Article

A Method of Ultrasonic Finger Gesture Recognition Based on the Micro-Doppler Effect

Qinglin Zeng ^{1,2}, Zheng Kuang ², Shuaibing Wu ² and Jun Yang ^{1,2,*}

¹ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China; zengqinglin16@mails.ucas.edu.cn

² Key Laboratory of Noise and Vibration Research, Institute of Acoustic, Chinese Academy of Sciences, Beijing 100190, China; kz@mail.ioa.ac.cn (Z.K.); wushuaibing@mail.ioa.ac.cn (S.W.)

* Correspondence: jyang@mail.ioa.ac.cn

Received: 6 May 2019; Accepted: 3 June 2019; Published: 5 June 2019

Abstract: With the popularity of small-screen smart mobile devices, gestures as a new type of human–computer interaction are highly demanded. Furthermore, finger gestures are more familiar to people in controlling devices. In this paper, a new method for recognizing finger gestures is proposed. Ultrasound was actively emitted to measure the micro-Doppler effect caused by finger motions and was obtained at high resolution. By micro-Doppler processing, micro-Doppler feature maps of finger gestures were generated. Since the feature map has a similar structure to the single channel color image, a recognition model based on a convolutional neural network was constructed for classification. The optimized recognition model achieved an average accuracy of 96.51% in the experiment.

Keywords: ultrasonic; finger gesture recognition; convolutional neural network; the micro-Doppler effect

1. Introduction

Touchscreen control is now used in most mobile devices, such as mobile phones and tablets. When a person uses it with a wet hand or gloved hand, the touch will not work well. With the rapid development of mobile devices, such as small-sized smart watches, it is also very inconvenient to control devices on a small screen. As a part of human communication, gestures can be used to express a wide variety of emotions and thoughts. Gestures are usually the second most natural method of interaction between humans and the environment, as well as among humans [1]. Gestures are convenient and have a vast interaction space and super high flexibility, providing excellent interactive experience. Therefore, gestures in human–computer interactions have gained greater attention in recent years [2].

A variety of gesture recognition methods have been proposed [3–10]. These methods can usually be divided into four categories: wearable sensor-based, optical vision-based, electromagnetic sensor-based and ultrasound-based.

Wearable sensor-based: Many gesture recognition methods based on wearable sensors have been reported [3,4,11,12]. John Weissmann et al. [3] explored a data glove as the input device and recognized five kinds of predefined hand gestures. Renqiang Xie and Juncheng Cao [12] presented an accelerometer-based pen-type sensing device and a user-independent hand gesture recognition algorithm that achieved almost perfect user-independent recognition accuracies. All of these methods require the user to wear additional sensors on the hand or arm. In contrast, ultrasonic-based gesture recognition methods directly sense the movements of the hand without any wearable device.

Optical vision-based: Optical vision sensors, including color cameras, depth cameras and infrared cameras, are the most widely used in gesture recognition because of the successful commercial products

Microsoft Kinect and Leap Motion, which can capture human body activities by cameras [5,6,13]. Guillaume Plouffe et al. [6] developed a natural gesture user interface that could track and recognize hand gesture based on depth data collected by a Kinect sensor. Aurelijus Vaitkevičius et al. [14] presented a system that is capable of learning gestures by using the data from the Leap Motion device and the Hidden Markov classification algorithm. Although optical vision-based methods have good recognition performance, they are susceptible to illumination conditions and ambient infrared radiation [15]. In addition, these methods also have to entail high computational costs.

Electromagnetic sensor-based: Electromagnetic sensors have been widely used for human activity classification [7,8,16,17] and gesture recognition [18–26]. Most of the proposed methods for gesture recognition based on electromagnetic sensors can only recognize hand gestures of big movements, but some methods for classifying finger gestures have also been reported, such as WiFinger's method [23] and project Soli [22,27]. WiFinger [23] presented a fine-grained finger gesture recognition system by using a single commodity WiFi device and achieved a recognition accuracy of over 93%. However, the system is impractical in the outdoors considering the availability of commercial wireless Access Points to connect the WiFi device. Soli [27] is a new gesture sensing technology designed by Google and designated to be an interaction sensor that uses radar for motion tracking of the human hand. The key part of Soli's design is a dedicated radar chip, incorporating the entire sensor and antenna array into an ultra-compact 8 mm × 10 mm package. In contrast, the proposed method in this paper just requires a pair of common commercial sensors, a simpler hardware setting and low-cost computing resources.

Ultrasound-based: Ultrasound-based methods of gesture recognition can be generally classified into three categories via different schemes. First, the Doppler effect has been used to sense gesture for a long time [9,10,28–32]. For example, with the speaker already embedded in the laptop, SoundWave [10] generates an inaudible tone and measures the frequency shift of echo reflected from the moving hand. Dolphin [28], leveraging the loudspeaker and microphone in smartphones, extracts features from the Doppler shift and recognizes a rich set of predefined hand gestures by combining manual recognition and machine learning methods. The obvious problem with these proposed methods is that they can only recognize hand gestures of big movements, and are limited to simple gestures or combinations of these simple gestures, such as: push, pull, slide left or right. Second, many methods use ultrasonic sensor array to estimate direction of arrival (DOA) or range [33–35], requiring a complex hardware setup and multiple (at least 3) sensors to form a specific geometry, such as: a triangle, a cross and a line. Third, though many works based on ultrasound tracking has been proposed [36–39]. These methods can track a hand or finger at a high accuracy, but these methods can not work well when it comes to multi-targets at the same time, especially targets moving in different directions. Hence, methods proposed by these works are inadequate for recognizing complex gestures.

In this paper, a method using only two ultrasonic sensors for recognizing finger gestures is proposed. One sensor is used to emit a single tone of 300 kHz, the other is used to receive echo reflected by moving fingers. Then, the micro-Doppler information from echos is processed into feature maps. According to the feature maps, a deep convolutional neural network (CNN) is built for finger gesture recognition and a competitive accuracy of recognition of five finger gestures is achieved.

2. Methods

2.1. The Micro-Doppler Effect

The micro-Doppler effect was firstly presented in coherent laser systems [40]. When the target or any part of the target has vibrations or rotations in addition to its bulk translation, it might cause additional frequency shifts on the returned signal. This phenomenon is referred to as the micro-Doppler effect [41]. In many early research, the micro-Doppler effect has been used to recognize the moving state of the human body [42,43]. Figure 1 is a schematic diagram of the micro-Doppler effect. The distance between the target and the ultrasonic transducer is $R(t)$. The displacement change caused by the

target's micro motion is $x(t)$. λ represents the wavelength of the ultrasonic wave, and the received signal $r(t)$ can be expressed by:

$$r(t) = A \cos(2\pi f_c t - \frac{4\pi R(t)}{\lambda} - \frac{4\pi x(t)}{\lambda} + \phi), \quad (1)$$

where f_c is the center frequency, ϕ is the initially phase of the transmit signal. Assuming that the distance $R(t)$ does not change in a short time, the phase change in the received signal $r(t)$ is caused by the target's micro motion $x(t)$. In this paper, the center frequency f_c is 300 kHz, and the speed of ultrasound travel in the air is 340 m/s, so the wavelength of transmit signal is 1.13 mm. If the micro motion of target is only 0.5 mm, it will induce 1.77π phase change.

Figure 1. The micro-Doppler effect.

2.2. Platform Design and Parameters Setting

Figure 2 shows a block diagram of the proposed method. It consists of two ultrasonic transducers, one signal conditioning circuit, one NI platform for signal sampling, and two post-processing modules, including micro-Doppler processing and recognition modules. The ultrasonic transducers are MA300D1-1 by Murata Manufacturing Co.. The MCU on the signal conditioning circuit generates a designed pulse signal. The pulse signal is amplified and stimulates the ultrasonic transceiver to emit ultrasonic waves. The echos with interference reflected from finger gestures are received and amplified by the signal conditioning circuit and then sampled by ADC in the NI platform [44]. The raw data is firstly saved on the NI platform and processed offline on a laptop.

Figure 2. The block diagram of the proposed method.

The center frequency of MA300D1-1 is 300 kHz and bandwidth is very small. So the transmit signal is designed as modulated pulses, showed in Figure 3. The τ is the width of pulse, T represents Pulse Repetition Interval (PRI). M is the total number of pulses for coherent processing, and T_{CPI} denotes Coherent Processing Interval (CPI). To obtain a high-quality micro-Doppler feature map requires high Doppler frequency resolution and range resolution. So the parameters of modulated pulses should be elaborately designed. The Doppler frequency resolution f_{dres} and the range resolution R_{res} are as [45]:

$$\begin{aligned} f_{dres} &= \frac{1}{M \cdot T} \\ R_{res} &= \frac{c \cdot \tau}{2} \end{aligned} \quad (2)$$

where c is the speed of ultrasonic wave travel in air and λ is wavelength of ultrasonic wave. In order to accurately recognize finger gestures, some parameters are designed carefully and are shown in the

left part of Table 1, and the performance indicators are calculated as shown in the right half of Table 1. As can be seen from the table, the Doppler frequency resolution is 15.625 Hz, the speed resolution reaches 9 mm/s and the distance resolution reaches 6.9 mm. Such high speed resolution and distance resolution are sufficient to clearly distinguish different finger motions.

Figure 3. The designed modulated pulses.

Table 1. System parameters setting and performance.

Parameters Setting		Performance	
Pulse carrier frequency (f_c)	300 kHz	Doppler frequency resolution (f_{dres})	15.625 Hz
Width of pulses (τ)	12	Velocity resolution (v_{res})	0.009 m/s
Pulse repetition interval (PRI)	1 ms	Range resolution (R_{res})	0.0069 m
Coherent processing interval (CPI)	64 ms	Unambiguous distance	0.172 m
Sampling frequency (f_s)	2 M/s	Unambiguous velocity	0.2867 m/s

2.3. Micro-Doppler Processing

2.3.1. Preprocessing

In Section 2.2, the designed modulated pulses can be expressed by:

$$T(t) = u(t - nT) \exp(j2\pi f_c t), \quad (3)$$

where $n = 0, 1, 2, \dots, N$, $T = PRI$ is the pulse repetition interval, f_c is the center frequency of the transmit signal, and $u(\cdot)$ is the complex envelope of modulated pulses. After reflected from finger motions, the received signal can be described as:

$$R(t, T') = u'(t - T') \exp \left[j2\pi f_c \left(t - \frac{2R(T') + 2x(T')}{c} \right) \right], \quad (4)$$

where $T' = nT$, $n = 0, 1, 2, \dots, N$. The received signal preprocessing flow as show in Figure 4 includes sampling, bandpass filtering, IQ demodulation, and lowpass filtering. After preprocessing, a two-dimensional pulse Doppler data matrix can be obtained. Each cell in the matrix can be represented by $y[l, n]$, and $y[l, n]$ is

$$y[l, n] = u' \exp(-j\frac{4\pi}{\lambda} R_l) \exp(j2\pi f_d nT), \quad (5)$$

where l is the range bin and n represents the nth pulse, $n = 0, 1, 2, \dots, N$. The larger n is, the longer the data collected is. The flow of preprocessing is shown in Figure 4.

Figure 4. The flow of preprocessing.

2.3.2. Micro-Doppler Processing

In order to improve the signal-to-noise ratio, the two-dimensional data matrix in the previous section must be coherently processed in the slow time dimension. The total number of pulses for coherent processing is set at 64. Based on the designed parameters of modulated pulses, target operating range is set to $R_1 \sim R_2$, that is in the range bin of $l_1 \sim l_2$. Accumulating all range bins within the operating range, Equation (5) can be rewritten as:

$$y[n] = u' \exp \sum_{R_1}^{R_2} \left(-j \frac{4\pi}{\lambda} R_l \right) \exp(j2\pi f_d n T), \quad (6)$$

Then $y[n]$ is divided into K segments, the length of each segment is M , and the incremental step between adjacent segments is D . The segmented data y_k is performed by fast Fourier transform (FFT) with a window size of $NFFT = 256$, and the result can be expressed as:

$$ST_k = \mathcal{F} \{y[D(k-1)], y[D(k-1)+1], \dots, y[D(k-1)+M-1]\}, \quad (7)$$

where $k = 1, 2, \dots, K$. Combining ST_k in order, the micro-Doppler feature map of finger gesture is acquired. In this paper, K is set at 45, and D is set at 14. Then the micro-Doppler feature map is in shape of 45×256 just like Figure 5, which is an example of five finger gestures. The whole flow of micro-Doppler processing illustrated in Figure 6.

Figure 5. The examples of five finger gestures. (a) Finger close; (b) Finger clockwise; (c) Finger slide; (d) Finger open; (e) Finger counterclockwise.

Figure 6. The flow of micro-Doppler processing.

2.4. Recognition Model

The gesture recognition model mainly uses machine learning algorithms, such as K-Nearest Neighbor (KNN) [28,32] and Support Vector Machine (SVM) [35], and deep learning algorithms, such as Recurrent Neural Network (RNN) [31] and convolution neural network (CNN) [2]. The machine learning algorithms need manual extraction of features, whose quality has a fatal impact on the results. While CNN shows excellent recognition performance in image recognition. The micro-Doppler feature map (after expanding it to a tensor of $45 \times 256 \times 1$) has the same structure as a single color channel image. Hence, the proposed method in this paper adopts a deep convolution neural network (CNN) for robust finger gesture recognition.

The network architecture of our recognition model is shown in Figure 7. The architecture is composed by three CNN layers for feature extraction, a Fully connected (FC) layer and a Softmax layer for classifying finger gestures.

Figure 8 illustrates the architecture of deep convolution neural network (DCNN) in this paper. Convolution Layers compute the dot product between the local region of the input image and the weight matrix of the filter, and the filter slides over the entire map, repeating the dot product operation. The convolution kernel (filter) size of each convolution layer is 4×4 and the number of kernel of three convolution layers is 64, 128 and 256, respectively. Generally, Batch normalization is cascaded after the convolution layer to prevent overfitting and accelerate deep network training. An activation function, which must be highly nonlinear, follows the Batch normalization. This paper employs Restricted Linear Units (ReLU), which describes the nonlinear input output. The pooling layer followed after ReLU, also known as downsampling, is designed to reduce data volume while preserving useful information. The Max pooling is used and Pooling size is set at 2×2 in the whole deep convolution network.

Figure 7. The network architecture of our recognition model.

Figure 8. The architecture of DCNN.

The Fully Connected (FC) layer flattens the features extracted by convolution Layers and computes the probability of different classes. Like traditional neural networks, all nodes in adjacent FC layers are connected by weights. In the proposed recognition model, an FC layer is followed after convolution layers and the size of the FC layer is 512. The ReLu activation function is used and local response normalization is applied after each layer. Dropouts of 0.5 are applied in the four FC layers. In the end, a Softmax output layer outputs the predicted label.

The proposed recognition model is implemented based on the Keras platform and trained from scratch because no pre-trained models can be used to our proposed model. Adam [46] is used to optimize the parameters of the model with respect to the loss function. The batch size is initially set at 128, learning rate is set at 0.0001, and the number of epochs is set at 500.

3. Results

3.1. Gestures Set

The proposed method is used to evaluated five finger gestures, including finger close, finger open, finger clockwise circle, finger counterclockwise circle, and finger slide, as shown in Figure 9. The selected gestures are all finger movements, excluding big movements of the palms or arms. These gestures are common gestures for people to control devices in daily life.

Figure 9. Illustrations of five finger gestures: (a) finger close, (b) finger clockwise circle, (c) finger slide, (d) finger open, (e) finger counterclockwise circle.

3.2. Data Acquisition

Training a deep neural network requires a large amount of training data containing enough variations of gestures. Six subjects including four males and two females were recruited to perform the five designated finger gestures. In order to obtain as many as variations, only minimal instruction on how to execute these gestures were given. Each feature map is marked with a class label. Each subject was asked to perform each gesture for 50 times. Thus, a total of $5 \times 6 \times 50 = 1500$ samples were recorded. The dataset was expanded to 4500 by time folding and superimposing Gaussian noise on feature maps. Finally, the dataset of total 4500 samples were applied to the experiments as raw input.

3.3. Discussion Parameters of the Recognition Model

(1) *Number of Convolution layers*: To further explore the impact of the number of convolution layers on the recognition accuracy, recognition models under configurations of different numbers of convolution layers are tested. The number of convolution layers is set at 2, 3, 4 and 5, respectively, in each model, and the other parameters of the models are consistent with the initial settings. The recognition accuracies of different models are as shown in Table 2. It can be seen from the table that the recognition rate is the highest when the number of layers is three. However, when the number of layers increase to 4 and 5, the recognition rate falls because of over-fitting. The reason is that the increase of convolution layers leads to over parameterization of the models, and consequently making it difficult to be well trained by relatively number of samples. So, the number of convolution layer is fixed at 3.

Table 2. Recognition accuracies vs. Number of convolution layers.

The Number of Convolution Layers	$3 \times 3 \times 64$	$3 \times 3 \times 64$	$3 \times 3 \times 64$	$3 \times 3 \times 64$
	$3 \times 3 \times 128$	$3 \times 3 \times 128$	$3 \times 3 \times 128$	$3 \times 3 \times 128$
		$3 \times 3 \times 256$	$3 \times 3 \times 256$	$3 \times 3 \times 256$
			$3 \times 3 \times 512$	$3 \times 3 \times 512$
				$3 \times 3 \times 1024$
Recognition accuracies	87.33%	95.48%	93.70%	92.67%

(2) *Size of the Convolution Kernel*: In this experiment, the performance of recognition model is tested under different sizes of the convolution kernel. The sizes of the convolution kernel in all convolution layers are set as 3×3 , 4×4 and 5×5 , respectively. Table 3 shows the recognition accuracies of different convolution kernel sizes. In Table 3, it can be found that the recognition accuracy increases with the size of the kernel, and the number of parameters needed to be trained also increases. Therefore, the kernel with size of 4×4 is selected.

Table 3. Recognition accuracies vs. Size of the convolution kernel.

The Size of the Convolution Kernel	$3 \times 3 \times 64$	$4 \times 4 \times 64$	$5 \times 5 \times 64$
	$3 \times 3 \times 128$	$4 \times 4 \times 128$	$5 \times 5 \times 128$
	$3 \times 3 \times 256$	$4 \times 4 \times 128$	$5 \times 5 \times 256$
Recognition accuracies	93.63%	96.00%	96.00%
Training parameters	173,011,333	173,298,501	173,667,717

(3) *Number of Convolution Kernels*: The more the number of convolution kernels is, the more features are extracted from the convolution layer, and the easier it is to overfit. On the contrary, the fewer the number of convolution kernels is, the fewer the features extracted from the convolution layer are, and the more prone to under-fitting. The experiment's aim is to find the suitable number of convolution kernels. The Table 4 illustrates recognition accuracies of four different configurations. The result of Table 4 shows that the first configuration has a comparable accuracy, but it has minimum training parameters and minimum training cost.

Table 4. Recognition accuracies vs. Number of convolution kernels.

The Number of Convolution Kernels	$4 \times 4 \times 16$	$4 \times 4 \times 32$	$4 \times 4 \times 64$	$4 \times 4 \times 128$
	$4 \times 4 \times 32$	$4 \times 4 \times 64$	$4 \times 4 \times 128$	$4 \times 4 \times 256$
	$4 \times 4 \times 64$	$4 \times 4 \times 128$	$4 \times 4 \times 256$	$4 \times 4 \times 512$
Recognition accuracies	96.78%	97.11%	96.00%	95.89%

(4) *Number of FC layers*: In order to optimize architecture of the recognition model, the impact of number of the FC layers is examined. Four different configurations of FC layers are tested. Table 5 shows the results under different numbers of FC layers. From the table, it can be concluded that one FC layer is enough.

Table 5. Recognition accuracies vs. Number of FC layers.

The Number of FC Layers	4096			
	1024	1024		
	512	512	512	
	256	256	256	256
Recognition accuracies	96.77%	97.33%	96.11%	96.67%

(5) *Size of the FC layer*: The impact of the size of the FC layer is also examined by our experiments. Table 6 shows the recognition accuracies with different FC layer sizes vary from 128 to 1024. From the table, it can be known that the performance enhances the increase of the FC layer size.

Table 6. Recognition accuracies vs. Size of the FC layer.

The Size of the FC Layer	128	256	512	1024
Recognition accuracies	94.67%	96.00%	96.78%	97.44%

(6) *Number of Epochs*: The number of epochs is not enough and the model is not trained to be optimal. Too many epochs will require a long training time. Based on the chosen parameters in previous sections, the recognition model is trained with different numbers of epochs of 300, 500, 800

and 1000. Table 7 shows the results. The results indicate that the recognition accuracy increases with the number of epochs, and the accuracy barely increased when the number of epochs reached 800. Therefore, the number of epochs is set to 800 in subsequent experiments.

Table 7. Recognition accuracies vs. Number of epochs.

The Number of Epochs	300	500	800	1000
Recognition accuracies	94.55%	96.00%	96.78%	97.44%

3.4. Performance Evaluation

For the gesture recognition training and validation experiments, the dataset is randomly divided into two parts, 70% for training and 30% for test, and a standard k-fold leave-one-subject approach is used, where k is five in the experiments. Table 8 shows the result of 5-fold cross validation. The average recognition accuracy is 97.11%. In terms of classification error analysis, the confusion matrix for all 5-folds is presented in Figure 10. The table shows that the accuracies of all finger gestures are above 96%.

Es Ac	a	b	c	d	e
a	96.28%	0.32%	1.55%	1.85%	0.00%
b	0.77%	97.57%	0.13%	0.53%	1.00%
c	0.44%	0.22%	98.24%	0.88%	0.22%
d	1.34%	0.37%	0.78%	97.41%	0.11%
e	0.00%	0.66%	1.35%	1.87%	96.12%

Figure 10. The confusion matrix of accuracy.

Table 8. The result of 5-fold cross validation.

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
98.67%	96.22%	97.67%	96.22%	96.78%	97.11%

4. Conclusions

In this paper, a new method of ultrasonic finger gesture recognition is proposed. Firstly, a hardware structure of ultrasonic active sensing, which uses only two ultrasonic transducers, is constructed. Second, based on the micor-Doppler effect of a moving finger, the received echos from finger gestures are processed in preprocessing stage, including sampling, bandpass filtering, IQ demodulation and lowpass filtering. In micro-Doppler processing, coherent processing and FFT are used to obtain high resolution micro-Doppler feature maps of finger gestures resulting from finger motions. Third, a recognition model based on convolution neural networks is built. Some experiments are designed to optimize some parameters of the model. The results of 5-fold cross validation show that our proposed method can recognize five finger gestures and achieve an average accuracy of 96.51%; the accuracy of each finger gesture is more than 96%.

Author Contributions: Q.Z. wrote the paper. Q.Z. and Z.K. investigated the method and designed the experiment. Q.Z. and S.W. performed the experiment and analyzed the data. J.Y. revised the paper.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to Han Jia for providing powerful data acquisition equipment, to Zhi Chen for his suggestions to this paper. The authors would like to thank the reviewers for the careful review and valuable suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- McNeill, D. *Gesture and Thought*; University of Chicago Press: Chicago, IL, USA, 2005. [\[CrossRef\]](#)
- Ling, K.; Dai, H.; Liu, Y.; Liu, A.X. UltraGesture: Fine-Grained Gesture Sensing and Recognition. In Proceedings of the 2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), Hong Kong, China, 11–13 June 2018; pp. 1–9. [\[CrossRef\]](#)
- Weissmann, J.; Salomon, R. Gesture Recognition for Virtual Reality Applications Using Data Gloves and Neural Networks. In Proceedings of the IJCNN'99, International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339), Washington, DC, USA, 10–16 July 1999; Volume 3, pp. 2043–2046. [\[CrossRef\]](#)
- Parate, A.; Chiu, M.C.; Chadowitz, C.; Ganesan, D.; Kalogerakis, E. RisQ: Recognizing Smoking Gestures with Inertial Sensors on a Wristband. In Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services—MobiSys '14, Bretton Woods, NH, USA, 16–19 June 2014; pp. 149–161. [\[CrossRef\]](#)
- Lu, W.; Tong, Z.; Chu, J. Dynamic Hand Gesture Recognition With Leap Motion Controller. *IEEE Signal Process. Lett.* **2016**, *23*, 1188–1192. [\[CrossRef\]](#)
- Plouffe, G.; Cretu, A.M. Static and Dynamic Hand Gesture Recognition in Depth Data Using Dynamic Time Warping. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 305–316. [\[CrossRef\]](#)
- Kim, Y.; Ling, H. Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1328–1337. [\[CrossRef\]](#)
- Javier, R.J.; Kim, Y. Application of Linear Predictive Coding for Human Activity Classification Based on Micro-Doppler Signatures. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1831–1834. [\[CrossRef\]](#)
- Kalgaonkar, K.; Raj, B. One-Handed Gesture Recognition Using Ultrasonic Doppler Sonar. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 1889–1892. [\[CrossRef\]](#)
- Gupta, S.; Morris, D.; Patel, S.; Tan, D. SoundWave: Using the Doppler Effect to Sense Gestures. In Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems—CHI '12, Austin, TX, USA, 5–10 May 2012; p. 1911. [\[CrossRef\]](#)
- Zhang, Y.; Harrison, C. Tomo: Wearable, Low-Cost Electrical Impedance Tomography for Hand Gesture Recognition. In Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology—UIST '15, Charlotte, NC, USA, 11–15 November 2015; pp. 167–173. [\[CrossRef\]](#)
- Xie, R.; Cao, J. Accelerometer-Based Hand Gesture Recognition by Neural Network and Similarity Matching. *IEEE Sens. J.* **2016**, *16*, 4537–4545. [\[CrossRef\]](#)
- Chin-Shyurng, F.; Lee, S.E.; Wu, M.L. Real-Time Musical Conducting Gesture Recognition Based on a Dynamic Time Warping Classifier Using a Single-Depth Camera. *Appl. Sci.* **2019**, *9*, 528. [\[CrossRef\]](#)
- Vaitkevičius, A.; Taroza, M.; Blažauskas, T.; Damaševičius, R.; Maskeliūnas, R.; Woźniak, M. Recognition of American Sign Language Gestures in a Virtual Reality Using Leap Motion. *Appl. Sci.* **2019**, *9*, 445. [\[CrossRef\]](#)
- Dahl, T.; Ealo, J.L.; Bang, H.J.; Holm, S.; Khuri-Yakub, P. Applications of airborne ultrasound in human–computer interaction. *Ultrasonics* **2014**, *54*, 1912–1921. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wang, Y.; Liu, J.; Chen, Y.; Gruteser, M.; Yang, J.; Liu, H. E-Eyes: Device-Free Location-Oriented Activity Identification Using Fine-Grained WiFi Signatures. In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking—MobiCom '14, Maui, HI, USA, 7–11 September 2014; pp. 617–628. [\[CrossRef\]](#)
- Cagliyan, B.; Gurbuz, S.Z. Micro-Doppler-Based Human Activity Classification Using the Mote-Scale BumbleBee Radar. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2135–2139. [\[CrossRef\]](#)
- Wan, Q.; Li, Y.; Li, C.; Pal, R. Gesture Recognition for Smart Home Applications Using Portable Radar Sensors. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; pp. 6414–6417. [\[CrossRef\]](#)
- Abdelnasser, H.; Youssef, M.; Harras, K.A. WiGest: A Ubiquitous WiFi-Based Gesture Recognition System. In Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM), Hong Kong, China, 26 April–1 May 2015; pp. 1472–1480. [\[CrossRef\]](#)

20. Molchanov, P.; Gupta, S.; Kim, K.; Pulli, K. Short-Range FMCW Monopulse Radar for Hand-Gesture Sensing. In Proceedings of the 2015 IEEE Radar Conference (RadarCon), Arlington, VA, USA, 10–15 May 2015; pp. 1491–1496. [\[CrossRef\]](#)
21. Kim, Y.; Toomajian, B. Hand Gesture Recognition Using Micro-Doppler Signatures With Convolutional Neural Network. *IEEE Access* **2016**, *4*, 7125–7130. [\[CrossRef\]](#)
22. Wang, S.; Song, J.; Lien, J.; Poupyrev, I.; Hilliges, O. Interacting with Soli: Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology—UIST '16, Tokyo, Japan, 16–19 October 2016; pp. 851–860. [\[CrossRef\]](#)
23. Tan, S.; Yang, J. WiFinger: Leveraging Commodity WiFi for Fine-Grained Finger Gesture Recognition. In Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing—MobiHoc '16, Paderborn, Germany, 5–8 July 2016; pp. 201–210. [\[CrossRef\]](#)
24. Dekker, B.; Jacobs, S.; Kossen, A.; Kruithof, M.; Huizing, A.; Geurts, M. Gesture Recognition with a Low Power FMCW Radar and a Deep Convolutional Neural Network. In Proceedings of the 2017 European Radar Conference (EURAD), Nuremberg, Germany, 11–13 October 2017; pp. 163–166. [\[CrossRef\]](#)
25. Sun, Y.; Fei, T.; Schliep, F.; Pohl, N. Gesture Classification with Handcrafted Micro-Doppler Features Using a FMCW Radar. In Proceedings of the 2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), Munich, Germany, 15–17 April 2018; pp. 1–4. [\[CrossRef\]](#)
26. Zhang, Z.; Tian, Z.; Zhou, M. Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor. *IEEE Sens. J.* **2018**, *18*, 3278–3289. [\[CrossRef\]](#)
27. Lien, J.; Gillian, N.; Karagozler, M.E.; Amihoud, P.; Schwesig, C.; Olson, E.; Raja, H.; Poupyrev, I. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Trans. Graph.* **2016**, *35*, 1–19. [\[CrossRef\]](#)
28. Yang, Q.; Tang, H.; Zhao, X.; Li, Y.; Zhang, S. Dolphin: Ultrasonic-Based Gesture Recognition on Smartphone Platform. In Proceedings of the 2014 IEEE 17th International Conference on Computational Science and Engineering, Chengdu, China, 19–21 December 2014; pp. 1461–1468. [\[CrossRef\]](#)
29. Pittman, C.; Wisniewski, P.; Brooks, C.; LaViola, J.J. Multiwave: Doppler Effect Based Gesture Recognition in Multiple Dimensions. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA '16, San Jose, CA, USA, 7–12 May 2016; pp. 1729–1736. [\[CrossRef\]](#)
30. Ruan, W.; Sheng, Q.Z.; Yang, L.; Gu, T.; Xu, P.; Shangguan, L. AudioGest: Enabling Fine-Grained Hand Gesture Detection by Decoding Echo Signal. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing—UbiComp '16, Heidelberg, Germany, 12–16 September 2016; pp. 474–485. [\[CrossRef\]](#)
31. Li, X.; Dai, H.; Cui, L.; Wang, Y. SonicOperator: Ultrasonic Gesture Recognition with Deep Neural Network on Mobiles. In Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 4–8 August 2017; pp. 1–7. [\[CrossRef\]](#)
32. Liu, Q.; Yang, W.; Xu, Y.; Hu, Y.; He, Q.; Huang, L. DopGest: Dual-Frequency Based Ultrasonic Gesture Recognition. In Proceedings of the 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Guangzhou, China, 8–12 October 2018; pp. 293–300. [\[CrossRef\]](#)
33. Przybyla, R.J.; Tang, H.; Guedes, A.; Shelton, S.E.; Horsley, D.A.; Boser, B.E. 3D Ultrasonic Rangefinder on a Chip. *IEEE J. Solid-State Circ.* **2015**, *50*, 320–334. [\[CrossRef\]](#)
34. Chen, H.; Ballal, T.; Saad, M.; Al-Naffouri, T.Y. Angle-of-Arrival-Based Gesture Recognition Using Ultrasonic Multi-Frequency Signals. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August–2 September 2017; pp. 16–20. [\[CrossRef\]](#)
35. Saad, M.; Bleakley, C.J.; Nigram, V.; Kettle, P. Ultrasonic Hand Gesture Recognition for Mobile Devices. *J. Multimodal User Interfaces* **2018**, *12*, 31–39. [\[CrossRef\]](#)
36. Yun, S.; Chen, Y.C.; Qiu, L. Turning a Mobile Device into a Mouse in the Air. In Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services—MobiSys '15, Florence, Italy, 18–22 May 2015; pp. 15–29. [\[CrossRef\]](#)

37. Nandakumar, R.; Iyer, V.; Tan, D.; Gollakota, S. FingerIO: Using Active Sonar for Fine-Grained Finger Tracking. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI '16, San Jose, CA, USA, 7–12 May 2016; pp. 1515–1525. [[CrossRef](#)]
38. Wang, W.; Liu, A.X.; Sun, K. Device-Free Gesture Tracking Using Acoustic Signals. In Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking—MobiCom '16, New York, NY, USA, 3–7 October 2016; pp. 82–94. [[CrossRef](#)]
39. Yun, S.; Chen, Y.C.; Zheng, H.; Qiu, L.; Mao, W. Strata: Fine-Grained Acoustic-Based Device-Free Tracking. In Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services—MobiSys '17, Niagara Falls, NY, USA, 19–23 June 2017; pp. 15–28. [[CrossRef](#)]
40. Zediker, M.S.; Rice, R.R.; Hollister, J.H. Method for Extending Range and Sensitivity of a Fiber Optic Micro-Doppler Ladar System and Apparatus Therefor. U.S. Patent 5,847,817, 8 December 1998.
41. Chen, V.C.; Li, F.; Ho, S.S.; Wechsler, H. Micro-Doppler Effect in Radar: Phenomenon, Model, and Simulation Study. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 2–21. [[CrossRef](#)]
42. Yang, Y.; Lei, J.; Zhang, W.; Lu, C. Target Classification and Pattern Recognition Using Micro-Doppler Radar Signatures. In Proceedings of the Seventh ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD'06), Las Vegas, NV, USA, 19–20 June 2006; pp. 213–217. [[CrossRef](#)]
43. Chen, V.C. Detection and Analysis of Human Motion by Radar. In Proceedings of the 2008 IEEE Radar Conference, Rome, Italy, 26–30 May 2008; pp. 1–4. [[CrossRef](#)]
44. Jeong, J.J.; Choi, H. An impedance measurement system for piezoelectric array element transducers. *Measurement* **2017**, *97*, 138–144. [[CrossRef](#)]
45. Richards, M.A.; Scheer, J.; Holm, W.A.; Melvin, W.L. *Principles of Modern Radar*; Citeseer, 2010.
46. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *Comput. Sci.* **2014**, arXiv:1412.6980.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).