

## Spectrum-Based Hand Gesture Recognition Using Millimeter-Wave Radar Parameter Measurements

Liu, Changjiang; Li, Yuanhao; Ao, Dongyang; Tian, Haiyan

**DOI**

[10.1109/ACCESS.2019.2923122](https://doi.org/10.1109/ACCESS.2019.2923122)

**Publication date**

2019

**Document Version**

Final published version

**Published in**

IEEE Access

**Citation (APA)**

Liu, C., Li, Y., Ao, D., & Tian, H. (2019). Spectrum-Based Hand Gesture Recognition Using Millimeter-Wave Radar Parameter Measurements. *IEEE Access*, 7, 79147-79158. Article 8736874. <https://doi.org/10.1109/ACCESS.2019.2923122>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

# Spectrum-based Hand Gesture Recognition using Millimeter-wave Radar Parameter Measurements

Changjiang Liu<sup>1</sup>, Yuanhao Li<sup>2</sup>, Member, IEEE, Dongyang Ao<sup>3</sup>, and Haiyan Tian<sup>4</sup>

<sup>1</sup>School of Transportation Science and Engineering, Beihang University, Beijing, 100191 China

<sup>2</sup>Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands

<sup>3</sup>School of Information and Electronics, Beijing Institute of Technology, Beijing, 100081 China

<sup>4</sup>School of Computer Science and Engineering, Beihang University, Beijing, 100191 China

Corresponding author: Dongyang Ao (e-mail: aodongyang@foxmail.com).

**ABSTRACT** Radar sensors offer several advantages over optical sensors in the gesture recognition for remote control of electronic devices. In this paper, we investigated the feasibility of human gesture recognition using the spectra of radar measurement parameters. With the combination of radar theory and classification methods, we found that the frequencies of different gestures' parameters could be utilized as features for gesture recognition. Six kinds of periodic dynamic gestures are designed to avoid the complexity of defining and extracting the start and end of the dynamic gesture. In addition to the frequency ratio, we also extracted some features related to motion range and detection coherence to eliminate the interferences brought by the unintended gestures. The decision tree classifier designed on the basis of experimental phenomena can guarantee effective classification between different gestures, and in general, the correct recognition rate of each gesture is higher than 90%. Finally, we collected the position and the Doppler velocity information of hand for classification by a W-band millimeter wave radar in the experiment and verified the usability of the proposed method.

**INDEX TERMS** Gesture recognition, millimeter-wave, feature extraction, decision tree.

## I. INTRODUCTION

Since the beginning of the 21st century, traditional control methods based on interactive hardware such as mice and keyboards have gradually been replaced by touch controls. However, touch is not the only way to interact. Other media, like voices and gestures, are also convenient and attractive ways to complete interactions because they do not require any auxiliary equipment or contact. Hand gesture recognition pertains to recognizing meaningful expressions of motion by fingers/hands/arms. It is of utmost importance in designing an intelligent and efficient human-computer interface. The applications of gesture recognition are manifold, ranging from sign language through medical rehabilitation to virtual reality [1]. In particular, with the development of artificial intelligence technology in recent years, the application potentials of gesture recognition in intelligent buildings, smart homes and smart driving have been explored [2]. However, one of the major problems is to achieve the stability and simplicity of direct contact

control. Therefore, gesture recognition has been a hot topic in the remote control field.

Hardware devices and data processing algorithms are two important parts for gesture recognition. The device system can operate on different sensors, including optical/infrared cameras [3]-[5], optical pulse equipment [6], data gloves [7], ultrasound devices [8], and radar systems [9]. Although optical/infrared-based gesture recognition systems provide reliable recognition rates, the limitation is that the sensors can be easily affected by brightness. Fortunately, the radar system is insusceptible to ambient brightness and low in price. Radar-based gesture recognition has aroused public interests. However, the radar system also needs to consider the high computing performance and power resources of the devices.

According to the recent published literatures, echo signals of gestures captured by the radar system are explored to extract useful features for subsequent recognition. The differences among radar-based gesture

recognitions are mainly manifested in the following three aspects: the radar system, the method of feature extraction, and the method of recognition/classification. Yet the classification approach is important, differences between the systems and the feature extraction methods affect the robustness, the miniaturization, the cost reduction, and real-time processing of the sensor system to a greater extent.

In terms of the radar system, several industry products have used the millimeter wave band, the broadband waveform, and the multi-antenna to make the radar smaller in size and more accurate in parameter estimation. For example, Google's Project Soli works at 60 GHz, and uses multiple antennas to capture the fine movements of the fingers [10]-[12]. Engineers from NVIDIA developed their Frequency Modulated Continuous Wave (FMCW) monopulse radar that works at 24GHz. Its high range resolution and capability of two-dimensional (2D) angle measurement can obtain the target's 4D information [13]. Texas Instruments (TI) designed a 76-81GHz 3D radar, which can also be applied to automotive detection, pedestrian monitoring, etc. [14]. Compared with the industry products, the academic researches are more diverse in the architecture of system. Their working frequencies are from S band [15], [16] to terahertz [17], and the transmitted waveform includes continuous wave [18], pulse [19], FMCW [20] and so on. However, most of these radars do not have the angle measurement capability and works in the low frequency band which is not conducive to the miniaturization of the radar or the application of high-resolution waveforms.

In order to obtain more useful information, the researchers had explored the feature differences between gestures in diverse information dimensions. The most popular topic is extracting the micro-motion features of different gestures in the time-frequency domain [15], [18], [21]-[25], [26]. According to these papers, it has been proved that the classification based on micro-motion features is effective. In addition, features such as target location, speed and strength have also been studied [11], [20], [27]-[30].

An effective and practical classification algorithm is the last step to achieve gesture recognition. The choice of the classification algorithm depends mainly on the types of the extracted features. When the features are the raw signals, the Range-Doppler Maps (RDMs), or the time-frequency diagrams, neural networks are usually used as classifiers [31]-[37]. If the extracted features are several gesture-related parameters, the traditional Nearest Neighbor (NN) classifier, K-means classifier, or Support Vector Machine (SVM) are generally faster and more efficient [38]-[42].

Although great progresses have been made, two problems still exist. The first is how to define and detect the start and end of the gesture. Since the movement of the human hand is coherent in both time and space, some unintended and redundant actions are quite likely to

interfere with gesture recognition [30]. The second is the robustness of the algorithms. The system should have the capability to filter out the non-exact gestures while be more tolerant to the natural spatial and temporal variations of each user and the various gesture motions [11].

Actually, the application context of gesture recognition determines the type of gesture and the level of accuracy. For example, in vehicular applications, it is not quite necessary for radar to identify particularly subtle, rapid motions, while considerable stability and reliability of identification are required [11]. In this paper, we propose a scheme for gesture recognition based on the differences in spectral characteristics of radar measurement parameters. In order to avoid the effects of unintended and redundant actions, we define six kinds of periodic gestures as examples. While the hand periodically performs these dynamic gestures, the parameters measured by the radar, i.e., the slant range, the Doppler, and the azimuth angle, also exhibit periodic changes. This periodicity is related to both the parameter itself and the gesture. Therefore, we tried to extract the spectral features of the radar parameters for gesture classification and recognition. In the experiment, we used a millimeter-wave radar operating at W-band to measure the target parameters required in our applications. The results of gesture classification based on experimental data showed that the proposed algorithm could correctly recognize the designed gestures as well as eliminate those invalid ones, which verified the feasibility of our scheme.

The organization of this paper is as follows. Section II mainly introduces the radar system, the experimental setup and the radar signal processing scheme. Section III analyzes the process and performance of the radar feature extraction in detail, and then gives the specific method of classification. The experimental results are given and discussed in Section IV. Our conclusions are given in the last section.

## II. EXPERIMENT SETUP AND RADAR MEASUREMENTS

To match the application requirements for radar gesture recognition, we used a millimeter-wave radar designed with TI's AWR1642 chip, which operates at 77-81 GHz. Compared to 24 GHz, the use of W-band for these applications enables finer range resolution and velocity resolution, and also results in a smaller form factor for the antennas, which is a significant advantage [43]. In our experiment, the radar sensor outputs parameters of gestures such as the Doppler velocity for further extraction of features. Since the recognition strategy utilizes spectrum features of dynamic gestures, the gestures were designed to manifest obvious periodicity.

### A. RADAR SENSOR SYSTEM AND PLACEMENT

The transmitting signal bandwidth of this FMCW radar is 3.5 GHz, so as to achieve a distance resolution of about 0.04

meter. The one-dimension Multiple-Input Multiple-Output (MIMO) antennas, two for transmitting and four for receiving, provide an azimuth angle resolution of about  $15^\circ$ . The Doppler velocity of the target can be obtained by Fast Fourier Transform (FFT) at the slow time. Parameters of the radar are listed in Table I.

TABLE I  
RADAR PARAMETERS

| Parameter                   | Value           |
|-----------------------------|-----------------|
| Chirp bandwidth             | 3.5 GHz         |
| Antennas                    | 2TX, 4RX        |
| Azimuth Field of View (FOV) | $\pm 55$ degree |
| Elevation FOV               | $\pm 10$ degree |
| Range resolution            | 0.043 m         |
| Maximum velocity            | $\pm 10$ m/s    |
| Velocity resolution         | 0.3m/s          |
| Azimuth resolution          | 15 degree       |
| Data periodicity            | 100 ms          |

TX is short for transmitter and RX is short for receiver.

The radar is basically designed with TI's AWR1642-based automotive radar demo [43]. As shown in Fig.1, the basic structure consists of an antenna array, a Monolithic Microwave Integrated Circuit (MMIC) chip, a Power Management Integrated Circuit (PMIC) chip and a communication interface. The equipped Universal Asynchronous Receiver/Transmitter (UART) communication interface supports direct data interaction between radar and computer via Universal Serial Bus (USB). Besides, data interaction with the outside through Controller Area Network (CAN) is also available. The entire radar is quite small with a lateral dimension of approximately 13 cm. The radar consumes only 4 W. If the radar waveform parameters are further optimized, there is still a lot of room for the reduction of power consumption. The signal processing algorithm on the radar also adopted the processing scheme disclosed by TI. For more detailed signal processing methods and principles, please refer to the relevant introduction on the TI website [43] or related literatures [44], [45].

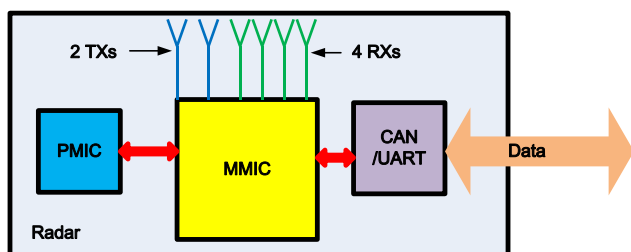


FIGURE 1. Radar system structure.

The experimental configuration for collecting gesture data is shown in Fig. 2. The radar with its beam vertically upward is placed flat on the table. A Cartesian coordinate system is established with the center of the radar antenna as the origin. The plane where the horizontal beam is located is defined as the  $x$ - $O$ - $y$  plane. During the experiment, the hand and the

forearm moved in the scope of the radar beam. The radar measured the slant range  $R$  of the scattering point and the azimuth angle  $\theta$  (defined as the angle between the radar line of sight and the  $y$ -axis in the  $x$ - $O$ - $y$  plane, and is positive in the first quadrant) and the Doppler velocity (is positive if the scattering point moves away from the radar). After the measurement process, the data was packaged and sent to the CAN interface card through the CAN, and then sent by the CAN interface card to the personal computer (PC) for data storage and subsequent processing.

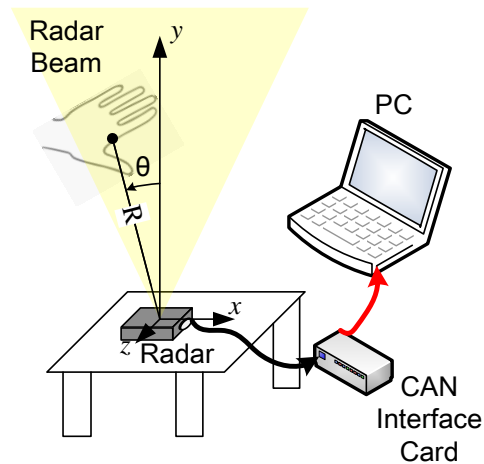


FIGURE 2. Experimental Configuration.

## B. GESTURE DESIGN

The main considerations when designing the gesture sets were the needs of the performer and characteristics of the system [11]. Our recognition model does not focus on tiny movements compared to traditional micro-Doppler-based models but on the periodicity of dynamic gestures. Therefore, we designed several periodic gestures with the large movements. As illustrated in Fig. 3, the six gestures are: 1) horizontal reciprocating motion with the  $y$ -axis as the axis of symmetry; 2) vertical reciprocating motion along the  $y$ -axis; 3) the arm remains stationary, the palm is swiftly swung up and down; 4) circular motion, the center of the circle is on the  $y$ -axis; 5) pendulum motion; 6) drawing Arabic numeral '8'. In the following paragraphs, the six gestures are assigned notations of G1-G6 for simplicity. The commonality of these six gestures is that when the motion of one cycle ends, it resumes the posture at the beginning. On the one hand, the gestures were designed to be distinguishable by the radar. On the other hand, they are memorable and easy to perform for the user.

It is worth noting that although we expect the motion of the gesture to be strictly in accordance with the ideal uniform acceleration/constant motion model and in accordance with the linear/circular motion, it is not guaranteed to meet the expectations. However, similar gestures are acceptable as long as they meet the requirements of periodicity and range. For example, the gesture of drawing a circle can be equivalent to drawing an ellipse, and the pendulum motion

can be either the upper semicircle or the lower semicircle. Those special cases will be analyzed in the experiments. By creating larger spatial zones for each gesture, and having gestures performed at varying speeds, non-exact gestures could be more acceptable, which decreases the need of visual attention of the user to the motion [11].

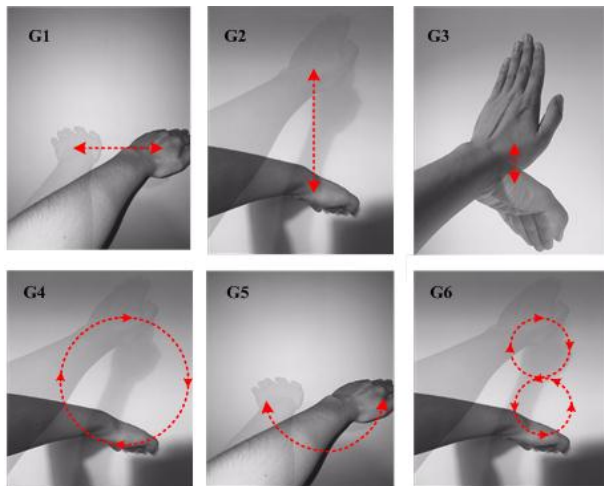


FIGURE 3. Illustration for designed dynamic hand gestures.

### C. MEASUREMENT DATA PREPROCESSING

The sensor periodically outputs parameter information of all detection points in the observation zone. Before feature extraction and gesture recognition, we need to pick out the gesture-related data from the recorded data. We first give the relevant parameter definitions. For each observation time (or frame), the parameters of the point cloud data collected by the radar are mainly the relative echo power  $P_k$ , the slant range  $R_k$ , the azimuth sine (sine value of the azimuth angle)  $A_k = \sin \theta_k$  and the Doppler velocity  $V_{r,k}$  of the  $k$ -th scattering point.

The first pre-processing work to be performed is to extract the scattering points of the hand from a plurality of scattering points. Since the close-range clutter in the scene is strong, even if the range resolution is relatively high, hands with zero Dopplers are difficult to distinguish from stationary targets. In another word, the Doppler velocity required for an effective detection should not be zero. In addition, to remove distant detections, the lateral distribution range of effective scattering points in the scene is limited to be  $|x_k| \leq 0.7m$ , and the longitudinal distribution range is  $0.1m \leq y_k \leq 1m$ , wherein the lateral coordinate  $x_k$  and the longitudinal coordinate  $y_k$  of the scattering point can be calculated according to  $R_k$  and  $A_k$ .

Secondly, the scattering center of the hand should be determined and the redundant data should be filtered out. Through frame-by-frame analysis, we found that due to the high range resolution of the radar, more than one scattering point of the moving hand/arm could be detected at certain time. Keeping in mind that a stronger power of a detected point in general means a higher signal-to-noise ratio and a

higher measurement accuracy, we selected the point with the strongest echo power among all the effective scattering points. Then, the parameters of this single point were considered as parameters of the dynamic gesture.

Finally, we need to assure the continuity of the data. When there is no effective scattering point (mostly the Doppler of the hand is 0 in this case), we set the slant range, the azimuth sine and the Doppler velocity to 0 values. Here, G1 is taken as the example. Fig. 4(a) shows the temporal curves of the radar parameters for two consecutive gesture periods of 2.7 s. It can be seen that there are multiple discontinuous points due to the loss of detection at zero Doppler. This kind of discontinuity is not conducive to the extraction of the spectrum and should be compensated. Therefore, the parameters of these missing detection points need to be reconstructed. There is no more processing on the Doppler velocity. We consider two possible cases for the slant range and azimuth sine. When the invalid detection point is at either ends of the signal, its parameters are set to the parameters of the nearest detected point. When there are effective detection points on both sides of the invalid detection point, the parameters of the two sides are used to linearly fit the missing parameters in the middle. The constructed parameters are shown in Fig. 4(b).

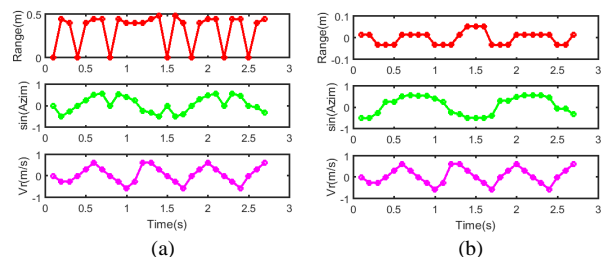


FIGURE 4. Curves of measured parameters of G1:(a) original results, (b) reconstructed result.

The whole preprocessing scheme is illustrated in Fig. 5. After the preprocessing, the raw data are transformed to a suitable format for feature extraction.

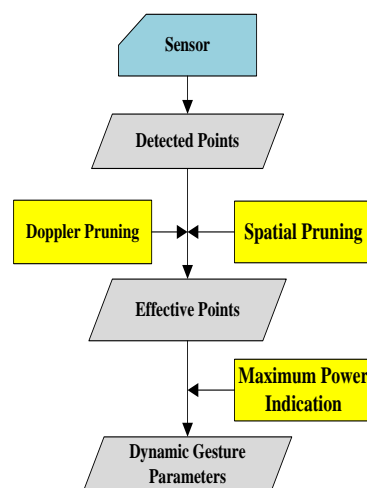


FIGURE 5. Measurement preprocessing scheme.



### III. FEATURE EXTRACTION AND CLASSIFICATION

With the processed measurement data, features will be extracted from the spectra, spatial distribution and continuities of the parameters. Different features play different roles in either the classification of different gestures or the recognition of valid gestures. We designed a classifier based on the decision tree. The decision tree can effectively utilize the extracted features and is quite convenient to implement [46]. This classification strategy is also designed to eliminate gestures that do not meet the requirements.

#### A. FEATURE EXTRACTION

##### 1) Frequencies of parameters

To extract useful features for the gesture classification, we need to analysis the characteristics of gesture radar parameters in the spectrum domain, the spatial domain, and the time domain. Firstly, several frames of continuous data were extracted. Then, a spectrum analysis of the continuous parameters was implemented based on the Fourier transform. Fig. 6(a) shows the spectra of parameters in Fig. 4(b). We can see that the frequencies of range and Doppler velocity are both twice that of the azimuth sine. In the subsequent processing, the peak positions in the spectrum of the slant range, the azimuth sine and the Doppler velocity are noted by  $f_r$ ,  $f_a$ , and  $f_v$  respectively and the unit is Hz. These peak positions are defined as the parameter frequencies.

To verify the correctness of our model, the results of the measured data will be compared with the theoretical model. We suppose that the hand moves horizontally (parallel to the  $x$ -axis) in the  $x$ - $O$ - $y$  plane. The center height of the motion is 0.4 m, the horizontal range is  $-0.3 - +0.3$  m, the period of one round-trip motion is 2 s, and the sampling interval is 0.1 s. When the hand moves from  $x=-0.3$  m to  $x=0$  m, an uniform acceleration motion is performed, and during the movement from  $x=0$  m to  $x=0.3$  m, the motion is uniformly decelerated. The absolute value of acceleration, which is  $2.4 \text{ m/s}^2$ , can be calculated via simple equation of motion. The right-to-left motion for the other half cycle is the replica of the first half of the cycle. Therefore, the position of the hand can be expressed by (unit of  $t$  is s):

$$x = \begin{cases} -0.3 + 1.2t^2 & t \in [0, 0.5] \\ 1.2(t-0.5) - 1.2(t-0.5)^2 & t \in (0.5, 1] \\ 0.3 - 1.2(t-1)^2 & t \in (1, 1.5] \\ -1.2(t-1.5) + 1.2(t-1.5)^2 & t \in (1.5, 2] \end{cases} \quad (1)$$

$$y = 0.4 \quad t \in [0, 2]$$

The simulated radar measurements including range, azimuth sine, and Doppler velocity can be calculated according to (1) and are depicted in Fig. 6(c). It can be seen that those curves quite resemble the measured curves shown in Fig. 4 regardless of the difference between periods. Then, we can give the spectra of the calculated radar parameters as shown in Fig. 6(b). The frequency of azimuth sine is the same as that of the gesture, while the frequencies of the slant

range and the Doppler velocity are twice the azimuth sine/gesture, which is consistent with the experimental results in Fig. 6(a). That is to say, although there are many differences in the speed and position of different hand movements, the proportional relationship between the frequencies of radar parameters is quite stable, which can be used as a potential feature for classifying different gestures.

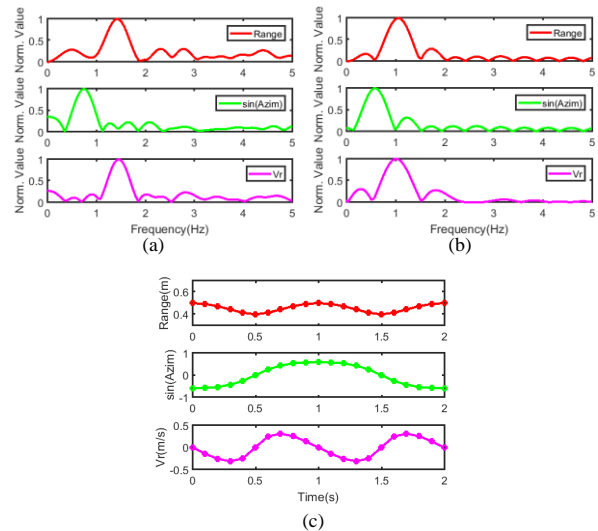


FIGURE 6. Measured results of G1:(a) experimental parameter spectra, (b) simulated parameter spectra, (c) curves of simulated parameters.

Further, we consider a common redundant gesture in which the hand moves in one direction but does not return during an observation period. In this case, although the data of the complete cycle is not collected, the parameter frequency estimation result is consistent with that of the designed round-trip motion. At this time, some one-way hand motions might be recognized as G1 or G2 relying solely on the frequency relationship, which is not our expectation. To solve this problem, one solution is to ensure that a complete round-trip period data is collected as much as possible. For example, for the motion modeling result of G1, under the condition of ensuring that at least one round-trip period of data is collected within 2 s (similar to [33]), the motion frequency of the hand is required to be greater than 0.5 Hz.

##### 2) Contrasts of spectra

Based on the above analysis, the first category of feature we extracted from radar data is the frequency of the radar parameter. However, as long as there is enough detection, we can always get the spectrum of the parameter, even if it is seriously interfered by certain factors. In order to ensure that the extracted spectral peaks are obvious and meaningful, the normalized contrast, which is usually used to check if the radar image is correctly focused, is applied here to define the validity of the spectrum. The normalized contrast (noted as  $C$ ) is calculated as [47]:

$$C = \frac{\sqrt{E\left\{\left[|S| - E(|S|)\right]^2\right\}}}{E(|S|)}, |S| = |S_f(n)|^2 / \left[\sum_{n=1}^N |S_f(n)|^2\right] \quad (2)$$

where  $S_f(n)$  is the parameter spectrum and  $E(\cdot)$  is the mean calculation. In the subsequent processing, the extracted peak position of the spectrum is considered to be valid only when  $C$  exceeds the threshold. Meanwhile, the normalized contrasts of the slant range, the azimuth sine, and the Doppler velocity are noted as  $C_r$ ,  $C_a$ , and  $C_v$ , respectively.

### 3) Spatial dispersion

Moreover, the spatial ranges of the six gestures specified above are different: the spatial range of the G1 in the lateral direction is significantly larger than that in the longitudinal direction, while G2 is exactly the opposite of G1. The spatial ranges of G3 in the longitudinal and lateral directions are both narrow. G4, G5, and G6 all have wide ranges of motion in either the longitudinal or the lateral direction. Although the recognition based solely on range of motion is insufficient, it can aid in recognition and help eliminate some unintended gestures. Thus, the differences between the maximum value and the minimum value of the  $x/y$  coordinates in a parameter waveform, defined as  $\Delta x$  and  $\Delta y$  respectively, are also calculated as potential features for gesture recognition.

### 4) Gesture validity

In addition to those features used to distinguish between different gestures, some features for distinguishing between valid gestures and invalid gestures are also critical in gesture recognition. Like  $\Delta x$  and  $\Delta y$  defined above, these features can help avoid identifying messy, irregular or redundant gestures as prescribed gestures. Based on this consideration, we define two parameters for gesture validation. Firstly, we give the parameters that can reflect the validity of the radar data, one of which is the ratio of valid detection in an analysis period, noted as  $K$ ,

$$K = N_d / N_a \quad (3)$$

where  $N_a$  is the number of frames in one observation period, and  $N_d$  is the number of frames with valid gesture detection for this observation period. Considering that the Doppler velocities are 0s in some positions during the hand's movement, we set the minimum acceptable value of  $K$  to 0.4 in the subsequent processing. Secondly, we denote the maximum number of frames of consecutive invalid detections throughout the analysis period as  $Q$ . Actually, when there are too many consecutive invalid detections, it means that the radar is unable to capture a complete gesture in this period. One possible reason is that the radial movement of the hand is too slow. On the one hand, the consecutive invalid detection does not meet the definition of continuous motion of the gesture. On the other hand, the large parameter reconstruction error may occur if this kind of data is used for feature extraction. Under the condition of  $N_a = 20$ , we believe that the data is valid only when the maximum number of consecutive invalid detections does not exceed 5, that is,  $Q \leq 5$ .

## 5) Summary

In summary, the extracted parameters for gesture recognition and classification are listed in Table II and the flowchart for parameter extraction is shown in Fig. 7.

TABLE II  
EXTRACTED PARAMETERS

| Parameter                            | Value                | Source                    |
|--------------------------------------|----------------------|---------------------------|
| Parameter frequency                  | $f_r, f_a, f_v$      | Parameter spectrum        |
| Normalized contrast of spectrum      | $C_r, C_a, C_v$      | Parameter spectrum        |
| Coordinate range                     | $\Delta x, \Delta y$ | Trajectory sequence       |
| Ratio of valid detection             | $K$                  | Doppler velocity sequence |
| Maximum number of invalid detections | $Q$                  | Doppler velocity sequence |

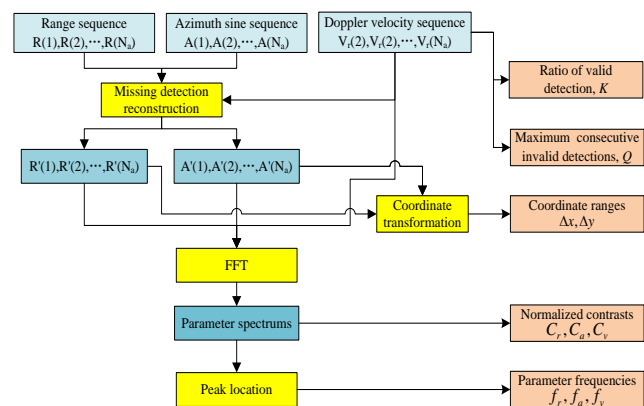


FIGURE 7. Flowchart of feature parameter extraction.

## B. ANALYSIS OF SPECTRUM FEATURES

As highlighted in the previous section, there is certain relationships between the frequencies of different parameters. And this kind of relationships has the potential for classification of gestures. In this section, we will statistically analyze the features based on a large number of measured data in order to find specific application criteria for these features in classification and recognition. During the process of collecting the measured data of each gesture, the experimenter continuously repeated a certain gesture according to its set standard, and the duration was about 1 to 2 minutes. The time length for a single sample sequence is 2 s. The collected radar parameter sequences are processed with a sliding window. The interval between adjacent sample sequences is 0.1 s, so there is a certain overlap. Considering that there is no need to determine the start or end of the gesture, this way of sliding window processing is feasible and effective.

Firstly, the extracted parameter frequencies of G1 are given in Fig. 8(a). A total of 869 samples were collected. It can be seen that the azimuth sine frequency and the Doppler velocity frequency, which are about 0.5 Hz and 1 Hz respectively, are relatively stable. However, there are some outliers in the frequency of the slant range, which is mainly because the slant range changes little during the observation

period and errors (such as the measurement error, the scattering center extraction error, and the reconstruction error) could have greater impacts on the range frequency estimation performance. Similarly, the normalized contrasts of azimuth sine and Doppler velocity shown in Fig. 8(b) are also larger than that of the slant range. The reason why the Doppler velocity contrast is not quite high is that the small Doppler variation range owing to the hand's laterally movement made the estimation of  $f_v$  susceptible to errors. Fig. 8(c) shows three frequency ratios of  $f_r/f_a$ ,  $f_r/f_v$ , and  $f_a/f_v$ . The unit of the frequency ratio is decibel (calculated by  $10\log_{10}$ ). Since the extraction results of  $f_a$  and  $f_v$  are more stable, the  $f_a/f_v$  is stably distributed around -3 dB. In contrast,  $f_r/f_a$  and  $f_r/f_v$  have some fluctuations outside the mean of their distribution due to the estimation error of  $f_r$ . The overall distribution of all three frequency ratios are consistent with the motion modeling results given earlier.

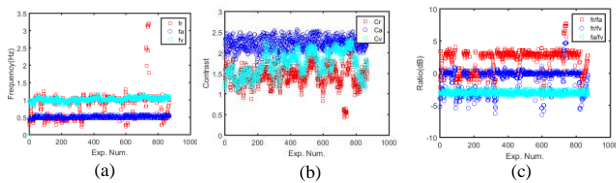


FIGURE 8. Feature parameter extraction result of G1:(a) parameter frequencies, (b) normalized contrasts, (c) frequency ratios.

Next, we will analyze the features of G2-G6 based on the measured frequency extraction results shown in Fig. 9. The motion characteristics of G2 determine that its azimuth sine frequency is mostly distributed near 0, while the velocity and the slant range have the same periodicity, so there is only a significant proportional relationship between  $f_r$  and  $f_v$ . G3 is a special case in these six gestures. Its motion period is short and the motion amplitude is small. Only the phase-sensitive Doppler velocity exhibits obvious periodicity. Theoretically, the slant range should also have a certain periodicity, but due to the instability of the motion, etc., the slant range frequency was only well extracted in the first half of the experiment. In particular, due to the gesture deformation caused by the fatigue of the experimenter, a period of unexpected Doppler velocity parameter appeared in the middle of the experimental results. Relatively, the swing ranges of G4, G5, and G6 are large and the fluctuations of the extracted parameters are also strong, which greatly facilitate the extraction of frequencies of radar parameters. The period of the three parameters of G4 is consistent, while the parameter frequency ratios of G5 is similar to those of G1. Though G6 is the most complex gesture in the gesture set, its parameter frequency ratios which show significant differences from the frequency ratios of other gestures were extracted steadily.

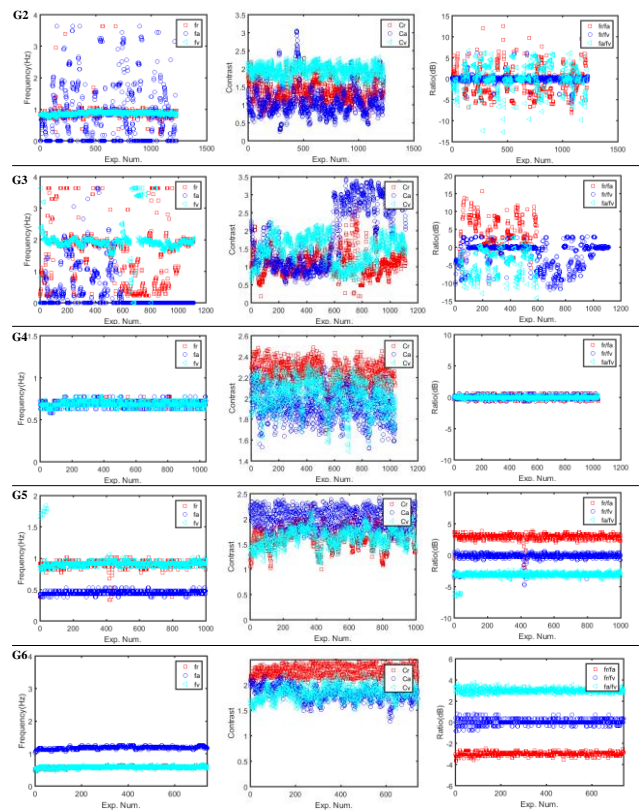


FIGURE 9. Feature parameter extraction results of G2-G6:(left column) parameter frequencies, (middle column) normalized contrasts, (right column) frequency ratios.

The major drawback of the application of the parameter frequency in classification is the dependence on the operator. Alternatively, the frequency ratio obtained from parameter frequency only depends on the gesture itself. Ideally, the ratio between the parameter frequencies of different gestures may be 1, 2 or 0.5. Based on the above analyses, we found that for different gestures, the frequency ratios are usually different. Therefore, the frequency ratio allows distinguishing between two gestures for most cases. For the rest cases, features such as the parameter frequency itself or the spatial range could be used as an aid.

### C. GESTURE CLASSIFICATION

We list the requirements for each parameter to identify various gestures in Table III. Here, a reliable parameter frequency extraction value corresponds to a normalized contrast greater than 0.8 and the estimation error of parameter frequency ratio is limited to less than 1 dB. Considering that the extractions of the slant range frequencies of G1 and G3 are not stable, the frequency ratios associated with slant range are not used for the classification. The requirements for the range of  $x$  and  $y$  are roughly set to satisfy most of the experimental data. In practical application, the range of motion should be given as a gesture requirement.



TABLE III  
FEATURE REQUIREMENTS FOR GESTURE CLASSIFICATION AND RECOGNITION

| Gesture | $f_r/f_a$<br>(dB) | $f_r/f_v$<br>(dB) | $f_a/f_v$<br>(dB) | $\Delta x$<br>(m) | $\Delta y$<br>(m) | Extra       |
|---------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------|
| G1      | -                 | -                 | $-3 \pm 1$        | $>0.4$            | $<0.15$           | -           |
| G2      | -                 | $0 \pm 1$         | -                 | $<0.15$           | $>0.3$            | -           |
| G3      | -                 | -                 | -                 | $<0.15$           | $<0.2$            | $f_v > 1.5$ |
| G4      | $0 \pm 1$         | $0 \pm 1$         | $0 \pm 1$         | $>0.3$            | $>0.3$            | -           |
| G5      | $3 \pm 1$         | $0 \pm 1$         | $-3 \pm 1$        | $>0.4$            | $>0.2$            | -           |
| G6      | $-3 \pm 1$        | $0 \pm 1$         | $3 \pm 1$         | $>0.3$            | $>0.3$            | -           |

Based on the requirements listed in Table III, it is not difficult to give a strategy for identifying these six gestures. A simple method is to first determine whether the conditions of each class are satisfied. If the conditions of all categories are not met, then the data is determined to be invalid. However, this method cannot directly reflect the mutual exclusion between condition sets of different gestures. In another word, a set of features cannot be attributed to two or more gestures at the same time. So here we utilized a more complex form, the decision tree, to achieve the validity discrimination and classification of gestures. The decision tree is a tree structure in which each internal node represents a test on an attribute, each branch represents a test output, and each leaf node represents a category. The decision tree had already been applied to camera-based hand gesture recognition [46], [48]. In general, the classifier is obtained based on learning. For our case it is quite intuitive and we directly give the decision tree classifier applied to this scheme, as shown in Fig. 10.

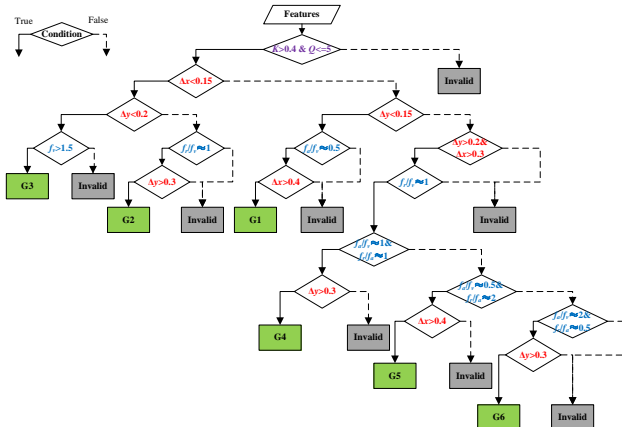


FIGURE 10. Decision tree classifier schematic.

In Fig. 10, in order to simplify the representation, we set the direction when the decision node condition is satisfied to the left, which is indicated by a solid line. Similarly, the direction when the condition is not satisfied is set to the right and is indicated by a broken line. First, it is necessary to check whether the given feature parameter set satisfies the requirements of the data validity parameters  $K$  and  $Q$ . According to the experimental data, the valid condition of the feature data is set to  $K > 0.4$  and  $Q \leq 5$ . When the valid

conditions are satisfied, we need to judge whether the requirements for coordinate ranges, the parameter frequencies and their ratios are satisfied. The data is finally discriminated as a certain type of gesture or invalid data. It can be seen that for each gesture, only one process can reach the leaf node it represents. That is to say, in the classification, the misjudgment between different gestures rarely occur, which greatly improves the stability in practical operation.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

##### A. TESTS OF DESIGNED GESTURES

Firstly, we test the recognition performance of the designed six gestures. All the collected experimental data shown in Fig. 9 were classified and identified based on the decision tree classifier. The results are shown in Fig. 11. In view of the characteristics of the classifier and the fact that there is no misclassification in the experimental results, only the probability of correct recognition is shown here. For ease of display, the correctly identified sample result is set to 0 in Fig. 11, while the invalid result is set to 1. The number of tested samples and the correct recognition probability for each gesture were counted as listed in Table IV. It can be seen that except for the recognition rates of G4 and G6 being 100%, the other four gestures are not fully recognized.

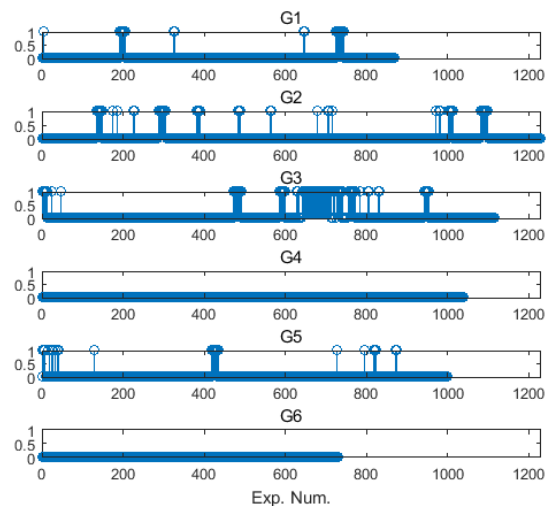


FIGURE 11. Recognition results for six gestures.

TABLE IV  
STATISTICS OF RECOGNITION RATES FOR DESIGNED GESTURES

| Gesture | Sample Number | Correct Recognition Rate |
|---------|---------------|--------------------------|
| G1      | 869           | 94.94%                   |
| G2      | 1129          | 92.11%                   |
| G3      | 1114          | 83.12%                   |
| G4      | 1039          | 100%                     |
| G5      | 999           | 96.0%                    |
| G6      | 730           | 100%                     |

By analyzing the performances of features in the classifier, we tried to find out the reason for each unrecognized case. For G1, its small Doppler velocity caused missed detection frequently, resulting in failure to meet the requirements for  $K$  or  $Q$ . Also, there are some moments when the  $y$ -direction motion range exceeds the threshold due to the non-ideal horizontal motion of the hand. For G2, the Doppler velocity is higher, so there are not too many invalid detections. However, since the Doppler velocities at the rising apex and at the falling valley are 0s, the according slant ranges can only be obtained by fitting. As a result, extraction of slant range frequency might be affected (Fig. 9) and hence the rate of correct recognition was decreased. For G3, due to the fact that the experimenter did not maintain the proper swinging speed of the hand during certain periods of data acquisition, continuous missed detection and incorrect estimation of Doppler velocity frequency occurred at that time, whereas the recognition rate was quite high during the rest of the time. For G5, it has invalid recognition results at some moments. After having looked up the original data, we found the main cause of the invalid recognition at these moments is that too many missed detections owing to small Doppler velocities led to errors in the extraction of slant range frequency and Doppler velocity frequency, as shown in Fig. 9.

### B. COMPARATIVE ANALYSIS OF ALOGRITHM PERFORMANCE

To further analyze the performance of the proposed method, we collected much more radar data from three other experimenters and the new data set of each gesture includes more than 3000 samples. In order to verify the recognition performance of our scheme for common gestures with irregular motions, more than 5000 samples of randomly moving and unintended gestures were collected as the data set of invalid gesture.

For gesture recognition based on parameter features, commonly used classifiers include NN, K-means, SVM, and the random forests, etc. [38]-[42]. Since the SVM classifier is more suitable for processing high-dimensional feature data [24], [25], [40], we also trained an SVM classifier for comparative analysis. The SVM classifier was trained based on a training set of 500 samples which were randomly selected from the whole data sets. The LIBSVM toolbox was used [49]. It should be mentioned that a sample is used for the training only if it satisfies the gesture validity criteria of  $K$  and  $Q$ . Six designed gestures and invalid gesture constituted the seven categories. After the training, a total of 644 supported vectors were obtained for the classification.

To evaluate the recognition performance, we defined three rates which are the correct recognition rate (R1), the misclassification rate (R2), and the recognition rate of invalid gestures (R3). R1 is the ratio of the number of correctly classified samples to total number of samples. R2 is the ratio of the number of the misclassified samples to the total number of samples. R3 is the ratio of the number of samples

classified as invalid gestures to the total number of samples. We noted that for G1-G6, when a sample satisfied the requirements of  $K$  and  $Q$  but was classified to invalid gesture in the SVM, it will not be considered as a misclassification. In particular, R1 is equal to R3 for those unintended gestures. The statistics of the recognition results of the experimental data are listed in Table V.

From Table V we found that for R1 of G1-G6, the difference between the SVM classifier and the decision tree classifier is not obvious. The soft classification boundaries of the SVM classifier sometimes may achieve a better recognition performance. However, values of R2 and R3 of G1-G6 indicate that it is difficult to avoid misclassification using the SVM classifier, while there are almost no misclassification problems using the proposed decision tree classifier. For certain specific application scenarios such as vehicular human-computer interaction, a lower misclassification rate may be preferred rather than a higher recognition rate. Another advantage of decision tree classifier over the SVM classifier is its high recognition rate for erratic gestures (99.5% vs. 93.1%), which can help reduce the interferences from unintended gestures as much as possible.

TABLE V  
COMPARISON OF RECOGNITION RATES BETWEEN OUR METHOD AND SVM

| Gesture | Sample Number | SVM   |      |      | Decision Tree |      |      |
|---------|---------------|-------|------|------|---------------|------|------|
|         |               | R1    | R2   | R3   | R1            | R2   | R3   |
| G1      | 3495          | 93.59 | 5.12 | 1.29 | 94.59         | 0.23 | 5.18 |
| G2      | 3560          | 98.96 | 0.34 | 0.70 | 97.11         | 0    | 2.89 |
| G3      | 3045          | 99.87 | 0    | 0.13 | 99.31         | 0    | 0.69 |
| G4      | 3880          | 95.72 | 3.25 | 0.95 | 97.45         | 0    | 2.55 |
| G5      | 3475          | 99.19 | 0.23 | 0.58 | 98.13         | 0    | 1.87 |
| G6      | 3730          | 90.70 | 1.39 | 0.21 | 97.64         | 0    | 2.36 |
| Random  | 5469          | 93.11 | 6.89 | -    | 99.51         | 0.49 | -    |

The unit of R1/R2/R3 is %.

Using the radar parameter data collected by the same type of radar chip, a deep-learning based gesture recognition was proposed by TI [44]. Another difference is that the temporal features were extracted in [44] instead of the spectrum features in our scheme. The tested classification rates of seven gestures, range from 95% to 99%, are similar to those of our method. However, its misclassification rates are much higher than those of our method.

### C. REAL-TIME TEST

In addition, we tested the real-time performance of the proposed method. The processing cycle is divided into two segments: feature extraction and classification. The feature extraction segment includes the preprocessing, the spectrum analysis, and the coordinate analysis, etc. We imported the collected radar data into Infineon's Aurix<sup>TM</sup> TC397 high-performance processing chip for real-time testing. The TC397 is a commonly used processing chip for automotive electronics systems with a main frequency of 300MHz.

According to the experimental results, the maximum processing times of different segments and their sum are listed in Table VI. The feature extraction processing time is 0.26 ms, of which the FFT operation and contrast calculation account for most of the time. The classification of the decision-tree-based method takes less 0.01 ms, which is almost negligible. Experimental results show that real-time processing of the proposed method is completely feasible. In contrast, the time consume of the SVM classification is 3.62 ms. If other processing functions need to be implemented on the same chip, the classification time of this magnitude might be an obstacle to the overall real-time performance. Moreover, the classification time of the deep-learning-based scheme has been tested in [44] with the radar chip AWR1642 and the result is about 1 ms. Considering that the processing ability of AWR1642 is stronger than that of TC397, our classification algorithm obviously has better real-time performance.

TABLE VI  
STATISTICS RESULTS OF THE REAL-TIME TEST

|                    | Decision Tree | SVM     |
|--------------------|---------------|---------|
| Feature Extraction | 0.26 ms       |         |
| Classification     | 0.01 ms       | 3.62 ms |
| Total              | 0.27 ms       | 3.88 ms |

#### D. ROBUSTNESS EVALUATION

Apart from the six gestures shown in Fig. 3, we also tested the classification performance of four non-standard periodic gestures, referred to as G7 to G10. As shown in Fig. 12, G7 is the rotation of G5 by 90°. G8 is to draw a triangle. G9 is to draw a square. G10 is the rotation of G5 by 180°. The statistics of classification rates for these four non-standard gestures are listed in Table VII. There are certain probabilities that G7, G8, and G9 may be recognized as G4, which means that periodic but non-standard gestures could be recognized as designed gestures. Since G9 has the highest probability and its similarity to the G4 is also the highest, we can infer that the corresponding probabilities basically depend on the similarity between non-standard gestures and standard gestures. The azimuth of G6 might be symmetrically distributed with respect to 0, in which case it could be recognized as G7. A special case is G10. It is totally classified to G5 owing to its mirror symmetry with respect to G5.

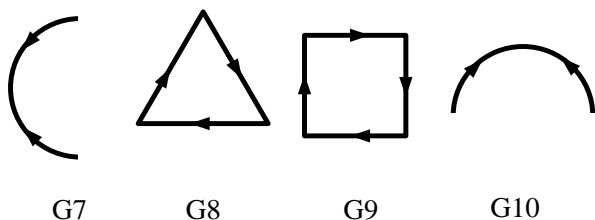


FIGURE 12. Illustration for non-standard gestures: G7-G10.

Though it is not guaranteed that all non-standard gestures with similar or equivalent features to the preset gestures are correctly identified, an important premise is that these non-standard gestures also have periodicity, which is not the usual case.

TABLE VII  
STATISTICS OF CLASSIFICATION RATES FOR NON-STANDARD GESTURES

| Gesture | Sample Number | G0    | G1 | G2 | G3 | G4    | G5    | G6    |
|---------|---------------|-------|----|----|----|-------|-------|-------|
| G7      | 1030          | 81.75 | 0  | 0  | 0  | 3.49  | 0     | 14.76 |
| G8      | 1110          | 48.47 | 0  | 0  | 0  | 51.53 | 0     | 0     |
| G9      | 1180          | 34.91 | 0  | 0  | 0  | 65.08 | 0.1   | 0     |
| G10     | 1030          | 0.19  | 0  | 0  | 0  | 0     | 99.81 | 0     |

G0 indicates invalid gesture. The unit of the classification rate is %.

#### E. DISCUSSIONS

The sensor system should accurately identify the correct gestures under the condition that actual operation might be different from standard gesture definition. Also, it needs to effectively eliminate those unintended gestures and the interference from other parts of the body. The above two points are often difficult to balance. The experimental results verified the feasibility of our method from three aspects. Firstly, if a well-performed gesture belongs to one of the designed gestures, it is recognized and classified as the correct gesture at a fairly high probability, and it is almost impossible to misclassify it as other gestures. Secondly, if a designed gesture is not well performed with respect to its trajectory rather than to its periodicity, there is still a certain probability that it can be correctly identified, which illustrates the robustness of the method to non-standard gesture conditions. Thirdly, when the collected radar data come from unintended gestures with irregular motions, we can basically guarantee that they will be judged as invalid gestures.

By qualitatively and quantitatively comparing the proposed method with methods in the existing literatures, it can be seen that the advantages of our method are mainly manifested in the following five points. First of all, compared with the gesture recognition based on trajectory or time-frequency distribution, our method does not need to extract the start and end points of the gesture, which reduces the complexity of processing. Secondly, only a few literatures considered the recognition performance for interference gestures [30], [39], [41]. When designing the classification criteria, we have considered quite a lot of invalid gestures, such as those with incomplete cycle, slow motion, and range not meeting the requirements. The experimental analysis results also show that the algorithm can effectively discriminate those gestures that do not meet the predetermined requirements. Thirdly, our recognition model does not focus on tiny movements compared to traditional micro-Doppler-based models that could differ from person to person. Also, the features based on the frequency ratio can also weaken the influence of the differences in the speeds, positions and frequencies of different persons' gestures. Fourthly, the decision tree classifier we designed can greatly

reduce the possibility of misclassification. This advantage is not available in many gesture recognition techniques, making our method play a better role in some special scenarios. Finally, the proposed method is quite simple and fully meets the requirements of real-time processing.

It is validated the system and algorithm can achieve their goals. But there will be limitations under the cases of non-standard gestures and system's limited detection ability for the motion near zero Doppler will also affect the its performance. This maybe be improved by using more complex algorithms in the reconstruction of parameter sequences. Besides, the radar system has much room for improvement in Doppler resolution and data rate. With proper system configuration, these two parameters can be optimized to achieve a better effect on motion detection while suppressing static clutter.

Future research may consider improvements in the bandwidth, the data rate, the Doppler resolution, and the applications of other classification algorithms, to identify smaller, faster actions as finger movements. Since gesture recognition based on the Artificial Neural Network (ANN) and temporal parameter sequences has been proven to be effective [44], we also expect the spectral parameter sequences to be used in neural-network-based gesture recognition, and attempts to compare the classification performances of different network structures. It is likely that this kind of multi-dimensional, high-resolution micro-radar will become the mainstream system for future gesture recognition, as its precise positioning and measurement capability provide enough information for identification without the need for complex micro-motion analysis.

## V. CONCLUSION

In this paper, we studied the feasibility of using the spectra of radar measurement parameters for dynamic gesture recognition and proposed a gesture recognition method based on the spectral characteristics of radar parameters and the decision tree. We defined six periodic motion gestures and used a millimeter-wave radar to capture the slant range, the azimuth angle, and the Doppler velocity parameters for each gesture from the radar echo. Our analyses indicate that, ideally, the ratio between the parameter frequencies of different gestures may be 1, 2 or 0.5. Based on this phenomenon, we extract the peak of the spectrum of each parameter as the main feature for gesture recognition, and combine the parameters such as the detection continuity, the motion range and the spectral contrast as auxiliary features for gesture recognition. Based on the designed decision tree classifier, we correctly classified the measured gestures. The data processing results show that if the normal motion is considered, the effective rate of the gesture recognition strategy can reach more than 90%. This spectrum-feature-based gesture recognition strategy can effectively avoid the influence of unintended gestures, and its processing becomes quite convenient without the need to extract the start and end

of the gesture. With the above advantages, it can be applied to certain occasions, such as driver gesture recognition in the car. In the future development, the physiological state and behavior of the driver can also be monitored while realizing the radar gesture recognition [43]. In the follow-up study, we expect to reconfigure the radar system, optimize the parameters and consider the combination of trajectories and spectral details for recognition.

## ACKNOWLEDGMENT

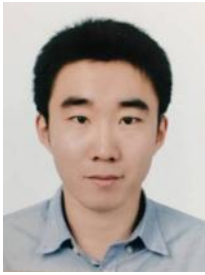
We would like to express our thanks to Mr. Kaisheng Liu from Beijing Information Science and Technology University for his contributes in the experiments.

## REFERENCES

- [1]. S. Mitra, and T. Acharya, "Gesture recognition: a survey," *IEEE Trans. Syst. Man & Cybern. Part C*, vol. 37, no. 3, pp. 311-324, May 2007.
- [2]. J. Francis, and B. K. Anoop, "Significance of hand gesture recognition systems in vehicular automation-a survey," *Int. J. Comput. Appl.*, vol. 99, no.7, pp. 50-55, Aug. 2014.
- [3]. R. R. Igorevich, P. Park, D. Min, Y. Park, J. Choi, and E. Choi, "Hand gesture recognition algorithm based on grayscale histogram of the image," in *Proc. IEEE Int. Conf. Appl. Informat. Commun. Techn., Uzbekistan*, Oct. 2010, pp. 1-4.
- [4]. H. Cheng, Z. Dai, and Z. Liu, "Image-to-class dynamic time warping for 3D hand gesture recognition," in *Proc. IEEE Int. Conf. Multimed. Expo (ICME)*, San Jose, CA, USA, Jul. 2013, pp. 1-6.
- [5]. C. C. Hsieh, D. H. Liou, and D. Lee, "A real time hand gesture recognition system using motion history image," in *Proc. IEEE Int. Conf. Signal Process. Syst.*, Dalian, China, Jul. 2010, pp. 394-398.
- [6]. S. Oprisescu, C. Rasche, and B. Su, "Automatic static hand gesture recognition using ToF cameras," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Bucharest, Romania, Aug. 2012, pp. 2748-2751.
- [7]. J. H. Kim, N. D. Thang, and T. S. Kim, "3-D hand motion tracking and gesture recognition using a data glove," in *Proc. IEEE Int. Symp. Indus. Electron.*, Seoul, South Korea, Jul. 2009, pp. 1013-1018.
- [8]. Y. Sang, L. Shi, and Y. Liu, "Micro hand gesture recognition system using ultrasonic active sensing," *IEEE Access*, vol. 6, pp. 49339-49347, Sep. 2018.
- [9]. C. Gu, and J. Lien, "A two-tone radar sensor for concurrent detection of absolute distance and relative movement for gesture sensing," *IEEE Sens. Lett.*, vol.1, no. 3, pp.1-4, Jun. 2017.
- [10]. Z. Flintoff, B. Johnston, and M. Liarokapis, "Single-grasp, model-free object classification using a hyper-adaptive hand, Google soli, and tactile sensors," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Madrid, Spain, Oct. 2018, pp. 1943-1950.
- [11]. K. A. Smith, C. Csech, D. Murdoch, and G. Shaker, "Gesture recognition using mm-wave sensor for human-car interface," *IEEE Sens. Lett.*, vol. 2, no. 2, pp. 1-4, Jun. 2018.
- [12]. S. Hazra, and A. Santra, "Robust gesture recognition using millimetric-wave radar system," *IEEE Sens. Lett.*, vol. 2, no. 4, pp. 1-4, Dec. 2018.
- [13]. P. Molchanov, S. Gupta, K. Kim, and K. Pulli, "Short-range FMCW monopulse radar for hand-gesture sensing," in *Proc. IEEE Radar Conf.*, Arlington, VA, USA, May 2015, pp. 1491-1496.
- [14]. S. Rao, A. Ahmad, J. C. Roh, and S. Bharadwa, "77GHz single chip radar sensor enables automotive body and chassis applications," Texas Instruments, [Online]. Available: <http://www.ti.com/lit/wp/spry315/spry315.pdf>
- [15]. Q. Wan, Y. Li, C. Li, and P. Pal, "Gesture recognition for smart home applications using portable radar sensors," in *Proc. 2014 36th Ann. Int. Conf. IEEE Eng. in Med. Bio. Society*, Chicago, IL, USA, Aug. 2014, pp. 6414-6417.
- [16]. X. Gao, J. Xu, A. Rahman, et al., "Barcode based hand gesture classification using AC coupled quadrature Doppler radar," in *Proc. IEEE MTT-S Int. Microw. Symp.*, 2016, San Francisco, CA, USA, May 2016, pp. 1-4.

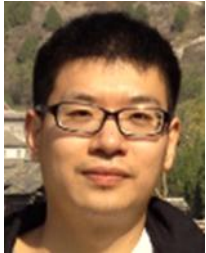


- [17]. Z. Zhou, Z. Cao, and Y. Pi, "Dynamic gesture recognition with a terahertz radar based on range profile sequences and Doppler signatures," *Sensors*, vol. 18, no. 10, pp. 1-15, Oct. 2018.
- [18]. J. Zhang, J. Tao, and Z. Shi, "Doppler-radar based hand gesture recognition system using convolutional neural networks," in *Proc. Int. Conf. Commun., Signal Process. Syst.*, Singapore, 2017, pp. 1096-1113.
- [19]. A. Arbabian, S. Callender S, S. Kang, M. Rangwala, and A M. Niknejad, "A 94 GHz mm-wave-to-baseband pulsed-radar transceiver with applications in imaging and gesture recognition," *IEEE J. Soli-St. Circ.*, vol. 48, no. 4, pp. 1055-1071, Apr. 2013.
- [20]. J. Fink, H. Guissouma, and F. Jondral, "Feature-based gesture classification by means of high resolution radar measurements," in *Proc. Int. Radar Symp.*, Prague, Czech Republic, Jun. 2017, pp. 1-10.
- [21]. Y Kim, and B. Toomajian, "Hand gesture recognition using micro-Doppler signatures with convolutional neural network," *IEEE Access*, vol. 4, pp. 7125-7130, Oct. 2016.
- [22]. G. Li, R. Zhang, M. Ritchie, and H. Griffiths, "Sparsity-based dynamic hand gesture recognition using micro-Doppler signatures," in *Proc. IEEE Radar Conf.*, Seattle, WA, USA, May 2017, pp. 928-931.
- [23]. G. Malysa, D. Wang, L. Netsch, and M. Ali, "Hidden Markov model-based gesture recognition with FMCW radar," in *Proc. IEEE Glob. Conf. Signal Informat. Process.*, Washington, DC, USA, Dec. 2016, pp. 1017-1021.
- [24]. G. Li, S. Zhang, F. Fioranelli, and H. Griffiths. "Effect of sparsity-aware time-frequency analysis on dynamic hand gesture classification with radar micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 12, no. 8, pp. 815-820, Aug. 2018.
- [25]. S. Zhang, G. Li, M. Ritchie, F. Fioranelli, and H. Griffiths, "Dynamic hand gesture classification based on radar micro-Doppler signatures," in *Proc. CIE Int. Radar Conf.*, Guangzhou, China, Oct. 2016, pp. 1-4.
- [26]. Z. Zhang, Z. Tian, and M. Zhou, "Latern: dynamic continuous hand gesture recognition using FMCW radar sensor," *IEEE Sens. J.*, vol. 18, no. 8, pp. 3278-3289, Apr. 2018.
- [27]. M. Ritchie, A. Jones, J. Brown, and H. Griffiths, "Hand gesture classification using 24 GHz FMCW dual polarised radar," in *Proc. Int. Conf. Radar Syst.*, Belfast, UK, Oct. 2017, pp. 1-6.
- [28]. P. Hügler, M. Geiger, and C. Waldschmidt, "RCS measurements of a human hand for radar-based gesture recognition at E-band," in *Proc. German Microw. Conf.*, Bochum, Germany, Mar. 2016, pp. 259-262.
- [29]. T. Fan, C. Ma, Z. Gu, *et al.*, "Wireless hand gesture recognition based on continuous-wave Doppler radar sensors," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 11, pp. 4012-4020, Nov. 2016.
- [30]. F. Khan, S. K. Leem, and S. H. Cho, "Hand-based gesture recognition for vehicular applications using IR-UWB radar," *Sensors*, vol. 17, pp. 1-18, Apr. 2017.
- [31]. B. Dekker, S. Jacobs, A. S. Kossen, M. C. Kruihof, A. G. Huizing, and M. Geurts, "Gesture recognition with a low power FMCW radar and a deep convolutional neural network," in *Proc. Euro. Radar Conf.*, Nuremberg, Germany, Oct. 2017, pp. 163-166.
- [32]. J. Suh; S. Ryu, B. Han, J. Choi, J. Kim, and S. Hong, "24 GHz FMCW radar system for real-time hand gesture recognition using LSTM," in *Proc. Asia-Pacific Microw. Conf.*, Kyoto, Japan, Nov. 2018, pp. 860-862.
- [33]. T. Sakamoto, 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 Sens. Lett.*, vol. 2, no. 3, pp. 1-4, Sep. 2018.
- [34]. Y. Sun, T. Fei, S. Gao, and N. Pohl, "Automatic radar-based gesture detection and classification via a region-based deep convolutional neural network," in *Proc. IEEE Int. Conf. Acous., Speech and Signal Process.*, Brighton, UK, May 2019, pp. 4300-4304.
- [35]. S. Skaria, A. Al-Hourani, M. Lech, and R. J. Evans, "Hand-gesture recognition using two-antenna Doppler radar with deep convolutional neural networks," *IEEE Sens. J.*, vol. 19, no. 8, pp. 3041-3048, Apr. 2019.
- [36]. J. Choi, S. Ryu, and J. Kim, "Short-range radar based real-time hand gesture recognition using LSTM encoder," *IEEE Access*, vol. 7, 33610-33618, Mar. 2019.
- [37]. Y. Wang, S. Wang, M. Zhou, Q. Jiang, and Z. Tian, "TS-I3D based hand gesture recognition method with radar sensor," *IEEE Access*, vol. 7, 22902-22913, Feb. 2019.
- [38]. S. Ryu, J. Suh, S. Baek, S. Hong, and J. Kim, "Feature-based hand gesture recognition using an FMCW radar and its temporal feature analysis," *IEEE Sens. J.*, vol. 18, no. 18, pp. 7593-7602, Sep. 2018.
- [39]. X. Lou, Z. Yu, Z. Wang, K. Zhang, and B. Guo, "Gesture-radar enabling natural human-computer interactions with radar-based adaptive and robust arm gesture recognition," in *Proc. IEEE Int. Conf. Syst., Man, and Cyber.*, Miyazaki, Japan, Oct. 2018, pp. 4291-4297.
- [40]. G. Li, R. Zhang, M. Ritchie, H. Griffiths, "Sparsity-driven micro-doppler feature extraction for dynamic hand gesture recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 655-665, Apr. 2018.
- [41]. L. Yang, and G. Li, "Sparsity aware dynamic gesture recognition using radar sensors with angular diversity," *IET Radar, Sonar Navigat.*, vol. 12, no. 10, pp. 1114-1120, Oct. 2018.
- [42]. M. Q. Nguyen, and C. Li, "Radar and ultrasound hybrid system for human computer interaction," in *Proc. IEEE Radar Conf.*, Oklahoma City, USA, Apr. 2018, pp. 1476-1480.
- [43]. AWR1642 evaluation module single-chip mmWave sensing solution user's guide (Rev. B), Texas Instruments, [Online]. Available: <http://www.ti.com/lit/ug/swru508b/swru508b.pdf>
- [44]. P. Goswami, S. Rao, S. Bharadwaj, and A. Nguyen, "Real-time multi-gesture recognition using 77 GHz FMCW MIMO single chip radar," in *Proc. IEEE Int. Conf. Consumer Electron.*, Las Vegas, USA, Jan. 2019, pp. 1-4.
- [45]. S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: a review of signal processing techniques," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 22-35, Mar. 2017.
- [46]. S. Nisar, A. A. Khan, and M. Y. Javed, "A statistical feature based decision tree approach for hand gesture recognition," in *Proc. Int. Conf. Front. Informat. Techn.*, Pakistan, Dec. 2009.
- [47]. Y. Luo, D. Hue, B. Luo, and Q. Zhang, "Motion compensation for SISAR based on contrast maximization," in *Proc. Asian. Pacif. Conf. SAR (APSAR)*, Nov. 2007, pp. 431-434.
- [48]. J. Oh, T. Kim, and H. Hong, "Using binary decision tree and multiclass SVM for human gesture recognition," in *Proc. Int. Conf. Informat. Sci. & Appl.*, Suwon, South Korea, Jun. 2013, pp. 1-4.
- [49]. Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1--27:27, 2011 .



**CHANGJIANG LIU** was born in Shandong, China, in 1990. He received his B.S. degree in information engineering from Beijing Institute of Technology in 2012 and his Ph.D. degree in signal and information processing from Beijing Institute of Technology in 2018. Since May 2018, he has been a post-doctoral researcher with the School of Transportation Science and Engineering, Beihang University, Beijing, China. His research interests include bistatic radar, forward scatter radar, insect radar, automotive radar, and radar-based intelligent

sensing.



**YUANHAO LI** (M' 18) received the B.S. degree in information engineering and the Ph.D. degree in Information and Communication Engineering from the Beijing Institute of Technology (BIT), Beijing, China, in 2012 and 2018, respectively.

From Dec. 2016 to Dec. 2017, he worked as a Visiting Researcher with the Dipartimento di Elettronica e Informazione, Politecnico di Milano (PoliMi), Italy. From Apr. to Sep. 2018, he was a Research Scientist with the Radar Research Lab, BIT, Beijing, China. Since Oct.

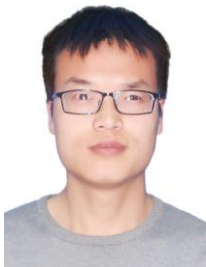
2018, he has been a Post-Doctoral Researcher with the Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands.

His research interests include radar system design and signal processing, InSAR/D-InSAR technique, ionospheric effects on radar systems, and radio frequency interference signal analysis. Dr. Li was a recipient of the Chinese Institute of Electronics Youth Conference Excellent Paper Award in 2014. He is a reviewer for many international peer-reviewed journals.



**DONGYANG AO** received the B.S. degree in information engineering and the Ph.D. degree in Information and Communication Engineering from the Beijing Institute of Technology (BIT), Beijing, China, in 2013 and 2019, respectively. From Oct. 2016 to Sep. 2018, he was a visiting scientist at the Data Intelligence and Knowledge Discovery group of German Aerospace Center (DLR). His research interests include signal processing and artificial intelligence, especially deep learning and its

applications in radar.



**HAIYAN TIAN** was born in Henan, China, in 1988. He received his M.S. degree from Chongqing University of Posts and Telecommunications in 2013 and his Ph.D. degree from Beijing University of Posts and Telecommunications in 2017. Since Dec. 2017, he has been a post-doctoral researcher with the School of Computer Science and Engineering, Beihang University, Beijing, China. His research interests include antenna design, PCB design and radar test.