# 3D Human Pose Tracking with RFID

Shiwen Mao(✉)

Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201, USA
smao@ieee.org

**Abstract.** RF-based human pose estimation has attracted considerable interest recently as an effective means of human-computer interaction (HCI). Compared with camera-based alternatives, RF sensing can better protect user's privacy and are robust to lighting, cluttered background, view angle, and non-line-of-sight conditions. However, due to complicated indoor propagation environments, most RF-based sensing methods are sensitive to the deployment environment, are hard to generalize, and rely heavily on large amounts of training data. In this talk, we present RFID-Pose, which uses RFID tags as wearable sensors to detect 3D human pose. In addition to introducing the RFID-Pose system, we also investigate various enhancements for (i) making the model more generalizable to various subjects and environments, (ii) making the model more generalizable to different wireless technologies, and (iii) reducing the dependence on large amounts of training data via data augmentation. Experiments conducted in various environments demonstrate the high performance and efficacy of the proposed approaches.

## 1 Introduction

Human posture detection and tracking are useful for a variety of applications, such as human-computer interaction (HCI), video surveillance, movie making, and somatosensory gaming. Nevertheless, the video data collected by camera-based methods for pose monitoring could be intercepted by attackers, thus raising considerable security and privacy concerns. Radio Frequency (RF) based approaches have been proposed to address the privacy concern, but they usually require large amounts of training data. They also suffer from poor generalization when applying a trained system to different environments. This is because the deep learning model used in such systems usually require considerable RF data for model training, and the well-trained model is usually hard to be applied to untrained data domains of new RF environments. To address these issues, various approaches, such as data augmentation, domain adversarial networks, and transfer learning, have been proposed recently.

In this talk, we first introduce the design of our proposed RFID-based 3D human pose tracking system, termed *RFID-Pose* [1]. RFID-Pose uses passive

This work is supported in part by the US NSF under Grants IIS-2306789 and CCSS-2245608.


J. L. Pons et al. (Eds.): ICNR 2024, BIOSYSROB 32, pp. 730–734, 2024.
https://doi.org/10.1007/978-3-031-77584-0_144

RFID tags as wearable sensors attached to human joints, and reconstruct 3D human poses from the phase angles in received RFID response signals. In addition, video camera captured data is used for supervised training of the deep learning model used in RFID-pose in the training stage, as well as ground truth in our performance evaluation. Next, we leverage a meta-learning strategy to address the environment/domain adaptation problem and present *Meta-Pose* [2,3], which is an environment-adaptive system. Meta-Pose is pretrained by meta-learning algorithms with training data sampled from a small number of known environments. The meta-learning algorithms are adopted to achieve an optimized parameter initialization, so the system is able to adapt to an untrained environment with few-shot fine-tuning. We then introduce two effective means to mitigate the reliance on training data. The first is a *technology-agnostic* approach [4]: by exploiting a domain adversarial neural network, our model can utilize the heterogeneous RF data collected by various wireless technologies. The second is data augmentation [5–8]: by using generative adversarial networks (GAN) and the more recent diffusion model [9], we synthesize large amounts of training data at a low cost, thus saving the huge efforts associated with RF data collection. The efficacy of the prototype systems are demonstrated by our experiments conducted in various indoor environments.

## 2 System Design

The RFID-Pose system [1] is proposed to perform 3D human pose tracking with phase data collected from the RFID tags attached to the joints of the test subject. Figure 1 presents an overview of the RFID-Pose system architecture, which is composed of three key modules, including RFID data collection, phase data preprocessing, and a multi-model, vision-assisted deep neural network for pose estimation.

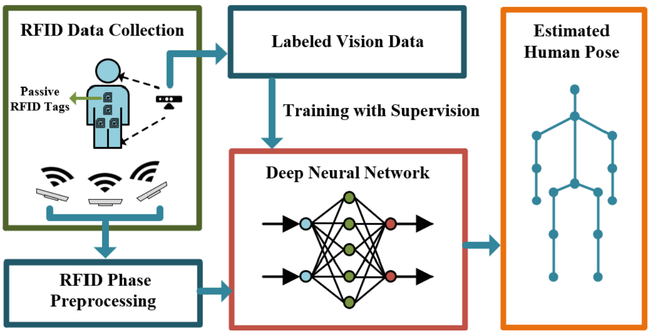


**Fig. 1.** Overview of the architecture of the proposed RFID-Pose system.

*1) RFID Data Collection and Preprocessing:* In RFID-Pose, the position in 3D space of human skeletal segment from the phase data measured through RFID communications, which follows the Low Level Reader Protocol (LLRP). The phase variation of two consecutively sampled phases is used to alleviate the impact of the phase offset caused by channel hopping. Such phase variation data effectively captures over time the dynamic tag-antenna distance caused by body movements, which is then translated into 3D human poses by a deep learning model.

*2) Vision-Assisted Deep Neural Network:* The translation from phase variation data to 3D human pose is achieved by a multi-modal vision-assisted deep neural network. The input to the network is the preprocessed phase variation data, and training is accomplished by the supervision of the 3D human pose generated by Kinect. A recurrent autoencoder is implemented in the network to extract body movement features from RFID data and translate it to the movements of human limbs [1]. The mapping from the autoencoder output to the 3D coordinates of human skeleton is accomplished in the Forward Kinematics layer, which is a classic motion generator used in 3D animation and robotics [10]. The training goal is to minimize the error between the estimated human pose and the corresponding label (i.e., vision data), so the network should perform effective translation from RFID data to human pose.

## 3 Challenges and Solutions

*1) Meta-learning for Environment Adaptation:* We leverage meta-learning to improve the adaptability of the model and overcome the barrier of deploying the system in various environments [11]. The deep learning model is initially pretrained with the datasets sampled from known data domains, using two representative algorithms, MAML [12] and Reptile [13]. Once the network is properly initialized, only a few new training data will be required from a new data domain for fine-tuning the model.

*2) Technology-Agnostic Design:* Most existing RF sensing systems are designed with a single wireless platform, which, however, constraints the deployment of the system. To reduce the cost and overcome the barrier of wide deployment, we propose to develop a single model that can work with various RF platforms, such as 900 MHz RFID, 2.4 GHz WiFi, 5 GHz WiFi, and 77 GHz mmWave Radar [4]. An additional benefit is a more robust performance due to the complementary technologies. This is achieved by incorporating a Domain Adversarial Neural Network to encourage the model to focus more on body movement features and ignore those features related to the specific wireless platform.

*3) Data Augmentation:* Since wireless data spans multiple dimensions, including spectrum, space, time, hardware design, and protocols, it is highly costly to collect RF datasets which is high fidelity and generalizable. We propose to develop GAN models [5,6] and diffusion models [7,8] to synthesize high quality RF data to train the deep learning models in downstream sensing models. Our investigation demonstrates that high-fidelity and high-diversity datasets can be effectively generated by the proposed models, which can train the downstream task models to achieve an RF sensing performance comparable to that trained by real RF data.

## 4 Prototype Systems and Evaluation

The prototype systems are implemented with a commodity Impinj R420 reader equipped with three polarized antennas (i.e., S9028PCR). As Fig. 2 shows, 12 ALN-9634 (HIGG-3) RFID tags are attached to the clothes of the test subject. The vision data used for vision-assisted training and performance evaluation is sampled by an Xbox Kinect 2.0 device. Eight different RF environments (named $D_1$ to $D_8$) are used in the evaluation with different antenna deployments and locations, where $D_1$ to $D_4$ are used for model pretraining, and $D_5$ to $D_8$ are new data domains for evaluating model adaptability. In addition to 900 MHz RFID, 2.4 GHz WiFi, 5 GHz WiFi, and 77 GHz millimeter wave (mmWave) Radar are also used to test the proposed technology-agnostic method. GAN [5,6], diffusion [8], and stable diffusion [7] models are used to synthesize various RF data for model training.

Experiments in various RF environments are conducted to assess the system performance, with respect to the RFID-Pose's accuracy in detection and tracking 3D human pose, its generalization performance with respect to various subjects and environments, as well as various RF platforms, and the efficacy of the proposed GAN and diffusion model-based data augmentation methods. Detailed results and discussions can be found in the corresponding papers.

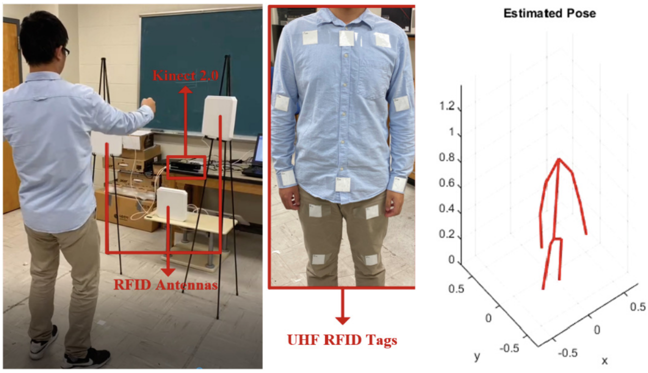


**Fig. 2.** Prototype systems and experiment set-up of the RFID-Pose system (unit: m).

## References

1. Yang, C., Wang, X., Mao, S.: RFID-pose: vision-aided 3D human pose estimation with RFID. IEEE Trans. Reliab. **70**(3), 1218–1231 (2021)

2. Yang, C., Wang, L., Wang, X., Mao, S.: Meta-pose: environment-adaptive human skeleton tracking with RFID. In: Proceedings of the IEEE GLOBECOM 2021, Madrid, Spain, pp. 1–6 (2021)
3. Yang, C., Wang, L., Wang, X., Mao, S.: Environment adaptive RFID based 3D human pose tracking with a meta-learning approach. IEEE J. Radio Frequency Identification **6**(1), 413–425 (2022)
4. Yang, C., Wang, X., Mao, S.: TARF: technology-agnostic RF sensing for human activity recognition. IEEE J. Biomed. Health Inform. **27**(2), 636–647 (2023)
5. Yang, C., Wang, Z., Mao, S.: RFPose-GAN: data augmentation for RFID based 3D human pose tracking. In: The 12th IEEE International Conference on RFID Technology and Applications (IEEE RFID-TA 2022), Cagliari, Italy, pp. 138–141 (2022)
6. Wang, Z., Yang, C., Mao, S.: Data augmentation for RFID-based 3D human pose tracking. In: Proceedings of the IEEE VTC-Fall 2022, London, UK (2022)
7. Wang, Z., Mao, S.: AIGC for RF sensing: the case of RFID-based human activity recognition. In: Proceedings of the 2024 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI (2024)
8. Wang, Z., Mao, S.: AIGC for wireless data: the case of RFID-based human activity recognition. In: Proceedings of the IEEE ICC 2024, Denver, CO (2024)
9. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. arXiv preprint arXiv:2006.11239 (2020)
10. Villegas, R., Yang, J., Ceylan, D., Lee, H.: Neural kinematic networks for unsupervised motion retargetting. In: Proceedings of the IEEE CVPR 2018, Salt Lake City, UT, pp. 8639–8648 (2018)
11. Yang, C., Wang, X., Mao, S.: RFID based 3D human pose tracking: a subject generalization approach. Elsevier/KeAi Digit. Commun. Netw. **8**(3), 178–288 (2022)
12. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: Proceedings of the ICML 2017, Sydney, Australia, pp. 1126–1135 (2017)
13. Nichol, A., Achiam, J., Schulman, J.: On first-order meta-learning algorithms. arXiv preprint arXiv:1803.02999 (2018)