Work in Progress: Enabling User Identification for mmWave-based Gesture Recognition Systems

Lilin Xu[1],Keyi Wang[2],Chaojie Gu[1,*], Shibo He[1], Jiming Chen[1]
[1]Zhejiang University, [2]University of California San Diego

ABSTRACT

The mmWave radar has been exploited for gesture recognition. However, existing mmWave-based gesture recognition methods cannot identify different users, which is important for ubiquitous gesture interaction in many applications. This paper proposes *GesturePrint*, which is the first to achieve person-independent gesture recognition and gesture-based user identification using a commodity mmWave radar sensor. *GesturePrint* features an effective pipeline that enables the gesture recognition system to identify users with a minor additional cost. By introducing an efficient signal preprocessing stage and a novel network architecture GesIDNet, which employs an attention-based adaptive multilevel feature fusion mechanism, *GesturePrint* extracts both unique characteristics of predefined gestures and effective features of personalized motion patterns. Experiments on our self-collected dataset and three public datasets demonstrate *GesturePrint*'s superior performance in enabling user identification for gesture recognition systems.

CCS CONCEPTS

• Human-centered computing → Ubiquitous and mobile computing systems and tools; **• Computing methodologies → Machine learning**.

ACM Reference Format:
Lilin Xu, Keyi Wang, Chaojie Gu, Shibo He, Jiming Chen. 2023. Work in Progress: Enabling User Identification for mmWave-based Gesture Recognition Systems. In *ACM Workshop on mmWave Sensing Systems and Applications (mmWaveSys '23), November 12, 2023, Istanbul, Turkiye.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3628357.3629711

1 INTRODUCTION

Nowadays, the mmWave radar has received increasing attention from industry and academia because of its low power, high spatial resolution, and robustness to temperature and lighting conditions. The mmWave radar has empowered plenty of applications in autonomous driving, human localization and tracking, and healthcare. In recent years, there is a trend of utilizing mmWave radar to implement gesture recognition systems, enabling ubiquitous gesture interaction in a broad spectrum of applications, including gaming control, Internet of Things (IoT), and virtual reality (VR). In the last several years, several mmWave-based gesture recognition solutions have been proposed. Although existing solutions can accurately recognize predefined gestures performed by different users, none of them can identify the user who performs the gestures. In practice, the capability of user identification can significantly improve the user experience in interacting with smart devices. For example, users can personalize the meaning of gestures according to their habits when operating smart devices.

To unleash the potential of the mmWave-based gesture recognition system, in this work, we propose *GesturePrint*, a reliable one-stop solution to enable the gesture recognition system to identify users with minimal extra cost. *GesturePrint* first obtains point clouds related to objects in the environment by the radar device. It then segments the gesture motions utilizing a parameter-adaptive sliding window method, subsequently removing the outlier points that are distinguished from points reflected from people. With the preprocessed data, *GesturePrint* can accurately recognize predefined gestures and identify the user who performs the gestures based on our specially designed network architecture GesIDNet.

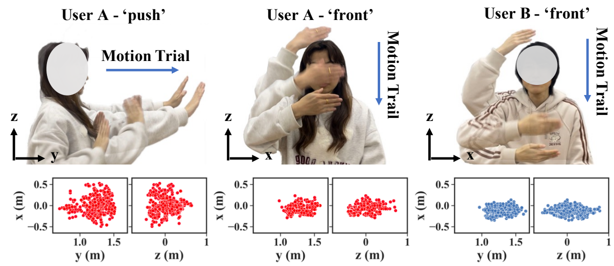

Figure 1: The visualization of gesture point clouds.

2 PRELIMINARIES AND DESIGN OVERVIEW

2.1 Preliminaries

The mmWave radar captures signals reflective of gesture motions as users perform gestures. As personal behavioral traits can serve as biometrics, we utilize mmWave radar to capture personal characteristics, including behavior manners and personalized unconscious motion styles, from gesture motions for identification.

Figure 1 shows the visualization of gesture point clouds captured from User A and User B when they performed American Sign Language (ASL) signs 'push' and 'front'. These two users have similar body shapes, with a height of around 160 cm and a weight of around 48 kg. The gesture point clouds exhibit distinct characteristics that can be utilized for gesture recognition and user identification. On the one hand, gesture point clouds demonstrate different shapes and movements of different ASL gestures. On the other hand, gesture point clouds differ in space and time between the same gesture

This research is supported by the National Natural Science Foundation of China under Grant No. U1909207.
*Chaojie Gu is the corresponding author.

mmWaveSys '23, November 12, 2023, Istanbul, Turkiye

ACM ISBN 979-8-4007-0441-3/23/11…$15.00
https://doi.org/10.1145/3628357.3629711

Table 1: Performance comparison. SOTA denotes previous state-of-the-art results. The best results are marked in bold.

Dataset	GesturePrint (ours)						Pantomime						mHomeGes			mTransSee		
Scenario	Office			Meeting Room			Office			Open			Home			Home		
Metrics	GRA	GRF1	GRAUC	GRA	GRF1	GRAUC	GRA	GRF1	GRAUC	GRA	GRF1	GRAUC	GRA	GRF1	GRAUC	GRA	GRF1	GRAUC
SOTA		/			/		0.9714	-	0.9994 [5]	0.9612	-	**0.9994** [3]	0.9800 [2]	-	-	0.9800 [1]	-	-
GesturePrint	**0.9822**	**0.9821**	**0.9908**	**0.9887**	**0.9885**	**0.9942**	**0.9854**	**0.9846**	**0.9997**	**0.9662**	**0.9633**	0.9993	**0.9960**	**0.9957**	**0.9966**	**0.9988**	**0.9988**	**0.9992**
Metrics	UIA	UIF1	UIAUC	UIA	UIF1	UIAUC	UIA	UIF1	UIAUC	UIA	UIF1	UIAUC	UIA	UIF1	UIAUC	UIA	UIF1	UIAUC
SOTA		/			/			/			/			/			/	
GesturePrint	**0.9926**	**0.9901**	**0.9947**	**0.9978**	**0.9972**	**0.9990**	**0.9985**	**0.9972**	**0.9987**	**0.9931**	**0.9902**	**0.9962**	**0.9933**	**0.9925**	**0.9969**	**0.9760**	**0.9707**	**0.9913**

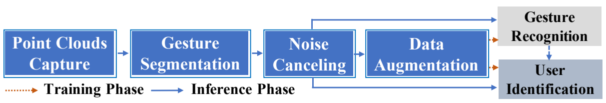

Figure 2: The overview of *GesturePrint*.

performed by different users, such as point number, coverage, and density. These differences are mainly caused by individual variations in arm length, motion speed, range of motion, and implicit motion habits. Thus, it is desirable to design an efficient data preprocessing method and an effective network for achieving reliable gesture recognition and user identification with mmWave radar.

2.2 Design Overview

GesturePrint is designed to extract effective features regarding specific gestures and users from sparse point clouds captured by the mmWave radar. Figure 2 shows the system pipeline of *GesturePrint*, which has two major stages with six modules. The data preprocessing stage includes *point clouds capture*, *gesture segmentation*, *noise canceling*, and *data augmenation*. The classification stage includes *gesture recognition* and *user identification*.

During the data preprocessing stage, *GesturePrint* works with a commodity mmWave radar sensor, after obtaining the points converted from signal data through the radar, it segments out gestures from temporal point cloud frames by using an adaptive sliding window. After gesture segmentation, *GesturePrint* further discards outlier noise points that are not reflected from the human body by utilizing DBScan. We aggregate points captured by radar in the whole gesture process, which are then fed into GesIDNet for gesture recognition. With the recognition result, GesIDNet further identifies the user performing the gesture with the gesture-corresponding recognition model. Finally, the gesture and the user are inferred by *GesturePrint*. In particular, during training, we augment the point cloud data by adding some random jitters to the points.

For the classification stage, we propose GesIDNet to address recognition and identification tasks based on gesture point clouds. PointNet++ [4] can extract details and features from data structured in the point cloud format. However, while PointNet++ is typically employed with large-scale dense point clouds, the gesture point clouds captured by the mmWave radar are usually sparse. Thus, to better extract features from the sparse gesture point clouds, we adopt the set abstraction block of PointNet++ and further design a multilevel feature fusion module with an attention mechanism.

On the one hand, GesIDNet uses the set abstraction block of PointNet++ to extract local spatial features at different scales from the aggregated gesture point clouds, and then these multiscale local features are combined for extracting high-level features. On the other hand, the aggregated point clouds comprise an unordered set of points with varying numbers and strong spatio-temporal correlations. To exploit the data characteristics, we introduce a novel multilevel feature fusion module with an attention mechanism to adaptively combine low-level features and high-level features extracted from point clouds. With the module, GesIDNet assigns large weights to the features reflecting effective spatio-temporal patterns. Based on the above design, GesIDNet can effectively extract unique gesture features for gesture recognition, as well as features containing personalized motion patterns for user identification.

3 EVALUATION

We evaluate *GesturePrint* on four mmWave-based gesture datasets that span diverse scenarios, user scales and predefined gestures. These datasets include our self-collected dataset, the GesturePrint dataset (including data from 17 participants performing 15 ASL gestures in two scenarios), and three public datasets, i.e., the Pantomime dataset [3], the mHomeGes dataset [2], and the mTransSee dataset [1]. We use four metrics, i.e., accuracy, F1-Score, AUC, and equal error rate (EER), to measure the performance of *GesturePrint* in both gesture recognition (GR) and user identification (UI).

As shown in Table 1, *GesturePrint* achieves accuracy above 96% for gesture recognition across all the datasets. Compared with the state-of-the-art results on the three public datasets [1–3, 5], *GesturePrint* achieves comparable recognition accuracy. Besides, the system consistently maintains GRF1 above 0.96 and GRAUC exceeding 0.99 across all the scenarios. For user identification, the overall accuracy of *GesturePrint* is over 97%, demonstrating that it is effective with different user scales. Besides, it consistently maintains reliable UIF1 and UIAUC. Moreover, *GesturePrint* achieves an average result of 0.75% EER across all the scenarios, with none exceeding 1.6% EER. All the results indicate the effectiveness of *GesturePrint* in gesture recognition and user identification.

REFERENCES

[1] Haipeng Liu, Kening Cui, Kaiyuan Hu, Yuheng Wang, Anfu Zhou, Liang Liu, and Huadong Ma. 2022. mTransSee: Enabling Environment-Independent mmWave Sensing Based Gesture Recognition via Transfer Learning. *ACM IMWUT* 6 (2022).

[2] Haipeng Liu, Yuheng Wang, Anfu Zhou, Hanyue He, Wei Wang, Kunpeng Wang, Peilin Pan, Yixuan Lu, Liang Liu, and Huadong Ma. 2020. Real-time arm gesture recognition in smart home scenarios via millimeter wave sensing. *ACM IMWUT* 4 (2020).

[3] Sameera Palipana, Dariush Salami, Luis A Leiva, and Stephan Sigg. 2021. Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds. *ACM IMWUT* 5, 1 (2021).

[4] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *31st NIPS* (2017).

[5] Dariush Salami, Ramin Hasibi, Sameera Palipana, Petar Popovski, Tom Michoel, and Stephan Sigg. 2022. Tesla-rapture: A lightweight gesture recognition system from mmwave radar sparse point clouds. *IEEE TMC* (2022).