IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Extracting Human Context through Receiver-End Beamforming

SAMEERA PALIPANA, STEPHAN SIGG
Department of Communications and Networking, Aalto University, Finland (e-mail: firstname.lastname@aalto.fi)
Corresponding author: Sameera Palipana (e-mail: sameera.palipana@aalto.fi).

'This work was funded by CHIST-ERA III, RadioSense.

ABSTRACT

Device-free passive sensing of the human targets using wireless signals have acquired much attention in the recent past because of its importance in many applications including security, heating, ventilation and air conditioning, activity recognition, and elderly care. In this paper, we use receiver-side beamforming to isolate the array response of a human target when the line of sight array response is several magnitudes stronger than the human response. The solution is implemented in a 5G testbed using a software-defined radio (SDR) platform. As beamforming with SDRs faces the challenge to train the beamformer to different azimuth angles, we present an algorithm to generate the steering vectors for all azimuth angles from a few training directions amidst imprecise prior information on the training steering vectors. We extract the direction of arrival (DoA) from the array response of the human target, and conducting experiments in a semi-anechoic chamber, we detect the DoAs of up to four stationary human targets and track the DoA of up to two walking persons simultaneously.

I. INTRODUCTION

Device free human sensing has gained interest because of its importance in many applications that includes human safety in industry settings [1], security and surveillance [2], heating, ventilation and air conditioning [3], and assisted living [4]. Diverse technologies have been considered for human sensing: cameras [5], infra-red [6], visible light [7] acoustics [8] or wearables [9]. Compared to these solutions, device-free wireless sensing has the advantages of reduced intrusion of privacy, performance in darkness and occlusion and not having the inconvenience of wearing a device.

Device-free passive human sensing is a challenging task because the signals scattered off the humans are several magnitudes weaker than that of the line of sight (LoS) signal. Previous attempts have looked at signal strength [10], direction of arrival (DoA) [11], time of flight [12] and Doppler shift [13], [14] or a combination of those [15] to extract the parameters of the channel for activity recognition, presence detection or localization and tracking. In this work, we use DoA as the channel parameter but extract the DoA from a novel approach. More specifically, we use beamforming at the receiver end and steer the beam to different directions to construct the array response to detect the human presence and to estimate the direction of that person. Unfortunately, current super-resolution algorithms that detect the DoA, e.g. multiple signal classification (MUSIC) [16], require the number of sources as prior information for an accurate estimation of the DoAs [17]. Using beamforming, such problems can be mitigated as the DoAs of the interested targets can be simultaneously estimated from the amplitude of the array response using our approach. Additionally, by focusing the beam towards the direction of the person, activity patterns can be simultaneously recognized for multiple persons [18]. However, beamforming has not been utilized for human sensing in the past mainly because of the requirement of transmitter (TX) and receiver (RX) phase synchronization and the requirement of large antenna arrays for high resolution.

Recent developments in 5G with massive MIMO have enabled beamforming with high resolutions. In our work, we utilize a 5G testbed [19] comprised of software-defined radios (SDRs). Beamforming with SDRs is still challenging because the beamformer has to be calibrated for different azimuth angles due to practical limitations like unknown wave propagation conditions and random phase shifts in RF chains. Moreover, training the beamformer for a larger number of azimuth angles is labor intensive and is not practical. We, therefore, use reference directions to train the beamformer

by transmitting from a limited number of azimuth angles and measure the steering vectors at the receiver for those directions using the relative phases of the antenna array. The measured steering vectors can still be far from ideal due to, e.g. unknown wavefield propagation conditions and antenna array orientation and geometry errors. Therefore, we sanitize these steering vectors and using them as reference points, we generate the steering vectors for all azimuth angles using a mathematical model.

As contributions, i) we propose an algorithm to estimate steering vectors for all azimuth angles using the reference points even when steering vector mismatches occur. ii) We also provide an approach to detect the human presence from the array response, and iii) a method to estimate the direction of the human when the direct path from the transmitter is several magnitudes stronger than the scattered paths from the human. iv) By conducting experiments in a semi-anechoic chamber of size 22.4m^2, we achieve a 100% detection accuracy and a maximum error for the median direction of arrival of 7° for a person performing in-place activities in five different locations of the area of interest. We perform experiments placing up to four persons in the given area and show that DoAs can be extracted from stationary persons when sufficient spatial separation among them is present. Finally, conducting experiments up to two walking persons we show that their DoAs can be simultaneously extracted.

The rest of the paper is organized as follows. Section II provides a summary of the most relevant device-free human sensing approaches. Section III introduces our signal model used for beamforming and explains how we construct the array response through beamforming using SDRs. In Section IV, we provide our solutions to problems that occurred due to practical limitations associated with SDR-based beamforming i) fixed phase offsets in RF chains, and ii) steering vector mismatch. In Section V we detect human presence from array response, extract the response of the human and the corresponding DoA. In Section VI we explain the configuration of our testbed, Section VII explains the experiment environment, data collection and evaluation of our algorithms. We discuss the scope and limitations of this work and provide our conclusions in Section VIII.

II. RELATED WORK

Our work is related to Device-free localization, tracking and activity recognition of humans using radio signals. Other competing technologies include vision-based solutions [5], thermal images [6], visible light [7] acoustics [8] or wearables [9]. Vision or optoelectronic sensors used for people monitoring exploit the properties of reflectivity with time-of-flight (ToF) cameras, emissivity with thermal images or acoustic sensors. They are, however, constrained by range, occlusions, environmental conditions (smoke/fire), and privacy concerns [20], [21]. Other solutions, which demand that people carry a device (Device-based), limit human subjects' mobility, visibility, and communication. Device-free RF-sensing solutions, which do not require people to wear any devices, and which rely on electromagnetic propagation, mitigate these challenges. The proliferation of radio devices inside the area of interest is therefore expected to boost the development of non-intrusive RF based localization, tracking and activity recognition [22], [23].

Previous Research has demonstrated the use of radio signal measurements for the inference of different characteristics of the human state. These measurements include time delay [24], phase [11], Doppler [25], and signal strength [26]. They have been used for various purposes including vital sign monitoring [27], activity and gesture recognition [28], localization [29], [30], tracking [31], gait identification [32], fall detection [4], [33], gesture control [34], and indoor human walking direction estimation [35].

Among the most relevant approaches that perform device-free localization and tracking, WiTrack [12] uses a custom made transceiver exploiting a wideband FMCW radar to estimate time delay of signals reflected from a human and a 5×5 antenna array with more than half wavelength separation between two elements to achieve spatial diversity. WiDeo [36] uses WiFi backscatter communication to jointly estimate ToFs and AoAs to localize humans. DynamicMusic [11] jointly estimates DoA and observed time of flight (reflected off human subjects) from channel state information (CSI) by a two dimensional MUSIC algorithm using commodity WiFi devices. However, the approach successfully detects only a walking person, whereas IndoTrack [31] extends this by incorporating AoA with Doppler information to localize a static and dynamic person. xd-Track [15] applies SAGE algorithm for multiple channel parameter estimation (signal strength, DoA, ToF and Doppler) using an SDR platform, thereby improving the resolution of multiple human target tracking despite heavy computation efforts and custom made hardware. Widar2.0 [37] extends this to commercial WiFi devices that uses CSI and further leverages multipath effects to enable absolute ToF estimation.

These approaches are relevant to our work in device-free estimation of DoA for human tracking, however, our approach is different to all these in the way DoA is estimated. While current state of the art approaches use either an extended version of MUSIC algorithm or a variation maximum likelihood estimation using the relative phases of antenna arrays, we use beamforming techniques and construct the array response to isolate the path corresponding to humans and estimate their DoAs. We show that the approach can estimate DoAs of multiple targets without the knowledge of the target count in advance and the target resolution is only limited by the number of antennas.

III. BEAMFORMING PRIMER

A. SIGNAL MODEL

Here we use a point source narrowaband signal model and a uniform linear array (ULA) of M omnidirectional antenna elements. The narrowband signal received at time instant t by

the ULA is modeled as

$$\mathbf{x}(t) = \mathbf{s}(t) + \mathbf{i}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{s}(t)$, $\mathbf{i}(t)$ and $\mathbf{n}(t)$ are $\mathbb{C}^{M\times 1}$ vectors representing the line of sight (LoS) signal, signal scattered off a person and noise respectively. Under these conditions $\mathbf{i}(t)$ is the desired signal and it can be amplified by finding the correct beamformer weight $\mathbf{w}$. However, as the person is not always stationary, we cannot use a fixed weight as in beamforming used for fixed transmitter and receiver positions. Therefore, we construct the array response using the beamformer weights for all azimuth directions.

B. ARRAY RESPONSE OF A DELAY AND SUM BEAMFORMER

In this work, we use a 5G testbed (Sec. VI) consisting of software defined radios (SDRs). In our testbed, a 12.48 MHz channel is divided into 52 OFDM subcarriers each having a bandwidth of 240 kHz. In OFDM, a channel is divided into multiple subcarriers and data is modulated in each subcarrier. The in-phase and quadrature (IQ) samples of the 52 subcarriers measured at runtime are fed to the beamformer as a post-processing step to obtain N spatial streams of the array response such that signals arriving from direction $i \in N$ are amplified at the i^{th} spatial stream.

Here, we implement a delay and sum beamformer using OFDM beacons. Even though, other commonly used beamformers like minimum variance distortionless response (MVDR) or linearly constrained minimum variance (LCMV) beamformers improve the signal to interference plus noise performance, they achieve this by having the nulls of the beam at interference directions [38]. This can be detrimental for human sensing, as we rely on the signals scattered off the humans behaving as the interference paths.

We implement the beamformer as follows. The transmitter sends a beacon with a known sequence $a(t)$ each subframe. IQ samples $x_m(t) \in \mathbb{C}$ are received by antenna element $m \in M$ where M is the number of receiver antenna elements. $x_m(t)$ can be represented as

$$X_m(f) = H_m(f) \cdot A(f) \tag{2}$$

where $X_m(f) = FFT\{x_m(t)\}$, $A(f) = FFT\{a(t)\}$ is the transmitted beacon and $H_m(f)$ is the wireless propagation channel between TX and receiving antenna m. Radio waves arrive at a receiver over multiple paths due to reflection and scattering caused by objects in the environment. A plane wave arriving from direction j introduces a phase rotation of $\phi_j^m = \frac{2\pi c \tau_j^m}{\lambda}$ due to the time delay τ_j^m to propagate to antenna element m where c is the speed of light and λ is the wavelength. We estimate $\tilde{\phi}_j^m \approx \phi_j^m$ by computing the cross-correlation $r(\phi_m)$,

$$r(\phi^m) = FFT^{-1}\{X_m(f) \cdot A(f)^*\} \tag{3}$$

and $\tilde{\phi}_j^m$ maximizes the correlation between transmitted and received beacons

$$\tilde{\phi}_j^m = \underset{\phi_m}{\text{argmax}} \ |r(\phi^m)|. \tag{4}$$

This approach is possible because a truncated version of frequency domain Zadoff-Chu sequence used in LTE is adopted here as $A(f)$ [39]. The sequence was truncated because it was longer than the number of available subcarriers.

The beamformer weight $\mathbf{w}_i$ is computed using $\mathbf{w}_i = \mathbf{p}_i/M$ where $\mathbf{p}_i$ is the steering vector for azimuth direction $i \in N$. We apply the weight vector on $\mathbf{r_j(t)}$ at time instant t

$$a_i(t) = \mathbf{w_i} \cdot \mathbf{r_j(t)} \tag{5}$$

where $\mathbf{r_j(t)} = [r(\tilde{\phi}_j^1, t), r(\tilde{\phi}_j^2, t), \ldots, r(\tilde{\phi}_j^M, t)]^T$ to obtain array response $\mathbf{a}(t)$ where $\mathbf{a}(t) = [a_1(t), a_2(2), \ldots, a_N(t)]$.

IV. PRACTICAL IMPAIRMENTS TO BEAMFORMING AND OUR SOLUTIONS

Even though the array response can be computed as above in an ideal setting, practical limitations in real world testbeds introduce errors on the relative phase of signals among the received streams and the steering vectors. To tackle these problems, we implemented calibration techniques.

A. PHASE OFFSETS OF THE RF CHAINS

The problem of time-varying phase-offset due to clock drift is solved by using a separate transmitter that sends reference signals before receivers compute the phase difference between RF chains for run time compensation by [39]. Phase offsets caused by non-identical antenna feed cables, connectors and other RF components are mitigated by first, performing phase measurements of each component via an external vector network analyzer and measuring a crude phase offset for each RF chain. Then, finding the phase offset α_m in each antenna element by using the line of sight (LoS) direction as the reference. If a plane wave impinges on the antenna array when a signal is emitted from the LoS, the relative phases among the antenna elements are almost zero. We use this assumption to calibrate the phase offsets. We first model the measured signal $\hat{x}_m(t)$ as $\hat{x}_m(t) = x_m(t) \cdot e^{j\alpha_m}$ where α_m is the unknown phase offset at antenna element m. Then, we calibrate α_m such that the steering vector $\mathbf{q}_{LoS}$ of the LoS direction is $\mathbf{q}_{LoS} = [0,\ 0,\ 0\ , \ldots 0]_{1\times M}$. This calibration helps us in two ways: i) to mitigate fixed phase offsets in RF chains, and ii) to use the LoS direction as the reference direction for a signal coming from 90° azimuth angle to the ULA.

B. STEERING VECTOR PROCESSING

In this work, we estimate the steering vectors and learn the beamformer weights from a particular azimuth direction by transmitting a signal from known anchor points. In practice, the steering vector associated with a plane wave impinging from a particular direction may not be known precisely. This results in a mismatch between the presumed and the

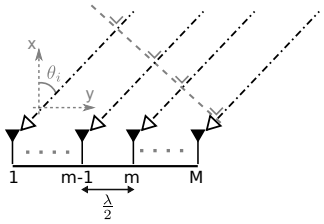


FIGURE 1: A plane wave impinging on a antenna array at an azimuth angle of θ_i.

measured steering vectors. This can occur due to imprecisely known wavefield propagation conditions, antenna calibration errors, antenna array geometry and orientation errors and/or signal pointing errors [40], [41]. Therefore, we choose to first measure the phase rotations from signals transmitted from k reference directions, convert them to steering vectors and assign them as anchor points. Thereby, we derive the directions of the signals arriving directly or from positions near to those anchor points. Next, we use a model of the delay and sum beamformer to sanitize the steering vectors and to assign azimuth angles to those anchor points.

1) Steering Vector Estimation using Anchor Points (calibration phase)

We estimate the steering vector $\mathbf{q_i}$ of receiver antenna elements by placing a transmitter at k anchor points, measure the azimuth angles $\Theta = [\theta_1, \theta_2 \dots \theta_k]$ and use them as reference angles for the k steering vectors. The phase shifts corresponding to the steering vectors for the k reference directions are estimated using the method explained Sec. III. The IQ samples are correlated with the TX beacon as in Eq. 3 and the phase rotation $\tilde{\phi}_i^m$ that maximizes the correlation $r(\phi^m)$ is estimated so that $\mathbf{q_i}$ normalized to the 1st antenna element is

$$\mathbf{q_i} = [1, e^{-j(\tilde{\phi}_i^2 - \tilde{\phi}_i^1)}, e^{-j(\tilde{\phi}_i^3 - \tilde{\phi}_i^1)}, \dots, e^{-j(\tilde{\phi}_i^M - \tilde{\phi}_i^1])} \quad (6)$$

As mentioned above, the steering vector derived from the anchor point is not ideal because of residual phase offsets in RF chains, measurement offsets and orientation errors in the antenna geometry.

2) Steering Vector Sanitation and Array Response Computation

As a solution for steering vector mismatch, we estimate the presumed steering vector using a method explained in [41]. Note that this method is executed for a minimum variance distortionless response (MVDR) beamformer under heavy interference conditions. Even though we use a delay and sum beamformer, this method can be applied to our case because the steering vector is determined by the physical properties of the channel independent of the type of beamformer being used. We adapt their approach to estimate the optimum steering vector as follows. The measured steering vector is modeled as $\mathbf{q}_i = \mathbf{p}_i + \mathbf{e}_i$ where $\mathbf{p}_i$ is the presumed steering vector and $\mathbf{e}_i$ is the error for the anchor point at direction i. Then we estimate the error $\mathbf{e}_i^{opt}$ by maximizing the beamformer output power of an MVDR beamformer $P(\mathbf{e}_i) = \frac{1}{|(\mathbf{p}_i+\mathbf{e}_i)^H \hat{\mathbf{R}}_\mathbf{i}^{-1}(\mathbf{p}_i+\mathbf{e}_i)|}$ [41] as a convex optimization problem

$$\mathbf{e}_i^{opt} = \underset{\mathbf{e}_i}{\operatorname{argmin}} \ |(\mathbf{p}_i + \mathbf{e}_i)^H \hat{\mathbf{R}}_\mathbf{i}^{-1}(\mathbf{p}_i + \mathbf{e}_i)| \quad (7)$$

Once the presumed steering vector is known, it is possible to estimate the corresponding azimuth angle of the anchor point by comparing with a modeled steering vector. We model the steering vector $\hat{\mathbf{p}}_\mathbf{i}$ for direction i as a function of azimuth angle

$$\hat{\mathbf{p}}_\mathbf{i} = [1, e^{-j2\pi f \tau_i^2}, e^{-j2\pi f \tau_i^3} \dots, e^{-j2\pi f \tau_i^M}] \quad (8)$$

where $\tau_i^m = \frac{\lambda}{2} \cdot (m-1) \cdot \sin\theta_i$ and $-90^\circ < \theta_i < 90^\circ$ is the azimuth angle in the xy plane measured from x axis for a signal arriving from the i^{th} direction. By minimizing the minimum mean square error (MMSE) between $\hat{\mathbf{p}}_\mathbf{i}$ and $\mathbf{p_i}$ using azimuth angle as a variable, we are able to obtain the optimum azimuth angle θ_i^{opt} of the anchor point and the corresponding steering vector $\hat{\mathbf{p}}_i^{opt} \approx \mathbf{p}_i$.

$$[\hat{\mathbf{p}}_i^{opt}, \theta_i^{opt}] = \underset{\theta_i}{\operatorname{argmin}} \Big(\frac{1}{m} \sum_m (\hat{p}_{i,m} - p_{i,m})^2 \Big) \quad (9)$$

Finally, we generate steering vectors for all azimuth angles using Eq. 8 and construct the array response using Eq. 5. Algorithm 1 summarizes this process. Fig. 2a and Fig. 2b illustrates the array response computed from a signal emitted for a duration of 15 s from a TX at A and and a TX at C. Fig. 2c illustrate the array response for a person at A and and Fig. 2d for a person at A and at B when the transmitter is at C. From the first two figures it is visible that the array response of a particular azimuth angle is constant. Compared to the first two figures, it can be noticed that there are more perturbations in the array response of the final two figures. Additionally, compared to Fig. 2c, Fig. 2d has more perturbations closer to the LoS response. However, the LoS response is quite strong in both situations which tends to shadow the response of the human.

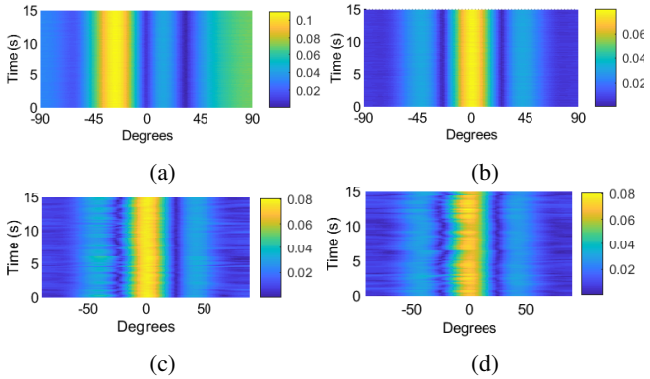


FIGURE 2: (a) Array response for a signal emitted from anchor point A at -28° during 15 s and array response when a signal is emitted from C at 0° (b) no human exists in the environment (c) a human exists at A and (d) a human exists at A and at B.

Result: Array response $\mathbf{a}(t)$
for $i \leftarrow 1$ *to* k **do**
  Estimate $\mathbf{e}_i^{opt}$ by solving (7)
  Update vector $\mathbf{p}_i$ using $\mathbf{q}_i = \mathbf{p}_i + \mathbf{e}_i$
  Estimate $\hat{p}_i^{opt}$ and θ_i^{opt} by solving (9)
end
Generate steering vectors using (8) for $-90 \leq \theta \leq 90$
Construct the array response using (5)

Algorithm 1: Array response computation.

V. DETECTION OF THE HUMAN IMPACT

In this section we analyze the array response in the presence of a human from the spatial streams computed according to Eq. 5 from the steering vectors processed according to Sec.IV-B corresponding to azimuth angles $-90° \leq \theta_i < 90°$.

A. HUMAN PRESENCE DETECTION

We infer human presence from the array response for a given time interval T as follows. We construct the array response $\mathbf{a_{hf}}(t)$ for a human-free setting using Eq.5. $\mathbf{r_j}(t)$ in Eq.5 is computed from the IQ samples collected from the testbed when no human exists. Then we find the difference, Δa_T, between $\mathbf{a_{hf}}(t_0)$ and the array response measured with a human subject present $\mathbf{a_{hp}(t)}$ for all θs during the interval $t_0 \leq t < t_0 + T$ (since the array response for the human free environment is constant over time, we use only the response at time instant t_0).

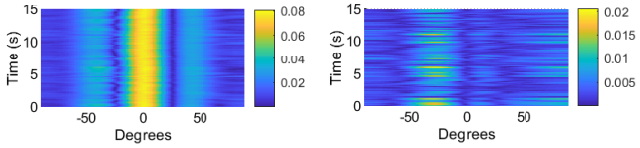
$$\Delta a_T = \sum_{t=t_0}^{t_0+T} \left\{ \frac{1}{181} \sum_{\theta=-90}^{90} |a_{hp,\theta}(t) - a_{hf,\theta}(t_0)| \right\} \tag{10}$$

Human presence is detected when $\Delta a_T > Th$, where Th is the maximum Δa for array responses measured in a human free environment $Th = \max(\Delta a_{T_1}, \Delta a_{T_2} \ldots \Delta a_{T_N})$ during N time intervals. Note that Δa_{T_n} where $n \in N$ is not zero due to noise and multipath effects.

B. DETECTION OF THE ARRAY RESPONSE OF HUMANS

When human presence is detected by the above approach, we remove the effect of the line of sight path from the array response $\mathbf{a_{hp}}(t)$ through subtraction and construct the array response of a human $\mathbf{a_h}(t)$. The intuition here is that the array response of the LoS has a constant gain for a particular angle over time. However, from the collected data we notice that when the human is closer to the LoS beam, the subtraction of a constant value from the array response of human and the LoS does not completely remove the LoS effect over a long duration. Therefore, we estimate the mean array response for a short time interval T and subtract that value from the array response such that

$$\mathbf{a_h(t)} = \mathbf{a_{hp}(t)} - \frac{1}{T} \sum_{t=t_0}^{t_0+T} \mathbf{a_{hp}(t)} \tag{11}$$



(a) (b)

FIGURE 3: (a) Array response of the LoS when a person exists at A. (b) Array response of the person when LoS response is removed.

where $t_0 \leq t \leq t_0 + T$. Fig. 3a illustrates the behaviour of $|\mathbf{a_{hp}}(t)|$ and Fig. 3b shows the behaviour of $|\mathbf{a_h}(t)|$ when a human is present at anchor point A for $T = 15$s. According to Fig. 3a, the LoS path overshadows the paths from the human, however, between azimuth angles $\{-50°, 0°\}$ perturbations can be observed in the array response than between the azimuth angles $\{0°, 90°\}$. When the effect of the LoS path is removed, human's impact can be observed as in Fig.3b. Interestingly, the two figures show that the highest amplitude of the human response is as much as 4 times smaller than the LoS response.

C. ESTIMATION OF A HUMAN'S DIRECTION

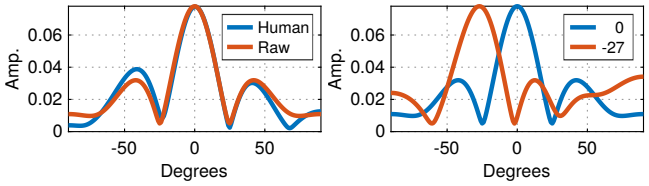
We use the array response of the human $\mathbf{a_h}(t)$ when the LoS response is subtracted to estimate the direction of arrival corresponding to the human. As the beamwidth of the unit response of a signal emitted from a particular direction is fairly wide (half power beam width of $\approx 25°$ for a ULA of 4 antenna elements in our case) and the residual response of LoS signal may still exist, the estimation of the human's direction is not trivial. As a naïve approach, we can detect the angle corresponding to the maximum gain at a particular time instant, however, this introduces noise due to large beamwidth, residuals of the LoS beam and sidelobes. Again, MVDR or MUSIC algorithms cannot be used because the number of sources are not known in advance. Therefore, we use the following procedure to estimate the DoA corresponding to a human.

We first model the array response of the ULA for a signal emitted from each azimuth angle. As an example, the array response $\hat{\mathbf{a}}_j$ of a signal emitted from direction j can be modeled by multiplying the beamformer weight $\mathbf{w}_j$ with all steering vectors of the interested K azimuth angles s.t.

$$\hat{\mathbf{a}}_j = \mathbf{w}_j^H \mathbf{P} \tag{12}$$

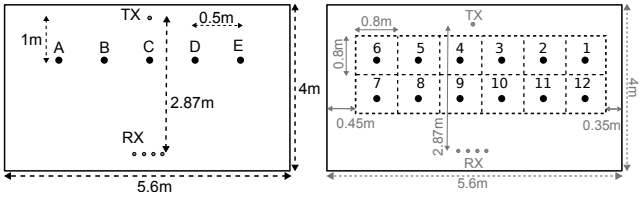
where $\mathbf{P} = [\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \ldots, \hat{\mathbf{p}}_K]$. Figure 4b illustrates the amplitude of the array response $|\hat{\mathbf{a}}_j|$ for a unit source impinging on the ULA from directions $\theta = 0°$ and $-27°$. Likewise we compute $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \ldots, \hat{\mathbf{a}}_J]$ for unit sources impinging on the ULA from azimuth angles $-90 \leq \theta \leq 90$. Then we correlate $|\hat{\mathbf{a}}_j|$ with the array response of the human $|\mathbf{a_h}(t)|$ and find the azimuth angle $\theta_{max}(t)$ that maximizes this correlation for time instant t.

$$\theta_{max}(t) = \underset{j}{\operatorname{argmax}} \{|\hat{\mathbf{a}}_\mathbf{j}|^T \cdot |\mathbf{a_h}(t)|\} \tag{13}$$



(a) (b)

FIGURE 4: (a) Comparison of the amplitude of the array response $|\mathbf{a_{hf}(t)}|$ when environment is human-free vs $|\mathbf{a_{hp}(t)}|$ where a human exists at -27° at a particular time instant. (b) Amplitude of the modeled array response $|\hat{\mathbf{a}}_j|$ for a unit source impinging on the ULA from directions $\theta = 0^\circ$ and -27°.



(a) (b)

FIGURE 5: Experiment Environment. # of RX antennas: 4, # of TX antennas: 1, antenna heights from ground: RX 0.89 m, TX 0.85 m, room size: 5.6 m $\times$ 4 m $\times$ 2.184 m. (a) A to E and (b) 1 to 12 are anchor points.

The idea here is that when the peaks of $|\mathbf{a}_h(t)|$ and $|\hat{\mathbf{a}}_j|$ are aligned for a signal incoming from a certain azimuth angle, this produces a high correlation. Here, we compute the matrix $\hat{\mathbf{A}}$ manually instead of shifting $|\hat{\mathbf{a}}_j|$ over different angles because the array responses for different azimuth angles are not equal to their shifted versions. As an example, according to Fig. 4b, $|\hat{\mathbf{a}}_j|$ for a signal arriving from -27° is not the same as when $|\hat{\mathbf{a}}_j|$ for a signal arriving from 0° is shifted by 27 degrees to the left.

VI. TESTBED DESCRIPTION

We used a software-defined radio platform consisting of a single-antenna transmitter and a receiver with a four element uniform linear array (ULA) with $\frac{\lambda}{2}$ inter-element spacing. Dipole antennas were used as array elements. The transmitter is comprised of a Universal Software Radio Peripheral (USRP) X300 series with UBX-160 and SBX RF-daughterboards as the radio front-ends connected to a host computer. The receiver was comprised of three USRPs, of which two served to collect the signal from the other high-level node (= transmitter). The third device is the reference USRP. It operates by synchronizing over the air with the primary two USRPs and then transmitting a reference signal to calibrate the starting offset between them. Phase coherence between the two receiver USRPs was accomplished using a clock distribution system. It provides both a pulse-per-second signal and a 10 MHz reference signal to discipline the local oscillators of the USRPs. The reference USRP was necessary due to inter-device random phase offset at start-up and component variability in each RF chain. Additionally, the reference USRP helps to ensure that phase coherence, and therefore beamforming performance is maintained throughout the measurement collection.

The air-interface configuration of the testbed is an OFDM frame structure at a carrier frequency of 3.42 GHz. The total bandwidth of the system is 15.36 MHz with 52 subcarriers each with a bandwidth of 240 kHz yielding a total bandwidth of 12.48 MHz. Each USRP stream IQ samples at a rate of 16.66 megasamples per second. This is subdivided into subframes of 3082 samples in length in time domain, yielding approximately a rate of 5408 subframes per second. This rate of the subframes is essentially the sampling rate of the IQ samples used as input for beamforming.

VII. EXPERIMENTS AND EVALUATION

We conducted experiments in an environment as shown in Fig. 5a which is a semi-anechoic chamber where ground reflections are still possible as the ground was not insulated with radio wave absorbents. Each experiment in below sections lasted for 15s. As the testbed provides samples at a rate in excess of 5400 samples/s we downsampled it by factor 10 and obtained 8112 samples for each experiment.

A. EVALUATION OF STEERING VECTOR SANITATION

As mentioned in Sec.IV-B we generate the beamformer weights using a semi-supervised learning approach where measurements are collected from anchor points and then assigning the closest steering vector generated from a model using the MMSE criterion. In this section we analyse the performance of this approach using the MMSE of 17 steering vectors collected during two measurement campaigns as shown in Fig. 5. In the first campaign we used 5 directions (indexed A, B, C, D & E) as shown in Fig. 5a and in the second campaign 12 reference directions (indexed 1-12) were used as shown in Fig 5b. Fig 6 illustrates the MMSE for each direction between the presumed steering vector $\mathbf{p_i}$ and modeled steering vector $\hat{\mathbf{p}}_\mathbf{i}$ together with the azimuth angle (θ_i) computed according to Sec. IV-B2.

For all the directions in both campaigns, except the direction having the index 7 in campaign 2, MMSE<0.2. Generally, the directions having a smaller |azimuth angle| tend to have low errors and vice versa, e.g., directions with indices 5, 6, 7 and 8 have the highest errors while C, D, 3, and 10 have low errors. For the direction indexed 7, the error is 2.65 and the corresponding azimuth angle is -20°. Given that direction 7 should be further away from directions with indices 6 and 8 (having angles of -55° and -52° respectively), it can be concluded that the measured steering vector for direction 7 has a very high error which cannot be sanitized by the proposed algorithm. Therefore estimation of the azimuth angle for that direction is also erroneous. We conjecture that the error is caused by the increase in azimuth angle where the direction with index 7 has the highest azimuth angle. It is well known that higher azimuth angles can introduce errors in DoA resolution.

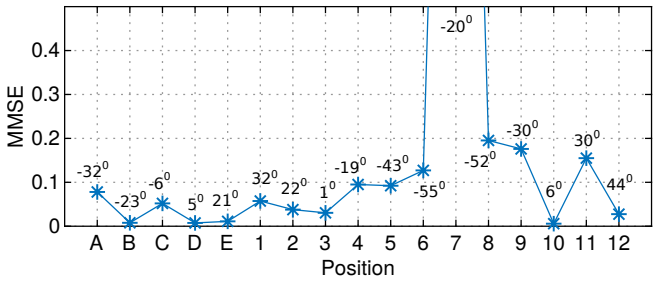


FIGURE 6: MMSE between modeled and presumed steering vectors of 17 directions from two measurement campaigns along with the computed azimuth angles.

B. EVALUATION OF HUMAN PRESENCE DETECTION

We conducted 5 experiments by placing a person in the positions corresponding to the anchor points A, B, C ,D, E, 4 experiments by placing 2 persons at AB, BC, CD and DE, 6 experiments with three persons at ABC, BCD and CDE, 4 persons at ABCD and BCDE and 5 persons at ABCDE. Fig. 7a shows the results comparing Δa of one person with Th corresponding to human-free environment. Similarly, Fig.7b illustrates the behavior of Δa for 2 persons and Fig.7c for Δa of 3,4 and 5 persons. The persons perform in-place random activities during the entire period the measurements are collected. All 15s samples were divided into intervals with a duration of $T = 1$ s. The duration for T was chosen to satisfy human response estimation according to Sec. V-B by comparing the performance of $\mathbf{a_h(t)}$ for different intervals ranging from 0.2 s to 2 s.

Comparing the five single person experiments, it can be observed that Δa is above Th during the entire period. The same can be observed with experiments involving 2-5 persons. The main difference is that the worst case performance of Δa for two persons (DE in Fig. 7b) is higher than that of 1 person experiments and it is even higher for experiments involving $3-5$ persons. This is expected behavior, more perturbations in Δa can occur, since more persons are present.

C. DIRECTION OF ARRIVAL ESTMATION OF A HUMAN

In this section we compare the impact of a person placed at directions A, B, C, D or E i) on the DoAs of the LoS signal and ii) on DoAs corresponding to the single person .

Fig. 8a shows a boxplot representation of the DoAs of the LoS signal when a human subject is at A, B, C, D or E. The median DoA of the LoS from A to E are $0°, 1°, 2°, -2°$ and $0°$. It is clear that LoS DoAs are not affected when the person is furthest from the LoS, e.g. A and E. Additionally, we observe that the outliers of experiments A and E are also low. When the person gets closer to the LoS, the low resolution of the beamformer causes the LoS angles to deviate slightly from the ideal, e.g., B ($-1°$), C ($2°$) and D ($2°$). Fig. 8b shows a boxplot representation of the raw DoAs corresponding to the human derived from the approach described in Sec. V. The median DoAs of the human are $-32°, -16°, -4°, 2°$, and $16°$. Compared to the DoAs of anchor points of A to E shown in Fig. 6, the absolute errors are $0°, 7°, 2°, 3°$ and $6°$ respectively which illustrates the accuracy of our DoA estimation approach. As positions furthest from the LoS have most outliers, we remove these outliers by applying a moving median filter. Finally, Fig.8c illustrates the result when those DoAs are filtered. The filtered DoAs have medians $-32°, -16°, -4°, 2°$, and $16°$.

D. EVALUATION OF THE RESOLUTION OF DOA OF MULTIPLE HUMAN TARGETS

1) Stationary Human subjects with 0.5 m Distance

Here we compare the distributions of the LoS signal with TX at A, B, C, D and E, single human subject at A, B, C, D, and E, two subjects at AB, BC and DE, three subjects at ABC and four subjects at ABCD.

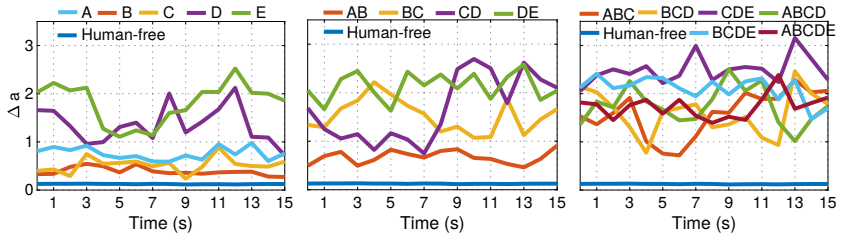
Fig 9a shows the distribution of angles corresponding to the LoS paths, Fig 9b for the single human subject positions and Fig 9c for the two human subjects case. Note that for these three figures, the experiments were performed separately and the results were merged to the three figures for ease of comparison. Figures 9d and 9e show the distributions for 3 and 4 human subjects respectively.

Clearly, the angles for the LoS signals have the most consistency. TX at A, B, C and E have azimuth angles $-32°$, $-23°$, $-6°$ and $20°$ 100% of the time and D at $5°$ around 65% of the time. However, when a human is placed in the same five positions, the distributions show a Gaussian behaviour. The mean angles can be computed as $-32°$, $-16°$, $-4°$, $5°$, and $14°$. This can be attributed to the paths scattering off of different points from the subject's body as opposed to the LoS path not going through such disturbance.

When two persons coexist in the environment as in Fig 9c 0.5 m away from each other, the distributions spread even further resembling a mixture of two Gaussians. However, the DoA corresponding to the two persons cannot be directly inferred from the mean or the peaks of the distributions. The distributions only provide a rough estimate of the area where the two persons are located, e.g. the distribution for AB is shifted to the left of BC and BC's distribution is shifted to the left of DE. When three (Fig. 9d) and four (Fig. 9e) persons coexist 0.5 m away from each other, the distributions are spread even further. Based on this information, it can be concluded that when people are stationary in a given area, the number of persons can be estimated by evaluating the spread of the distribution even though there DoAs cannot be directly estimated.

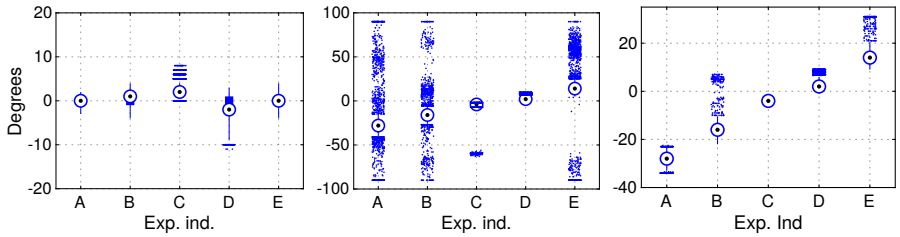
2) Human Subjects with increased Distance

Next, we conducted two experiments placing two persons at positions *2* and *4* marked in Fig. 5b, 1.6 m away from each other and *1* and *6*, 4 m away from each other. Figures 10a and 10b show the distributions of the DoAs for the two experiments. Unlike in the previous case, the distributions for the two humans have a clear separation. For the positions *2* and *4*, the mean DoAs can be estimated as $-26°$ and $25°$. The DoAs of the anchor points for the two positions were estimated



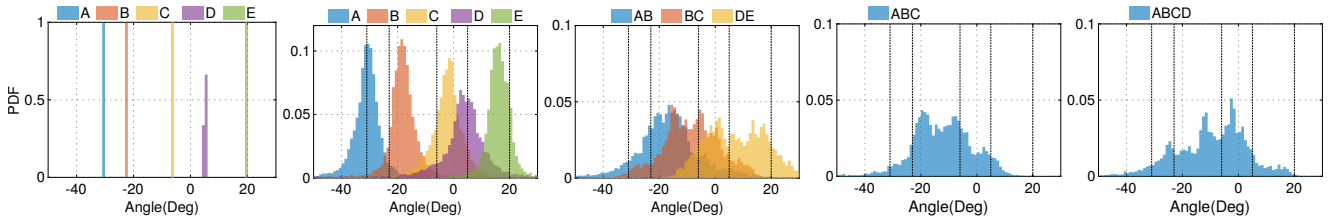
(a) (b) (c)

FIGURE 7: Human presence detection from the array response (a) single person at A, B, C, D and E, compared with human-free case, (b) two persons at AB, BC, CD, DE compard with human-free case, (c) three persons at ABC, BCD, CDE, four persons at ABCD and BCDE and five persons at ABCDE compared with human-free case.



(a) (b) (c)

FIGURE 8: Resuts of the experiments with and w/o human. (a) LoS AoA, (b) AoA of the human, (c) Moving median AoA of the human.



(a) (b) (c) (d) (e)

FIGURE 9: Direction of arrival distribution in different scenarios. (a) No human exists, a TX is placed at A, B, C, D & E in separate measurement campaigns, (b) one human exists at A, B, C, D & E in separate measurement campaigns, (c) two humans exist at AB, BC or DE within 0.5 m in three separate measurement campaigns, (d) three humans exist at ABC 0.5 m from each other and (e) four humans exist at ABCD 0.5 m from each other.

to be $-19°$ and $22°$ which results in a DoA difference of $7°$ and $3°$. As shown in Fig. 10b, again the distributions for positions *1* and *6* has clear separation as expected, however, the resolution is also diminished compared to the former experiment given that the distributions have a larger spread. The peak DoA of *1* is around $45°$ and the peak of *6* is around $-50°$ while the anchor points have DoAs of $32°$ and $-55°$ respectively.

Finally, two experiments are carried out by placing four persons simultaneously. In the first experiment, two of them are at *1* and *2* and the other two are at *5* and *6*, and in the second experiment, two persons are placed at *1* and *12* and another two at *6* and *7* as marked in Fig. 5b. The DoAs for those anchor points were estimated (Fig. 6) as *1*: $32°$, *2*: $22°$, *12*: $44°$, *5*: $-43°$, *6*: $-55°$. Note that the DoA of *7* was erroneous and based on the DoAs of *6* and *8*, it is approximated as $-60°$. The distributions of DoAs for the two experiments are shown in Figs. 10c and 10d. As in Section VII-D1, the DoAs of the two adjacent humans in experiment 1 at *5* & *6* and *1* & *2* have now formed two clusters. As in Fig. 10a, the wide separation between *2* and *5* has allowed two DoA distributions to develop peaks at $-28°$ and $-45°$. The trends are similar for the second experiment where adjacent humans have formed two clusters. However, the distance between the two clusters has widened further with peaks at $-47°$ and $48°$.

Summarizing the trends in Sec. VII-D1 and VII-D2, even though, the four element ULA is capable of distinguishing a single person and estimating the DoA accurately, it struggles in distinguishing when two persons are close specially up to around $20°$. Therefore, DoAs of nearby persons are identified as clusters. Additionally, based on the width of the distributions human induced DoAs, we can observe that the DoA resolution is high when the azimuth angles are closer to $0°$ as opposed to angles in excess of $\pm 50°$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2945476, IEEE Access

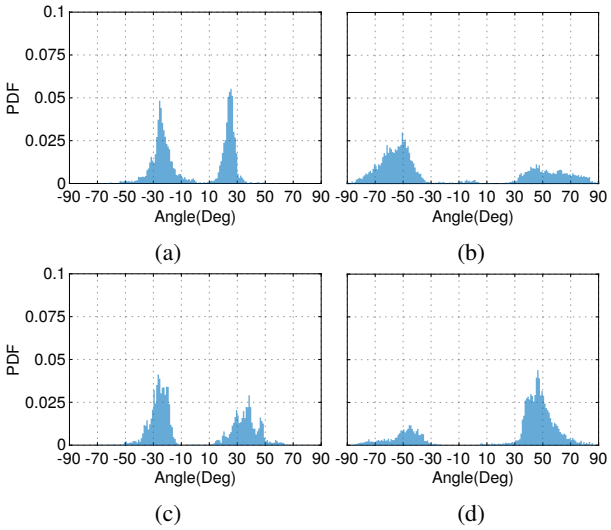


FIGURE 10: Direction of arrival distribution of humans at different distances according to positions (Fig 5b). (a) Two humans at 2 and 5, (b) Two humans at 12 and 7, (b) four humans at 1, 2, 5 and 6 and (e) four humans at 1, 12, 6 and 7

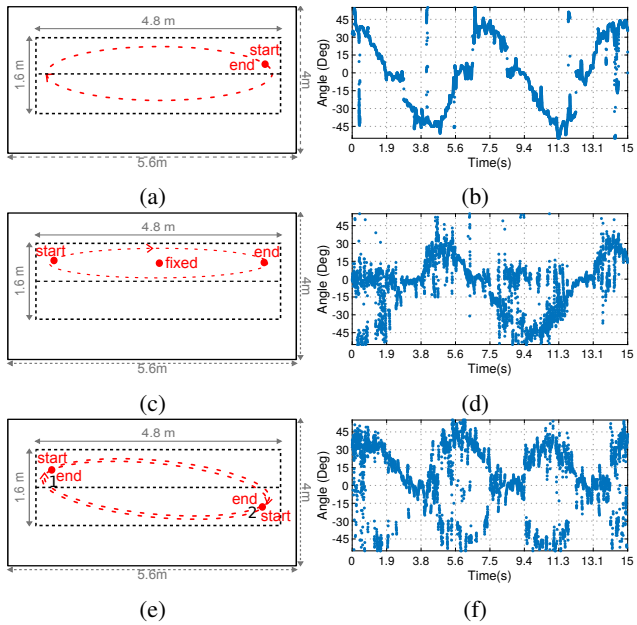


FIGURE 11: (a) A person starts moving from position 6 marked in Fig 5b, walks two rounds along the path marked in red and stops at 6, (b) DoA of the person from (a), (c) A person starts moving from position 1, walks between positions 1 and 6 along the path marked in red for two rounds and stops at 3. Another person is fixed at 3, (d) DoA of the two persons from c, (e) person 1 starts from position 6 and person 2 starts from position 12 and walk along the paths marked in red for two rounds, (f) DoAs of the two persons from (e).

E. DOA ESTIMATION OF WALKING PERSONS

Finally, we perform three experiments with 1 person walking, 1 person walking while another person is standing still (Fig. 11c) and two persons walking at two opposite ends 11e. Route of the first experiment is marked in Fig. 11a. The person starts at position *1* (as in Fig 5b), walks two rounds in the marked route and ends at position *1*. Figure 11b llustrates the DoAs captured for the person during his entire walk. From the figure it can be observed that the DoAs correctly follow the route of the person. The DoAs fluctuate between $\approx +45°$ and $\approx -50°$. Given that the maximum ($+44°$) and minimum angles ($-55°$) corresponding to the anchor points in Fig 5b, the DoA path is accurate. However, we also notice discontinuities closer to the LoS, e.g. 1.9-3.8 s and 5.6-7.5 s.

Fig. 11d illustrates the DoA estimation when one person walks while the other person stands in LoS as shown in Fig. 11c. Again, it can be observed that it is possible to capture both the persons at the same time. Note that the maximum DoA this time is $\approx 30°$ while the minimum is around $\approx -50°$. This can be attributed to the path that the person is taking where the anchor points of the two extremes have DoAs of $32°$ and $-55°$. Fig. 11f shows the DoAs when two persons walk simultaneously this time between $\approx +50°$ and $\approx -50°$. Interestingly, both paths were captured while discontinuities exist when a person moves closer to the LoS e.g. 1.9-3.8s. Another important observation is that the noise levels when two persons exist are higher compared to when a single person walks. We note that this is caused by low TX power levels used in the two experiments involving the two persons.

VIII. DISCUSSION AND CONCLUSIONS

This work presents an approach to capture a human and the direction of arrival through beamforming utilizing SDRs as the hardware. The main contributions are an algorithm to generate the steering vectors from few reference points and imprecise prior information, and detection of the human and the direction of that person from human induced paths amidst strong LoS signal. From experiments we detect human presence 100% of the times, estimate DoA up to a maximum error of $7°$ of a single person, show that detection of multiple stationary persons is possible when sufficient separation among them is present, and finally demonstrate the capability of tracking up to two walking persons.

Currently, one of the main limitations of our approach is the requirement of a calibration step to calculate the steering vector for a few reference directions. The proposed solution relies on the beam sweeping capabilities of the receiver as a post processing step which relies on accurate steering vector estimation. Even though, new wireless standards such as IEEE 802.11ac, ax and 5G support such capabilities, currently there is no commodity hardware that can be utilized for our purpose. Therefore, we use USRP hardware for this work. However, with the current expansion of research on device-free human sensing, we expect chipset manufacturers to expose such capabilities in a wide range of products.

The experiments are performed in a semi-anechoic chamber which has only the ground reflection, this setup is more similar to an open area that has no reflections. In future work, we intend to perform experiments in a cluttered environment with more reflections. However, as the reflections

from stationary objects remain static, we expect to subtract those effects using the same approach used to remove the LoS effect. For the experiments, currently a 4 antenna ULA is used at the receiver and the minimum distance between two nearby human subjects is 0.5m. In the current setup, the 0.5m distance between humans is not enough to separate them with the current ULA. It is well known that, as the number of antenna elements increase in the linear array, the beamwidth can further be reduced so that the interference on adjacent beams can be further attenuated. We will, in further investigations, study settings with 8 and 16 antennas in order to increase the directional perception accuracy and the count of people that can be recognized simultaneously. In future work we expect to extend this approach to detect and localize multiple persons and simultaneously recognize their activities with a larger ULA that supports a higher resolution.

REFERENCES

[1] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," Mechatronics, vol. 55, pp. 248–266, 2018.

[2] Y. Qiao, O. Zhang, W. Zhou, K. Srinivasan, and A. Arora, "PhyCloak: obfuscating sensing from communication signals," in Proc. of USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2016.

[3] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. Agarwal, "Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings," in Proc. of the 11th Conf. on Embedded Network Sensor Systems (Sensys), ACM, 2013.

[4] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," Proc. of the Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 1, no. 4, p. 155, 2018.

[5] D. T. Nguyen, W. Li, and P. O. Ogunbona, "Human detection from images and videos: A survey," Pattern Recognition, vol. 51, pp. 148–175, 2016.

[6] R. Gade and T. B. Moeslund, "Thermal cameras and applications: a survey," Machine vision and applications, vol. 25, no. 1, pp. 245–262, 2014.

[7] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human Sensing Using Visible Light Communication," in Proc. of the 21st Annual Int. Conf. on Mobile Computing and Networking (Mobicom), 2015.

[8] W. Mao, J. He, and L. Qiu, "CAT: high-precision acoustic motion tracking," in Proc. of the 22nd Annual Int. Conf. on Mobile Computing and Networking (Mobicom), ACM, 2016.

[9] A. Brajdic and R. Harle, "Scalable indoor pedestrian localisation using inertial sensing and parallel particle filters," in In Proc. of IPIN, 2012.

[10] S. Sigg, S. Shi, F. Buesching, Y. Ji, and L. Wolf, "Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features," in Proc. of Int. Conf. on Advances in Mobile Computing & Multimedia (MoMM), ACM, 2013.

[11] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-music: accurate device-free indoor localization," in Proc. of the Int. Joint Conf. on Pervasive and Ubiquitous Computing (Ubicomp), pp. 196–207, ACM, 2016.

[12] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person Localization via RF Body Reflections," in Proc. of USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2015.

[13] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home Gesture Recognition Using Wireless Signals," in Proc. of the Int. Conf. on Mobile Computing and Networking (Mobicom), ACM, 2013.

[14] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Inferring motion direction using commodity wi-fi for interactive exergames," in Proc. of CHI Conference on Human Factors in Computing Systems, ACM, 2017.

[15] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "xD-Track: Leveraging multi-dimensional information for passive wi-fi tracking," in Proceedings of the 3rd Workshop on Hot Topics in Wireless, pp. 39–43, ACM, 2016.

[16] R. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE transactions on antennas and propagation, vol. 34, no. 3, pp. 276–280, 1986.

[17] K. Han and A. Nehorai, "Improved source number detection and direction estimation with nested arrays and ulas using jackknifing," IEEE Transactions on Signal Processing, vol. 61, no. 23, pp. 6118–6128, 2013.

[18] S. Sigg, S. Palipana, S. Savazzi, and S. Kiyanoush, "Capturing human-machine interaction events from radio sensors in industry 4.0 environments," in Business Process Management Workshops, Springer, 2019.

[19] N. Malm, L. Zhou, E. Menta, K. Ruttik, R. Jäntti, O. Tirkkonen, M. Costa, and K. Leppänen, "User localization enabled ultra-dense network testbed," in Proc. of the Int. Conf. on 5G World Forum (5GWF), IEEE, 2018.

[20] J. Kruger, T. Lien, and A. Verl, "Cooperation of human and machines in assembly lines," CIRP Annals - Manufacturing Technology, vol. 58, no. 2, pp. 628 – 646, 2009.

[21] J. Fryman and B. Matthias, "Safety of industrial robots: From conventional to collaborative applications," in Proc. of the German Conf. on Robotics, pp. 1–5, 2012.

[22] L. Wang, "Collaborative robot monitoring and control for enhanced sustainability," The Int. Journ. of Advanced Manufacturing Technology, vol. 81, no. 9-12, pp. 1433–1445, 2015.

[23] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl, "Rf-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals," IEEE Trans. on Mobile Comput., vol. 13, no. 4, pp. 907–920, 2013.

[24] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart Homes That Monitor Breathing and Heart Rate," in Proc. of the CHI Conference on Human Factors in Computing Systems, 2015.

[25] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home Gesture Recognition Using Wireless Signals," in Proc. of the Int. Conf. on Mobile Computing and Networking (Mobicom), 2013.

[26] N. Patwari and P. Agrawal, "Effects of Correlated Shadowing: Connectivity, Localization, and RF Tomography ," in Proc. of the Int. Conf. on Information Processing in Sensor Networks (IPSN), 2008.

[27] H. Wang, D. Zhang, J. Ma, Y. Wang, Y. Wang, D. Wu, and T. Gu, "Human Respiration Detection with Commodity WiFi Devices: Do User Location and Body Orientation Matter?," in Proc. of the Int. Joint Conf. on Pervasive and Ubiquitous Computing (Ubicomp), ACM, 2016.

[28] S. Sigg, U. Blanke, and G. Tröster, "The telepathic phone: Frictionless activity recognition from wifi-rssi," in Proc. of the Int. Conf. on Pervasive Computing and Communications (Percom), IEEE, 2014.

[29] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," IEEE Access, vol. 5, pp. 18066–18074, 2017.

[30] S. Shi, S. Sigg, L. Chen, and Y. Ji, "Accurate location tracking from CSI-based passive device-free probabilistic fingerprinting," IEEE Trans. on Veh. Technol., vol. 67, no. 6, pp. 5217–5230, 2018.

[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. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 1, no. 3, p. 72, 2017.

[32] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in Proc. of the Int. Joint Conf. on Pervasive and Ubiquitous Computing (Ubicomp), ACM, 2016.

[33] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices," IEEE Trans. on Mobile Comput., vol. 16, no. 2, pp. 511–526, 2016.

[34] P. Melgarejo, X. Zhang, P. Ramanathan, and D. Chu, "Leveraging Directional Antenna Capabilities for Fine-grained Gesture Recognition," in Proc. of the Int. Joint Conf. on Pervasive and Ubiquitous Computing, Ubicomp, 2014.

[35] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "WiDir: walking direction estimation using wireless signals," in Proc. of the Int. joint Conf. on pervasive and ubiquitous computing (Ubicomp), ACM, 2016.

[36] K. Joshi, D. Bharadia, M. Kotaru, and S. Katti, "WiDeo: Fine-grained Device-free Motion Tracing using RF Backscatter," in Proc. of USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2015.

[37] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2. 0: Passive human tracking with a single wi-fi link," in Proc. of the 16th Annual Int. Conf. on Mobile Systems, Applications, and Services (Mobisys), ACM, 2018.

[38] H. L. V. Trees, Optimum Array Processing . Wiley, 2002.

[39] E. Y. Menta et al., "On the Performance of AoA-Based Localization in 5G Ultra-Dense Networks," IEEE Access, vol. 7, pp. 33870–33880, 2019.

[40] A. Vural, "Effects of perturbations on the performance of optimum/adaptive arrays," IEEE Trans. Aerosp. Electron. Syst., no. 1, pp. 76–87, 1979.

[41] A. Hassanien, S. A. Vorobyov, and K. M. Wong, "Robust adaptive beamforming using sequential quadratic programming: An iterative solution to the mismatch problem," IEEE Signal Process. Lett., vol. 15, pp. 733–736, 2008.