Cong YU, Dongheng ZHANG, Zhi WU, Zhi LU, Chunyang XIE, Yang HU, Yan CHEN, 2023. RFPose-OT: RF-based 3D human pose estimation via optimal transport theory. *Frontiers of Information Technology & Electronic Engineering*, 24(10):1445-1457. https://doi.org/10.1631/FITEE.2200550

RFPose-OT: RF-based 3D human pose estimation via optimal transport theory

Key words: Radio frequency sensing; Human pose estimation; Optimal transport; Deep learning

Corresponding author: Yan CHEN
E-mail: eecyan@ustc.edu.cn
ORCID: https://orcid.org/0000-0002-3227-4562

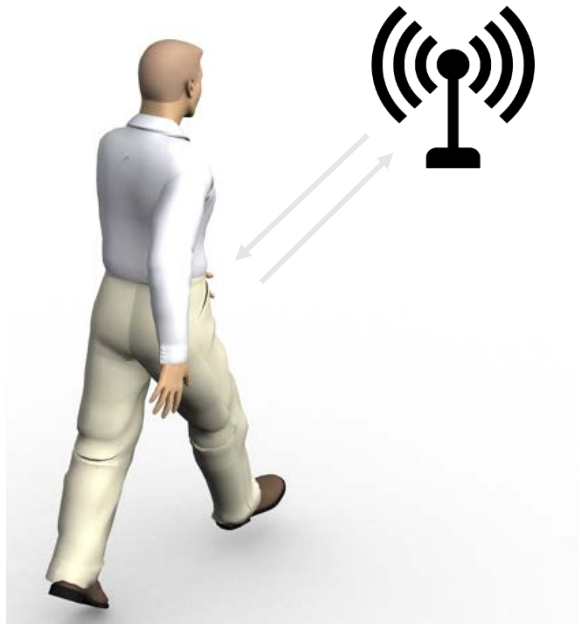
Motivation

Compared with camera-based and wearable-devices-based methods, radio frequency (RF) based human pose estimation can obtain human behavior and posture information with all-weather, non-contact, and non-line-of-sight characteristics.

Camera

Wearable devices

RF signals



Method

RF signals are transformed to the target pose feature domain based on the optimal transport (OT) theory, and then pose keypoints are generated from the transformed features.

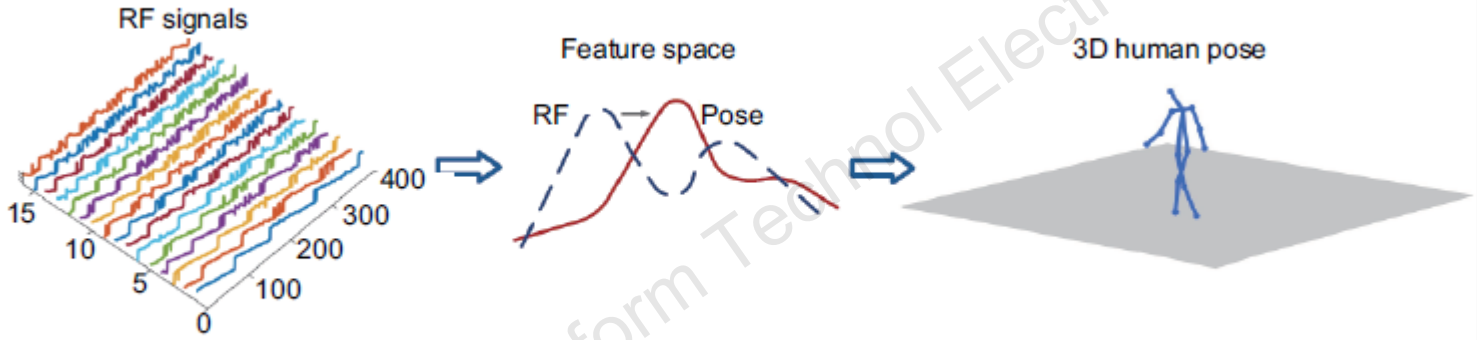


Fig. 1 RFPose-OT transforms the radio frequency (RF) signals to the pose domain to enable fine-grained three-dimensional (3D) human pose estimation

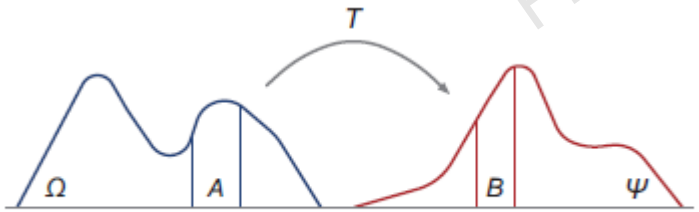


Fig. 2 Transport map

$$\mathcal{K}(\gamma) = \int_{\Omega \times \Psi} C(z, \hat{z}) \mathrm{d}\gamma(z, \hat{z})$$

$$\begin{cases} \gamma(A \times \Psi) = \mu(A), \\ \gamma(\Omega \times B) = \nu(B) \end{cases}$$

Method

The RFPose-OT architecture consists of a pose encoder, an RF encoder, and a keypoint predictor.

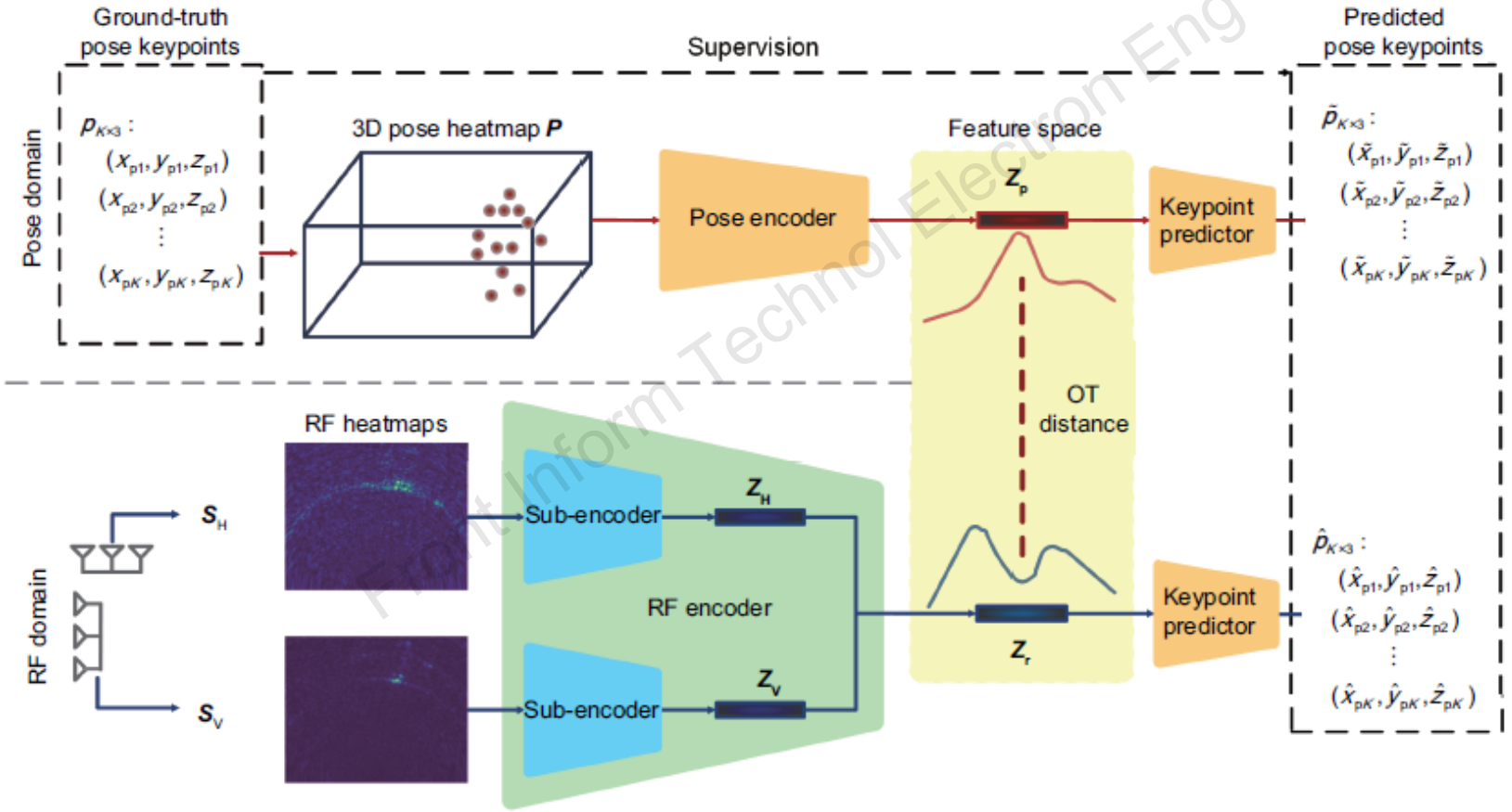


Fig. 4 The RFPose-OT architecture which consists of a pose encoder, a radio frequency (RF) encoder, and a keypoint predictor. Once trained, only the RF encoder and the keypoint predictor are retained to predict 3D human poses from RF heatmaps

Experiments

The spatial location error (SLE) between the predicted keypoints and the corresponding ground-truth keypoints is used as the metric:

$$\mathrm{SLE}_k = \frac{1}{U}\sum_{u=1}^{U}\left\|\hat{\boldsymbol{p}}_k^{(u)} - \boldsymbol{p}_k^{(u)}\right\|_2$$

Table 4 Quantitative evaluation results of different methods in the basic a and occlusion b indoor environments

Env	Method	SLE (cm)								
		Nose	Neck	Shoulders	Elbows	Wrists	Hips	Knees	Ankles	Overall
a	RF-Pose3D (Zhao et al., 2018a)	8.11	**5.21**	7.57	9.92	15.74	6.64	11.31	21.10	11.27
	mm-Pose (Sengupta et al., 2020)	8.19	5.30	7.23	9.67	15.29	**6.20**	10.83	19.04	10.72
	RFPose-OT	**7.90**	6.14	**6.76**	**7.99**	**11.67**	6.39	**8.34**	**12.60**	**8.68**
b	RF-Pose3D (Zhao et al., 2018a)	**6.53**	4.86	6.65	8.75	14.05	6.95	11.26	21.52	10.70
	mm-Pose (Sengupta et al., 2020)	6.64	**3.88**	**6.34**	9.16	14.84	6.98	11.28	19.28	10.45
	RFPose-OT	7.85	6.42	6.78	**7.90**	**11.41**	**6.82**	**9.35**	**14.05**	**9.07**

Bold numbers represent the minimum spatial location error (SLE) values. Env: environment

Experiments

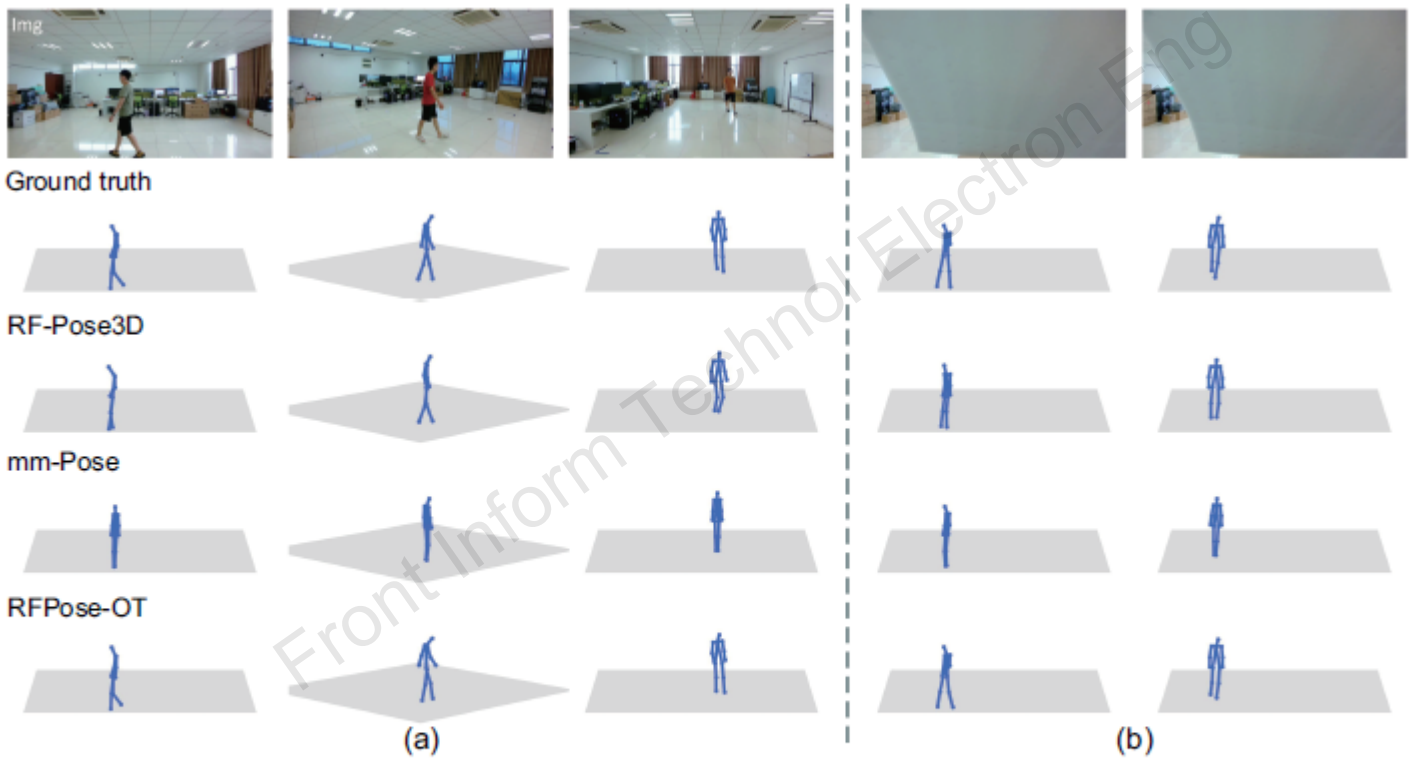
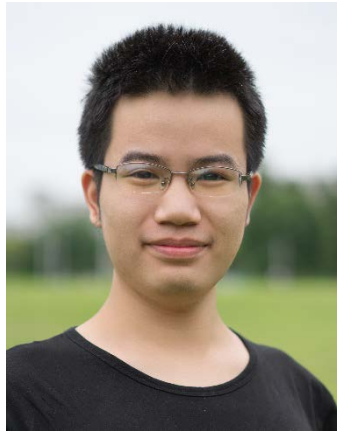


Fig. 5 Qualitative results of different methods in the basic (a) and occlusion (b) indoor environments

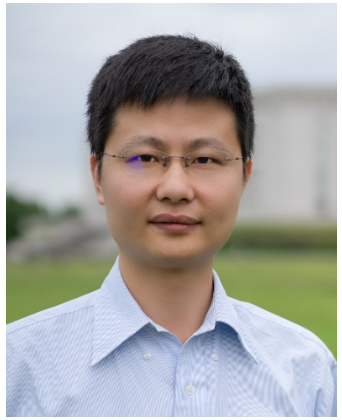
The 1st row shows the images captured by a camera that is attached to the radio system, the 2nd row shows the ground-truth 3D human poses, and the 3rd, 4th, and 5th rows show the 3D human poses estimated by RF-Pose3D (Zhao et al., 2018a), mm-Pose (Sengupta et al., 2020), and our proposed RFPose-OT, respectively

Conclusions

1. We propose an RF-based accurate human pose estimation method via the optimal transport theory.
2. The proposed RFPose-OT can be applied to a variety of scenarios.
3. RFPose-OT provides a new and valid framework to tackle related RF human sensing tasks.



Cong YU received his BS and PhD degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2019 and 2023, respectively. He is currently an assistant researcher at the Institute of Electronic Engineering, China Academy of Engineering Physics, Mianyang, China. His research interests include lightweight model, object detection, and wireless sensing.



Yan CHEN received his bachelor degree from the University of Science and Technology of China in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology in 2007, and the Ph.D. degree from the University of Maryland, College Park, MD, USA, in 2011. He was with Origin Wireless Inc. as a founding principal technologist. From Sept. 2015 to Feb. 2020, he was a professor with the School of Information and Communication Engineering at the University of Electronic Science and Technology of China. He is currently a professor and Vice Dean with the School of Cyber Science and Technology at the University of Science and Technology of China.

Dr. CHEN'S research interests include multimodal sensing and imaging, multimedia signal processing, and wireless multimedia. He is a co-author of *Reciprocity, Evolution, and Decision Games in Network and Data Science* (Cambridge University Press, 2021) and *Behavior and Evolutionary Dynamics in Crowd Networks: an Evolutionary Game Approach* (Springer, 2020), as well as co-author of over 200 technical papers including more than 100 IEEE journal papers. He is the Associate Editor for *IEEE Transactions on Network Science and Engineering* (TNSE) *and IEEE Transactions on Signal and Information Processing over Networks* (TSIPN). He was the Chair for APSIPA Signal and Information Processing Theory and Methods (SIPTM) Technical Committee, a Distinguished Lecturer for APSIPA, and the Secretary-General for the CES Young Scientist Network Multimedia Technical Committee. He was an Organizing Co-Chair of PCM 2017, a Special Session Co-Chair of APSIPA ASC 2017, the 10K Best Paper Award Committee Member of ICME 2017, the Multimedia Communications Symposium Lead Chair of WCSP 2019, an Area Chair for ACM Multimedia 2021, and a TPC Co-Chair of APSIPA ASC 2021, 2022. He was the recipient of multiple honors and awards, including an Excellent Editor for IEEE TNSE in 2021 and 2022, the best paper award runner-up at the IEEE MMSP in 2022, the best paper award at the APSIPA ASC in 2020, the best student paper award at the PCM in 2017, the best student paper award at the IEEE ICASSP in 2016, and the best paper award at the IEEE GLOBECOM in 2013.