Progress In Electromagnetics Research B, Vol. 115, 63–77, 2025

(Received 6 June 2025, Accepted 8 August 2025, Scheduled 31 August 2025)

Wireless Dual-Hand Motion Perception Based on Millimeter-Wave FMCW MIMO Radar

Haipeng Wang[1,3,*], Zhongfang Ren[1,2], Wei Pan[1], Zheng Xiao[1], and Yun Bo Li[3]

[1]*Research Center of Applied Electromagnetics*
Nanjing University of Information Science and Technology, Nanjing 210044, China
[2]*Tianchang Research Institute, Nanjing University of Information Science & Technology*
No. 28, Zonger Road, Tianchang 239300, China
[3]*State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China*

ABSTRACT: Radar-based hand gestures recognition has played an important role in developing human-computer interaction (HCI). However, when radar-based hand gesture recognition techniques are applied to multi-target scenarios, the challenges mainly involve problems of mutual interference and inaccurate recognition of hand motion when both hands move within the same plane. Here, we propose a dual-hand trajectory perception prototype based on a 60 GHz frequency-modulated continuous-wave (FMCW) multiple-input multiple-output (MIMO) radar sensor platform with an L-shaped virtual antenna array. To address the challenges, the approach involves estimating azimuth and elevation angles respectively from the two data components derived from the L-shaped array through multiple signal classification (MUSIC) algorithm, incorporating spatial division techniques combined with digital beam formation (DBF). Dual-hand applications mainly include angle targets at two distinct distances or dual-angle targets at the same distance. Therefore, the target distances are first determined using range fast Fourier transform (range FFT). If a single target distance is identified, we proceed to solve for the angles of two targets. Alternatively, if two distinct target distances are distinguished, we individually solve for the single-angle target corresponding to each distance. Furthermore, to mitigate noise inherent in the raw data of visualization, a frame point removal and smoothing algorithm is devised to refine the trajectories. Experimental verifications prove that the proposed multi-target motion perception algorithm by using a MIMO FMCW radar sensor platform can realize accurate recognition of air-writing gestures and enable tracking the trajectories of both single-handed and dual-handed targets in three-dimensional space. It also gives a new option for controlling the HCI.

1. INTRODUCTION

With the development of machine learning, hand gestures recognition techniques have played an important role in developing human-computer interface (HCI), including virtual reality control, medical health monitoring, indoor positioning and tracking. Traditional methods of gesture recognition mainly adopt wearable sensors [1], optical or depth cameras [2, 3] for data collection. Owing to the characteristics of non-invasiveness and low cost, electromyography sensors have been widely studied and applied in contact-based gesture recognition systems [4, 5], but the wearable usage will reduce the user experience. Vision-based systems have achieved remarkable success in gesture recognition research: employing multi-camera [6], structured light [7], and time-of-flight cameras [8] to reconstruct three-dimensional (3D) images or acquire 3D information. However, they can be easily affected by lighting conditions, background colors, configuration complexity, occlusion problems, and other optical factors, leading to inaccurate recognition results and limiting the applicability of optical technology in certain environments and conditions.

Due to the advantages of non-contact, penetrating obstructions, capabilities of being unaffected by lighting and smoke, supplying better user privacy, and full-time working conditions of electromagnetic (EM) waves, contactless gesture human target perception (or gesture recognition) systems based on short-range radar sensors have been developed rapidly, particularly those operating in the microwave and millimeter-wave frequency range. These technologies have significant applications in healthcare monitoring, elderly care, and smart homes, providing accurate and convenient ways to monitor vital signs without physical contact. The radar systems used for gesture recognition primarily fall into three categories: pulse radar [9, 10], single-frequency continuous wave (CW) radar (also known as Doppler radar) [11–16], and frequency-modulated continuous-wave (FMCW) radar [17–20]. In 2019, a Sweden team employed a pulse millimeter-wave radar to acquire high-frame-rate distance-time envelope feature data for 12 common gesture recognition tasks [10]. By combining transfer learning algorithms on pre-trained convolutional neural networks, they achieved an impressive accuracy of 99.5% on the validation set. Researchers introduced single-input multiple-output (SIMO) and multiple-input multiple-output (MIMO) structures into radar systems, using fewer transceiver channels to create the effect of large arrays, achieving higher angular resolution and more virtual channels. A novel Doppler radar using ambient wireless communication signals for posture sensing is proposed in [10]. Fan et al. designed a dual-channel 5.8 GHz Doppler radar sensor system [12]. They

---

* Corresponding author: Haipeng Wang (hpwang@nuist.edu.cn).

doi:10.2528/PIERB25060605

Published by THE ELECTROMAGNETIC ACADEMY

utilized an inverse sine algorithm and motion imaging algorithm to linearly reconstruct the demodulated Doppler phase shift for hand and finger movements in a two-dimensional (2D) plane. Zhang et al. [14], based on continuous wave MIMO Doppler radar sensors and a time-division multiplexing scheme, realized high-sensitivity 3D imaging in multipath indoor environments. Since Google Inc. designed a 60 GHz millimeter-wave radar chip embedded in smartphones to recognize various hand gestures [17], the research on gesture-target recognition methods using FMCW radar has gained substantial attention. Based on a 24 GHz SIMO FMCW radar, Wang et al. [18] reconstructed air-writing symbols within the observation plane through a series of signal processing, feature extraction, and smoothing algorithms. Gu et al. proposed a novel 4D gesture sensing technology based on 60 GHz FMCW radar with spatial-temporal detection to realize the detection of subtle gesture movements and air-writing [19, 20]. By utilizing MIMO technology to synthesize a virtual array and increase the array aperture, they improved spatial resolution for gesture tracking in the spatial domain.

Recently, researchers began employing cascaded multiple radars. Multiple Doppler radar sensors [21] and ultra-wideband (UWB) radars [22, 23] have been introduced to sense mid-air hand trajectories, to capture and recognize arm movements, and to locate finger movements in space. However, these works required large arrays or cascading multiple radars for posture sensing, or they relied on backend signal fusion and filtering algorithms for trajectory prediction, making the systems more complex. In addition, with the rising popularity of artificial intelligence, more scholars applied these technologies to radar-based hand movement perception systems. Typical gesture recognition processing algorithms include hidden Markov models (HMMs) [24, 25], random forests [17], convolutional neural networks (CNNs) [26–29], recurrent neural networks (RNNs) [30], transfer learning methods [31, 32], and self-attention neural networks [33]. Compared with mathematical modeling of continuous radar signals, neural networks (NNs) instead only model their statistical characteristics. Therefore, NNs are used for the features extraction of radar echo signal or range-Doppler map, modeling and classification for gesture recognition and tracking. However, the extensive use of deep learning generally requires more hardware and computing resources, increasing the cost and inconvenience of the system. To achieve a more refined depiction of hand movements, many studies fused various types of data (such as optical images) to form higher-dimensional data, but it is still challenging to operate in real-time environments.

The above studies focused on either static or dynamic gestures of single hand, and numerous deep learning methods have been applied to radar-based hand perception. For single-target scenarios, [20] focuses on necessitating algorithms for solving single-target angle estimation. However, when radar-based hand gesture recognition techniques are applied to multi-target scenarios, the challenges and complexities mainly involve the mutual interference of both hands moving within the same plane and the inaccuracy in recognizing hand movements. For instance, at time t_0, we can obtain the distance between the two hands as well as the azimuth and elevation angles, essentially treating them as two point targets. At the subsequent time t_1, we can still detect two point targets, but we are unable to determine the temporal continuity between the two points. In addition, the approach to tracking dual-hand movement relies solely on radar signal processing techniques, abstaining from the employment of intricate neural networks and enabling real-time operation, has not been investigated. To address the aforementioned issues, we propose the following solutions: 1) Using a spatial division method to directly constrain the movement of both hands within the azimuth angle, thereby obtaining the temporal continuity of the point targets of the hands; 2) Employing the high-resolution multiple signal classification (MUSIC) angle estimation algorithm to capture all possible angle spaces for the hand targets; 3) Then applying the digital beam formation (DBF) algorithm to identify the correct two-dimensional angle combinations from the angle space, thereby determining the 3D spatial coordinates of the hands. Finally, at each observation time point, the 3D coordinates of both hands can be obtained, allowing for the acquisition of the spatial movement trajectory of the hands.

Therefore, a dual-hand trajectory perception prototype based on a 60 GHz FMCW MIMO radar sensor platform with an L-shaped virtual antenna array is presented in this paper. The contributions of this paper are summarized as follows: The approach involves estimating azimuth and elevation angles separately from the two data components derived from the L-shaped array through multiple signal classification (MUSIC) algorithm, incorporating spatial division techniques combined with DBF. The proposed technique can divide the operational space into distinct zones, each assigned to a potential hand motion, thereby mitigating interference and enhancing tracking accuracy. The dual-hand applications mainly include angle targets at two distinct distances or dual-angle targets at the same distance. Thus, the target distances are first determined using range fast Fourier transform (range FFT). If a single target distance is identified, we proceed to solve for the angles of two targets. Alternatively, if two distinct target distances are distinguished, we individually solve for the single-angle target corresponding to each distance. Through the MUSIC algorithm to obtain precise target angle subspaces, subsequently the DBF is within the target angle subspaces to achieve accurate target angle estimation. Finally, after frame point data removal and smoothing processing, clear visualization of the spatial hand motion trajectories can be obtained.

2. METHODS

2.1. MIMO Radar Architecture and Virtual Array

The system diagram of the dual-hand motion perception system based on the FMCW MIMO radar sensor is shown in Fig. 1. The EM radiating signal is transmitted from transmitting antennas, reflected by both hands, and then received by receiving antennas. The received signal passes through devices such as low-noise amplifiers, mixers, and filters, and is finally collected by an analog-to-digital converter (ADC). MIMO radar refers to radar systems with multiple transmitting and receiving antennas. As shown in Fig. 1(a), three transmitters and four

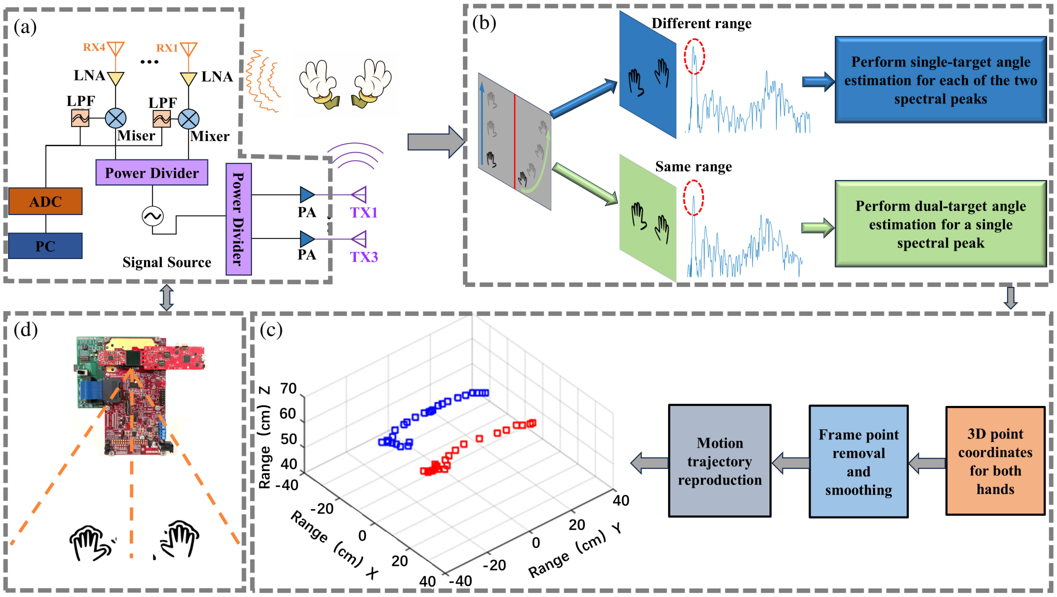

FIGURE 1. System architecture of dual-hand motion perception system based on FMCW MIMO radar sensor. (a) Block diagram of the motion sensing prototype. (b), (c) The motion trajectory recognition method for both hands. (d) The schematic representation of bilateral hand movement based on spatial segmentation.

receivers (3T/4R configuration) are employed in the prototype. Figs. 1(b)–(c) illustrate the method for recognizing the motion trajectories of both hands, where Fig. 1(b) depicts two scenarios arising when hands move within a vertical plane, each corresponding to a distinct peak in the distance spectrum, necessitating the adoption of specific processing strategies to acquire the targets of both hands. Fig. 1(d) provides a schematic diagram of dual-hand motion based on spatial segmentation, wherein the left hand is confined to the left-half space of the radar, and the right hand to the right-half space, with the division criterion being the azimuth angle of 0° of the radar serving as the separation line.

2.2. The Principles of Range Estimation and MUSIC Algorithm

1) Range estimation

In this radar system, the FMCW signal is internally generated by a phase-locked loop (PLL). The mathematical representation of the complex chirp signal within one pulse repetition interval can be expressed as

$$S_{Tx}(t) = \exp\left(j\left(2\pi f_c t + \pi\gamma t^2 + \varphi_0\right)\right), \quad t \in (-T_0/2, T_0/2) \tag{1}$$

where f_c is the center frequency of the FMCW signal, t the so-called 'fast time', $\gamma = B/T_0$ the slope of the ramp, B the bandwidth, T_0 the pulse repetition time (PRT), and φ_0 the initial phase. The transmitted signal propagates through space and undergoes reflection at the target. The reflected signal received by the radar has a time delay relative to the transmitted signal $t = 2R(\tau)/c$. $R(\tau)$ is the motion of the target including nominal distance d_0 and relative displacement $x(\tau)$, and τ is the slow time. The equation for the received signal is given by

$$s_{Rx}(t) = \rho s_{Tx}(t - \Delta t) \tag{2}$$

where ρ is the reflection coefficient. The transmitted and received signals are then mixed to obtain the differential beat signal as

$$\begin{aligned} s_b(t,\tau) &= s_{Tx}(t)\, s^*_{Rx}(t) \\ &= \rho \exp\left(j\left(\frac{4\pi\gamma R(\tau)}{c}t + \frac{4\pi f_c R(\tau)}{c} - \frac{4\pi\gamma R^2(\tau)}{c^2}\right)\right) \end{aligned} \tag{3}$$

where $4\pi\gamma R^2(\tau)/c^2$ is the residual phase that can be ignored, and '$*$' denotes the complex conjugate operation.

In the digital domain, the discrete form of the differential beat signal is:

$$s_b[n,m] = \rho \exp\left(j\left(\frac{4\pi B R[m]}{Nc}n + \frac{4\pi f_c R[m]}{c}\right)\right) \tag{4}$$

Here, n corresponds to the fast time; m corresponds to the slow time; and N is the number of samples in an FMCW period. The discrete differential beat signals are stored as a 2-D matrix. The FFT is applied to the discrete signal in fast time to obtain the frequency f at the peak of the FFT spectrum. The target distance is then calculated using the formula $d = \gamma{\cdot}f{\cdot}c/2$.

2) Direction of Arrival (DOA) estimation

Assuming a SIMO transmission model comprising an array of M uniformly spaced linear receiving antennas, with a spacing of d between them, and contemplating a far-field signal emission scenario, the signal path is elongated by an additional length of $d\sin\theta$ to reach the second antenna element relative to the first. Similarly, each subsequent antenna element experiences a progressive increase in the path length by $d\sin\theta$. Consequently, the time delay incurred between the signal's arrival at one antenna element and the subsequent one can be mathematically formulated as:

$$\delta t = \frac{d \sin\theta}{c} \tag{5}$$

where the velocity of EM waves is defined as C, and the wavelength is λ. Hence, considering the frequency denoted as f_0, the phase difference between two contiguous antenna elements can be derived as:

$$\delta\phi = \frac{2\pi f_0 d \sin\theta}{c} = \frac{2\pi d \sin\theta}{\lambda} \tag{6}$$

Assuming that $s(t)$ represents the far-field transmitted signal, the signal received by the entire antenna array can be formulated as:

$$x(t) = \left[1, e^{j2\pi\frac{d\sin\theta}{\lambda}}, \cdots, e^{j2\pi\frac{(M-1)d\sin\theta}{\lambda}}\right] s(t) + n(t) \tag{7}$$

Then, we define the steering vector A as follows:

$$A = \left[1, e^{j2\pi\frac{d\sin\theta}{\lambda}}, \cdots, e^{j2\pi\frac{(M-1)d\sin\theta}{\lambda}}\right] \tag{8}$$

The noise associated with the ith array element at time t is represented by $n_i(t)$. Extending the framework from the SIMO setup, a MIMO spatial model can be formulated. Supposing D transmitting signals, the received array signal within the MIMO configuration can be mathematically expressed as:

$$x(t) = \begin{bmatrix} \mathrm{e}^{j2\pi\frac{0d\sin\theta_1}{\lambda}} & \cdots & \mathrm{e}^{j2\pi\frac{0d\sin\theta_D}{\lambda}} \\ \mathrm{e}^{j2\pi\frac{d\sin\theta_1}{\lambda}} & \cdots & \mathrm{e}^{j2\pi\frac{d\sin\theta_D}{\lambda}} \\ \vdots & \vdots & \vdots \\ \mathrm{e}^{j2\pi\frac{(M-1)d\sin\theta_1}{\lambda}} & \cdots & \mathrm{e}^{j2\pi\frac{(M-1)d\sin\theta_D}{\lambda}} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_D(t) \end{bmatrix} \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \\ \vdots \\ \eta_M(t) \end{bmatrix} \tag{9}$$

Therefore, the received signal can be expressed as Equation (9). Consequently, estimating $x(t)$ enables the inference of the DOA. We can also derive Equation (10) and Equation (11).

$$A = \left[a(\theta_1), \cdots, a(\theta_D)\right] \tag{10}$$

$$a(\theta) = \left[1, \mathrm{e}^{j2\pi\frac{d\sin\theta}{i}}, \cdots, \mathrm{e}^{j2\pi\frac{(M-1)d\sin\theta}{\lambda}}\right]^H \tag{11}$$

It should be noted that the vector $a(\theta)$ is designated as the steering vector, fundamentally characterizing the spatial phase and being intimately tied to the configuration and relative positioning of the array elements. Here, the superscript H denotes the Hermitian transpose (conjugate transpose) operation. Based on the N snapshot data received by the array, the received data is sampled to obtain its covariance matrix estimate, denoted as $\hat{R}_x$. Subsequently, the estimated covariance matrix $\hat{R}_x$ undergoes eigenvalue decomposition. The resulting eigenvalues $\hat{U}_N$ are then utilized to calculate the corresponding eigenvectors, which are subsequently sorted according to the magnitude of the eigenvalues to delineate the signal subspace from the noise subspace. Finally, a spectral peak search is conducted over θ to locate the desired parameter. The MUSIC algorithm spectrum is scanned by iterating over potential directions θ using in Equation (12) to locate the spectral peaks. The corresponding angle is determined from the spectral peak, yielding an estimate of the DOA.

$$P(\theta) = \frac{1}{a(\theta)^H \hat{U}_N \hat{U}_N^H a(\theta)} \tag{12}$$

The classical DBF algorithm attains the suppression or enhancement of target signals in diverse directions through weighting and phase adjustment of array signals, thereby facilitating the derivation of the target's angle of arrival. Nevertheless, achieving high angular resolution with a limited number of antenna array elements poses a challenge. To address this issue, a super-resolution angle estimation algorithm is adopted to promptly identify potential target angles, effectively generating probable 2D angle combinations within the target angle space. Subsequently, the steering vectors previously introduced are constructed from the target angle space. The DBF algorithm is then employed to assist in identifying the target angle combination within this space that most closely resembles the actual target angle. This selected combination serves as the final target angle estimation, thereby enabling the attainment of a more accurate target angle as feasible.

2.3. Recognition Method of Two-Hand Motion Trajectories

Figure 2 illustrates the processing flowchart for recognizing the motion trajectories of both hands. When hands move within the same plane, the distinct motions of the left and right hands alternate their distances from the radar between identical and disparate values. Considering the varying positions of the hands within this plane, it is imperative to adopt distinct angle estimation strategies. Following the parsing and preprocessing of raw data, a range FFT is initially performed to obtain the distance from the spectral peak estimation. Within the corresponding range, a peak search is conducted to determine the number of spectral peaks. In the presence of a single peak, it is inferred that the distance between both hands and the radar is identical, necessitating the application of an angle estimation for two targets at the same distance. Conversely, when multiple peaks (greater than one) are detected, indicating that the hands are positioned at different distances from the radar, an angle estimation strategy tailored for a single target at varying distances should be used.

1) Acquisition of target angle space

Figure 3 illustrates the process diagram of obtaining target angle space. The numerals 1∼12 in the figure signify the 12-channel data virtually generated via time-division multiplexing (TDM) MIMO, where Channels 1, 3, 5, 9, and Channels 2, 4, 6,

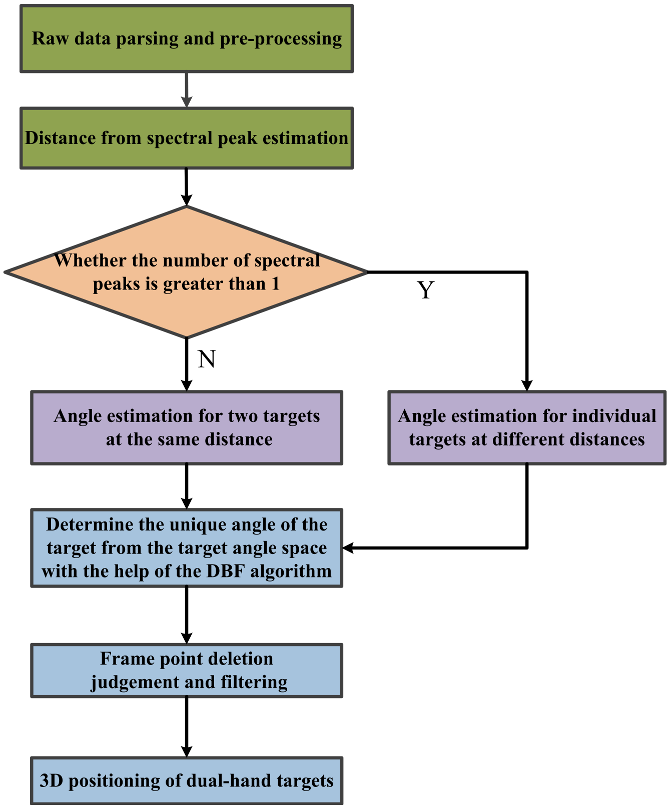

FIGURE 2. Flowchart of the dual-hand motion trajectory recognition method.

10 represent the channel data of virtual array elements in the elevation plane, utilized for elevation angle acquisition; whereas Channels 5, 6, 7, 8, and Channels 9, 10, 11, 12 correspond to the channel data of virtual array elements in the azimuth plane, employed for azimuth angle estimation. The two raw data presented in Fig. 3 represent the fast and slow time-domain data of selected equivalent virtual receiving channels within a single frame, corresponding to their respective planes. Performing range FFT on these samples enables the extraction of target distances. Notably, the red and purple portions in the figure signify the extracted data at the target range index, which are subsequently utilized for elevation and azimuth angle estimation, respectively. The MUSIC algorithm is employed to independently estimate the azimuth and elevation angular spectra of the target. Then, the angle values corresponding to each local maximum in the spectrum (note: not directly taking the angle value at the peak) are obtained to form a larger set of angle values for the potential angle space of the target.

2) 2D angle estimation strategies for individual targets at varying distances

When the hands move on a plane relative to the radar, at certain stages, they may differ in distance from the radar, necessitating the application of distinct 2D angle estimation strategies for individual targets at varying distances. The DBF algorithm process flow for estimating the 2D angles of a single target at different distances is illustrated in Fig. 4. Initially, a single (θ, φ) pair is sequentially extracted from the previously acquired target angle space to construct azimuth and elevation angle steering vectors (green vectors in Fig. 4), which are then multiplied with the original data matrix. This results in a corresponding value for each (θ, φ) pair processed by the DBF algorithm. However, due to the L-shaped layout of the virtual antenna array, the original data cannot naturally form a matrix. Hence, it is restructured into two matrices, as shown in Fig. 4. The raw data within the yellow dashed box contains more elevation angle information, leading to the selection of the elevation angle corresponding to the maximum value. Conversely, the raw data within the purple dashed box encapsulates more azimuth angle information, and thus, the azimuth angle associated with its maximum value is chosen. Finally, this method yields the estimated azimuth and elevation angles for individual targets at varying distances, which are then fed into a frame point deletion smoothing algorithm for further processing.

3) 2D angle estimation for dual targets at the same distance

When both hands move within a plane at an equal distance from the radar during certain phases, a two-dimensional angle estimation strategy tailored for dual targets at the same distance should be adopted. The DBF algorithm for this case deviates from that for a single target, as illustrated in Fig. 5. Given the constraint of hand movements within two sections of space, symmetrically positioned around the azimuth angle of 0°, the MUSIC algorithm is directly applied to estimate the DOA for the azimuth angles of both targets, yielding an estimated azimuth angle for Target 1 (less than 0°) and an estimated azimuth angle for Target 2 (greater than 0°). Subsequently, the elevation angles are sequentially extracted from the estimated elevation angle space, forming elevation angle steering vectors (green vectors in Fig. 5) along with the already estimated azimuth angles of each target, which constitute azimuth angle steering vectors. These vectors are then multiplied with the data matrix enclosed by the yellow dashed box, and the maximum value of each row is taken to correspond to the respective elevation angle. This process culminates in the estimation of both azimuth and elevation angles for both targets, which are also then subjected to frame point deletion and smoothing algorithms for further refinement.

4) Frame point removal and smoothing algorithm

The flowchart of the frame point removal and smoothing algorithm is shown in Fig. 6. The obtained distance estimation, estimated azimuth angle, and elevation angle described above are further processed as follows: First, frame index determination: Several frames of target data are retained without processing and stored in a smoothing array as the initial values. Starting from the next frame of target data, the algorithm proceeds to the next step. Second, frame point judgement: For the current frame point data, the algorithm checks if the distance between the target and the radar falls within a predefined threshold range. It is determined by comparing the distance between the current frame point and the previous frame point. If the distance does not meet the threshold criteria, the frame point is considered an invalid data and excluded from the smoothing array, which means that this point will not be smoothed and visualized. If the distance meets the threshold criteria, the frame point is included in the smoothing array. Therefore, the main approach is to perform difference comparisons between

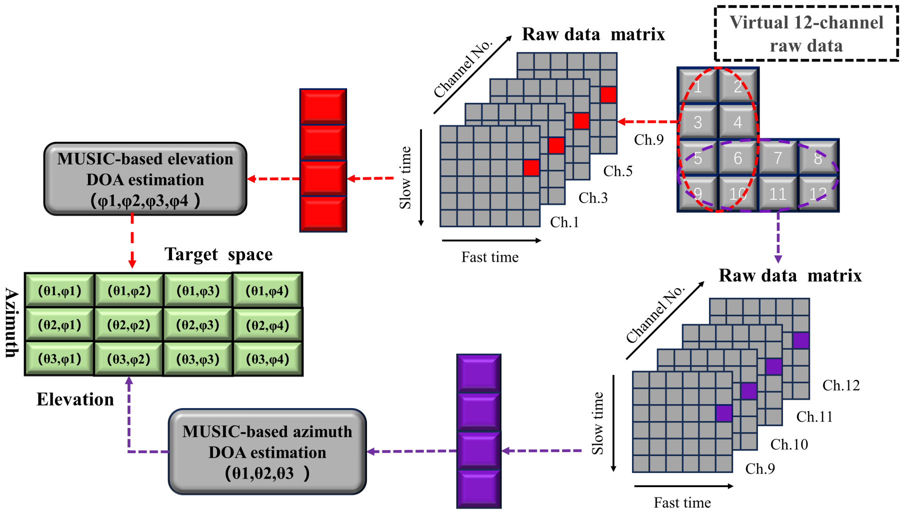

FIGURE 3. Flowchart of the target angle space acquisition process.

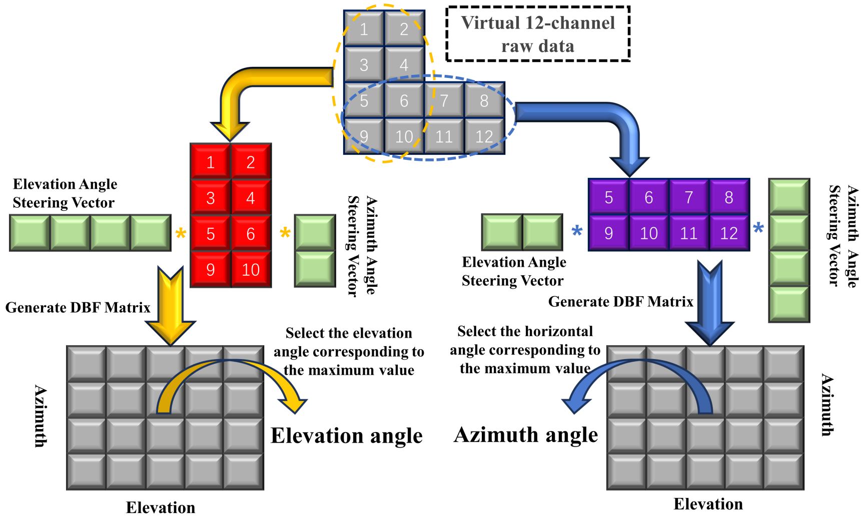

FIGURE 4. Processing flow of DBF algorithm applied to single-target motion perception.

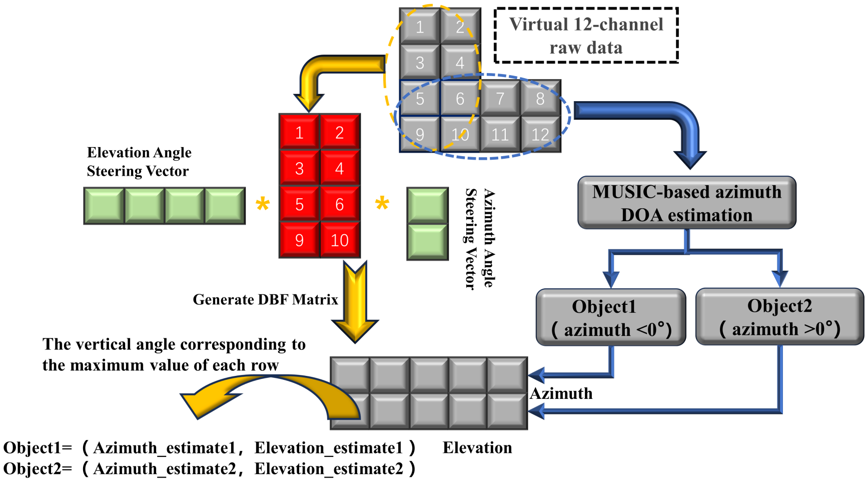

FIGURE 5. Processing flow of DBF algorithm applied to dual-target motion perception.

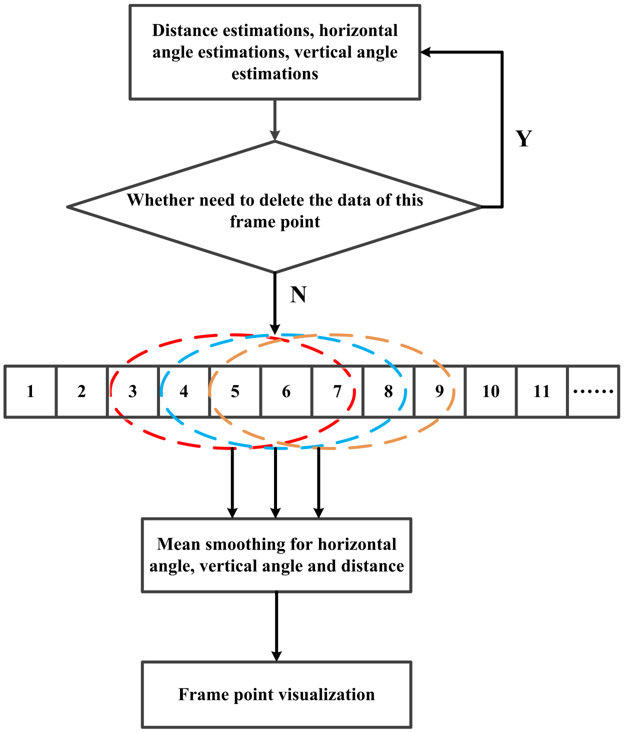

FIGURE 6. Flowchart of frame point removal and smoothing algorithm.

frame point data at adjacent time intervals, with thresholds set for differences smaller than 1 cm or greater than 6 cm. Since the removal of frame points is primarily aimed at handling excessively large outliers and closely spaced neighboring points, any points exceeding the defined thresholds are deleted. Third, smoothing and visualization: For each frame point data included in the smoothing array, the algorithm calculates the average data (distance, azimuth angle, and elevation angle) by taking the mean values of the current frame point and the previous M frame points (M is taken as 5 or 6). The resulting smoothed values are then visualized and displayed. The value of M has a significant impact on the smoothing effect. If M is too small, it will not achieve sufficient smoothing, whereas if M is too large, it can lead to over-smoothing. This may result in the loss of accurate depiction of certain gesture movements. Fourth, the above steps are repeated for each subsequent frame point data in the sequence. This algorithm can automatically remove invalid data and smooth the target data over consecutive frames, providing more stable and visually appealing results.

3. EXPERIMENTS AND RESULTS

The experiments were conducted using our customized MIMO radar prototype system, which incorporates a 60 GHz radar SoC with L-shaped virtual antenna array configuration. A photograph of the radar platform and antenna array is presented in Fig. 7(a). The radar system consists of three transmitters (Tx) and four receivers (Rx). The transmitter module consists of three parallel transmitter chains, each with independent binary phase and amplitude control units. Each transmitter channel can also be individually enabled or disabled. The receiver module consists of four parallel channels. Each receiver channel consists of a Low-Noise Amplifier (LNA), a passive mixer, a high-pass filter, and a Variable Gain Amplifier (VGA). All four receiver channels can work simultaneously. Time-division multiplexing (TDM) method is used to change the state of control units of each transmitter channel in a specific order. The FMCW radar operates at a frequency range of 60 to 64 GHz, with a transmission power of 0 dBm, a center frequency of 62 GHz, and an FMCW bandwidth of 4 GHz. The repetition time for each chirp is 60 µs; the sampling frequency (f_s) is 10 MHz; and 512 samples are taken per chirp pulse period. The data of the intermediate frequency (IF) signal from the radar is recorded by a data acquisition board connected to the radar platform, and the baseband data is transmitted to the computer via an Ethernet cable.

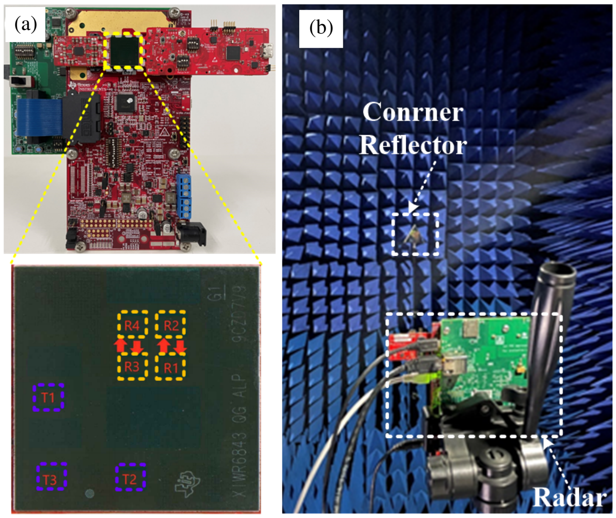

FIGURE 7. TI 60-GHz MIMO radar platform. (a) Photograph of the radar platform and radar SoC IWR6843AOP with antennas in package. (b) Photograph of the radar calibration setup.

3.1. Calibration

Given the variability in signal path lengths and losses or gains across different transmitter and receiver channels, which introduces inter-channel consistency deviations, calibration is essential to compensating these discrepancies. The calibration process is conducted within an anechoic chamber equipped with corner reflector (CR), as depicted in Fig. 7(b). The radar platform is mounted on a stand 80 cm above the ground. The basic idea of calibration is that under far-field conditions, each channel in the same plane receives signals with the same amplitude and phase when the signals are incident from the normal direction of the array. The far-field condition for both the Tx and Rx antennas is defined as $L_F = 2D^2/\lambda$, where LF represents the minimum distance for far-field conditions, D the antenna aperture, and λ the wavelength. In this experiment, with $\lambda = 5$ mm and $D = 7.5$ mm, LF is calculated to be 22.5 mm. Within the calibration setup, the CR is positioned 1.85 m away from the radar, centered and aligned accordingly.

Figures 8(a) and (b) illustrate the amplitude and phase of range profiles for 12-channel data in azimuth plane before calibration, respectively. The data from the first channel is used

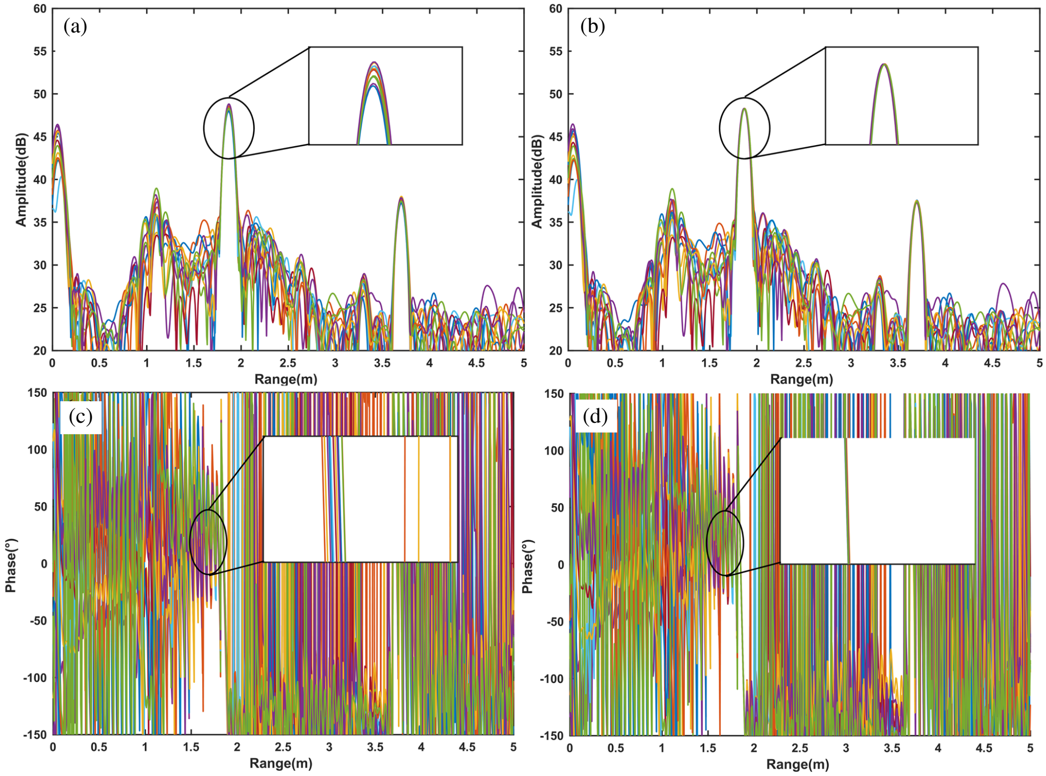

FIGURE 8. Range profiles of 12-channel data in azimuth plane. (a) Amplitude and (b) phase of uncalibrated data. (c) Amplitude and (d) phase of calibrated data.

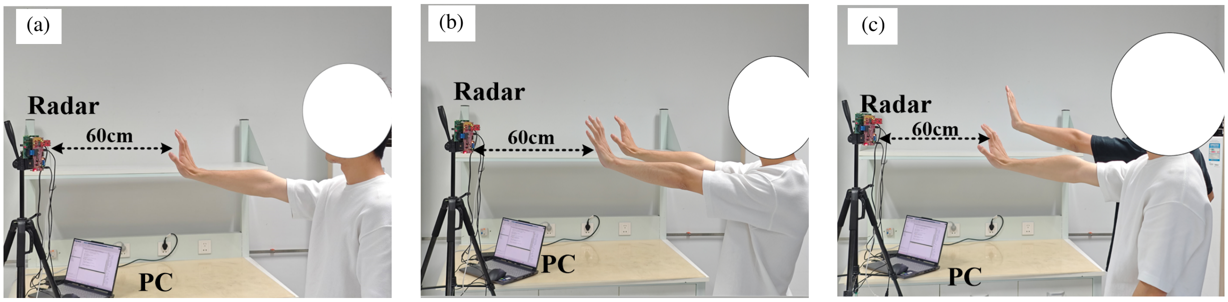

FIGURE 9. Photograph of experimental setup for motion perception measurements. (a) Scenario with one person and one hand. (b) Scenario with one individual and a pair of hands. (c) Scenario with two hands and two individuals.

as a reference to normalize the amplitude and align the range and phase of the distance distribution for each channel at the location of the CR. The calibrated range profiles for amplitude and phase are shown in Figs. 8(c) and (d), respectively. The results show that the amplitude and phase of each channel after calibration are well aligned with the position of the CR. This data calibration matrix is further utilized in subsequent 3D positioning data processing.

3.2. Air-Writing Gestures Setup

This experiment demonstrates the capability of writing digits, letters, and symbols in air using hands, with the setup illustrated in Fig. 9. Throughout the test, the movements of the hands were kept within an elevation plane at approximately the same distance from the radar. For some complicated digits such as ‘2’ and ‘3’, slower movements were suggested to capture more details, facilitating subsequent frame point deletion and smoothing algorithms for deriving clear hand motion trajectories. The experiment comprises three distinct scenarios:

1) Unimanual motion scenario (Fig. 9(a)). It involves a single individual using a single hand, with the hand’s motion plane maintained at 60 cm from the radar and a motion range about 20 cm. This scenario primarily serves to validate the angle estimation performance for a single target at varying distances.

2) Bimanual mirror motion scenario (Fig. 9(b)) features a single individual using both hands. The distance between movement plane of the hand and the radar is kept at about 60 cm, with a hand motion range about 20 cm. The results are mainly used to verify the effect of angle estimation for dual targets under the same distance.

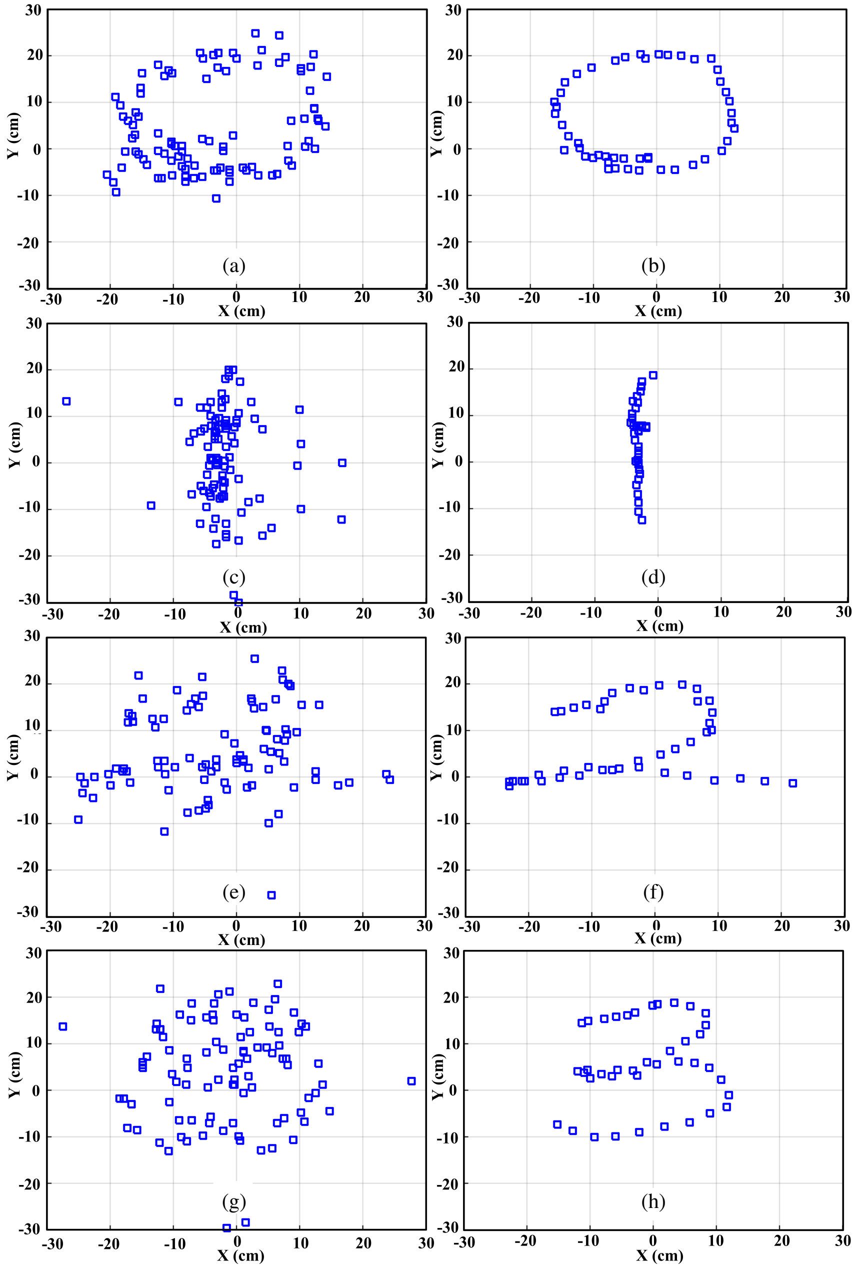

FIGURE 10. Experimental results of single-hand motion perception. Subfigures (a), (c), (e), and (g) show the results without frame point removal and smoothing processing. Subfigures (b), (d), (f), and (h) represent the results obtained after applying the frame point removal and smoothing algorithm.

3) Bimanual asynchronous motion scenario (Fig. 9(c)) is a two-handed scene for two individuals, with the hands' motion planes maintained at 60 cm from the radar and a motion range about 20 cm. This scenario primarily aims to assess the handling of distance switching when two sets of hands move within the same plane, thereby combining the challenges presented in the previous two scenarios.

3.3. Results of Frame Point Removal and Smoothing Algorithm

1) Unimanual motion perception processing

In the single hand motion perception experiments, the demonstration involved tracing the digits '0', '1', '2', and '3' in the air. Figs. 10(a), (c), (e), and (g) illustrate the results obtained directly from the DBF algorithm for unilateral motion perception, without applying frame point deletion or smoothing. These results exhibit numerous discrete and repetitive points, rendering the hand motion trajectories difficult to discern. Conversely, Figs. 10(b), (d), (f), and (h) display the outcomes following the application of a frame point deletion and smoothing algorithm. It is evident that the processed hand motion trajectories are clearer and more readily identifiable.

2) Bimanual motion perception processing

In the experiments of dual-hand motion perception, demonstrations involved simultaneously drawing patterns of '1-1', '(-

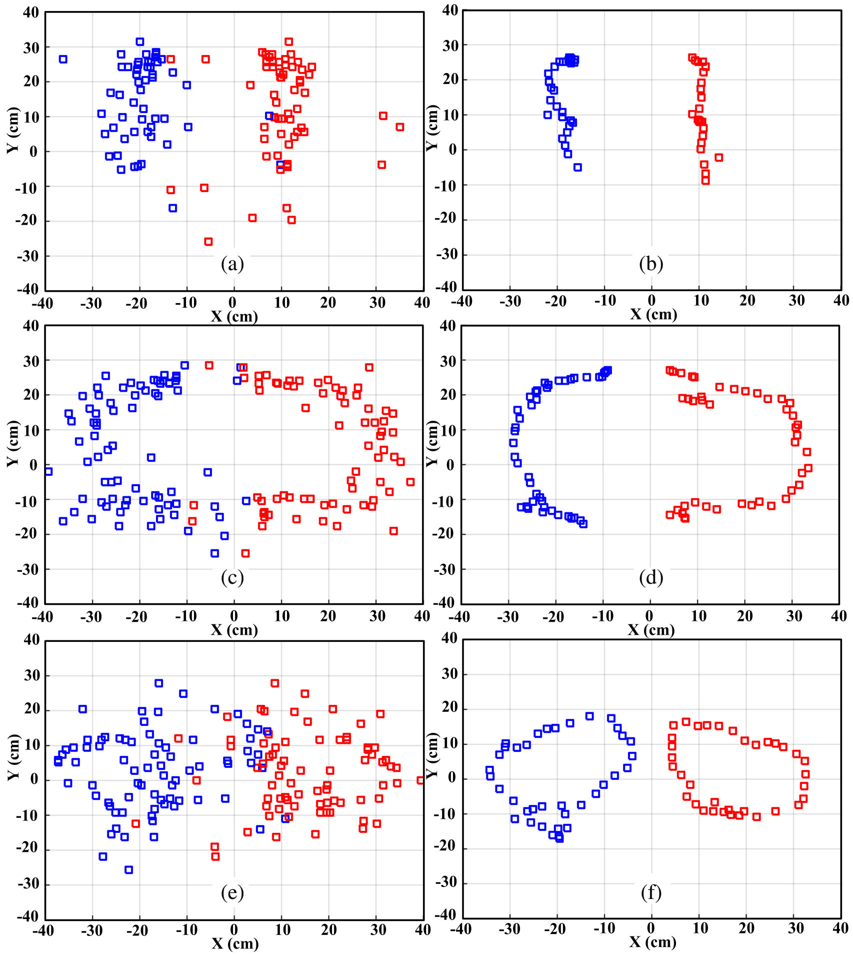

FIGURE 11. Experimental results of dual-hand motion perception. Subfigures (a), (c), and (e) show the results without frame point removal and smoothing processing. Subfigures (b), (d), and (f) represent the results obtained after applying the frame point removal and smoothing algorithm.

)', and 'O-O', with both hands on a plane at an equal distance. Figs. 11(a), (c), and (e) present the raw data obtained directly from the DBF algorithm for bilateral motion perception, without applying frame point deletion or smoothing. It can be observed that there are many scattered and repeated points, and the hand motion trajectories of both hands are not easily recognizable. Figs. 11(b), (d), and (f) present the results after applying the frame point removal and smoothing algorithm. The hand motion trajectories become clearer and easier to identify after the processing. Furthermore, in the example of simultaneously drawing circles with both hands, in Fig. 11(e) the hand motion trajectories are barely discernible for identification, whereas Fig. 11(f) effectively reveals the spatial motion trajectories of both hands.

3.4. Results of Air-Writing Gestures

1) Scenario 1: Unimanual motion perception

This section emphasizes the 3D results for single-hand motions involving a broader range of gestures. The processed data will display the refined trajectories of various hand gestures after they are subjected to frame point elimination and filtering smoothing operations to ensure the most accurate and clean representation of the hand's motion through 3D space. Fig. 12 presents composite illustrations of the 3D and 2D results for single-hand motion perception. The top-right corner of each subfigure displays the 2D projection of the 3D trajectory onto the XY-plane, representing the outcomes of various digits and letters drawn by the hand at a distance of 50 cm from the radar. Figs. 12(a)–(j) depict the digits '0' through '9', and Figs. 12(k)–(o) illustrate the letters 'A' through 'D', along with an intersection symbol. The figure distinctly outlines the precise trajectory of each hand movement, even capturing intricate details such as the continuous and closely adjacent segments of the more complex letters 'A' and 'B'. Notably, even for the intricate hand motions involved in drawing digits like '4', '5', and '8', clear trajectories are observable and closely match handwritten numbers.

2) Scenario 2: Bimanual mirror motion perception

Figure 13 shows the 2D and 3D results of hand movement perception for bimanual mirror activities, where Figs. 13(a)–(h) display symmetrical drawing of '1', '2', '3', a elevation line with a diagonal slash, '()', '7', 'z', and 'D', respectively. Fig. 13(i) represents the symmetrical drawing of digit '3' with an additional azimuth line using both hands, where the '3' is

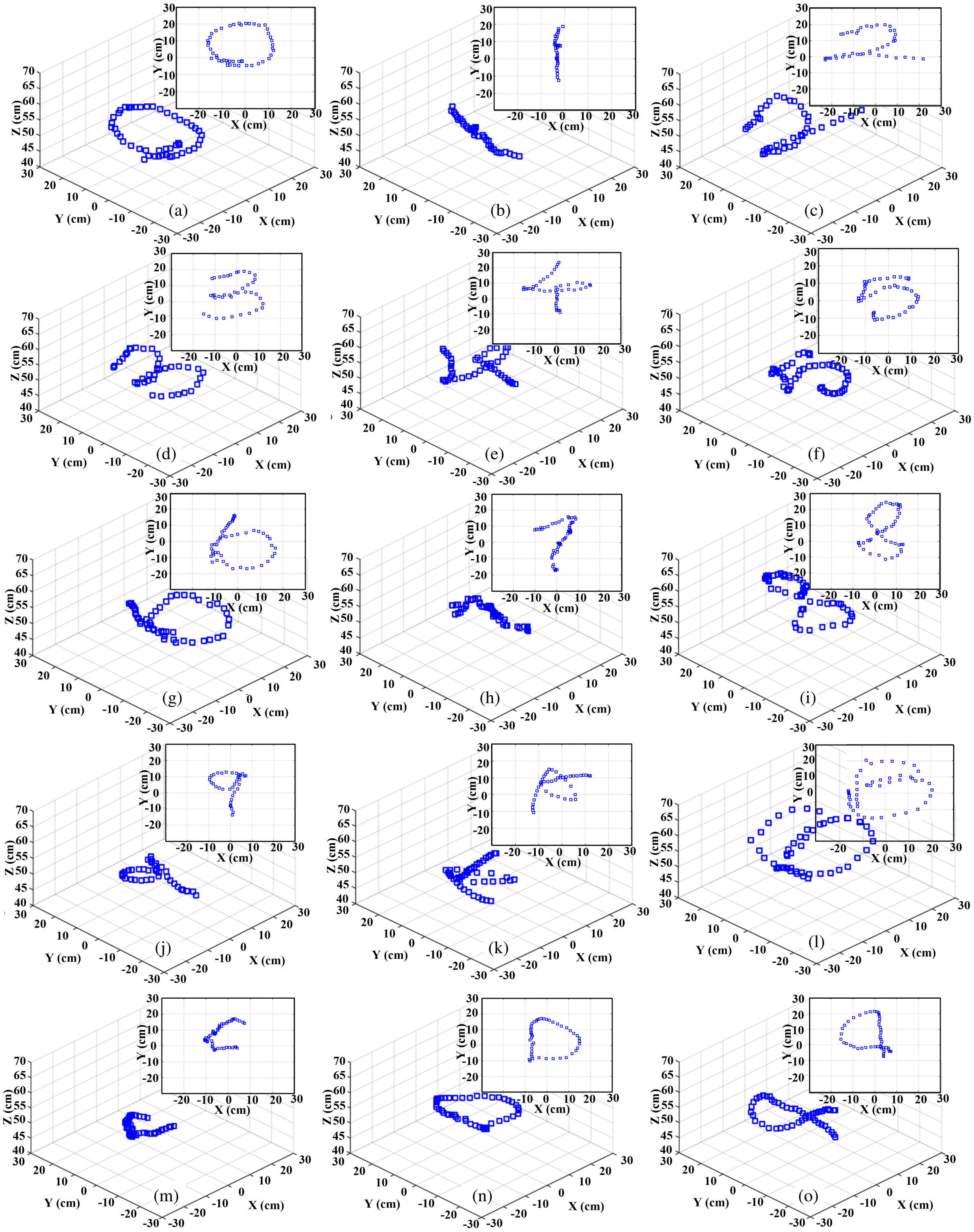

FIGURE 12. Composite illustrations of the 3D and 2D experimental results for single-hand motion perception, with subfigures (a) to (j) representing numbers '0'-'9', subfigures (k) to (o) depicting letters 'A' to 'D', and an intersection symbol.

drawn as a zigzag, distinguishing it from the semicircular form in the mentioned Fig. 13(c). In this figure, the blue trajectory represents the left-hand movement, while the red trajectory corresponds to the right-hand movement. By separating and plotting symmetrical patterns for each hand on the same plane (with the left hand moving in the left half of the XY plane and the right hand in the right half), we achieve bimanual hand mirror movement perception. The experimental results clearly demonstrate that the traced hand movement trajectories align with actual hand movements, thereby validating the feasibility of the proposed spatial division-based approach for realizing bimanual hand movement perception.

3) Scenario 3: Asynchronous bimanual motion perception

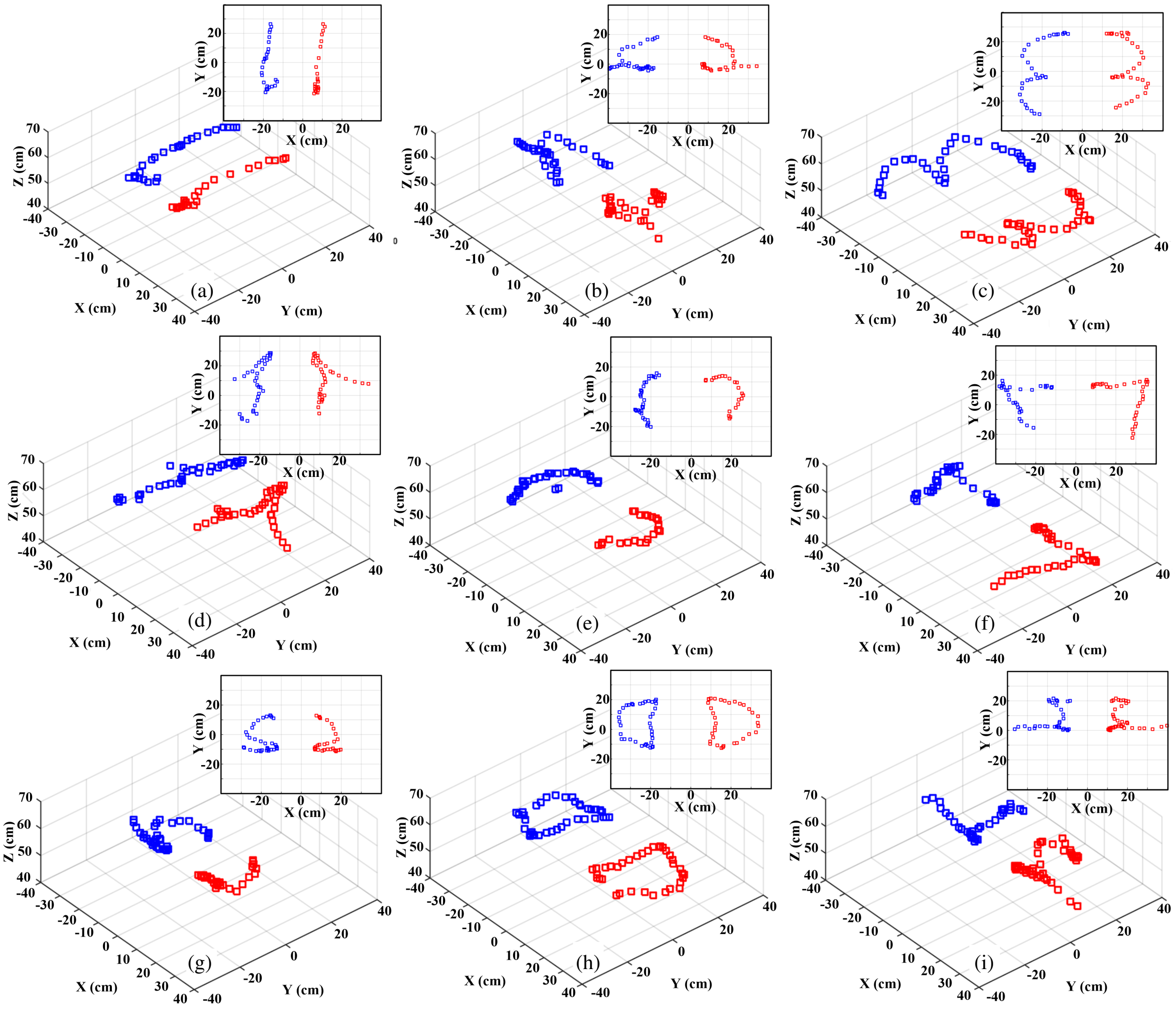

FIGURE 13. Composite illustrations of the 3D and 2D experimental results for the perception of bimanual mirror activities. Subfigures (a) to (j) depict symmetrical drawings of the following patterns: '1', '2', '3', an elevation line with a diagonal slash, parentheses '()', '7', 'z', and 'D'. Notably, subfigure (i) shows the digit '3' drawn symmetrically with an additional azimuth line..

Figure 14 illustrates the recognition results of asynchronous hand motion trajectories performed by both hands from two individuals. The blue color in each subfigure represents the detected trajectory of the gesture drawn by the left hand, while the red color corresponds to that of the right hand. According to the 3D and 2D plot shown in Fig. 14(a), it can be easily obtained that the digits '1' and '0' are drawn by left and right hands, respectively. The result can be represented as '1-0'. Similarly, Figs. 14(b)–(f) are the trajectories of '1-7' gesture, '1-9' gesture, '1-2' gesture, '1-4' gesture, and 'triangle-square' gesture. The 2D trajectory plots indicate that the recognized results closely resemble the actual hand motion trajectories, which can be effectively distinguished between the distinct gesture motions performed by the left and right hands.

In addition, a comparative analysis with other similar works is presented in Table 1. It reveals that the proposed method achieves superior range and angle resolution, particularly when employing a small array and notably under the L-shaped virtual MIMO array configuration. Furthermore, the introduced sensing technique enables 3D imaging of both single-handed and dual-handed targets, along with tracking the trajectories of gestures in 3D space, thereby presenting a novel solution for dual-hand gesture recognition.

4. DISCUSSION

1) This study demonstrates the feasibility of our millimeter-wave radar-based dual-target motion perception approach through qualitative trajectory analysis. The core innovation is a novel signal processing framework for concurrent hand movement localization. Future work will include: (1) quantitative benchmarking against optical motion capture systems (RMSE, maximum deviation), and (2) standardized classification tasks (e.g., character recognition) with performance metrics (accuracy, F1-score). These evaluations will further validate clinical applicability while maintaining advantages in occlusion robustness and privacy preservation over vision-based systems.

2) The current study employs offline processing for algorithm validation, while maintaining a lightweight design prin-

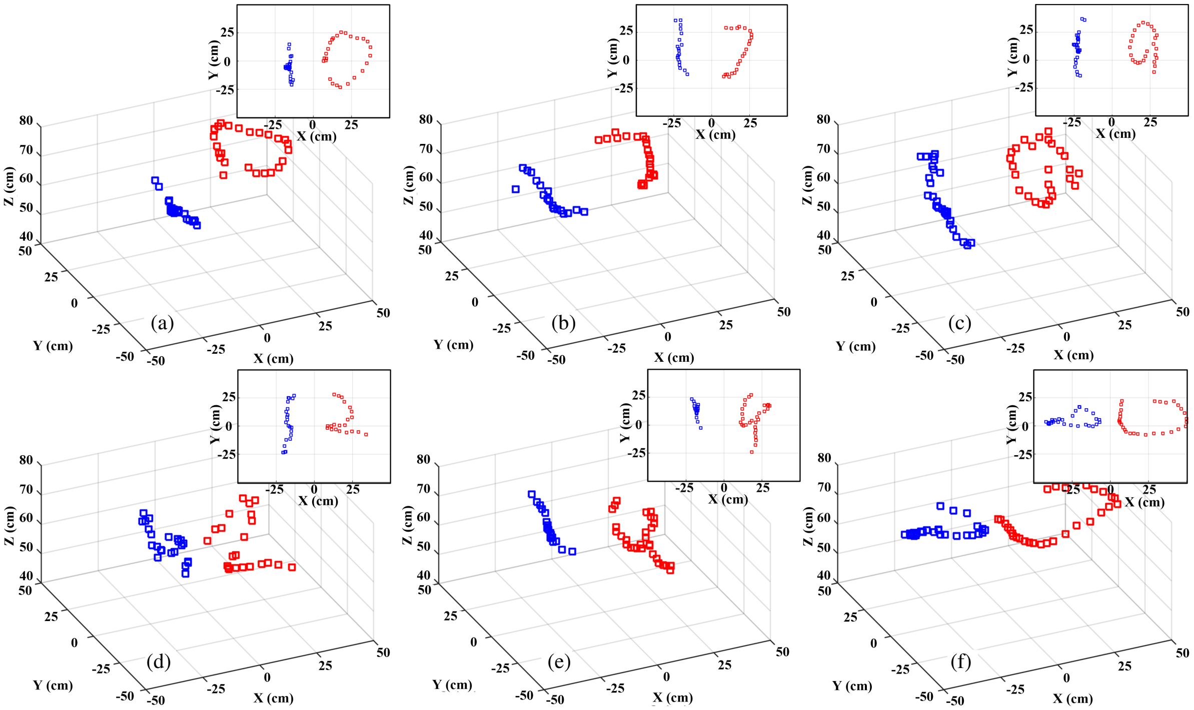

FIGURE 14. Composite illustrations of the 3D and 2D experimental recognition results of asynchronous hand motion trajectories performed by both hands from two individuals. (a) ‘1-0’ gesture, (b) ‘1-7’ gesture, (c) ‘1-9’ gesture, (d) ‘1-2’ gesture, (e) ‘1-4’ gesture, and (f) ‘triangle-square’ gesture.

TABLE 1. Experimental results of angular resolution using two corner reflectors (CRS) at different distances to radar.

	[29]	[20]	this work
Radar Type	FMCW MIMO	FMCW MIMO	FMCW MIMO
MIMO Array	16T16R	3T4R	3T4R
Working Frequency	72–80 GHz	60–64 GHz	60–64 GHz
Max. Range	2 m	1 m	1 m
Range Res.	0.02 m	0.07 m	0.1 m
Angular Res.	2.3°	2°	2°
Radar Performance	3D Imaging	3D Imaging + 20 hand gestures	3D Imaging + 15 single-hand gestures + 15 dual hand gestures

ciple throughout. Although the MUSIC and DBF modules exhibit relatively high computational complexity, their execution efficiency can be significantly improved through matrix operation optimization on embedded platforms. Notably, the frame trajectory smoothing algorithm utilizes moving average computation with minimal computational overhead, meeting real-time processing requirements.

Regarding power characteristics, the experimental TI IWR6843AOP radar SoC demonstrates a typical power consumption of 1.2 W at 60 GHz operation. As the current implementation relies on MATLAB platform, comprehensive system-level power analysis remains unavailable. Future work will focus on hardware implementation of the algorithm on FPGA/DSP embedded platforms to conduct systematic real-time performance testing and precise power consumption measurements.

3) This study adopts a signal processing-based trajectory reconstruction approach to provide more flexible underlying representations for gesture classification and integrated applications. The current solution preserves complete spatial motion information, offering rich feature foundations for subsequent classifier design, while its unsupervised characteristics effectively address cross-user adaptation challenges. Future work will employ multimodal data fusion to establish a hierarchical processing architecture that maintains real-time trajectory detection advantages: bottom-layer signal processing ensures robust motion tracking, while upper-layer few-shot learning enables efficient semantic understanding, ultimately forming a comprehensive radar-based gesture interaction solution.

4) This study presents an innovative millimeter-wave radar-based algorithm for dual-target motion trajectory perception. Through a meticulously designed bimanual interaction exper-

imental framework, we have successfully validated the algorithm's superior performance in three-dimensional spatial localization. At the current stage, we have focused on addressing key technical challenges in radar signal processing and motion trajectory reconstruction, with these fundamental breakthroughs establishing a solid technical foundation for subsequent application development. It is noteworthy that while the present research achievements remain in the theoretical verification and system optimization phase, with practical scenario validation yet to be conducted, these interim results have already demonstrated significant application potential across multiple important domains, including precision control for smart homes, human-computer interaction in virtual reality, and auxiliary diagnosis in medical rehabilitation. This progressive research paradigm, evolving from fundamental investigation to applied development, not only ensures the reliability of the technical solution but also provides substantial technical reserves for future industrial applications.

5. CONCLUSION

We propose a multi-target motion perception algorithm by using a MIMO FMCW radar based on an L-shaped virtual antenna array to realize dual-hand gesture recognition and trajectories tracking. The approach involves estimating azimuth and elevation angles separately from the two data components derived from the L-shaped array through MUSIC algorithm, incorporating spatial division techniques combined with DBF to address the challenges of dual-hand motion perception. To mitigate noise inherent in the raw frame-point data, a corresponding frame point removal and smoothing algorithm is devised to refine the trajectories. After the calibration of radar platform using corner reflectors, the experiments with three distinct scenarios, including unimanual, motion, bimanual mirror motion, and bimanual asynchronous motion, are carried out to validate the recognition of air-writing (digits, letters, and symbols) gestures and 3D trajectories detection. Therefore, compared with other similar works, the proposed method achieves superior range and angle resolution, particularly when employing a small array and notably under the L-shaped virtual MIMO array configuration. It effectively enables millimeter-wave wireless perception of dual-hand motion, yielding clear spatial trajectories of both single-handed and dual-handed gestures, and it gives a new option for controlling the HCI.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China (61801262), the Open Research Program of the State Key Laboratory of Millimeter Waves (K202312), and the Startup Foundation to Introduce Talent of NUIST (2022r071).

REFERENCES

[1] Li, G., A. E. Schultz, and T. A. Kuiken, "Quantifying pattern recognition — Based myoelectric control of multifunctional transradial prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 18, No. 2, 185–192, 2010.

[2] Li, Y., Q. Miao, K. Tian, Y. Fan, X. Xu, R. Li, and J. Song, "Large-scale gesture recognition with a fusion of RGB-D data based on the C3D model," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 25–30, Cancun, Mexico, 2016.

[3] Almasre, M. A. and H. Al-Nuaim, "Recognizing arabic sign language gestures using depth sensors and a KSVM classifier," in *2016 8th Computer Science and Electronic Engineering (CEEC)*, 146–151, Colchester, UK, 2016.

[4] Farina, D., N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 4, 797–809, 2014.

[5] Amsüss, S., P. M. Goebel, N. Jiang, B. Graimann, L. Paredes, and D. Farina, "Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control," *IEEE Transactions on Biomedical Engineering*, Vol. 61, No. 4, 1167–1176, 2014.

[6] Zhang, X., X. Chen, F. Farzadpour, and Y. Fang, "A visual distance approach for multicamera deployment with coverage optimization," *IEEE/ASME Transactions on Mechatronics*, Vol. 23, No. 3, 1007–1018, 2018.

[7] Song, Z., R. Chung, and X.-T. Zhang, "An accurate and robust strip-edge-based structured light means for shiny surface micromeasurement in 3-D," *IEEE Transactions on Industrial Electronics*, Vol. 60, No. 3, 1023–1032, 2013.

[8] May, S., B. Werner, H. Surmann, and K. Pervolz, "3D time-of-flight cameras for mobile robotics," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 790–795, Beijing, China, 2006.

[9] Fhager, L. O., S. Heunisch, H. Dahlberg, A. Evertsson, and L.-E. Wernersson, "Pulsed millimeter wave radar for hand gesture sensing and classification," *IEEE Sensors Letters*, Vol. 3, No. 12, 1–4, 2019.

[10] Heunisch, S., L. O. Fhager, and L.-E. Wernersson, "Millimeter-wave pulse radar scattering measurements on the human hand," *IEEE Antennas and Wireless Propagation Letters*, Vol. 18, No. 7, 1377–1380, 2019.

[11] Wang, F.-K., M.-C. Tang, Y.-C. Chiu, and T.-S. Horng, "Gesture sensing using retransmitted wireless communication signals based on Doppler radar technology," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 63, No. 12, 4592–4602, 2015.

[12] Fan, T., C. Ma, Z. Gu, Q. Lv, J. Chen, D. Ye, J. Huangfu, Y. Sun, C. Li, and L. Ran, "Wireless hand gesture recognition based on continuous-wave Doppler radar sensors," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 64, No. 11, 4012–4020, 2016.

[13] Sun, Y., T. Fei, X. Li, A. Warnecke, E. Warsitz, and N. Pohl, "Real-time radar-based gesture detection and recognition built in an edge-computing platform," *IEEE Sensors Journal*, Vol. 20, No. 18, 10 706–10 716, 2020.

[14] Zhang, Y., C. Zhu, S. Dong, Z. Gu, M. Balle, B. Zhang, C. Li, and L. Ran, "3-D motion imaging in a multipath coordinate space based on a TDM-MIMO radar sensor," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 68, No. 11, 4642–4651, 2020.

[15] Amin, M. G., Z. Zeng, and T. Shan, "Hand gesture recognition based on radar micro-Doppler signature envelopes," in *2019 IEEE Radar Conference (RadarConf)*, 1–6, Boston, MA, USA, 2019.

[16] Pramudita, A. A., *et al.*, "Contactless hand gesture sensor based on array of CW radar for human to machine interface," *IEEE Sensors Journal*, Vol. 21, No. 13, 15 196–15 208, 2021.

[17] Lien, J., N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Transactions on Graphics (TOG)*, Vol. 35, No. 4, 1–19, 2016.

[18] Wang, P., J. Lin, F. Wang, J. Xiu, Y. Lin, N. Yan, and H. Xu, "A gesture air-writing tracking method that uses 24 GHz SIMO radar SoC," *IEEE Access*, Vol. 8, 152 728–152 741, 2020.

[19] Gu, C., J. Wang, and J. Lien, "Motion sensing using radar: Gesture interaction and beyond," *IEEE Microwave Magazine*, Vol. 20, No. 8, 44–57, 2019.

[20] Li, Y., C. Gu, and J. Mao, "4-D gesture sensing using reconfigurable virtual array based on a 60-GHz FMCW MIMO radar sensor," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 70, No. 7, 3652–3665, 2022.

[21] Wang, Z., Z. Yu, X. Lou, B. Guo, and L. Chen, "Gesture-radar: A dual Doppler radar based system for robust recognition and quantitative profiling of human gestures," *IEEE Transactions on Human-Machine Systems*, Vol. 51, No. 1, 32–43, 2021.

[22] Leem, S. K., F. Khan, and S. H. Cho, "Detecting mid-air gestures for digit writing with radio sensors and a CNN," *IEEE Transactions on Instrumentation and Measurement*, Vol. 69, No. 4, 1066–1081, 2020.

[23] Khan, F., S. K. Leem, and S. H. Cho, "In-air continuous writing using UWB impulse radar sensors," *IEEE Access*, Vol. 8, 99 302–99 311, 2020.

[24] Wang, Z., G. Li, and L. Yang, "Dynamic hand gesture recognition based on micro-Doppler radar signatures using hidden Gauss-Markov models," *IEEE Geoscience and Remote Sensing Letters*, Vol. 18, No. 2, 291–295, 2021.

[25] Malysa, G., D. Wang, L. Netsch, and M. Ali, "Hidden Markov model-based gesture recognition with FMCW radar," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 1017–1021, Washington, DC, USA, 2016.

[26] Sakamoto, T., X. Gao, E. Yavari, A. Rahman, O. Boric-Lubecke, and V. M. Lubecke, "Hand gesture recognition using a radar echo I-Q plot and a convolutional neural network," *IEEE Sensors Letters*, Vol. 2, No. 3, 1–4, 2018.

[27] Sun, Y., T. Fei, S. Gao, and N. Pohl, "Automatic radar-based gesture detection and classification via a region-based deep convolutional neural network," in *ICASSP 2019 — 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4300–4304, Brighton, UK, 2019.

[28] Skaria, S., A. Al-Hourani, M. Lech, and R. J. Evans, "Hand-gesture recognition using two-antenna Doppler radar with deep convolutional neural networks," *IEEE Sensors Journal*, Vol. 19, No. 8, 3041–3048, 2019.

[29] Zhang, Z., Z. Tian, and M. Zhou, "Latern: Dynamic continuous hand gesture recognition using FMCW radar sensor," *IEEE Sensors Journal*, Vol. 18, No. 8, 3278–3289, 2018.

[30] Arsalan, M. and A. Santra, "Character recognition in air-writing based on network of radars for human-machine interface," *IEEE Sensors Journal*, Vol. 19, No. 19, 8855–8864, 2019.

[31] Du, H., Y. He, and T. Jin, "Transfer learning for human activities classification using micro-Doppler spectrograms," in *2018 IEEE International Conference on Computational Electromagnetics (ICCEM)*, 1–3, Chengdu, China, 2018.

[32] Yin, W., L.-F. Shi, and Y. Shi, "Indoor human action recognition based on millimeter-wave radar micro-Doppler signature," *Measurement*, Vol. 235, 114939, 2024.

[33] Jin, B., Y. Peng, X. Kuang, Z. Zhang, Z. Lian, and B. Wang, "Robust dynamic hand gesture recognition based on millimeter wave radar using Atten-TsNN," *IEEE Sensors Journal*, Vol. 22, No. 11, 10 861–10 869, 2022.