Check for updates

Poster Abstract: Person Identification Under Heavy Occlusions Using mmWave Radar

Tao Wang
Harbin Institute of Technology
Shenzhen, Guangdong, China
hitszwangtao@gmail.com

Yang Zhao
Harbin Institute of Technology
Shenzhen, Guangdong, China
yang.zhao@hit.edu.cn

Jie Liu
Harbin Institute of Technology
Shenzhen, Guangdong, China
jieliu@hit.edu.cn

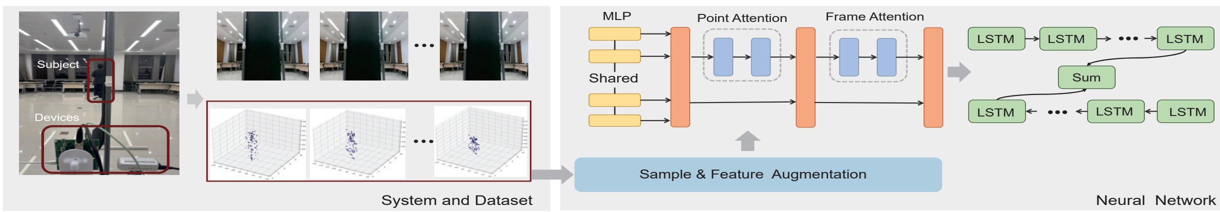

Figure 1: Overview of PID under heavy occlusion using mmWave radar.

ABSTRACT

We propose *mmWave-ocPID*, a person identification (PID) method with millimeter-wave radar to identify individuals even when they are heavily occluded by obstacles. We collect a multi-modal dataset comprising mmWave radar point clouds and RGB images obtained from 9 human subjects, with over 180,000 frames for each modality. The *mmWave-ocPID* prototype employs a novel Neural Network integrated with two augmentation strategies for learning. Our initial experimental results show that *mmWave-ocPID* can achieve high identification accuracy, even when most of the human body of an individual is occluded in a controlled environment.

KEYWORDS

Person Identification, Millimeter Wave Radar,Occluded Conditions

ACM Reference Format:
Tao Wang, Yang Zhao, and Jie Liu. 2023. Poster Abstract: Person Identification Under Heavy Occlusions Using mmWave Radar. In *The 21st ACM Conference on Embedded Networked Sensor Systems (SenSys '23), November 12–17, 2023, Istanbul, Turkiye.* ACM, Istanbul, Turkiye, 2 pages. https://doi.org/10.1145/3625687.3628404

1 INTRODUCTION

Person identification (PID) aims to distinguish individuals based on distinct features extracted from their body characteristics, such as facial appearance, gait, etc. PID has various applications in security, smart environments, and many other areas. Traditional PID systems relying on visual sensors assume that the entire human body or part of an individual is visible [1], yet in real-world home and office scenarios, people are often occluded by various obstacles, such as chairs and other furniture. In such situations, visual sensor-based PID methods perform poorly or fail completely.

As an alternative sensing modality to visual sensors, millimeter wave (mmWave) radar has several advantages, such as its ability to function in low-light conditions, adverse weather, and detect targets by penetrating non-metallic objects. Previous studies show mmWave radar's capability of identifying individuals in open space, without much occlusion between people and sensors [4]. In this paper, we show research efforts in PID using mmWave radar even under heavily occluded conditions, i.e., most of the human body is occluded by obstacles.

As shown in Figure 1, we build a multi-sensor system and collect a new dataset for mmWave radar-based PID under various occlusions, which we call the **mmWave-ocPID** dataset. Our dataset is accessible on **https://zenodo.org/record/8377254**. The dataset consists of mmWave radar point clouds and synchronized RGB image sequences, collected under partially or heavily occluded conditions. The radar point clouds roughly outline a person's silhouette and provide velocity information, emphasizing mmWave radar's potential in identifying obstructed individuals. Thus, we propose a novel neural network with two modules: (1) A feature extractor with two attention mechanisms takes a sequence of radar point cloud frames as input, (2) A spatio-temporal network to extract complete gait features. Because point clouds are sparse, we propose two data augmentation strategies to improve PID accuracy. Finally, experimental results show that our data augmentation method can increase accuracy by 10.72%, and the overall method can achieve an accuracy of 94.17% on average.

2 SYSTEM AND METHOD

2.1 System

To conduct PID under diverse occlusion conditions, We use a COTS mmWave radar and an RGB-D camera to capture the sequential data. Specifically, We utilize a TI IWR6843ISK-ODS radar for the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SenSys '23, November 12–17, 2023, Istanbul, Turkiye
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0414-7/23/11…$15.00
https://doi.org/10.1145/3625687.3628404

transmission and reception of radar signals; and we use an Intel RealSense D435 camera to collect RGB images with 424×240 resolution. The radar achieves a range resolution of $4.5cm$ and a maximum unambiguous range of $11.52m$. It can measure a maximum radial velocity of $3m/s$ with a resolution of $0.049m/s$. Sampling rate for radar frames and images is $15Hz$. The raw signals are transported from the radar to a PC using a DCA1000EVM data capture card. Our dataset comprises 9 recruited individuals instructed to walk behind the obstacle in an inbound/outbound manner, each subject completing 5 minutes of walking. Our experiment includes three types of obstacles: paper box, sponge and plant. Additionally, we deploy obstacles at three different orientations: $0°$, $45°$, and $90°$. To generate point clouds from raw radar signals, we adopt the signal processing method proposed by [2].

2.2 Method

To enhance PID accuracy, we implement two augmentation strategies on the original point clouds before feeding them to the neural network. The *sample augmentation* individually inverts the sign of each component of radar points (x, y, z, v), where (x, y, z) represents the Cartesian coordinates and v is the recorded velocity. Then, the *feature augmentation* computes the differences between a point in the current frame and three closest points in the previous frame using Euclidean distance. These difference values are concatenated with original components of the radar point to increase features. Our *sample augmentation* strategy can enlarge the amount of training data 5 times larger than originally acquired data. Our *feature augmentation* effectively exploits the spatio-temporal relationship between consecutive point cloud frames.

The proposed network comprises two main modules. The *feature extractor* module maps the augmented point cloud sequence into a high-dimensional feature space using three shared linear layers with input sizes of 16, 64, 256, respectively. To aggregate the global feature for a point cloud frame, we employ a point attention mechanism. This mechanism is constructed with two linear layers, having input sizes of 256 and 50, respectively, and it calculates a weighted sum of all points within the current frame. Meanwhile, we employ a frame attention mechanism, structured similarly to the point attention mechanism, to dynamically adjust the contribution of distinct point cloud frames. This involves generating weights for each point cloud and performing multiplication with corresponding the point cloud. Subsequently, the *spatio-temporal extractor* module is based on a Bidirectional Long Short-Term Memory (Bi-LSTM) neural network to capture the spatio-temporal relationship of successive point cloud frames, followed by two linear layers and a Softmax function are used to determine the person identity.

3 EVALUATION

3.1 Experiments

We use a point cloud sequence consisting of 45 frames as an input sample for each subject. The number of points in each point cloud is predefined as 128. If the actual number of points is fewer than 128, the points in the current point cloud are used as padding points. The training-to-testing ration is maintained at 8:2.

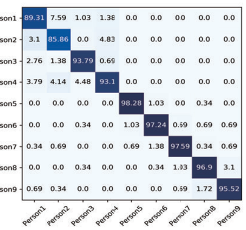

Figure 2: Confusion Matrix.

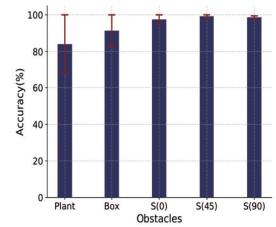

Figure 3: Occluded Scenarios.

3.2 Initial Results

Identification accuracy. The experimental result shows that our system can achieve an average accuracy of 94.17% under various occluded conditions for the 9 human subjects. Meanwhile, the accuracy witnesses an increase from 84.35% to 94.17% by data augmentation. Moreover, Figure 2 depicts a percentage confusion matrix, highlighting 7 individuals with accuracy surpassing 93.00%.

Impact of Various Occluded Scenarios. We then compare the model performance across occluded conditions. Figure 3 demonstrates that our method consistently maintains an average accuracy of over 91.00% across 4 occluded conditions. In the scenario with complete occlusion caused by the presence of plants, our method achieves an accuracy of 84.10%, slightly lower than in other scenarios. Note that, the S(0), S(45) and S(90) correspond to the sponge that is revolved at $0°$, $45°$, and $90°$, respectively.

3.3 Discussions

As our future work, we plan to expand the current mmWave radar dataset with more human subjects with a variety of gaits and motion patterns for a more thorough investigation and evaluation of our mmWave-ocPID method. We also plan to explore the open-set PID task [3], which requires identification of not only the "Known" targets but also the ability to reject the "Unknown" targets.

4 CONCLUSION

We study the feasibility of using mmWave radar for person identification in heavily or even completely occluded scenarios, by recognizing the subject's spatial-temporal gaits. We collect a mmWave radar point clouds dataset, with 5 occlusion configurations. We also develop a novel mmWave identification method incorporating two data augmentation strategies and a neural network. Experiments show that our method achieves 94.17% average accuracy.

REFERENCES

[1] Dhritimaan Das, Ayush Agarwal, and Pratik Chattopadhyay. 2022. Gait Recognition from Occluded Sequences in Surveillance Sites. In *European Conference on Computer Vision*. Springer, 703–719.
[2] Dmytro Iatsenko, Peter V. E. McClintock, and Aneta Stefanovska. 2016. Extraction of instantaneous frequencies from ridges in time-frequency representations of signals. *Signal Process.* 125 (2016), 290–303.
[3] Sagar Vaze, Kai Han, Andrea Vedaldi, and Andrew Zisserman. 2022. Open-Set Recognition: A Good Closed-Set Classifier is All You Need. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.
[4] Peijun Zhao, Chris Xiaoxuan Lu, Jianan Wang, Changhao Chen, Wei Wang, Niki Trigoni, and Andrew Markham. 2019. mID: Tracking and Identifying People with Millimeter Wave Radar. In *DCOSS*. IEEE, 33–40.