Quantitative Assessment of People Tracking with FMCW MIMO Radar

Wang, Dingyang; Fioranelli, Francesco; Yarovoy, Alexander

DOI
10.23919/EuRAD61604.2024.10734901

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 2024 21st European Radar Conference (EuRAD)

Quantitative Assessment of People Tracking with FMCW MIMO Radar

Dingyang Wang, Francesco Fioranelli, Alexander Yarovoy
Microwave Sensing Signals & Systems (MS3) Group, Department of Microelectronics, TU Delft, The Netherlands
{D.Wang-6, F.Fioranelli, A.Yarovoy}@tudelft.nl

***Abstract*— In this paper the problem of tracking multiple people in an indoor environment is formulated and analyzed with the usage of a Multiple Input Multiple Output (MIMO) Frequency Modulated Continuous Wave (FMCW) radar. The objective is to evaluate the performance of FMCW MIMO radar with relatively limited bandwidth for accurately tracking single and multiple individuals in various scenarios. Three different tracking approaches are compared and quantitatively analyzed with single and multi-target scenarios. As the angular resolution is significant for distinguishing multiple targets, the effect of the number of MIMO channels used is compared among different trackers. The performance is analyzed by metrics of distance error and cardinality error in track association/assignment.**

***Keywords*— Multi-target tracking, FMCW Radar, Human Monitoring, Indoor Tracking**

I. Introduction

Radar-based indoor human tracking is a very active research field, leveraging on the advantages of using radar sensors, which operate contactless and do not require additional equipment to be attached or worn by the users. They are also more respectful of personal privacy than cameras and are insensitive to ambient light. Research activities in this field began with a focus on single or distributed mono-static radar due to its low cost and ease of manufacturing. However, with advancements in electronic and manufacturing technologies, each radar sensor in this research domain is typically a multiple input multiple output (MIMO) radar.

The development of MIMO radar has expanded capabilities by adding azimuth & elevation information to the existing range and Doppler measurements, enabling more accurate estimations of location and even postures of people moving in indoor environments. Processing pipelines for human monitoring such as tracking and classification can operate starting from range-azimuth (RA) or range-Doppler (RD) maps [1]. The former may provide a more stable signature of multiple people in close proximity over frames by avoiding Doppler spread effects, but are heavily influenced by the angular resolution of the radar.

Achieving fine angular resolution often requires an unfeasible number of MIMO channels in compact devices, and combined with the usage of relatively low bandwidth and carrier frequency, can make multi-target tracking a significant challenge in indoor environments with clutter and multipath. This paper presents a quantitative assessment of tracking results with three algorithms, with different numbers of MIMO channels used for finer/coarser angular resolution in different cases. The considered scenarios include 1 person randomly walking, 2 people walking side by side with 0.5 m spacing between them, and 3 people randomly walking. Quantitative metrics used for performance assessment include distance errors and the OSPA (Optimal SubPattern Assignment Metric) metric, which also accounts for the cardinality error in assigning wrong track numbers.

The rest of the paper is organized as follows. In Section II, the tracking methods are presented. The experimental setup is shown in Section III, with results and performance analysis also provided. Finally, the paper is concluded in Section IV.

II. Tracking approaches

In our previous work [1], a general data processing pipeline for human monitoring using MIMO FMCW radar was proposed. Specifically, there were two possible variants, one pipeline extracts point cloud from Range-Azimuth (RA) maps [2] which is shown in Figure 1, and another extracts point cloud from range-Doppler (RD) maps [3]–[5]. For the problem of tracking multiple people moving in not-predefined directions, it was shown that detections from RA maps contained more stable information. After detection, a Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [6] was applied for grouping the detected points belonging to targets. DBSCAN is particularly well-suited for extended targets due to its ability to handle irregularly shaped clusters and automatically identify outlier points within the data. This step can help differentiate between multiple targets and reduce false alarms, even if its performance can be reduced when multiple targets are close to each other, within one resolution cell.

Following clustering, in this paper different tracking algorithms are implemented and assessed to find the optimal match between new detections and current tracks. The update and correct process steps are in general based on the Kalman filter framework and included in the tracking step. The three tracking algorithms are discussed as follows.

A. GNN

Global nearest-neighbor (GNN)[7]–[9] is a multi-target version of Nearest-Neighbor (NN) algorithm. The algorithm considers all possible pairings between new detections and current tracks, and selects the most likely association according to a distance cost function. When considering a multi-target scenario, the tracker often involves strategies about new

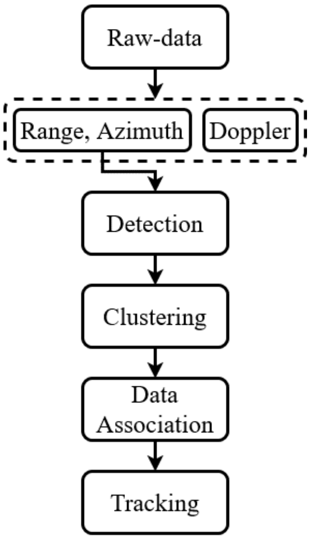


Fig. 1. Proposed processing pipeline for people tracking with MIMO radar

track initialization, updating current track and handling track termination. A track management system is required based on history logic of 'm out of n' method.

B. JPDA

Joint probabilities data association (JPDA) [10] is the extension of the probabilities data association (PDA), which obtains probabilities from measurement to target jointly across targets. The key feature of JPDA is that it evaluates the conditional probabilities of the following joint association events:

$$A(k) = \cap_{j=1}^{m} A_{jt_j}(k) \tag{1}$$

where $A_{jt}(k)$ is the event of measurement j at time k originated to track t, $j = 1, ..., m$, $t = 0, 1, ..., N_T$; t_j is the index of target associated to measurement j; t starts from 0 meaning that miss detection is considered; N_T is the track number.

C. MHT

Multiple hypothesis tracker (MHT) [8], [11], [12] makes hard association decisions over multi-scan of frames. It creates multiple potential tracks and continually evaluates hypotheses to determine the most likely track configurations. MHT algorithm is robust, but computationally expensive, especially in scenarios with dense clutter. Unlike traditional data association methods that consider a single hypothesis for each measurement-track association, MHT considers multiple possible hypotheses simultaneously and evaluates them in a probabilistic manner to determine the most likely set of tracks.

The algorithm allows for multiple hypotheses about the assignment of detections to tracks. Assignments of detections to tracks create branches for the assigned tracks. Tracks with no assigned detections are coasted (i.e., predicted). All track branches are scored. To manage computational complexity, branches with low initial scores are pruned. This reduces the number of hypotheses to be considered in subsequent steps. Clusters of branches that share detections (incompatible branches) in their history are then generated. Global hypotheses of compatible branches are formulated and scored. Branches are scored based on their existence in the global hypotheses. Low-scored branches are pruned. Additional pruning is performed based on N-scan history. All tracks are corrected and predicted to the input time. Moreover, MHT includes strategies for track management, such as track pruning, track merging, and track termination, to maintain track consistency and reliability over time.

III. Experimental results and comparison

To assess and compare different tracking algorithms in the proposed pipeline, a dedicated dataset was collected in the laboratory room of the MS3 group at TU Delft. The room is a cluttered environment with pieces of furniture such as tables, chairs, and metallic cabinets. A commercial FMCW radar (by Joby Austria, former INRAS) with a relatively low bandwidth of 250 MHz was used to evaluate the performance, with detailed parameter settings listed in Table 1. The range resolution is limited to 60 cm due to the relatively narrow bandwidth. Therefore, a target can occupy a large area in the range profile, and this may cause overlaps and merging for multiple target scenarios. To quantitatively analyze the tracking performance, an auxiliary RGBD camera was used to collect ground truth data at the same time as the radar measurements. Notably, due to the maximum range being limited at 6m, tracking ID switches can happen in the ground truth if people exceed that range during measurements.

For a single target scenario, assessing distance errors (e.g., median error, mean absolute error and root mean square error) is sufficient. To evaluate the accuracy of multi-target tracking systems, the Optimal Sub-pattern Assignment (OSPA) [13] metric is also considered. A cut-off distance of 1m from the radar is applied. For each activity, data are specifically analyzed in the 'steady-state', excluding the starting and the end of the recording by taking samples #101 to #500, i.e., a total of 400 samples equivalent to 40 seconds. This avoids the uncertain data at the beginning and the end where the participants might not be moving along the desired trajectory.

Table 1. JOBY (former INRAS) FMCW radar parameters

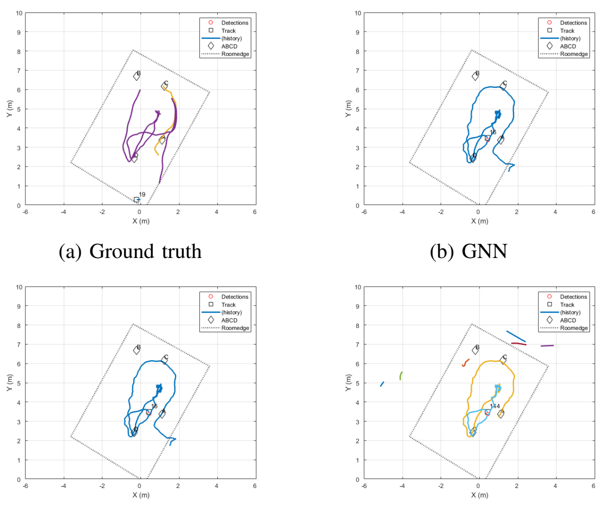
FMCW radar model	RadarBook2 (RBK2)
Operating frequency	24 GHz
Sweep bandwidth	250 MHz
ADC sampling rate	120 ksps
ADC samples	56
Up chirp duration	467 μs
Chirp repetition interval	483 μs
Number of chirps in a frame	90
Slow-time sampling frequency	10 Hz
Number of TX & RX channels	2 x 8
Antenna horizontal 3 dB beamwidth	76.5°

A. Tracking results of 1 person scenario

The numeric results of 1 person randomly walking are listed in Table 2. The corresponding tracking results are shown in Figure 2 when using 6 channels for angle estimation. First, the distance error is increased when using fewer channels as the angular resolution is coarser. The median error is around 23 cm, to be considered in the context of a range resolution of 60 cm (related to 250MHz bandwidth). The performance of GNN and JPDA tracking algorithms are quite close to each other for 1 target scenario. The MHT tracker has a lower median error as it takes multiple hypotheses into account which gives a more accurate assignment. However, MHT method shows more false trajectories as in Figure 2d Unassigned tracks show a similar trend. In Figure 3, the OSPA is shown for each measurement time step. The red dotted line is the localization error component, whereas the blue dotted line is the cardinality/assignment error.

Table 2. Quantitative metrics for 1 person random walking and sitting

Tracker	Channels	RMSE (cm)	MAE (cm)	Median (cm)	Mean of OSPA	Unassigned Tracks (%)	Miss Detection (%)
GNN	4	26.16	23.81	24.78	0.24	0	0
	6	24.53	22.62	22.07	0.23	0	0
	8	23.57	21.73	21.74	0.25	5.48	0
	12	22.59	20.73	20.7	0.25	8.36	0
	15	22.46	20.44	20.34	0.27	11.75	0
JPDA	4	26.16	23.81	24.78	0.24	0	0
	6	24.53	22.62	22.07	0.23	0	0
	8	23.57	21.73	21.74	0.25	5.48	0
	12	22.59	20.73	20.7	0.25	8.36	0
	15	22.46	20.44	20.34	0.27	11.75	0
MHT	4	26.09	23.76	24.12	0.24	1.31	0
	6	24.14	22.24	21.16	0.25	4.96	0
	8	22.14	20.11	18.97	0.38	28.72	0
	12	21.75	19.76	19.07	0.33	23.50	0
	15	20.89	18.74	18	0.41	37.34	0



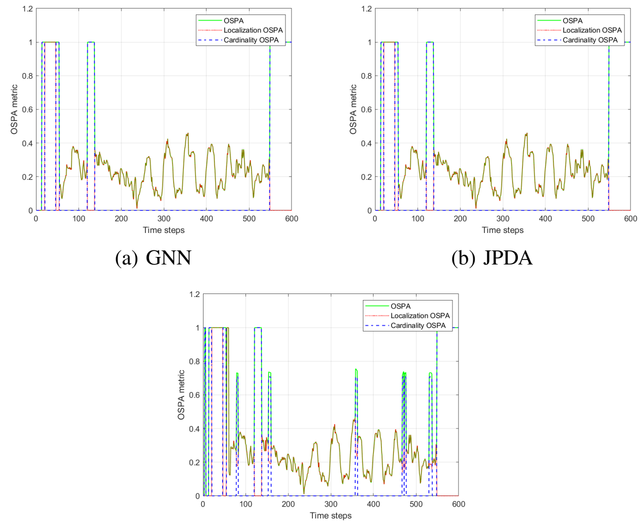
(a) Ground truth (b) GNN

(c) JPDA (d) MHT

Fig. 2. Ground truth and tracking results of 1 person random walking and sitting when using 6 MIMO channels. It should be noted that the ground truth has a break-point, indicated by different colors in the two tracks; however, the two trajectories belong to the same target.

B. *Tracking results of 2 people scenario*

Figure 4 shows the trajectories of 2 people walking in parallel, side by side, with 0.5m spacing between each other, with 12 channels used for angle estimation. The trajectory history changes in color for the same person (i.e., miss detection event), but it is still recognizable that two people



(a) GNN (b) JPDA

(c) MHT

Fig. 3. OSPA metric results of 1 person random walking and sitting when using 6 MIMO channels.

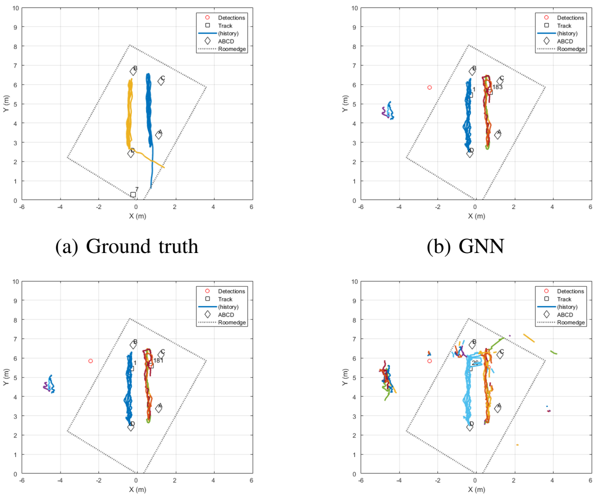
are moving in a radial direction with respect to the radar line of sight. The mean OSPA is around 0.32 for 2 people. MHT tracker seems to generate more false trajectories than the other two approaches. Tracker performance improvements are expected if the two individuals walk more randomly which reduces the impact of coarser angular resolution.

Table 3. Quantitative metrics for 2 people parallel walking side by side with 0.5m spacing

Tracker	Channels	RMSE (cm)	MAE (cm)	Median (cm)	Mean of OSPA	Unassigned Tracks (%)	Miss Detection (%)
GNN	4	17.58	16.49	17.07	0.55	66.5	2.25
	6	16.86	14.66	14.6	0.49	53.75	8.75
	8	23.97	17.3	14.21	0.55	63.75	0
	12	28.31	20.12	12.23	0.32	23.25	0
	15	23.86	17.71	14.35	0.39	40.25	0
JPDA	4	34.59	27.21	19.77	0.58	57.5	0
	6	33.32	24.45	17.24	0.54	49.25	7
	8	23.97	17.3	14.21	0.55	63.75	0
	12	28.32	20.13	12.21	0.32	23.25	0
	15	23.85	17.7	14.35	0.39	40.5	0
MHT	4	18.66	17.39	16.83	0.54	64	0
	6	24.02	19.64	17.27	0.46	47.25	0
	8	46.04	37.52	29.41	0.66	48.25	0
	12	27.44	19.88	13.88	0.49	54.5	0
	15	29.04	21.4	14.73	0.53	57	0

C. *Tracking results of 3 people scenario*

The trajectory of 3 people randomly walking is shown in Figure 5 when using 12 channels for angle estimation. Firstly, the trajectories of all 3 people are recognizable. The GNN and JPDA methods present an ID switching between the person in the middle and the one on the left. The MHT result shows more false trajectories than for GNN and JDPA. Quantitative results across the different metrics are shown in Table 4. The mean OSPA is approximately 0.3 when using 12 channels for GNN and JPDA, but increases significantly if fewer channels are used.



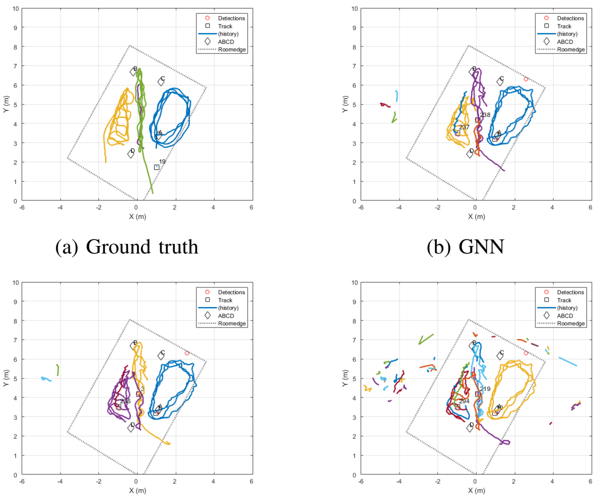
(a) Ground truth (b) GNN

(c) JPDA (d) MHT

Fig. 4. Ground truth and tracking results of 2 people parallel walking side by side with 0.5m spacing when using 12 MIMO channels.

Table 4. Quantitative metrics for 3 people random walking (2 Side 1 Middle)

Tracker	Channels	RMSE (cm)	MAE (cm)	Median (cm)	Mean of OSPA	Unassigned Tracks (%)	Miss Detection (%)
GNN	4	25.55	22.29	18.81	0.65	81.25	3.5
	6	24.76	22.13	19.14	0.39	36.25	0
	8	26.93	23.86	19.4	0.38	33.25	0
	12	25.48	23.38	21.24	0.3	18	0
	15	28.06	24.76	21.28	0.33	24	0
JPDA	4	25.95	22.56	18.92	0.65	82.5	3.5
	6	24.76	22.13	19.14	0.39	36.25	0
	8	27.04	23.92	19.53	0.37	30	0
	12	25.45	23.43	21.5	0.3	18	0
	15	24.84	23.22	21.27	0.32	25.75	0
MHT	4	32.72	29.13	24.4	0.57	63.25	0
	6	27.41	23.97	20.2	0.35	29.5	0
	8	27.23	23.17	18.44	0.52	65	0
	12	29.12	24.86	20.34	0.44	46.5	0
	15	25.74	22.76	19.02	0.46	57	0



(a) Ground truth (b) GNN

(c) JPDA (d) MHT

Fig. 5. Ground truth and tracking results of 3 people random walking (2 Side 1 Middle) when using 12 MIMO channels.

IV. Conclusion

In this paper, a processing pipeline for indoor people tracking with three different algorithms is implemented and evaluated using experimental data from an FMCW MIMO radar with relatively low bandwidth. Scenarios with up to 3 people walking in a furnished room were collected. For each tracker, tracking results from various numbers of MIMO channels used for angular estimation are also compared.

The results demonstrate that both the GNN and JPDA methods achieve comparable performance, with localization MAE around 23 cm and OSPA 0.23 when using 6 channels for single person scenarios. For 2 people walking in parallel, finer angular resolution is required to distinguish the people, and when using 12 channels the MAE is around 21 cm and OSPA 0.32. For 3 people randomly walking, the MAE is around 24 cm and the OSPA is 0.3 when using 12 channels.

Acknowledgment

This research was in part financially supported by Huawei Sweden Gothenburg Research Center. The authors thank the Huawei team Zhong Chen, Yanming Wu and Jingjing Chen for the technical discussions, and are grateful to the volunteers at TU Delft who helped with the data collection.

References

[1] D. Wang, F. Fioranelli, and A. Yarovoy, *Analysis of Processing Pipelines for Indoor Human Tracking using FMCW radar*, Accepted for presentation at IEEE RadarConf'24, Denver, USA, Feb. 2024. arXiv: 2402.19124 [eess].

[2] Texas Instruments, *People Tracking and Counting Reference Design Using mmWave Radar Sensor*. [Online]. Available: https://www.ti.com/lit/ug/tidue71d/tidue71d.pdf.

[3] A. Ninos, J. Hasch, M. Heizmann, and T. Zwick, "Radar-Based Robust People Tracking and Consumer Applications," *IEEE Sensors Journal*, vol. 22, no. 4, pp. 3726–3735, Feb. 2022, ISSN: 1558-1748.

[4] D. Wang, J. Park, H.-J. Kim, K. Lee, and S. H. Cho, "Noncontact Extraction of Biomechanical Parameters in Gait Analysis Using a Multi-Input and Multi-Output Radar Sensor," *IEEE Access*, vol. 9, pp. 138 496–138 508, 2021, ISSN: 2169-3536.

[5] L. Ren, A. Yarovoy, and F. Fioranelli, "Grouped People Counting Using mm-wave FMCW MIMO Radar," *IEEE Internet of Things Journal*, pp. 1–1, 2023, ISSN: 2327-4662, 2372-2541.

[6] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, ser. KDD'96, Portland, Oregon: AAAI Press, Aug. 1996, pp. 226–231.

[7] I. J. Cox, "A review of statistical data association techniques for motion correspondence," *International Journal of Computer Vision*, vol. 10, no. 1, pp. 53–66, Feb. 1993, ISSN: 0920-5691, 1573-1405.

[8] S. S. Blackman and R. Popoli, "Design and analysis of modern tracking systems," *(No Title)*, 1999.

[9] S. S. Blackman, "Multiple-target tracking with radar applications," *Dedham*, 1986.

[10] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82–100, 2009.

[11] J. R. Werthmann, "Step-by-step description of a computationally efficient version of multiple hypothesis tracking," vol. 1698, pp. 288–300, Aug. 1992. DOI: 10.1117/12.139379.

[12] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, "Multiple hypothesis tracking revisited," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 4696–4704.

[13] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE transactions on signal processing*, vol. 56, no. 8, pp. 3447–3457, 2008.