in Figure 3(b), with sampling rate $F_s = 2000 \; packets/s$. The filtered waveform is shown in Figure 3(c). Obviously, Butterworth low-pass filter successfully removes most high frequency noise. However, the noise cannot be completely eliminated, hence, we have to do further processing towards the CSI stream.

Weighted Moving Average

Figure 3(c) shows the CSI stream of two consecutive gestures, which is extracted from the third subcarrier of finger gesture No.1. It can be observed that the CSI waveform is still noisy for gesture extraction. WiFinger further introduces weighted moving average method to the filtered CSI stream. Specifically, as for the CSI values of the 3th subcarrier $\mathbf{c}_3 = [h_{t_1}, h_{t_2}, \cdots, h_{t_L}]$, the CSI value at time t_i is averaged by previous m values. The latest CSI value has the highest weight m, thus

$$h_{t_i} = h'_{t_i} = \frac{1}{m + (m-1) + \cdots + 1} \cdot$$
$$[m \cdot h_{t_i} + (m-1) \cdot h_{t_{i-1}} + \ldots + h_{t_{i-m+1}}], \quad (2)$$

where h'_{t_i} is the averaged new value at time t_i. The value of m decides in what degree the current value is related to the historical records [26]. Figure 3(d) shows the weighted moving average waveform. We empirically set $m = 30$ in WiFinger to achieve preferable denoising results. In the following, the processed CSI stream is utilized to extract gestures.

GESTURE EXTRACTION

We show how WiFinger extracts the CSI stream for each finger gestures. The main effects of human finger motions on the received CSI stream are either rising edges, falling edges, or pausing. These variations are critical for detecting finger gestures and the uniqueness of different variation patterns are exploited to classify finger motions. To detect the starting and finishing points of a gesture, we split the processed CSI stream into segments.

Gesture Detection

To detect the finger gestures from the CSI stream, we refer to [40] and design an adaptive method to detect the starting and finishing points of finger gestures. We explain the steps to get *sign indicator*, and then, we dynamically get the
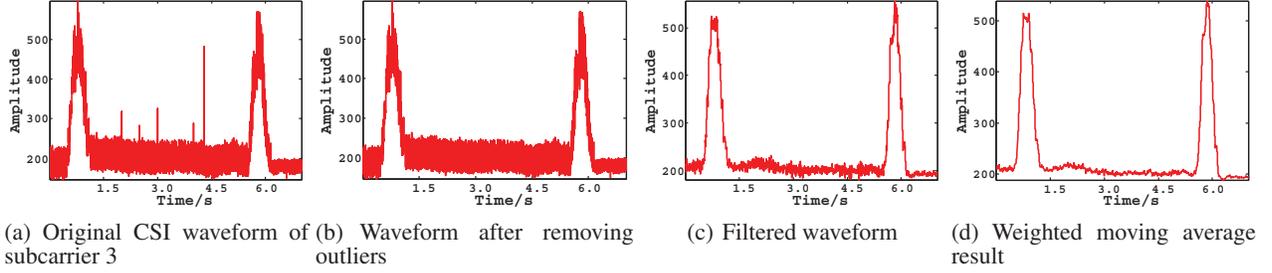
(a) Original CSI waveform of subcarrier 3

(b) Waveform after removing outliers

(c) Filtered waveform

(d) Weighted moving average result

Figure 3. The procedure of signal preprocessing.

threshold from *sign indicators* which is further applied to extract finger gestures. The extraction process is arranged as the following four steps.

1. *Preprocessing*: Firstly, WiFinger removes the Direct Current (DC) component of every subcarrier by subtracting the corresponding constant offsets from the CSI stream. The constant offsets can be calculated through a long-term averaging over different subcarriers. Then WiFinger cuts the CSI stream $\mathbf{C}$ into bins with a sliding window. The window size is 500 and the step length is set as 400. Each bin is a matrix of size 30×500, represented as $\mathbf{M}$.

2. *Correlation estimation*: WiFinger calculates the correlation matrix as $\mathbf{M}^T \times \mathbf{M}$. The dimension of the correlation matrix equals the number of CSI subcarriers extracted from Intel 5300 NIC.

3. *Sign indicator calculation*: WiFinger calculates the eigenvectors and eigenvalues of the correlation matrix base on a key observation. Namely, when there is no gesture or motion in the experimental environment, the second eigenvector $\mathbf{q}_2$ varies randomly over neighboring subcarriers. However, when a finger gesture happens, the CSI subcarriers become correlated, $\mathbf{q}_2$ varies smoothly and its mean of first difference $\delta_{\mathbf{q}_2} = \frac{1}{N_c - 1} \sum_{l=2}^{N_c} |\mathbf{q}_2(l) - \mathbf{q}_2(l-1)|$ becomes smaller, where N_c represents the number of CSI subcarriers and $|\mathbf{q}_2(l) - \mathbf{q}_2(l-1)|$ represents the difference in coefficients for neighboring subcarriers. Besides, the occurrence of gestures results in a higher variance $\mathbb{E}\{\mathbf{h}_2^2\}$ of the principal component $\mathbf{h}_2$. Thus we define $\mathbb{E}\{\mathbf{h}_2^2\}/\delta_{\mathbf{q}_2}$, as a *sign indicator* to indicate the occurrence of finger motions.

4. *Smoothing*: As is shown in Figure 4(a), there are some abrupt changes in the waveform of *sign indicator* which are likely to lead to fallacious results of detection. Therefore, we apply a 5-point median filter to the set of *sign indicator* to facilitate the detection of finger motions.

Finger Gesture Profile Extraction

Considering the characteristics of *sign indicator* values calculated above, we empirically select the value of the third quartile of the *sign indicator* values as a threshold to automatically detect and extract finger gestures. The threshold can be dynamically adjusted in accordance with variation of CSI values in real time. Figure 4(a) and 4(b) plot the

sign indicator before and after applying 5-point median filter. We observe that 5-point median filter facilitates finger gesture extraction in WiFinger.

In the process of gesture extraction, WiFinger sets a guard interval T^b on both side of the estimated finger gesture profile. The algorithm compares the values in the set of *sign indicator* with threshold to obtain their intersections as the candidates of starting or finishing points of finger gestures. For example, says the candidate set is $\{t_1^s, t_1^e, t_2^s, t_2^e, \cdots t_n^s, t_n^e\}$. Then the starting points should minus T^b and the finishing points plus T^b. The purpose of guard interval is to better obtain the complete finger gesture profiles. Hence the candidate set turns into $\{t_1^s - T^b, t_1^e + T^b, t_2^s - T^b, t_2^e + T^b, \cdots t_n^s - T^b, t_n^e + T^b\}$. The CSI values between each pair of starting and finishing points are mentioned as finger gesture profiles, represented as $\mathbf{P}_i = [\mathbf{H}_{t_i^s - T^b}, \cdots \mathbf{H}_{t_i^e + T^b}](i = 1, 2, \cdots n)$. $\mathbf{P}_i$ has a size of $N_c \times l$, where $N_c = 30$ represents the amount of CSI subcarriers, and $l = t_i^e - t_i^s + 2 \cdot T^b$ represents gesture duration. The profiles of different subcarriers is represented as $\mathbf{p}_i (i = 1, 2, \cdots n)$.

FEATURE EXTRACTION

After finger gesture profile extraction, WiFinger obtains motion profiles for different finger gestures. Figure 5(a) to 5(c) show the CSI values variation in 30 subcarriers of digits 1, 4, 7 in ASL. According to characteristics depicted by these time series patterns of different finger gestures, we find that simply utilizing a single subcarrier or several subcarriers cannot fully represent the distinguishable features of each finger motion. The finger motion features distribute in all 30 subcarriers, WiFinger must integrate those distributed features to identify different finger motions. Hence WiFinger combines 30 subcarriers by averaging every 6 subcarriers, and then WiFinger concatenates them to form a synthetic waveform, mentioned as a feature vector $\mathbf{F} = [\mathbf{f}_{1\sim6}, \mathbf{f}_{7\sim12}, \mathbf{f}_{13\sim18}, \mathbf{f}_{19\sim24}, \mathbf{f}_{25\sim30}]$, where $\mathbf{f}_{k\sim l} = \frac{1}{l-k+1} \sum_{i=k}^{l} \mathbf{p}_i$. Figure 5(d) to Figure 5(f) show the compressed synthetic feature vector for finger gesture No.1, No.4 and No.7, respectively. The component in every colored bar of each figure is the averaged result of every 6 subcarriers. The details of compressing synthetic feature vector will be addressed next.

CLASSIFICATION

Considering the feature vectors of different finger gestures contain thousands of CSI values, directly using the synthetic waveforms as the input of classification algorithm leads to
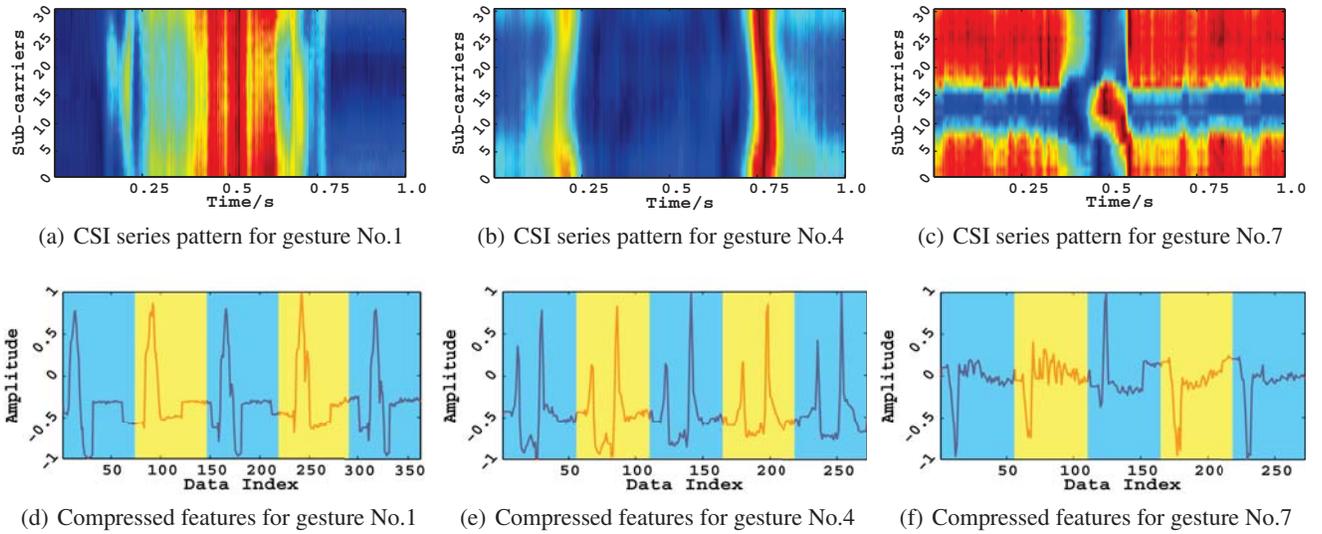
(a) CSI series pattern for gesture No.1   (b) CSI series pattern for gesture No.4   (c) CSI series pattern for gesture No.7

(d) Compressed features for gesture No.1   (e) Compressed features for gesture No.4   (f) Compressed features for gesture No.7

Figure 5. The features of different finger gestures.

high computational costs. WiFinger compresses the feature vectors by utilizing DWT. The advantage of DWT are two-folds: 1) It compresses the original signal while preserving both time and frequency domain information. This attribute facilitates the signal analysis and reduces the computational cost in the process of classification. 2) In WiFinger, the finger gestures are finer-grained movement, DWT realizes multi-scale analysis towards finger gesture profiles. Owing to the lengths of finger gesture features are not equal, Euclidean distance is not suitable in this scenario. To distinguish the minute difference among finger movements, WiFinger exploits DTW to calculate the distance between two finger gesture features.

Discrete Wavelet Transformation
The classical Fourier transform and short-time Fourier transform do not provide multiple resolution in time and frequency, which is an important characteristic for analyzing transient signals containing both high and low frequency components. The Discrete Wavelet Transformation overcomes those limitations by employing analyzing functions that are local both in time and frequency. The DWT is defined by the following equation:

$$W(j,k) = \sum_j \sum_k x(k) 2^{-j/2} \psi(2^{-j}n - k) \qquad (3)$$

where $\psi(t)$ is a time function with finite energy and fast decay called the mother wavelet. The DWT analysis can be performed using a fast, pyramidal algorithm related to multirate filterbanks [41]. In the pyramidal algorithm the signal is analyzed at different frequency bands with different resolution by decomposing the signal into a coarse approximation coefficients and detail coefficients. The coarse approximation coefficients is then further decomposed using the same wavelet decomposition step. The complete binary tree in the decomposition process can be shown in Figure 6.

The efficacy of wavelet transformation depends on selecting appropriate wavelet basis. In WiFinger, several wavelet families have been tested such as Daubechies, Coiflets, Symlets. Due to the classification performance, the Daubechies D4 coefficient wavelet family is selected.

Dynamic Time Warping
WiFinger utilizes kNN classifier to recognize different finger gestures. Traditional kNN classifier adopts Euclidean distance as the measurement criteria between samples. Considering the feature vectors of gestures might not share the same length, we use DTW to calculate the distances among features. In contrast to Euclidean distance, DTW provides intuitive distance between two waveform and can be resilient to signal distortion or shift. DTW distance is the Euclidean distance of the optimal warping path between two waveforms calculated under boundary conditions and local path constraints [42]. The objective of DTW is to compare two (time-dependent) series $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ of length $n \in \mathbf{N}^+$ and $\mathbf{Y} = (y_1, y_2, \cdots, y_m)$ of length $m \in \mathbf{N}^+$. These sequences may be discrete signals (time series) or, more generally, feature sequences sampled at equidistant points in time. DTW can handle waveforms with different lengths and calculates distance by recovering optimal alignments between the two time series. WiFinger feeds the finger gesture features to kNN classifier and obtains a decision. The kNN classifier uses DTW to calculate the distance between finger gesture
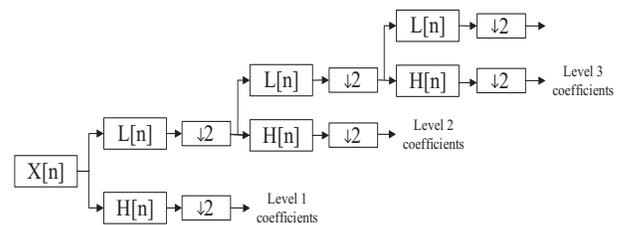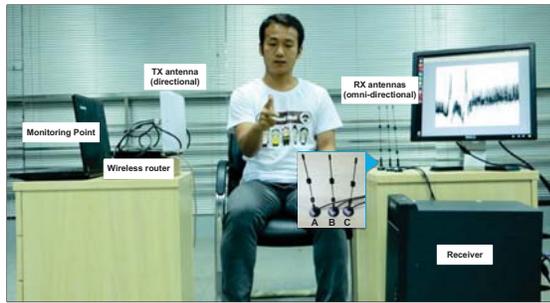

Figure 6. Discrete wavelet transformation structure

Figure 7. The experimental setup of WiFinger.

Figure 8. The way of user performing finger gestures.

Figure 9. Finger gesture extraction accuracy per gesture for users 1-10

features, and searches for the class majority label among k nearest neighbors of the corresponding finger gesture feature.

IMPLEMENTATION & EVALUATION

Hardware Setup

In the experiment, we use one 3.2 GHz dual core CPU, 4GB ROM desktop, which is equipped with Intel 5300 NIC, as the receiver, and its operating system is Ubuntu 12.04. A TP-LINK TLWDR4300 wireless router connecting to a laptop with a cable is deployed as the transmitter, and the router is set as a AP. The AP possesses one DB-Link directional antenna and operates in IEEE 802.11n AP mode at 5.745 GHz (channel 149). The receiver has 3 omni-directional antennas and its firmware is modified as in [43] to report CSI to upper layers. In addition, the laptop also serves as a Monitor Point (MP) and connects to the receiver by SSH, which realizes remote control to minimize the interferences from other operators. The receiver and the AP are posited on cabinets at the same height. The distance between the receiver and the AP is 50 cm, and the user performs finger gestures in the middle of Line-Of-Sight (LOS) path between TX and RX antennas. We set the directional antenna facing the receiver, which enhances the received signals to capture the subtle signal changes caused by finger gestures of the target user. Figure 7 illustrates the experimental setup of WiFinger. During the measurement campaign, the MP continuously sends packets with a high data rate of about 2000 $packets/s$ towards receiver using PKTGEN tool through the WiFi router. Setting a higher sending packets frequency leads to a higher sampling rate of CSI, which ensures the time resolution of CSI values for capturing the subtle changes in CSI stream and maximizes the details of different finger motions.

Controlled Gesturing

To accurately measure the CSI variations caused by finger motions during data collection, we ask users to sit on a fixed chair aside the middle of LOS path between transceivers. For the sake of avoiding the interferences caused by the movement of other human body parts, we instruct the users not to move their heads or other body parts in significant scale. Users perform the gestures naturally while keeping the interval time of gestures between 1 to 2 seconds to make the starting and finishing points of finger motions easily to be identified. When the user performs finger gestures, his/her hand remains 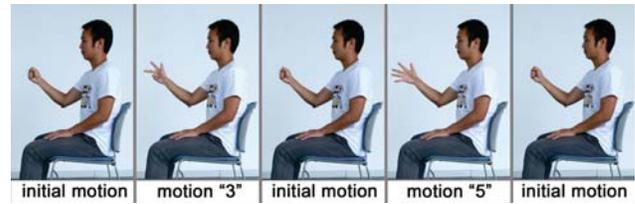in the fixed position continuously and only the fingers move. Figure 8 shows how the user performs gesture "3" and gesture "5" successively.
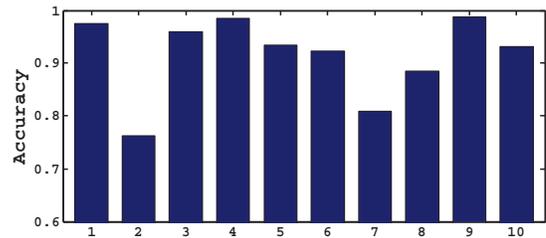
Data Collection

We collect training and testing data from 10 users to feed the classification algorithm. Those users are college students who volunteered to collect experiment data for WiFinger. We give them a brief introduction about the finger gestures digit 1 to 9 in ASL. Users 2-10 perform these gestures each for 35 times under our environment to build the training data sets. To evaluate the effects of training samples in different size. We collect 70 instances for each of the 9 finger-grained gestures from user 1. The users perform the gestures sequences naturally with an average time interval at 2 seconds. We totally acquire 3465 instances for WiFinger. After that, we randomly generate 20 sequences using matlab with 18 digits per sequence, and each user chooses 5 sequences to perform. WiFinger selects the CSI stream of the above instances, and processes the original CSI values following three steps: signal preprocessing, gesture extraction and then classification.

Automatic Gesture Extraction Accuracy

We evaluate the accuracy of the finger motion detection algorithm in WiFinger. The detection accuracy is defined as the ratio of total number of correctly extracted finger gesture profiles to the total number of finger motions performed by each user. Figure 9 plots the percentage of correctly extracted finger motions of 9 digits from 10 users.

As the figure shows, the detection accuracy varies among users. This is caused by the habits of performing finger motions. Different users probably perform the same gesture in slightly different ways or orientations. The behaviors of performing finger motions affect the detection rates of WiFinger. For example, we observed that user 2 and user 7 separately have average gesture detection ratio of 76% and 81%. We also find that these two users sometimes cannot fully stretch their fingers in some finger gestures. For instance, when they perform finger gesture 3 and 9, their fingers stretch in small
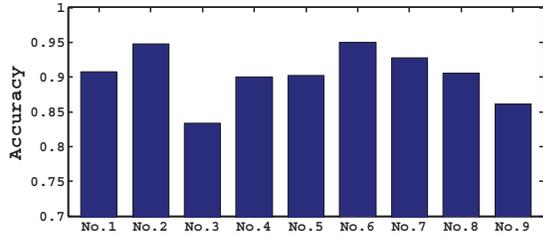
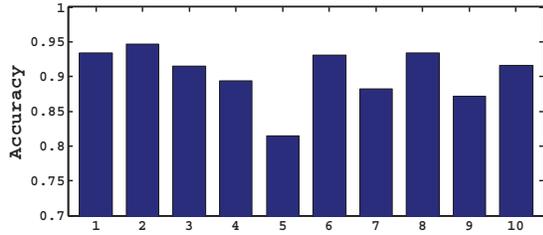Figure 10. Recognition accuracy per gesture for users 1-10

Figure 11. Finger gesture recognition accuracy per user

Figure 12. The average accuracy for user 1 using different instances per gesture

Figure 13. The average accuracy per sequence for all users

scale, which results in relatively inapparent variations in CSI waveform. However, user 4 and 9 possess comparably a right performing behaviors. Hence the gesture detection ratios of user 4 and 9 are both above 98%. The extraction accuracy can be further improved for the given users by adjusting the parameters of algorithm in WiFinger.

Gesture Recognition Accuracy
Here, we elaborate the recognition accuracy of finger gesture in WiFinger. We evaluate WiFinger with two kinds of experiments. In the first experiment, we separately feed the user-specific kNN classifiers using 35 profiles per finger gesture of 10 users. In all these experiments, we set $k = 5$, and then we perform 10-fold cross validation on the finger gesture features in each classifier to acquire the recognition accuracy of each finger gesture. In the second experiment, we test each classifier using corresponding three randomly selected 18 digit sequences performed by each user. As the results show, WiFinger provides a recognition accuracy more than 90% for the continuously random 18 digit sequences. Next, we give detailed explanations for the evaluation process of WiFinger.

Average Recognition Accuracy
We evaluate the recognition accuracy of WiFinger using average recognition accuracy per user and average recognition accuracy per finger gesture. For each user, we calculate the mean value of the total recognition accuracy of 9 digits finger gestures. And we also get the average accuracy per finger gesture of all 10 users. Figure 10 shows the average recognition accuracy per gesture. We observe that the recognition accuracies of gesture 3 and 9 are around 85% which are relatively lower than the other gestures. This is because we set the initial gesture as a fist, and gesture 3 and 9 need to open several consecutive fingers by a large margin. Some users hardly perform these gestures with a fully opened palm, hence lead to a relatively low recognition accuracy. However, the overall finger gesture recognition accuracy is still more than 90%. Figure 11 plots the average recognition accuracy per user. As the
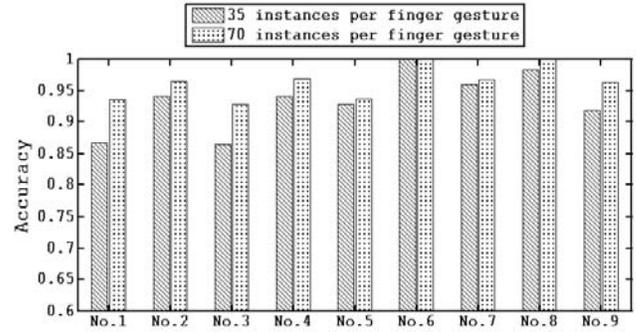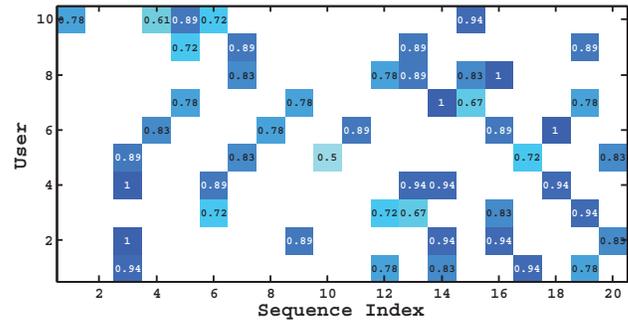
result shows that the recognition accuracy of user 5 is around 80%, this is related to his behaviors when he performs finger gestures. The sitting posture and the behaviors of performing gestures will affect the recognition accuracy in our scheme.

Impacts of Training Set Size
To evaluate the impact of the training set size. We increase the number of training samples per finger gesture from 35 to 70. And we perform 10-fold cross validation towards the 9 finger gestures when 35 and 70 training samples are used per gesture. Figure 12 plots the recognition accuracy for user 1 with two different training set size. We observe that the average recognition accuracy increased from 93.37% to 96.25% for user 1 with the number of training samples increasing from 35 to 70.

Number Text Input Using WiFinger
To evaluate the accuracy of continuous text input in WiFinger, we collected CSI values for the randomly selected 18 digits sequences performed by different users as mentioned before. We use the dataset of individual finger gesture features as training set, and the extracted finger gesture features from the corresponding three randomly selected 18 digit sequences of the same user as test data to feed the classifier. As is shown by Figure 13. WiFinger achieves average recognition accuracy of 82.67% for all finger gesture sequences.

DISCUSSTION AND LIMITATIONS
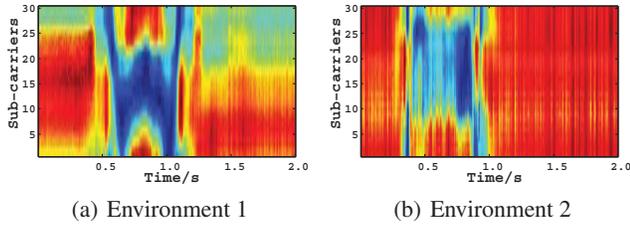In our current experimental environment such as laboratory and dormitory, WiFinger runs well under a relatively stable

(a) Environment 1   (b) Environment 2

Figure 14. Finger gesture No.3 in two different multipath environment.

(a) Stream 1-2 (RX antenna B)   (b) Stream 1-3 (RX antenna C)
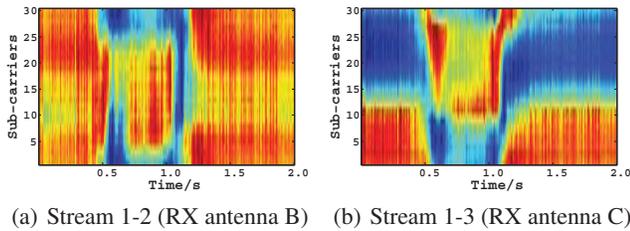
Figure 15. CSI variations caused by finger gesture No.3 in different CSI streams.

(a) 2.4G-No.3   (b) 5G-No.3

(c) 2.4G-No.3   (d) 5G-No.3

(e) 2.4G-No.6   (f) 5G-No.6

Figure 17. Features comparison at 2.4 GHz and 5 GHz.

environment with only two occupants, a target user and a system controller. The accuracy of WiFinger is related to the variations of environment such as human moving around the target user, the orientation and distance between transmitter and receiver. The following illustrations are some discussions of WiFinger.

Differences among TX-RX Antenna Pairs

Antennas separated by more than half a wavelength ($>$ $2.6\ cm$ at 5.745 GHz) have independent fading channels [38]. In WiFinger, the distances between each two RX antennas are greater than $2.6\ cm$, and we totally obtain 1×3 CSI streams. To understand the differences among TX-RX antenna pairs, we analyzed all 3 CSI streams in our experiment. We have the following observation: (1) Figure 14(a), Figure 15(a) and Figure 15(b) illustrate the CSI variations caused by the same finger gesture No.3 in all 3 CSI streams, respectively, and they are dramatically different. (2) It improves finger gesture recognition by integrating features resulted from finger gesture in all 3 CSI streams, because combining all 3 streams can obtain more detailed and distinguishable features for finger gestures. (3) However, it absolutely increases the time and space complexity during DTW. WiFinger finds a trade-off between accuracy and computational complexity and selects only one stream to recognize finger gestures.

Interference with Multipath Reflection

In our experiment, the variations in the environment may affect the accuracy of WiFinger, such as human body motions and moving objects around the target user. WiFinger is designed only for single target user with a system administrator who guides the target user during the data collection. Figure 14(a) and 14(b) show the CSI stream patterns of two different environments in the same 5 GHz band. Environment 1 is our original experimental environment. In addition, we add two chairs and a cabinet around the receiver to form Environment 2. Thus, there are different multipath reflectors in these
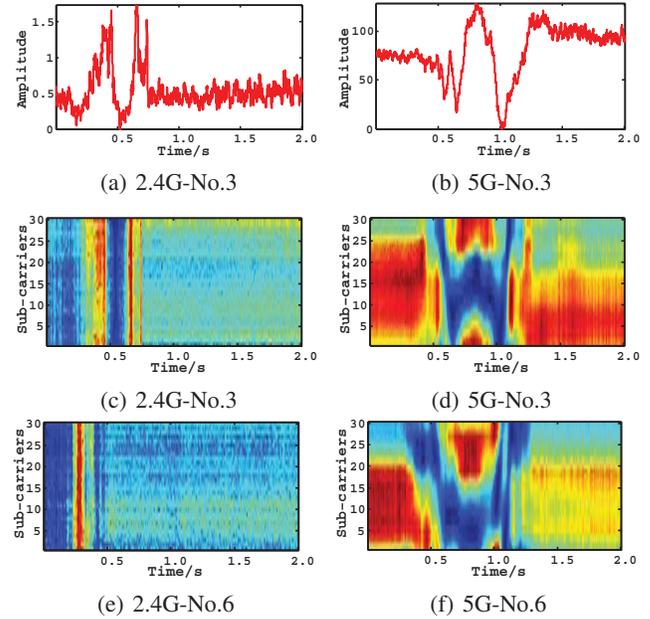
two environments. We observe that they almost have approximate patterns in the two different environments. Moreover, owing to the directivity of our TX antenna, a irrelevant person who walks or twists his body out of the coverage range has little interference on WiFinger. Hence, WiFinger is resilient to the static reflections or some marginal interferences. However, when another person moves around the target user, his activities can disturb or even cover the minute CSI fluctuations induced by finger gestures, which leads to unsatisfactory results.

Devices Positioning

WiFinger applies DB-Link directional antenna operating in the 5GHz frequency band. We test the classification results under different distances of transceivers with the target user sitting aside in the middle between TX and RX antennas. We put the receiver at a fixed position and adjust the position of the transmitter. Figure 16(a), 16(b) and 16(c) show the waveforms of finger gesture No.2 at different distances between transceivers. We find that the detected finger gesture patterns get weaker with the distance between transceivers increasing. And when the distance exceeds $1m$, the patterns caused by finger gestures in CSI stream nearly disappeared. Therefore, in order to capture obvious and stable finger gesture patterns in CSI stream, we should place transceivers at a relatively close distance.

Different Frequency Band

Due to the large number of devices operating at 2.4 GHz in our surrounding environment, 2.4 GHz frequency band is crowded and more vulnerable to be interfered. To validate the impacts of different frequency bands, we install the transmitter with another directional antenna TL-ANT2406A and operate it in IEEE 802.11n AP mode at 2.432 GHz (channel 5) and collect CSI values of the same pre-set finger gestures.
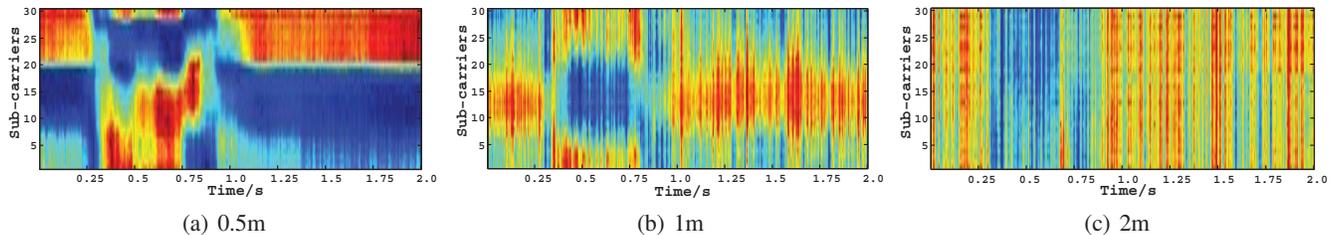
(a) 0.5m   (b) 1m   (c) 2m

Figure 16. Different distances between transceivers.

After comparing the differences of CSI streams modulated by finger motions in 2.4 GHz and 5 GHz frequency band, we have the following observations. First, the period of finger gestures and their amplitudes at 2.4 GHz are relatively smaller than that at 5 GHz. The waveform possesses more noise component at 2.4 GHz, which can be shown in Figure 17(a) and 17(b). Second, as Figure 17(c) to 17(f) show, compared with the CSI streams at 5 GHz, the CSI streams at 2.4 GHz are highly correlated. This is because the channels in 2.4 GHz band have a high overlapping ratio, however, they are not overlapped in 5 GHz [44, 45]. Third, we find that the CSI variations caused by different finger gestures are weaker at 2.4 GHz. Considering 2.4 GHz frequency band suffers more serious interferences, and the frequencies of phase change caused by the same finger gestures are lower. Specifically, a finger gesture moving at a speed of v m/s can result in a maximum Doppler shift of $\frac{2f}{c}v$, where f is the frequency of one OFDM subcarrier [46]. For example, a 0.5 m/s finger gesture results in a 8 Hz Doppler shift in a 2.4 GHz frequency band and a 17 Hz Doppler shift in a 5 GHz band.

CSI Sampling Rate

Sampling rate is extremely important for finer-grained gesture recognition, because the changes caused by finger motions are subtle. The accuracies of finger motion detection and extraction both rely on a high time resolution of CSI values. Higher sampling rate means more CSI values between the starting and finishing point of each gesture, which contains more information in CSI waveform, hence increases the finger motion extraction and classification accuracy. To verify this assumption, we also test WiFinger at sampling rate of 500 $packets/s$, and ask user 1 to perform 35 times per finger gesture. We find that the periods of finger gestures in collected CSI streams are obviously shorter. Therefore, we adjust the parameters in WiFinger such as segment window length, sliding step length and guard interval. We perform 10-fold cross validation on the data obtained. After evaluation, the average recognition accuracy with a sampling rate of 500 $packets/s$ is 80.95%. However, the average recognition accuracy is 93.38% when the sampling rate reaches 2000 $packets/s$. Thus, CSI sampling rate is especially important for finer-grained gesture recognition in WiFi environment.

Context-based Error Correction

For the the environment and users that WiFinger has been trained on, it achieves an average accuracy of more than 90%. Currently, WiFinger enables the recognition and number text input using 9 digit finger gestures from ASL. In future work, we prepare to enlarge the finger gesture set and achieve text input of any characters or sentences with high accuracy. We note that the recognition accuracy could be further improved by utilizing context-based error correction. This idea is inspired by widely used context-aware approaches in automatic speech recognition [47]. For example, when the target user continuously performs finger gestures to achieve continuous text input using WiFinger. When WiFinger detects a wrong word like "gestuer", WiFinger can automatically distinguish and recognize "gesture" instead of "gestuer" based on the context. This method reduces the recognition mistakes and facilitates the recognition of gestures in WiFinger.

CONCLUSION

In this paper, we present WiFinger, a novel system that enables WiFi signals to realize continuously number text input in WiFi devices by recognizing finger-grained gestures from American Sign Language. WiFinger exploits the ubiquitous WiFi signals to sense finger-grained gesture, and it does not need any specialized hardware devices. Hence it can be extended easily to commercial Wi-Fi products. The preliminary results show that WiFinger achieves high recognition accuracy with an average recognition accuracy 90.4% per user. WiFinger collects channel state information available in the existing WiFi devices. We benefit from the observation that different human activities, for example finger gestures, lead to a number of unique CSI time-series patterns. Thus, we are inspired to explore the potential of sensing finger-grained gestures under wireless environment. Compared with other gesture recognition schemes, WiFinger is a non-intrusive and device-free solutions in finger-grained gesture recognition. Given that our scheme is able to identify subtle changes caused by tiny movement of human body, therefore, it can be extended and improved to wide areas such as fine-grained gesture recognition, human activities detection, or any motion recognition scenarios.

REFERENCES

1. Ali, K., Liu, A. X., Wang, W., and Shahzad, M. Keystroke recognition using wifi signals. In *Proc of ACM MobiCom* (2015), 90–102.

2. Wang, G., Zou, Y., Zhou, Z., k. wu, and Ni, L. We can hear you with wi-fi! In *Proc of ACM MobiCom* (2014), 593–604.

3. Gupta, S., Morris, D., Patel, S., and Tan, D. Soundwave: using the doppler effect to sense gestures. In *Proc of the SIGCHI Conference on Human Factors in Computing Systems* (2012), 1911–1914.

4. Leap Motion. `https://www.leapmotion.com`.

5. Microsoft Kinect. `http://www.roborealm.com/help/MicrosoftKinect.php`.

6. Pu, Q., Gupta, S., Gollakota, S., and Patel, S. Whole-home gesture recognition using wireless signals. In *Proc of ACM MobiCom* (2013), 27–38.

7. Adib, F., and Katabi, D. See through walls with wifi! In *Proc of ACM SIGCOMM* (2013), 75–86.

8. Melgarejo, P., Zhang, X., Ramanathan, P., and Chu, D. Leveraging directional antenna capabilities for fine-grained gesture recognition. In *Proc of ACM UbiComp* (2014), 541–551.

9. Halperin, D., Hu, W., Sheth, A., and Wetherall, D. Tool release: Gathering 802.11n traces with channel state information. *ACM SIGCOMM Computer Communication Review 41*, 1 (2011), 53–53.

10. Sturman, D. J., and Zeltzer, D. A survey of glove-based input. *IEEE Computer Graphics and Applications 14*, 1 (1994), 30–39.

11. Dipietro, L., Sabatini, A. M., and Dario, P. A survey of glove-based systems and their applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 38*, 4 (2008), 461–482.

12. Finger Gesture. `http://www.lifeprint.com/dictionary.htm`.

13. Scholz, M., Riedel, T., Hock, M., and Beigl, M. Device-free and device-bound activity recognition using radio signal strength. In *Proc of the 4th Augmented Human International Conference*, ACM (2013), 100–107.

14. Tarzia, S. P., Dick, R. P., Dinda, P. A., and Memik, G. Sonar-based measurement of user presence and attention. In *Proc of ACM UbiComp* (2009), 89–92.

15. Zhao, Y., Patwari, N., Phillips, J. M., and Venkatasubramanian, S. Radio tomographic imaging and tracking of stationary and moving people via kernel distance. In *Proc of the 12th international conference on Information processing in sensor networks* (2013), 229–240.

16. Starner, T., and Pentland, A. Real-time american sign language recognition from video using hidden markov models. In *Motion-Based Recognition*. Springer, 1997, 227–243.

17. Juan, P. W., Kölsch, M., Stern, H., and Edan, Y. Vision-based hand-gesture applications. *Communications of the ACM 54*, 2 (2011), 60–71.

18. Rehg, J. M., and Kanade, T. Visual tracking of high dof articulated structures: an application to human hand tracking. In *Computer Vision ECCV'94*. Springer, 1994, 35–46.

19. Ketabdar, H., Roshandel, M., and Kamer, A. Y. Towards using embedded magnetic field sensor for around mobile device 3d interaction. In *Proc of the 12th international conference on Human computer interaction with mobile devices and services* (2010), 153–156.

20. Ketabdar, H., Moghadam, P., Naderi, B., and Roshandel, M. Magnetic signatures in air for mobile devices. In *Proc of the 14th international conference on Human-computer interaction with mobile devices and services companion* (2012), 185–188.

21. Agrawal, S., Constandache, I., Gaonkar, S., Roy, C. R., Caves, K., and DeRuyter, F. Using mobile phones to write in air. In *Proc of ACM Mobisys* (2011), 15–28.

22. Park, T., Lee, J., Hwang, I., Yoo, C., Nachman, L., and Song, J. E-gesture: a collaborative architecture for energy-efficient gesture recognition with hand-worn sensor and mobile devices. In *Proc of ACM SenSys* (2011), 260–273.

23. Sigg, S., Shi, S., Buesching, F., Ji, Y., and Wolf, L. Leveraging rf-channel fluctuation for activity recognition: Active and passive systems, continuous and rssi-based signal features. In *Proc of International Conference on Advances in Mobile Computing & Multimedia* (2013), 43.

24. Sigg, S., Scholz, M., Shi, S., Ji, Y., and Beigl, M. Rf-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals. *Mobile Computing, IEEE Transactions on 13*, 4 (2014), 907–920.

25. Abdelnasser, H., Youssef, M., and Harras, K. A. Wigest: A ubiquitous wifi-based gesture recognition system. In *Computer Communications (INFOCOM), 2015 IEEE Conference on* (2015), 1472–1480.

26. Han, C., Wu, K., Wang, Y., and Ni, L. M. Wifall: Device-free fall detection by wireless networks. In *Proc of IEEE INFOCOM* (2014), 271–279.

27. Wang, Y., Liu, J., Chen, Y., Gruteser, M., Yang, J., and Liu, H. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *Proc of ACM MobiCom* (2014), 617–628.

28. Zhou, Z., Yang, Z., Wu, C., Shangguan, L., and Liu, Y. Towards omnidirectional passive human detection. In *Proc of IEEE INFOCOM* (2013), 3057–3065.

29. Xi, W., Zhao, J., Li, X.-Y., Zhao, K., Tang, S., Liu, X., and Jiang, Z. Electronic frog eye: Counting crowd using wifi. In *Proc of IEEE INFOCOM* (2014), 361–369.

30. Nandakumar, R., Kellogg, B., and Gollakota, S. Wi-fi gesture recognition on existing devices. *CoRR abs/1411.5394* (2014).

31. Sen, S., Lee, J., Kim, K.-H., and Congdon, P. Avoiding multipath to revive inbuilding wifi localization. In *Proc of the 11th annual international conference on Mobile systems, applications, and services* (2013), 249–262.

32. Xiao, J., Wu, K., Yi, Y., and Ni, L. M. Fifs: Fine-grained indoor fingerprinting system. In *Computer Communications and Networks (ICCCN), 2012 21st International Conference on* (2012), 1–7.

33. Yang, Z., Zhou, Z., and Liu, Y. From rssi to csi: Indoor localization via channel response. *ACM Computing Surveys (CSUR) 46*, 2 (2013), 25.

34. Kellogg, B., Talla, V., and Gollakota, S. Bringing gesture recognition to all devices. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)* (2014), 303–316.

35. Lyonnet, B., Ioana, C., and Amin, M. G. Human gait classification using microdoppler time-frequency signal representations. In *Radar Conference, 2010 IEEE* (2010), 915–919.

36. Adib, F., Kabelac, Z., Katabi, D., and Miller, R. C. 3d tracking via body radio reflections. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)* (2014), 317–329.

37. Chen, B., Yenamandra, V., and Srinivasan, K. Tracking keystrokes using wireless signals. In *Proc of the 13th Annual International Conference on Mobile Systems, Applications, and Services* (2015), 31–44.

38. Halperin, D., Hu, W., Sheth, A., and Wetherall, D. Two antennas are better than one: A measurement study of 802.11 n.

39. Davies, L., and Gather, U. The identification of multiple outliers. *Journal of the American Statistical Association 88*, 423 (1993), 782–792.

40. Wang, W., Liu, A. X., Shahzad, M., Ling, K., and Lu, S. Understanding and modeling of wifi signal based human activity recognition. In *Proc of ACM MobiCom* (2015), 65–76.

41. Mallat, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 11*, 7 (1989), 674–693.

42. Müller, M. Dynamic time warping. *Information retrieval for music and motion* (2007), 69–84.

43. Halperin, D., Hu, W., Sheth, A., and Wetherall, D. Predictable 802.11 packet delivery from wireless channel measurements. In *Proc of ACM SIGCOMM* (2010), 159–170.

44. Committee, I. C. S. L. M. S., et al. *Wireless LAN medium access control (MAC) and physical layer (PHY) specifications*. IEEE Std, 1997.

45. Perahia, E., and Stacey, R. *Next Generation Wireless LANS: 802.11 n and 802.11 ac*. Cambridge university press, 2013.

46. Quegan, S. Spotlight synthetic aperture radar: Signal processing algorithms. *International Journal for Numerical Methods in Engineering 83*, 11 (2010), 1498–1517.

47. Baldauf, M., Dustdar, S., and Rosenberg, F. A survey on context-aware systems. *International Journal of Ad Hoc and Ubiquitous Computing 2*, 4 (2007), 263–277.