Exploring Passive Activity Recognition using Multi-shot UWB CIRs

HAN LIN[1] ATSUSHI NOMURA[1] KOTA TSUBOUCHI[2] NOBUHIKO NISHIO[3]
MASAMICHI SHIMOSAKA[1]

Abstract: As a promising device-free solution, wireless sensing techniques have been applied to various applications while addressing privacy concerns associated with traditional computer vision-based methods. There has been limited research focusing on the use of UWB CIR for activity recognition and showing improved performance over Wi-Fi CSI. However, previous studies using one-shot CIR cannot correctly capture the variations inherent in dynamic activities. This research proposes a device-free activity recognition approach by utilizing multi-shot CIRs, consisting of several CIRs arranged in a time series, in order to cover the transition of activities. And three variants of wavelet denoising across various dimensions are introduced to remove noise on signal. Experiments were conducted with horizontal and vertical device settings to test the sensing performance in different scenarios. The results were also compared with Wi-Fi CSI at the same frequency to benchmark its effectiveness.

Keywords: UWB, device-free, human activity recognition

1. Introduction

Human Activity Recognition (HAR) plays a pivotal role in the development of intelligent systems, with applications spanning from healthcare monitoring to smart home automation. Traditionally, HAR research has been bifurcated into sensor-based and vision-based methodologies. Sensor-based approaches, while precise, often suffer from the inconvenience of requiring users to wear devices continuously. Vision-based methods, on the other hand, are impeded by environmental variables such as lighting conditions and physical obstructions, and significantly, they raise serious privacy concerns due to their intrusive nature [1].

In light of these limitations, wireless sensing techniques have emerged as a promising device-free alternative, garnering attention for their ability to circumvent the privacy issues inherent in vision-based systems without necessitating the use of wearable sensors. Among these, RF-based methods have demonstrated satisfactory performance in indoor localization and HAR, and methods based on Wi-Fi Received Signal Strength Indicator (RSSI) or Channel State Information (CSI), when coupled with advanced deep learning techniques, also showed promising results in individual behaviors recognition [2–4].

More recently, the focus has shifted towards leveraging Ultra-wideband (UWB) technology for HAR, driven by its growing popularity in smart home automation and indoor localization services [5]. UWB technology, characterized by its transmission of signals across a bandwidth exceeding 500MHz using minimal energy, allows for the conveyance of substantial signal energy without causing interference with existing narrowband and carrier wave transmissions within the same frequency band. The incorporation of UWB beacons in consumer devices, such as iPhones, heralds widespread application possibilities.

There has been limited research focusing on the use of UWB Channel Impulse response(CIR) for activity recognition. Existing studies have indicated that UWB CIR can outperform Wi-Fi CSI in activity recognition tasks by analyzing the time-series data from CIR to discern human behaviors through multi-path effects [6]. However, these studies primarily utilized a one-shot CIR approach, which fails to adequately capture the dynamic nature of human activities due to the brief duration of each CIR snapshot. This recognition gap underscores the necessity for a multi-shot CIR approach, which employs a series of CIR snapshots over time to comprehensively capture the evolving nature of activities.

This research is dedicated to exploring a passive activity recognition system that employs multi-shot UWB CIRs in order to extract enhanced information about the CIR variations resulting from different activities. Additionally, our system aims to to mitigate the noise introduced by hardware and environmental factors in the CIRs data. To achieve this, we employ calibration procedures along with wavelet denoising techniques. The preprocessed CIRs are subsequently used as input for a neural network designed to accurately recognize various activities.

The contributions of this article include the following:

- We propose a device-free activity recognition approach by utilizing multi-shot CIRs to cover the transition of activities.
- Three types of wavelet denoising across various dimensions are introduced to remove noise on signal.

[1] Department of Computer Science, School of Engineering, Tokyo Institute of Technology
[2] LY Corporation
[3] Ritsumeikan University

- We conduct the experiments to test the sensing performance in the two typical application scenarios. The results demonstrate the limitations of one-shot CIR and highlight the enhanced performance of multi-shot CIRs and various wavelet denoising techniques. Further, we benchmark our approach against Wi-Fi CSI at the same frequency to assess its effectiveness.

Related work

Passive Activity recognition based on Wi-Fi

Passive activity recognition using Wi-Fi signals, particularly through RSSI and CSI, has garnered significant interest in the research community. RSSI-based methodologies, as explored by Sigg et al., have proven capable of recognizing a range of activities, gestures, and environmental contexts using signals captured by mobile phones [7]. Similarly, Gu et al. have introduced an adaptable online activity recognition architecture that utilizes a novel fusion algorithm and a classification tree to enhance the differentiation of activities with closely related signal patterns [8]. Despite these advances, RSSI's reliance on signal fingerprinting struggles to achieve high accuracy due to the limited channel information it provides.

The advent of CSI at the physical layer has marked a turning point, offering finer granularity and thereby enabling more successful applications in CSI-based HAR. Noteworthy among these are the neural network framework proposed by Chen et al., which integrates bidirectional and attention mechanisms for Wi-Fi CSI data processing [4], and the WiAct prototype by Yan et al., which employs an Adaptive Activity Cutting Algorithm (AACA) to distinguish between active and inactive signal variances [9]. Sch"afer et al.'s use of the Nexmon tool to obtain comprehensive HAR results further underscores the potential of CSI-based approaches, achieving performance on par with or surpassing other machine learning-based studies [10].

Nonetheless, the effectiveness of these deep learning models can be compromised by individual variances and environmental dynamics, which may introduce intrinsic noise into the CSI data, ultimately affecting performance [5].

Passive Activity recognition based on Ultra-Wideband Radios

UWB technology, known for its high temporal resolution and broad frequency bandwidth, offers significant advantages for passive activity recognition due to its sensitivity to human-induced signal variations and low power consumption. UWB has been effectively applied to a range of activities, from hand gesture recognition to basic activity classification, demonstrating superior performance [11].

A key focus within UWB-based HAR is the use of Channel Impulse Response, which, akin to Wi-Fi CSI, captures the multipath propagation of signals in indoor environments. Sharma et al.'s exploration into the use of UWB CIR for activity recognition has shown high accuracy in identifying basic activities such as standing, sitting, and lying [12]. Furthermore, Bocus et al. have conducted an analysis comparing the classification performance using high-resolution CIR from UWB modules against CSI data extracted from Wi-Fi, illustrating the distinct advantages of UWB in certain scenarios [6].

However, challenges remain, particularly regarding the temporal limitations of CIR data. The inherently short duration of CIR captures can hinder the accurate reflection of all human-induced signal reflections, leading to potential misrecognitions in activities with subtle differences, such as jogging versus walking in our experiments.

2. Activity recognition using UWB CIR

2.1 Channel Impulse Response

The Channel Impulse Response of UWB represents encapsulates the channel characteristics of a transmitted signal. The widely accepted UWB channel model, inspired by the Saleh-Valenzuela model, describes the signal as comprising multiple path components grouped into L clusters, each containing K paths [13]. The model is mathematically represented as follows:

$$h(t) = \sum_{l=0}^{L} \sum_{k=0}^{K} \alpha_{k,l} e^{j\theta_{k,l}} \delta(t - T_l - \tau_{k,l}) \qquad (1)$$

Here, $\alpha_{k,l}$, $\theta_{k,l}$, and $\tau_{k,l}$ represent the amplitude, phase shift, and time delay of the k^{th} path in the l^{th} cluster, respectively. Tl signifies the arrival time or the delay of the l^{th} cluster, and $\delta(\cdot)$ is the Dirac delta function.

In scenarios like non-line-of-sight indoor office or industrial environments, characterized by a dense, continuous arrival of multipath components, a simplified single-cluster model can be adopted:

$$h(t) = \sum_{k=0}^{K} \alpha_k e^{j\theta_k} \delta(t - \tau_k) \qquad (2)$$

This CIR adequately captures the multipath propagation effect, allowing for the detection of signal variations caused by human movement. By analyzing these variations, one can infer the type of activity being performed.

2.2 Limitations of one-shot CIR

Prior research on passive activity recognition leveraging CIR has shown promising results, utilizing one-shot CIR measurements combined with machine learning and deep learning techniques to recognize activities [6, 12].

The high temporal resolution of CIR is capable of detecting minute movements. However, the brevity of CIR measurements (approximately 1ns) limits their ability to encompass entire activities, which may last several seconds. For instance, activities such as walking or standing up cannot be fully captured in such a short snapshot, leading to inaccuracies in activity recognition when movements share similar postures but differ in speed or pattern.

Moreover, the equation for the UWB system can be expanded to include the received signal, represented as follows:

$$r(t) = \sum_{k=0}^{K} \alpha_k e^{j\theta_k} s(t - \tau_k) + n(t) \qquad (3)$$

Here, $s(\cdot)$ denotes the transmitted signal, and $n(\cdot)$ represents additive Gaussian noise. Merely selecting specific portions of the CIR does not inherently eliminate Gaussian noise, necessitating

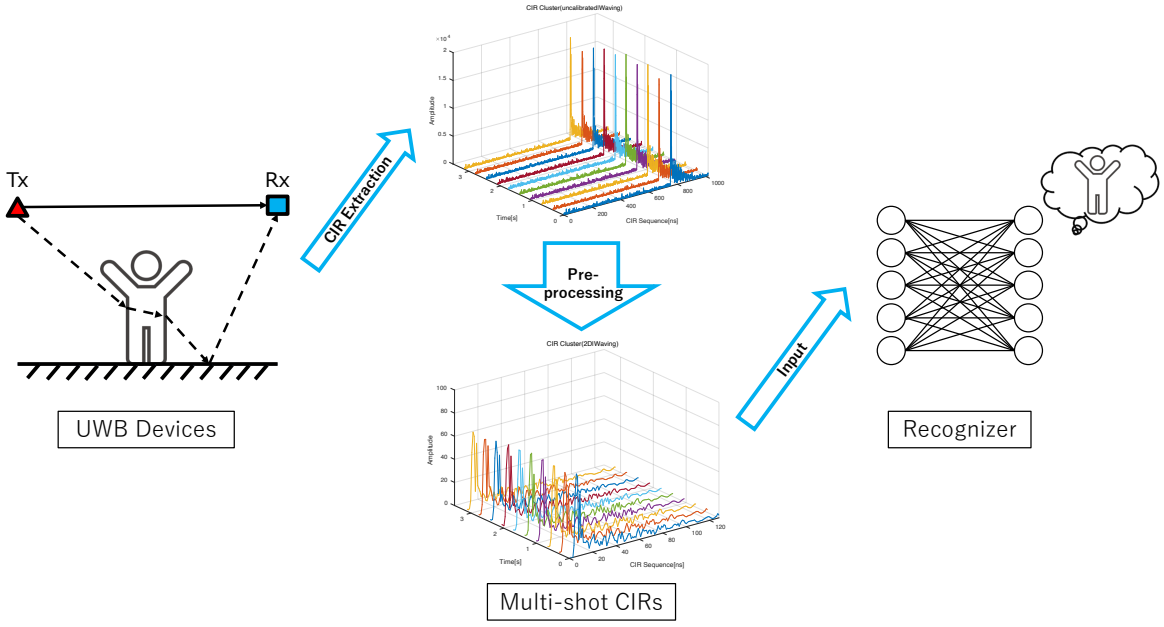


Fig. 1 System Overview, consisting of extracting CIRs from UWB devices, pre-processing on multi-shot CIRs and recognition.

the application of denoising techniques.

To address these limitations, we introduce a multi-shot CIR approach complemented by pre-processing techniques, which we will explore in the subsequent chapter. This method aims to enhance the accuracy and reliability of activity recognition by overcoming the constraints posed by one-shot CIR measurements.

3. Proposed method: Multi-shot UWB CIRs based activity recognition

3.1 Overview

Our method utilizes multi-shot UWB CIRs, comprising several CIRs arranged in a time series to span the entire activity period. Due to the redundancy and noise inherent in the raw signal data, we implement a two-step preprocessing procedure that includes calibration and wavelet denoising.

In the calibration step, we calibrate each CIR by first path index and normalize it using the Preamble Accumulation Count (PAC) value, which is calculated by the built-in chip. For denoising, we explore three types of wavelet denoising techniques applied across various dimensions to eliminate Gaussian noise originating from the environment or the devices themselves.

Following preprocessing, we employ an Attention-Based Long Short-Term Memory (ABLSTM) network as the recognizer to classify activities. An overview of our proposed method is illustrated in Fig.1.

3.2 Multi-shot CIRs

In this study, we direct our attention to analyzing the amplitude of CIRs. The multi-shot CIRs matrix at time t, denoted by $\boldsymbol{M}^{(t)}$, is formulated as follows:

$$\boldsymbol{M}^{(t)} = \begin{bmatrix} \boldsymbol{x}_{t-iT} \\ \vdots \\ \boldsymbol{x}_{t-T} \\ \boldsymbol{x}_{t} \end{bmatrix} \tag{4}$$

In the above equation, T signifies the interval between successive CIR measurements, while i represent the count of measurements taken before the current measurement, respectively. The term $\boldsymbol{x}_t$ denotes the amplitude row vector of a single CIR measurement at time t, which is defined as:

$$\boldsymbol{x}_t = [|r(t)|, |r(t+1)|, ... |r(t+N-1)|] \tag{5}$$

where N is the length of one-shot CIR sequence and equal to 1016 in our devices.

3.3 System mechanism

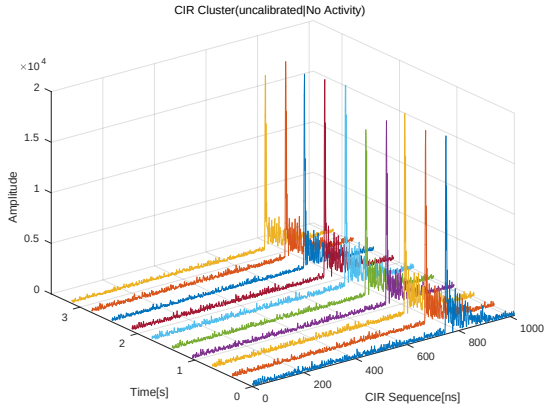
3.3.1 Calibration

As illustrated in Fig.2(a), the raw CIRs data exhibits redundancy along with significant noise, attributed to environmental factors and device output. The noise is predominantly Gaussian and affects both the initial and final segments of the CIR sequence.

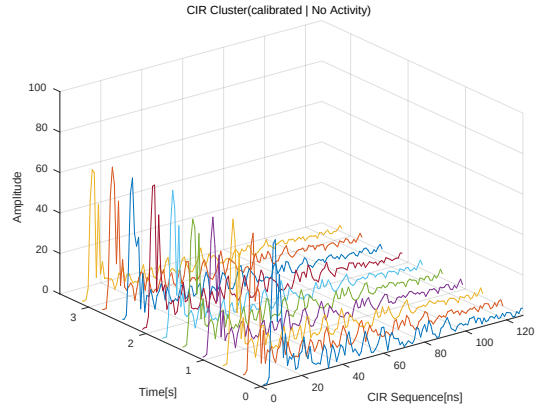
To address this issue, we leverage the estimated first index value computed by the device firmware to discard the noisy segments, concentrating our analysis on the fluctuating segments attributable to the multi-path effect. Consequently, the refined amplitude vector for a single CIR measurement is represented as:

$$\boldsymbol{x}'_t = [|r(t+l-3)|, ... |r(t+l)|, ... |r(t+l+124)|] \tag{6}$$

where the l indicates the first path index in the current shot of CIR measurement, and we empirically select a subset of 128 samples, including three samples preceding and 125 samples following the first path index, to capture the relevant signal information.



(a) Uncalibrated multi-shot CIRs



(b) Calibrated multi-shot CIRs

Fig. 2 Calibration of Multi-Shot CIRs. CIR Sequence represents a single CIR measurement, measured in nanoseconds (ns), while Time denotes the interval between consecutive CIRs.

Furthermore, we employ the Preamble Accumulation Count (PAC) value for normalize purposes. The PAC value, which reflects the accumulated preamble symbols, correlates with the received signal's quality [14]. Consequently, the CIR's amplitude is adjusted based on the PAC value to mitigate the effects of the transmission environment.

After these two steps, the calibrated multi-shot CIRs matrix is represented as:

$$\boldsymbol{M}'^{(t)} = \begin{bmatrix} \boldsymbol{x}'_{t-iT}/p_i \\ \vdots \\ \boldsymbol{x}'_{t-T}/p_1 \\ \boldsymbol{x}'_{t}/p_0 \end{bmatrix} \tag{7}$$

Here, the p_i denotes the PAC value associated with the i^{th} CIR measurement that precedes the CIR at time t. The calibrated signal is shown in Fig.2(b).

3.3.2 Wavelet denoising

In the calibration process, we select the vicinity segment of the estimated first path index as feature. However, addressing the issue of in-band noise removal is still critical. By observing time-oriented overlapped CIR curves, as illustrated in Fig.3, Gaussian noise persists within the CIRs even in empty environments. Considering the significant challenge posed by noise from environmental reflections to activity recognition, we have chosen Wavelet denoising as our method for removing Gaussian noise in the CIRs.

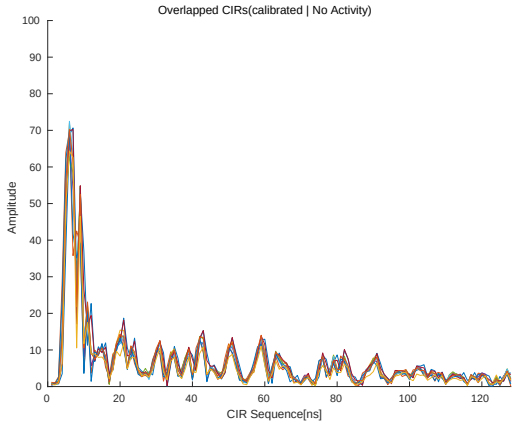


Fig. 3 Overlapped figure of calibrated multi-shot CIRs across the Time axis.

Wavelet-based denoising is a renowned preprocessing method across various fields. For instance, in Wi-Fi gesture recognition researches [15, 16], this technique has been instrumental in filtering out high-frequency noise while preserving the critical details of CSI patterns necessary for distinguishing similar gestures. Similarly, in UWB positioning systems [17, 18], it has been effectively utilized to mitigate the effects of personnel motion noise, thereby enhancing positioning accuracy and motion localization.

Discrete Wavelet Transform

This method utilizes the Discrete Wavelet Transform (DWT) to transform the signal into the time-frequency domain, without assuming any prior knowledge about the signal's original nature. Offering a significant advantage over the Short-Time Fourier Transform (STFT), the DWT facilitates variable resolution between the time and frequency domains, providing a multi-scale analysis of the signal [19].

The DWT performs an iterative decomposition, breaking down the signal $f(n)$ into two components: an approximation coefficients component and a detail coefficients component. This decomposition process is then repeated on the approximation component to extract finer details from the signal. The process is applied iteratively in several steps or levels, denoted by j. The approximation and detail coefficients at different levels are calculated using a series of scaling function coefficients $\boldsymbol{\phi}_{j,k}(n)$ and wavelet function coefficients $\boldsymbol{\psi}_{j,k}(n)$ according to the following equations:

$$\boldsymbol{c}_{j,k} = < f(n), \boldsymbol{\phi}_{j,k}(n) > = \sum_{n\in\mathbb{Z}} f(n)\boldsymbol{\phi}_{j,k}(n) \tag{8}$$

$$\boldsymbol{d}_{j,k} = < f(n), \boldsymbol{\psi}_{j,k}(n) > = \sum_{n\in\mathbb{Z}} f(n)\boldsymbol{\psi}_{j,k}(n) \tag{9}$$

Here, the $< \cdot >$ denotes the dot product operation. Through multiple decompositions, we ultimately obtain a series of decomposition coefficients at the coarsest level J. As a result, the original signal can be expressed by:

$$f(n) = \sum_{k\in\mathbb{Z}} \boldsymbol{c}_{j,k}\, \boldsymbol{\phi}_{J,k}(n) + \sum_{j=1}^{J}\sum_{k\in\mathbb{Z}} \boldsymbol{d}_{j,k}\, \boldsymbol{\psi}_{j,k}(n) \tag{10}$$

Thresholding and Reconstruction

Thresholding is then applied to the wavelet detail coefficients to remove their noisy component. The threshold rule is based on the an empirical Bayes approach to the estimation of possibly sparse sequences observed in Gaussian white noise [20]. Finally, the signal is reconstructed by combining the coefficients of the last approximation level with all thresholded details to get the denoised signal.

Variants of denoising

In our research, we explore three types of wavelet denoising techniques to tackle noise across various dimensions effectively.

(1) Seq-Oriented denoising: This approach is widely used to eliminate in-band noise from individual CIR measurements. We apply wavelet denoising to each calibrated CIR measurement, i.e., the row vectors of $\boldsymbol{M}'^{(t)}$. This method is particularly effective in addressing noise inherent to the single measurement.

(2) Time-Oriented denoising: In this method, wavelet denoising is applied to each index of CIR measurements, i.e., the column vectors of $\boldsymbol{M}'^{(t)}$, to mitigate Gaussian noise as aforementioned in Fig.x. This noise may be attributed to environmental factors or hardware limitations. Since the amplitude variations in CIR caused by human interference are significantly larger than those caused by environmental or hardware issues, this approach proves beneficial in preserving the integrity of human-induced signal variations.

(3) Seq-Time-Oriented denoising: This innovative approach utilizes 2D wavelet denoising on the entire cluster of CIRs, $\boldsymbol{M}'^{(t)}$, akin to denoising techniques used in computer vision. It is designed to remove noise along both the sequence and time axes of CIR, offering a comprehensive solution to noise reduction that considers the complex interplay between temporal and sequential dimensions of the data.

3.3.3 ABLSTM

The ABLSTM framework is widely recognized for its application in Wi-Fi CSI-based passive activity recognition [4]. This neural network also excels in extracting time-series features from each CIR frame, making it an ideal candidate for our comparison with Wi-Fi CSI performance. As a result, it is integrated into our system with minor modifications.

As depicted in Fig.4, our implementation leverages a Bidirectional Long Short-Term Memory (BLSTM) network for feature learning. Subsequently, an attention mechanism generates an attention matrix, highlighting the significance of features and time steps. Here, an element-wise multiplication is used to integrate the learned features with the attention matrix, and we incorporate a shortcut connection into the architecture to mitigate the degradation problem [21]. The procedure concludes with a flatten layer that transforms the feature matrix into a vector, which is then classified using a softmax layer to distinguish between different activities.

4. Experiment

In this chapter, we delve into a comparative analysis of performance across various activities, comparing with one-shot CIR within two typical application scenarios. Additionally, we extend our comparison between Wi-Fi CSI and UWB CIR to highlight the advantages offered by UWB CIR.

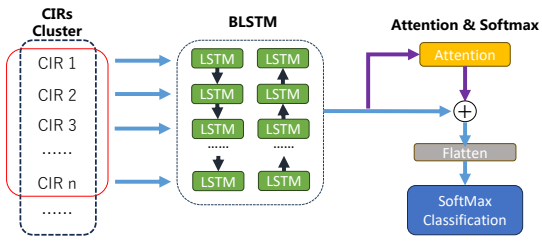


Fig. 4 The ABLSTM framework for multi-shot CIRs

4.1 Experiment settings

In our experiments, we utilized pairs of Decawave EVK1000 boards, which leverage the Two-Way Ranging (TWR) protocol to provide accurate distance measurements in accordance with the 802.15.4a standard. These boards were configured in Mode 3, setting the DW1000 chip to operate with a bandwidth of 500MHz, a carrier frequency of 4.0GHz, a Pulse Repetition Frequency (PRF) of 64MHz, a preamble length of 1024 bits, and a data rate of 110kbps [22]. With this configuration, CIRs comprising 1016 samples were extracted by interfacing with a laptop.

For CSI data collection, we employed the Raspberry Pi equipped with the Nexmon CSI toolkit [23]. The Wi-Fi configuration was set to a bandwidth of 40MHz, encompassing 128 subcarriers. The transmission frequency for both CIR and CSI data was standardized at 3Hz.

4.2 Experiment results

4.2.1 Horizontal device setting

The implementation of horizontal device placement has been a prevalent methodology in prior research, particularly in scenarios where device alignment at specific levels captures abundant human reflections effectively. In our study, devices were positioned at a height of 0.7m and spaced 1.8m apart. This setup was designed to emulate a typical environment conducive to activity recognition research. Within this arrangement, participants were instructed to perform a series of activities at different locations, as depicted in Fig.5.

As Fig.6 demonstrates, distinguishing between jogging and walking activities using one-shot CIR is challenging, primarily due to the similarity in movement. Given that each CIR measurement encapsulates data from a mere 1ms window, lacking sequential frame variation knowledge complicates the differentiation of activities with similar postures. Therefore, it becomes natural to consider multi-shot CIRs to enhance recognition accuracy.

Utilizing multi-shot CIRs, notably with the incorporation of wavelet denoising techniques, significantly improves recognition performance. As evidenced in Table 1, the results of 10-fold validation, the accuracy of raw CIRs averaged at 92.6%, outperforming the one-shot approach. Moreover, Time-Oriented and Seq-Time-Oriented denoising methods contributed approximately a 2% increase in accuracy, contrasting with the Seq-Oriented approach. This discrepancy underscores the minimal noise impact in single CIR measurements within horizontal setups, where

(a) No Activity (b) Jogging

(c) Walking (d) Waving

(e) Standing (f) Sitting

Fig. 5 Activities in horizontal scenario

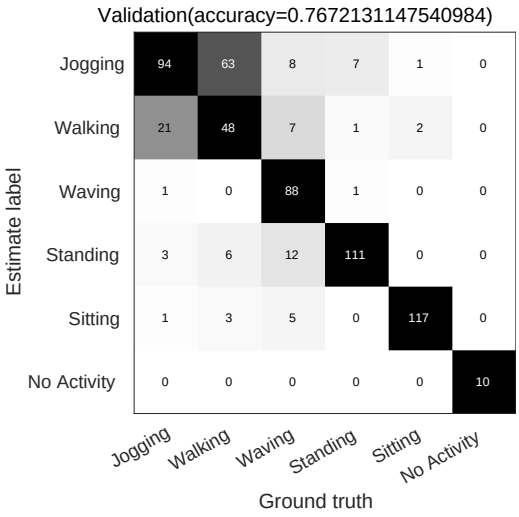


Fig. 6 Recognition result in horizontal scenario with one-shot CIR

Table 1 Results of wavelet denoising with multi-shot CIRs in the horizontal setting

Method	Top1 Accuracy
Calibrated Data	92.6% ± 4.9%
Seq-Oriented denoising	91.7% ± 3.6%
Time-Oriented denoising	94.1% ± 3.4%
Seq-Time-Oriented denoising	**94.9% ± 2.2 %**

human-induced multipath effects are prevalent [6].

4.2.2 Vertical device setting

Contrary to the horizontal arrangement, this setup involves mounting the devices on the ceiling, at a height of 2.5m, as shown in Fig.7, where human reflections are considerably reduced. This deployment was selected to evaluate the impact of device positioning in scenarios characterized by varying degrees of reflections. We let the subject shake its shoulder under the device to

(a) No Activity (b) Shaking

Fig. 7 Vertical scenario

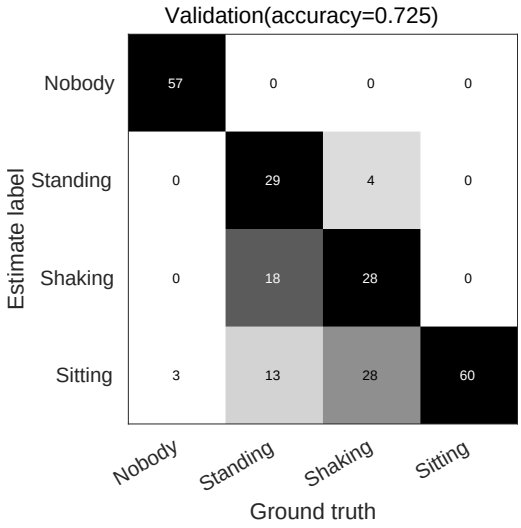


Fig. 8 Recognition result in horizontal scenario with one-shot CIR

Table 2 Results of wavelet denoising with multi-shot CIRs in the vertical setting

Method	Top1 Accuracy
Calibrated Data	86.8% ± 7.3%
Seq-Oriented denoising	89.6% ± 5.5%
Time-Oriented denoising	93.9% ± 3.7%
Seq-Time-Oriented denoising	**94.1% ± 3.9 %**

detect the sensitivitiy of system.

Contrary to the horizontal arrangement, the vertical device arrangement, entailing device placement on the ceiling at a 2.5m height, significantly diminishes human reflections, as shown in Fig.7. This setup was chosen to assess device positioning effects under varied reflection intensities. Subjects were asked to perform shoulder shaking movements and other static posture beneath the device to evaluate system sensitivity in such configurations.

Fig.8 indicates the difficulty of one-shot CIR in distinguishing between shaking and standing activities. Nonetheless, Table 2 shows that the accuracy achieved in multi-shot CIR scenarios was markedly higher, with wavelet denoising further enhancing performance. This improvement of denoising is attributed to the lower human reflection intensity, making it more susceptible to environmental noise interference.

Table 3 Comparison between UWB CIR and Wi-FI CSI in the horizontal setting

Signal	Top1 Accuracy
UWB CIR	92.6% ± 4.9%
Wi-Fi CSI	61.6% ± 8.3%

4.2.3 Comparison with Wi-Fi CSI

Comparative analysis between Wi-Fi CSI and UWB CIR, specifically without preprocessing in the horizontal scenario, underscores a notable disparity in performance. According to Table 3, the effectiveness of Wi-Fi CSI significantly lags behind that of UWB CIR. This performance gap is primarily attributed to the inherent noise present in CSI data, which necessitates extensive preprocessing efforts to mitigate. Additionally, synchronizing the transmission frequency to that of CIR fails to address the low resolution of CSI data, thereby unable to replicate the high levels of efficacy observed in previous researches.

5. Conclusion

In this study, we explored the potential of utilizing multi-shot UWB CIRs for passive activity recognition without the need for additional sensors. Our experimental results demonstrate the efficacy of our approach, particularly in distinguishing between activities such as jogging and walking—a task that one-shot CIR struggle with. Notably, our method achieved an accuracy rate exceeding 90% in both horizontal and vertical device deployment. We observed a significant enhancement in performance through the application of wavelet denoising, especially in vertical device configurations where signal reflection is inherently lower compared to horizontal setups.

Furthermore, we conducted a comparative analysis between Wi-Fi CSI and UWB CIR, highlighting the superior capabilities of CIR technology. However, it is important to acknowledge that the comparative experiments with CSI were somewhat limited. The constraints included maintaining the same transmission rate as that of the CIR and the absence of pre-processing on the noisy CSI data.

For the future work, we will focus on conducting more comprehensive experiments involving both Wi-Fi CSI and UWB CIR, without the constraints of frequency limitations for Wi-Fi CSI. This will include implementing full preprocessing measures to ensure a more accurate assessment of Wi-Fi CSI's potential and a fairer comparison with UWB CIR. By exploring higher frequencies, we aim to refine the clarity of CIRs' reflection of different actions, thereby optimizing the system for real-time activity recognition. Additionally, while wavelet denoising has shown promise in improving performance in certain scenarios, its effectiveness varies. Therefore, we plan to explore other preprocessing and signal analysis techniques to further enhance performance, particularly in environments where human reflections predominantly affect CIRs data.

References

[1] Wang, J., Chen, Y., Hao, S., Peng, X. and Hu, L.: Deep learning for sensor-based activity recognition: A survey, *Pattern recognition letters*, Vol. 119, pp. 3–11 (2019).

[2] Zhang, D., Ma, J., Chen, Q. and Ni, L. M.: An RF-based system for tracking transceiver-free objects, *Fifth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom'07)*, IEEE, pp. 135–144 (2007).

[3] Sigg, S., Shi, S. and Ji, Y.: Rf-based device-free recognition of simultaneously conducted activities, *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pp. 531–540 (2013).

[4] Chen, Z., Zhang, L., Jiang, C., Cao, Z. and Cui, W.: WiFi CSI based passive human activity recognition using attention based BLSTM, *IEEE Transactions on Mobile Computing*, Vol. 18, No. 11, pp. 2714–2724 (2018).

[5] Yang, J., Xu, Y., Cao, H., Zou, H. and Xie, L.: Deep learning and transfer learning for device-free human activity recognition: A survey, *Journal of Automation and Intelligence*, Vol. 1, No. 1, p. 100007 (2022).

[6] Bocus, M. J., Chetty, K. and Piechocki, R. J.: UWB and WiFi Systems as Passive Opportunistic Activity Sensing Radars, *2021 IEEE Radar Conference (RadarConf21)*, pp. 1–6 (online), DOI: 10.1109/RadarConf2147009.2021.9455175 (2021).

[7] Sigg, S., Blanke, U. and Tröster, G.: The telepathic phone: Frictionless activity recognition from WiFi-RSSI, *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 148–155 (online), DOI: 10.1109/PerCom.2014.6813955 (2014).

[8] Gu, Y., Ren, F. and Li, J.: PAWS: Passive Human Activity Recognition Based on WiFi Ambient Signals, *IEEE Internet of Things Journal*, Vol. 3, No. 5, pp. 796–805 (online), DOI: 10.1109/JIOT.2015.2511805 (2016).

[9] Yan, H., Zhang, Y., Wang, Y. and Xu, K.: WiAct: A Passive WiFi-Based Human Activity Recognition System, *IEEE Sensors Journal*, Vol. 20, No. 1, pp. 296–305 (online), DOI: 10.1109/JSEN.2019.2938245 (2020).

[10] Schäfer, J., Barrsiwal, B. R., Kokhkharova, M., Adil, H. and Liebehenschel, J.: Human activity recognition using CSI information with nexmon, *Applied Sciences*, Vol. 11, No. 19, p. 8860 (2021).

[11] Maitre, J., Bouchard, K., Bertuglia, C. and Gaboury, S.: Recognizing activities of daily living from UWB radars and deep learning, *Expert Systems with Applications*, Vol. 164, p. 113994 (2021).

[12] Sharma, S., Mohammadmoradi, H., Heydariaan, M. and Gnawali, O.: Device-Free Activity Recognition Using Ultra-Wideband Radios, *2019 International Conference on Computing, Networking and Communications (ICNC)*, pp. 1029–1033 (online), DOI: 10.1109/ICCNC.2019.8685504 (2019).

[13] Saleh, A. and Valenzuela, R.: A Statistical Model for Indoor Multipath Propagation, *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 2, pp. 128–137 (online), DOI: 10.1109/JSAC.1987.1146527 (1987).

[14] Bocus, M. J. and Piechocki, R.: A comprehensive ultra-wideband dataset for non-cooperative contextual sensing, *Scientific Data*, Vol. 9, No. 1, p. 650 (2022).

[15] Abdelnasser, H., Youssef, M. and Harras, K. A.: Wigest: A ubiquitous wifi-based gesture recognition system, *2015 IEEE conference on computer communications (INFOCOM)*, IEEE, pp. 1472–1480 (2015).

[16] Tan, S. and Yang, J.: WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition, *Proceedings of the 17th ACM international symposium on mobile ad hoc networking and computing*, pp. 201–210 (2016).

[17] Hongchun, Q., Jisheng, S. and Yongbo, W.: An adaptive wavelet denoising algorithm for UWB indoor localization, *Proceedings of the 2017 International Conference on Data Mining, Communications and Information Technology*, pp. 1–5 (2017).

[18] Liu, N., Zhang, R., Su, Z., Fu, G. and He, J.: Research on wavelet threshold denoising method for UWB tunnel personnel motion location, *Mathematical Problems in Engineering*, Vol. 2020, pp. 1–14 (2020).

[19] Mallat, S.: *A wavelet tour of signal processing*, Elsevier (1999).

[20] Johnstone, I. M. and Silverman, B. W.: Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse sequences (2004).

[21] He, K., Zhang, X., Ren, S. and Sun, J.: Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778 (2016).

[22] : DW1000 USER MANUAL, `https://www.qorvo.com/products/d/da007967`.

[23] Schulz, M., Wegemer, D. and Hollick, M.: Nexmon: The C-based Firmware Patching Framework (2017).