# A Real-time Hand Gesture Recognition System using 24 GHz Radar Array

Guiyuan Zhang, Kang Zhang, Shengchang Lan, Yuanxun Liu, Lijia Chen
Dept. Microwave Enginnering
Harbin Institute of Technology
Harbin, China
1150550114@stu.hit.edu.cn, 1160501020@stu.hit.edu.cn, lansc1015@hit.edu.cn,
1160500726@stu.hit.edu.cn, ljchen@hit.edu.cn

***Abstract*—This paper presents a description of a real-time hand gesture recognition system. This system consists of three commercial modules perpendicular mounted in an three-dimensional array to provide six-channel baseband I/Q signals. The I/Q signals are pre-processed by the doppler signal amplitude threshold detection and spectral analysis. A convolutional neural network consisting in two convolutional layers and two fully connected layers is constructed as the recognition classifier with less dependence of feature extraction. The network is trained with 1000 groups of datasets and verified by testing recognized results as the customized shortcut keys. Results show that this system could achieve a high recognition accuracy rate higher than 95% in the real-time test.**

## I. Introduction

Hand gesture recognition is an efficient way for Human Machine Interaction. The key problem in human gesture recognition is feature extraction and classification. Rather than cameral imaging and inertial sensors, radar has been suggested as an alternate method for recognizing hand gestures[1]. In this paper we use doppler radars to detect the velocity information of hand gesture and to reach a real-time control of a windows computer based on 24 GHz commercial doppler radars. The essential work of this project is to recognize the five hand gestures and later to implement the computer control as such as page up and down, window switch, and window close according to the previous recognition result. There are lots of developments about human gesture recognition using millimeter wave radars. Most impressive progress has been made, since 2016, by Google's milestone work "Soli" using a 60 GHz Frequency Modulated Continuous Wave radar to fulfill hand recognition [2]. Our project aims to investigate the feasibility of recognizing different hand gesture based on commercial lost-cost radars with a three-dimensional array configuration. The classification is based on a convolutional neural network with the optimization in gradient descending and anti-over-fitting. The prototype has been developed and tested with higher accuracy than 95% in the real-time scenario. This result verifies the proposed method with a promising vision in the modern human-computer interactions.

## II. System overview

We use K-LC2, from RFBeam Switzerland to construct the radar array. K-LC2 is a $2 \times 4$ patch module operating at 24.125 GHz with a bandwidth of 50 MHz [3]. The patch antenna has an asymmetric beam with $80^{\circ}$ in horizontal and $34^{\circ}$ in vertical. The reflected electromagnetic waves are heterodyned by the integrated local oscillator inside. K-LC2 has two output channels of the baseband I/Q signals. DAQ USB-6211 from National Instruments is used as the analog-digital converter. A baseband signal amplifier circuit is inserted between the radar baseband output and the analog-digital converter. The amplifier is implemented on an LMV722 chip providing two operational amplifier channels with a gain of 10 dB. A statistic study of the average hand gesture time duration studied in this project is between 0.4 and 0.5 seconds at most. Therefore, the sampling frequency is defined as 1500 Hz and the sampling duration of one frame data is 1 second. Fig. 1 shows the radar array architecture designed in the project.

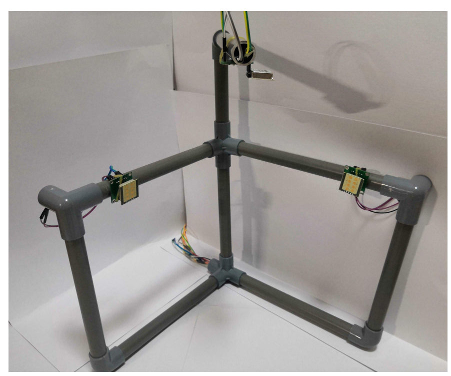

Fig. 1. The 24GHz radar array for hand gesture recognition.

Users prefer nature and fluent hand gestures to control electronics. We also use five frequently used hand gestures shown in Fig. 2 as the target hand gestures. The five hand gestures in this project : (a)hand right, (b)hand left, (c)open hand , (d)close hand , (e)finger snap. We construct a three-dimensional radar array with three commercial radar modules mounted perpendicular to each other. In the coverage of three modules' Field of View (FoV), any hand gesture can reflect the electromagnetic waves back to three radar modules in the array.

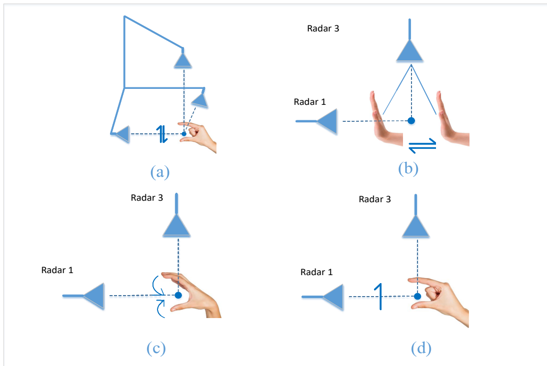


Fig. 2. The gestures supported in this project.(a)the position of radars and gesture,(b)hand right and left,(c)open and close hand,(d)finger snap

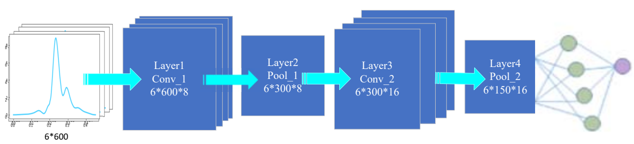


Fig. 3. The structure of the CNN Network.

## III. Feature Extraction and Classification

We use doppler effect to obtain the velocity of hand gestures. By using these radar modules, gesture information is translated into the time domain. To unveil the features, short-time energy spectral analysis in speech signal processing is applied to the temporal radar signals. The I/Q signal is processed in frames. For real-time recognition, we apply amplitude threshold detection on the I/Q signal to find frames containing gesture signals. Two adjacent frames are processed together to maintain the temporal peak integrity. After spectral analysis, the results are used as input data for the Convolutional Neural Network (CNN) to complete the gesture recognition. In the classification, it avoids the complicated pre-processing, directly utilizes the original data as the input, and decreases the complexity of neural network efficiently under its unique network structure. The structure of our CNN is shown in Fig. 3. It has two convolutional layers and two fully connected layers. Input data is formed into a one-dimensional vector. The value of each element in this vector represents the probability of accurate recognition. Adam optimizer is used to avoid the over-fitting before the output layer, it can iteratively update the neural network weights based on the training datasets [4]. This optimizer originally proposed in 2015 has advantages in solving non-convex optimization problems including efficient calculation, less memory required. The optimizer in the project uses a learning rate of 0.001 and meanwhile minimizes the cost function in terms of cross entropy.

Real-time demonstration include acquiring the data, processing the data, recognizing the data and finally using the data for desired functions by the cooperation of LabVIEW, TensorFlow. Radar signal is sampled based on LabVIEW and updated every second and the data is read and processed to the input data of CNN by Python, then the trained CNN model recognise the data, finally the result of recognition is used to accomplish the pre-defined functions.

TABLE I
Different input data and accuracy.

| Input Data | Accuracy |
|---|---|
| Time domain waveform | 75% |
| Detected time domain feature | 95% |
| Frequency domain feature | 95% |

## IV. Experiments and Result Analysis

To verify the project, we implemented an experiment of gesture recognition. For each gesture listed, we provided 200 training datasets. Then the experiment calculated the results of each hand gesture in real-time demonstration 200 times. The accuracy of the slide control can reach to 95%. Also, we compared the performance of the proposed method with other analogue methods such as non-preprocessed temporal signals input, intercepted temporal signals input based on the same CNN network. Table I shows the accuracy of different input data types. The results show that the I/Q signal processing help to increase the recognition from 75% to 95%.

## V. Conclusion

This paper proposed a real-time solution of using a Doppler radar array to implement the non-contact hand gesture recognition. A short-term energy spectrum analysis on the I/Q signals was to extract the hand gesture features. We also designed a CNN network to classify the proposed hand gestures and insert the recognition result into the Microsoft Office as the shortcut keys. The final result showed a good performance in five gesture with an average accuracy of 95%.

## Acknowledgment


We give our sincere thanks to the supports from National Natural Science Foundation of China, Harbin Institute of Technology and RFBeam company. This paper is contributed to China National Natural Science Foundation Project (Grant No. 61601141) and the Fundamental Research Funds for the Central Universities (Grant No. HIT.NSRIF.201817).


## References


[1] Y. Kim and B. Toomajian, “Application of doppler radar for the recognition of hand gestures using optimized deep convolutional neural networks,” in *2017 11th European Conference on Antennas and Propagation (EUCAP)*, March 2017, pp. 1258–1260.

[2] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, “Soli: Ubiquitous gesture sensing with millimeter wave radar,” *ACM Transactions on Graphics (TOG)*, vol. 35, no. 4, p. 142, 2016.

[3] *Standard product K-LC2 Dual Channel Radar Transceiver*, RfBeam, 2018. [Online]. Available: https://www.rfbeam.ch/product?id=5

[4] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.