LT-Fall: The Design and Implementation of a Life-threatening Fall Detection and Alarming System

DUO ZHANG, Peking University, China
XUSHENG ZHANG, Peking University, China
SHENGJIE LI, JD.com, China
YAXIONG XIE, University at Buffalo, United States
YANG LI, Peking University, China
XUANZHI WANG, Peking University, China
DAQING ZHANG*, Peking University, China and Institut Polytechnique de Paris, France

Falls are the leading cause of fatal injuries to elders in modern society, which has motivated researchers to propose various fall detection technologies. We observe that most of the existing fall detection solutions are diverging from the purpose of fall detection: timely alarming the family members, medical staff or first responders to save the life of the human with severe injury caused by fall. Instead, they focus on detecting the behavior of human falls, which does not necessarily mean a human is in real danger. The real critical situation is when a human cannot get up without assistance and is thus lying on the ground after the fall because of losing consciousness or becoming incapacitated due to severe injury. In this paper, we define a life-threatening fall as a behavior that involves a falling down followed by a long-lie of humans on the ground, and for the first time point out that a fall detection system should focus on detecting life-threatening falls instead of detecting any random falls. Accordingly, we design and implement LT-Fall, a mmWave-based life-threatening fall detection and alarming system. LT-Fall detects and reports both fall and fall-like behaviors in the first stage and then identifies life-threatening falls by continuously monitoring the human status after fall in the second stage. We propose a joint spatio-temporal localization technique to detect and locate the micro-motions of the human, which solves the challenge of mmWave's insufficient spatial resolution when the human is static, *i.e.*, lying on the ground. Extensive evaluation on 15 volunteers demonstrates that compared to the state-of-the-art work (92% precision and 94% recall), LT-Fall achieves zero false alarms as well as a precision of 100% and a recall of 98.8%.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**; • **Applied computing → Health care information systems**.

Additional Key Words and Phrases: Fall Detection, mmWave Radar, wireless sensing, localization

*This is the corresponding author.

Authors' addresses: Duo Zhang, zhangduo@stu.pku.edu.cn, Key Laboratory of High Confidence Software Technologies (Ministry of Education), school of Computer Science, Peking University, Beijing, China, 100871; Xusheng Zhang, zhangxusheng@stu.pku.edu.cn, Key Laboratory of High Confidence Software Technologies (Ministry of Education), school of Computer Science, Peking University, Beijing, China, 100871; Shengjie Li, lishengjie1@jd.com, JD.com, Beijing, China, 101111; Yaxiong Xie, University at Buffalo, Buffalo, New York, United States, yaxiongx@buffalo.edu; Yang Li, Key Laboratory of High Confidence Software Technologies (Ministry of Education), school of Computer Science, Peking University, Beijing, China, liyang@stu.pku.edu.cn; Xuanzhi Wang, Key Laboratory of High Confidence Software Technologies (Ministry of Education), school of Computer Science, Peking University, Beijing, China, xuanzhiwang@stu.pku.edu.cn; Daqing Zhang, Key Laboratory of High Confidence Software Technologies (Ministry of Education), school of Computer Science, Peking University, Beijing, China, Telecom SudParis and Institut Polytechnique de Paris, Paris, France, dqzhang@sei.pku.edu.cn.

2474-9567/2023/3-ART40 $15.00
https://doi.org/10.1145/3580835

ACM Reference Format:
Duo Zhang, Xusheng Zhang, Shengjie Li, Yaxiong Xie, Yang Li, Xuanzhi Wang, and Daqing Zhang. 2023. LT-Fall: The Design and Implementation of a Life-threatening Fall Detection and Alarming System. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 1, Article 40 (March 2023), 24 pages. https://doi.org/10.1145/3580835

1 INTRODUCTION

Falls are the leading cause of fatal and nonfatal injuries to elders in modern society[57]. Human falls happen frequently, especially for elder people, *e.g*, according to the Centers for Disease Control and Prevention, one out of three adults aged 65 and above falls every year[57]. Accurately detecting human falls of elderly people is becoming critically important for modern society, which has motivated researchers to propose various fall detection technologies [6–9, 24, 27, 29, 36, 38, 44, 47, 49, 52, 55, 57, 60]. Current fall detection technologies can be broadly categorized into two types: device-based methods and device-free methods. Device-based methods require the elderly to wear certain types of sensors twenty-four seven. The sensors only work when they are properly worn and powered. Charging the sensors regularly is extremely cumbersome, especially for the elderly at home. Among device-free technologies, computer vision-based technologies are usually accurate, however, they pose severe privacy concerns. [9, 52].

RF-based device-free fall detections have attracted huge attention in the past decade because of its privacy-preserving nature [7, 8, 24, 27, 29, 38, 44, 47, 49, 55, 57, 60]. RF-based fall detection leverages RF signals, such as Wi-Fi signals and FMCW radar signals, to continuously monitor human motion and detect human falls by analyzing the signal reflections from the human body. The detection algorithms of most RF-based solutions are built atop one key observation: the human fall involves a sudden body motion that causes large human body displacement within a short time, which results in a short but strong RF signal fluctuation. Accordingly, time-frequency analysis based [24, 44, 49, 57, 58, 60] and deep-learning based algorithms [12, 23, 29, 55] are proposed to detect such a signal fluctuation and thus identify any instance of human fall.

We observe that most of the current fall detection solutions are diverging from the **purpose of fall detection**: timely alarming the family members, medical staff, or first responders to save the life of humans with severe injuries. Instead, such solutions focus on detecting the behavior of falling down itself, which does not necessarily mean a human is in real danger. As long as the human can get up after the fall, the human is either not injured or able to seek for help on their own. An alarm in such a situation is not necessary and will waste precious community resources if the medical staff or first responders react after receiving the alarm.

The critical situation is when the human cannot get up without assistance and is thus lying on the ground after the fall due to loss of consciousness or becoming incapacitated due to severe injury, which results in delayed medical treatment if not timely discovered and thus significant increases of mortality risk in some clinical conditions [62]. Studies show that half of the elders who lay on the floor for more than one hour after a fall died in the following six months [10, 62]. Accordingly, we model a life-threatening fall as a two-stage behavior: a sudden height drop of the human body (*i.e.*, the falling) followed by the human lying down on the ground for a prolonged period of time (*i.e.*, a long-lie).

In this paper, we design LT-Fall , a mmWave-based fall detection system that generates alarms only when detecting a life-threatening fall. LT-Fall consists of two stages: a fall detection stage where we aim at detecting the existence of potential falls and a verification stage where we validate whether the potential falls are indeed life-threatening. In the first stage, we leverage the mmWave radar to extract the 4D point cloud (the coordinates in 3D space plus the velocity) of the human body [40, 45]. Based on the extracted 4D point cloud, we continuously monitor the human height (of the head or shoulders) and report a potential fall after detecting a sudden height drop.

In the second stage, we verify whether the reported fall behavior is indeed life-threatening by monitoring the human status after the fall: positive if the human lies on the ground for a long time. We would like to emphasize

that accurately locating a motionless human lying on the ground is challenging even for mmWave radar with a bandwidth of gigahertz. The COTS mmWave radar can easily extract the 4D point cloud of a moving object since it has sufficient velocity resolution to separate the signal reflecting off the object and its surrounding environment in the velocity domain. When the object becomes static, mmWave radar does not have enough spatial resolution to resolve signal reflections from targets in the spatial domain. We observe that although the human body no longer has limb movements with detectable velocity when lying on the ground, there are still micro-motions like breathing and heartbeat, which differ from static environments. We propose a joint spatio-temporal localization technique to detect and locate the micro-motions of the motionless human by appropriately combining the multiple radar frames over a long period of time to improve the Doppler shift resolution. We use the height of the micro-motions (chest) to roughly estimate the height of the human, and identify a fall as life-threatening if human height is close to the ground for a predefined time period.

Compared to the existing fall detection systems, LT-Fall significantly reduces the false positives. It is known that there exist a large number of fall-like activities in our daily life that may induce similar signal fluctuations as a real fall and thus confuse the detection system, such as picking up something from the floor, sitting on a chair or lying on the bed. Obviously, these fall-like behaviors are much more likely to occur than falls, easily leading to false alarms. LT-Fall filters out both fall-like behaviors and falls that are not life-threatening in the second stage by continuously monitoring the status of the human after the reported fall. LT-Fall is a purely model-based fall detection system without offline dataset collection and training, so it is capable of working across various environments and for different human subjects. The demo of our system is present at: https://www.youtube.com/watch?v=lUBGOeFOhME.

In summary, the main contributions of LT-Fall can be summarized as follows:

- To the best of our knowledge, for the first time, we point out that the fall detection system should focus on detecting life-threatening falls instead of detecting any random falls. Accordingly, we propose a life-threatening falls detection system that consists of two stages: the first stage, similar to existing solutions, detects both falls and fall-like behaviors; the second stage filters out falls that are not life-threatening and fall-like behaviors by continuously monitoring the human status after the fall.
- A joint spatio-temporal localization technique is proposed to effectively localize the micro-motion of the human body to determine his height after a fall. Apart from fall detection, such technology has broad prospects for other wireless sensing applications, such as motionless human presence detection, CPD (child presence detection in the car), etc.
- We implement LT-Fall on the COTS mmWave radar and recruit 15 volunteers to evaluate its performance. To fully verify the effectiveness of LT-Fall, extensive evaluations are conducted in seven different environments using three different sets of radar systems and four different sets of radar parameter settings. Compared to the state-of-the-art work (92% precision and 94% recall) [55], LT-Fall achieves zero false alarms as well as a precision of 100% and a recall of 98.8%.

2 RELATED WORK

Fall-related healthcare information systems can be divided into two categories: fall prediction systems and fall detection systems. Fall prediction systems aim to alert high-risk individuals before the occurrence of a fall, and increase their awareness of fall protection [20, 48]. In contrast, fall detection systems intend to alert the user and medical staff in a timely manner after the fall to expedite and improve the medical care provided. In this paper, we focus on the fall detection systems.

The fall detection systems can be classified into two broad categories: wearable and non-wearable technologies. Among wearable technologies, a range of sensors have been explored for fall detection, ranging from accelerometers [31, 39], gyroscopes [33], RFID [14] to smart phones [1]. These systems can only work on the premise that

all the sensors are properly worn and powered. However, elderly people usually feel encumbered by wearable devices. Apart from the inconvenience, aging seniors may also forget to wear or charge their devices[55]. The limitations mentioned above can be overcome by using non-wearable technologies. Computer vision approaches perform fall detection by analyzing images or video sequences containing human activity captured by cameras. However, it is subjected to various problems, such as serious privacy concerns and high demand for computational resources [22, 64]. Ambient senor-based approaches make use of ambient information caused by falls to detect the risky activity. The ambient information being used includes pressure [11, 13] and floor vibration [6]. However, other sources of pressure or vibration around the subject in the environment account for a large proportion of false alarms [43].

As wireless sensing technology has made great progress [2, 3, 34, 37, 56, 59, 61, 65], researchers have developed many fall detection systems based on wireless signals. These systems can be categorized into time-frequency analysis based systems and time-frequency-space feature based systems. The time-frequency analysis based systems mainly include methods on Wi-Fi [44, 57, 60], Doppler radar [7, 21, 24, 32, 49, 53], and ultrasound [38], which mainly leverages the Doppler shift of targets for fall detection. These methods work on the assumption that falls happen at a high speed or acceleration, resulting in high-frequency pulses in the wireless signal. Based on this assumption, time-frequency analysis methods such as short-time Fourier transform (STFT) and wavelet transform can be utilized to extract features of falls. After that, the extracted time-frequency features are fed into classifiers (e.g., SVM or Random Forest) to distinguish high-speed fall patterns from non-fall events. However, the elderly may not fall as fast as they assume. Individuals always instinctively grab nearby objects when falling, which reduces the speed of the fall. Furthermore, some fall-like behaviors such as sudden sitting or standing motions which occur at rapid speeds can easily lead to false alarms in such systems [7]. Apart from time-frequency analysis based systems, some recent works have taken spatial information (e.g., distance and angle) into account. In particular, FMCW radar (especially mmWave) [15, 16, 29, 30, 35, 55] and UWB radar [25, 41, 42] are the main sensing devices used in such methods due to their ability to measure distance and angle of objects. To achieve fall detection, they usually preprocess the radar signal to acquire range-velocity spectrum, range-angle spectrum, point clouds, dynamic range-doppler trajectory [15], etc. The preprocessing results are then fed into traditional classifiers (e,g, KNN) or deep neural networks (e.g., CNN) to learn features and achieve behavior classification. It is worth noting that these features used to detect falls are only extracted from the height drop process of the fall, which is similar to fall-like behaviors. Therefore, they are also susceptible to fall-like behaviors and cannot accurately identify life-threatening falls in real scenarios. To solve this problem, a case study has tried to harness the after-fall breath as an aid to detect falls[8]. However, like most works that utilize radar to detect breathing and heartbeats [4, 5, 66], it can only measure the distance of the person from the radar but provide no angle information. As a result, they have to mount the radar on the ceiling and can only detect a fall right under the radar, which greatly limits its application scenarios. Compared with methods based on machine learning, deep learning methods can handle a more diverse set of falls given sufficient training data. However, collecting large-scale fall training data is labor-intensive, and it seems impossible to collect a complete dataset covering all kinds of falls.

In contrast, LT-Fall is a model-based fall detection system that does not require any offline dataset collection or training. It adopts a two-stage algorithm to detect falls, first using 4D point cloud for fall and fall-like behavior detection. Then, by determining whether the person is lying on the floor, LT-Fall can effectively distinguish life-threatening falls from non life-threatening falls and fall-like behaviours.

3 SYSTEM OVERVIEW

LT-Fall is a two-stage fall detection system that builds upon the COTS mmWave radar, whose architecture is shown in Figure 1. During the data collection stage, the mmWave radar frontend sends the FMCW signal

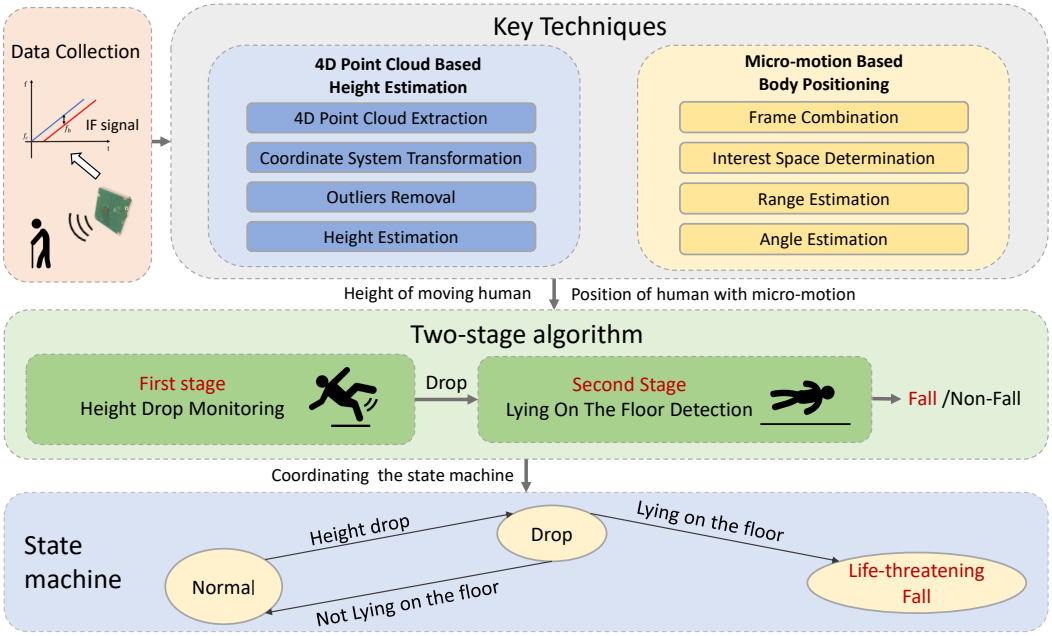


Fig. 1. An overview of LT-Fall. The data collection module captures the reflections of environments and sends the IF signal to the two key techniques. These two key techniques are used for the localization of dynamic and stationary human bodies, respectively. The two-stage algorithm is responsible for detecting falls based on people's spatial information provided by two key techniques. To further clarify the system, a state machine is designed to track a life-threatening fall event.

and receives reflected signals from the environment. LT-Fall processes the received IF signal to generate 4D point clouds of the human body, based on which LT-Fall estimates the height of the human and detects any micro-motions of the human body. We are able to locate the human if any micro-motions (e.g., breath) of the target are detected. LT-Fall also implements a two-stage fall detection algorithm. In the first stage of the algorithm, we detect human falls and fall-like behaviors by identifying the human height drop according to the 4D point cloud. In the second stage, we keep monitoring the status of the human target and identify a life-threatening fall if the height of the human body is low and close to the ground for a period of time, such as ten seconds.

LT-Fall implements a state machine to track the status of the human target and then accurately identify any life-threatening fall events. The state machine includes three states: a normal state, a drop state, and a long-lie fall state, where state transitions are triggered by the corresponding event detected by the two-stage fall detection algorithm. When a height drop is detected, the state is transitioned from the normal state to the drop state. Once in the drop state, the system will determine whether the person is lying on the floor. The state switches to the long-lie fall state if the person is lying on the floor for a pre-defined period of time. Otherwise, the state will transition back to the normal state.

4 4D POINT CLOUD BASED HEIGHT ESTIMATION

In this section, we present the algorithm to extract a body's height based on the 4D point cloud. We first introduce the preliminaries of mmWave radar, then describe the process to extract 4D point clouds from radar signals, and lastly introduce the height estimation algorithm.

4.1 Preliminaries of mmWave Radar

An FMCW mmWave radar is usually equipped with both transmitting antennas and receiving antennas. When the radio signals are transmitted into the environment and reflected back to receiving antennas, there is a frequency difference between the transmitting signal and receiving signal due to the time-of-flight. If we mix the receiving signal with the transmitting signal, we can get the intermediate frequency (chirp) signal whose frequency is proportional to the range between the radar and the target. In terms of a moving object, the chirp's phase varies with the range between the radar and the target. The phase difference between adjacent chirps is proportional to the radial velocity of the object. In other words, one can estimate the distance and radial velocity of a reflector by estimating the frequencies of sample points in a chirp and across chirps respectively.

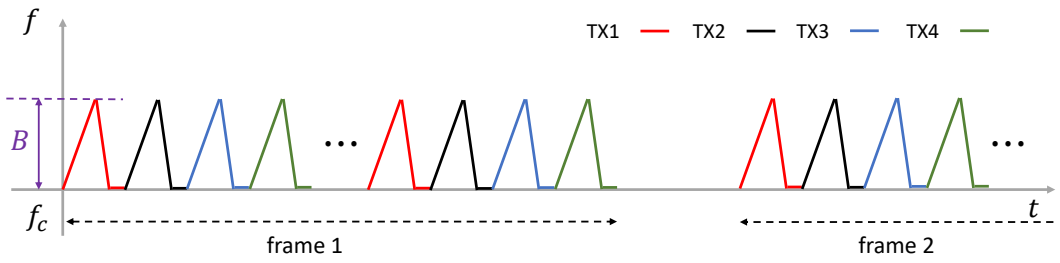


Fig. 2. A radar frame consists of several groups of chirps.

4.1.1 Range and Velocity Resolution. As shown in Figure 2, a typical mmWave radar performs range and radial velocity measurement by sending back-to-back chirps. Different transmitting antennas take turns to send chirps, forming a chirp group. Several chirp groups form a radar frame, which is the basic unit for estimating range and radial velocity. When a radar frame ends, there is an idle time until the next frame transmits. Given a radar with bandwith B, its range and radial velocity resolutions are determined as:

$$R_{res} = \frac{C}{2B} \tag{1}$$

$$V_{res} = \frac{\lambda}{2T_{frame}} \tag{2}$$

where C is the light speed ($3 \times 10^8 m/s$), λ is the wavelength for the starting frequency of the radar, and T_{frame} is the duration of a frame.

We see that, given the wavelength, the longer the duration of a frame, the finer the velocity resolution. However, the radar frame duration is restricted to the transmission efficiency (i.e., frames per second, FPS). Hence, there is a trade-off between better velocity resolution and higher FPS. For a FMCW mmWave radar operating at 60 GHz, its finest velocity resolution is about $v_{res} = 0.08m/s$ under the FPS limit, which can easily separate moving humans from the static environment. Thus, we can extract point clouds of people with moving speed greater than $0.08m/s$. We note that our empirical results show that such a velocity resolution cannot support detecting the millimeter-scale micro-motions of people, such as breathing and heartbeat.

4.1.2 Angle Resolution. Besides the range and radial velocity, FMCW mmWave radar also provides angle information, indicating the azimuth and elevation directions of the target in 3D space. Since targets in different directions will introduce different phase differences among the antennas, the direction of signal arrival can be

estimated by the phase differences. Similar to the resolution of range and velocity, the angle resolution is given by:

$$\theta_{res} = \frac{\lambda}{Md\cos(\theta)} \tag{3}$$

where M is the number of antennas, d is the distance between antennas in the antenna array, and θ is the angle of targets relative to the radar.

From equation 3, we can see the angle resolution is related to the angle of the target. The best angle resolution is obviously when the direction of the target is 0 degree relative to the radar. To improve the angular resolution, we can virtualize x transmitting antennas and y receiving antennas into $x \times y$ virtual antennas using the Time-Division-Multiplexing (TDM) Multiple-Input-Multiple-Output (MIMO) technique [50]. Since the signals sent by multiple transmitting antennas at the same time are not easily separated at reception, TDM-MIMO lets the transmitting antennas take turns sending signals in chronological order (as shown in Figure 2). Our radar is equipped with 4 transmitting and 4 receiving antennas (as shown in Figure 11). However, even with MIMO technology, the optimal angular resolution for azimuth and elevation of our radar is only 14.32° and 38.20°, which is far from enough to separate the human body from the environment.

4.2 4D Point Cloud Extraction of Moving Human Body

The 4D point cloud characterizes the location of one point in the 3D space plus its radial velocity, which is given as:

$$\vec{p} = (x_i, y_i, z_i, v_i) \tag{4}$$

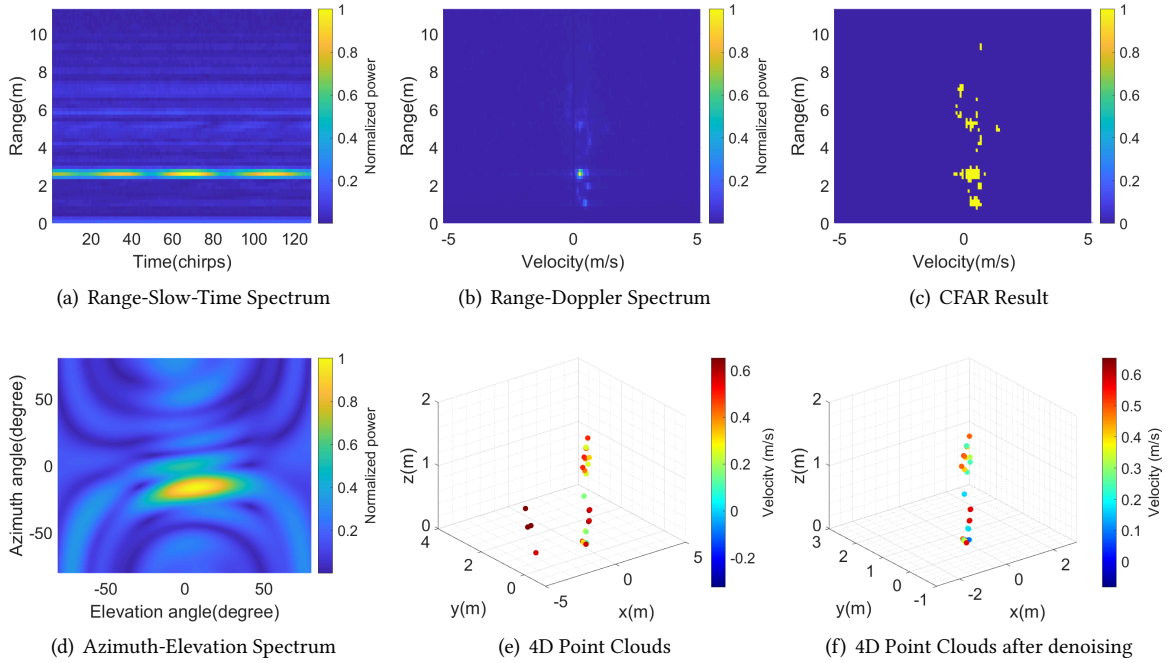
where i is point index in the current frame, x_i, y_i, z_i are the point's spatial coordinates, and v_i is the radial velocity. To obtain 4D points of signal reflectors from a radar frame, we design the following pipeline.

4.2.1 Range Estimation. As the range between the moving person and radar is linearly proportional to the frequencies of the IF signal, we first use Fast-Fourier Transform (FFT) to compute the frequency composition of each chirp [12, 63]. Specifically, the bandwidth of the mm-Wave radar we use is $1GHz$, which provides a distance resolution of $15cm$. When concatenating the FFT results of all the chirps in a frame, we can obtain the range-slow-time spectrum as shown in Figure 3(a) whose vertical axis represents range and the horizontal axis represents time.

4.2.2 Velocity Estimation. As mentioned in the previous subsection, the radial velocity is proportional to the frequency of the signal along the slow time. Hence, we can conduct the second FFT along the slow time dimension for each range bin in the range-time spectrum to extract the radial velocity [12, 63]. In detail, a frame contains 512 chirps (128 per TX), where one chirp has a duration of $60\mu s$, corresponding to a velocity resolution of $8cm/s$. Then, a range-doppler spectrum can be obtained as shown in Figure 3(b). A larger value in the spectrum means a higher probability of a moving target at the corresponding range and velocity.

4.2.3 Constant False Alarm Rate (CFAR).. To further clarify the range-doppler spectrum, we adopt an algorithm called CFAR [18]. CFAR combines range-doppler spectrums from different antennas into one 2D-array called Range-Doppler Matrix (RDM) and then uses an adaptive thresholding approach to classify each element on the RDM into targeted and untargeted. As shown in Figure 3(c), the yellow targeted elements indicate the presence of the targets at the corresponding range and velocity. Since our goal is to extract the point cloud of a moving human body, we classify all elements with zero velocity as non-target elements.

4.2.4 Phase Compensation in Virtual Array. We leverage the phase difference introduced by the target at different antennas to measure the angle. However, due to the adoption of TDM-MIMO, additional phase differences are



(a) Range-Slow-Time Spectrum (b) Range-Doppler Spectrum (c) CFAR Result

(d) Azimuth-Elevation Spectrum (e) 4D Point Clouds (f) 4D Point Clouds after denoising

Fig. 3. Results for each module in 4D point cloud extraction.It is worth noting that the normalized power in these heatmaps only represents the relative relationship of magnitudes (1 is the largest and 0 is the smallest), without specific units.

created by the person's movement in the interval between the chirps sent by different transmitting antennas [51]. The additional phase difference can be expressed as:

$$\Delta\varphi_{doppler} = 2\pi f_{doppler} T_{itv} \tag{5}$$

where $f_{doppler}$ denotes the Doppler shift of a target, T_{itv} denotes the time interval for transmitting antennas to transmit signal. In order to calculate accurate angles of targets, we need to eliminate such a phase difference. Fortunately, we observe that T_{itv} is determined at the time of radar configuration, so we are able to calculate $f_{doppler}$ during the *Velocity Estimation* stage.

4.2.5 Angle Estimation. The 3D FFT method is a widely used angle estimation algorithm, whose performance is, however, limited by the resolution of the signal. Instead, we adopt 2D digital beamforming (DBF) [46] and estimate the angle by controlling the beam to search the 2-dimensional plane consisting of azimuth and elevation directions step by step. Specifically, for both azimuth and elevation, the search step we employed here is one degree. A 2D angle spectrum can be obtained as shown in Figure 3(d), where the peak values indicate the presence of the target. Combined with the results of CFAR, we can get the azimuth and elevation of all moving objects.

4.2.6 4D Point Cloud. Finally, combining the range, azimuth, and elevation information above, we calculate the coordinates of each point of the moving object, forming 3D point clouds. Further considering the velocity information of each point, the 4D point cloud can be constructed.

4.3 Height Estimation

Ideally, we can estimate the human height if we have the 4D point cloud of the human head which contains the location of the human head inside 3D space. The estimation, however, faces significant challenges in practice due to the following two reasons:

- **Coordinate system.** The distance and angle of these point clouds represent the relative position relationship between the human and the radar, while the body's height represents the distance of the human body from the ground.
- **Multipaths affect.** The point clouds may contain noise caused by multipath in cluttered environments [45].

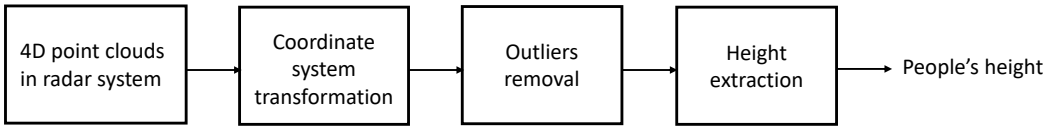
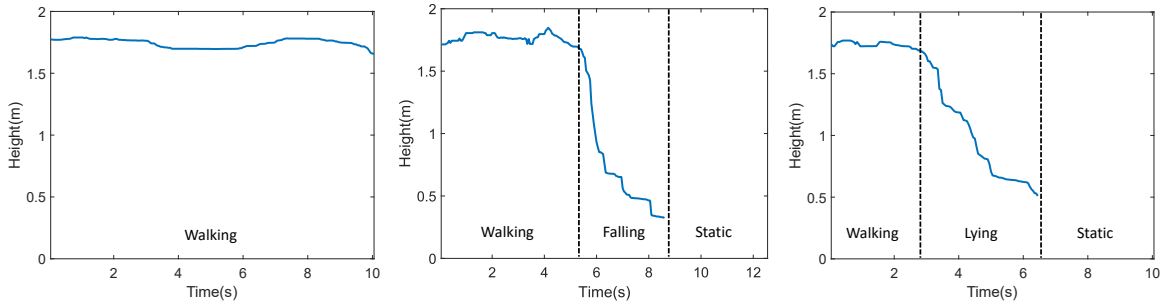


Fig. 4. Steps of height estimation.

To counter the above two challenges, we designed the height extraction algorithm, whose workflow is shown in Figure 4. The first step of the algorithm is to transform point clouds into the room coordinate system. To implement this transformation, we need the orientation and position of the mmWave radar in the room coordinate system, which can be measured by the extrinsic calibration of the radar [17]. According to the radar's orientation and position, we calculate the rotation matrix and translation vector of the transformation [17] and obtain 4D point clouds of the human body in the room coordinate system, as shown in Figure 3(e).

The second step is to remove the outlier point clouds introduced by the multipath effect. We observe that, compared to the point cloud of the human body, noise caused by multipath is sparse and can thus be removed based on point density. Specifically, we apply the DBSCAN algorithm [19] to remove the outliers and set the minimum number of points in the cluster to 10. Figure 3(f) shows the 4D point cloud after outliers removal. Besides, DBSCAN also groups the points when there are multiple moving objects [29], such as pets, robots, and people. Therefore, applying the DBSCAN algorithm not only effectively removes outliers, but also ensures that our algorithm can perform well in presence of multiple moving targets.



(a) Height while walking (b) Height while falling (c) Height while going tp the bed

Fig. 5. Height estimation results of walking, going to the bed and falling

At last, we estimate the human height by selecting three points inside the cleaned 4D point clouds that have the largest z-coordinates and use the median of their z-coordinates as the final estimation result. Due to the radar resolution and the variability of the reflection point, the estimated height may not be exactly the same as the true height, resulting in height estimation error. However, such an error won't affect the detection of human falls, since the height drop during a fall is significantly larger than the height estimation error.

We are now able to describe the height changes of a human fall by putting the height estimation of each frame in chronological order. For example, Figure 5(a) shows the height change when a person actually walks, from which we see that the height fluctuates around the person's real height (1.78m). Figures 5(b) and 5(c) illustrate the change in height for a person falling while walking and a person lying down on the bed, respectively. We note that, in both cases, the human point cloud cannot be extracted after a person lying/falling down and thus no longer has detectable movement. We see from these two figures that, utilizing the height change, we are able to distinguish fall and fall-like behaviors from other daily activities such as walking and running. However, because they share a similar pattern of height drop, it is hard to differentiate between falls and fall-like behaviors such as lying down.

5 MICRO-MOTION BASED BODY POSITIONING

After the height drops, a person may lie on a bed, sit on a chair, or fall on the floor, during which the person is likely to keep still. With no obvious movements, we no longer have enough velocity resolution to separate the 4D point cloud of the human target and thus cannot locate the person. In this section, we first introduce the intuition of how to locate the person without obvious movements in 5.1. We discuss in detail how to determine the range of people in 5.2. Then, we present the joint spatio-temporal angle estimation algorithm.

5.1 Intuition for Micro-motion Body Positioning

Observation. We observe that although the body no longer has limb movements with detectable velocity, there are still micro-motions like breathing and heartbeat, which are different from static environments.

Goal. Based on this observation, we aim to distinguish the person from the environment through such micro-motion so that we can locate the person again. The chest vibration of breathing is only about $5mm$, which is much smaller than the velocity resolution ($8cm/s$). Here comes another challenge: how can we distinguish such tiny movements from environments?

Insight. The inspiration stems from the trade-off between frequency resolution and temporal resolution in FFT [19]. When extracting the 4D point cloud, the velocity resolution is proportional to the Doppler frequency resolution, which is determined by observation time (the duration of a frame). So, can we extract the fine-grained 4D point cloud of micro-motion objects by extending the duration of a frame? Unfortunately, the duration of a frame cannot be extended indefinitely. The reason is that the 4D point cloud extraction algorithm is based on the assumption that the range and angle of the moving human body within a frame remain unchanged. Therefore, the 4D point cloud is an instantaneous representation of human spatial information and velocity. If the duration of a frame is too long, the human body may have a large movement and the extracted 4D point cloud will be meaningless. But for micro-motions such as breathing and heartbeat, the range and angle of the person remain constant for a considerable period of time. The relatively fixed position gives us the possibility of extending observation time, which will allow us to separate people from the environment by positioning micro-motion.

However, the duration of the frame is configured at the time of the radar starts. If we directly extend the duration of a frame, we will not be able to extract the 4D point cloud of moving people. To be compatible with the 4D point cloud based height estimation, we propose a joint spatio-temporal localization technique that extends the observation time by combining multiple frames instead of modifying the duration of a single frame. This

compatibility ensures that LT-Fall can not only capture targeted height drops in time but also estimate body height when there are only micro-motions.

5.2 Range Determination

In this section, we introduce the algorithm to determine the range between the body with micro-motions and the radar by combining multiple frames.

5.2.1 Long Time Window Data Acquisition. Simply stitching the data from multiple frames is the most straightforward approach. We note that, the difference between the inter-frame interval and the inter-chirp interval (chirps inside the same frame) results in unevenly distributed chirp in time, introducing significant error in the subsequent slow-time FFT.

We leverage the intra-frame averaging to solve such a problem. Because the time of a frame is very short, the micro-motion target can be considered as stationary within a frame. All the chirps in a frame can be considered as samples of the same environment. Specifically, chirps in a frame can be expressed as:

$$chirp_i = s(t) + n_i(t), i = 1, 2, ..., C \tag{6}$$

$chirp_i$ is the i_{th} chirp in the frame. $n_i(t)$ is the Gaussian noise in i_{th} chirp and C is the total number of chirps in a frame. $s(t)$ is the received reflection signal, which can be considered constant within a frame. We can average all chirps in the same frame to get a synthetic chirp for each frame. The synthetic chirp can be expressed as:

$$SC(t) = s(t) + n(t) \tag{7}$$

where $s(t)$ is the averaged chirp in the frame. It is worth noting that the power of $n(t)$ is significantly reduced because the Gaussian noises cancel each other out when averaging [37]. Since the time interval of the frames is uniform, the synthetic chirp is also uniform in time. In other words, intra-frame averaging not only solves the problem of uniform sampling, but also enhances the signal-to-noise ratio. After averaging, the sampling rate equals to the FPS, which is sufficient for micro-motions detection.

We then stitch together the synthetic chirps of different frames within a time window for subsequent processing. A time window data can be expressed as follows:

$$window = \{SC_j(t) | j = 1, 2, ..., F\} \tag{8}$$

where SC_j is the synthetic chirp of j_th frame in the time window, F is the total number of frames in the time window. In LT-Fall, we chose a time window size of 6 seconds, corresponding to a frequency resolution of 0.17Hz, which is sufficient to separate the breathing (0.2Hz to 0.4Hz) from the static environments. It is worth noting that the data in $window$ comes from only one channel. Our radar has four transmitting antennas and four receiving antennas, so there are sixteen channels, and each channel can get its corresponding $window$.

5.2.2 Range Bin Selection. The mmWave radar divides the space into a number of range bins by range FFT. Range determination is therefore translated into the problem of range bin selection. The signal magnitude of a range bin measures the intensity of the signal reflected by the objects at the corresponding range. So conventional methods select the range bin corresponding to the peak on the range spectrum as the human's location. However, except the human, there are many static objects that also reflect radar waves. The power of these static objects in the spectrum could be strong and thus prevents us from directly using the signal magnitude to select the range bin.

To address the above challenge, we separate them in the frequency domain by leveraging a long-time window data. Specifically, we perform range FFT for all synthetic chirps in $window$ (From Equation 8) to get the range-time spectrum, as shown in Figure 7. In order to select the range bin where the human body is located, we analyze the pattern of signal variation along slow time for each range bin, which can be expressed as:

$$H_i(t) = S_i + D_i(t) + N_i(t) \tag{9}$$

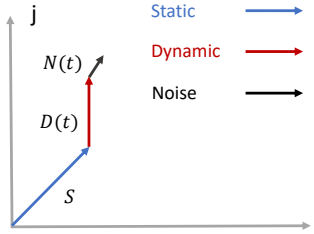


Fig. 6. $H_i(t)$ in complex plane

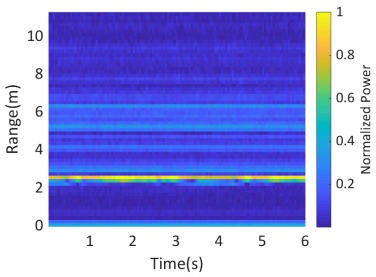


Fig. 7. Range-Time Spectrum

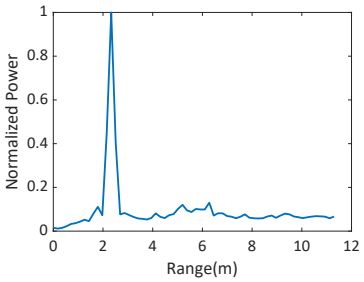


Fig. 8. Dynamic-Power-Range Spectrum

where i is the index of range bins, S_i is static component corresponding to the static environment, $D_i(t)$ is dynamic component corresponding to micro-motion people and $N_i(t)$ is the Gaussian noise. These three components can be represented by three complex vectors, as shown in Figure 6. The static vector does not change with time, while the dynamic vector may change in both phase and amplitude due to the micro-motions of the people.

The static vector and dynamic vector are mixed together in time domain. But, we can easily separate them in frequency domain because the large time window ensures that we have sufficient frequency resolution to distinguish the micro-motion body from the environment. Basically, the static environment contributes to the power of the zero-frequency part of the spectrum, while the power of the non-zero frequency part (i.e., dynamic power) is largely determined by the micro-motion of the body. As a consequence, if the micro-motion body is in a range bin, the dynamic power of the range bin that contains the human with micro-motions is maximized.

Based on the above analysis, we calculate the dynamic power of each range bin and select the range bin according to the strength the dynamic power. Dynamic power is calculated by accumulating the power of non-zero frequencies in the frequency spectrum obtained by FFT over time. Figure 8 shows the variation of dynamic power with range bins. The range bin corresponding to the highest peak represents the bin of the human body with micro-motions.

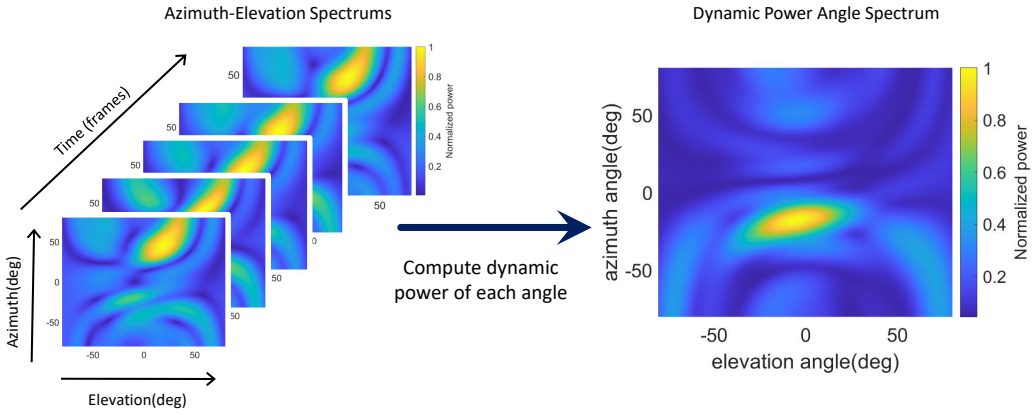


Fig. 9. An illustration of joint spatio-temporal angle estimation. Azimuth-elevation spectrums are first calculated. Both static environments and people are represented in these spectrums. To separate people from static environments, a dynamic power spectrum is calculated. In this spectrum, the peak implies the presence of human with micro-motion.

5.3 Joint Spatio-temporal Angle Estimation

To estimate the angle, our basic idea is to calculate the dynamic power of each angle and select the angle with the maximum dynamic power as the angle corresponding to the location of the person. This algorithm is shown in Figure 9. In order to calculate the angle spectrum of dynamic power, we adopt the 2D digital beamforming (DBF) approach to acquire the azimuth-elevation spectrums for each sample within the time window.

Based on the data of the people's range bin selected in 5.2.2, we control the beam to search the 2-dimensional plane consisting of azimuth and elevation directions step by step. To determine the steer vector, we establish a three-dimensional radar coordinate system, where the radar array is located in the $x-z$ plane and the first antenna is located at the origin of the coordinate system. Suppose there is a target whose azimuth angle is θ_{azi} and elevation angle is θ_{ele}, the phase difference between each antenna and the first antenna can be expressed as follows [26]:

$$\Delta\varphi_i = \frac{2\pi}{\lambda}(x_i \sin\theta_{azi} \cos\theta_{ele} - z_i \sin\theta_{ele}) \tag{10}$$

where $\Delta\varphi_i$ is the difference between i_{th} antenna and the first antenna, λ is the wavelength, x_i is x coordinate of i_{th} antenna and z_i is z coordinate of i_{th} antenna. Therefore, the steer vector can be expressed as:

$$a(\theta_{azi}, \theta_{ele}) = [1, e^{-\Delta\varphi_1}, ..., e^{-\Delta\varphi_n}] \tag{11}$$

where n is the number of antennas in the radar array.

Then, the signal vector consisting of the signals of each antenna in the range bin of the people at time t is expressed as:

$$s(t) = [s_1, s_2, ..., s_i, ..., s_n]^T \tag{12}$$

where s_i is the signal of i_{th} antenna. The value of the azimuth-elevation spectrum at (θ_{azi},θ_{ele}) is calculated by:

$$Spectrum(\theta_{azi}, \theta_{ele}) = a(\theta_{azi}, \theta_{ele}) * s(t) \tag{13}$$

The azimuth-elevation spectrum can be obtained by searching for each azimuth and elevation angle using the above method. The azimuth and elevation range we used are respectively $[-80°, 80°]$ and $[-80°, 80°]$, both with a step size of $1°$.

For each sample, we calculate an azimuth-elevation spectrum and obtain F spectrums within one time window. Both static objects as well as people with micro-motions are captured in these spectrums. In order to eliminate the influence of static objects, we stitch these F spectrums in time order to form a three-dimensional cube. Then we perform FFT over the cube along the time dimension, and accumulate the non-zero frequency power to calculate the dynamic power for each azimuth-elevation angle pair. Eventually, the dynamic power at each azimuth-elevation angle pair forms the angle spectrum of dynamic power as shown in Figure 9. The highest peak in the spectrum indicates the presence of a human with micro-motions at the corresponding angle.

Up to this point, the range and angle of the micro-motion body have been calculated. Based on these, the three-dimensional coordinates of the micro-motion body in the radar coordinate system can be easily computed. Then, the coordinates in the room coordinate system are obtained by coordinate system transformation. By calculating the position of the thorax - the body part where micro-motion occurs most often at - when the height of the thorax remains on the floor for a prolonged period of time, we can determine the person is lying on the ground. Besides breathing, other micro-motions such as slight struggles are also applicable under the present algorithm, since all these micro-motions generate dynamic power at the corresponding ranges and angles.

In real-life scenarios, there may be other objects other than the human moving in the room, such as pets, robots and other people, etc. With a novel mechanism, LT-Fall works well even under the influence of moving objects. The micro-motion human positioning algorithm is adopted only when the 4D point cloud extraction algorithm does not work. After the point cloud algorithm fails, the human only displays micro-motions, hence his position will not differ to a large extent. It is worth noting that DBSCAN is used in 4D point cloud extraction, which

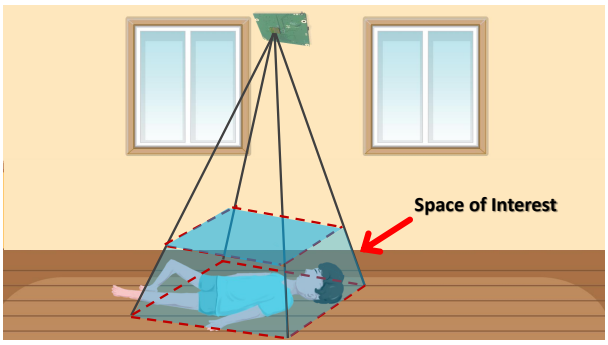


Fig. 10. An illustration of our interest space. When performing micro-motion people localization, we only search the space of interest. Moving objects outside this space will not affect our system. It is worth noting that the radar size in the picture is for display purposes only and does not represent the actual size.

could separate the point clouds of multiple moving targets. Therefore, the algorithm switch occurs when the 4D point cloud of any of the targets cannot be extracted. Based on this observation, we can record the approximate coordinates of the person before the algorithms switch. Then, at the stage of micro-motion positioning, the focus will be concentrated on a small area around the recorded coordinates. Other moving objects outside the area of interest will not affect the system, as shown in Figure 10.

6 EVALUATION

To evaluate the performance of LT-Fall, we will first introduce our system implementation and experimental setup, then study the performance of two key techniques respectively: 4D point cloud-based height estimation and micro-motion human localization. Last but not least, we will present the results of our investigation of the overall system performance of LT-Fall with 15 volunteers under different parameter settings in different environments.

6.1 Experiment Setup

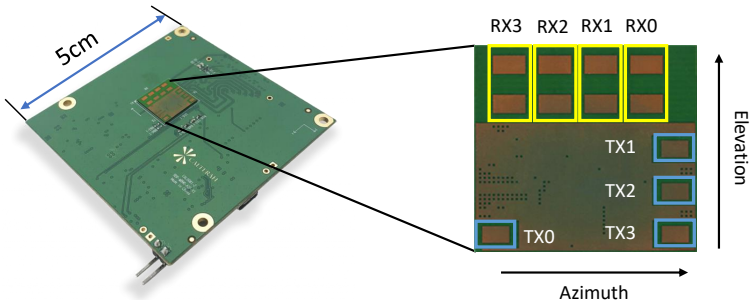


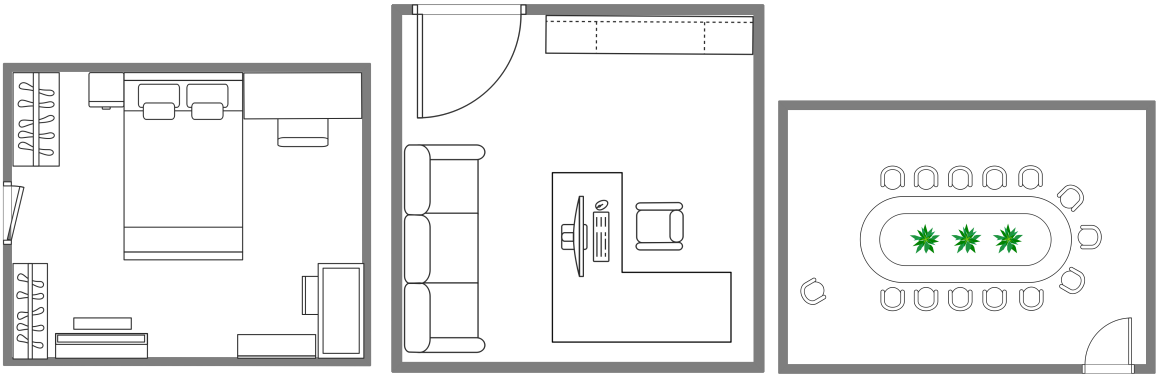
Fig. 11. CAL60S244-IB radar. On the left is the physical appearance of the CAL60S244-IB. On the right is the arrangement of CAL60S244-IB antenna array. [54]

LT-Fall is based on the Calterah's CAL60S244-IB, an antenna-in-package (AiP) radar sensor chip with 4 TX and 4 RX antenna array embedded in the package. It integrates a single-chip FMCW radar sensor capable of

operation in 60 to 64 GHz band. Figure 11 illustrates the physical appearance of the CAL60S244-IB radar devices and the arrangement of its antenna array.

Through TXs, the radar transmits radio signals into the environment. And signals reflected by the targets are received by RXs, then mixed to form the IF signal. The sampled IF signal is sent to a laptop for further analysis. Our default transmitting signal bandwidth is $1GHz$, which corresponds to a distance resolution of $15cm$. A frame contains 512 chirps (128 per TX), where one chirp has a duration of $60\mu s$, corresponding to a velocity resolution of $8cm/s$.

6.2 Performance of Height Estimation



(a) bedroom (3.8m × 4.6m) (b) office (3.8m × 3.8m) (c) meeting room (4.5m × 6.3m)

Fig. 12. The layouts of the three experimental rooms

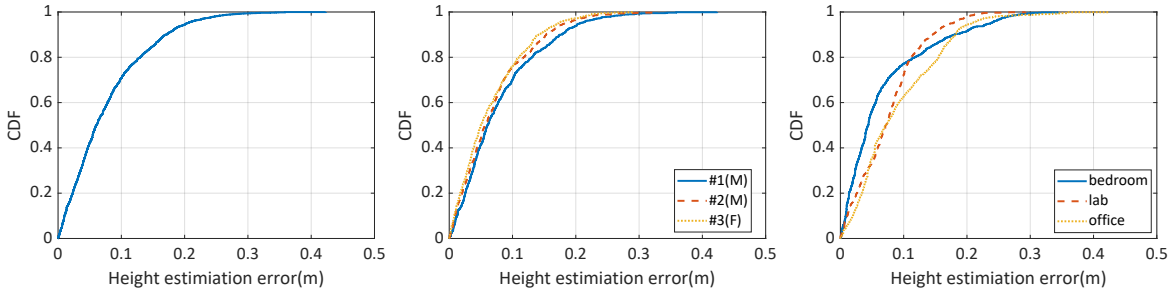
6.2.1 Data Collection. To fully evaluate the algorithm, the system is deployed in three typical environments: 1) a bedroom, 2) an office room and 3) a meeting room. The layouts of the three environmental rooms are shown in Figure 12. All three environments are furnished with common furniture such as sofas, tables, and chairs. Three volunteers have performed various activities in different environments, including walking, sitting down, working and calling on the phone. In terms of height estimation, these activities can be classified into sitting-height activities and standing-height activities. Table 1 shows the two kinds of height information of three volunteers.

Table 1. Information of the three volunteers

#Person	1	2	3
Standing height(cm)	178	172	165
Sitting height(cm)	120	116	113
Gender	M	M	F

6.2.2 Height Estimation. As shown in Figure 13(a), in about 80% of the cases, the height estimation errors were within $0.13m$. The mean value of the height estimation error is $0.0797m$ and the median is $0.061m$, which is sufficient to detect a height drop during a fall. To further demonstrate the system performance against individual diversity, we compared the system performance across different volunteers as shown in Figure 13(b). Evidently, the approach is robust to people of different heights, which benefits from our model-based height estimation

mechanism. Like how our method applies to different human subjects, it can adapt to different environments without any training effort as shown in Figure 13(c).



(a) Overall results (b) Impact of different people (c) Impact of different environments

Fig. 13. Performance of height estimation

6.3 Performance of Micro-motion Positioning

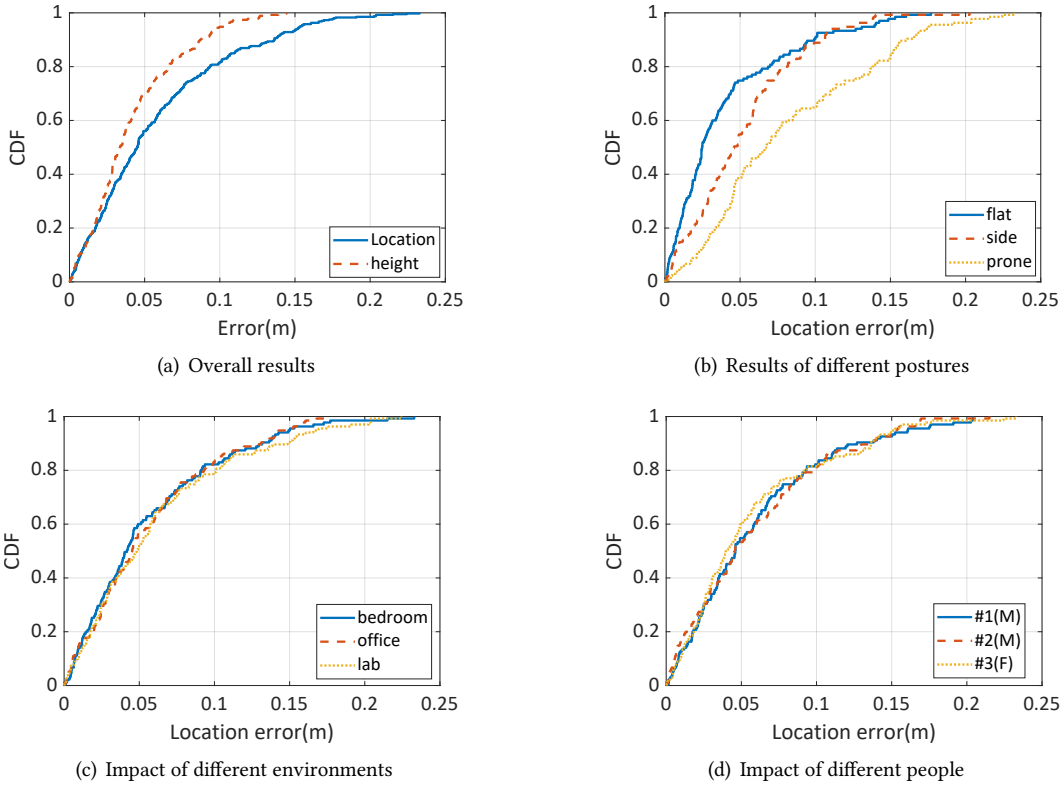
6.3.1 Data Collection. To evaluate the performance of the proposed micro-motion based body positioning algorithm, experiments were also conducted in the above three environments. As the posture of the person when falling is uncontrollable, volunteers are asked to perform micro-motion activities in eight different locations with three different postures: lying on the side, on the stomach and on the back. For each volunteer, 72 sets of data are collected. In total, 216 sets of data are collected. To obtain the ground truth, we used a laser rangefinder to measure the 3D coordinates (i.e., $x - y - z$) of the center point of the chest.

(a) lie on the back (b) lie on the side (c) lie on the stomach

Fig. 14. Different postures in each location.

6.3.2 Positioning Results. As shown in Figure 15(a), in 80% of the cases, the positioning error (i.e., $x - y$) of our algorithm is within $10cm$. In particular, the height error (i.e., z) is less than $10cm$ in 94% of the cases, successfully distinguishing falls from fall-like behavior. Figure 15(b) shows the results of the algorithm for different body postures, and it can be seen that the most accurate results are obtained when lying on the back. This is due to the fact that the micro-motions (breathing and heart beating) of the chest are the largest when lying on the back, and thus the error is the smallest. However, even lying on the stomach, with the largest error, is within $15cm$ of the

positioning error in 80% of cases, which is still much smaller than the height difference between falls and fall-like behaviors. Figures 15(c) and 15(d) demonstrate the excellent generalizability of the algorithm to different people as well as different environments.



(a) Overall results

(b) Results of different postures

(c) Impact of different environments

(d) Impact of different people

Fig. 15. Performance of micro-motion body positioning

6.4 Evaluation of Fall Detection

6.4.1 Data Collection. In order to comprehensively evaluate the overall system performance of LT-Fall, we considered a wide range of locations where the elderly may fall, and deploy the LT-Fall system in 7 different environments, including two different bedrooms, a shower room, a kitchen, an office room, a meeting room, and a laboratory. We recruit 15 volunteers of different heights, ages, and weights, 11 of whom were male and 4 were female, All volunteers are free to carry out their daily activities in 7 environments. Among these activities, they are asked to perform different kinds of long-lie falls and fall-like behaviors. Types of activities we have collected are listed in Table 2. In total, we have collected 500 sets of falls, 500 sets of fall-like behaviors, and 300 sets of other daily activities.

Moreover, we also build a real-time system and keep it continuously running for 74.2 hours in two volunteers' own bedrooms. These data are segmented into segments of 30*s* time windows, and we finally obtain a total of 8904 segments. During the system long-term run, 6 sets of falls and 62 sets of fall-like behaviors and 8526 sets

Table 2. Activities types in our dataset

	Types	scale
Falls	trip, slip, loss of consciousness, loss of balance, falling from the chair, falling when getting up	506
Fall-like behaviors	sit-down, picking up, lying on the bed, large objects falling	562
Daily activities	walking, sleeping, eating, drinking, working on the computer, washing and rinsing, jogging, watering flowers, playing the piano, opening door, etc.	8826

of daily activities are captured. As shown in Table 2, we collected a total of 506 sets of falls, 562 sets of fall-like behaviors and 8826 sets of daily activities in our evaluation.

6.4.2 Evaluation Metrics. To explain the metrics, we abbreviate True Positives, False Negatives, True Negatives and False Positives as TP, FN, TN and FP, respectively. Due to falls occurring much less frequently than non-fall behaviors in real-life scenarios, accuracy is no longer a reliable metric for fall detection. This is because even if all the data are classified as non-falls, accuracy is still high (94.88% in our dataset). Similarly, the false alarm rate is defined as $\frac{FP}{FP+TN}$. TN will dominate the calculation as long as the daily activities are correctly classified. For example, in our dataset, even if all fall-like behaviors are misclassified as falls, the false alarm rate is only 5.98%. Compared to fall-like behaviors, daily activities can easily be classified correctly, so TN should be excluded from the calculation formula in the evaluation metric.

- *Precision:* which is the fraction of corrected detected falls over all detected falls, i.e., $p = \frac{TP}{TP+FP}$.
- *Recall:* which is the fraction of corrected detected falls over the total falls, i.e., $r = \frac{TP}{TP+FN}$.
- *False discovery rate(FDR):* which is the fraction of incorrect detected falls over all detected falls, i.e., $p = \frac{FP}{TP+FP}$. From this equation, false discovery rate is equal to 1 minus precision. This metric measures the proportion of false alarms out of all detected falls. In this paper, we leverage this metric to measure the robustness of the system to false alarms.
- *F1 score:* which is the harmonic mean of precision and recall, i.e., $\frac{2pr}{p+r}$

6.4.3 LT-Fall's Overall Performance. Table 3 shows the overall system performance in terms of the above metrics. As we can see, LT-Fall achieves excellent performance for fall detection, with the precision, recall, FDR, and F1 score of 100%, 98.8%, 0, 99.4% respectively. We investigate several false negatives and found that they are mainly due to the blockage of furniture, which can be solved by deploying multiple radars in the room. However, thanks to our two-stage fall detection mechanism, the false alarms significantly reduced in LT-Fall.

Table 3. LT-Fall's Overall Performance.

	Precision	Recall	FDR	F1 score
LT-Fall	1	0.988	0	0.994
Aryokee[55]	0.919	0.938	0.081	0.929

In comparison, the previous SOTA Aryokee[55] achieves the precision of 91.9% and a recall of 93.8% . We note that the Aryokee is evaluated on another dataset. The comparison between their dataset and ours is demonstrated in Table 4, which shows that our dataset is more diverse. In addition, Aryokee is based on deep learning. Therefore, a large amount of labeled data is required to achieve its claimed performance, which is difficult in practice.

Table 4. The dataset comparison between LT-Fall and Aryokee

	different people	different environments	different radar products	different radar parameter settings
LT-Fall	✓	✓	✓	✓
Aryokee	✓	✓	✗	✗

6.4.4 Impact of Different Devices. To verify LT-Fall's generalization across different devices, in addition to the CAL60S244-IB, we also testify the system performance with another two radars, CAL70S244-AB (from Calterah) and AWR2944 (from Ti). CAL77S244-AB radar is also from Calterah, which operates at 76GHz to 81GHz. AWR2944 is a commercial mmWave radar manufactured by Texas Instruments operating at 60GHz to 64GHz. Table 5 shows the fall detection results of three different radar products. We observe that LT-Fall achieves zero false alarms as well as high recall on three radars, which demonstrates the generalization of our system across different radar products.

Table 5. LT-Fall's Performance over different radar products

Radar	Precision	Recall	FDR	F1 score
CAL60S244-IB	1	0.993	0	0.996
CAL77S244-IB	1	0.980	0	0.990
AWR2944	1	0.960	0	0.980

6.4.5 Impact of Different Bandwidths. For a radar system, the larger the bandwidth, the better the range resolution and the smaller the sensing distance. Thus, the evaluation of different bandwidths is conducted to demonstrate the robustness of LT-Fall. With a trade-off between sensing range and range resolution, we investigate the system performance at four bandwidths of 1G, 2G, 3G, and 4G. The experimental results are shown in Table 6.

Table 6. LT-Fall's performance over different bandwidths

Bandwidth(Hz)	Precision	Recall	FDR	F1 score
1G	1	0.990	0	0.995
2G	1	1	0	1
3G	1	0.975	0	0.987
4G	1	0.975	0	0.987

In order to cover the entire room, a large bandwidth requires a high sampling rate. However, the high sampling rate increases the amount of transmission data, which causes difficulties in subsequent signal processing, especially in real-time systems. As the results in Table 6 show that LT-Fall is working properly at all four bandwidths, we selected 1GHz as the default bandwidth for real-time data processing. Unless otherwise mentioned, most of our experiments are conducted at 1GHz bandwidth.

6.4.6 Impact of Other Moving Objects. When an elderly person falls, there may also be other moving objects in the room, such as pets, sweeping robots or other people. We conduct experiments to verify the system's performance in the presence of other moving objects. For convenience, we collect a dataset of people walking nearby while a person falls. The dataset is collected in two different environments with three people. For comparison, we use the

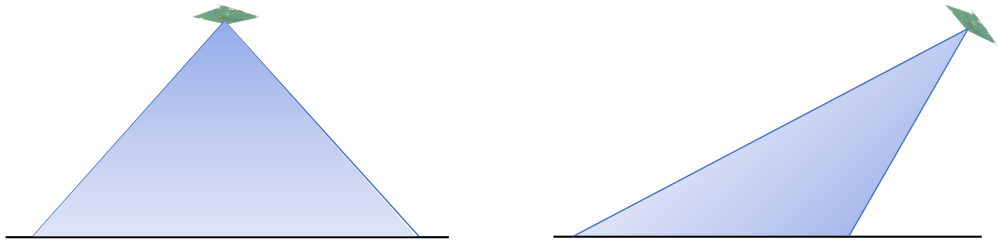
results presented in 6.4.3 and exclude the examples when multiple people exist in the environment. The results are reported in Table 7.

Table 7. LT-Fall's performance in the presence of other moving objects.

	Precision	Recall	FDR	F1 score
With other moving people	1	0.971	0	0.985
Without other moving people	1	0.991	0	0.995

By investigation, it is found that *Recall* decrease slightly due to the moving human body blocking the line of sight (LOS) of the radar and the faller. Overall, the system can still maintain good performance in the presence of multiple moving people.

6.4.7 Factors Affecting Radar Sensing Coverage. There are many factors that affect mmWave radar sensing coverage, such as bandwidth, FOV, and radar placement. At the same sampling rate, the smaller the bandwidth, the longer the sensing distance. For example, in LT-Fall, the sensing distance is approximately 11m when the bandwidth is configured to $1GHz$, which is suitable for indoor scenes. For FOV, of course, the larger it is, the greater the coverage. However, the FOV of a radar is determined by the radar's hardware, which is fixed at the time of radar production. According to Calterah's technical documentation and our own testing, the FOV of the CAL60S244-IB is ±75°in both azimuth and elevation. At the same time, the SNR decreases from 0 degrees to both sides of the FOV.



(a) Vertically downward placement.

(b) Diagonally downward placement.

Fig. 16. Two placement methods. The shaded area is the radar sensing area. When mounted diagonally downward, a blind zone of sensing will appear directly under the radar.

Once the bandwidth, FOV, etc. are determined, the factor that has the greatest impact on the system coverage is the placement of the radar. We have designed experiments to evaluate the coverage of two common placement methods: vertically downward (Figure 16(a)) and diagonally downward (Figure 16(b)). Our experiments show that a vertical downward placement can cover a square area with a side length of $6.5m$ centered on the radar projection on the ground at a radar height of $3m$. In contrast, it is found that the diagonal downward placement covers a square area with a side length of $8m$, but creates a sense blind zone directly under the radar.

Based on the above experimental results, we recommend the vertical downward placement method when the room area is comparatively small. When the vertical downward placement cannot cover the whole room, the diagonal downward placement can be adopted, while multiple radars can be considered to cooperate to eliminate the sensing blind area.

6.4.8 Ablation Study. LT-Fall is a two-stage fall detection system. To better understand the effectiveness of this design, we conduct the ablation experiments using the dataset in Section 6.4.3. The results are presented in Table 8. Using only the first stage of the LT-Fall, the *Precision* was only 54.6%. This is because fall-like behaviors also occur with height drop. By incorporating the second stage, the *Precision* is significantly improved to 100%. Noticeably, the increase in precision is accompanied by a slight drop in recall. This is due to the fact that people are more likely to be blocked when they are on the floor.

Table 8. Ablation study on each stage of LT-Fall.

	Precision	Recall	FDR	F1 score
Without second stage	0.546	0.992	0.454	70.5
Two-stage	1	0.988	0	0.994

7 DISCUSSION

In this section, we discuss the limitations and potential direction of our future work.

Blockage. We note that most of the undetected falls in our system are due to the blockage of furniture in the room. This is due to the weak penetration ability of mmWave radar and only falls in the LOS can be effectively detected. This problem can be solved by deploying multiple mmWave radars in the room. How to organically combine multiple radars to achieve fall detection without dead ends will be a very interesting problem. We leave this direction for our future work.

Effect of age factor on LT-Fall. We note that the participants (20 to 42 years old) recruited in our experiments are younger than the fall-prone elderly. However, when the subjects are changed to the elderly, the system will still function. This is because the age factor does not affect the operation of two key techniques of LT-Fall. For point cloud based height estimation, although the walking speed of the elderly is relatively slow, it is still much larger than the velocity resolution of the radar (8cm/s), therefore the point cloud can be extracted as normal. For micro-motion human positioning, the breathing amplitude of the elderly is significantly larger than the micron-level sensitivity of the mmWave radar [28], thus the micro-motion of the elderly can be detected and localized. The proper functioning of these two key techniques will ensure that LT-Fall can work properly even when users are the elderly.

Other applications of micro-motion body positioning. In many scenarios, human presence detection and positioning are the basic needs. However, existing technologies can only sense moving people. By using micro-motion body positioning, the scenarios of human presence detection and positioning can be fully expanded. For example, this technique can be used to detect the presence of babies unknowingly left in the car, potentially saving the lives of countless infants.

8 CONCLUSION

Traditional fall detection aims at detecting every possible human fall. For the first time, we point out that, for elderly caring applications, it is only meaningful to detect life-threatening fall, *i.e.,* the motion of falling down followed by the human lying down on the ground for a prolonged period of time instead of detecting every single falling behaviour. Accordingly, we implement LT-Fall, a system that detects life-threatening fall by leveraging mmWave radar. We solve the challenge of insufficient spatial resolution when the human is motionless. Extensive experimental results demonstrate that our system outperforms the state-of-the-art.

ACKNOWLEDGMENTS

This work is supported by NSFC A3 Foresight Program Grant 62061146001.

REFERENCES

[1] Stefano Abbate, Marco Avvenuti, Francesco Bonatesta, Guglielmo Cola, Paolo Corsini, and Alessio Vecchio. 2012. A smartphone-based fall detection system. *Pervasive and Mobile Computing* 8, 6 (2012), 883–899.

[2] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D tracking via body radio reflections. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*. 317–329.

[3] Fadel Adib and Dina Katabi. 2013. See through walls with WiFi!. In *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*. 75–86.

[4] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 837–846.

[5] Mostafa Alizadeh, George Shaker, João Carlos Martins De Almeida, Plinio Pelegrini Morita, and Safeddin Safavi-Naeini. 2019. Remote monitoring of human vital signs using mm-wave FMCW radar. *IEEE Access* 7 (2019), 54958–54968.

[6] Majd Alwan, Prabhu Jude Rajendran, Steve Kell, David Mack, Siddharth Dalal, Matt Wolfe, and Robin Felder. 2006. A smart and passive floor-vibration based fall detector for elderly. In *2006 2nd International Conference on Information & Communication Technologies*, Vol. 1. IEEE, 1003–1007.

[7] Moeness G Amin, Yimin D Zhang, Fauzia Ahmad, and KC Dominic Ho. 2016. Radar signal processing for elderly fall detection: The future for in-home monitoring. *IEEE Signal Processing Magazine* 33, 2 (2016), 71–80.

[8] Abhijit Bhattacharya and Rodney Vaughan. 2020. Deep learning radar design for breathing and fall detection. *IEEE Sensors Journal* 20, 9 (2020), 5072–5085.

[9] Zhen-Peng Bian, Junhui Hou, Lap-Pui Chau, and Nadia Magnenat-Thalmann. 2014. Fall detection based on body part tracking using a depth camera. *IEEE journal of biomedical and health informatics* 19, 2 (2014), 430–439.

[10] Etienne J Bisson, Elizabeth W Peterson, and Marcia Finlayson. 2015. Delayed initial recovery and long lie after a fall among middle-aged and older people with multiple sclerosis. *Archives of physical medicine and rehabilitation* 96, 8 (2015), 1499–1505.

[11] Jit Biswas, Daqing Zhang, Guopei Qiao, Victor Foo, Qiang Qui, and Philip Yap. 2006. A system for activity monitoring and patient tracking in a smart hospital. In *Proceedings of the International Conference on Smart Homes and Health Telematics, ICOST*. 196–203.

[12] Graham M Brooker et al. 2005. Understanding millimetre wave FMCW radars. In *1st international Conference on Sensing Technology*, Vol. 1.

[13] Kabalan Chaccour, Rony Darazi, Amir Hajjam el Hassans, and Emmanuel Andres. 2015. Smart carpet using differential piezoresistive pressure sensors for elderly fall detection. In *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 225–229.

[14] Yung-Chin Chen and Yi-Wen Lin. 2010. Indoor RFID gait monitoring system for fall detection. In *2010 2nd International Symposium on Aware Computing*. IEEE, 207–212.

[15] Chuanwei Ding, Hong Hong, Yu Zou, Hui Chu, Xiaohua Zhu, Francesco Fioranelli, Julien Le Kernec, and Changzhi Li. 2019. Continuous human motion recognition with a dynamic range-Doppler trajectory method based on FMCW radar. *IEEE Transactions on Geoscience and Remote Sensing* 57, 9 (2019), 6821–6831.

[16] Chuanwei Ding, Yu Zou, Li Sun, Hong Hong, Xiaohua Zhu, and Changzhi Li. 2019. Fall detection with multi-domain features by a portable FMCW radar. In *2019 IEEE MTT-S International Wireless Symposium (IWS)*. IEEE, 1–3.

[17] Joris Domhof, Julian FP Kooij, and Dariu M Gavrila. 2021. A joint extrinsic calibration tool for radar, camera and LiDAR. *IEEE Transactions on Intelligent Vehicles* 6, 3 (2021), 571–582.

[18] Jerry Eaves and Edward Reedy. 2012. *Principles of modern radar*. Springer Science & Business Media.

[19] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *kdd*, Vol. 96. 226–231.

[20] Jonathon Fagert, Mostafa Mirshekari, Shijia Pan, Pei Zhang, and Hae Young Noh. 2017. Characterizing left-right gait balance using footstep-induced structural vibrations. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2017*, Vol. 10168. SPIE, 357–365.

[21] Dustin P Fairchild and Ram M Narayanan. 2014. Classification of human motions using empirical mode decomposition of human micro-Doppler signatures. *IET Radar, Sonar & Navigation* 8, 5 (2014), 425–434.

[22] Homa Foroughi, Baharak Shakeri Aski, and Hamidreza Pourreza. 2008. Intelligent video surveillance for monitoring fall detection of elderly in home environments. In *2008 11th international conference on computer and information technology*. IEEE, 219–224.

[23] Jana Fortes, Michal Švingál, Tamás Porteleky, Patrik Jurík, and Miloš Drutarovskỳ. 2022. Positioning and Tracking of Multiple Humans Moving in Small Rooms Based on a One-Transmitter–Two-Receiver UWB Radar Configuration. *Sensors* 22, 14 (2022), 5228.

[24] Ajay Gadde, Moeness G Amin, Yimin D Zhang, and Fauzia Ahmad. 2014. Fall detection and classifications based on time-scale radar signal characteristics. In *Radar Sensor Technology XVIII*, Vol. 9077. SPIE, 330–338.

[25] Taekjin Han, Wonho Kang, and Gyunghyun Choi. 2020. IR-UWB sensor based fall detection method using CNN algorithm. *Sensors* 20, 20 (2020), 5948.

[26] Marlene Harter, Tobias Mahler, Tom Schipper, Andreas Ziroff, and Thomas Zwick. 2013. 2-D antenna array geometries for MIMO radar imaging by Digital Beamforming. In *2013 European Microwave Conference*. IEEE, 1695–1698.
[27] Sijie Ji, Yaxiong Xie, and Mo Li. 2022. SiFall: Practical Online Fall Detection with RF Sensing. (2022).
[28] Chengkun Jiang, Junchen Guo, Yuan He, Meng Jin, Shuai Li, and Yunhao Liu. 2020. mmVib: micrometer-level vibration measurement with mmwave radar. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–13.
[29] Feng Jin, Arindam Sengupta, and Siyang Cao. 2020. mmfall: Fall detection using 4-d mmwave radar and a hybrid variational rnn autoencoder. *IEEE Transactions on Automation Science and Engineering* (2020).
[30] Branka Jokanović and Moeness Amin. 2017. Fall detection using deep learning in range-Doppler radars. *IEEE Trans. Aerospace Electron. Systems* 54, 1 (2017), 180–189.
[31] Maarit Kangas, Antti Konttila, Per Lindgren, Ilkka Winblad, and Timo Jämsä. 2008. Comparison of low-complexity fall detection algorithms for body attached accelerometers. *Gait & posture* 28, 2 (2008), 285–291.
[32] Youngwook Kim and Hao Ling. 2009. Human activity classification based on micro-Doppler signatures using a support vector machine. *IEEE transactions on geoscience and remote sensing* 47, 5 (2009), 1328–1337.
[33] Qiang Li, John A Stankovic, Mark A Hanson, Adam T Barth, John Lach, and Gang Zhou. 2009. Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information. In *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*. IEEE, 138–143.
[34] Shengjie Li, Zhaopeng Liu, Yue Zhang, Qin Lv, Xiaopeng Niu, Leye Wang, and Daqing Zhang. 2020. WiBorder: Precise Wi-Fi based boundary sensing via through-wall discrimination. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–30.
[35] Wenxuan Li, Dongheng Zhang, Yadong Li, Zhi Wu, Jinbo Chen, Dong Zhang, Yang Hu, Qibin Sun, and Yan Chen. 2022. Real-Time Fall Detection Using Mmwave Radar. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 16–20.
[36] Yun Li, KC Ho, and Mihail Popescu. 2012. A microphone array system for automatic fall detection. *IEEE Transactions on Biomedical Engineering* 59, 5 (2012), 1291–1301.
[37] Yang Li, Dan Wu, Jie Zhang, Xuhai Xu, Yaxiong Xie, Tao Gu, and Daqing Zhang. 2022. DiverSense: Maximizing Wi-Fi Sensing Range Leveraging Signal Diversity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–28.
[38] Jie Lian, Xu Yuan, Ming Li, and Nian-Feng Tzeng. 2021. Fall Detection via Inaudible Acoustic Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–21.
[39] Ulrich Lindemann, A Hock, M Stuber, W Keck, and Clemens Becker. 2005. Evaluation of a fall detector based on accelerometers: A pilot study. *Medical and Biological engineering and computing* 43, 5 (2005), 548–551.
[40] Haipeng Liu, Yuheng Wang, Anfu Zhou, Hanyue He, Wei Wang, Kunpeng Wang, Peilin Pan, Yixuan Lu, Liang Liu, and Huadong Ma. 2020. Real-time arm gesture recognition in smart home scenarios via millimeter wave sensing. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 4, 4 (2020), 1–28.
[41] Liang Ma, Meng Liu, Na Wang, Lu Wang, Yang Yang, and Hongjun Wang. 2020. Room-level fall detection based on ultra-wideband (UWB) monostatic radar and convolutional long short-term memory (LSTM). *Sensors* 20, 4 (2020), 1105.
[42] Julien Maitre, Kévin Bouchard, and Sébastien Gaboury. 2021. Fall Detection With UWB Radars and CNN-LSTM Architecture. *IEEE Journal of Biomedical and Health Informatics* 25, 4 (2021), 1273–1283. https://doi.org/10.1109/JBHI.2020.3027967
[43] Muhammad Mubashir, Ling Shao, and Luke Seed. 2013. A survey on fall detection: Principles and approaches. *Neurocomputing* 100 (2013), 144–152.
[44] Sameera Palipana, David Rojas, Piyush Agrawal, and Dirk Pesch. 2018. FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–25.
[45] Sameera Palipana, Dariush Salami, Luis A Leiva, and Stephan Sigg. 2021. Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–27.
[46] S Unnikrishna Pillai. 2012. *Array signal processing*. Springer Science & Business Media.
[47] J Pena Queralta, Tuan Nguyen Gia, Hannu Tenhunen, and Tomi Westerlund. 2019. Edge-AI in LoRa-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks. In *2019 42nd international conference on telecommunications and signal processing (TSP)*. IEEE, 601–604.
[48] Ramesh Rajagopalan, Irene Litvan, and Tzyy-Ping Jung. 2017. Fall prediction and prevention systems: recent trends, challenges, and future research directions. *Sensors* 17, 11 (2017), 2509.
[49] Luis Ramirez Rivera, Eric Ulmer, Yimin D Zhang, Wenbing Tao, and Moeness G Amin. 2014. Radar-based fall detection exploiting time-frequency features. In *2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP)*. IEEE, 713–717.
[50] Frank C Robey, Scott Coutts, Dennis Weikle, Jeffrey C McHarg, and Kevin Cuomo. 2004. MIMO radar theory and experimental results. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, Vol. 1. IEEE, 300–304.

[51] Christian M. Schmid, Reinhard Feger, Clemens Pfeffer, and Andreas Stelzer. 2012. Motion compensation and efficient array design for TDMA FMCW MIMO radar systems. In *2012 6th European Conference on Antennas and Propagation (EUCAP)*. 1746–1750. https://doi.org/10.1109/EuCAP.2012.6206605

[52] Erik E Stone and Marjorie Skubic. 2014. Fall detection in homes of older adults using the Microsoft Kinect. *IEEE journal of biomedical and health informatics* 19, 1 (2014), 290–301.

[53] Bo Yu Su, KC Ho, Marilyn J Rantz, and Marjorie Skubic. 2014. Doppler radar fall activity detection using the wavelet transform. *IEEE Transactions on Biomedical Engineering* 62, 3 (2014), 865–875.

[54] Calterh Semiconductor Technology. 2017. Calterh Semiconductor Technology‘s Official Website. (2017). Accessed 1 November 2022. https://www.calterah.com/.

[55] Yonglong Tian, Guang-He Lee, Hao He, Chen-Yu Hsu, and Dina Katabi. 2018. RF-based fall monitoring using convolutional neural networks. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–24.

[56] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human respiration detection with commodity wifi devices: do user location and body orientation matter?. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 25–36.

[57] Hao Wang, Daqing Zhang, Yasha Wang, Junyi Ma, Yuxiang Wang, and Shengjie Li. 2016. RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices. *IEEE Transactions on Mobile Computing* 16, 2 (2016), 511–526.

[58] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking*. 65–76.

[59] Xuanzhi Wang, Kai Niu, Jie Xiong, Bochong Qian, Zhiyun Yao, Tairong Lou, and Daqing Zhang. 2022. Placement Matters: Understanding the Effects of Device Placement for WiFi Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–25.

[60] Yuxi Wang, Kaishun Wu, and Lionel M Ni. 2016. Wifall: Device-free fall detection by wireless networks. *IEEE Transactions on Mobile Computing* 16, 2 (2016), 581–594.

[61] Haowen Wei, Ziheng Li, Alexander D Galvan, Zhuoran Su, Xiao Zhang, Kaveh Pahlavan, and Erin T Solovey. 2022. IndexPen: Two-Finger Text Input with Millimeter-Wave Radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–39.

[62] Deidre Wild, US Nayak, and B Isaacs. 1981. How dangerous are falls in old people at home? *Br Med J (Clin Res Ed)* 282, 6260 (1981), 266–268.

[63] Volker Winkler. 2007. Range Doppler detection for automotive FMCW radars. In *2007 European Radar Conference*. IEEE, 166–169.

[64] Xinguo Yu. 2008. Approaches and principles of fall detection for elderly and patient. In *HealthCom 2008-10th International Conference on e-health Networking, Applications and Services*. IEEE, 42–47.

[65] Youwei Zeng, Dan Wu, Jie Xiong, Enze Yi, Ruiyang Gao, and Daqing Zhang. 2019. FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[66] Wenyu Zhang, Gang Li, Zetao Wang, and Hao Wu. 2022. Non-contact monitoring of human heartbeat signals using mm-wave frequency-modulated continuous-wave radar under low signal-to-noise ratio conditions. *IET Radar, Sonar & Navigation* 16, 3 (2022), 456–469.