Demo Abstract: Vision-aided 3D Human Pose Estimation with RFID

†Chao Yang, ‡Xuyu Wang, and †Shiwen Mao
†*Dept. of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201*
‡*Dept. of Computer Science, California State University, Sacramento, CA 95819-6021*
Email: czy0017@auburn.edu, xuyu.wang@csus.edu, smao@ieee.org

1. Introduction

Radio Frequency (RF) based human pose estimation techniques have been proposed to generate human pose without using a camera, so people will no longer worry about their privacy. Compared with other RF sensing based systems, Radio Frequency Identification (RFID) provides a promising solution for RF based human pose estimation. RFID tags can be used as wearable sensors because of their small size. The interference caused by the multipath effect is much smaller in the RFID system. The cost of RFID systems is also lower than the advanced radar based systems such as FMCW radar. Thus, we propose the RFID-Pose system for tracking the movements of multiple human limbs in realtime [1]. In the proposed system, RFID tags are attached to the target human joints. The movement of the tags are captured by the phase variations in the responses from each tag. The human pose is reconstructed by estimating rotation angles from RFID data and the initial human skeleton. The vision data will not be needed anymore in the testing phase, so the user's privacy can be well protected.

2. RFID-Pose System Overview

In this demo, we present an RFID based sensing system, termed RFID-Pose, to estimate 3D human pose in realtime. The RFID-Pose system can sense the 3D positions of all the RFID tags attached to human body by exploiting the phase data collected by the reader antennas. Human pose can be effectively constructed by mapping the positions of the RFID tags into 3D coordinates. An overview of the RFID-Pose system is presented in Fig. 1, which is mainly composed of four components, including (i) RFID phase data collection, (ii) Kinect skeleton data collection, (iii) RFID data preprocessing, and (iv) Skeleton reconstruction using a deep kinematic neural network.

2.1 RFID and Kinect Data Collection: In the proposed system, training data is sampled by both RFID antennas and a Kinect 2.0 device simultaneously. The collected RFID data will be used to train the deep kinematic neural network, and the Kinect 3D pose data will be used as labels for supervised training. To collect RFID data, we attach passive RFID tags to 12 joints of the human body. Three reader antennas are used to collect the phase and timestamp information from all attached RFID tags. Kinect 2.0 is a depth camera widely used for capturing 3D poses in interactive video games. The 3D position of each human joint is estimated by both an RGB camera and infrared sensors, and all measured joint positions are stored as 3D coordinates.

2.2 Data Preprocessing: Since the sampled RFID raw phase data suffers from considerable distortions caused by channel hopping and phase wrapping, calibration must be applied to cleanse the data before using it to train the deep neural network. We first calibrate the phase variation to mitigate the influence of channel hopping and phase wrapping. Next, we downsample the calibrated RFID data and synchronize it with the 3D pose time sequence obtained by Kinect 2.0. However, because of the slotted ALOHA-like transmission in the RFID system, tags are not evenly interrogated by the antennas. In order to synchronize the RFID data with the collected pose data from Kinect 2.0, we obtain the phase for all tags corresponding to each Kinect data frame. To this end, we propose to employ low rank tensor completion to estimate the missing phase values from the tags. Finally, the calibrated phase data is used to train the deep neural network for human skeleton reconstruction.

2.2 Deep Kinematic Neural Network: In RFID-Pose, we incorporate a deep kinematic neural network to learn the features of RFID phase data. Unlike monitoring one particular limb movement as in traditional RFID based skeleton tracking systems [2], [3], the deep kinematic neural network is designed to simultaneously estimate the spatial rotations of all

978-1-7281-9916-0/20/$31.00 ©2020 IEEE
DOI 10.1109/MSN50589.2020.00104

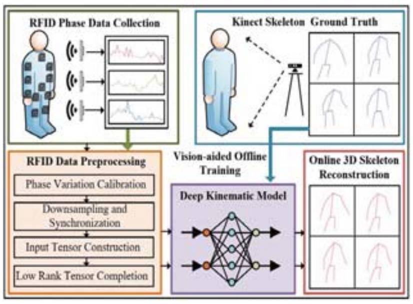

Figure 1. RFID-Pose system architecture.

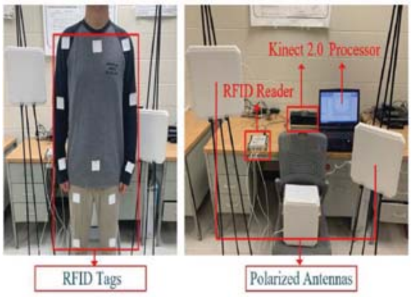

Figure 2. Illustration of the system setup for 3D pose estimation.

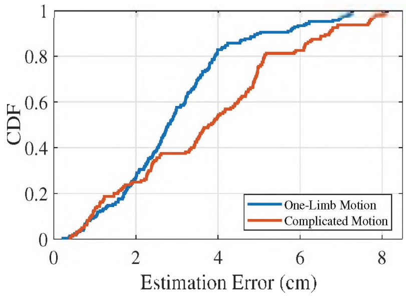

Figure 3. CDF of estimation errors.

human joints relative to their parent joints. Once the initial human skeleton (i.e., the length of the limbs of target) is given, the network could effectively learn the features of calibrated RFID tensor data, and reconstruct the positions of human joints with estimated rotation angles. In RFID-Pose, the Kinect pose data is only used as benchmark for evaluating the accuracy of 3D pose reconstruction in the online testing process.

The existing RFID based techniques are mostly focused on estimating the movement of a particular limb movement, such as the front arm, the front leg, and thighs [2], [3]. The RFID-pose system is designed to effectively monitor multiple human joints simultaneously in realtime.

3. Evaluation

To evaluate the performance of RFID-pose, we develop a prototype system with an off-the-shelf Impinj R420 reader equipped with three S9028PCR polarized antennas. The RFID tags used for tracking human joint movements are ALN-9634 (HIGG-3). The vision data used for training supervision and test accuracy evaluation is collected with an Xbox Kinect 2.0 device. The setup of the system is illustrated in Fig. 2. As the figure shows, we attach RFID tags to 12 joints of the human body. The antennas are placed at different altitude positions to ensure that all the tags can be interrogated. An MSI laptop with a Nvidia GTX 1080 GPU and an Intel Core i7-6820HK CPU is used as processor for data preprocessing and training. The frequency used by the prototype system hops among 50 channels from 902 MHz to 928 MHz, and each time it remains on a channel for 0.2 second.

We train the proposed deep kinematic neural network with different types of motions. The first type are simple motions, which involve the movement of a single-limb. The second type of motions are complicated motions, which are composed of movements of the entire body, such as body twisting, deep squat, boxing, and walking. The mean error of all the joints for each time slot T is calculated as: $\epsilon(T) = \frac{1}{12}\sum_{n=1}^{12} ||\hat{P}_n^T - \dot{P}_n^T||$, where $\hat{P}_n^T$ denotes the estimated position and $\dot{P}_n^T$ is the ground truth position collected by Kinect 2.0 in the 3D space for joint n at time T; and $||\hat{P}_n^T - \dot{P}_n^T||$ is the Euclidean distance between the two 3D vectors. The overall accuracy of human pose estimation is presented in the form of cumulative distribution function (CDF) of estimation errors in Fig. 3. From the CDF curves, we can see that the median estimation error is 2.83cm for the single-limb motion test and 3.75cm for the complicated motion test. The results show that the estimation accuracy of the entire body motion is lower than one-limb motion, because more moving joints need to be reconstructed in realtime in the former case. However, RFID-Pose achieves very high accuracy for all the complicated motions, and the largest error among all the tests is 8.12cm. The estimation results validate that the proposed RFID-Pose system can estimate the joint positions accurately and can effectively reconstruct the pose of the entire moving body in realtime through RFID phase data.

Acknowledgment This work is supported in part by the NSF (ECCS-1923163 and ECCS-1923717).

References

[1] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, to appear.

[2] H. Jin, Z. Yang, S. Kumar, and J. I. Hong, "Towards wearable everyday body-frame tracking using passive RFIDs," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–23, Dec. 2018.

[3] C. Wang, J. Liu, Y. Chen, L. Xie, H. B. Liu, and S. Lu, "RF-Kinect: A wearable RFID-based approach towards 3D body movement tracking," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–28, Mar. 2018.