ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-7605-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICASSP39728.2021.9414837

RECOGNITION OF DYNAMIC HAND GESTURE BASED ON MM-WAVE FMCW RADAR MICRO-DOPPLER SIGNATURES

Wen Jiang[1] *Yihui Ren*[1] *Ying Liu*[1] *Ziao Wang*[2] *Xinghua Wang*[2]

[1]School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, China

[2]School of Information and Electronics, Beijing Institute of Technology, Beijing, China

ABSTRACT

Radar-based sensors provide an attractive choice for hand gesture recognition (HGR). The very challenging problems in radar-based HGR are radar echo data preprocessing and recognition accuracy. In this paper, we propose a convolutional neural network (CNN) for dynamic HGR based on a millimeter-wave Frequency Modulated Continuous Wave (FMCW) radar which operates at 77GHz. Six different dynamic hand gestures are designed and the time-frequency analysis of micro-Doppler signatures are adopted as the input to CNN. The measured data of the dynamic hand gestures are collected in different experimental scenarios. The recognition accuracy of the six gestures based on the measured data reached 95.2%. The experimental results demonstrate that the proposed method is effective in the measured data and the micro-Doppler signature is effective for dynamic HGR.

Index Terms— Hand gesture recognition, millimeter wave radar, FMCW radar, micro-Doppler signatures, convolutional neural network

1. INTRODUCTION

Hand gesture, as one of the most natural and instinctive form of interpersonal communication, has long been idealized as a method of human-computer interaction (HCI). Hand gesture recognition (HGR) can be equipped with miniature electronic apparatus instead of physical contact with buttons or touch screens. With the development of computer science, the devices and techniques of HGR such as wearable sensors [1], vision-based sensors [2] have been widely used in gaming, driver assistance, smart home, medical rehabilitation and other HCI fields [3]. Initially, wearable devices which depend on gyroscope can digitize hand movements into multi-parametric data, but these devices are quite expensive and bring a lot of complicated operations and distraction to the users [4]. Then, numerous noncontact vision-based solutions have been proposed [5], which present high performance in recognizing hand gesture and provide a more comfortable experience than wearable devices. However, vision-based methods [5,6] would significantly consume computational resources and highly rely on the environment conditions, and they cannot work under low visibility. Recently, radar-based sensors provide an attractive choice for HGR [7-9]. Although the accuracy of radar-based sensor is not as good as that of vision-based sensor, it still has advantages in privacy protection, illumination robustness, easy embeddable and low cost.

Currently, the most common radar-based methods for HGR contain 3 steps: 1) hand gesture detection, 2) feature extraction, 3) recognition. In terms of the hand gesture detection, several radar products have used the multi-antenna, the millimeter wave band, and the broadband waveform to make the radar sensor smaller in size and more accurate in performance. NVIDIA developed a 24GHz monopulse radar which can obtain the multi-dimensional information of targets [10]. In [11], a dynamic HGR system based on 60GHz radar sensor was designed by Google Soli team for fine-grained HGR. Texas Instruments designed a 76-81GHz FMCW radar sensor which can be applied to vital sign detection, automotive detection, etc. [12].

Micro-Doppler effect is caused by the micro-motion of an object and its parts. Radar micro-Doppler signatures have been widely used in human movement analysis. Kim et al. [13] used micro-Doppler signatures to classify six human activities including running, walking, holding, crawling, boxing and sitting. Le et al. [14] classified a person walking with arm swinging and without swinging from micro-Doppler signals. Seyfioglu et al. [15] exploited human gait features from radar micro-Doppler spectrograms. According to these research, it has been proved that classification based on micro-Doppler signatures is effective.

The selection of classification algorithm mainly depends on the types of the extracted features from hand gesture signals. In traditional two-phase classification methods, empirical features [16], the principal component analysis features [17], and the sparse features [18] are first extracted and then fed into a classifier, such as support vector machine (SVM), random forest or decision tree to accomplish dynamic HGR. The emerging deep learning algorithm, such as convolutional neural networks (CNN), has brought a breakthrough in various fields [19], which is regarded as an effective method for dynamic HGR. In [20], a fused CNN architecture is designed for HGR based on multistatic radar micro-Doppler signatures. Zhang et al. [21] proposed a CNN approach for dynamic HGR in driver assistance system. Park et al. [22] used the 1-dimension CNN and long short-term memory (LSTM) to learn the waveforms in the 33GHz radar-based HGR system.

Although great progress has been made in radar sensors and classification algorithms for dynamic HGR, the frequencies of the above radar sensors for HGR are mainly focusing on the U-band or lower band which results in poor recognition accuracy. Moreover, the movement of the hand is coherent in spatial and temporal dimension, and the unintended actions are highly possible to interfere with recognition performance. To improve recognition accuracy, in this paper, we present a dynamic HGR system based on 77GHz FMCW radar sensor. In order to avoid the effects of redundant actions, we select six kinds of gestures as examples and process them periodically. Gesture data is collected by the millimeter-wave radar with high resolution. Finally, a CNN is utilized to classify the six hand gestures by extracting the micro-Doppler signatures from the time-frequency spectrograms.

2. MEASUREMENT OF DYNAMIC HAND GESTURES

2.1. FMCW radar system

The 77GHz FMCW radar adopted for dynamic HGR in this paper is the AWR1642 device, which is an integrated single-chip mm-wave radar manufactured by Texas Instruments and specifically designed for automotive applications. This sensor includes onboard etched antennas with 2 transmitters and 4 receivers and it can operate in the band of 76 to 81 GHz. Fig.1 presents the AWR1642 radar board and a sawtooth modulated FMCW. The radar board consists of radio frequency subsystem, radio processor subsystem and sensing evaluation module. The tasks of radio frequency subsystem are signal transmission and reception through onboard antennas. The radio processor subsystem mainly performs signal generation and processing. The sensing evaluation performs real-time data storage and extension.

In this FMCW radar system, the transmit signal is a chirp signal. A group of chirps form a frame and then multiple frames are integrated as a processing unit. The transmitted FMCW signal is given by

$$s(t) = \exp(j2\pi f_c t + j\pi k t^2) \qquad (1)$$

where f_c is the carrier frequency, $k = B/T$ denotes the sweep frequency slope, B is the bandwidth, and T is the sweep time. Assuming a target with the distance R from the radar and moving radical velocity v, the reflected signal is described as

$$r(t) = \exp(j2\pi f_c (t-\tau) + j\pi k (t-\tau)^2) \qquad (2)$$

where $\tau = 2(R+vt)/c$ denotes the round-trip delay. In the FMCW radar system, the output beat signal is obtained by mixing the transmitted signal with the received signal as

$$b(t) = \exp(j2\pi f_c \tau + j2\pi k t \tau - j\pi k \tau^2) \qquad (3)$$

The beat signal consists of multiple tones. The frequency of each tone is proportional to the distance of the corresponding object. The experimental parameters are listed in Table 1.

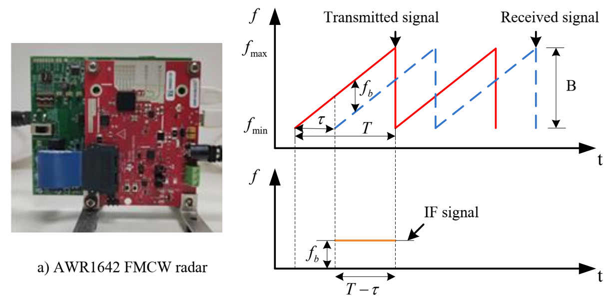

Fig.1. AWR1642 radar and FMCW signal

Table 1. Radar waveform settings and sensing parameters

Waveform Settings	
Sweep bandwidth B	3.96GHz
Sweep frequency slope k	65.9MHz/us
Signal period time T	60us
Frame time	40.96ms
Number of frames	256
Sampling frequency f_s	10MHz
Sensing parameters	
Range resolution ΔR	3.75cm
Velocity resolution Δv	4.75cm/s
Maximum detection range R_{max}	22.7m
Maximum detection velocity v_{max}	6.08m/s

2.2. Measurement setup and gesture design

In this paper, a novel radar-based dynamic HGR system is proposed, which is shown in Fig.2. The AWR1642 radar sensor is placed vertically upward on the test-bed. Experimental subjects perform different hand gestures in front of the antenna. The beat signal from radar sensor is collected and then processed by a filer to remove the static clutter of the experimental environment. To obtain the micro-Doppler signatures from the measured data, the data is processed to time-frequency spectrum by short-time Fourier transform (STFT) and then fed into a deep CNN classifier for dynamic HGR.

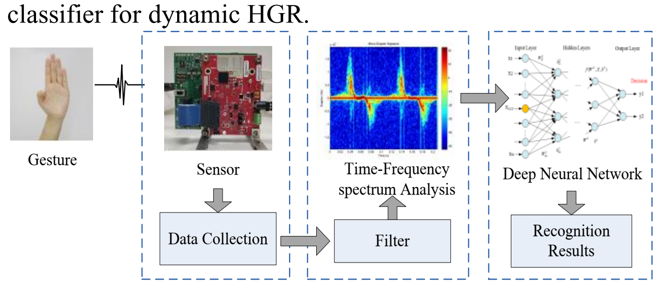

Fig.2. The diagram of the proposed radar-based HGR system

Six different dynamic hand gestures are considered which are presented in Fig. 3, and the descriptions of the hand gestures are given as follows.

a). Swiping forward: waving the hand towards and away from the radar while keeping the wrist static.

b). Swiping backward: waving the hand from front to back.

c). Holding: opening the hand initially then closing the palm as if holding something.

d). Rotating: rotating the hand clockwise and keeping the wrist static over the radar.

e). Flapping: waving the hand up and down to the radar boresight.

f). Snapping finger: pressing the thumb and the middle finger together and flinging the thumb forward while the middle finger sliding into the palm rapidly.

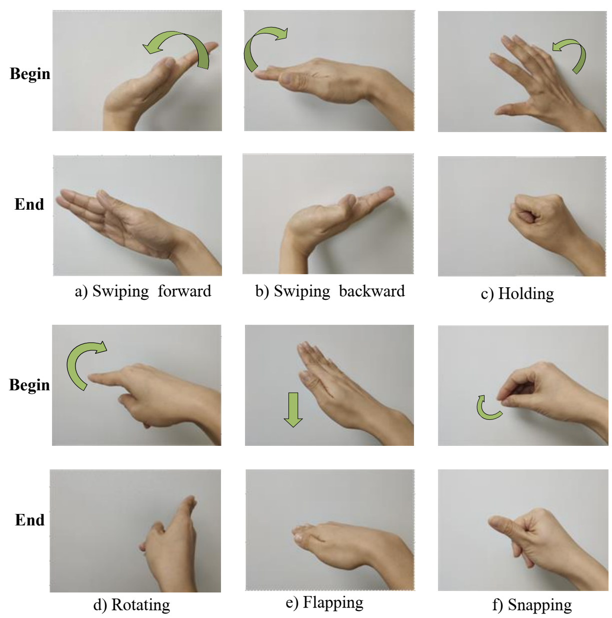

Fig.3. Illustrations of six kinds of dynamic hand gestures

3. METHODOLOGY

3.1. Time-frequency analysis

To obtain the micro-Doppler characteristics from the dynamic hand gestures, the STFT is applied to the measured data. The time-frequency spectrograms of six dynamic hand gestures from one experimental subject are shown in Fig. 4.

It is clear to see that the time-frequency distributions of six hand gestures are different from each other. To help understanding, colors indicate the speed and strength of each hand movement. The strongest component which is around 0Hz in each spectrogram corresponds to the static body or arm of the experimental subject. From Fig. 4 a) and b) we can see that when the subject is performing the gesture of swiping forward, the radial velocity of the hand changes from positive to negative or vice versa. As to the gesture of holding in Fig. 4 c), a high positive frequency peak and a following negative frequency peak, which corresponds to curling up the fingers onto the palm continuously. In Fig. 4 d), we can find that the gesture of hand rotation is a continuous motion. The micro-Doppler frequency changes along with the time slightly, like a sine wave. The gesture of flapping has two motions just as a bird flapping its wings. The spectrogram indicates the motion of waving the palm up and down. In Fig. 4 f), the gesture of snapping finger has two discontinuous motions in each period and the Doppler frequency of this gesture is close to zero on account of the motion is slight in the radial direction. In practice, the hand is not a point-like target and the micro-motion trajectories of hand are unlikely perfectly symmetric.

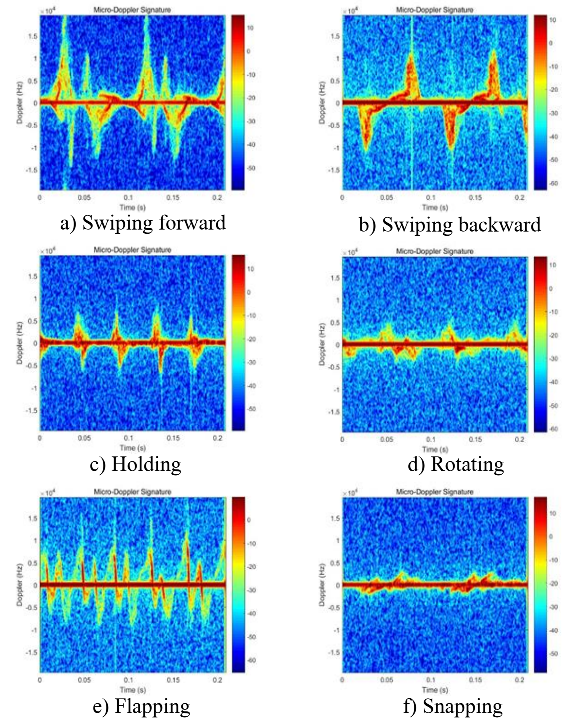

Fig.4.The time-frequency spectrograms of six hand gestures

3.1. The proposed CNN classifier

The proposed CNN classifier is evaluated on the measured hand gesture dataset described in Fig. 3. As shown in Fig.5, the CNN used for dynamic HGR consists of 5 convolutional (Conv) layers, 5 max-pooling layers, 2 fully connected layers and batch normalization. The activation function is ReLU. The size of input is 901×1201.

We trained the parameters of the CNN model to recognize the dynamic hand gesture. The kernel size of each Conv layer is 3×3, the kernel number of the first 5 Conv layers increases from 6, 16, 32, 64 to 128. The Conv layers are followed by max-pooling layers and each max-pooling layer has a 2×2 spatial pooling area with a stride of 2. The network model is trained with back-propagation and stochastic gradient descent optimizer with initial learning rate 0.001. In our experiment, the dropout rate is 0.4 and the batch size is 100. PyTorch is used for model implementation.

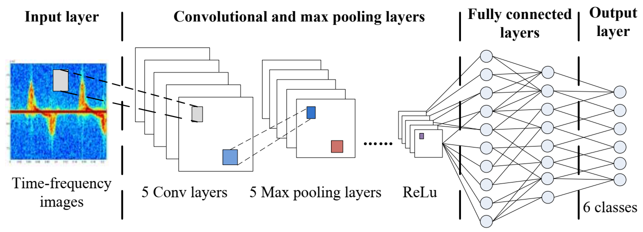

Fig.5. The CNN architecture for dynamic HGR

4. EXPERIMENTAL RESULTS

Five right-handed adults are invited to perform each hand gestures for 60 times, generating a dataset containing 5(subjects) × 6(gestures) ×60(repetitions) = 1800 recordings. The measured data of four subjects is used for training, and the other subject for testing. Such procedure is repeated for all the subjects and the average recognition accuracy is reported.

Experimental subjects perform different hand gestures in front of the radar at a range of about 20cm. The body of the subjects should remain static during the process of data collection. For the six hand gestures, the average accuracy is 95.2%. The confusion matrix of six hand gestures is shown in Fig.6. The labels in column correspond to true hand gestures and the labels in row are the gestures predicted by CNN. Observed from the confusion matrix, a relatively large misjudgment rate between rotating and snapping, due to the similarity between these two gestures for radar.

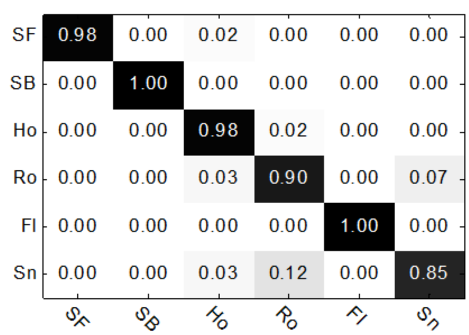

Fig.6. Confusion matrix of six hand gestures

In practical applications, there is a lot of uncertainty in the process of HGR. Since subjects are highly possible perform hand gestures at different locations, it is necessary to test the classification accuracy of gestures in different scenarios. Therefore, we designed two experimental scenarios with different distances and angles which are described in Fig.7. We used the previously trained CNN model which is compared with SVM [16]. Table 2 and Table 3 show the experimental results of the average accuracy at different distances and angles respectively.

a) Comparison between different distances

In order to verify the effect of distance on the classification accuracy, we tested the data in the normal radial direction of the radar at distances of 20cm, 50cm, and 100cm. The accuracy decreases slightly as the distance between human and radar increases. The recognition accuracy of six hand gestures remains at about 90% within a range of 100cm. The increase in distance only makes the signal-to-noise ratio decrease, but the profile of the time-frequency spectrogram is similar to the initial one.

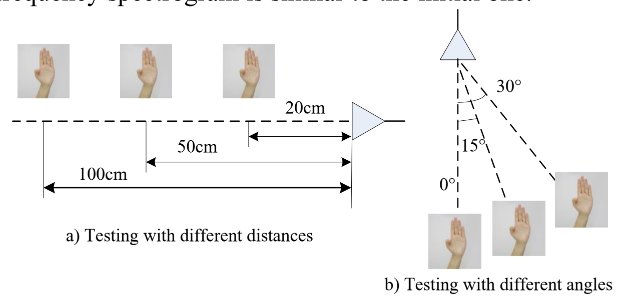

Fig.7. Gestures testing in different scenarios

Table 2. Average accuracy of HGR at different distances

Accuracy	R=20cm	R=50cm	R=100cm
SVM [16]	91.4%	87.3%	81.5%
CNN	95.2%	92.3%	88.7%

b) Comparison between different angles

We then performed hand gestures at different angles and the distance from the radar remained at 20cm. Considering the beam width of main lobe of the radar, the measured angles to the radar were 0°, 15°, and 30°. The recognition accuracy of hand gestures drops as the angle becomes larger. The main lobe has a certain range. As the angle increases, the energy of the received signal decreases. Furthermore, since Doppler frequency only depends on radial velocity, the change of angle also affects the micro-Doppler signatures of the hand gestures.

Table 3. Average accuracy of HGR at different angles

Accuracy	Angle=0°	Angle=15°	Angle=30°
SVM [16]	91.4%	77.8%	60.3%
CNN	95.2%	81.7%	68.4%

In summary, the recognition accuracy was mainly affected by the angle between the hand and the radar. Compared with angular variation, the drop in accuracy caused by distance variation is not obvious. Thus, the angle between a radar device and a subject is particularly important for the HGR system.

5. CONCLUSION

In this paper, we demonstrated the feasibility of a dynamic HGR system based on a 77GHz FMCW radar. Six dynamic hand gestures are defined and the micro-Doppler signatures are analyzed. The CNN is adopted to classify the time-frequency spectrograms of the gestures. Experimental results verify that the proposed classifier can work well in HGR system. Further work will focus on extending the types of the gestures and performing methods to augment the data set. Additionally, it is quite necessary to explore more advanced algorithms to overcome the variations in the same gestures in different application scenarios, thereby improving the robustness of the HGR system.

6. REFERENCES

[1] A. Dekate, A. Kamal, and K. S. Surekha, "Magic-Glove wireless hand gesture hardware controller," in *International Conference on Electronics and Communication Systems*, pp. 1-4, 2014.

[2] G. G. Rogez, J. S. Supancic, and D. Ramanan, "Understanding everyday hands in action from RGB-D Images," in *IEEE International Conference on Computer Vision*, pp. 3889-3897, 2015.

[3] M. J. Cheok, Z. Omar, and M. H. Jaward, "A review of hand gesture and sign language recognition techniques," *Int. J. Mach. Learn. & Cyber.*, vol. 10, no. 1, pp. 131–153, Jan. 2019.

[4] C. Xu, P. H. Pathak, and P. Mohapatra, "Finger-writing with smartwatch: A case for finger and hand gesture recognition using smartwatch," in *International Workshop on Mobile Computing Systems and Applications*, pp. 9-14, 2015.

[5] S. Y. Cheng and M. M. Trivedi, "Vision-based infotainment User Determination by Hand Recognition for Driver Assistance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 3, pp. 759-764, 2010.

[6] E. Ohn-Bar and M. M. Trivedi, "Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations," *IEEE Transactions on Transportation Systems*, vol. 15, no. 6, pp. 2368-2377, 2014.

[7] Q. Wan, Y. Li, C. Li, and R. Pal, "Gesture recognition for smart home applications using portable radar sensors," in *Proceeding of 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6414-6417, August 2014.

[8] Z. Zhang, Z. Tian, M. Zhou, "Latern: Dynamic continuous hand gesture recognition using FMCW radar sensor," *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3278-3289, 2018.

[9] K. Zhang, Z. Yu, D. Zhang, "RaCon: A gesture recognition approach via Doppler radar for intelligent human-robot interaction," in *2020 IEEE International Conference on Pervasive Computing and Communications Workshops*, 2020.

[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, pp. 1943–1950, Oct. 2018.

[11] 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, pp. 1491–1496, May 2015.

[12] S. Rao, A. Ahmad, J. C. Roh, and S. Bharadwa. "77GHz single chip radar sensor enables automotive body and chassis applications." *Texas Instruments*, Accessed: Dec. 20, 2018. [Online]. Available: http://www.ti.com/lit/wp/spry315/spry315.pdf

[13] Y. Kim, H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, 47: 1328-1337, May 2009.

[14] H. T. Le, S. L. Phung, A. Bouzerdoum, and F. H. Tivive, "Human motion classification with micro-Doppler radar and bayesian-optimized convolutional neural networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, pp. 7640–7644, 2018.

[15] M. S. Seyfioglu, B. Erol, S. Z. Gurbuz, and M. G. Amin, "DNN transfer learning from diversified micro-Doppler for motion classification", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 55, No. 5, October 2019.

[16] S. Zhang, G. Li, M. Ritchie, F. Fioranelli, and H. Griffiths, "Dynamic hand gesture classification based on radar micro-Doppler signatures," in *2016 International Conference on Radar*, IEEE, pp. 1-4, 2016.

[17] G. Li, R. Zhang, M. Ritchie, and H. Griffiths, "Sparsity-based dynamic hand gesture recognition using micro-Doppler signatures," in *2017 IEEE Radar Conference*, IEEE, pp. 0918-0931, 2017.

[18] L. Yang, G. Li, "Sparsity Aware Dynamic Gesture Classification Using Dual-band Radar," in *2018 19th International Radar Symposium*, IEEE, pp. 1-6, 2018.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, May 2015.

[20] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, "Dynamic hand Gesture classification based on multistatic radar micro-Doppler signatures using convolutional neural network," in *2019 IEEE Radar Conference*, IEEE, pp.1-5, 2019.

[21] X. Zhang, Q. Wu, D. Zhao, "Dynamic hand gesture classification based on radar micro-Doppler signatures," in *2018 10th International Conference on Wireless Communications and Signal Processing*, October 2018.

[22] J. Park, J. Jang, G. Lee, H. Koh, C. Kim, and T. W. Kim, "A time domain artificial intelligence radar for hand gesture recognition using 33-GHz direct sampling," in *2019 Symposium on VLSI Circuits Digest of Technical*, IEEE, 2019.