Filterbank-Based Micro-Doppler Spectrogram Transformation for Human Activity Recognition

Do-Hyun Park
Department of Electrical and Electronics Engineering
Pusan National University
Busan, Republic of Korea
dohpark@pusan.ac.kr

Hyoung-Nam Kim
School of Electrical and Electronics Engineering
Pusan National University
Busan, Republic of Korea
hnkim@pusan.ac.kr

***Abstract*—Human activity recognition (HAR) utilizing radar sensors has garnered significant attention due to its ability to robustly recognize activities in various environments without infringing on privacy concerns. In prior radar-based HAR studies, recognition has been predominantly conducted using spectrograms generated from received signals. In this paper, we propose a novel time-frequency domain signal representation method specifically designed for the effective analysis of micro-Doppler signatures arising from human activities. The proposed method applies a filterbank capable of nonlinearly transforming the frequency bands of conventional spectrograms, which possess linear frequency bands. This transformation allows for a more precise analysis of micro-Doppler patterns generated by human activities. We evaluated the recognition performance by applying convolutional neural networks to datasets that utilized the proposed filterbank-based spectrograms in comparison with those employing conventional spectrograms. Our results demonstrate that the modified spectrograms achieved up to 5.02% improvement in recognition performance.**

***Index Terms*—human activity recognition, micro-Doppler, filterbank, deep learning**

I. Introduction

Recently, radar systems have been actively utilized in tracking moving humans [1], recognizing various human activities [2], monitoring vital signs [3], and recognizing hand gestures [4]. Among these, radar-based human activity recognition (HAR) holds significant societal value. For instance, as illustrated in Fig. 1, HAR systems can be applied in elderly care to detect life-threatening activities such as falls. Traditional HAR systems primarily acquire data through camera, LiDAR, or wearable sensors. However, camera-based surveillance suffers from low resolution in low-light conditions and poses privacy concerns. LiDAR sensors are easily affected by environmental conditions, and wearable sensors can be uncomfortable for users due to the sensation of wearing them. In contrast, HAR utilizing radar sensors has emerged as an important research topic due to its ability to overcome these limitations effectively.

Generally, radar-based HAR is achieved by exploiting the features of human movement embedded in the spectrograms, which have a time-Doppler domain. Numerous studies have focused on using spectrograms for HAR because spectrograms can capture the movement characteristics of individual body parts [2]. Traditional radar-based HAR typically employs manually extracted features from spectrograms and uses a support vector machine [5] or multi-layer perceptron [6] as the classifier. However, this manual feature extraction approach often struggles to capture highly discriminative feature information from spectrograms, leading to reduced accuracy, especially when dealing with low-power target echo signals. To address these limitations, methods utilizing deep-learning models have been proposed, where features within spectrograms are automatically extracted and classified. In these studies, convolutional neural networks (CNNs) are employed, using convolution layers to effectively extract features embedded within the two-dimensional spectrogram data [2].

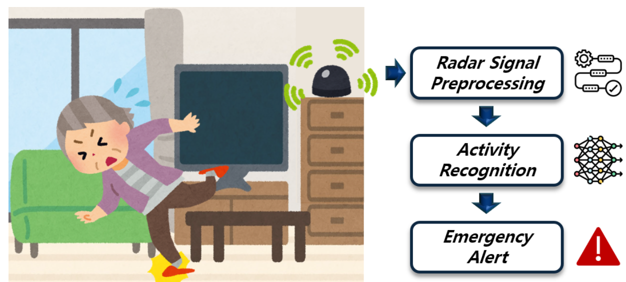


Fig. 1. Processes of the HAR system for elderly care.

While spectrograms are useful for visualizing various frequency components, they have limitations in analyzing the delicate Doppler patterns associated with human activities. The short-time Fourier transform (STFT) used in spectrogram generation applies a linear scale in time-frequency analysis, which makes it challenging to capture detailed features arising from micro-Doppler effects caused by subtle vibrations or rotations of human moving.

To overcome these limitations, this study proposes a novel time-frequency representation technique optimized for analyzing the distinctive micro-Doppler patterns of human activities. The proposed method applies a filterbank that adjusts the frequency resolution of conventional spectrograms to capture micro-Doppler features more precisely. Based on data obtained from real-world experiments, we validate the superiority of the

979-8-3315-0694-0/25/$31.00 ©2025 IEEE

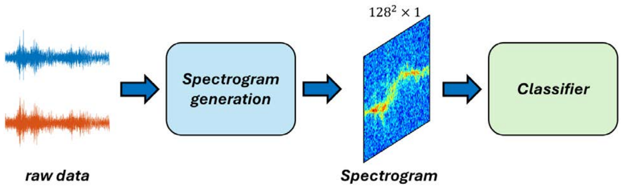


Fig. 2. Flowchart of the spectrogram-based conventional HAR system.

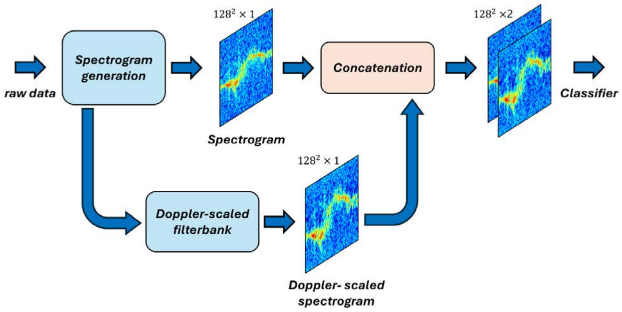


Fig. 3. Flowchart of the proposed HAR system.

proposed method by comparing the recognition performance between deep-learning models that utilize conventional spectrograms and those that use the proposed spectrogram data.

II. Spectrogram-Based HAR method

The raw data obtained from radar sensors consists of complex in-phase/quadrature (I/Q) time-series data. The first step in generating a spectrogram involves creating a range profile from the raw I/Q data. This profile shows the distance of the target echo over time. Since the received signal in a radar system contains both target echo signals and static clutter, it is necessary to remove the DC component from the received signal. In this paper, we utilize a moving target indication (MTI) filter for static clutter suppression. After the MTI filter is applied to the range profile, the STFT is applied to generate the spectrogram. As shown in Fig. 2, the received signal is converted into a two-dimensional spectrogram, which is then input into a deep learning-based classification model to determine human activity.

III. Proposed HAR Method

Unlike traditional recognition methods that solely rely on conventional spectrograms, we propose a methodology that utilizes modified spectrograms. Fig. 3 illustrates an overview of the proposed system. Our HAR system is composed of four blocks: 1) spectrogram generation, 2) Doppler-scaled spectrogram generation via filterbank, 3) data concatenation, and 4) activity recognition using a deep learning-based classification model.

The Doppler-scaled spectrogram can be generated by nonlinearly transforming the frequency domain of the original spectrogram. To achieve this, a Doppler-scaled filterbank is applied to the spectrogram, converting the linear frequency spectrum into a Doppler-scaled frequency spectrum. Typically,

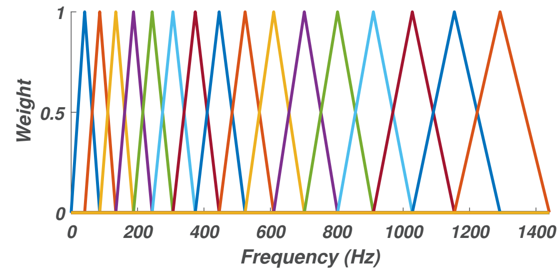


Fig. 4. Example of the Doppler-scaled filterbank.

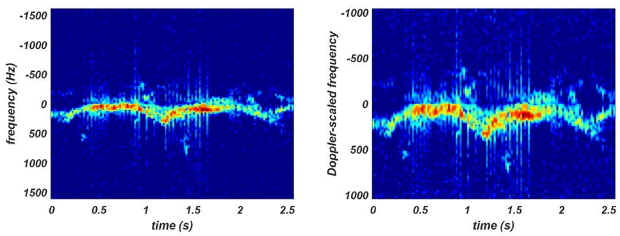
the key features of Doppler frequencies resulting from human activities are concentrated in the lower frequency range. In this study, the proposed Doppler scale utilizes a frequency metric that allows for a more detailed examination of frequency regions relevant to activity recognition. The proposed Doppler scale is mathematically modeled to account for nonlinear frequency perception and is defined as follows:

$$D(f) = \frac{f_c}{\log_{10}(2)} \log_{10}\left(1 + \frac{f}{f_c}\right), \quad (1)$$

where f is the original frequency (in Hz) and $D(f)$ is the corresponding Doppler-scaled frequency. The parameter f_c is the corner frequency, which determines the frequency range for detailed analysis. By applying the Doppler scale defined in (1), the frequency region below f_c is analyzed linearly, providing higher resolution, while the region above f_c is analyzed nonlinearly with less detail.

A filterbank is generated to apply this nonlinear Doppler-scale transformation to the spectrogram. For filterbank generation, the minimum frequency f_{min} and maximum frequency f_{max} in the original spectrogram are transformed into $D(f_{min})$ and $D(f_{max})$ using (1). The range between $D(f_{min})$ and $D(f_{max})$ is then divided into n equally spaced intervals, where n represents the number of bands used in the Doppler-scale transformation, which also determines the size of the resulting Doppler-scaled spectrogram. The n divided values are converted back to the original frequency domain to determine the center frequencies of each band. Filters are then generated for each band based on these Doppler-scaled-transformed center frequencies. Each filter is shaped as a triangular function, linearly increasing from 0 to 1 between the previous center frequency and the current center frequency, and linearly decreasing from 1 to 0 between the current center frequency and the next center frequency. Fig. 4 shows an example of these filters when the corner frequency f_c is 500 Hz and n is 16.

After generating a filterbank for frequencies ranging from 0 Hz to half of the sampling frequency, this filterbank is applied symmetrically to both the positive and negative frequency range of the original spectrogram simultaneously. Each frequency component of the spectrogram is weighted by the corresponding filter in its frequency band, resulting in the Doppler-scaled spectrogram. Fig. 5 shows examples of spectrograms obtained from a limping action. The Doppler-



(a) (b)

Fig. 5. Examples of (a) conventional spectrogram and (b) Doppler-scaled spectrogram with f_c of 500 Hz and n of 64.

scaled spectrogram example shows that the resolution increases in regions where micro-Doppler frequencies exist while it decreases in other regions. As a result, the micro-Doppler effects generated by human activities become more distinct and pronounced.

In the proposed HAR method, we utilize both the original spectrogram and the Doppler-scaled spectrogram to perform HAR. To efficiently use these two types of time-frequency domain data, we concatenate them into a two-channel input, which is then fed into the deep learning-based classification model.

IV. Performance Analysis

To analyze the performance of the proposed method, we collected data that captured various daily activities. The data was acquired using a Texas Instruments AWR1642BOOST frequency-modulated continuous wave (FMCW) radar. Participants performed a total of 6 different daily activities (falling, walking, running, picking up objects, limping, sitting) at various distances within 4m from the radar. To ensure diversity in the data, 5 participants were involved in the data collection process, and a total of 2,250 activity data were gathered. For training the deep learning-based HAR model, data from four participants was used, while data from one participant not included in the training dataset was used for testing.

To compare the performance of the spectrogram-based HAR method with the novel HAR method using the two-channel spectrogram dataset, we selected VGG16, VGG19 [7], and ResNet18 [8] as classification models. Table I presents the HAR accuracy of various classification models using different datasets. When using only spectrograms for classification, the average accuracy across the three models was 92.7%. The datasets using the proposed Doppler-scaled spectrograms were created with various f_c values. The average classification accuracies when f_c was set to 400, 500, and 600 were 92.1% for ResNet18, 95.5% for VGG16, and 97.1% for VGG19. These results show that the models using the Doppler-scaled spectrogram dataset achieved superior classification performance compared to the models using only spectrograms, with improvements of 0.52% for ResNet18, 3.35% for VGG16, and 2.72% for VGG19. These performance analysis results indicate that the proposed Doppler-scaled spectrogram can significantly enhance the performance of deep learning-based HAR models.

TABLE I
Recognition Accuracy of HAR models on various datasets

Model	Spectrogram	Proposed		
	Acc.	Acc. (f_c = 400 Hz)	Acc. (f_c = 500 Hz)	Acc. (f_c = 600 Hz)
ResNet18	91.54	**93.42**	91.54	91.22
VGG16	92.16	**93.42**	**95.92**	**97.18**
VGG19	94.36	**97.18**	**97.81**	**96.24**

V. Conclusion

In this paper, we proposed a HAR method that utilizes Doppler-scaled spectrograms generated by applying a filterbank to conventional spectrograms. The Doppler-scaled spectrogram offers enhanced resolution in regions where human activities produce micro-Doppler frequencies, thereby facilitating the efficient extraction of micro-Doppler signatures. The proposed method employs a two-channel input by concatenating the conventional spectrogram and Doppler-scaled spectrogram to enhance HAR performance. Experimental results show that the recognition accuracy improved by up to 5.02% compared to HAR models using only conventional spectrograms. This result demonstrates that the proposed Doppler-scaled spectrogram can significantly contribute to HAR research.

This study proposed a spectrogram transformation method where the corner frequencies were selected manually. In future work, research will focus on developing a method to automatically select the optimal corner frequencies for spectrogram transformation.

References

[1] Y. He, P. Aubry, F. L. Chevalier and A. Yarovoy, "Decentralised tracking for human target in multistatic ultra-wideband radar," *IET Radar, Sonar & Navigation*, vol. 8, no. 9, pp. 1215-1223, Dec. 2014.

[2] S. Zhu, R. G. Guendel, A. Yarovoy and F. Fioranelli, "Continuous Human Activity Recognition With Distributed Radar Sensor Networks and CNN–RNN Architectures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, Jul. 2022, Art no. 5115215.

[3] T. Sakamoto, P. J. Aubry, S. Okumura, H. Taki, T. Sato and A. G. Yarovoy, "Noncontact measurement of the instantaneous heart rate in a multi-person scenario using X-band array radar and adaptive array processing," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 2, pp. 280-293, Jun. 2018.

[4] Y. Sun, T. Fei, X. Li, A. Warnecke, E. Warsitz and N. Pohl, "Real-time radar-based gesture detection and recognition built in an edge-computing platform," *IEEE Sensors Journal*, vol. 20, no. 18, pp. 10706-10716, 2020.

[5] H. Li, A. Shrestha, H. Heidari, J. L. Kernec and F. Fioranelli, "Activities Recognition and Fall Detection in Continuous Data Streams Using Radar Sensor," *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, Nanjing, China, 2019.

[6] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using an artificial neural network," *2008 IEEE Antennas and Propagation Society International Symposium*, San Diego, CA, USA, Jul. 2008, pp. 1–14.

[7] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *3rd International Conference on Learning Representations (ICLR 2015)*, San Diego, CA, USA, May 2015, pp. 1–4.

[8] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.