

PosMonitor: Fine-Grained Sleep Posture Recognition With mmWave Radar

Xiulong Liu^{ID}, Member, IEEE, Wei Jiang, Sheng Chen^{ID}, Member, IEEE, Xin Xie^{ID}, Hankai Liu^{ID}, Qixuan Cai^{ID}, Graduate Student Member, IEEE, Xinyu Tong^{ID}, Member, IEEE, Tuo Shi^{ID}, and Wenyu Qu^{ID}

Abstract—Sleep posture recognition is practically important in various scenarios, such as sleep healthcare, bedridden patient care, and chronic disease diagnosis. With concerns of user privacy preserving, we prefer the wireless sensing methods to computer vision methods when dealing with sleep posture recognition. However, the existing wireless sensing methods suffer from at least one of the following major limitations: 1) difficult to deploy in practice; 2) few posture categories; 3) insufficient accuracy; and 4) poor generalization ability. In this article, we use commercial-off-the-shelf (COTS) millimeter-wave radar to implement a sleep posture recognition system called PosMonitor. When designing the PosMonitor system, we need to address the following challenging issues. First, we propose an angle purification method based on multiframe joint analysis to alleviate the sparsity and instability of the point cloud. Then, we endow the point cloud with respiratory features to enhance its representation of the sleep posture. Further, to make the system applicable to different users, we extract relative respiratory features by normalization to overcome individual differences. Extensive experimental results show that our PosMonitor system can achieve 98% accuracy on average in recognizing 6 typical sleep postures and has good reliability across different conditions.

Index Terms—Millimeter-wave (mmWave) radar, sleep posture recognition.

I. INTRODUCTION

A. Background and Motivation

MANY clinical studies have shown that monitoring sleep postures is essential in sleep healthcare and diagnosing some chronic diseases. For example, sleep posture is an important indicator of positional obstructive sleep apnea (P-OSA) diagnosis [1]; in hospitals and communities, caregivers need to ensure that patients change postures regularly to prevent prolonged immobilization from causing pressure ulcers [2], a chronic wound prone to recurrence; and for people with epilepsy, improper sleep postures can increase the risk of sudden death [3]. Existing sleep posture recognition systems can be generally divided into two categories: 1) wearable-based systems and 2) nonwearable-based systems.

Manuscript received 17 August 2023; revised 30 September 2023; accepted 17 October 2023. Date of publication 31 October 2023; date of current version 26 March 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62002259. (Corresponding authors: Sheng Chen; Xin Xie.)

The authors are with the College of Intelligence and Computing, Tianjin University, Tianjin 300350, China (e-mail: xiulong_liu@tju.edu.cn; jiang_wei@tju.edu.cn; chensheng@tju.edu.cn; xinxie@tju.edu.cn; hkliu@tju.edu.cn; qixuan@tju.edu.cn; xytong@tju.edu.cn; shituo@tju.edu.cn; wenyu.qu@tju.edu.cn).

Digital Object Identifier 10.1109/JIOT.2023.3328866

- 1) *Wearable-Based Systems*: Traditional sleep monitoring systems, such as polysomnography (PSG), rely on specialized medical equipments and professionals, which are costly and difficult to widely promote. To make sleep posture monitoring services more accessible, many works [4], [5] have proposed to use lightweight and inexpensive wearable devices (e.g., inertial measurement units) to enable posture recognition. However, these approaches require the user to wear some devices on different body parts, which is inconvenient and may cause discomfort.
- 2) *Nonwearable-Based Systems*: Computer vision methods [6], [7] are seriously affected by visual factors, such as quilt occlusion and dark environments, and also begets the privacy leakage risk. As a result, radio frequency (RF) signal-based approaches [8], [9], [10], [11] have gained significant attention owing to their attractive proprieties, including nonintrusive, robust to light conditions, no requirement on line-of-sight, and privacy preserving. RFID-based systems [9], [12] attach large-scale RFID tag arrays under the bed sheets, thus can be affected by the wrinkles of the RFID sheets and may affect sleep comfort. WiFi-based system performance [10], [13] depends on the way WiFi devices and access points are deployed, which is difficult to guarantee in real-world scenarios. In contrast to other RF technologies, radar-based human sensing has attracted tremendous attention for its high precision and ease of deployment.

Therefore, this article focuses on radar-based sleep posture recognition.

B. Limitations of Prior Art

Several existing works have explored using millimeter-wave (mmWave) radar to monitor sleep posture. However, they have the following limitations. BodyCompass [8] used frequency modulated continuous wave (FMCW) radio to recognize human sleep orientation. This work inferred the angle of the torso relative to the bed by analyzing the reflected signal. However, this method requires a long training data acquisition time to ensure better recognition accuracy. Argosleep [11] used spatial information acquired by two mmWave devices to reconstruct the human skeleton. Argosleep can determine the state the body is in, including the tumbling state and the resting state, and predicts information about the body's



Fig. 1. Deployment scenario of the PosMonitor system.

skeleton in the resting state. Argosleep used the distance error between the predicted human skeleton and the person's real joint position as a judgment criterion and does not include common symmetrical postures, such as supine and prone. Zhou et al. [14] extracted multidimensional features from echo signals of mmWave radar to recognize postures. However, the sleep postures in this work refers to some dynamic sleep-related processes, such as getting up and turning over. It cannot recognition static sleep postures.

C. Our Approach

To overcome these limitations, this article aims to perform accurate and reliable sleep posture recognition with low-cost, easy-to-deploy commercially device. We propose PosMonitor, a mmWave radar-based sleep posture recognition system. As shown in Fig. 1, the proposed PosMonitor system uses radar above the bed to detect the target and return the signal sampling data, which is then processed by the backend server to generate the point cloud. Next, the PosMonitor system estimates the respiratory feature of each point in the point cloud. Finally, the PosMonitor system employs a deep neural network (DNN) with point cloud containing respiratory information as the input feature to recognize the sleep posture of the user. In addition, we propose a multiframe joint analysis method to improve the quality of the point cloud and eliminate individual differences by normalizing the respiratory features, making the PosMonitor system highly accurate and well migratory.

D. Challenges and Solutions

There are three major challenges that we need to address when implementing our PosMonitor system.

The first challenge is to obtain abundant and stable point cloud information. Due to the limited angular resolution of radar, the point cloud is sparse in space. Moreover, the number and location of points in the point cloud are disturbed by random noise, which increases the instability. These problems make the point cloud unavailable as a reliable feature to describe sleep posture. In this work, we propose a point cloud purification method based on multiframe angular occurrence

probability clustering analysis. At each distance, we cluster the estimated angles of multiple frames with a certain radius. Then, we determine whether each angle cluster is valid or not based on the probability and extract the centroid angles of the valid clusters. By utilizing the statistical properties of multiple frames, the method can suppress the drift of reflection points due to noise. Meanwhile, increasing the length of timing analysis can alleviate the problem of partial reflection point omission due to reflection characteristics.

The second challenge is to make the point cloud describe sleep postures better. The reflected signal received by the radar is a superposition of the human body and the surrounding objects. The reflection points of objects in the environment vary from scene to scene, which can reduce the reliability of the point cloud. Therefore, it is important to highlight the clusters of the human body in the point cloud. During sleep, the upper part of the human body is slightly displaced due to respiratory behavior, resulting in significant changes in the phase features of the reflected points. Moreover, we also observe differences in the phase changes of reflection points in different postures. We obtain the respiration features of the point cloud by extracting the phase change of each reflection point, which enhances the representation of the point cloud on the sleeping posture.

The third challenge is to overcome individual differences. We observe that the measured respiratory intensities of the human point cloud differed in different sleep postures. This phenomenon is due to the fact that the displacement caused by respiration is more significant in the direction normal to the body and more minor in the lateral direction, and this difference provides an opportunity for sleep posture recognition. However, the value of respiratory intensity cannot be directly used as a determinant of different sleep postures due to the different respiratory intensities of each individual. In this work, we normalize the respiratory intensity of each point accumulated from a single measurement to obtain the relative respiratory features. The normalization eliminates the individual differences in the mapping of respiratory intensity to sleep posture and makes the PosMonitor system perform well for different users.

E. Contributions

Our main contributions can be summarized as follows.

- 1) We propose a noncontact sleep posture recognition system called PosMonitor, which utilizes a single commercial mmWave radar to accurately identify six common sleep postures. PosMonitor employs a novel recognition framework that combines point cloud and respiratory features, introducing respiratory characteristics to traditional spatial point cloud for the first time.
- 2) We propose a new angle refinement method based on multiframe probability density to obtain reliable point cloud estimates. We investigate the correlation between respiratory amplitude features and sleep postures and introduce a method that uses relative respiratory features in combination with spatial point cloud, allowing the system to perform well across different users.

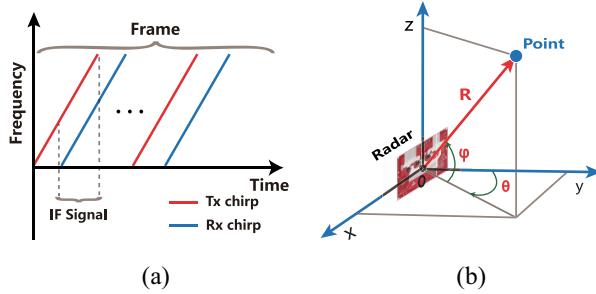


Fig. 2. Illustration of FMCW radar data dimensionality and a single point generation. (a) FMCW chirp signals. (b) Point in 3-D space.

- 3) We conduct a comprehensive experimental evaluation of the PosMonitor system. In basic scenario, the system achieves a classification accuracy of 98% for the six common sleep postures. In scenarios with blanket occlusions, different users, and varying environments, the average recognition accuracy remains above 90%, with only minimal additional data cost required to approach the performance of typical scenarios.

The remainder of this article is organized as follows. In Section II, we introduce some preliminary knowledge about mmWave radar. Section III presents the system model and the overall workflow. Section IV describes the system design methodology of PosMonitor in detail. We evaluate the performance of PosMonitor in Section V. Related works are reviewed in Section VI. Section VII concludes this article.

II. PRELIMINARIES

The mmWave radar usually refers to FMCW radar in the millimeter wave band, which is widely used in target detection. In this article, we use mmWave radar to extract point cloud and respiratory features to recognize sleep postures. The basic principles of measurement are outlined as follows.

A. Point Cloud Generation

FMCW radar is used to calculate the distance and angle of an object relative to the radar to generate point cloud. The radar first estimates the distance of the object, the basic principle of which is to transmit a linear FM wave signal called chirp, which will be reflected by the object and received by the radar. The received signal is a delayed version of the transmitted signal. By bypassing the transmit signal and the received signal through a mixer, we can obtain the intermediate frequency (IF) signal. The frequency of the IF signal (f_{IF}) is the frequency difference between the received signal and the transmitted signal, as shown in Fig. 2(a). The ratio of f_{IF} to the FM slope s indicates the signal propagation time, so we can get the distance of the object by $R = (c f_{IF} / 2s)$, where c is the speed of light. Different distances of reflection points correspond to different frequency components of the IF signal. Through fast Fourier transform (FFT) analysis of the frequency component of the IF signal, we can distinguish multiple reflectors at different distances. This processing step is called Range-FFT, and each result cell is called Range-bin. As an object moves, its displacement causes a change in

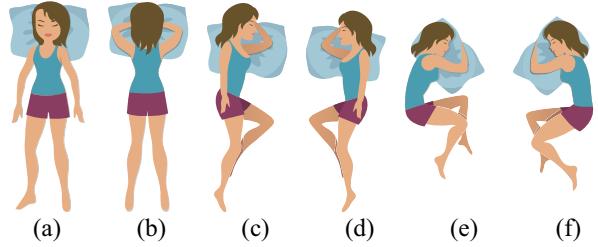


Fig. 3. Six common sleep postures [16]: (a) supine, (b) prone, (c) left log, (d) right log, (e) left foetus, and (f) right foetus.

the phase of the reflected chirp signal. By performing FFT analysis on multiple consecutive chirps, we can measure the velocity of the object. The FFT processing here is called Doppler-FFT. Further, to obtain the spatial position of the reflector, we need to estimate the signal direction of arrival (DoA). This opportunity comes from the multitransmitter multireceiver antenna array of the radar. The difference in distance from the reflector to the antenna array element leads to the phase difference in the received signal of each array element, which is determined by the DoA. By performing FFT on the received signal of the 2-D array, we can obtain the azimuth θ and the elevation φ . This processing step is called Angle-FFT. As shown in Fig. 2(b), the position of the point can be calculated by $x = R \cos(\varphi)\sin(\theta)$, $y = R \cos(\varphi)\cos(\theta)$, $z = R \sin(\varphi)$.

B. Respiration Estimation

Human respiration drives regular expansion and contraction of the abdominal and thoracic cavities. This displacement is usually between 1 and 12 mm [15]. Due to the short wavelength of the mmWave signal, this micro displacement can be captured using the phase information of the reflected signal. FMCW radar detection is usually performed in frames, with multiple chirps emitted. We sample the phase of human reflected signal at each frame and unwrap the obtained phase sequence. After band-pass filtering, the unwrapped signal according to the typical human respiratory frequency, the respiratory information, such as frequency and intensity, can be estimated using FFT and waveform analysis.

In this article, we use a Texas instruments (TIs) MMWCAS radar to capture the reflected signals from the target to further estimate the point cloud and respiratory information. We use six transmitting (Tx) antennas and eight receiving (Rx) antennas, which form a 2-D virtual antenna array with a size of 12.5×3 wavelength. The range, velocity, azimuth, and elevation resolution of the radar are about 4.22 cm, 0.79 m/s, 5° , and 18° , respectively.

III. SYSTEM OVERVIEW

In this article, we consider using mmWave radar to recognize accurately six common sleep postures: supine, prone, left log, right log, left foetus, and right foetus (shown in Fig. 3), which represent the most common cases in daily life [9], [16]. For brevity, we denote left log, right log, left foetus, and right foetus as log-l, log-r, foetus-l, and foetus-r, respectively. Fig. 4

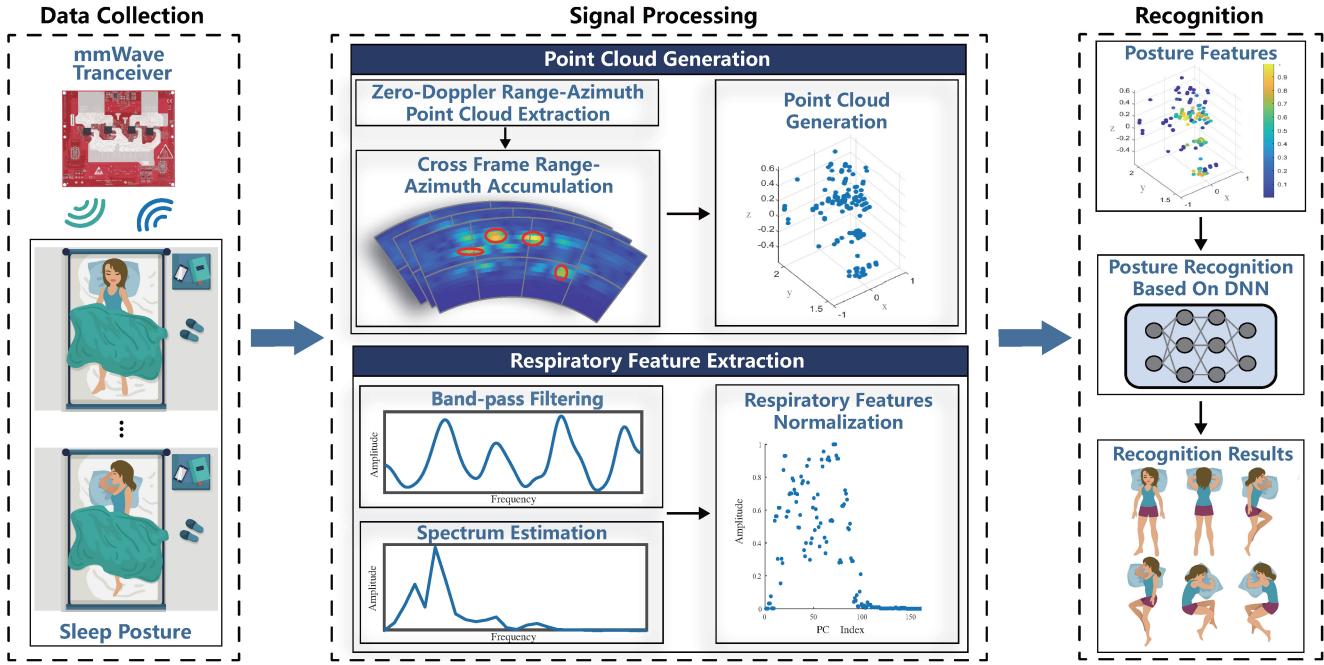


Fig. 4. Overview of the PosMonitor system.

shows the overview of our PosMonitor system, which uses mmWave radar to collect data, signal processing to extract features, and DNN to determine the sleep posture.

A. Data Collection

The mmWave radar transmits signals according to preset parameters, and the Rx antenna captures the reflected signals. The radar RF front-end amplifies, mixes, and filters the received signal, gets the digital signal by sampling, and sends it to the backend server.

B. Signal Processing

This component is designed to extract features related to sleep postures from the collected data, including point cloud generation and respiratory features extraction.

Point Cloud Generation: First, we perform Range-FFT and Doppler-FFT on the received signal of each frame to obtain the range-velocity data of the reflection points. Then, we estimate the azimuth corresponding to each Range-bin on zero-Doppler chirp. The same processing is done for all frames of the signal to obtain multiframe Range-Azimuth data. Next, for each Range-bin, a multiframe accumulative clustering of the azimuth is performed with a certain error range as the radius to obtain a reliable azimuth angle estimate. Finally, the corresponding elevations are estimated under the selected azimuths, and the Cartesian coordinates of the points are further calculated to generate the point cloud.

Respiratory Feature Extraction: We perform respiration feature extraction in the reflection point dimension, rather than the Range-bin dimension. This opportunity comes from the fact that 2-D-Angle-FFT is a spatial spectral decomposition of the reflected signal in the Range-bin dimension, and the decomposition preserves the phase correlation information of the reflected points. First, we obtain the phase profile of each

reflection point with a sampling interval of frames. Then, the respiratory signal is obtained by performing unwrapping and band-pass filtering on the phase profile. Next, we use FFT to analyze the spectrum of the respiratory signal and take its peak as the respiratory feature, representing the chest or abdomen displacement. Finally, to eliminate individual differences, we normalize the respiratory features of the point cloud generated from a single measurement to obtain the relative respiratory features.

C. Sleep Posture Recognition

The purpose of this component is to recognize the sleep posture of the users using the point cloud containing respiratory features obtained from signal processing. We designed a lightweight DNN to further extract the spatial distribution information of the point cloud containing respiratory features and to recognize sleep postures.

IV. METHODOLOGY

This section describes the technical details of the system, including the generation of stable point cloud, the extraction of respiratory features, and the structure of the DNN.

A. Stable Point Cloud Generation

The backend server receives the data returned from the radar and converts its format to $\{\text{ADC}, \text{Chirp}, \text{Horizontal}, \text{Vertical}, \text{Frame}\}$, where ADC is the sampling result of the IF signal, Horizontal represents the horizontal dimension of the 2-D virtual antenna array, and Vertical represents the vertical dimension. Using Range-FFT to estimate the distance information, and we have data of $\{\text{Range}, \text{Chirp}, \text{Horizontal}, \text{Vertical}, \text{Frame}\}$. Then, we need to extract the Range-bins of the human body. In sleep monitoring, the bed and the floor

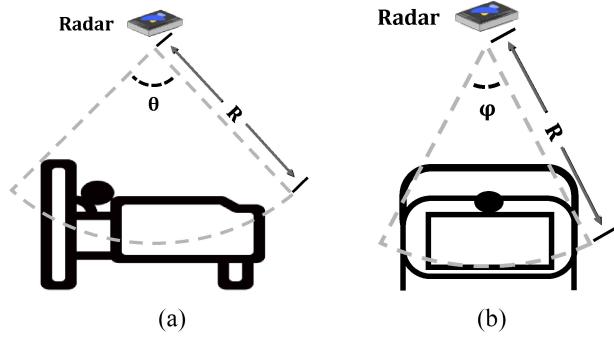


Fig. 5. Illustrating the effective sensing range of radar. (a) θ is the field of view of azimuth. (b) φ is the field of view of elevation.

are conspicuous reflectors. We do not use the CFAR algorithm to extract Range-bins for the human body. The principle of the CFAR algorithm is to adaptively estimate the threshold for target detection in a radar system based on the statistical characteristics of the background clutter. It is suitable for situations in which the background clutter reflections are weaker than the effective target, otherwise the effective target may be obscured. We first determine the approximate range R of the bed and then extract its corresponding Range-bins, as shown in Fig. 5. This coarse-grained extraction method may contain Range-bins that lack valid reflection information, such as beds and floors. The presence of these static reflective objects does not affect the functionality of the system since we use respiratory features to highlight the human point cloud, as will be proved later. Next, the velocity information is estimated using Doppler-FFT on the extracted Range-bins, we have data of {Range, Doppler, Horizontal, Vertical, Frame}. The human body is in a relatively stationary state during sleep, with almost only the upper body moving in periodic undulations with respiration. The speed of this small movement is very low, less than the speed interval indicated by a Doppler-bin, which is contained within the zero-Doppler-bin. Specifically, based on our parameter settings for the radar, the zero-Doppler-bin covers the velocity range of $[-0.395, 0.395]$ m/s. This makes it impossible to use static clutter filtering to extract the human signal because it loses part of the human signal while filtering out static reflectors, such as the bed. On the other hand, other moving objects in the environment (e.g., walking people) can also impact the detection of sleepers. The speed of these moving objects is usually much higher than the respiratory speed, i.e., outside the zero-Doppler-bin. Moreover, we extract an effective sensing range, which prevents objects outside the bed from causing interference. Therefore, we extract the data from the zero-Doppler-bin to filter them out, and the data format changes to {Range, Horizontal, Vertical, Frame}. Then, we estimate the azimuth for each extracted Range-bin from the signal of the horizontal antenna array. The reflected signal of the target within Range-bin r arriving at the m th horizontal antenna can be expressed as

$$\left\{ y_r^{[m]} = A^{[m]} \times \exp \left[j \frac{4\pi f_c}{c} (d_r + \Delta d_m) \right] \right\}_{r \in R, m \in [1, M]} \quad (1)$$

Algorithm 1 Multiframe Azimuth Occurrence Probability Clustering

Input: Angle set of Range-bin r : θ_r
Cluster radius: γ
Number of extracted angles: K

Output: Purified angles: θ'_r

- 1: $i \leftarrow 1$
- 2: **while** θ_r is not empty **do**
- 3: Search the angles with the maximum frequency angle as the center and γ as the radius
- 4: $\theta_{Ci} \leftarrow$ Searched angles
- 5: Delete searched angles from θ_r
- 6: $i++$
- 7: $\{C'_1, C'_2, \dots, C'_K\} \leftarrow$ Top K categories with the highest number of angles
- 8: $\theta'_r \leftarrow \{\text{avg}(\theta_{C'_1}), \text{avg}(\theta_{C'_2}), \dots, \text{avg}(\theta_{C'_K})\}$
- 9: **return** θ'_r

where A is the amplitude, d_r is the distance from the reflector to the reference array element, Δd_m is the distance difference between the reflector to the m th array element and the reference array element, f_c is the starting frequency of chirp, and M is the number of horizontal array elements [17]. To balance the angle estimation accuracy and computational burden, we use Angle-FFT instead of the Capon beamforming method which has higher accuracy. Capon beamforming is based on the minimum mean square error criterion, so it requires a sufficiently large number of snapshots (i.e., chirp number) to guarantee reliable estimation performance. This implies a considerable amount of data and high-computational cost. In contrast, Angle-FFT requires only one snapshot of the received data from the antenna array to achieve acceptable angle estimation performance. Performing Angle-FFT in the horizontal dimension, the data format changes to {Range, Azimuth, Vertical, Frame}. The Range-Azimuth estimation results are denoted as

$$\left\{ y_{r,\theta} = A \times \exp \left[j \frac{4\pi f_c}{c} (d_r + \Delta d_\theta) \right] \right\}_{r \in R, \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]} \quad (2)$$

where $y_{r,\theta}$ represents the reflected signal of the target located at Range-bin r and Angle-bin θ , and Δd_θ is the wave path-difference over the θ . Then, we estimate the actual angle of incidence of the reflected signal located at r . For the angular spectrum $\rho = |y_r|$ of r , we perform a wave peaks searching and obtain n peaks $\rho' = [|y_{r,\theta_1}|, |y_{r,\theta_2}|, \dots, |y_{r,\theta_n}|]$. To reduce the effects of sidelobe and multipath, we use a threshold method based on the signal-to-noise ratio SNR. We first calculate the maximum value ρ'_{\max} of ρ' , then determine the threshold $TH = \rho'_{\max} \cdot 10^{-\text{SNR}}$ by combining the preset SNR, and take the angle satisfying $\rho' > TH$ as the azimuth at distance r . We set an effective azimuth extraction range to filter out part of the background static object interference, as shown in Fig. 5(a).

However, the angles estimated from a single snapshot are biased by random noise interference, which can cause instability of the point cloud feature. To suppress this interference,

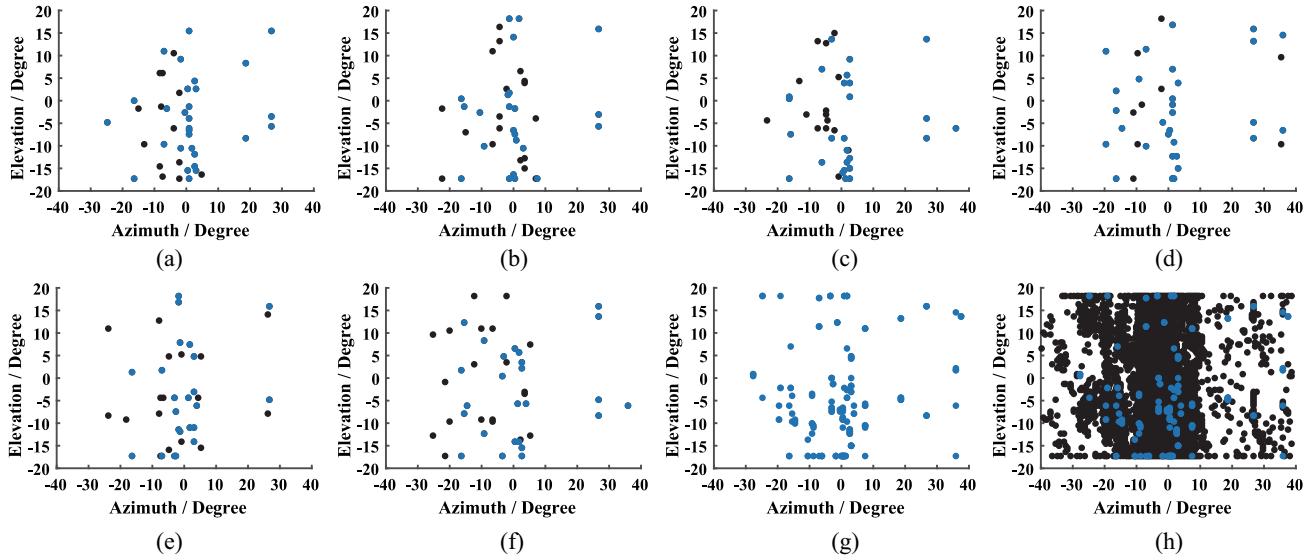


Fig. 6. Comparison of DoA estimation results using different methods. (a)–(f) To use only the reflected signal of one frame to estimate DoA, the result obtained has many noise points, and some effective reflection points are lost. (g) Result of a purification method using multiframe angular occurrence probability cluster analysis, and the estimated DoA is stable and reliable. (h) Simple accumulation of the DoA calculated for all frames, highlighting that the DoA estimated by common methods are unstable.

we design the multiframe azimuth occurrence probability clustering algorithm. Let the total number of frames be N_{frame} , and we do Angle-FFT for each frame of the snapshot to get the azimuth $\theta_r = [\theta_r^{[1]}, \theta_r^{[2]}, \dots, \theta_r^{[N_{\text{frame}}]}]$ of all frames. We first statistics the probability of occurrence of each angle in θ_r , and rank them in descending order. Then, we take the angle with the maximum probability as the cluster center, assign a new category C_1 to it, and traverse the remaining angles with a preset noise interference tolerance as the cluster radius γ . We consider the angle to belong to the same category when the difference $\Delta\theta$ between the traversed angle and the cluster center is less than the cluster radius and change its category to C_1 . After traversing all the remaining angles, the first round of clustering is completed, and the angles with category C_1 are removed from θ_r . Then, the angle with the highest probability in the current angle set becomes the new clustering center, is given the category C_2 , and a new round of clustering is performed. Repeat this process until the angle set is empty. At this point, we purify all the estimated azimuth into several categories. To keep the number of points contained in the point cloud input to the DNN the same, we rank the categories according to the number of angles they contain, and finally, we take the center angle of several (e.g., 3) categories that contain the most angles as the extracted result. If the number of categories is not enough, we copy the first-ranked category to supplement. The algorithm of azimuth purification is provided in Algorithm 1. Then, we do Angle-FFT in the vertical antenna dimension to get the elevation estimates for each Range-Azimuth point. Same as azimuth, the elevation is also extracted within the valid range, as shown in Fig. 5(b). And now we have data of {Range, Azimuth, Elevation, Frame}. Finally, the point coordinates are converted from the polar coordinate system to the Cartesian coordinate system. The same process is done on each Range-bin to generate the whole point cloud. Fig. 6 shows the DoA estimation results of the reflection

point for a user in the supine posture. The DoA estimation in a single frame is susceptible to noise interference. On the one hand, there are more noise points. On the other hand, part of the effective reflection points does not appear. As shown in Fig. 6(a)–(f), the angles are different for each frame. In Fig. 6, valid reflection points are indicated using blue color and randomly occurring noise points are indicated using black color. The impact of noise on DoA estimation is terrible, as shown in Fig. 6(h). When the DoA obtained from all the frames sampled at one time were accumulated together, the point cloud could no longer reflect the sleeping posture correctly, which tends to cause errors in sleep posture recognition, especially when distinguishing between similar postures. Moreover, the reflection points obtained from a single frame signal are relatively sparse, which is insufficient to describe the sleep posture adequately. In contrast, the DoA obtained by angle purification is relatively ideal. As shown in Fig. 6(g), effective reflection points are retained after angle purification, and the noise points with low probability are filtered out, making point cloud features more reliable in describing sleep postures.

B. Respiratory Features Extraction

We set the signal sequence of the i th point P_i as $y_i = [y_i^{[1]}, y_i^{[2]}, \dots, y_i^{[N_{\text{frame}}]}]$, after unwrapping, we can get the phase sequence $\phi_i = \text{unwrap}(\angle y_i)$. We obtain the respiratory phase profile ϕ'_i by band-pass filtering on ϕ_i with the typical human respiratory frequency (0.1–0.6 Hz) [18] as the passband. The peak-to-peak value $\Delta\phi'_i$ of ϕ'_i can reflect the small displacement ΔR caused by human respiration, in that $\Delta\phi'_i = (4\pi/\lambda)\Delta R$. Fig. 7 depicts a comparison between the respiratory point and the static point. Compared to the static point, the phases of the point that are affected by respiration have a significant undulating change. However, due to the

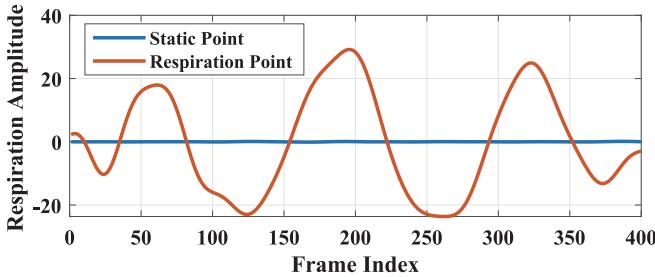


Fig. 7. Comparison of the phase profile: respiratory point *versus* static point.

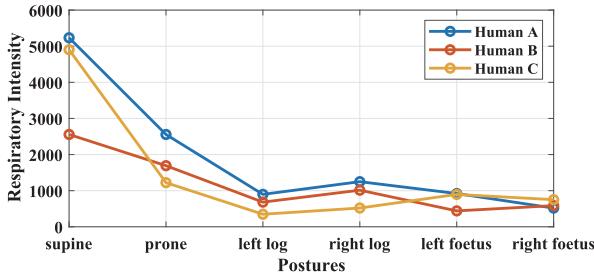


Fig. 8. Comparison of respiratory intensity in different postures.

instability of the respiratory pattern, the displacement caused by each respiration is different, that is, $\Delta\phi'_i$ is a fluctuating value. Therefore, we use FFT to analyze the spectrum of the respiratory signal and take the maximum of the spectral peak to represent the respiratory intensity and use it as the respiratory feature. That is, for each P_i , we assign the respiratory feature γ_i to it.

To verify the validity of the respiratory features, we measure them of different users in different sleep postures, and the results are shown in Fig. 8. The value of the vertical coordinate is the peak of the FFT result of the phase signal, which is a representation of the respiratory intensity. We observe that the respiratory features of the same user are somewhat differentiated across different sleep postures. This provides an opportunity to use respiratory features to recognize the postures. However, the different respiratory intensities of different people lead to two detrimental effects. On the one hand, the respiratory features measured for different people in the same sleep posture may differ significantly. On the other hand, one respiratory feature may be confused with other sleep postures. To overcome the individual differences, we normalized the respiratory features of a single measurement through dividing them by the maximum respiratory intensity of all points in the point cloud to obtain the relative respiratory features, i.e., $\gamma_i \leftarrow (\gamma_i / [\max(\gamma)])$. Fig. 9 compares the relative respiratory features of three different users in six postures. To avoid losing generality, we randomly select three different users, and the relative respiratory features of each posture are averaged over the data set. The distribution trends of relative respiratory features are similar for different people in the same sleep posture. At the same time, there is a clear distinction in different asymmetrical postures, as shown in Fig. 9(a)–(d). The range of body regions displaced due to respiratory movements is almost the same across symmetrical postures,

and the distribution trends of the relative respiratory features of the point clouds are also approximate, as shown in Fig. 9(b) and (e) and Fig. 9(c) and (f). This approximate tendency may make the system confused in recognizing symmetrical postures. However, there is a difference in the location of the point cloud distribution between different symmetrical postures, e.g., the point cloud generated by limbs of opposite orientations are located differently in space. Our PosMonitor system extracts information related to the spatial distribution of point clouds containing respiratory features to recognize different postures.

One should note that DC offset is a common problem in respiratory measurements [19], [20], causing distortions and inaccuracies in the phase, and this aberration can significantly affect the measurement of respiratory amplitude. However, in our work, we used the peak of the normalized respiratory frequency spectrum as the respiratory signature, which is less affected by DC offset, as corroborated by the respiratory signature legend in Fig. 9. Reliably estimating vital signs in different sleep postures is a challenging problem that would be explored in depth in our future work.

C. DNN Structure

In this article, we design a DNN to perform the task of sleep postures recognition, as shown in Fig. 10. The DNN mainly contains two convolutional layers and a transformer Encoder. The input of the DNN is a tensor of dimension $8 \times 170 \times 4$, where 8, 170, and 4 are batch size, point number, and feature number, respectively. The features are the x , y , and z coordinates of points in the point cloud and their relative respiratory features r , respectively. The DNN first performs 1-D convolution in each feature dimension to extract the spatial distribution features. The size of the convolution kernel is 170×1 with 1×1 strides. And we obtain a feature vector of size 20×4 . The feature vector is sent to the next layer after normalization and activation by the ReLu function. Then, the DNN performs a 1-D convolution in the spatial dimension of the feature vector to correlate the spatial distribution trends of each feature. The size of the convolution kernel is 4×1 with 1×1 strides. And we obtain an intermediate feature of size 20×32 . We use batch normalization followed by the ReLU functions after the layer. To better correlate the spatial distribution information among different features, we use a Transformer Encoder encoder, which mainly contains a multiheaded attention layer and a feed-forward neural network layer. Transformer Encoder has achieved good performance in many existing sensing works [21], [22], [23]. The multiheaded attention layer of the Transformer Encoder also includes a residual structure that extends the depth of the network and improves the convergence performance of the network. Next, two fully connected layers are used to complete the final recognition task. We first train the model using processed data of multiple individuals with different postures and then feed the currently collected data into the model after extracting features

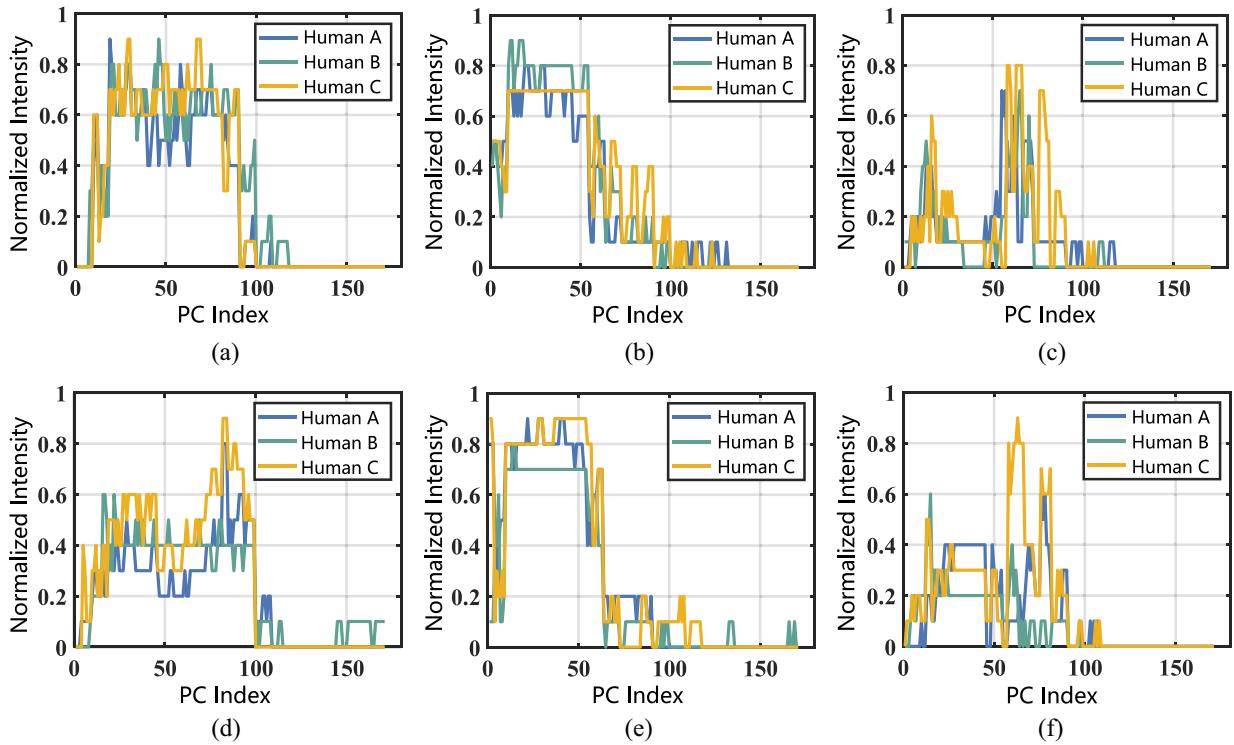


Fig. 9. Comparing the relative respiratory features of the point cloud in different postures. The distribution trends of relative respiratory features are similar for different people in the same sleep posture. At the same time, there is a clear distinction in different asymmetrical postures. (a) Supine. (b) Log-I. (c) Foetus-I. (d) Prone. (e) Log-r. (f) Foetus-r.

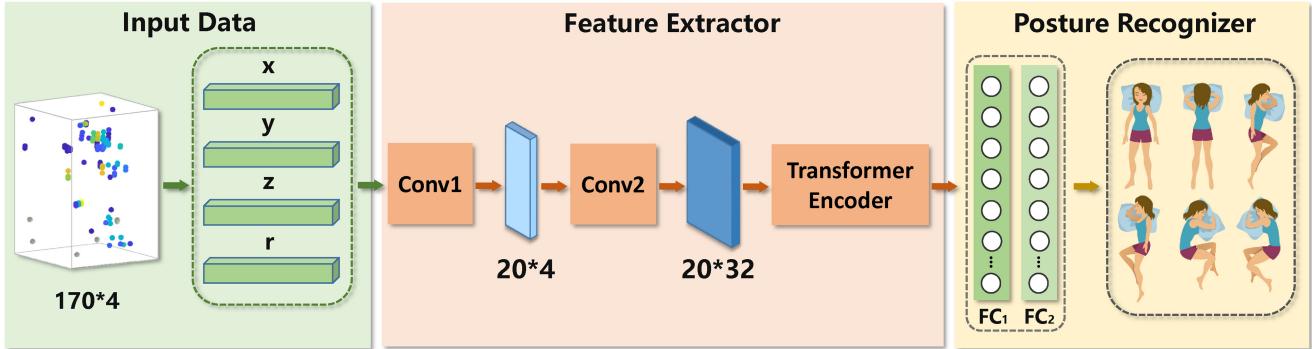


Fig. 10. Illustrating the DNN for sleep posture recognition.

through signal processing to obtain the results of sleep posture recognition.

V. EVALUATION

In this section, we first introduce the deployment of the system in detail. Then, we introduced the data set used in our experiment. Finally, we conducted extensive experiments to evaluate the performance of the PosMonitor system.

A. Implementation Details

All our experiments are implemented indoors, as shown in Fig. 11. We use a bed of $180 \text{ cm} \times 118 \text{ cm} \times 40 \text{ cm}$ and place the radar at a vertical distance of 1.72 m from the bed, which is in line with most scenarios. We adopt a

commercial set of mmWave radar, TI MMWCAS EVM [24], to evaluate the PosMonitor system. The radar works on the 77 GHz mmWave frequency band 77–81 GHz. The specific parameters of the radar detection are configured as follows: the ADC sampling frequency and number of samples are 7.2 MHz and 256, the ramp end time of the chirp is 40 μs , and the effective bandwidth is about 3.56 GHz. A single data acquisition contains 400 frames, each of the Tx antenna emits 8 chirps, with a frame period of 50 ms. The detailed parameter configuration of the device is shown in Table I. The raw data is collected via 1 Gigabit Ethernet offloading to the backend server (AMD R7-4800H CPU @ 2.90-GHz, 16-GB memory). Our signal processing algorithm is executed on MATLAB 2021b and the DNN is deployed on PyTorch framework.

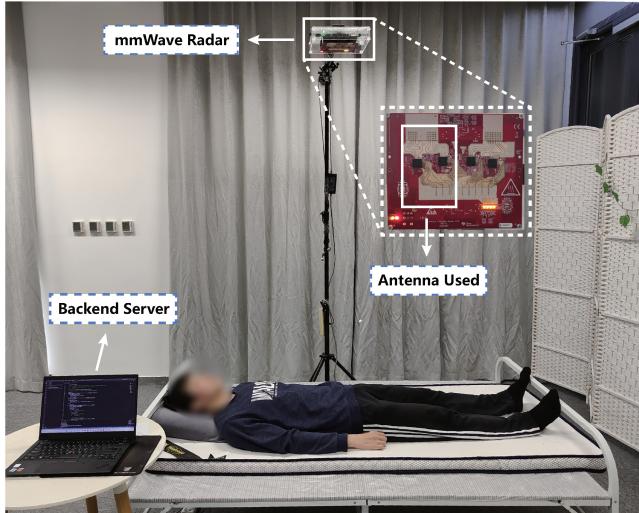


Fig. 11. Experiment setup.

TABLE I
PARAMETER SETTING OF MMWAVE RADAR

Parameters	Value	Parameters	Value
Start Frequency	77GHz	Range Resolution	4.22cm
Frequency Slope	100MHz/us	Chirps Per Frame	48
Idle Time	7us	Frame Interval	48ms
ADC Start Time	3us	Number of Frames	400
Sampling Rate	7.2Msps	Number of Tx	6
ADC Samples	256	Number of Rx	8
Ramp End Time	40us	Azimuth Resolution	5°
Bandwidth	3.56GHz	Elevation Resolution	18°

B. Sleep Posture Data Set

We collect data from 10 volunteers whose height ranges from 160 to 180 cm, and weight ranges from 50 kg to 80 kg. All follow-up experimental data are obtained from these ten volunteers. The volunteers are fully informed of the experimental contents. Each volunteer simulates 6 sleep postures in sequence for about one hour. For each posture, the radar acquires 20 s to obtain a set of data. Within a 20-s timeframe, the volunteers' trunk posture remains constant, while some volunteers may exhibit slight limb (arms and legs) movements or tremors during the 20-s data collection. Then, the features are extracted by signal processing on raw data, and finally, we have the point cloud with respiratory features. We collect data in three different scenarios, described as follows.

1) *Basic Scenario*: The radar sensing range set is only the user and the bed, without other static reflective objects. All ten volunteers participated in the data acquisition, and radar collected 12 sets of data for each individual in each posture. As a result, we obtained a total of 720 data on sleep postures.

2) *Quilt Covered Scenario*: To verify the performance of our system in a more realistic scenario, we collected data from users in different postures with the quilt covered. Six volunteers participated in the data acquisition, and radar collected ten sets of data for each individual in each posture. As a result, we obtained a total of 180 data on sleep postures.

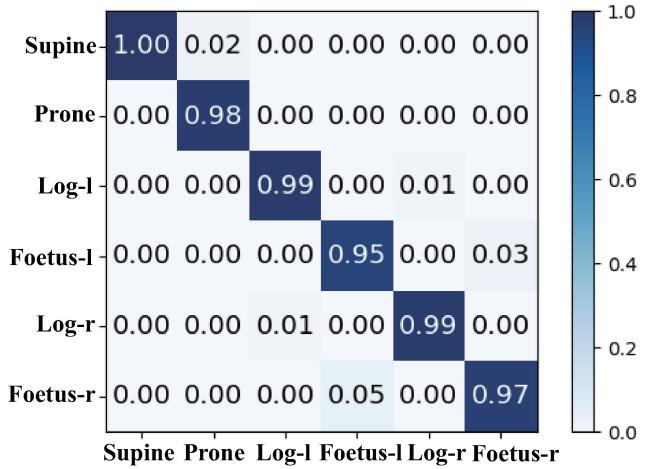


Fig. 12. Confusion matrix of recognizing the six sleep postures.

- 3) *Static Reflector Interference Scenario*: In the actual scenario, some static reflectors may be placed around the bed, such as bedside tables, arm chairs, etc. We set three possible placement scenarios, as shown in Fig. 16. Four volunteers participated in the data acquisition, and radar collected 10 sets of data for each individual in each posture. As a result, we obtained a total of 240 data on sleep postures.

C. Basic Scenario Performance

In this section, we use the data collected in the basic scenario to verify the performance of the system. We randomly divide all the collected data into training and test sets according to 8:2 and use the DNN for model training and posture recognition. Fig. 12 reveals the confusion matrix of 6 sleep postures recognition. The PosMonitor system achieves an average accuracy of 98%. The result reveals that the confusion samples only occur in the grid of symmetric postures, which is mainly due to the fact that the respiratory features of symmetric postures are similar, and the recognition depends more on the point cloud feature. However, it also indicates that our PosMonitor system can achieve higher performance in recognizing asymmetric postures. We compare the performance of PosMonitor system with existing representative RF-based works, including RFID-based [9], WiFi-based [10], and radar-based [8], and the results are shown in Table II. Due to the data and methods used being significantly different, we directly quote the evaluation results claimed in their papers. WiFi-based and radar-based cannot distinguish the left and right orientation of the human body, and we use the left orientation as its lateral sleeping direction. Compared to RFID-based and WiFi-based, we achieve higher recognition accuracy with lower device deployment difficulty. The radar-based performs well, but our system can recognize finer grained postures thanks to the refined design of the signal processing flow, which allows the point cloud to more adequately reflect the sleep posture. In addition, Argosleep [11] is generating skeleton information for each sleep posture and determining the distance deviation of the skeleton information from the

TABLE II
COMPARISON THE PERFORMANCE OF POSMONTIOR
WITH REPRESENTATIVE WORKS

Work	Technology	Posture types	Avg. Acc.
[10] 2018	Wi-Fi	4	98%
[12] 2018	RFID	4	86.24%
TagSheet [9] 2019	RFID	6	96.70%
BodyCompass [8] 2020	mmWave	NA	94%
Our PosMonitor	mmWave	6	98%

TABLE III
COMPARISON OF COMPUTATIONAL COMPLEXITY AND RECOGNITION
ACCURACY OF POINT CLOUD GENERATION USING CAPON
ALGORITHM AND ANGLE-FFT

Algorithm	400-Frame Avg. Time	Avg. Acc.
Angle-FFT	6.64s	98.00%
Capon	238.56s	93.67%
Capon Purification	242.46s	97.52%

real human joint positions, and does not give the accuracy of sleep posture recognition.

D. Validation of the Method

In this section, we validate the effectiveness of the proposed method on the data set under the basic scenario, including the feature extraction method and DNN method.

1) *Effectiveness of Feature Extraction Methods:* In this section, we verify the effectiveness of our proposed feature extraction method.

To validate the effectiveness of our point cloud generation scheme, we conduct a comparison between the computation time and recognition accuracy of the Capon beamforming algorithm, known for its high resolution and angle-FFT capabilities. The Capon algorithm requires a higher number of snapshots to achieve better classification accuracy, which corresponds to the number of chirps received in each antenna of the mmWave radar. In our capture setup, each frame has eight snapshots, which may cause Capon's DoA estimation results containing some information of false reflection points. We use the designed angle purification algorithm with multiframe Azimuth occurrence probability to optimize the DoA and further estimate the point cloud and respiration features. The extracted features are then fed into the DNN to calculate the recognition accuracy and the experimental results are shown in Table III. We only replace the algorithm for angle estimation in our feature extraction process, so we only count the time required for angle computation in 400 frames of data. It can be seen that even though the computation is performed on a PC, Capon still takes 238.56 s to compute the DoA for 400 frames of data, while Angle-FFT takes only 6.64 s. In terms of classification accuracy, using only the Capon algorithm has the lowest recognition accuracy of 96.52%. When the purification algorithm is used, the Capon algorithm achieves a classification accuracy of 97.52%, which is very close to the results of Angle-FFT. Therefore, due to the low-computational complexity of Angle-FFT, it is more suitable to be used in sleep monitoring system with some real-time requirements.

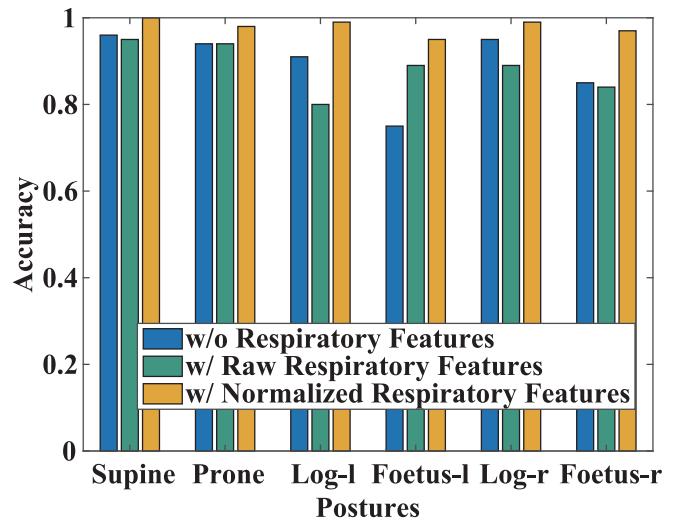


Fig. 13. Verify the effectiveness of feature extraction.

To verify the validity of the respiration feature, we set up three methods: using point cloud without respiratory features, point cloud with original respiratory features, and point cloud with normalized respiratory features. All three methods use the same data for feature extraction and adopt the same training settings, including parameters for data set division and model training. Fig. 13 shows the recognition results of the three methods in 6 postures, and the average accuracy is 89%, 88%, and 98%, respectively. Our final proposed solution, using the point cloud with normalized respiratory features, is far superior to the other two methods in terms of average accuracy and per posture recognition accuracy. Furthermore, the results of the point cloud without the respiratory features are slightly better than the point cloud with the original respiratory features because it does not overcome individual differences well, validating our previous conclusion.

2) *Effectiveness of DNN:* To verify the effectiveness of our designed DNN, we reimplement three common machine learning methods, SVMs with linear kernel, k-NNs ($K = 3$), and RFs. In addition, we reimplement mmGaitNet [25], which uses point cloud features containing intensity information for gait recognition. This work approximates our data format and has achieved good performance in gait recognition work. Fig. 14 reveals the overall performance of these classification methods on the data set and reveals that our designed DNN is significantly better than the other methods. Moreover, the overall accuracy of these machine learning classifiers is above 88%, with the recognition accuracy of SVMs and RFs exceeding 90%, which validates the effectiveness of our proposed method for feature extraction.

E. Impact of Quilt Occlusion

Due to the occlusion of the quilt, the signal collected by the radar is somewhat altered, which may affect the performance of PosMonitor system. We design two types of evaluation experiments: one in which the covered data are used for testing only, the other in which a small number of covered data for training, and the rest for testing. Fig. 15 reveals that the

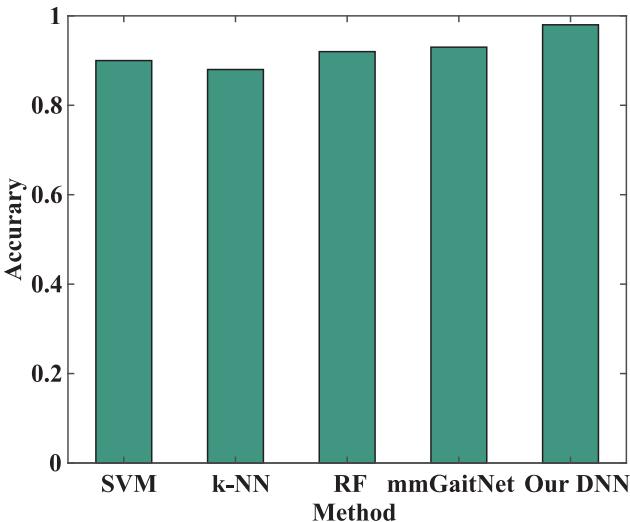


Fig. 14. Verify the effectiveness of DNN.

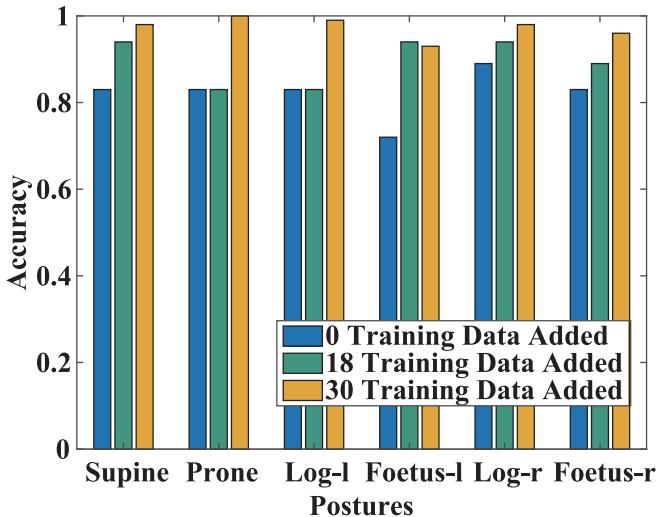


Fig. 15. Impact of quilt occlusion.

average accuracy of the system for direct posture recognition under quilt-covered conditions is 83%, which indicates that the extracted features change a bit when users cover the quilt. It is normal for the accuracy of the predictions to drop as the model has not seen data for this scenario. The overall performance of the system improves significantly when a small number of covered postures are added for training. The overall recognition accuracy improves to 90% when 18 (three per posture) training data are added per subject. Moreover, when 30 (five per posture) training data are added per subject, the accuracy achieves 97%. This demonstrates that our system can easily handle the quilt occlusion problem, and only a small number of new data is required to be collected to supplement the training to achieve satisfactory results.

F. Impact of Static Reflectors

We limit the sensing range of the radar in the feature extraction method to filter out the interference from other

TABLE IV
EVALUATION RESULTS FOR THREE DIFFERENT SCENARIOS WITH STATIC REFLECTORS AROUND THE BED

	Scenario 1	Scenario 2	Scenario 3	Average
Supine	95%	90%	95%	93.3%
Prone	95%	95%	90%	93.3%
Log-l	90%	85%	90%	88.3%
Foetus-l	75%	95%	85%	85%
Log-r	85%	85%	95%	88.3%
Foetus-r	100%	95%	90%	95%
Avg. Acc.	90%	90.8%	90.8%	90.6%

reflectors in the environment. However, we cannot accurately restrict the sensing area to include only the human body and the bed. Other objects around the bed will be included in our extracted point cloud features. These objects may also exacerbate multipath interference. To verify the performance of our system in the existence of interfering objects, we set up three evaluation scenarios, as shown in Fig. 16. We chose the round table and armchair as the reflectors, which is in line with the actual setting of the scene. Reflected objects are placed at 10 cm from the bed to ensure that they are within the sensing range of the radar. We invited four volunteers to participate in this evaluation, and the results of each scenario are averaged across them. The recognition results of each posture in the three scenarios are shown in Table IV. Confusion in classification occurs mainly among symmetric postures, mainly because distinguishing such features relies more on the spatial distribution features of the point cloud. The average accuracy of the system exceeds 90% in all three scenarios, which is acceptable.

G. User-Independent Performance

In this section, we evaluate the performance of PosMonitor system for the user-independent case. Considering the actual deployment needs, this case is essential. First, we use the data set of the basic scenario for this evaluation. Specifically, we use data from nine volunteers to train the model and another volunteer's data to test the model, which is often used in real-world development scenarios. We perform ten cross-validations, and the results are shown in Fig. 17. The overall accuracy of PosMonitor system is above 70%, and five of the volunteers have an accuracy of over 80%. Compared with the results in Fig. 12, there is a significant degradation in the recognition performance of PosMonitor in the user-independent case, which is a normal phenomenon. Enabling the system to work reliably across different user directly is challenging for the following reasons.

- 1) Different people have different heights, weights, and other form factors.
- 2) Different people have different limb details in the same given posture. For example, there are differences in the placement of the arms or legs, and the position of the body on the bed relative to the radar varies.

All of these factors can lead to variations in RF reflection patterns. Third, it is difficult for the model to handle this situation without a large amount of labeled data. However,

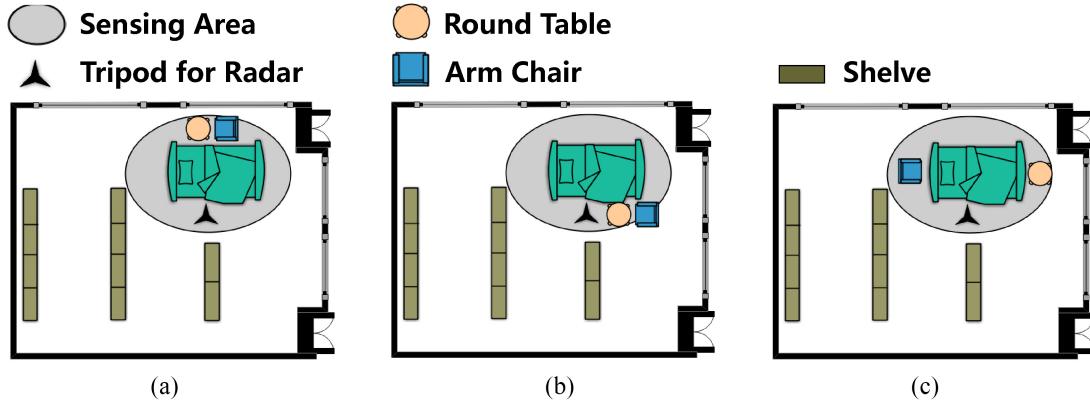


Fig. 16. Layouts of our three evaluation environments in static reflectors scenarios. (a) Scenario 1. (b) Scenario 2. (c) Scenario 3.

TABLE V
PERFORMANCE OF THE SYSTEM AFTER FINE-TUNING THE MODEL BY ADDING DIFFERENT AMOUNTS OF NEW USER DATA

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Average
Original Accuracy	75%	79.2%	81.7%	75%	75%	81.3%	72%	91.2%	70.8%	81.3%	78.3%
2 Data per Pos. (8-min)	86%	83.3%	84.3%	75%	86.1%	91.7%	83.3%	100%	85.7%	86%	86.1%
3 Data per Pos. (12-min)	93.3%	91%	91.2%	86.7%	93.3%	96.7%	93.3%	96.7%	86.1%	83.3%	91.2%
4 Data per Pos. (16-min)	96.7%	93.6%	83.3%	86.7%	93.3%	100%	96.7%	100%	92.4%	86.7%	92.9%
5 Data per Pos. (20-min)	96.7%	96.2%	92.2%	92.1%	96.7%	96.7%	96.7%	100%	95%	90%	95.2%

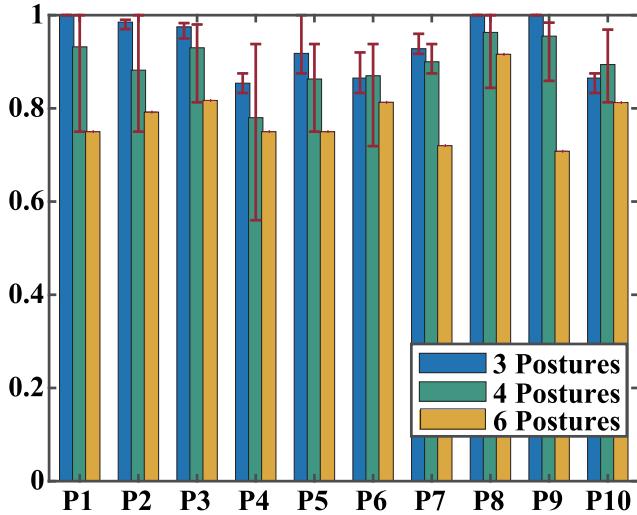


Fig. 17. Impact of user-independent case.

the extracted features can balance the effect of some individual differences. To prove this idea, we try to reduce the categories of recognized postures. Specifically, we examine the performance of the system in recognizing three postures and four postures in the person-independent case. We keep the two postures of supine and prone, and the rest of the posture permutations are tested separately for recognition accuracy. As shown in Fig. 17, the performance of the system improves significantly as the number of recognition categories is reduced. Confusion of postures mainly happens between symmetric postures, which are more dependent on point cloud

features. We believe our system will perform superior on the device with higher spatial resolution.

An acceptable solution for performance degradation on new users is to collect little data to correct the model. Since data collection is a very time-consuming and laborious task, we expect the system to achieve well with minimal use of data from new users. Specifically, we evaluate adding two to five new data items for each posture separately to the training data and using the original test data set. It took us about thirty seconds to collect data for one posture and about 4 min for a new user to collect data of the complete six postures. Table V shows the evaluation results. There is a significant improvement in the performance of the system as the target subject data increases. We only need to spend 8 min collecting two pieces of data for each posture to achieve an average accuracy of 86%. As a contrast, BodyCompass [8], which also uses a transfer learning strategy, achieves an average accuracy of 83.7% and 86.7% in 16 min and one night, respectively. The system approaches the initial performance after adding five sets of data, which took only 20 min, while BodyCompass takes a week. The results show that our system has good generalization ability and can quickly adapt to new users.

H. Prolonged Sleep Monitoring Performance

To validate the system's performance in a real-world environment, we collect data from volunteers while they sleep and calculate the error rate of sleep posture recognition. Specifically, we collect data from a volunteer during nap for six consecutive days, with nap durations ranging from approximately 45 to 80 min. The collected data includes six sleep postures, with 30 samples collected for each posture.

TABLE VI
POSTURE RECOGNITION ERROR RATE UNDER PROLONGED SLEEP

	Supine	Prone	Log-l	Log-r	Foetus-l	Foetus-r
Error Rate	0%	3.33%	13.33%	10%	16.67%	10%

After collecting each data sample, the system recognizes the posture, and the results are compared to the manually labeled ground truth. Table VI shows the evaluation results. The Supine posture is recognized correctly in all cases because it exhibits distinct respiratory feature compared to other postures. The Prone posture has only one instance recognized as Supine. Recognition errors are more likely to occur on symmetrical postures, with the worst performance being for the Foetus-l posture, occasionally being classified as Foetus-r or Log-l. The average recognition error rate is approximately 8.89%. With an increase in the number of model samples and test data, the recognition error rate is expected to further decrease.

VI. RELATED WORK

Existing works on sleep posture recognition can be categorized into wearable-based and nonwearable-based approaches. In this section, we will discuss them in each category.

Wearable-Based: SleepMontior [26] used the accelerometer in a smartwatch to detect the 3-D tilt angle and position of the wrist and therefore recognize four sleep postures. Fallmann et al. [4] implemented the recognition of eight sleep postures using a generalized matrix learning vector quantization classifier by binding accelerometers to the human ankle and chest. SleepGuard [5] used a smartwatch containing an IMU to monitor different sleep postures, with determination based on arm position. SleepGuard monitors four basic sleep postures and other sleep-related behaviors such as body rolls. However, all of these methods require the user to wear some devices that can affect sleep comfort.

Nonwearable-Based: Torres et al. [6] proposed MASH, which used three RGB-D cameras to monitor sleep posture and the stages of sleep posture transitions in a multiview. Liu and Ostadabbas [7] used a long infrared camera to implement thermal diffusion imaging, which can recognize the sleep posture of a fully covered person in complete darkness. However, such methods may lead to privacy leakage of users, therefore RF-based methods have attracted a lot of attention and research. Hu et al. [12] used an array of RFID tags embedded in a bed sheet to recognize sleep postures and detect sleep activities. Using low-pass filtering to filter the signal backscattered by the RFID array, and then using CNN to recognize the postures. In contrast, TagSheet [9] used an RFID tag array to image sleep posture, and the method requires no training model and is plug-and-play. Sleepy [13] modeled channel data from commercial WiFi devices as CSI with human factors and further analyzed it to achieve sleep posture recognition. Liu et al. [10] used a pair of WiFi devices and a wireless AP to detect the vital signs of the user and used machine learning to recognize the postures.

mmWave-Based: Zhou et al. [14] used mmWave radar to recognize eight sleep-related behaviors. The work did not

consider the static sleep posture of the user. BodyCompass [8] used the angle between the normal vectors of the bed surface and the human forehead surface to describe the sleep postures, but can only recognize the body orientation rather than the posture. Argosleep [11] reconstructed the human skeleton using two mmWave radars and calibrated the data using an RGB-D camera during training. However, it is difficult to use skeleton information to show the difference between symmetrical postures, such as supine and prone.

mmWave Human Sensing: The large bandwidth and short wavelength of mmWave radar make it uniquely suited for human sensing. Existing work has carried out extensive research, such as vital sign monitoring [27], [28], [29], [30], [31], imaging [32], [33], [34], [35], tracking [36], [37], [38], and gesture recognition [39], [40], [41], [42], [43]. In this article, our proposed point cloud method containing respiratory features has full applicability in human sensing, especially in static scenes such as sleep monitoring.

VII. CONCLUSION

In this article, we presented the PosMonitor system based on mmWave radar that can accurately recognize six common sleep postures. We addressed three main technical challenges to endow the PosMonitor system with excellent performance.

- 1) We proposed a multiframe joint analysis method to acquire stable point cloud.
- 2) We estimated respiratory features for each point to enhance the ability of the point cloud to represent the postures.
- 3) We extracted relative respiratory features to overcome individual differences for different users.

The experimental results reveal that the PosMonitor system achieves an overall recognition accuracy of 98% on sleep posture, which shows great potential in smart homes and intelligent medical applications.

REFERENCES

- [1] A. Menon and M. Kumar, "Influence of body position on severity of obstructive sleep apnea: A systematic review," *Int. Scholarly Res. Notices*, vol. 2013, Oct. 2013, Art. no. 670381.
- [2] M. Reddy, S. S. Gill, and P. A. Rochon, "Preventing pressure ulcers: A systematic review," *JAMA*, vol. 296, no. 8, pp. 974–984, 2006.
- [3] S. D. Lhatoo, L. Nashef, T. Tomson, and P. Ryvlin, "Association of prone position with sudden unexpected death in epilepsy," *Mortality Epilepsy Monit. Units Study*, vol. 85, no. 13, p. 1185, 2015.
- [4] S. Fallmann, R. van Veen, L. Chen, D. Walker, F. Chen, and C. Pan, "Wearable accelerometer based extended sleep position recognition," in *Proc. 19th IEEE Healthcom*, 2017, pp. 1–6.
- [5] L. Chang et al., "SleepGuard: Capturing rich sleep information using Smartwatch sensing data," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT)*, vol. 2, no. 3, pp. 1–34, 2018.
- [6] C. Torres, J. C. Fried, K. Rose, and B. S. Manjunath, "A multi-view multimodal system for monitoring patient sleep," *IEEE Trans. Multimedia*, vol. 20, no. 11, pp. 3057–3068, Nov. 2018.
- [7] S. Liu and S. Ostadabbas, "Seeing under the cover: A physics guided learning approach for in-bed pose estimation," in *Proc. MICCA*, 2019, pp. 236–245.
- [8] S. Yue, Y. Yang, H. Wang, H. Rahul, and D. Katabi, "BodyCompass: Monitoring sleep posture with wireless signals," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT)*, vol. 4, no. 2, pp. 1–25, 2020.
- [9] J. Liu, X. Chen, S. Chen, X. Liu, Y. Wang, and L. Chen, "TagSheet: Sleeping posture recognition with an unobtrusive passive tag matrix," in *Proc. IEEE INFOCOM*, 2019, pp. 874–882.

- [10] J. Liu, Y. Chen, Y. Wang, X. Chen, J. Cheng, and J. Yang, "Monitoring vital signs and postures during sleep using WiFi signals," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2071–2084, Jun. 2018.
- [11] A. Adhikari and S. Sur, "Argosleep: Monitoring sleep posture from commodity millimeter-wave devices," in *Proc. IEEE INFOCOM*, 2023, pp. 1–10.
- [12] X. Hu, K. Naya, P. Li, T. Miyazaki, K. Wang, and Y. Sun, "Non-invasive sleeping posture recognition and body movement detection based on RFID," in *Proc. IEEE IThings, GreenCom, CPSCom, SmartData*, 2018, pp. 1817–1820.
- [13] Y. Gu, Y. Zhang, J. Li, Y. Ji, X. An, and F. Ren, "Sleepy: Wireless channel data driven sleep monitoring via commodity WiFi devices," *IEEE Trans. Big Data*, vol. 6, no. 2, pp. 258–268, Jun. 2020.
- [14] T. Zhou, Z. Xia, X. Wang, and F. Xu, "Human sleep posture recognition based on millimeter-wave radar," in *Proc. IEEE SPSymo*, 2021, pp. 316–321.
- [15] T. F. Weber, R. Tetzlaff, F. Rengier, P. Geisbüsch, and A. Kopp-Schneider, "Respiratory displacement of the thoracic aorta: Physiological phenomenon with potential implications for thoracic endovascular repair," *Cardiovasc Intervent Radiol*, vol. 32, no. 4, pp. 658–665, 2009.
- [16] C. Idzikowski, *Beating Insomnia: How to Get a Good Night's Sleep*. Dublin, Ireland: Gill and Macmillan Ltd, 2003.
- [17] J. Guo, M. Jin, Y. He, W. Wang, and Y. Liu, "Dancing waltz with ghosts: Measuring sub-Mm-level 2D rotor orbit with a single mmWave radar," in *Proc. ACM/IEEE IPSN*, 2021, pp. 77–92.
- [18] X. Xu et al., "mmECG: Monitoring human cardiac cycle in driving environments leveraging Millimeter wave," in *Proc. IEEE INFOCOM*, 2022, pp. 90–99.
- [19] J. Liu, C. Gu, and J. Mao, "DC-independent high-linear motion sensing based on a novel ACAA algorithm," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 3, pp. 261–264, Mar. 2022.
- [20] F. Lin et al., "SleepSense: A noncontact and cost-effective sleep monitoring system," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 1, pp. 189–202, Feb. 2017.
- [21] Z. Cao, W. Ding, R. Chen, J. Zhang, X. Guo, and G. Wang, "A joint global-local network for human pose estimation with Millimeter wave radar," *IEEE Internet Things J.*, vol. 10, no. 1, pp. 434–446, 2023.
- [22] D. Hong et al., "SpectralFormer: Rethinking hyperspectral image classification with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Nov. 2022, Art no. 5518615.
- [23] Y. Gao, X. Liu, J. Li, Z. Fang, X. Jiang, and K. M. S. Huq, "LFTnet: Local feature transformer network for point clouds analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2158–2168, Feb. 2023.
- [24] "AWRx cascaded radar RF evaluation module," Texas Instrum., Dallas, TX, USA, Rep. SWRU553A, 2020.
- [25] Z. Meng et al., "Gait recognition for co-existing multiple people using millimeter wave sensing," *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 1, pp. 849–856, 2020.
- [26] X. Sun, L. Qiu, Y. Wu, Y. Tang, and G. Cao, "SleepMonitor: Monitoring respiratory rate and body position during sleep using Smartwatch," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT)*, vol. 1, no. 3, pp. 1–22, 2017.
- [27] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. R. Liu, "ViMo: Multiperson vital sign monitoring using commodity millimeter-wave radio," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1294–1307, Feb. 2021.
- [28] F. Wang, X. Zeng, C. Wu, B. Wang, and K. R. Liu, "mmHRV: Contactless heart rate variability monitoring using millimeter-wave radio," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16623–16636, Nov. 2021.
- [29] F. Wang, X. Zeng, C. Wu, B. Wang, and K. J. R. Liu, "Driver vital signs monitoring using Millimeter wave radio," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11283–11298, Jul. 2022.
- [30] Y. Wang, T. Gu, T. H. Luan, M. Lyu, and Y. Li, "HeartPrint: Exploring a heartbeat-based multiuser authentication with single mmWave radar," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 25324–25336, Dec. 2022.
- [31] H. Liu, X. Liu, X. Xie, X. Tong, T. Shi, and K. Li, "Application-oriented privacy filter for mmWave radar," *IEEE Commun. Mag.*, early access, Oct. 2, 2023, doi: [10.1109/MCOM.011.2200580](https://doi.org/10.1109/MCOM.011.2200580).
- [32] F. Zhang, C. Wu, B. Wang, and K. J. R. Liu, "mmEye: Super-resolution Millimeter wave imaging," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6995–7008, Apr. 2021.
- [33] H. Xue et al., "mmMesh: Towards 3D real-time dynamic human mesh construction using millimeter-wave," in *Proc. ACM MobiSys*, 2021, pp. 269–282.
- [34] H. Kong et al., "m3Track: MmWave-based multi-user 3D posture tracking," in *Proc. ACM MobiSys*, 2022, pp. 491–503.
- [35] J. W. Smith, M. E. Yanik, and M. Torlak, "Near-field MIMO-ISAR millimeter-wave imaging," in *Proc. IEEE RadarConf*, 2020, pp. 1–6.
- [36] J. Pegoraro, F. Meneghelli, and M. Rossi, "Multiperson continuous tracking and identification from mm-Wave micro-doppler signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 2994–3009, Apr. 2021.
- [37] A. Pearce, J. A. Zhang, and R. Xu, "Regional trajectory analysis through multi-person tracking with mmWave radar," in *Proc. IEEE RadarConf*, 2022, pp. 1–6.
- [38] P. Zhao et al., "Human tracking and identification through a millimeter wave radar," *Ad Hoc Netw.*, vol. 116, May 2021, Art. no. 102475.
- [39] J. Wang, C. Wang, D. Yin, Q. Gao, X. Liu, and M. Pan, "Cross-scenario device-free gesture recognition based on self-adaptive adversarial learning," *IEEE Internet Things J.*, vol. 9, no. 9, pp. 7080–7090, May 2022.
- [40] H. Liu et al., "M-Gesture: Person-independent real-time in-air gesture recognition using commodity Millimeter wave radar," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3397–3415, Mar. 2022.
- [41] S. D. Regani, C. Wu, B. Wang, M. Wu, and K. J. R. Liu, "mmWrite: Passive handwriting tracking using a single millimeter-wave radio," *IEEE Internet Things J.*, vol. 8, no. 17, pp. 13291–13305, Sep. 2021.
- [42] J. Wu, J. Wang, Q. Gao, M. Cheng, M. Pan, and H. Zhang, "Toward robust device-free gesture recognition based on intrinsic spectrogram of mmWave signals," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 19318–19329, Oct. 2022.
- [43] H. Liu et al., "mTransSee: Enabling environment-independent mmWave sensing based gesture recognition via transfer learning," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT)*, vol. 6, no. 1, pp. 1–28, 2022.



Xiulong Liu (Member, IEEE) received the B.E. and Ph.D. degrees from Dalian University of Technology, Dalian, China, in 2010 and 2016, respectively.

He is currently a Professor with the College of Intelligence and Computing, Tianjin University, Tianjin, China. He was also a Visiting Researcher with Aizu University, Aizuwakamatsu, Japan; a Postdoctoral Fellow with The Hong Kong Polytechnic University, Hong Kong; and a Postdoctoral Fellow with the School of Computing Science, Simon Fraser University, Burnaby, Canada. His research papers were published in many prestigious journals and conferences, including IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE TRANSACTIONS ON COMPUTERS, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, IEEE TRANSACTIONS ON COMMUNICATIONS, INFOCOM, and ICNP. His research interests include wireless sensing and communication, indoor localization, and networking.

Prof. Liu received the Best Paper Awards from ICA3PP 2014 and IEEE SYSTEM JOURNAL 2017. He is also the recipient of the CCF Outstanding Doctoral Dissertation Award 2017.



Wei Jiang received the B.E. degree in electrical information engineering from Northeastern University, Shenyang, China, in 2021. He is currently pursuing the master's degree with Tianjin University, Tianjin, China.

His current research interests include mmWave radar and wireless sensing.



Sheng Chen (Member, IEEE) received the bachelor's degree from Dalian Maritime University, Dalian, China, in 2015, the master's degree from Dalian University of Technology, Dalian, in 2017, and the Ph.D. degree from the College of Intelligence and Computing, Tianjin University, Tianjin, China, in 2023.

He is currently a Postdoctoral Researcher with the College of Intelligence and Computing, Tianjin University. His research interests include data center network, edge computing, wireless sensing, and indoor localization.



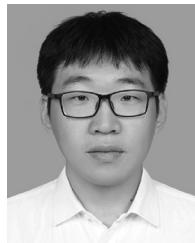
Xinyu Tong (Member, IEEE) received the B.E and Ph.D. degrees from the Department of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2015 and 2020, respectively.

He is currently a Postdoctoral Researcher with the College of Intelligence and Computing, Tianjin University, Tianjin, China. His research papers were published in many prestigious journals and conferences, including the IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE/ACM TRANSACTIONS ON NETWORKING, MobiCom, UbiComp, and INFOCOM. His research interests include wireless sensor network and wireless localization.



Xin Xie received the B.E. and Ph.D. degrees from the School of Computer Science and Technology, Dalian University of Technology, Dalian, China, in 2013 and 2019, respectively.

He is currently an Associate Professor with the College of Intelligence and Computing, Tianjin University, Tianjin, China. He was also a Visiting Scholar with the Department of Computer Sciences, Purdue University, West Lafayette, IN, USA; a Postdoctoral Fellow with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong; and a Researcher with the Theory Lab of 2012 Labs, Huawei, Hong Kong. His research interests includes IoT, edge intelligence, and network system.



Tuo Shi received the B.S., M.S., and Ph.D. degrees in computer science from Harbin Institute of Technology, Harbin, China, in 2015, 2016, and 2021, respectively.

He is currently a Research Associate with the College of Intelligence and Computing, Tianjin University, Tianjin, China. He has published papers in refereed journals and conferences, including IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE INTERNET OF THINGS JOURNAL, ACM TRANSACTIONS ON SENSOR NETWORK, IEEE International Conference on Computer Communications, and IEEE International Conference on Distributed Computing Systems. His research interests include sensor networks, battery-free networks, and mobile edge computing.

Dr. Shi is also the reviewer of distinguished journals, including IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, and *Journal of Computer Science and Technology*.



Hankai Liu received the B.E. and M.S. degrees from Tianjin University, Tianjin, China, in 2018 and 2021, respectively, where he is currently pursuing the Ph.D. degree.

His research interests include mmWave radar and wireless sensing security.



Wenyu Qu received the bachelor's and master's degrees from Dalian University of Technology, Dalian, China, in 1994 and 1997, respectively, and the Ph.D. degree from Japan Advanced Institute of Science and Technology, Nomi, Japan, in 2006.

She was a Professor with Dalian Maritime University, Dalian, from 2007 to 2015. She was an Assistant Professor with Dalian University of Technology from 1997 to 2003. She is currently a Professor with the College of Intelligence and Computing, Tianjin University, Tianjin, China. She has authored over 80 technical articles in international journals and conferences. Her research interests include cloud computing, computer networks, and information retrieval.

Prof. Qu is on the committee board for a couple of international conferences.



Qixuan Cai (Graduate Student Member, IEEE) received the B.E. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2021. He is currently pursuing the Ph.D. degree with Tianjin University, Tianjin, China.

His current research interests include embedded deep learning, multimodal learning, and ubiquitous computing.