Poster Abstract: Enhancing Human Motion Sensing with synthesized Millimeter-Waves

Xiaotong Zhang[1,2], Kun Wang[1], Zhenjiang Li[1], Jin Zhang[2,3]
[1]Department of Computer Science, City University of Hong Kong, Hong Kong, China
[2]Research Institute of Trustworthy Autonomous Systems and Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China
[3]Peng Cheng Laboratory, Shenzhen, China

ABSTRACT

This poster introduces SynMotion, a novel mmWave-based human motion sensing system addressing the scarcity of training datasets. By synthesizing mmWave signals using existing vision-based human motion datasets, this system overcomes the challenge of collecting and labeling mmWave data, facilitating wider adoption of mmWave technology for applications like activity recognition, skeleton tracking and radar placement recommendation.

KEYWORDS

Human Motion Sensing; Millimeter Wave; Body Skeleton Tracking; Activity Recognition

1 INTRODUCTION

Human motion sensing, crucial for various applications, is effectively achieved by vision-based methods utilizing cameras [4]. These methods rely on rich labeled datasets and machine learning innovations for accuracy. However, inherent constraints such as line-of-sight views, light conditions, and privacy issues limit the effectiveness of vision-based approaches [5, 7].

To address these limitations, recent breakthroughs explore Radio Frequency (RF) solutions [3, 6], with millimeter-wave (mmWave) from frequency-modulated continuous-wave (FMCW) radars being a notable example. Unlike vision, mmWave-based training datasets are extremely scarce. Collecting and labeling mmWave data for fine-grained tracking services present challenges and expenses that hinder the widespread adoption of mmWave sensing in practice. Recent designs have explored synthesizing sensing signatures from mmWave signals [1]. However, limited research has focused on synthesizing mmWave itself—the common source of various sensing signatures. If other sensing signatures are required, or even new signatures are proposed, they can be directly derived from the synthesized mmWaves. In addition to the data scarcity issue, mmWave radar's capability is also limited to perceive only radial displacement, emphasizing the importance of selecting an appropriate placement position for the radar.

Therefore, we introduces SynMotion, a novel mmWave-sensing system to address the scarcity of training data by leveraging existing vision-based human motion datasets to synthesize mmWave sensing signals. The synthesized signals inherit labels, including body skeleton coordinates and motion names, directly from the vision-based datasets. SynMotion enables three practical sensing services with commercial radars: 1) zero-shot activity recognition, where the classifier is trained solely on synthesized data but recognizes real signals, 2) few-shot body skeleton tracking, serving as a seed for future zero-shot skeleton tracking designs and 3) radar placement recommendation. To synthesize realistic mmWaves, we propose a novel software pipeline to emulate the entire procedure from transmission to reception of the mmWave signals reflecting off the human body by discretizing the human body into a manageable set of parameters (Figure 1). For more diversity, we incorporate the diffusion model [2] to augment the motion capture dataset and environmental reflections. We also design a novel training framework to deal with the micro-level signal differences between synthesized signals and real ones. With our proposed synthesizing software, we can pre-simulate mmWave signals collected from radar placed in various locations. This enables us to recommend radar placement positions and optimize the performance of human motion sensing.

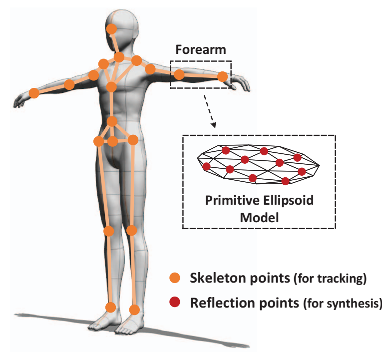

Figure 1: For the signal synthesis, we adopt a pictorial body model composed of primitive ellipsoids, *e.g.*, illustrated for the fore-arm.

2 SYSTEM DESIGN

2.1 Synthesizing Sensing Signals

1) FMCW radio. FMCW radar transmits a signal called a **chirp**, which is a sinusoid with the frequency linearly increasing from f_0 in bandwidth B over time duration T. Multiple chirps further form one **frame**. Each received (Rx) chirp represents a time-delayed version of the transmitted (Tx) chirp. The radar circuit produces an intermediate frequency (IF) signal with a constant frequency equal to the frequency difference between the Tx and Rx chirps.

2) Modeling FMCW signals. To model FMCW sensing signals, we consider at time t, the $(m+1)^{th}$ chirp is being transmitted for time t_m already. Thus, we can mathematically derive $S_{IF}(t)$ from a single reflection point at distance D as follows:

$$S_{IF}(t) = Ae^{j2\pi(f_0\tau - \frac{B\tau^2}{2T} + \frac{Bt_m\tau}{T})}, \tag{1}$$

979-8-3503-6201-5/24/$31.00 ©2024 IEEE
DOI 10.1109/IPSN61024.2024.00054

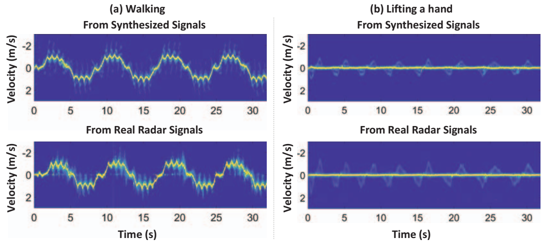

Figure 2: Visualization of micro-Doppler spectrums derived from our synthesized results and real signals.

where A is the attenuated amplitude related to radar cross section (RCS) σ, τ is $\frac{2D}{c}$, c is the light speed.

3) Reflections from human body. For each primitive ellipsoids part k of our body, we consider a finite number of reflection points, denoted as n. For reflection points from an ellipsoid, the estimation of RCS can be written as:

$$\sqrt{\sigma} = \left[\frac{\frac{1}{4}\pi R_k^4 H_k^4}{R_k^2 sin^2\theta_k + \frac{1}{4}H_k^2 cos^2\theta_k}\right]^{\frac{1}{2}} e^{-j\frac{2\pi}{\lambda}2D}, \tag{2}$$

where θ_k is the angle of the incident wave relative to the height axis of the ellipsoid with height H_k and radius R_k. After σ is obtained, all the parameters in $S_{IF}(t)$ from a single reflection point are known, and the final sensing signal can be obtained by:

$$S(t) = \sum\nolimits_{k\in\mathcal{K}} \sum\nolimits_{i=1}^{n} S_{IF}^{i}(t). \tag{3}$$

2.2 Augment with the Diffusion Model

In the formulation above, the distance D is calculated from predefined radar position and skeleton points of in the human motion dataset. Therefore, the diversity of the human motion dataset enables us to obtain more enriched synthetic mmWave data, thereby augmenting the training dataset. However, despite the convenience of collecting motion capture data, its diversity remains limited in both the human body model and motion patterns. Furthermore, the above pipeline is limited to synthesizing signals reflections on the human body, ignoring the environmental reflections.

To address these two issues and obtain more realistic synthetic mmWave signals, we include two distinct diffusion models with different functions in SynMotion. One is designed to generate diverse human motion data, while the other is responsible for generating environmental reflection noises. Figure 2 shows the comparison between the micro-Doppler spectrums derived from our synthesized signals and those from real ones.

2.3 Training Framework

Figure 3 depicts the training framework in SynMotion. Even though our software pipeline can accurately delineating mmWave signals, subtle differences persist compared to real signals due to high-order reflections, discrete human models, and other factors. Applying a skeleton tracker trained on synthesized data directly to real mmWave signals results in significant tracking errors. To tackle this challenge, we introduce a variant for the skeleton tracker with the same network structure, incorporating the user's initial pose for each motion as an additional input. Both trackers train initially with synthesized data. By feeding a few mmWave signals and initial poses from the target users to the variant tracker, we use the estimated coordinates as pseudo labels to fine-tune the actual tracker, enabling generalization to real mmWave signals.

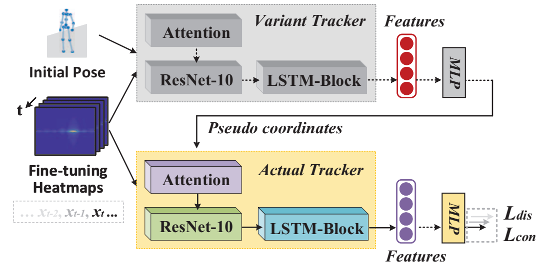

Figure 3: SynMotion's training framework.

2.4 Radar Placement Recommendation

Millimeter-wave human motion sensing is constrained by radar's capability to perceive only radial displacement. When users move along the circumference with the radar as the center, the radar can hardly perceive their motion. Therefore, the radar needs to be placed in a position that maximizes its ability to perceive radial displacement of the users. To this end, we first synthesize the signals received by the radar at different positions, then evaluate the sensing quality for each position using a specialized metric. The metric comprises two components: 1) the sensing accuracy directly derived from neural networks and 2) the assessment of motion-related information such as velocity and displacement extracted from the sensing signature generated by synthesized signals.

3 CONCLUSION

We present SynMotion, a new system to address the scarcity issue of training dataset for mmWave-based human motion sensing in this poster. This poster is related to our recent work [8].

ACKNOWLEDGEMENTS

This work is supported by the GRF grant from Research Grants Council of Hong Kong (CityU 11213622). Corresponding authors: Zhenjiang Li and Jin Zhang.

REFERENCES

[1] K. Ahuja, Y. Jiang, M. Goel, and C. Harrison. Vid2doppler: Synthesizing doppler radar data from videos for training privacy-preserving activity recognition. In *Proceedings of ACM CHI*, 2021.

[2] J. Ho, A. Jain, and P. Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 2020.

[3] W. Jiang, H. Xue, C. Miao, S. Wang, S. Lin, C. Tian, S. Murali, H. Hu, Z. Sun, and L. Su. Towards 3d human pose construction using wifi. In *Proceedings of ACM MobiCom*, 2020.

[4] Y. Liu, S. Zhang, and M. Gowda. When video meets inertial sensors: zero-shot domain adaptation for finger motion analytics with inertial sensors. In *Proceedings of ACM/IEEE IoTDI*, 2021.

[5] H. Truong, S. Zhang, U. Muncuk, P. Nguyen, N. Bui, A. Nguyen, Q. Lv, K. Chowdhury, T. Dinh, and T. Vu. Capband: Battery-free successive capacitance sensing wristband for hand gesture recognition. In *Proceedings of ACM SenSys*, 2018.

[6] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang. Person-in-wifi: Fine-grained person perception using wifi. In *Proceedings of IEEE/CVF ICCV*, 2019.

[7] H. Xue, Y. Ju, C. Miao, Y. Wang, S. Wang, A. Zhang, and L. Su. mmmesh: towards 3d real-time dynamic human mesh construction using millimeter-wave. In *Proceedings of ACM MobiSys*, 2021.

[8] X. Zhang, Z. Li, and J. Zhang. Synthesized millimeter-waves for human motion sensing. In *Proceedings of ACM SenSys*, 2022.