WiImage: Fine-grained Human Imaging from Commodity WiFi

Shengkai Zhang, Gaoyang Liu, Keping Yu, and Shaohua Wan*

Abstract—Privacy-preserving wireless-based human imaging technologies attract much attention in academia and industry. They can serve in surveillance, security inspection, and health monitoring, while preventing the privacy leakage from current camera-based surveillance system. However, previous solutions either require dedicated hardware or costly infrastructure deployment, which hurts their practicality. In this paper, we propose WiImage, a low-cost, instantly-deployed sensing system that can capture a fine-grained human image and infer his contour using ubiquitous WiFi. WiImage is free of pre-training or placing any marker on a user’s body. WiImage consists of two technical components. First, a lightweight multipath resolution algorithm exploits the spatial and temporal correlation of WiFi packets to address the multipath and extract reflected signals from a human body. Second, an imaging algorithm infers the contour of a human from WiFi reflections. We prototype WiImage using off-the-shelf WiFi chips and conduct experiments in several typical indoor settings. The results show that WiImage can recover a representative human figure and use it to identify the user precisely.

I. INTRODUCTION

Radio frequency (RF) sensing is an emerging research field and has attracted considerable attentions. It impels many remarkable applications, such as indoor localization [1], [2], [3], [4], gesture recognition [5], [6], fall detection [7], robot navigation [8], [9], [10] and human motion tracking [11], [12]. Despite the success in diverse applications, recent years have witnessed the rapid proliferation of RF-based human imaging technologies, which provide privacy-preserving services for applications like surveillance, security inspection, and health monitoring.

There have been many advances on human imaging using RF signals. Noticeable approaches include RF-Capture [13] and Tagscan [14]. However, they either require dedicated hardware that generates ultra-wideband Frequency-Modulated Continuous-Wave (FMCW) signals [13], or need to deploy costly infrastructure [14] with RFID readers. To improve the practicality of RF-based human imaging, the design should be low-cost, easy-to-deploy, and accurate. WiFi-based sensing, e.g., localization [3], gait recognition [15], and breath monitoring [16], holds the potential to meet the above requirements due to that 1) WiFi is ubiquitously available in modern cities, 2) WiFi devices are very cheap (costs 10-30 US dollars), and 3) multi-antenna transceivers and WiFi OFDM signals with multiple subcarriers encode rich spatial and temporal information of signal propagation, and thus can improve the accuracy. Existing WiFi-based imaging systems [17] suffers from lackluster performance, leaving the design space to improve the imaging accuracy.

This paper proposes WiImage, a passive sensing system that captures a fine-grained human image and recover his contour using commodity WiFi. WiImage analyzes the signal reflections from a human between WiFi transceivers to enable the human imaging and identification. We leverage the spatial difference of multipath to resolve the reflections from the human body and other objects. Then, WiImage profiles the body reflections and recover the human image.

Realizing the above idea poses two challenges. First, a WiFi receiver will obtain a combination of the reflections from multipath in an indoor environment. Human imaging requires the ability to separate reflected signals from a human body. To address this problem, we propose a human reflection extraction algorithm that leverages the spatial diversity of multiple antennas and the temporal continuity of the human reflection signals. The key idea is that the phase difference among multiple antennas encode the angle-of-arrival (AoA) and the multiple subcarriers of an OFDM signal encode the time-of-flight (ToF). This makes the WiFi channel state information (CSI) on different antennas from a strong reflection path exhibit a slight change, indicating the additional propagation distance. Therefore, we can eliminate the multipath effect by comparing the ToF difference. The human reflection signal will show a temporally stable spatial constraint.

Second, the large wavelength of WiFi signals (12 cm for 2.4 GHz band) severely limits the sensing resolution, making it hard to determine the sharp contour of a human body. To address this challenge, we propose a contour recovery algorithm that uses the reflection signals’ power spectrum. We find that the power of the human reflection signal depends on the body size. Thus, we design a contour indicator that relates to the ratio of the signal power and the change rate of the power, which is very effective for human contour recovery.

We prototype WiImage using only one WiFi router with an Intel 5300 WiFi card. Three directional antennas are attached on the WiFi card in a linear form. We evaluate WiImage in terms of the imaging performance and the accuracy of human identification in two venues: an office in our lab and a corridor

S. Zhang is with the School of Information Engineering, Wuhan University of Technology, Wuhan 430074, China. (E-mail: shengkai.zhang@gmail.com).
G. Liu is with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China. (E-mail: liu-gao-yang@hust.edu.cn).
K. Yu is with Global Information and Telecommunication Institute, Waseda University, Shinjuku, Tokyo, 169-8555, Japan. (E-mail: keping.yu@aoni.waseda.jp).
S. Wan is with School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China. (Email: shaohua.wan@ieee.org).

*The corresponding author is Shaohua Wan (shaohua.wan@ieee.org).
in our academic building. To evaluate whether these RF-based human images have enough information to distinguish different persons, we take the images into a classifier to perform the identification of participants. In the corridor environment, the human identification accuracy is higher than 88.5% for 5 different participants. WiImage can also extract the contour of a human body accurately that can help extract biological features, including height, weight and body surface area (BSA) [18].

**Contribution:** The main contributions of this paper are summarized as follows:

- We propose a human reflection extraction algorithm that eliminates the multipath effect and extracts the human reflection component.
- We design a contour recovery algorithm to determine the contour of a human from the body reflected signals’ power spectrum. This algorithm works well no matter the target is overweighted or slim.
- We implement WiImage with commercial WiFi devices and conduct extensive experiments to demonstrate the effectiveness under various conditions.

The organization of this paper is as follows. We introduce the preliminary of signal processing in Section II. Section III elaborates on the design of WiImage. The implementation is detailed in Section IV followed by the evaluation in Section V. We further discuss the limitation of WiImage in Section VI and summarize the related work in Section VI. Finally, we conclude this work in Section VII.

II. PRELIMINARY

In this section, we introduce the background of CSI and AoA estimation.

A. Channel States Information

In wireless communication, CSI manifests the channel property of a wireless link, indicating how wireless signals propagate from a transmitter to a receiver. In WiFi systems, the channel model between a pair of transmitting (TX) and receiving antennas (RX) is shown below:

\[ y = H \times x + n \]

where \( x \) denotes the transmitted signal, \( n \) the noise, \( y \) the received signal, and \( H \) the complex matrix of Channel Frequency Response (CFR). WiFi OFDM signals have multiple subcarriers at different frequencies [2]. We denote \( N_t \) and \( N_r \) as the number of TX and RX antennas respectively. We further denote \( N_s \) as the number of subcarriers. Then \( H \) can be a CSI matrix, whose size is \( N_t \times N_s \times N_r \). A WiFi card (e.g., Intel 5300) can report the CSI matrix for each packet. WiFi CSI provides fine-grained channel measurements across multiple subcarriers, including not only signal strengths but also phases.

B. AoA Estimation

When an RF signal propagates in the air, the phase keeps rotating and one wavelength of transmitting distance corresponds to a \( 2\pi \) phase rotation. Due to the spatial differences of RX antennas, received signals experience different traveling distances across antennas, and thus have different phase rotations. For an instance, as shown in Fig.1, the antennas are spaced with an interval of half wavelength. An incident signal arrives at the array with AoA \( \theta \). The signal phases at Ant.1 and Ant.2 are \( \phi_1 \) and \( \phi_2 \). Then AoA \( \theta \) can be estimated by following equation:

\[ \theta = \arccos\left(\frac{\phi_1 - \phi_2}{2\pi \cdot d}\right) \]

where \( \lambda \) is the wavelength and \( d \) is the spacing distance of adjacent antennas.

The classic MUSIC algorithm [19] has been used to estimate AoA. To understand how MUSIC work, we consider there are \( L \) propagation paths and a uniform linear array with \( M \) antennas as shown in Fig.1. Signal \( S = [s_1, s_2, \ldots, s_L]^T \) arrive at the array at directions of \( \Theta = [\theta_1, \theta_2, \ldots, \theta_L] \). The received signal is mixture of all \( L \) signals, then the measured signal \( x_m \) at antenna \( m \) can be expressed as:

\[ x_m = \sum_{l=1}^{L} s_l \cdot e^{-j(m-1)\frac{2\pi d}{\lambda} \cos(\theta_l)} \]

The received signal \( X = [x_1, x_2, \ldots, x_M]^T \) is:

\[ X = AS + n \]

where \( n \) is the white noise, \( A = [a(\theta_1), a(\theta_2), \ldots, a(\theta_L)] \) is the steering matrix and \( a(\theta_l) = [1, \Phi(\theta_l), \Phi(\theta_l)^2, \ldots, \Phi(\theta_l)^{M-1}]^T \) is the steering vector, \( \Phi(\theta_l) = e^{-j\frac{2\pi dl}{\lambda} \cos(\theta_l)} \).

\( X \) can be easily measured by the RX antenna array. Steering matrix \( A \) is determined by parameters including wavelength, antenna spacing, and the number of antennas. The AoA of a signal propagating in multiple paths can be derived by Eqn. (3). MUSIC analyzes the eigenstructure of a signal’s correlation matrix \( R \) to estimate the AoA. Based on Eqn. (3), the correlation matrix \( R \) can be expressed as following:

\[ R = \mathbb{E}[XX^H] = AE[SS^H]A^H + \sigma^2 I \]

where \( \mathbb{E}[SS^H] \) is the source correlation matrix. The matrix \( R \) has \( M \) eigenvalues \( \lambda_1, \ldots, \lambda_M \) associated with eigenvectors \( U = [u_1, \ldots, u_M] \). The eigenvectors corresponding to the largest \( L \) eigenvalues forms the signal space and the rest eigenvectors denote the null space of noise. Thus, the eigenvectors...
WiFi transmits signals with multiple subcarriers. The dimension
of measurement matrix $X$ can become “fat” enough to distinguish the signal space and the noise space.

$U$ can be separated into parts of signal and noise space:

$$U = [U_s, U_n],$$
$$U_s = [u_1, \ldots, u_L],$$
$$U_n = [u_{L+1}, \ldots, u_M].$$

where $U_s$ is the signal space and $U_n$ is the noise space. It can be seen that $M$ must be larger than the number of propagation paths. Otherwise, MUSIC cannot solve out all the paths. This means that the number of antennas must be larger than the number of paths. Satisfying this constraints we have the AOA spectrum:

$$P(\theta) = \frac{1}{a^H(\theta)U_NU_N^HAL(\theta)}. \quad (6)$$

Finding the maximal $P(\theta)$ by searching $\theta$ can obtain the AOA of paths.

**Super-resolution AoA Estimation** From the above discussion, the multipath resolution of the MUSIC algorithm subjects to the number of receiving antennas. A commercial WiFi network card usually supports no more than 3 antennas (e.g., Intel 5300, Atheros). But in an indoor environment, the number of signal propagation paths is usually more than 4. The MUSIC algorithm cannot be used by commercial WiFi cards.

The key to break the curse of MUSIC is to make the measurement matrix $X$ “fat”. According to 802.11n standard, WiFi transmits signals with multiple subcarriers. The dimension of measurement matrix $X$ can be expanded by a smooth operation on subcarriers. For example, as shown in Fig. 2, there are three antennas attached on an Intel 5300 card. The smoothed measurement matrix can become $32 \times 30$. MUSIC algorithm can be directly applied on the new CSI matrix to estimate the AoAs of all propagation paths. Please refer to [1] for the detailed mathematical derivation.

### III. Design

#### A. System Overview

WiImage is a human contour profiling system using commodity WiFi. It operates through moving received and transmitted antennas along a human body vertically. The system captures reflection signals off the human body at every vertical point and analyzes these signals to extract human contour features. Finally, WiImage uses the above measurements to derive human body features including height, weight, BSA, BMI, and profile the target human contour. The system pipeline is shown in Fig. 3.

#### B. CSI Preprocessing

Due to the imperfection of WiFi cards, the reported CSI has amplitude offsets and phase rotations. We cannot directly use the reported CSI to do the sensing. Previous studies [1], [20] have reported the measurement error sources including power control uncertainty, packet detection delay (PDD), sampling frequency offset (SFO), carrier frequency offset (CFO), random initial phase offset, and phase ambiguity.

First, we mitigate the impact of above errors on CSI phase measurements. For a pair of TX and RX antennas, the measured phase $\phi_i$ for $i_{th}$ subcarrier can be expressed as:

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N_c} B\delta + \beta + Z, \quad (7)$$

where $\hat{\phi}_i$ denotes the phase shift from signal propagation, $\delta$ is the processing time at the receiver, including time shift due to PDD and SFO, $\beta$ is an unknown phase offset, and $Z$ is the measurement noise. $k_i$ is the subcarrier index and $N_c$ denotes the total number of subcarriers ($N_c = 30$ in our implementation). $B$ denotes the signal bandwidth. To mitigate the impact of CSI phase errors, we perform a linear transformation on the raw CSI phases [21]:

$$\hat{\phi}_i = \phi_i - ak_i - b$$

$$= \phi_i - \frac{\phi_i N_c - \phi_1}{N_c} k_i - \frac{1}{N_c} \sum_{j=1}^{N_c} \phi_j, \quad (8)$$

$$a = \frac{\phi_1 - \bar{\phi}_i}{N_c}, \quad b = \frac{1}{N_c} \sum_{j=1}^{N_c} \bar{\phi}_j,$$

where the $\bar{\phi}_i$ is a linear combination of true phases. By subtracting the linear term $ak_i + b$ from the raw phase, time offset $\delta$ and phase offset $\beta$ are eliminated.

To mitigate the impact of CSI amplitude errors, we utilize the Hampel identifier [22] to remove the outliers. The CSI amplitude errors can be removed by averaging a sufficient number of CSI measurements. However, taking too many CSI measurements requires a user to be stationary for a long time, which is impractical. To balance the effectiveness of error elimination and the real-time processing, we set a sliding window over time to bound the number of CSI measurements. The phase linear transformation and the sliding window averaging remove the noise while ensuring the real-time processing. Fig. 4 illustrates an example of CSI phases and amplitudes after the preprocessing.

#### C. Multipath Signal Elimination

In indoor environments, RF signals that reflect off a human body are mixed with reflections from other objects, e.g., furniture, tables, walls, etc. In this section, we describe how we eliminate the multipath effect.

In our system, we use a WiFi router with 2 antennas to send packets to a receiver equipped an Intel 5300 WiFi
card that attaches 3 antennas. We obtain the CSI matrices through an open-source linux CSI tool [23]. After the CSI preprocessing, we utilize the MUSIC algorithm on the reshaped CSI matrix [11] to obtain the AoA spectrum.

Our key observation is that the signals emitted by different TX antennas propagate through different reflection paths because of the spatial diversity of TX antennas. When a signal reflects only once in propagation, the AoA change among TX antennas won’t be very large. On the other hand, a signal reflected many times exhibits a drastic change. We illustrate this observation through an example as shown in Fig. 5. Consider a human standing near the antenna array, a WiFi signal transmitted by different TX antennas only reflect once through Path 1 and path 2 respectively. At the receiving end, the AoA change is \( \Delta \alpha = \alpha_1 - \alpha_2 \), where the \( \alpha_1 \) and \( \alpha_2 \) are the AoAs of the signals corresponding to the TX 1 and TX 2. The AoA change from signals propagating in Multipath 1 and Multipath 2 is \( \Delta \omega = \omega_1 - \omega_2 \). After reflecting several times, the AoA change of multipath reflections \( \Delta \omega \) is much larger than \( \Delta \alpha \). We can use this observation to eliminate the multipath reflections.

**AoA Change Model.** The threshold of the AoA change is key to WiImage. We first conduct a geometrical analysis of the AoA change among TX antennas.

\[
\Delta \alpha_{\text{change}} = \arctan\left( \frac{2h}{d} \right) - \arctan\left( \frac{2h}{d+l} \right),
\]

where \( \Delta \alpha_{\text{change}} \) denotes the theoretical AoA change, \( h \) the vertical distance from the human body to the antenna array, \( d \) the distance between a transmitter and a receiver. \( l \) denotes the spacing of TX antennas. From Eqn. (9), once \( h \) is fixed, \( \Delta \alpha_{\text{change}} \) becomes a constant. However, a human body is a scatterer, the AoA difference can’t always follow Eqn. (9). We empirically set the threshold of the AoA change to be \( \Delta \alpha_{\text{thr}} = 2 \times \Delta \alpha_{\text{change}} \).
**Signal Spectrum Selection.** WiImage filters out the AoAs of multipath reflections from MUSIC spectrum series \( P_{s,n} \) based on the threshold \( \Delta\alpha_{thr} \). In our experiment, we found that some AoA spectrums are relatively flat. The AoA estimations of these spectrums are not always precise by the peak searching, leading to inaccurate selections. To address this problem, our system performs dynamic time warping (DTW) to find the AoA change among received signals from different TX antennas. We keep the spectrum \( P_{1,n} \) constant, then translate \( P_{2,n} \) with a shifting degree \( \Delta\alpha \). When a shifting degree minimizes the DTW distance, we assign this degree to \( \Delta\alpha_n \), denoting the AoA change of \( n^{th} \) CSI measurement. The mathematical expression is:

\[
\Delta\alpha_n = \arg\min_{\Delta\alpha} DTW(P_{1,n}, P_{2,n}, \Delta\alpha). \tag{10}
\]

Next, we filter out the AoA spectrum whose \( \Delta\alpha_n \) is greater than the threshold \( \Delta\alpha_{thr} \). We denote \( P_{m,s} \) as \( s^{th} \) horizontal AoA spectrum in the reserved signal power azimuth spectrum (PAS) series \( \hat{P} \). We need to calibrate the AoA extracted from the reserved PAS series for further processing. We denote \( \hat{\alpha}_{m,s} \) as the AoA series of the retained PAS \( \hat{P}_{m,s} \). Since most objects are scatters with a reflection surface, the AoA spectrum peak should vary within a certain range. Then we use the K-means algorithm to cluster the AoA series \( \hat{\alpha}_{m,s} \) and find the center of each cluster. We denote \( \hat{\alpha}^{(u)}_{m,s} \) as the center of \( u^{th} \) cluster of the AoA series corresponding to the WiFi signal from \( m^{th} \) TX antenna.

Then, we compare the data in \( \left[ \hat{\alpha}^{(1)}_{1}, \hat{\alpha}^{(2)}_{1}, \ldots, \hat{\alpha}^{(U)}_{1} \right] \) and \( \left[ \hat{\alpha}^{(1)}_{1}, \hat{\alpha}^{(2)}_{1}, \ldots, \hat{\alpha}^{(V)}_{1} \right] \), where \( U \) is the number of AoA clusters of \( \hat{P}_{1,s} \) and \( V \) is the number of AoA clusters of \( \hat{P}_{2,s} \). A center of an AoA cluster can represent a stationary object in the environment. Due to the spacial difference of TX antennas, we reserve the pair of AoA cluster centers whose AoA change is smaller than \( \Delta\alpha_{thr} \), and denote \( \Psi_1 \) and \( \Psi_2 \) as the reserved AoAs corresponding to different TX antennas respectively.

### D. Human Reflection Extraction

So far, we have eliminated most multipath reflections. However, in a typical indoor environment, there are stable reflectors such as walls, floors, and tables, which may be positioned at some places where can reflect RF signals in a stable manner. The signals reflected from a human body are still mixed with such reflections. WiImage has to extract the human body reflection for further processing. Before we describe how our system separates a human body from other reflectors, let’s take a look at the RF signal propagation model.

**RF Signal Fading Model.** According to wireless communication principles [2], the power fading is mainly related to the propagation fading and reflection fading. In an indoor environment, the power of received signals can be expressed as follows:

\[
P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^\nu L} \tag{11}
\]

where \( P_r(d) \) is the received signal power affected by the propagation distance \( d \). \( P_t \) is the transmitting power, \( G_t \) and \( G_r \) are the antenna gains of TX and RX respectively. \( L \) is a constant about the system loss. \( n \) is the path loss fading exponent. Generally, \( n \) varies from 2 to 5. Its value depends on the distinctive feature of an indoor environment, which can be learned by a few tests. Under the condition that the received RF signals are reflected by an obstacle, the receiving power attenuates with the power of 4 to distance \( d \).

Different materials have different abilities to reflect RF signals, and the main influencing factor is the permittivities of materials. The typical materials of an indoor office include plywood, glass, plasterboard, marble and cement. Compared with water, which is the main component of human body, the permittivities of these materials are much smaller. The permittivity of water is 50, and other materials’ permittivities are listed in Table I. The bigger the permittivity is, the ability of reflecting RF signals is stronger. Therefore, chairs, tables, or plasterboard walls can only reflect a small part of RF signals. Most signals continue to propagate passing through the obstacles that are made of the above materials. If we set an antenna array to a target human body at a certain distance, the strength of the body reflection is likely to be equal to the strength of signals reflection off other stable reflectors.

| Material | Human | Concrete | Wood | Glass |
|----------|-------|----------|------|-------|
| \( \epsilon_r \) | 50    | 2.0      | 2.5  | 6.4   |

Based on above observation, WiImage set the distance from the human to the RX antennas smaller than any other stable reflectors to the RX antennas. With this setup, the reflections from a human body are easy to extract by utilizing the signal strength. During each measurement, the signals reflected from a human body persists stably and its power is greater than the reflections from other reflectors. We use the Principal Component Analysis (PCA) technique to extract the principal components from the series \( \hat{P} \) so as to reduce the effect of the noise from different reflectors.

To extract the human reflection, we keep the first ten PCA components from the series \( \hat{P} \) and discard the rest, which are noise components. However, there is only one component corresponding to the human reflection, we need to find the correct component from these candidates. The correct component that corresponds to the human reflection has the following properties: 1) the strength is greater than other components; 2) the angle of the peak is in \( \Psi_1 \). Based on the properties, WiImage can choose the proper PCA component that represents the human reflection. We denote \( \tilde{P} \) as the selected PCA component. So far, We have removed the interference and preserved the human reflection.

### E. Contour Detection

According to the human model in Through-Wall Radar Imaging, the main component of human reflections is caused by torso centre. The signal strength corresponding to the contour of human body in \( \tilde{P} \) is lower than the torso centre but still higher than the background environment. One important observation is that the signal attenuation of the contour position is significantly faster than that of the center position.

**TABLE I**

The dielectric coefficient of common materials

| Material | Human | Concrete | Wood | Glass |
|----------|-------|----------|------|-------|
| \( \epsilon_r \) | 50    | 2.0      | 2.5  | 6.4   |
These parameters should be adjusted for better performance of the contour extraction. Automatically tuning the parameters using the Modified Varri Method and the Savitzky-Golay filter can be accomplished by applying a genetic algorithm, which is key to achieve better performance [25]. Since the genetic algorithm is heuristic, there is no guaranteed time complexity. In our system, we tune above parameters experimentally to ensure the real-time processing.

A human head is mainly composed of non-fat components including bones and blood, and a human torso is composed of fat and muscle tissues. Therefore, we choose two sets of parameters, which are used to detect the contours of the head and torso separately. By searching for the local maximum of $E(\alpha)$, we can extract the contours easily. The human body horizontal width $W$ can be derived by the following equation:

$$W = h \times \Delta \Theta$$

where $h$ is the vertical distance between the human and the antenna array, $\Delta \Theta$ is the AoA change of the left and right contours.

In addition, human body is a scatter, not all reflections from a body can be received. The horizontal width derived from Eqn. (14) only represents a part of the human body where reflections are received. The real horizontal width $W_r$ should be greater than $W$:

$$W_r = \sigma \times h \times \Delta \Theta$$

where $\sigma$ is the magnification between $W_r$ and $W$. According to the radar theory, the radar reflection coefficient of the same kind of objects is a constant. For human bodies, the coefficient varies between $-4$ and $0$ dBm. When receiving the body reflections, the radar reflection coefficient is $0$ dBm. WiImage chooses a typical value of radar reflection coefficient in indoor environment, i.e. $\sigma = 2$.

F. Human Image Splicing

The previous approach can obtain a horizontal part of a body with respect to the linear antenna array. To profile the whole surface of a human body, we move the antenna array vertically with an interval of $2$ cm to emulate a two-dimensional antenna array that allows 3D sensing. While moving the antenna array, WiImage extracts the reflections and determines the human body contours at each height. Then, integrating the pieces of contours obtains a fine-grained human RF image.

Because a human body is a rigid object, the contour should be continuous and smooth. We use the following smoothing method to improve the contour:

$$C_i = \frac{C_{i-2} + C_{i-1} + C_i + C_{i+1} + C_{i+2}}{5}$$

where $C_i$ is $i^{th}$ data in the contour data series.

IV. IMPLEMENTATION AND EXPERIMENT SETUP

We implemented our prototype on an off-the-shelf Intel 5300 WiFi NICs with 3 antennas. The system consists of a receiver with the Intel 5300 card and a transmitter using a Tenda router. The distance between TX antennas and RX antennas is 0.6 m. The transceivers are using linear antenna
arrays. The RX antennas are separated with an interval of half wave-length. We use the Linux CSI Tool [23] to obtain the channel states information of WiFi. We choose 5 G WiFi band, rather than 2.4GHZ, because the signal with a higher frequency attenuates much faster. Thus, we can reduce the influence of multipath and improve the performance of our system. The software algorithms, including the scanning, reconstruction and recognition algorithms, are implemented on a laptop running Ubuntu 14.04 with i7 processor and 8GB of RAM.

To demonstrate the effectiveness of WiImage, we conduct experiments in two different environments. The first is a 8 times5 m² office, containing furnitures in front of concrete walls. The second is a 3 × 5 m² corridor outside our lab. The corridor has concrete walls and floors. In each experiment, we move the antenna array in a bottom-top manner until the whole body is scanned. WiImage requires the target human standing in front of the antenna array with a distance of 0.6 m, which is smaller than the distance from most obstacles to the antenna array.

V. PERFORMANCE EVALUATION

WiImage delivers two sets of functions: the ability to capture the human figure and the ability to profile the human's contour.

A. Human Contour Profiling

1) AoA Estimation Accuracy: We first evaluate the accuracy of AoA estimation by WiImage. We use a digital protractor to provide the ground truth of the direct-path AoA between the human body and the receiver.

In this experiment, we set the distance between the transmitter and the receiver to be 1.5 m, which is smaller than the distance from most obstacles to the receiver. We compare with the state-of-the-art indoor AoA estimation algorithm, i.e., SpotFi [1]. Fig. 7 plots the CDFs with respect to the error of direct-path AoA estimation. In the office, our algorithm achieves a median error of 5.3 degrees, while SpotFi has a median error of 8.7 degrees. In the corridor, our algorithm has a median error of 3.4 degrees while SpotFi’s median error is 5.5 degrees. WiImage works better in the corridor because the multipath reflections are fewer in the corridor.

2) Contour Estimation Accuracy: Because a human body is a scatter, the AoAs of the reflections vary within a certain range that depends on the human body’s width. Here we evaluate the contour extraction accuracy to show the effectiveness of our contour detection algorithm.

Fig. 8 shows the contour estimation accuracy for two scheme in the corridor and the lab office. Our algorithm achieves a median error of 4cm and 6cm in different scenarios. From Fig. 8, we can see that our width estimation algorithm works better in the corridor than in the lab office. In the corridor, there is noting behind the human body so that the reflections extracted by WiImage are only caused by the body. In the lab office, there are walls, windows and chairs behind the human body. A part of signals reflected by these objects transmit through the human body are received by RX antennas finally.

3) Width Estimation Accuracy: In this subsection, we answer the following two questions: First, what is the accuracy of width estimation? Second, what is the performance of width estimation under varying conditions? To answer above questions, we evaluate the performance of WiImage on 5 persons with distinct body shapes. Fig. 9 illustrates the width estimation errors with respect to different human targets. Specifically, the weights and heights of the five persons are ID1 (an overweight boy whose weight 85 kg and height 170 cm), ID2 (a boy whose weight 75 kg and height 174 cm), ID3 (a boy whose weight 63 kg and height 168 cm), ID4 (a slim girl whose weight 45 kg and height 158 cm), ID5 (an adult whose weight 72 kg and height 173 cm). WiImage performs well for all the persons. The average errors are 8.5 cm and 12 cm in the corridor and in the lab office, respectively.

To examine the accuracy impact of the CSI series length (the amount of CSI measurements), we increase the measurement time from 10 s to 60 s with a step size of 10 s. As shown in Fig. 10, the average width estimation errors become smaller with the increasing CSI series length. The intuition is that, with more measurements, the noise are easier to be filtered out so as to obtain the stable human reflections by PCA analysis.

B. Human Image Capture

To fully demonstrate the ability of WiImage, we now evaluate the quality of the constructed human RF images as well as the contours of the human body. In addition, we also show the human identification accuracy using such RF images.

1) Human Feature Extraction: To understand how the RF images relate to a human’s height and body shape, Fig. 11 shows the RF heat-maps and contours of two subjects. We can see that WiImage captures the height of a person. The upper contours of the contours are around the top of the heads of human A and B. We can see that the height of subject ID1 is higher than subject ID4, which is true that their heights are 174 cm and 158 cm, respectively. From the contours and RF heat-maps, we can go further to distinguish the body shapes of these two subjects. Fig. 11 shows that the width of human ID1 is much wider than human ID4. In truth, human ID1 has a standard body shape while human ID4 is a slim girl.

Whole Body Surface Area (WBSA) [18] is a significant indicator of anesthesia usage in clinical surgery and accurate calculation of the WBSA is a topic that has been actively studied over the last century. The method that is widely used for determining WBSA makes use of a well know formulation with two parameters, weight and height:

\[ S_A = W \times 0.0061 + H \times 0.0124 - 0.0099 \]  

WBSA is highly relative to human’s height and weight. From the contour of a human body, we can obtain the height and the body surface area of a human by integral calculation. Based on Eqn. (17), we acquire the weight of a target human from RF images. Table II shows the experiment results of weights, heights and BSAs. The average error of the estimated weights is 10.1kg, greater than the ground truth. In terms of the estimated BSAs, the average error is 0.126 m². The human’s features, including weight, height and BSA, can be extracted precisely from the RF images (heat-maps and contours).
2) Human Identification: In this subsection, we would like to gain a deeper understanding of WiImage’s RF images. We evaluate whether these human RF images can identify people. Here we conduct experiments with 5 participants in the corridor for 5 times. We divide our data into a training set and a testing set. In particular, five images selected randomly from the experimental results are used for training and the remaining images are used for testing.

To obtain the feature vectors for classification, we normalize the 2-D human figures to a 1-D feature vector by concatenating the rows. We add the heights, weights and BSAs of participants to a feature vector. Then we use KNN classification algorithm to identify them. Since the number of samples for each person is only five, we set $k = 3$. The classification algorithm is running in MATLAB with the RF images generated from WiImage. Then we repeat the identification processing for 100 times.

Table III shows the classification confusion matrix whose rows represent the actual labels and columns represent the classified results. When WiImage classifies ID1 and ID4, the accuracies are 94.2% and 96.1% respectively. We notice that the accuracies of ID1 and ID4 are higher than other participants. This is because ID1 is overweighted while ID4 is slim. The body shapes of ID1 and ID4 are distinct from other participants, making them easy to be identified. In addition, ID2 is always confused with ID5. This is because ID2 and ID5 have similar weights and heights. The identification accuracy of ID2 and ID5 are 82.7% and 80.4% respectively.

VI. RELATED WORKS

A. Device-based Sensing System

Past works for gesture recognizing and motion tracking usually operates in direct line-of-sight scenarios or instruments users with markers. The recognition systems based on cameras or infrared, like Kinect and Leap Motion, require a direct line-of-sight from people to sensing devices. The system performance will degrade severely when the surrounding is dark or smoky. Device-based solutions place various kind of markers on the target, using RF, infrared, acoustic, or ultrasonic sensors to track these markers. RF-IDraw [26] that enables a virtual touch screen requires attaching a special RFID tag on the user’s finger. Wang et al. [27] enables human activity recognition but needs to wear an RFID reader on the user’s body. WiCapture [11], Rover [3] and WINS [8] require the user to hold an active WiFi radio to enable the WiFi sensing. Such instrumentations of users hinder these approaches from widely deployments. However, WiImage provides a fine-grained imaging capability in a device-free manner.

B. Device-free Sensing System

Human bodies and human motions will induce modulated RF signal fluctuation remarkably. Past works leverage this phenomenon to sense human gestures, activities and vital signs [2], [15], [22], [28], [29], [30], [31], [32], [33], [34]. The main difference between past proposals is to use various RF signals to conduct the sensing. Radar systems are the first to analyze the signal reflection to detect and track objects. Some can draw the terrain profiles. Inspired by radar techniques, many works utilize Doppler effect and FMCW signals to obtain human gait, respiration rate, and human emotion status. Although these methods can achieve a high accuracy, they require specialized hardware such as ultra-wideband signal
generator to sweep from 5.46 to 7.25 GHz or dedicated USRPs to receive and process RF signals. Gesture recognition and motion tracking based on WiFi is an emerging field that attracts much attention in both industry and academia. By analyzing the WiFi CSI, LiFS [2] and IndoTrack [31] are able to localize a target in a sub-meter level accuracy, even without pre-training. Besides, CSI measurements can also be used for gesture recognition and motion tracking. WiFinger [32] leverages subtle changes of CSI to capture finger movements for fine-grained gesture recognition and this system can achieve 93% recognition accuracy. WifiU [15] extracts human gait features, including walking speed, gait cycle time, and footstep length, from CSI spectrograms using Doppler radar techniques. Then it uses gait features to recognize people. Liu et al. [22] can extract respiration information of a person successfully even target is in different posture.

VII. CONCLUSION

We presented WiImage, a fine-grained human imaging system with commodity WiFi support. WiImage consisted of a lightweight multipath resolution algorithm and a human-contour imaging algorithm. WINS is low-cost, real-time and instantly deployable. It uses WiFi as a new device-free sensing modality to provide privacy-preserving security services such as surveillance and inspection. We implemented WiImage using off-the-shelf WiFi cards equipped with a three-antenna linear array. The real-world experiments demonstrate WiImage’s effectiveness in human contour imaging and human identification. We believe that extending our system to simply the operations for 3D imaging is an important task for future work.

Authors’ contributions. Shengkai Zhang is the principal author who has performed this research work, including algorithm design and implementation. Shaohua Wan is the corresponding author who shepherd the manuscript. All the authors contributed to technical writing and evaluations.

DECLARATIONS

Conflicts of interest. Authors have no conflict of interest.

Ethics approval. ‘Not applicable’.

Consent to participate. All authors participated in this study with their consent.

REFERENCES

[1] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, “Spotfi: Decimeter level localization using wifi,” in ACM SIGCOMM Computer Communication Review, vol. 45, no. 4, 2015, pp. 269–282.

[2] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie, “Lifs: low human-effort, device-free localization with fine-grained subcarrier information,” in Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking. ACM, 2016, pp. 243–256.

[3] S. Zhang, W. Wang, S. Tang, and T. Jiang, “Robot-assisted backscatter localization for iot applications,” IEEE Transactions on Wireless Communications, vol. 19, no. 9, pp. 5807–5818, 2020.

[4] H. Jiang, J. Li, P. Zhao, F. Zeng, and A. Iyengar, “Location privacy-preserving mechanisms in location-based services: A comprehensive survey,” ACM Computing Surveys (CSUR), vol. 54, no. 1, pp. 1–36, 2021.
[5] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, “Understanding and modeling of wifi signal based human activity recognition,” in Proceedings of the 21st Annual International Conference on Mobile Computing and Networking. ACM, 2015, pp. 65–76.

[6] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, “E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures,” in Proceedings of the 20th annual international conference on Mobile computing and networking. ACM, 2014, pp. 617–628.

[7] Y. Wang, K. Wu, and L. M. Ni, “Wifail: Device-free fall detection by wireless networks,” IEEE Transactions on Mobile Computing, vol. 16, no. 2, pp. 581–594, 2017.

[8] S. Zhang, W. Wang, and T. Jiang, “Wi-fi-inertial indoor pose estimation for microaerial vehicles,” IEEE Transactions on Industrial Electronics, vol. 64, no. 5, pp. 3331–3430, 2017.

[9] S. Zhang, W. Wang, N. Zhang, and T. Jiang, “RF backscatter-based state estimation for micro aerial vehicles,” in Proc. IEEE INFOCOM, 2020.

[10] S. Zhang, S. Tang, W. Wang, T. Jiang, and Q. Zhang, “Conquering textureless with rf-referenced monocular vision for mav state estimation,” in Proc. IEEE ICRA, 2021.

[11] M. Kotaru and S. Katti, “Position tracking for virtual reality using commodity wifi,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[12] J. Wang, J. Xiong, H. Jiang, X. Chen, and D. Fang, “D-watch: Embracing bad multipaths for device-free localization with cots rfid devices,” IEEE/ACM Transactions on Networking, vol. 25, no. 6, pp. 3559–3572, 2017.

[13] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, “Capturing the human figure through a wall,” ACM Transactions on Graphics (TOG), vol. 34, no. 6, p. 219, 2015.

[14] J. Wang, J. Xiong, X. Chen, H. Jiang, R. K. Balan, and D. Fang, “Tagscan: Simultaneous target imaging and material identification with commodity rfid devices,” in Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking. ACM, 2017, pp. 288–300.

[15] W. Wang, A. X. Liu, and M. Shahzad, “Gait recognition using wifi signals,” in Proc. ACM UbiComp. ACM, 2016, pp. 363–373.

[16] R. Abdelnasser, K. A. Harras, and M. Youssef, “Ubibreath: A ubiquitous non-invasive wifi-based breathing estimator,” in Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing. ACM, 2015, pp. 277–286.

[17] D. Huang, R. Nandakumar, and S. Gollakota, “Feasibility and limits of wi-fi imaging,” in Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems. ACM, 2014, pp. 266–279.

[18] E. A. Gehan and S. L. George, “Estimation of human body surface area from height and weight.” Cancer chemotherapy reports, vol. 54, no. 4, pp. 225–235, 1970.

[19] R. Schmidt, “Multiple emitters location and signal parameter estimation,” IEEE Trans. Antennas Propag., vol. 34, no. 3, pp. 276–280, 1986.

[20] K. Qian, C. Wu, Z. Yang, Y. Liu, F. He, and T. Xing, “Enabling contactless detection of moving humans with dynamic speeds using csi,” ACM Transactions on Embedded Computing Systems, vol. 17, no. 2, 2018.

[21] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, “Pads: Passive detection of moving targets with dynamic speed using phy layer information,” in Proc. IEEE International Conference on Parallel and Distributed Systems (ICPADS), 2014, pp. 1–8.

[22] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, “Contactless respiration monitoring via off-the-shelf wifi devices,” IEEE Transactions on Mobile Computing, vol. 15, no. 10, pp. 2466–2479, 2016.

[23] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, “Tool release: Gathering 802.11n traces with channel state information,” ACM SIGCOMM CCR, vol. 41, no. 1, p. 53, 2011.

[24] K. Kosar, L. Lhotsk, and V. Krajca, “Classification of long-term eeg recordings,” in Biological & Medical Data Analysis, International Symposium, 2004.

[25] Y. Zou, X. Jiang, J. Han, K. Wu, L. Yun, and L. M. Ni, “GRFID: A device-free rfid-based gesture recognition system,” IEEE Transactions on Mobile Computing, vol. 16, no. 2, pp. 381–393, 2017.

[26] J. Wang, D. Vasisht, and D. Katabi, “Rf-idraw: Virtual touch screen in the air using rf signals,” ACM SIGCOMM Computer Communication Review, vol. 44, no. 4, pp. 235–246, 2014.

[27] L. Wang, T. Gu, X. Tao, and J. Lu, “Toward a wearable rfid system for real-time activity recognition using radio patterns,” IEEE Transactions on Mobile Computing, vol. 16, no. 1, pp. 228–242, 2017.

[28] K. Joshi, D. Bharadia, M. Kotaru, and S. Katti, “Wideoe: Fine-grained device-free motion tracing using rf backscatter,” in Proc. USENIX NSDI, 2015, pp. 189–204.

[29] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, “Widear: Enabling hands-free drawing in the air on commodity wifi devices,” in Proc. ACM MOBICOM, 2015, pp. 77–89.

[30] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, “Non-invasive detection of moving and stationary human with wifi,” IEEE Journal on Selected Areas in Communications, vol. 33, no. 11, pp. 2329–2342, 2015.

[31] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, “Indotrack: Device-free indoor human tracking with commodity wi-fi,” Proc. ACM UbiComp, vol. 1, no. 3, 2017.

[32] S. Tan and J. Yang, “Wifinger: Leveraging commodity wifi for fine-grained finger gesture recognition,” in Proc. ACM MobiHoc, 2016, pp. 201–210.

[33] L. Fan, T. Li, R. Fang, R. Hristov, Y. Yuan, and D. Katabi, “Learning longterm representations for person re-identification using radio signals,” in Proc. IEEE CVPR, 2020.

[34] M. Zhao, T. Li, M. Abu Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, “Through-wall human pose estimation using radio signals,” in Proc. IEEE CVPR, 2018.