Device-free occupant activity recognition in smart offices using intrinsic Wi-Fi components

Qizhen Zhou [a,b], Jianchun Xing [a,b], Qiliang Yang [a,*]

[a] *College of Defense Engineering, Army Engineering University of PLA, No. 1, Haifu Xiang, Nanjing, 210007, China*
[b] *School of Computer Science and Engineering, Nanjing University of Science and Engineering, No. 200, Xiaolingwei Road, Nanjing, 210094, China*

ARTICLE INFO

ABSTRACT

Occupant activity recognition (OAR) is essential for building management systems (BMS) to provide occupants with intelligent and comfort environments. Conventional sensing methodologies rely on burdensome wearables, privacy-risking cameras or specialized wireless devices. Pervasive existing Wi-Fi signals are a promising alternative and enable ubiquitous occupant sensing. In this paper, we propose a Wi-Fi-based OAR system called Wi-OAR that enables energy-efficient and user-centric services in smart offices. Its technical novelties are twofold. First, to recover activity-induced information, we innovatively present the fast and robust target component separation (FRTCS) algorithm regarding both time efficiency and high accuracy. Second, noting that handcrafted features can be inefficient and redundant, we develop an efficient feature selection algorithm based on class differences and information entropy. We prototyped the Wi-OAR system with only a pair of commercial Wi-Fi devices and implemented it in diverse office environments. The experimental results illustrate a consistent accuracy of over 96% in the different scenarios with considerable time cost savings. Further studies compare the system performance with prior approaches and discuss the influences of variables, which demonstrate the superiority of Wi-OAR.

1. Introduction

Concerns about energy efficiency and occupant comfort have inspired the development of building management systems (BMS) in smart buildings [1,2]. For office workers who spend excessive hours at the workplace, it is more appealing to sustainably work in smart offices [3]. Smart offices could interconnect smart devices indoors, integrate useful information and provide occupants with intelligent and comfort environments [4,5]. To establish energy-efficient and user-oriented smart offices, an occupant activity recognition (OAR) system is the key enabler [6–13]. For instance, an off-duty employee leaving the office in a hurry could be reminded to turn off lights and Heating Ventilation Air Conditioning (HVAC) systems for energy conservation. A sedentary white-collar worker could be regularly recommended to take a walk periodically depending on both health and comfort settings. To date, most innovations of OAR systems concentrate on wearable sensors [14, 15] and computer vision systems [16,17]. These two techniques have earned great reputations in both commercial and residential fields. However, occupants may feel encumbered by wearables or be afraid of privacy risks caused by cameras. Therefore, a contact-free,

privacy-preserving and ubiquitous OAR system is urgently needed.

With recent developments in wireless technologies, many researchers have exploited ambient signals from passive infrared (PIR) sensors [18], Bluetooth low energy (BLE) sensors [19], radio frequency identification (RFID) tags [20] and Wi-Fi devices [7–9] to realize OAR. Among these signals, Wi-Fi holds the most promise, as in-air Wi-Fi signals can noninvasively depict the physical environment and be easily collected from existing Wi-Fi devices. Hence, a Wi-Fi network can be empowered with sensing ability and pioneering new applications beyond communication. Mainstream efforts are devoted to Wi-Fi channel state information (CSI) exploitation because CSI measures the multipath variations influenced by the occupant at the granularity of the subcarrier level, and thus CSI-based OAR systems are inherently informative and fine-grained [21]. To utilize the hidden activity information involved in the CSI measurements, two pivotal blocks, i.e., a signal processing block and a feature processing block, deserve to be emphasized. For signal processing, the pursuits of target signal separation never stop, and several algorithms are proposed to alleviate this problem to some extent, including bandpass filtering [22], singular value decomposition (SVD) [23], wavelet transformation (WT) [24] and

principal component analysis (PCA) [25–30]. Unfortunately, few of them can simultaneously separate activity-induced signals from diverse backgrounds with relatively high computational speed. For feature processing, massive explorations on feature engineering (i.e., hand-crafted features based on domain knowledge [25–27,29–33]) and deep learning (i.e., automatic high-level feature extraction from multiple perception layers [34–38]) are inspiring, but either hardly distinguish redundant and inefficient features or rely on sophisticated hardware and large training sets.

In this paper, we propose a Wi-Fi CSI-based OAR system, called Wi-OAR, to enable energy-efficient and occupant-centric services in smart offices. Technically, we first describe the occupant dynamics using subcarrier correlations from noisy measurements and then innovatively present a fast and robust target component separation (FRTCS) method considering both the time efficiency and accuracy. To circumvent the shortcomings of feature engineering, the feature pool is then elaborately designed with distinct physical meanings. In addition, we develop a feature selection algorithm to search for representative and independent features based on class differences and information entropy. We prototyped the Wi-OAR system using only a pair of commercial off-the-shelf (COTS) Wi-Fi devices and implemented it in three office scenarios in a school building. According to the experimental results, Wi-OAR achieved constant high accuracy of 98.82%, 97.98% and 96.23% in the corridor, meeting room and student office, respectively. More comprehensive studies are conducted to compare our method with those of prior works and to discuss the influences of the algorithm parameters as well as the experimental variables. We envision Wi-OAR as a vital preamble to enlighten other user-centric revolutions in construction.

Our contributions are summarized as follows:

- We propose a Wi-Fi CSI-based OAR system based on existing COTS Wi-Fi devices to enable energy-efficient and occupant-centric services in smart offices.
- We are the first to present a novel FRTCS algorithm for accurate signal recovery in diverse indoor spaces that ensures high accuracy, time efficiency and robustness.
- A feature selection algorithm is innovatively designed to avoid inefficient and redundant features so as to improve the accuracy with lowest computational effort.
- Before purification, a robust indicator based on raw subcarrier correlation is designed to monitor the occupant activity and thus avoids the manipulation of the entire CSI measurements.
- Wi-OAR was implemented in diverse office scenarios, and extensive experiments were conducted to illustrate the superior performance of Wi-OAR in terms of the accuracy, time efficiency and robustness.

The subsequent sections are organized as follows. We first review the recent literature in Section 2 and then provide an overview of Wi-OAR in Section 3. Technical details of the signal processing and feature processing steps are illustrated in Section 4 and Section 5, respectively, followed by a discussion of the experimental evaluations in Section 6. Finally, we conclude this paper in Section 7.

2. Literature review

Humans spend more than 80% of their lives indoors and 6 h per day in the office [3,6]. Activity-based information on office workers is therefore the key enabler for building management systems (BMS) that can provide energy -efficient and user-oriented services, such as context-aware applications [7,8], intelligent environmental controls [9], construction energy reduction [10,11] and workforce safety guarantees [12]. In other words, if OAR is not included in BMS, smart offices will no longer be 'smart' because more manual efforts will be devoted to reducing energy consumption and improving working enthusiasm [13]. Recent advances in OAR partly fall into the wearable sensor category. For example, Akhavian et al. [14] captured a construction workforce's

activity information from inertial measurement units of carry-on smart phones. Cha et al. [15] investigated the feasibility of office activity recognition in the workplace using an accelerator. Despite their high accuracy, wearable sensors require always-on-body equipment, which renders them not ubiquitously applicable among occupants. Another branch of OAR is computer vision-based techniques. By collecting videos from depth cameras and analyzing continuous frames on a local server or in the cloud, high-precision occupant information can be obtained [16,17]. However, this method inherently suffers from a limited surveillance range (e.g., 4 m for Kinect), illumination interference, extra expenses and privacy-leakage problems. All of the above shortcomings have strongly motivated the investigation of appealing alternatives that can alleviate these issues.

Tremendous progress in wireless sensors has shown great potential in OAR. Basically, it provides a noninvasive option to depict physical environments using in-air signals, which can be generated by PIR sensors [18], BLE sensors [19], RFID tags [20] and Wi-Fi devices [7–9]. For instance, Dodier et al. [18] detected when and where residents occupied a commercial building with preinstalled low-cost PIR sensors. However, PIR sensors cannot perceive fine-grained activities. Wang et al. [19] utilized BLE signals for occupants' spatial distribution monitoring to automatically guide the operation of HVAC systems. Li et al. [20] presented an RFID-based multi-person detection scheme to reduce construction energy consumption. Nevertheless, both BLE sensors and RFID tags require specialized hardware and dense deployment. Wi-Fi-based OAR systems are the most appropriate choice for smart offices. By reusing pervasive communication signals from existing infrastructure, a Wi-Fi-based OAR system could deliver low-cost, noninvasive, omnidirectional and fine-grained detection. As a priority, we conduct a comprehensive survey on Wi-Fi CSI-based OAR in smart buildings.

Since body motions in a wireless environment can induce fluctuations in the received signal, how to process signals and features for robust activity-related information separation and preservation remains critical and challenging [21]. For signal processing blocks, several algorithms have been investigated to reveal the essence of activities in wireless environments, such as bandpass filtering [22], SVD [23], WT [24] and PCA [25–30]. Among them, PCA is the most attractive for its efficiency in revealing the significant part of motion data. Wang et al. [25] leveraged the second component of PCA to discern daily activities, gait patterns [26] and keystrokes [27]. Qian et al. [28] utilized the first dominant part of PCA to track residents' locations. However, its brittleness with respect to principal component selection puts its robustness in jeopardy. Some methods seek prompt validity by selecting sensitive subcarriers [29] or the cumulate variance [30], but they are vulnerable to environmental changes. This paper aims to develop a FRTCS method by eliminating both gross noise and background.

Another critical part is feature extraction. Some studies have resorted to dynamic time warping (DTW) to compare the similarity between waveforms, yet DTW increases the computational cost and is sensitive to slight changes [27,31]. Therefore, to ensure validity, most current studies compress the original waveforms with several representative features and train efficient classifiers. Early attempts computed the statistical features for time-frequency analysis [32]. Pu et al. [33] exploited Doppler frequency shifts (DFSs) to indicate human motions and aid in gesture recognition. Wang et al. [25] resorted to a discrete WT to analyze activities in the frequency bands of interest. Wang et al. [26] applied the short-time Fourier transform (STFT) algorithm to acquire abundant gait information. The above features or fusion of them are widely adopted, and a few studies emphasize the importance of feature selection [7,24,30]. Nevertheless, few studies have simultaneously considered feature robustness and redundancy. Wang et al. [34,35] first introduced deep learning models (DLMs) to extract invariant representations automatically. Wang et al. [36] reprocessed feature maps via self-organizing map (SOM) networks. However, DLMs require high computing power to handle a large amount of training data, which are inapplicable in the majority of office settings. Jiang et al. [37] and Zhang

et al. [38] borrowed the idea of transfer learning to reduce the cost of cross-site Wi-Fi sensing, but the implementation of these schemes is still limited by high-performance hardware and drastic environmental changes. This paper tries to select efficient and independent features with precise physical meanings.

3. Wi-OAR overview

In this section, the concept of Wi-Fi CSI is firstly introduced, followed by the technical challenges raised in practice. The system architecture of Wi-OAR is then provided to guide the solutions.

3.1. Background and challenges

With the booming demand for wireless data traffic, Wi-Fi networks are now pervasive in smart offices, allowing device-free ubiquitous OAR. The underlying rationale is that when an occupant engages in an activity, his/her body movements affect the Wi-Fi propagation and change the multi-path profile. To acquire activity-induced variation, CSI is extracted from Wi-Fi-enabled IoT devices equipped with a 5300 Network interface card (NIC) [39]. CSI is capable of representing the combined effect of, for example, scattering, fading, and power decay with distance. Mathematically, CSI can be modeled as

$$y = Hx + Noise, \tag{1}$$

where y and x denote the received and transmitted signal, respectively, and H is the estimated CSI matrix. As N_{tx} transmitting antennas broadcast continuous Wi-Fi signals into the air, complex CSI with total K subcarriers collected by N_{rx} receiving antennas at time t is written as

$$H_t = \left[H_t^1, ..., H_t^K \right]^T, K = N_{tx} \times N_{rx} \times 30 \tag{2}$$

where $H_t^n = |H_t^n| e^{j \angle H_t^n}$ with $|H_t^n|$ and $\angle H_t^n$ being the amplitude and phase values, respectively, of the n-th subcarrier. The interference caused by occupant activities would modify the channel response and thus induce temporal fluctuations in the amplitude and phase of the CSI measurements in (2). In this paper, since the 802.11n CSI tool [39] provides only 30 subcarriers for a 20-MHz Wi-Fi channel, we obtain a total of 90 subcarriers from 3 channels composed of 1 transmitter antenna and 3 receiver antennas. In addition, we employ only the amplitude information because precise phase information is extremely complex to recover due to hardware imperfections [40].

To accomplish an accurate and robust Wi-Fi CSI-based OAR system in smart offices, the following technical challenges should be addressed. (1) *How to sensitively detect the state changes in occupant movements from continuously received data.* It is nontrivial to indicate real changes from noisy measurements, and noise removal for the entire time series could incur extra computational costs. (2) *How to separate the precise activity-induced signal component from corrupted measurements.* This problem is challenging since disturbances in physical environmental changes and abrupt noise could generate similar fluctuations. In addition, irrelevant body movements may entangle the received signals. (3) *How to select efficient and independent features for robust activity classification.* The pursuit of representative features helps to reduce the dimensions, reveal the precise physical meanings and boost the system validity. However, it still remains a crux due to the unknown feature efficiency and redundancy.

3.2. System architecture

To alleviate the above issues, we innovatively present Wi-OAR, which mainly includes three blocks, i.e., a data collection block, a signal processing block and a feature processing block, as shown in Fig. 1.

3.2.1. Data collection block
This block is responsible for CSI data collection from each monitoring point (MP) antenna. For the time required for T sample packets, a $K \times T$ CSI amplitude frame would be obtained from continuous Wi-Fi measurements.

3.2.2. Signal processing block
This block processes the raw CSI amplitude in three steps. First, we calculate the CSI cross-correlation values (CCVs) in each second to judge the occupant status and compare the differences in adjacent time windows to refine the detection. Second, we develop a FRTCS method to recover precise activity information from the raw measurements. Third, we divide the signals into diverse frequency bands using the WT algorithm.

3.2.3. Feature processing block
This block aims to extract representative time-frequency domain and STFT image features from wavelet signals. Afterwards, we set the selection criterion to choose features with high-efficiency and low-redundancy scores in the offline phase. Finally, a softmax-based classifier is adopted to estimate the probabilities that online-extracted features fall into every activity class.

4. Wi-Fi signal processing

In this section, we present the details of the Wi-Fi signal segmentation, separation and transformation methods.

4.1. Signal segmentation

To accomplish intelligent OAR, developing a real-time, automatic and efficient signal segmentation method is the foremost task. Previous works using accumulated differential [41] or threshold-based slide windows [42] always implemented segmentation after the signal denoising step because these indicators are sensitive to noise [31]. However, this module must handle all CSI measurements and thus yields extra computational costs. Inspired by the key observation depicted in Fig. 2, we notice that the CCVs of adjacent subcarriers deserve exploration. When no occupant appears in the current environment, high CCVs are concentrated only on the matrix diagonal, and the distribution of the remaining values is relatively random. Once an occupant performs activities, the five adjacent subcarriers show similar variations, and most the CCVs exceed 0.9. Therefore, CSI CCVs can be utilized as robust indicators. To further refine the start and end times, we compute the mean average difference (MAD) between contiguous windows.

In Algorithm 1, we explain how CCV-based segmentation works. The input parameters are $A_{k \times T}$, δ_1, δ_2, and W. $A_{k \times T}$ is the collected amplitude frame with K subcarriers and time duration T. δ_1 and δ_2 denote the thresholds of coarse detection and fine detection, respectively. W is the

Fig. 1. Wi-OAR system architecture.
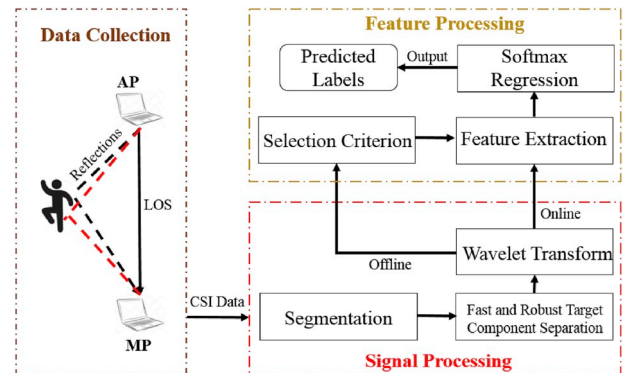
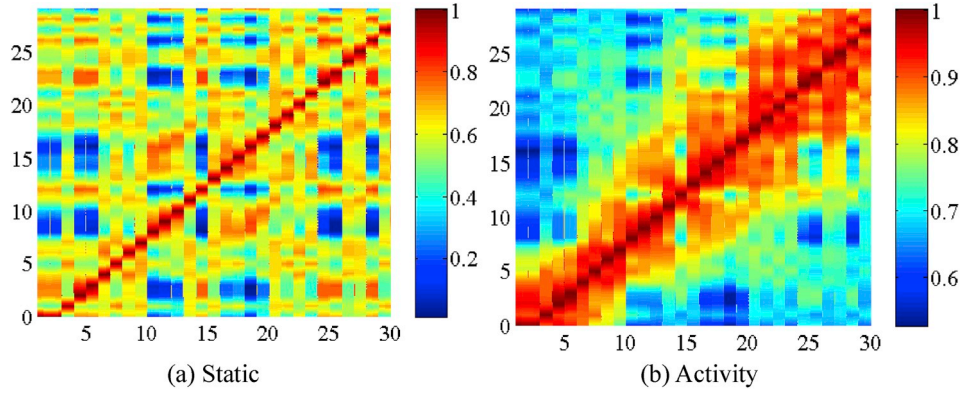(a) Static   (b) Activity

Fig. 2. Key observation: the cross-correlation value of adjacent subcarriers increases dramatically when an activity occurs.

fixed window size in the CSI frame without overlapping. Steps 3–4 and 11–13 are responsible for motion status detection by computing the mean CCV of all 30 subcarriers. Only when the motion status changes do steps 5–8 and 14–17 begin to determine the proper start time and end time through MAD comparisons. If the latter (former) window is obviously larger than the former (latter) window, then the latter (former) window is termed the exact start (end) time. In this paper, W, δ_1, and δ_2 are experimentally set to 0.2, 0.8 and 0.18, respectively.

Algorithm 1: Pseudocode for CCV-based segmentation

Input parameters: $A_{k \times T}$, δ_1, δ_2, W

---

1. t=1; start_point=false; [m,n]=size($A_{k \times T}$);

2. **while** $t \leq n$, calculate $\overline{CCV}$ for every second

3. **if** start_point==false, // Detect whether the activity has begun

4.   **if** $\overline{CCV}$>δ_1, start_point==true; // activity begins

5.   **for** j=1:1/W, // Determine the exact start time

6. $U(j)$=MAD(A(t+j*W)- A(t+(j-1)*W));

7.   **if** $U(j)$> δ_2, then t_start=t+j*W;

8.   **else** t_start=t, end if; end for

9.   **else** t+=1, return to step 2 // No activity occurs, keep monitoring

10. **else**, // Detect whether the activity has ended

11.   **if** $\overline{CCV}$>δ_1, t+=1, return to step 2; // continue

12.   **else**, // Activity ends

13.   start_point==false;

14.   **for** j=1:1/W, // determine exact end time

15. V(j)=MAD(A(t-(j-1)*W)- A(t-j*W));

16.   **if** $V(j)$> δ_2, then t_end= t-j*W;

17.   **else** t_end=t, t+=1, return to step 2;

18.   end if; end for;

19. **Output** t_start, t_end;

4.2. FRTCS-based separation

PCA enhances applications including data analytics, signal retrieval and noise reduction. It captures the majority of motion-induced variability using only a small number of orthogonal basis vectors from Wi-Fi signals [43]. However, canonical PCA is error-prone due to the presence

of interruptions [44]. In addition, there is no consensus on which component prevails over the others. Inspired by the development in low-rank decomposition, robust PCA (RPCA) is presented to decompose the given amplitude matrix A into a low-rank component $L \in \mathbf{R}^{k \times t}$, which denotes the stable background, and a sparse component $S \in \mathbf{R}^{k \times t}$, which acquires dynamic changes. The separation problem can also be expressed as a convex optimization problem:

$$\min_{L,S} \|L\|_* + \gamma \|S\|_1, \text{s.t. } A = L + S \qquad (3)$$

where $\|.\|_*$ and $\|.\|_1$ correspond to the nuclear norm and the sum of the elementwise absolute values of the matrix, respectively. This problem is commonly solved by the alternating direction method under the augmented Lagrangian multiplier (ADM-ALM) framework.

As pioneers, Wu et al. [45] first leveraged the above algorithm to handle Wi-Fi sensing through walls. We try to reproduce this state-of-the-art algorithm in our environmental settings. Specifically, the transmitter is set to broadcast Wi-Fi signals via the line-of-sight (LOS) path, and a volunteer walks along the midperpendicular LOS path. γ, the tolerance error ε and the maximum number of iterations are set to 0.1, 10^{-6} and 10^3 by default. The final results of RPCA are shown in the form of 3D mesh matrix. In Fig. 3(a), a trend of amplitude variation similar to that in Ref. [45] is observed. After separation, the low-rank component is supposed to record the properties of the physical environment, while the sparse component is prone to be affected by the activities that are occurring. However, we observe entangled low-rank components in Fig. 3(b), which reflect the multi-path profile caused by both physical environments and physical motions, while the sparse component shown in Fig. 3(c) consists of leaked activity information, random noises and irregular variations. Similar observations of signal separation could be obtained when other forms of RPCA (e.g., mixture of Gaussians RPCA (MoG-RPCA) [46] and motion-assisted matrix restoration MAMR [47]) methods are adopted and tested [48]. We argue that in Ref. [45], the RPCA algorithm was implemented in through-the-wall scenarios, where the direct propagation path is blocked. In such cases, the motions performed away from the Line-of-Sight (LOS) path could generate observable and correlated fluctuations, inducing distinguishable CSI separations in signal strength, sparsity and rank. In contrast, in smart office, occupants appear randomly that establishing diverse wireless environments, and their activities reveal diverse performance regarding the obstacle settings. Therefore, both the physical environment and occupant activities could yield similar low-rank information with a small number of singular values, while the sparse components mainly preserve the irrelevant variations. To separate the background, activity and noise information, we must restrict the rank of low-rank matrix L and cardinality of the sparse matrix S simultaneously.

Motivated by previous work [49], we present the FRTCS method to recover activity-induced signal variations efficiently. The separation problem can be alleviated by minimizing the decomposition error under
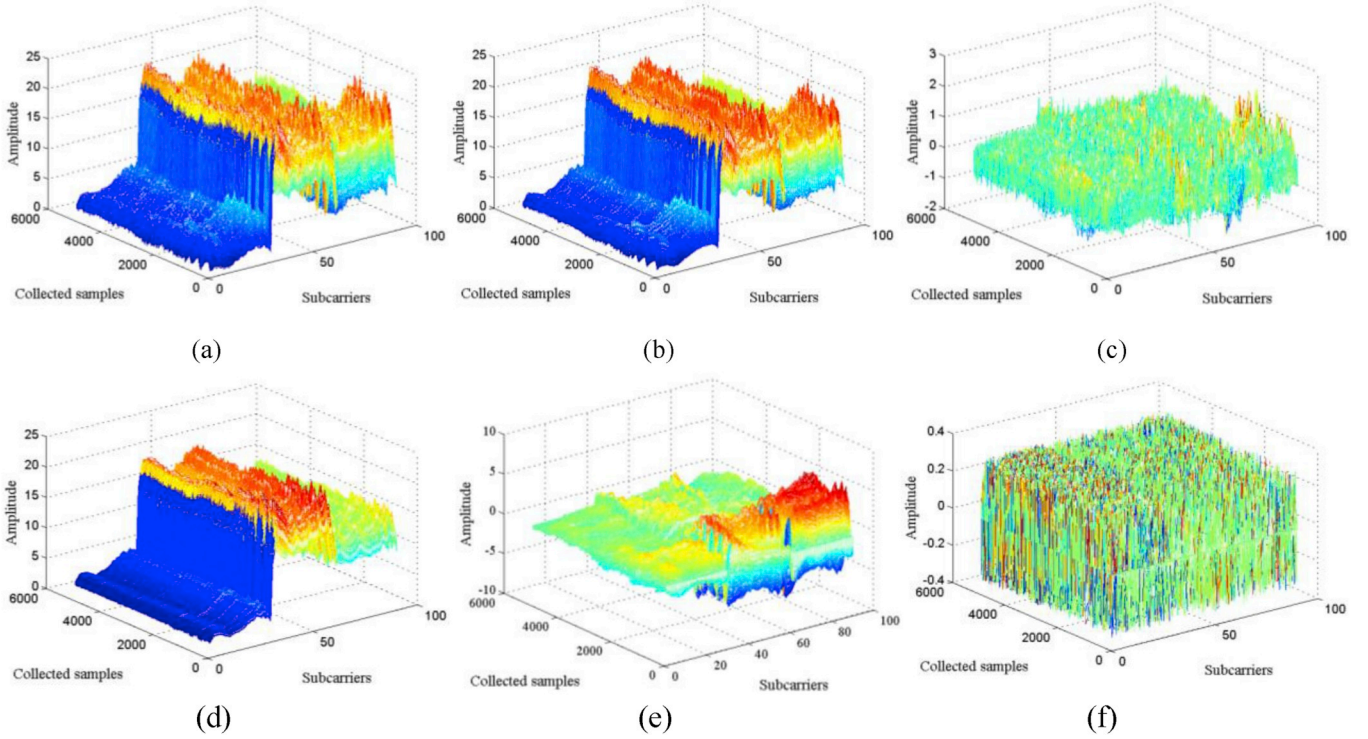
Fig. 3. Experimental studies on the TW-See method and our FRTCS method: (a) raw amplitude, (b) low-rank component separation using the TW-See method, (c) sparse component separation using the TW-See method, (d) low-rank component separation of the FRTCS method contains environmental information, (e) sparse component separation of the FRTCS method finds the dominant motion information and (f) random noise.

predefined hard constraints:

$$\min_{L,S}\|A - L - S\|_F^2, \text{ s.t. } A = L + S + G, \ rank(L) \le r, \ card(S) \le c, \quad (4)$$

where c denotes the nonzero numbers in S. Different from the above algorithms, the FRTCS method reduces the model complexity by constraining r and c and approximates the decomposition by alternatively solving the following two subproblems:

$$L_t = \min_{rank(L) \le r}\|A - L - S_{t-1}\|_F^2, \ S_t = \min_{card(S) \le c}\|A - L_t - S\|_F^2 \quad (5)$$

where the main computation efforts exist in the SVD operation towards $A\text{-}S_{t-1}$ in the updating of the L_t sequence (i.e., requiring a minimum of $k^2 t$ flops). To facilitate the process, bilateral random projection (BRP) of a dense matrix A is considered, and its rank-r approximation L is:

$$L = Y_1\left(B_2^T Y_1\right)^{-1} Y_2^T, \quad (6)$$

where $Y_1 = AB_1$ and $Y_2 = A^T B_2$ represent the BRP operation of A and $B_1 \in \mathbf{R}^{t \times r}$ and $B_2 \in \mathbf{R}^{k \times r}$ are independent random Gaussian/subsampled random Fourier transform (SRFT) matrices. Considering that the singular values of A may deteriorate, we apply the power scheme method [50] and construct $\tilde{A} = (AA^T)^q A$ with the same singular vector. Therefore, the BRP of $\tilde{A}$ updates as $Y_1 = \tilde{A}B_1$ and $Y_2 = \tilde{A}^T B_2$, and the low-rank matrix is recovered based on BRP as:

$$\tilde{L} = Y_1\left(B_2^T Y_1\right)^{-1} Y_2^T \quad (7)$$

Through the fast QR decomposition towards Y_1 and Y_2, the approximate low-rank matrix L of A is easily obtained:

$$L = \left(\tilde{L}\right)^{1/(2q+1)} = Q_1\left[R_1\left(B_2^T Y_1\right)^{-1} R_2^T\right]^{1/(2q+1)} Y_2^T \quad (8)$$

In Fig. 4, we summarize the flow chart of the FRTCS method. ε controls the reconstruction errors, q is an adjustable power scheme

parameter to avoid the degeneration of singular values, and P_Ω is responsible for the reservation of c nonzero elements in $A\text{-}L_t$. In this paper, q, r, and ε are set to 2, 3 and 10^{-6}, respectively. Fig. 3(d)–(f) displays the low-rank, sparse and noise matrix separation of raw amplitude, respectively, via the FRTCS method. Compared with RPCA,
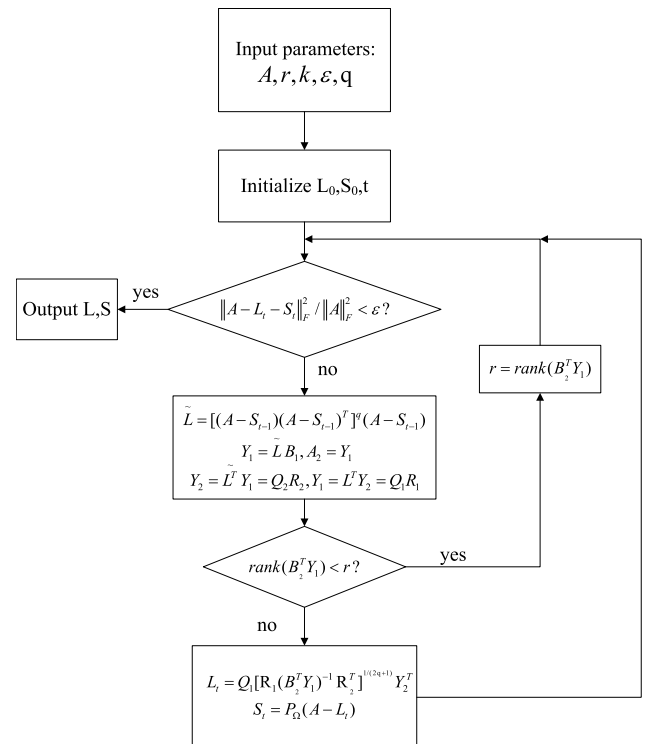

Fig. 4. Flow chart of the FRTCS method.

FRTCS provides more stable and smooth waveforms across the time (i.e., collected samples) and frequency domains (i.e., subcarriers) in Fig. 3(d). Fig. 3(e) further indicates that observable activity information can be recovered from overwhelmed signals, reducing the irrelevant disturbances shown in Fig. 3(f). In addition to better separation results, the FRTCS method also avoids the time-consuming process (i.e., SVD (A-S_{t-1})$_{k \times t}$), which requires only an SVD and inversion of an $r \times r$ matrix and five matrix multiplication operations in (8). The experimental studies and discussions of time cost are presented in Section 6. Afterwards, we use the first component for the subsequent process, which constantly incorporates over 95% of the signal variations from the sparse matrix.

4.3. Division into diverse frequency components

Before extracting useful features from the FRTCS signals, many promising solutions try to analyze activity signals from a multiscale perspective using WT [29], empirical mode decomposition (EMD) [51] and tensor decomposition [52]. In this paper, we use WT to decompose mixed signals into diverse frequency components. The basic intuition is that different body parts generate distinct reflections, which belong to certain frequency bands. Therefore, each decomposed component has its own clear physical meaning. In practice, we refer to Mallat's algorithm [53] and divide the i-th approximation signal into the (i+1)-th detailed signal (indicating high-frequency components reflected from legs and hands) and another approximation signal (denoting low-frequency components reflected from the torso). Specifically, we select the Daubechies 6 wavelet (db6) and divide the raw FRTCS signal into 6 frequency bands, which are shown in Fig. 5. As the sampling rate is set to 2000 Hz, the major part of motion-induced variation is recorded by the Detail 4, Detail 5, Detail 6 and Approximation 6 layers, respectively, while the remainder mainly contains the correlated but useless high-frequency outliers. Hence, we extract features from only the 4 informative signals below.

5. Activity feature generation

In this section, we present several handcrafted features and then recommend which features should be used. Finally, a softmax classifier is introduced to predict the activity labels.

5.1. Original feature extraction

Intrinsic features could reflect the physical meanings of signal changes and further clarify the motion patterns. To investigate which features should be adopted and constructed, we examine the related areas closely and summarize the feature categories [7,24]. In general, these features can be classified into common statistical features (i.e., time- and frequency-domain signals), transformation-based features (i.e., STFT images and DFSs), and deep features. Deep features are not considered in this work since they have abstract physical meanings and require a massive amount of training samples. Table 1 summarizes all the features that are commonly used in recent state-of-the-art methods. For statistical feature extraction, the *mean*, *minimum*, *maximum*, *standard deviation*, *median*, *entropy*, *skewness*, *kurtosis* and *histogram distribution* values of each time interval are calculated for each wavelet signal to accomplish light and efficient time-frequency analysis. For STFT image feature extraction, we create the STFT image from the entire FRTCS signal and calculate the *spectrogram entropy*, *torso speed*, *frequency ratio* and *fractal dimension* to indicate body movements. We refer interested readers to our previous work for more technical details in the STFT process [29].

5.2. Feature selection

With the summarized features above, $\partial \times \beta \times \eta$ features for an activity have been pooled:

Table 1
Original feature pool.

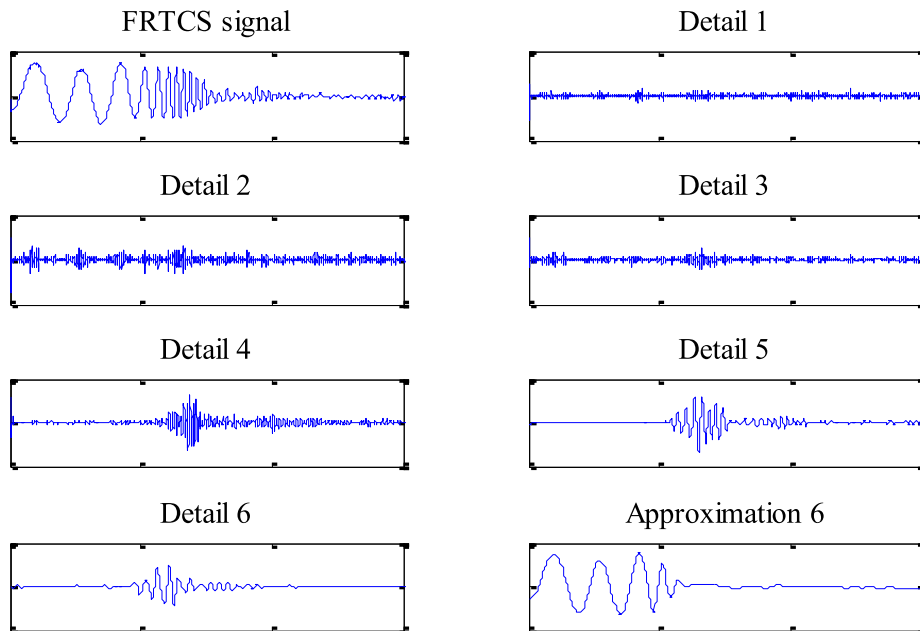| Feature Category	Feature Description
Statistics features of the time-frequency signals	mean[1], min[2], max[3], std[4], median[5], entropy[6], skewness[7], kurtosis[8] histogram distribution[9]
STFT image features	spectrogram entropy[10] torso speed (mean[11], min[12] max[13], std[14]) frequency ratio[15] fractal dimension[16]

Fig. 5. Wavelet transformation of the FRTCS signal.

$$\Phi = \begin{bmatrix} F_{1,1}^1 & \cdots & F_{1,\eta}^1 & \cdots & F_{\beta,1}^1 & \cdots & F_{\beta,\eta}^1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ F_{1,1}^\alpha & \cdots & F_{1,\eta}^\alpha & \cdots & F_{\beta,1}^\alpha & \cdots & F_{\beta,\eta}^\alpha \end{bmatrix}, \tag{9}$$

Each row contains one feature extracted from diverse time intervals of the wavelet signals, where ∂, β and η indicate the index of the features, waveforms and time intervals, respectively. However, not all features included in Φ characterize the essence of the signals, and redundant features may induce the curse of dimensionality as well as decrease both the time efficiency and accuracy. Hence, we should select proper features based on the following three principles: (1) *low cost* (achieving the highest accuracy with the smallest number of features), (2) *robustness* (the minimum differences with interferences, e.g., scenario changes and occupant diversity, in the same activity class) and (3) *uniqueness* (the maximum differences between activity classes, e.g., '*walk*' and '*run*'). The relief-based algorithm (RBA), as a kind of filtering method that offers a fast and independent instance-based learning solution to explore feature interactions [24]. However, it has been noted that reduced power is needed to handle multiclass attributes of activities (e.g., scenario categories, occupant diversity and performance times) [54]. Therefore, a novel modified information entropy-based multiclass RBA (MIE-MRBA) is presented to address the entire training set J. First, we reformulate the intra-activity variation $J[N(C_H)]$ as:

$$\sum_{\substack{C_H = \\ \text{class}(R_i)}} \left[\frac{P(C_H)}{P(\text{class}(R_i))} \sum_{j=1}^{\kappa} \text{diff}\left(F, R_i, N_j(C_H)\right) \right] \Big/ (m \times \kappa) \tag{10}$$

and the interactivity difference $J[N(C_M)]$ equals:

$$\sum_{\substack{C_M \neq \\ \text{class}(R_i)}} \left[\frac{P(C_M)}{1 - P(\text{class}(R_i))} \sum_{j=1}^{\kappa} \text{diff}\left(F, R_i, M_j(CM)\right) \right] \Big/ (m \times \kappa) \tag{11}$$

where $N_j(C_H)$ and $M_j(C_M)$ represent the j-th nearest neighbor in the same activity category C_H and different activity category C_M, respectively, with a target sample R_i. Subformula diff $(F, R_i, N_j(C_H))$, for example, calculates the distance between the target sample and the intraclass j-th nearest sample. Parameters m and κ denote the number of sampling times and the number of nearest neighbors, respectively. The prior probability of target samples is used to weight the feature difference. Then, the vector $W(F)$ recurrently updates the feature scores to estimate the contribution of features:

$$W(F) = W(F) - J[N(C_H)] + J[M(C_M)], \tag{12}$$

Features with scores greater than the threshold ξ are preserved in the subset, while those with scores less than ξ are excluded. To further remove redundancies, we develop an information entropy $E(.)$-based indicator θ for a feature subset ν:

$$\begin{aligned} \theta(\nu) &= \frac{1}{1 + e^{-\sum W(\nu)}} - Red(\nu) \, Red(\nu) \\ &= \frac{1}{|\nu|} \sum_{\chi_i, \chi_j \in \nu} \left(\frac{2I(\chi_i, \chi_j)}{E(\chi_i) + E(\chi_j)} \right) I(\chi_i, \chi_j) \\ &= E(\chi_i) + E(\chi_j) - E(\chi_i, \chi_j) \end{aligned} \tag{13}$$

where $I(.)$ shows the relevance between variables χ_i and χ_j and $|\nu|$ denotes the number of features in ν. With low contribution and high redundancy, features are more likely to be estimated as redundant. We summarize the pseudocode of the MIE-MRBA in Algorithm 2, which denotes multiclass feature selection from step 2 to step 7, and further searches the redundant features from step 10 to step 17. In particular, we

set the tolerance value σ to seek the feature subset ν_f, which could maximize the overall score of θ. Thus, our proposed algorithm could take both the operational efficiency and redundancies into consideration.

Algorithm 2: Pseudocode of the MIE-MRBA

Input parameters: J, m, k, ξ, v_t, v_f, σ

1. Initialize $W(F=1:\alpha)=0.0$, $i=0$, $j=0$, $v_t=0$, $v_f=0$;

2. **for** i=1 to m

3.   Choose a sample R_i at random

4.   **for** every C_H=class(R_i), find the k-nearest samples $N(C_H)$; end for;

5.   **for** every $C_M \neq$class(R_i), find the k-nearest samples $M(C_H)$; end for;

6.   **for** F=1: ∂, $W(F)=W(F)$-$J(N(C_H))$+$J(M(C_H))$;

7.   end for;

8. end for;

9. **Output** features over threshold ξ, obtain v_0 and W_0

10. **while** $|v_0| >0$ // *the number of features exceeds zero*

11.   **for** j=$|v_0|$ to 1, // *backward sequence*

12. $v_t=v_0-F_i$, // *delete the feature with a relatively low score*

13.   **if** $\theta(v_t)$- $\theta(v_0)> \sigma$, $v_t \rightarrow v_f$; //*set tolerance σ to avoid the local optimum*

14.   **else if** $\theta(v_t)$- $\theta(v_0)> 0$, save v_t and $\theta(v_t)$, $v_t \rightarrow v_0$,

15.   **else** $v_t \rightarrow v_0$, end for

16. find all the saved v_t in step 14 to compute argmax[$\theta(v_t)$], $v_t \rightarrow v_f$

17. **Output** v_f and W_f;

In this paper, we normalize the values in the feature pool from 0 to 1 for machine learning and select the most discriminative features in the offline training phase. Experimentally, the final feature subset ν_f including 10 features of the d-th activity for the e-th sample (i.e., the *standard deviation*, *median*, *skewness*, *kurtosis* and *histogram distribution* of wavelet signals, and *spectrogram entropy*, *mean and standard torso speed*, *frequency ratio* and *fractal dimension* of the STFT image) is constructed.

5.3. Activity classification

To manipulate the selected features, several classification techniques have been exploited in previous studies, such as softmax regression (SR) [36], naïve Bayes (NB) [55], support vector machine (SVM) [29], K-nearest neighbor (KNN) [31] and random forest (RF) classifiers [23]. Considering that the multiclass classification problem can be simplified as a minimization problem,

$$\overline{d} = \underset{d,e}{\text{argmin}} ||\nu - \nu_f(d,e)||_2, \tag{14}$$

where ν denotes the feature subset extracted online and $\nu_f(d,e)$ represents the stored fingerprint for the d-th activity of the e-th sample. The SR algorithm, which estimates the possibility $\overline{d}$ that falls into every activity class, is an ideal choice to handle the above problem [36]. In the offline training phase, we utilize the training samples $\nu_f(d,e)$ to estimate the parameter matrix Θ of the SR model, and in the online phase, we can calculate the probabilities using the selected feature subset ν and parameters to indicate which activity an occupant performs:

$$h_\Theta(\nu) = \Theta \times \nu, \tag{15}$$

The class that has the largest element value is estimated as the true activity.

6. Experimental study

6.1. Experimental configuration

6.1.1. Hardware and software setup

We establish a prototype system based on available Wi-Fi devices to evaluate the performance of Wi-OAR. Specifically, two ThinkPad laptops equipped with Intel 5300 NICs and the Ubuntu 11.04 operating system are employed: one with an external antenna acts as the access point (AP) to transmit Wi-Fi signals in air, and the other one with three external antennas acts as the MP responsible for data collection. Both *trans*-receivers are preinstalled 802.11n CSI tools to record CSI measurements on channel 165 at 5.825 GHz in monitoring mode. The MP antennas are linearly arranged with a spacing distance of 2.6 cm, which is equal to half of the wavelength. The sampling rate is set to 2000 Hz, and further study on the effect of sampling rate is discussed through downsampling. During the training phase, a domestic laptop equipped with only an Intel i7-5700HQ 2.70 GHz CPU is employed to run MAT-LAB 2017b to validate the time efficiency.

We illustrate the layout and real photos of the three experimental scenarios in Fig. 6. In the corridor scenario, there is no block in the signal path in Fig. 6(b). In Fig. 6(c), we carry out the implementation in a 6 m × 6 m meeting room with multiple chairs and a long table nearby. In Fig. 6(d), the *trans*-receiver pair is built in a 6 m × 12 m student office, where the signal reflections are severely affected by the blocks. The distances between the AP and MP are set to 2 m, 3 m and 4 m, respectively.

6.1.2. Dataset description

In total, 10 volunteers from the College of Civil Engineering were recruited: 3 staff members and 7 students. They had no prior knowledge about so-called benchmark activities and diverse activity habits in their daily lives. During the data collection phase, both laptops worked in monitoring mode and continuously sent packets. All the participants were asked to perform 10 predefined common activities in the wireless environment, including 5 whole-body movements: (1) walking, (2) running, (3) jumping, (4) standing up, (5) sitting down, and another 5 partial-body, seated activities: (6) typing words on the keyboard, (7) writing, (8) calling, (9) drinking, and (10) stretching. Each activity was performed ten times in each scenario within a day, and we carried out the experiments on different days in a week. The samples collected from Monday to Wednesday were utilized to train the machine learning models, while the last four days of samples (i.e., 12,000 samples in total) composed the validation test.

6.2. Performance evaluation

6.2.1. Evaluation metrics

First, we determine the proper metrics to evaluate the system performance. (1) *Confusion matrices* are used to visualize the number of instances that an estimated activity corresponds to each class. The columns indicate the true labels of the activity, and the rows indicate the activities as predicted by Wi-OAR. Each element of the matrix represent the activity in the column is identified as the activity in the row. (2) *Accuracy* is the fraction of instances of activity X that are correctly identified as activity X [13]. High accuracy means a sufficient number of activity samples are correctly identified, while robustness denotes the high-accuracy performance under diverse conditions (e.g., testing scenarios, user diversity and sampling rates) [56]. (3) *Time cost* records the exact running time. Time efficiency means that the system achieves high-accuracy performance with the least time cost.

6.2.2. Overall performance

Taking all technical components into consideration, the capability of the system for high accuracy and robustness is the top concern. As the confusion matrices in Fig. 7 show, Wi-OAR achieves consistently high
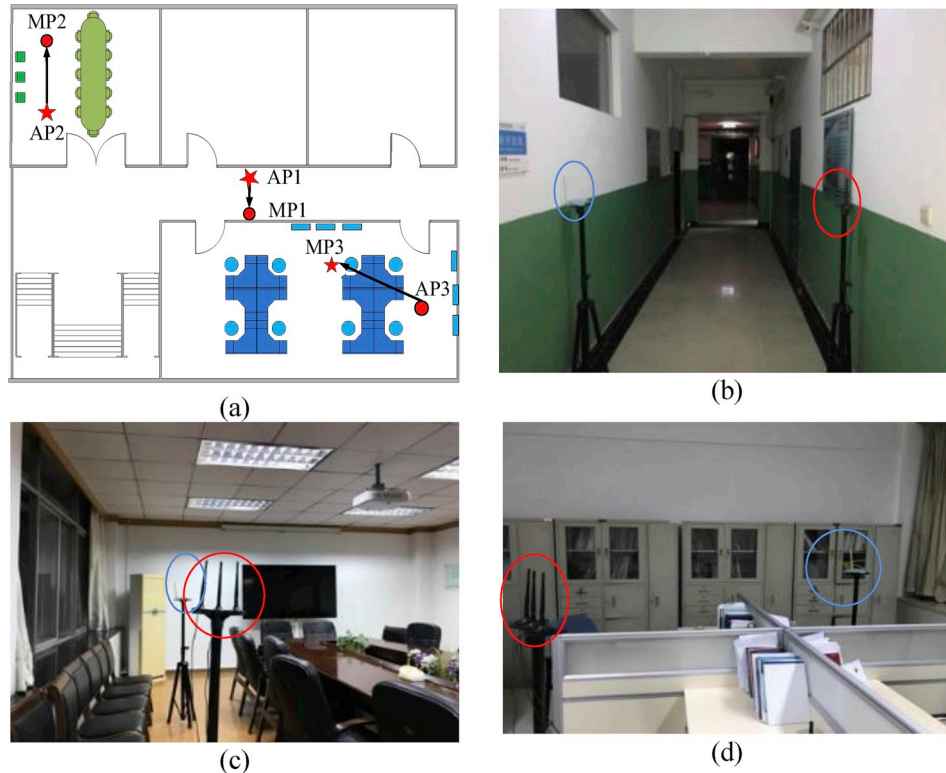

Fig. 6. Layout and photos of the three experimental scenarios. (a) Overall layout of the experimental scenarios on the 6th floor of a school building. (b) Photo of the corridor. (c) Photo of the meeting room. (d) Photo of the student office. Red circles and blue circles denote the locations of the receiving and transmitting antennas, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
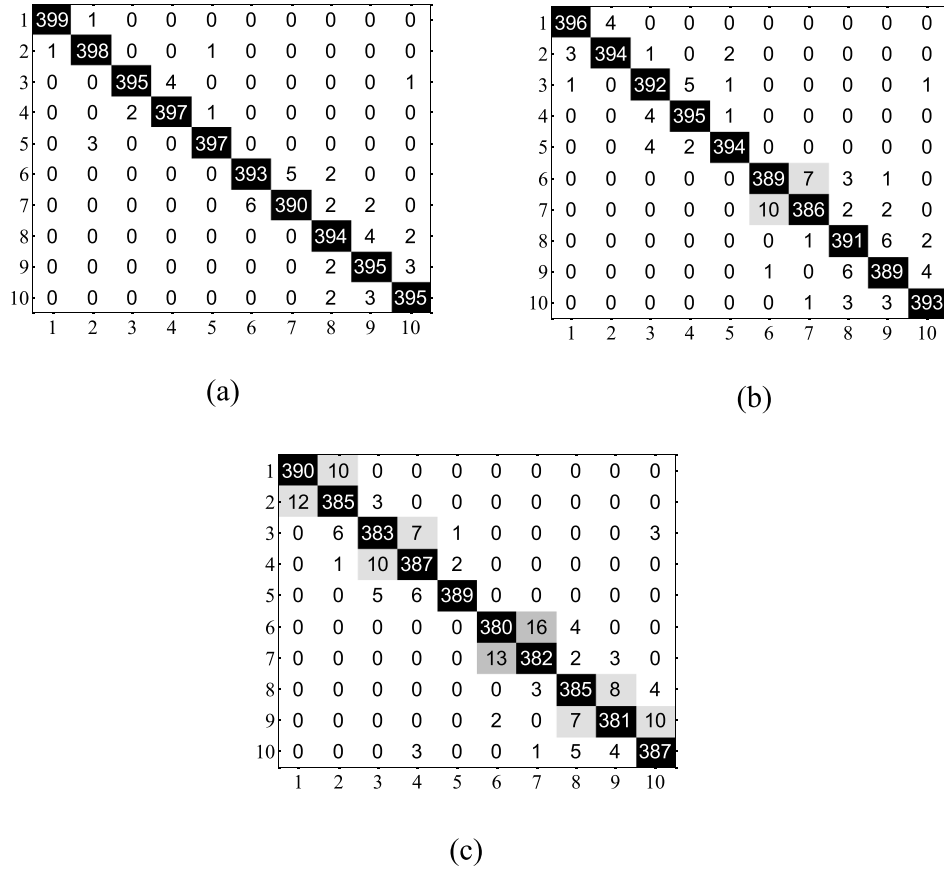
(a)

(b)

(c)

Fig. 7. Confusion matrix of (a) the corridor, (b) the meeting room and (c) the student office. The X-axis and Y-axis denote the estimated activity and ground truth, respectively.

performance for all ten activities across the three scenarios. In Fig. 7 (a), for example, 399 instances of 'walking' instances (i.e., label 1) are correctly identified, while only 1 instance is mis-identified as 'running' (i.e., label 2). 'Typing words' (i.e., label 6) and "writing" (i.e., label 7) are the most similar activities but only produce 7 and 10 incorrect labels, respectively. In Fig. 7 (b), the system performance drops slightly when the indoor space becomes complex. More similar activities are prone to misclassification, but the performance is still satisfactory. For example, the recognition accuracy of "writing" is at 96.5% (i.e., 386/400 × 100%). In Fig. 7 (c), the system performance decreases because the Wi-Fi propagation path is blocked but is still sufficient to identify similar activities. For example, there are 382 instances in which 'writing' is correctly identified, yielding an accuracy at 95.5%, while only 18 instances are incorrectly identified. Based on the above observations, we draw three important conclusions. First, if the direct propagation path is clear and the surroundings remain relatively stable, the signal fluctuations could convey real activity information, and thus the system could produce the most correct activity labels. Second, whole-body movements are more distinguishable than partial-body movements, whose accuracy is 0.95%, 1.5% and 0.95% higher. This is because whole-body movements could affect more propagation paths than partial-body movements and carry more information in signals. Third, similar activities are easily confused in highly variable wireless environments. This is because similar movements generate similar signal patterns, and the discriminative part could be violated due to background noises. The results of the confusion matrices illustrate that our system could handle diverse environmental influences with high accuracy and robustness.

6.2.3. Performance of the FRTCS method

To fairly illustrate the performance of FRTCS, we implement the FRTCS and three other RPCA algorithms (i.e., cardinal RPCA [45],

MoG-RPCA [46] and MAMR [47]) for comparison. Cardinal RPCA [45] works well with respect to grossly corrupted observations through low-rank decomposition, while MoG-RPCA [46] and MAMR [47] improve accuracy by modeling the background and motion information. In Fig. 8 and Table 2, we refer to the settings of TW-See Ref. [45], which considers the sparse component as the activity-induced part and low-rank component as the background and shows competitive performance in the third experiment scenario (non-line-of-sight (NLOS)), approximately 90%. Due to the random appearances induced by occupant activities in the first and second LOS scenarios, the background and activity variations are entangled in the low-rank matrix and produce the lowest accuracy. MoG-RPCA is expected to model the background noises with the MoG noise, and we set the parameters as the original paper suggests [46]. In both graphics, the MoG-RPCA produces a balance between accuracy and time cost, with the third highest average accuracy and the second lowest time cost; however, the occurrence of non-Gaussian strikes and the settings of multiple hyper parameters could deteriorate the performance. MAMR seems theoretically possible
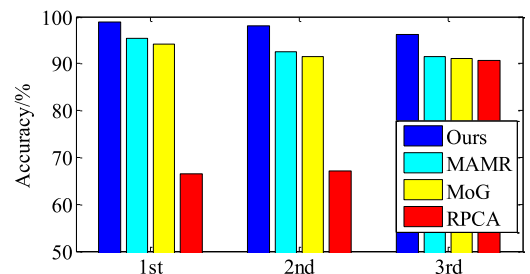

Fig. 8. Accuracy comparison for the MAMR, MoG, RPCA and FRTCS methods.

Table 2
Time cost comparisons for 1 s of CSI.

	FRTCS	MAMR	MoG	RPCA
Time/s	0.86	5.69	2.32	3.82

because it leverages the optical flow (i.e., gradient descent values between adjacent 1-s frames) to enhance the motion information in the observation matrix, and indeed achieves the second highest accuracy in all 3 scenarios. However, it has a large time cost due to the time-consuming frame-by-frame calculation. Compared with the other three state-of-the-art methods, with restricted rank and sparsity, the FRTCS method achieves consistently high accuracies of 98.82%, 97.98% and 96.23% in the first, second and third experimental scenarios, respectively, a significant advance compared with the other forms of RPCA. Meanwhile, the BRP-based FRTCS only takes 0.86 s to separate the background, activity and noise matrix for 1 s of CSI segments, markedly faster than the other SVD-based decomposition methods.

6.2.4. Performance of the MIE-MRBA

To evaluate the performance of MIE-MRBA implemented in Wi-OAR, we compare MIE-MRBA with the original ReliefF algorithm [24] and without feature selection. Specifically, the 'without selection' (W/O) means all handcraft features are utilized for activity classification, while the ReliefF adopted in literature [24] ranks the feature scores in descending order and selects the top-10 efficient features. In Fig. 9, we notice that without feature selection, the system accuracy is the lowest in each scenario. The main reason is that different signal features show varying sensitivity to environmental changes. Thus, some features are error-prone and some features are helpful. Applying the original ReliefF algorithm and selecting the most efficient features is a straightforward means of exploiting these efficient features. The observations in Fig. 9 show the amelioration of the performance. However, we argue that the top-10 efficient features may not be totally independent and are vulnerable to the same kind of interferences. Therefore, if we could remove the redundant features and supplement different kinds of efficient features, the system robustness and accuracy could both be ensured. Based on this motivation, we design MIE-MRBA, which considers both feature efficiency and redundancy and yields 1.0%, 3.5% and 3.7% higher accuracy than the original ReliefF. Moreover, MIE-MRBA provides large increases in accuracy in the complex wireless environment, which indicates the importance of feature diversity in handling environmental disturbances.

6.3. Parameter studies

6.3.1. Segmentation parameters δ_1 and δ_2

The former threshold controls the approximate length of segments, while the latter determines the detailed start points and end points of activities. Large δ_1 and δ_2 values help reduce the data size and amount of misinformation but increase the risk of negligence when short-term activity occurs. Therefore, we should carefully select thresholds
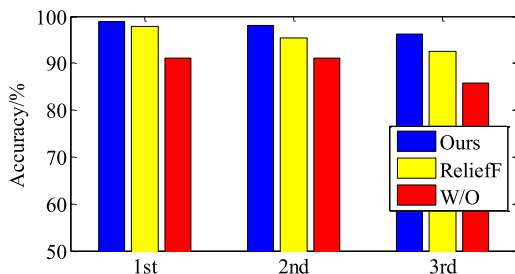
Fig. 9. Comparison of the accuracy of the MIE-MRBA with that of ReliefF and that without feature selection.

through experimental study. As shown in Fig. 10, we note that a proper δ_1 contributes greatly to the system, while δ_2 just fine-tunes the accuracy, which is responsible only for millisecond adjustments. In this paper, we set δ_1 to 0.80 and δ_2 to 0.18 due to the relatively high accuracy and considerable time efficiency.

6.3.2. Parameters r and c in the FRTCS method

First, we illustrate the time costs when we recover the signal matrices for different rank r and cardinality c values. In Table 3, we notice that with the increase in r and c, Wi-OAR spends more time separating the low-rank matrix as well as the sparse matrix because, on the one hand, the background matrix is naturally low-rank owing to the limited antenna number. On the other hand, since occupant activities cannot exert an impact on every packet and every subcarrier, the search for too many nonzero elements in the sparse matrix is unnecessary. To reduce the time cost, we only consider c values of $0.9^4_{\times 10}$ and $1.1^4_{\times 10}$ because all these parameter settings control the time cost within 1 s.

It is also noteworthy that in Fig. 11, $c = 1.1^4_{\times 10}$ produces an average accuracy over 93% with rank changes and achieves the highest accuracy when $r = 3$. Although $c = 0.9^4_{\times 10}$ has better time efficiency (i.e., 60 ms faster than $c = 1.1^4_{\times 10}$), its accuracy is obviously lower, at approximately 90%. To ensure both time efficiency and accuracy, we set $r = 3$ and $c = 1.1^4_{\times 10}$.

6.3.3. Number of features in ν_0

Intuitively, more efficient features could convey more information and contribute to the overall performance. As we expect, Fig. 12 shows an increase when the number of features increases from 8 to 10 and achieves the best accuracy with 10 features, followed by a slight drop when the number of features exceeds 10. This phenomenon is because insufficient features cannot depict activities from multiple aspects and further result in unstable performance, while superabundant features may be redundant and are prone to overfitting. Hence, 10 features are sufficient to realize a robust OAR system.

6.4. Discussion

6.4.1. User diversity

To demonstrate the aptness of Wi-OAR for diverse occupants, we recruited 10 participants (8 males and 2 females) to test Wi-OAR. Body mass index (BMI) was employed to indicate the height, weight and body shape, as shown in Table 4. Note that the participants also had different health conditions, and they did not receive special trained and performed activities according to their natural habits. Fig. 13 displays the accuracy affected by the users' BMI and sex. As shown, there exists no distinct correlation between the system accuracy and user ID. In other
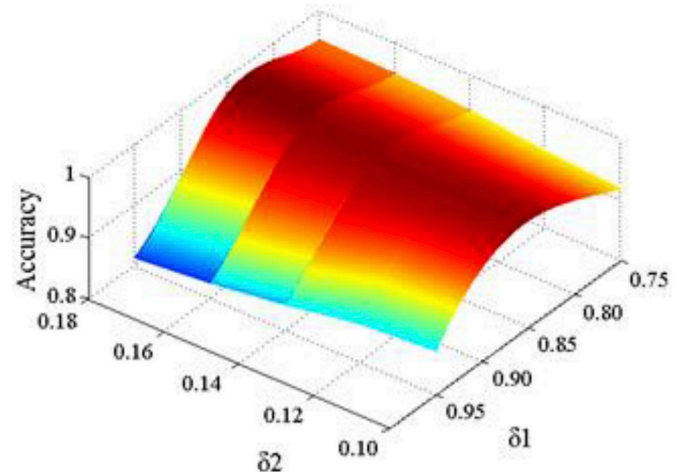
Fig. 10. Accuracy affected by the segmentation thresholds δ_1 and δ_2.

words, Wi-OAR can achieve a consistently high performance when encountering diverse occupants.

6.4.2. Sampling rate

A proper sampling rate resolves the contradiction between energy cost and information abundance. To determine the minimum sampling rate, we initialized the sampling rate at the MP as 2000 Hz and conducted downsampling operations towards the CSI collections to 1500 Hz, 1000 Hz, and 500 Hz. From Fig. 14, we note that whole-body activity recognition is robust to a low sampling rate because the fluctuations caused by whole-body movements convey more information in 1 s than partial-body movements. However, partial-body activities can generate fast and slight variations, and thus, a low sampling rate may miss some key information, which accounts for the deterioration in the system performance. In this paper, to ensure both whole-body and partial-body activity recognition performance, we use a 2000-Hz sampling rate, which is also reasonable in practice. However, regarding the

Table 3
Time cost comparisons for different r and c values.

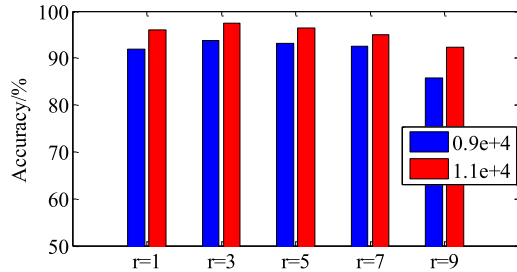| r	c				
	$0.9^4_{\times10}$	$1.1^4_{\times10}$	$1.3^4_{\times10}$	$1.5^4_{\times10}$	$1.7^4_{\times10}$
1	0.686 s	0.768 s	0.967 s	1.028 s	1.189 s
3	0.783 s	0.821 s	0.972 s	1.121 s	1.248 s
5	0.851 s	0.901 s	1.006 s	1.178 s	1.311 s
7	0.898 s	0.934 s	1.022 s	1.225 s	1.523 s
9	0.912 s	0.985 s	1.262 s	1.267 s	1.583 s

Fig. 11. Accuracy affected by the rank r and cardinality c.
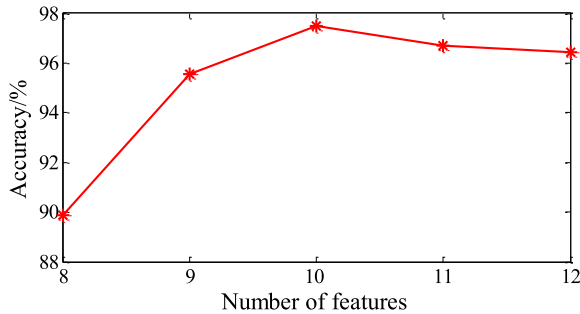

Fig. 12. Accuracy affected by the number of features.

Table 4
Diverse occupant identities in terms of BMI and sex.

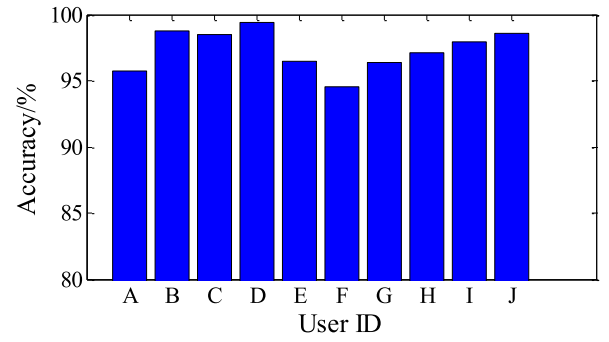| ID	BMI	Sex	ID	BMI	Sex
A	20.2	Female	F	21.7	Male
B	18.7	Female	G	20.8	Male
C	18.6	Male	H	20.5	Male
D	24.2	Male	I	21.6	Male
E	20.8	Male	J	21.8	Male

Fig. 13. Impact of user diversity.

problem of long-term behavior monitoring, we suggest lowering the sampling rate to avoid a time-consuming process.

6.4.3. Passerby interferences

To evaluate the interference of passersby, during the single-occupant experiment, we asked another volunteer to stand still, perform micromovements (e.g., random hand gestures) and perform macromovements (e.g., squatting and standing up) in the middle of the direct propagation path. We observe comparable results, which are shown in Fig. 15; the static body blocks the direct path but induces few errors in the final results. The performance drops significantly when dynamic passerby interferences occur, especially for the recognition of partial-body movements because the reflections from limbs are weak and are easily disturbed by irrelevant but dominant movements. For example, when the target occupant performing a partial-body activity encounters another occupant performing macromovements, the system accuracy would drop to only 57%. Since concurrent activity recognition remains an open challenge, we leave it as our future work.

7. Conclusion

In this paper, we propose an OAR system using existing Wi-Fi devices named Wi-OAR that enables ubiquitous occupant-centric services in smart offices. Our technical contributions are threefold. First, a novel target component separation algorithm for Wi-Fi-based OAR is presented that ensures high accuracy, robustness and time efficiency. Second, a novel feature selection algorithm is designed to avoid feature inefficiency and redundancy. Third, a subcarrier correlation-based indicator is developed to sensitively monitor the presence of occupant activity. We implemented the Wi-OAR in real scenarios and validated the system performance through extensive experiments. The results show the high accuracy and robustness of Wi-OAR, with 98.82%, 97.98% and 96.23% accuracy in a corridor, meeting room and student office, although the performance slightly drops when identifying similar activities in a severe indoor environment. Taking 'writing in the student office' as an example, 382 instances are correctly identified, and only 18
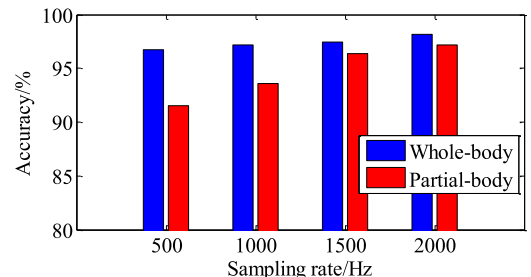

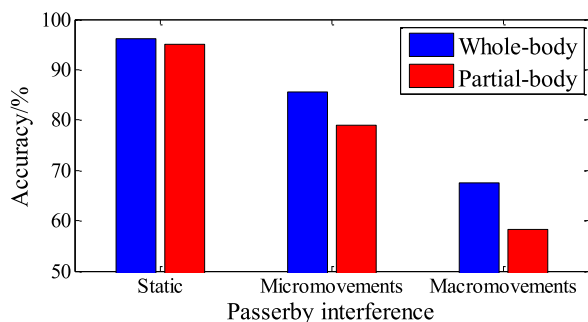**Fig. 14.** Impact of the sampling rate.

Fig. 15. Impact of passerby interference.

instances are incorrectly identified. To further overcome the challenge of similar activity types, more distinguishable features should be discovered, and more manual efforts are needed. In such cases, we expect that a combination of deep learning methods can realize automatic feature extraction and selection with the least human intervention. In addition, the system is vulnerable to multi-person occurrence or multi-variable environments. In the future, we will attempt to integrate multi-type smart sensors indoors for multi-modal awareness.

Funding

Declarations of competing interest

None.

Acknowledgements

References

[1] C. Bolchini, A. Geronazzo, E. Quintarelli, Smart buildings: a monitoring and data analysis methodological framework, Build. Environ. 121 (2017) 93–105, https://doi.org/10.1016/j.buildenv.2017.05.014.

[2] T.M. Lawrence, M.C. Boudreau, L. Helsen, G. Henze, J. Mohammadpour, D. Noonan, D. Patteeuw, S. Pless, R.T. Watson, Ten questions concerning integrating smart buildings into the smart grid, Build. Environ. 108 (2016) 273–283, https://doi.org/10.1016/j.buildenv.2016.08.022.

[3] C. Haslam, A. Kazi, M. Duncan, Process evaluation of A tailored workplace intervention designed to promote sustainable working in A rapidly changing world, Ergonomics (2019) 1–10, https://doi.org/10.1080/00140139.2019.1614212.

[4] C. Le Gal, J. Martin, A. Lux, J.L. Crowley, Smart office: design of an intelligent environment, IEEE Intell. Syst. 16 (2001) 60–66, https://doi.org/10.1109/5254.941359.

[5] M. Tehseen, H. Javed, A. Mehmood, M. Amin, I. Hussain, B. Jan, Multi modal aptitude detection system for smart office, IEEE Access 7 (2019) 24559–24570, https://doi.org/10.1109/access.2019.2893202.

[6] N.E. Klepeis, W.C. Nelson, W.R. Ott, J.P. Robinson, A.M. Tsang, P. Switzer, J. V. Behar, S.C. Hern, W.H. Engelmann, The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants, J. Expo. Sci. Environ. Epidemiol. 11 (2001) 231–252, https://doi.org/10.1038/sj.jea.7500165.

[7] H. Zou, Y. Zhou, J. Yang, C.J. Spanos, Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT, Energy Build. 174 (2018) 309–322, https://doi.org/10.1016/j.enbuild.2018.06.040.

[8] H. Zou, Y. Zhou, J. Yang, C.J. Spanos, Towards occupant activity driven smart buildings via WiFi-enabled IoT devices and deep learning, Energy Build. 177 (2018) 12–22, https://doi.org/10.1016/j.enbuild.2018.08.010.

[9] H. Zou, Y. Zhou, H. Jiang, S.-C. Chien, L. Xie, C.J. Spanos, WinLight: a WiFi-based occupancy-driven lighting control system for smart building, Energy Build. 158 (2018) 924–938, https://doi.org/10.1016/j.enbuild.2017.09.001.

[10] T.A. Nguyen, M. Aiello, Energy intelligent buildings based on user activity: a survey, Energy Build. 56 (2013) 244–257, https://doi.org/10.1016/j.enbuild.2012.09.005.

[11] J. Wang, N.C.F. Tse, J.Y.C. Chan, Wi-Fi based occupancy detection in a complex indoor space under discontinuous wireless communication: a robust filtering based on event-triggered updating, Build. Environ. 151 (2019) 228–239, https://doi.org/10.1016/j.buildenv.2019.01.043.

[12] J. Park, X. Yang, Y.K. Cho, J. Seo, Improving dynamic proximity sensing and processing for smart work-zone safety, Autom. ConStruct. 84 (2017) 111–120, https://doi.org/10.1016/j.autcon.2017.08.025.

[13] H. Chen, S.H. Cha, T.W. Kim, A framework for group activity detection and recognition using smartphone sensors and beacons, Build. Environ. 158 (2019) 205–216, https://doi.org/10.1016/j.buildenv.2019.05.016.

[14] R. Akhavian, A.H. Behzadan, Smartphone-based construction workers' activity recognition and classification, Autom. ConStruct. 71 (2016) 198–209, https://doi.org/10.1016/j.autcon.2016.08.015.

[15] S.H. Cha, J. Seo, S.H. Baek, C. Koo, Towards a well-planned, activity-based work environment: automated recognition of office activities using accelerometers, Build. Environ. 144 (2018) 86–93, https://doi.org/10.1016/j.buildenv.2018.07.051.

[16] S. Han, S. Lee, F. Peña-Mora, Vision-based detection of unsafe actions of a construction worker: case study of ladder climbing, J. Comput. Civ. Eng. 27 (2013) 635–644, https://doi.org/10.1061/(asce)cp.1943-5487.0000279.

[17] J.W. Dziedzic, Y. Da, V. Novakovic, Indoor occupant behaviour monitoring with the use of a depth registration camera, Build. Environ. 148 (2019) 44–54, https://doi.org/10.1016/j.buildenv.2018.10.032.

[18] R.H. Dodier, G.P. Henze, D.K. Tiller, X. Guo, Building occupancy detection through sensor belief networks, Energy Build. 38 (2006) 1033–1043, https://doi.org/10.1016/j.enbuild.2005.12.001.

[19] W. Wang, J. Chen, Y. Lu, H.-H. Wei, Energy conservation through flexible HVAC management in large spaces: an IPS-based demand-driven control (IDC) system, Autom. ConStruct. 83 (2017) 91–107, https://doi.org/10.1016/j.autcon.2017.08.021.

[20] N. Li, G. Calis, B. Becerik-Gerber, Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations, Autom. ConStruct. 24 (2012) 89–99, https://doi.org/10.1016/j.autcon.2012.02.013.

[21] H. Jiang, C. Cai, X. Ma, Y. Yang, J. Liu, Smart home based on WiFi sensing: a survey, IEEE Access 6 (2018) 13317–13325, https://doi.org/10.1109/access.2018.2812887.

[22] Y. Zeng, P.H. Pathak, P. Mohapatra, Analyzing shopper's behavior through WiFi Signals, in: Proceedings of the 2nd Workshop on Workshop on Physical Analytics - WPA '15, ACM Press, New York, NY, 2015, pp. 13–18.

[23] Y. Wang, K. Wu, L.M. Ni, WiFall: device-free fall detection by wireless networks, IEEE Trans. Mobile Comput. 16 (2016) 581–594, https://doi.org/10.1109/tmc.2016.2557792.

[24] J. Zhang, B. Wei, W. Hu, S.S. Kanhere, WiFi-ID: human identification using WiFi signal, in: 2016 International Conference on Distributed Computing in Sensor Systems (DCOSS), IEEE, Washington, DC, USA, 2016, pp. 75–82.

[25] W. Wang, A.X. Liu, M. Shahzad, K. Ling, 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-MobiCom '15, ACM Press, New York, 2015, pp. 65–76.

[26] W. Wang, A.X. Liu, M. Shahzad, Gait recognition using wifi signals, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16, ACM Press, New York, NY, USA, 2016, pp. 363–373.

[27] K. Ali, A.X. Liu, W. Wang, M. Shahzad, Recognizing keystrokes using WiFi devices, IEEE J. Sel. Area. Commun. 35 (2017) 1175–1190, https://doi.org/10.1109/jsac.2017.2680998.

[28] K. Qian, C. Wu, Z. Yang, Y. Liu, K. Jamieson, Widar: decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi, in: Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing-Mobihoc '17, ACM Press, New York, 2017, p. 6.

[29] Q. Zhou, C. Wu, J. Xing, S. Zhao, Q. Yang, Enabling noninvasive physical assault monitoring in smart school with commercial Wi-Fi devices, Wireless Commun. Mobile Comput. (2019) 1–14, https://doi.org/10.1155/2019/8186573, 2019.

[30] S. Palipana, D. Rojas, P. Agrawal, D. Pesch, FallDeFi: ubiquitous fall detection using commodity Wi-Fi devices, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 1 (2018) 1–25, https://doi.org/10.1145/3161183.

[31] H. Li, W. Yang, J. Wang, Y. Xu, L. Huang, WiFinger: talk to your smart devices with finger-grained gesture, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16, ACM Press, Heidelberg, Germany, 2016, pp. 250–261.

[32] S. Sigg, S. Shi, F. Buesching, Y. Ji, L. Wolf, Leveraging RF-channel fluctuation for activity recognition: active and passive systems, continuous and RSSI-based signal features, in: Proceedings of International Conference on Advances in Mobile Computing & Multimedia-MoMM '13, ACM Press, Vienna, Austria, 2013, pp. 43–52.

[33] Q. Pu, S. Gupta, S. Gollakota, S. Patel, Whole-home gesture recognition using wireless signals, in: Proceedings of the 19th Annual International Conference on Mobile Computing & Networking - MobiCom '13, ACM Press, Miami, USA, 2013, pp. 27–38.

[34] X. Wang, L. Gao, S. Mao, S. Pandey, CSI-based fingerprinting for indoor localization: a deep learning approach, IEEE Trans. Veh. Technol. 66 (2016) 763–776, https://doi.org/10.1109/tvt.2016.2545523.

[35] X. Wang, X. Wang, S. Mao, RF sensing in the internet of things: a general deep learning framework, IEEE Commun. Mag. 56 (2018) 62–67, https://doi.org/10.1109/mcom.2018.1701277.

[36] J. Wang, L. Zhang, Q. Gao, M. Pan, H. Wang, Device-free wireless sensing in complex scenarios using spatial structural information, IEEE Trans. Wireless Commun. 17 (2018) 2432–2442, https://doi.org/10.1109/twc.2018.2796086.

[37] W. Jiang, D. Koutsonikolas, W. Xu, L. Su, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, Towards environment independent device free human activity recognition, in: Proceedings of the 24th Annual International Conference on Mobile Computing and Networking-MobiCom '18, ACM Press, New York, NY, USA, 2018, pp. 289–304.

[38] J. Zhang, Z. Tang, M. Li, D. Fang, P. Nurmi, Z. Wang, CrossSense: towards cross-site and large-scale WiFi sensing, in: Proceedings of the 24th Annual International Conference on Mobile Computing and Networking-MobiCom '18, ACM Press, New York, NY, 2018, pp. 305–320.

[39] D. Halperin, W. Hu, A. Sheth, D. Wetherall, Tool release: gathering 802.11 n traces with channel state information, ACM SIGCOMM Comput, Commun. Rev. 41 (2011) 53, https://doi.org/10.1145/1925861.1925870.

[40] J. Gjengset, J. Xiong, G. McPhillips, K. Jamieson, Phaser: enabling phased array signal processing on commodity WiFi access points, in: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking-MobiCom'14, ACM Press, New York, NY, 2014, pp. 153–164.

[41] S. Tan, J. Yang, WiFinger: leveraging commodity WiFi for fine-grained finger gesture recognition, in: Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing - MobiHoc '16, ACM Press, New York, NY, USA, 2016, pp. 201–210.

[42] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, S. Li, RT-Fall: a real-time and contactless fall detection system with commodity WiFi devices, IEEE Trans. Mobile Comput. 16 (2017) 511–526, https://doi.org/10.1109/tmc.2016.2557795.

[43] N. Vaswani, T. Bouwmans, S. Javed, P. Narayanamurthy, Robust subspace learning: robust PCA, robust subspace tracking, and robust subspace recovery, IEEE Signal Process. Mag. 35 (2018) 32–55, https://doi.org/10.1109/msp.2018.2826566.

[44] H. Xu, C. Caramanis, S. Sanghavi, Robust PCA via outlier pursuit, Adv. Neural Inf. Process. Syst. (2010) 2496–2504.

[45] X. Wu, Z. Chu, P. Yang, C. Xiang, X. Zheng, W. Huang, TW-see: human activity recognition through the wall with commodity Wi-Fi devices, IEEE Trans. Veh. Technol. 68 (2018) 306–319, https://doi.org/10.1109/tvt.2018.2878754.

[46] Q. Zhao, D. Meng, Z. Xu, W. Zuo, L. Zhang, Robust principal component analysis with complex noise, in: International Conference on Machine Learning, IEEE, Beijing, China, 2014, pp. 55–63.

[47] X. Ye, J. Yang, X. Sun, K. Li, C. Hou, Y. Wang, Foreground–background separation from video clips via motion-assisted matrix restoration, IEEE Trans. Circ. Syst. Video Technol. 25 (2015) 1721–1734, https://doi.org/10.1109/tcsvt.2015.2392491.

[48] T. Bouwmans, Background subtraction website. https://sites.google.com/site/backgroundsubtraction/available-implementation/recent-background-modeling/background-modeling-via-rpca.

[49] T. Zhou, D. Tao, GoDec: randomized low-rank & sparse matrix decomposition in noisy case, in: Proceedings of the 28th International Conference on Machine Learning (ICML-2011), Omnipress, Madison, WI, USA, 2011, pp. 33–40.

[50] S.T. Roweis, EM algorithms for PCA and SPCA, in: M.I. Jordan, M.J. Kearns, S. A. Solla (Eds.), Advances in Neural Information Processing Systems, The MIT Press, Cambridge, 2007, pp. 626–632.

[51] J. Wang, Y. Zhao, X. Fan, Q. Gao, X. Ma, H. Wang, Device-free identification using intrinsic CSI features, IEEE Trans. Veh. Technol. 67 (2018) 8571–8581, https://doi.org/10.1109/tvt.2018.2853185.

[52] L. Zhang, M. Liu, L. Lu, L. Gong, Wi-Run: multi-runner step estimation using commodity Wi-Fi, in: 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), IEEE, Hong Kong, China, 2018, pp. 1–9, 2018.

[53] S. Mallat, A Wavelet Tour of Signal Processing, Elsevier, San Diego, 1999.

[54] R.J. Urbanowicz, M. Meeker, W. La Cava, R.S. Olson, J.H. Moore, Relief-based feature selection: introduction and review, J. Biomed. Inf. 85 (2018) 189–203, https://doi.org/10.1016/j.jbi.2018.07.014.

[55] S.D. Domenico, M.D. Sanctis, E. Cianca, F. Giuliano, G. Bianchi, Exploring training options for RF sensing using CSI-IEEE Commun, What Mag. 56 (2018) 116–123, https://doi.org/10.1109/mcom.2018.1700145.

[56] J. Kim, S. Chi, Action recognition of earthmoving excavators based on sequential pattern analysis of visual features and operation cycles, Autom. ConStruct. 104 (2019) 255–264, https://doi.org/10.1016/j.autcon.2019.03.025.