mmSCALE: Self-Calibration of mmWave Radar Networks from Human Movement Trajectories

Anish Shastri*§, Marco Canil‡, Jacopo Pegoraro‡, Paolo Casari*, and Michele Rossi‡

* Department of Information Engineering and Computer Science, University of Trento, Italy

‡ Department of Information Engineering, University of Padova, Italy

***Abstract*—We present mmSCALE, a practical self-calibration method that automatically estimates the relative position and orientation of a network of millimeter wave (mmWave) radars by post-processing the trajectories of detected targets that move within the radars' fields of view (FoVs). This is a key component of multi-device mmWave radar deployments for indoor human sensing. As commercial mmWave radars have limited range (up to 6-8 m) and are subject to occlusion, covering large indoor spaces requires multiple radars. A fully self-contained system should estimate the location and orientation of each radar with no intervention by a human operator. To solve this problem, mmSCALE fuses target detections from multiple radars, yielding median errors of 0.18 m and 2.86° for radar location and orientation estimates, respectively. For this, mmSCALE requires no specific target trajectories or controlled conditions, it autonomously assesses the calibration quality over time, and is robust to occlusion and to the presence of multiple subjects.**

***Index Terms*—Indoor sensing; mmWave radar; target tracking; self calibration; self registration**

I. Introduction and Related Work

The increasingly growing interest in the use of millimeter waves (mmWaves) for human tracking [1], [2] and activity recognition [3], [4] demands solutions to improve the usability and practicality of such systems. Frequency-Modulated Continuous-Wave (FMCW) radars working in the mmWave band have emerged as valid alternatives to cameras for indoor monitoring, as they are robust to changing and poor lighting conditions and do not raise privacy concerns [5]. However, commercial mmWave radars have limited range [1] (up to 6-8 m) and are subject to occlusion [2]. Covering medium-to-large indoor spaces thus requires multiple radars (radar networks), with *known* position and orientation. This raises the problem of how to automatically obtain the positions and orientations of the radars, as it is often impossible to know them in advance, and it is impractical to manually input these settings at deployment time. This problem remains unsolved in the existing literature. To the best of our knowledge, only two works have tackled it, i.e., [6], [7]. In [6], the walking trajectory of a single person moving along a straight line (estimated via a Kalman filter (KF) [8]) is used to compute the relative position and orientation of two radars. Even though results are accurate (< 10 cm position and $< 1^\circ$ orientation errors), the system has limitations, as *(i)* only straight-line human movement trajectories are supported, which is unrealistic, and *(ii)* it only works when a single person is being tracked. Both aspects require that the radar network performs a dedicated calibration phase in controlled conditions, which is impractical and time-consuming. The authors of [7] propose a similar algorithm for multiple sitting subjects. Here, data association between different radars is accomplished by pairing the reflections from the same target seen by two radars. The pairing employs a measure of similarity between the respiration waveform of a subject. While this method tackles the multiple target problem, it only works when the subjects are stationary, significantly limiting its application scope.

In this work, we propose mmSCALE, a practical algorithm that automatically estimates the locations and orientations of multiple radars with respect to a reference coordinate system by leveraging the movement trajectories of people in the environment. Our method solves the problem of handling free movement trajectories and multiple subjects using *(i)* a standard KF-based tracking routine, *(ii)* an efficient, singular value decomposition (SVD) least-squares (LS) approach to obtain the relative position and orientation of an arbitrary number of mmWave radars, and *(iii)* the Hungarian algorithm, applied to an originally designed cost function, to associate the tracks detected by different radars. The original contributions of the present work are

1) We propose mmSCALE, a fully automated method for the self-calibration of multiple mmWave radars. The algorithm estimates the relative positions and orientations of the radars with a median error of 0.18 m and 2.86°, respectively, using the information obtained by tracking people moving freely in the FoV of multiple radars.
2) mmSCALE is highly practical and requires no controlled conditions for the calibration. It can support an arbitrary number of radars. Thanks to a novel data association cost function, it handles multiple moving subjects in the environment even when occlusions occur. Its fast convergence time enables accurate calibration in less than 6 seconds.
3) We evaluate our method via an extensive measurement campaign involving up to 4 commercial mmWave radars and multiple subjects, deployed in realistic conditions including possibly challenging human motion.

§ Corresponding author email: anish.shastri@unitn.it

This work has been supported by the EU MSCA ITN project MINTS "Millimeter-wave networking and sensing for beyond-5G" (Grant no. 861222) and by MIUR (Italian Ministry of Education, University and Research) through the "Departments of Excellence" initiative (Law 232/2016).

The authors thank Enver Bashirov for his help in the realization of the radar sensing experiments.

The remainder of this paper is organized as follows: Section II gives preliminary description of mmWave radar-based human tracking; Section III formalizes the self-calibration problem; Section IV presents and discusses our approach; Section V describes the experimental results on our testbed; finally, we draw concluding remarks in Section VI.

II. Preliminaries on mmWave indoor radar

A Multiple-Input Multiple-Output (MIMO) FMCW radar jointly estimates the distance, the radial velocity, and the angular position of the targets with respect to the radar device [9]. During the sensing process, the radar transmits sequences of linear chirp signals with bandwidth B. A full sequence, or "radar frame", is repeated with period T seconds.

1) Distance, velocity and angle estimation: The distance, r, and velocity, v, of the targets are computed from the frequency shift induced by the delay of each reflection, usually applying discrete Fourier transform (DFT) processing. The FMCW radar distance resolution is related to the bandwidth B by $\Delta r = c/(2B)$, where c is the speed of light. This makes mmWave devices accurate to the level of a few centimeters using a bandwidth of $2-4$ GHz [2], [10]. Furthermore, using a 2D array of multiple receiving antennas makes it possible to obtain the angle-of-arrival (AoA) of the reflections along the azimuth (θ), and the elevation (ϕ) dimensions, by leveraging phase shifts across different antenna elements. The azimuthal AoA resolution depends on the number of antennas N in the array and is given by $\Delta\theta = \lambda/(Nd\cos\theta)$, where d is the spacing between the antennas.

2) Radar point-clouds: A human presence in the environment generates a large number of reflections in the form of points detected by the radar. This set of points, usually termed *radar point-cloud*, can be transformed into the 3-dimensional Cartesian space using the distance, azimuth, and elevation angle information of the multiple body parts. Each point is described by vector $[x, y, z]^T$, including its spatial coordinates x, y, z obtained transforming r, θ and ϕ.

3) People tracking: the common approach to people tracking from mmWave radar point-clouds [1], [2], [11] includes *(i)* a detection phase via density-based clustering algorithms (e.g., DBSCAN [12]) to separate the reflections from multiple subjects, *(ii)* applying Kalman filtering techniques [8] on each cluster centroid to track the movement trajectory of each subject in space.

In this paper, we estimate the subjects' trajectories in the (x, y) horizontal plane. We define the *state* of each subject at time k as $\mathbf{s}_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]^T$, where $\dot{x}_k$ and $\dot{y}_k$ are the velocity components along the two axes. We assume that the state evolution obeys $\mathbf{s}_k = \mathbf{C}\mathbf{s}_{k-1}$, where the transition matrix $\mathbf{C}$ represents a constant-velocity (CV) model [13]. The KF estimates the state $\hat{\mathbf{s}}_k$ for a target subject at time k.

III. Problem outline

In this work, we tackle the problem of automatically estimating the position and orientation of multiple radar devices on the *azimuth* (horizontal) plane with respect to a common reference system by leveraging the movement trajectories of one or more targets moving in the radars' FoV.

A. System self-calibration

Assume S radars are deployed in the space and call F_i, $i = 1, \dots, S$, their reference systems (RSs). Each RS is composed of a pair $F_i = (\mathbf{t}_i, \mathbf{R}_i)$, where $\mathbf{t}_i = [x, y]^T$ represents the position of the origin of the i-th RS and $\mathbf{R}_i$ corresponds to the 2×2 rotation matrix identifying its orientation.

We elect the RS of one radar to be the reference RS, and refer all other RSs to the reference one. Therefore, the common RS is known, whereas all the others are unknown. In particular, assuming F_1 is chosen as the reference device, it holds that $F_1 = (\mathbf{t}_1, \mathbf{R}_1)$, with $\mathbf{t}_1 = [0, 0]^T$ and $\mathbf{R}_1 = \mathbf{I}_2$, the 2×2 identity matrix. Hence, our objective is to estimate F_i, $i = 2, \dots, S$, since F_1 is known.

Consider, now, a set $\mathcal{Q} = \{\mathbf{Q}_1, \dots, \mathbf{Q}_S\}$ of S tracks obtained by synchronously tracking the same target at each radar device. Each track is a matrix containing the position vectors of the target across time, as seen from a different perspective (i.e., from a different radar). Denote by $\mathbf{q}_k$ the (x, y) position of the target (i.e., the first two components of the KF state estimate $\hat{\boldsymbol{s}}_k$) with respect to F_1 at time k, and with $\mathbf{u}_k$ the (x, y) position of the target with respect to the generic RS F_i, at the same time instant. It holds that

$$\mathbf{q}_k = \mathbf{R}_i \mathbf{u}_k + \mathbf{t}_i, \tag{1}$$

where $\mathbf{R}_i$ and $\mathbf{t}_i$ represent, respectively, the rotation matrix and translation vector to move from F_i to F_1. Considering a number P of consecutive pairs of estimated positions in tracks $\mathbf{Q}_1 \in \mathbb{R}^{2\times P}$ and $\mathbf{Q}_i \in \mathbb{R}^{2\times P}$, we get an over-determined system of $2P$ independent equations corrupted by noise, due to the imperfect estimation of the positions by the KF. The desired pair $(\mathbf{R}_i, \mathbf{t}_i)$ can be found solving the LS problem

$$(\mathbf{R}_i, \mathbf{t}_i) = \underset{\substack{\mathbf{R}_i \in SO(2),\\ \mathbf{t}_i \in \mathbb{R}^2}}{\arg\min} \sum_{k=1}^{P} ||(\mathbf{R}_i \mathbf{u}_k + \mathbf{t}_i) - \mathbf{q}_k||_2, \; i = 2, \dots, S, \tag{2}$$

where $SO(2)$ denotes the special orthogonal group in dimension 2 and $||\cdot||_2$ is the Euclidean norm. In Section IV, we present an algorithm to efficiently solve this problem.

B. Challenges of self-calibration with mmWave real data

The automatic self-calibration of a radar network using targets moving in their FoV in a real-world scenario poses several challenges concerning *(i)* the presence of multiple tracks, *(ii)* the time synchronization of the radars, *(iii)* the association between the tracks, forming pairs to be used for the calibration, and *(iv)* how to assess the quality of the calibration once deployed in a previously unseen, uncontrolled scenario. All these points are briefly discussed below.

Multiple tracks. While in Section III-A we only considered the case of every radar providing only one track, in a real-world scenario, every device would likely provide a list of tracks. This is due to the presence of multiple targets,

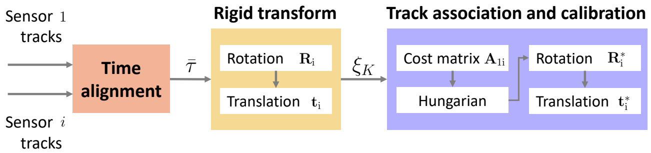


Fig. 1. mmSCALE worflow.

spurious ghost targets related to mmWave reflections on highly reflective surfaces, or the splitting of the trajectory of a single target into multiple tracks caused by occlusion events. Thus, we need a method to understand the correspondence between the tracks acquired by different radars. Moreover, not all the track pairs provide the same calibration quality. For instance, longer and accurately synchronized tracks are expected to perform better. In the following, focusing on a single time instant, we denote by N_1 the number of tracks from radar 1 and by N_i the number of tracks from radar i, indicized by $\ell = 1, \dots, N_1$ and $m = 1, \dots, N_i$, respectively.

Time synchronization. Some level of time synchronization between radars is needed, as it allows us to associate the pairs $(\mathbf{q}_k, \mathbf{u}_k)$ to be used in Eq. (2), which are obtained by different devices. To this end, once a data frame is acquired by one radar, we attach a timestamp and then use the timestamps to match data coming from different devices. We call τ_k^ℓ and τ_k^m the timestamps (in seconds) of the k-th position vector in tracks ℓ (of the reference radar 1) and m (of the i-th radar), respectively. As we empirically assess in Section V, synchronization on the radar frame duration level (typical frame rates range from 10 to 30 Hz) is sufficient for mmSCALE to work, so errors of a few milliseconds can be tolerated.

Assessment of the calibration process. As we envision a system that should generalize to unseen, uncontrolled, real-world scenarios, a procedure to evaluate the calibration quality in such conditions is required. To this end, we leverage the fact that, after calibration, associated tracks belonging to the same target should match closely. The same metric can also be used to verify the validity of the calibration across time and to decide whether re-calibration is required.

As further explained in Section IV-C, all these aspects are accounted for in the proposed cost function (Eq. (6)).

IV. mmSCALE WORKFLOW

From a high-level perspective, mmSCALE solves the calibration problem and the above challenges by operating in three steps, as shown in Fig. 1.

(1) Time alignment. Considering all the possible pairings of the reference radar 1 with radars $i = 2, \dots, S$, a time alignment between the N_1 tracks maintained by radar 1 and the N_i tracks from radar i is sought, by minimizing the difference between timestamps τ_k^ℓ and τ_k^m.

(2) Rigid transformation. Using the time alignment from point (1), we solve the problem in Eq. (2) for all aligned track pairs, obtaining the corresponding rotation matrices and translation vectors.

(3) Tracks association and radar calibration. Radar calibration is performed using the best matching track pairs in terms of time alignment and residual rigid transformation error. The key idea is that track pairs that are well aligned in time and that leads to low residuals in Eq. (2) represent the same target, as seen by radar 1 and i, respectively. Following this rationale, we design a new track-to-track association cost function and find the one-to-one track association yielding the minimum cost. Finally, the calibration of radar i is carried out by solving the rigid transformation problem Eq. (2) using all the points in the associated track pairs.

A. Time alignment

The ℓ-th and m-th tracks from radar 1 and i, respectively, are aligned in time by exploiting timestamps that are attached to each frame. This alignment is performed so that every element of track ℓ is associated with the element of track m that minimizes the time difference between the two acquisition instants, reducing the tracks to a common length of K time-aligned positions. Elements of track ℓ that do not have a corresponding element of track m within T seconds (recall that T is the duration of a time frame) are discarded. Due to this last operation, our time alignment procedure selects only the portions of the tracks that are sufficiently well synchronized, in order to avoid performing the rigid transformation on wrongly associated points. We define the *mean time alignment* of the (ℓ, m) pair as

$$\bar{\tau}(\ell, m) = \frac{1}{K} \sum_{k=1}^{K} |\tau_k^\ell - \tau_k^m|. \quad (3)$$

The value of $\bar{\tau}(\ell, m)$ represents the quality of the time alignment between the two tracks, and will be used in the track association step (see Section IV-C).

B. Rigid transformation

Let the subject's trajectories observed from the reference radar 1 and the i-th radar (after time alignment) respectively be $\mathbf{Q}_1 \in \mathbb{R}^{2\times K}$ and $\mathbf{Q}_i \in \mathbb{R}^{2\times K}$. Then, the transformed trajectory $\mathbf{Q}_{i1}$ is given by $\mathbf{Q}_{i1} = \mathbf{R}_i\mathbf{Q}_i + \mathbf{t}_i$. We use the subscript $_{xy}$ to represent a transformed trajectory from the RS of radar x to that of radar y. To find the optimal $\mathbf{R}_i$ and $\mathbf{t}_i$, we solve the LS problem in Eq. (2), that minimizes the square of the Euclidean norm between $\mathbf{Q}_{i1}$ and $\mathbf{Q}_1$. The solution is obtained, in closed form, as [14]

$$\mathbf{t}_i = \bar{\mathbf{q}}_1 - \mathbf{R}_i\bar{\mathbf{q}}_i \ , \ \text{ with } \mathbf{R}_i = \mathbf{U}\mathbf{V}^T \ , \quad (4)$$

where, $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_i$ are the mean positions of the trajectories $\mathbf{Q}_1$ and $\mathbf{Q}_i$, respectively, and the columns of $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors obtained after the SVD of the covariance matrix $\mathbf{X}\mathbf{Y}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$. Here, $\mathbf{X} = \mathbf{Q}_i - \bar{\mathbf{q}}_i$ and $\mathbf{Y} = \mathbf{Q}_1 - \bar{\mathbf{q}}_1$. The translation vector obtained after the rigid transformation gives the position of radar i in the RS of the reference radar. We can also compute the orientation angle of radar i using the $\mathbf{R}_i$ matrix. This orientation angle, defined by θ_i, is given as $\theta_i = \cos^{-1}\left(\frac{1}{2}tr(\mathbf{R}_i)\right)$, where $tr(\cdot)$ is the trace of the matrix. mmSCALE computes the rigid transformation parameters $(\mathbf{R}_i^{(\ell,m)}, \mathbf{t}_i^{(\ell,m)})$ for all the aligned

(ℓ, m) track pairs from radars 1 and i, $i = 2, \ldots, S$, as part of the following association step.

C. Tracks association and radar calibration

Tracks association. Our data association strategy consists in *(i)* computing a *cost* for each association ($\ell \leftrightarrow m$), and *(ii)* solving the resulting combinatorial cost minimization problem to associate the best matching track pairs. Since we operate in real-world conditions, as mentioned in Section III, our cost function needs to incorporate different aspects: *(a)* the length of the tracks, as longer tracks are assumed to provide a better calibration; *(b)* the time alignment of the tracks, as we should compare positions acquired almost simultaneously by the different radars; and *(c)* the quality of the rigid transformation, in terms of residual error in superimposing tracks form the different radars. We define

$$\xi(\ell, m) = \sum_{k=1}^{K} || \, \mathbf{q}_k^{\ell} - \mathbf{R}_i^{(\ell,m)} \mathbf{u}_k^m - \mathbf{t}_i^{(\ell,m)} \, ||_2, \tag{5}$$

where $\mathbf{q}_k^{\ell}$ and $\mathbf{u}_k^m$ are the (x, y) positions at time k in tracks ℓ and m, respectively. $\xi(\ell, m)$ is the residual sum of errors in the trajectories after applying the time alignment and the rigid transformation and measures the quality of the solution to the LS problem. Then, the association cost, A, for the tracks pair (ℓ, m), is defined as

$$A(\ell, m) = -\rho(K, \bar{\tau}) \left(1 + \xi(\ell, m)\right)^{-1}, \tag{6}$$

where $\rho(K, \bar{\tau})$ is a corrective term that depends on the length of the tracks and on their mean time alignment. Recalling that T is the sampling interval of the system, the corrective term is formalized as

$$\rho(K, \bar{\tau}) = \frac{\log(KT)}{1 + \bar{\tau}(\ell, m)}, \tag{7}$$

where $\log$ denotes the natural logarithm. Note that this corrective term favors tracks which are longer and better time aligned, and penalizes the others. Costs $A(\ell, m)$, $\ell = 1, \ldots, N_1$, $m = 1, \ldots, N_i$, are arranged into a $N_1 \times N_i$ cost matrix, $\mathbf{A}_{1i}$, and the optimal association of tracks is obtained by minimizing the overall cost, computed through the Hungarian algorithm [15].

Radar calibration. The Hungarian algorithm yields $\min(N_1, N_i)$ pairs of associated tracks, which, according to the mmSCALE rationale, are possibly the same targets seen by the two radars. Due to the presence of spurious tracks, ghost targets and clutter, we select a subset of the associated track pairs that have a cost below a threshold $A_{\rm th}$, which represents a confidence value under which the pair is truly a track pair generated by a human. The N_t track pairs which are selected in this way are then stacked together and used to set up a rigid transformation problem as in Eq. (2) that includes all the information available from multiple subjects. The problem is solved with the same procedure described in Section IV-B, obtaining the final rotation matrix and translation vector to be used to calibrate radar i, namely $(\mathbf{R}_i^*, \mathbf{t}_i^*)$. This step exploits all the available information from multiple subjects, improving the calibration accuracy by increasing the number of useful measurements per time frame. Note that even though target occlusion events may split a trajectory into multiple components, our algorithm still works by exploiting each resulting sub-trajectory.

V. Experimental results

We implemented mmSCALE using Texas Instruments IWR1843BOOST mmWave radars[1] connected to two NVIDIA Jetson TX2 edge computing devices[2] communicating via Ethernet. The radars operate in the 77-81 GHz band in real-time at a frame rate of $1/T = 15$ Hz with a FoV of $\pm 60^\circ$ and $\pm 15^\circ$ over the azimuth and elevation planes, respectively. In this section, we present the experimental results obtained by testing the system in several different scenarios, with 2 and 4 radar devices and up to 2 concurrently moving subjects.

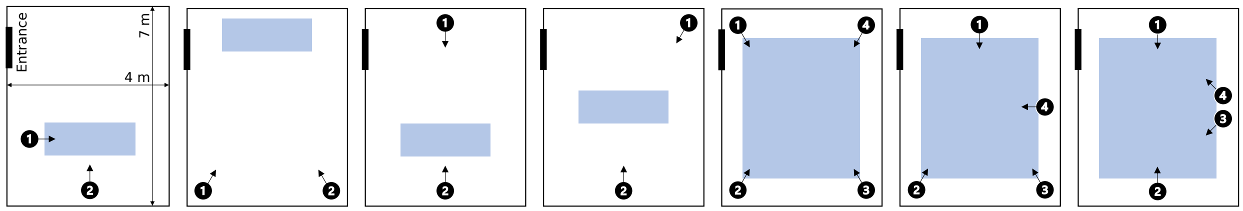
A. Measurements setup and Dataset

To assess the performance of mmSCALE, we conducted tests in a 7×4 m research laboratory (see Fig. 2a) equipped with a motion tracking system featuring 10 cameras. This provides the ground truth (GT) 3D localization of a set of four markers placed on each radar[3] with millimeter-level accuracy. We considered 4 scenarios with 2 radars and 1 or 2 moving targets, and 3 scenarios with 4 radars and 1 moving target. The locations and orientations of the radars in the different setups are shown in Fig. 2, where the black circles represent the radars, the arrows identify their orientations, and the blue area the region of the laboratory where the subjects were allowed to move. We also asked the subjects to move according to 3 possible different trajectories: *(i)* straight, identifying a movement along a straight line, *(ii)* 8-shape, and *(iii)* free, corresponding to an arbitrary trajectory. In the following, we use the notation xR-yT to identify a measurement sequence involving x radars and y targets.

B. Performance evaluation

I. Qualitative results. The input of mmSCALE is a set of tracks obtained from different radars, and the output is the estimation of their position and orientation with respect to a common reference system, which, for convenience, we consider to be coinciding with that of radar 1. Fig. 3 shows a qualitative example of the calibration process with a target moving freely. Here, after finding the optimal rotation and translation parameters, we applied the rigid transformation to the trajectory seen by radar 2 (blue line), so as to superimpose it with the one of radar 1 (orange line). The transformed trajectory (yellow line) matches well with the reference one, showing a good calibration result. We represent the reference radar with a purple triangle (located at $[0, 0]^T$), while the black square and the pink triangle mark the estimated position of radar 2 and its ground truth, respectively.

[1] https://www.ti.com/tool/IWR1843BOOST
[2] https://developer.nvidia.com/embedded/jetson-tx2
[3] Using four markers allows us to estimate the orientation angle.



(a) (b) (c) (d) (e) (f) (g)

Fig. 2. Illustration of the setups used to test mmSCALE. The numbered dots represent the radar devices, with the arrow identifying the pointing direction of the radars. The blue area corresponds to the region where the target moved.

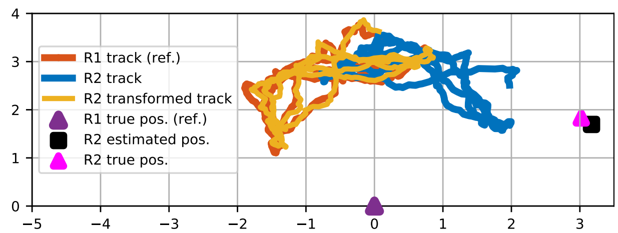
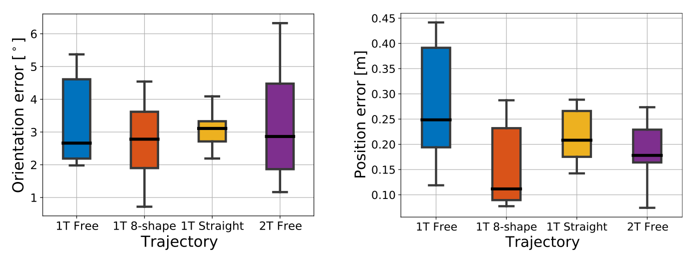


Fig. 3. Example of the transformation, via mmSCALE, of a free movement trajectory using a 2R-1T setup.



(a) Orientation error. (b) Position error.

Fig. 4. Orientation and position errors with 2 radars and one (1T) or two (2T) subjects on the scene.

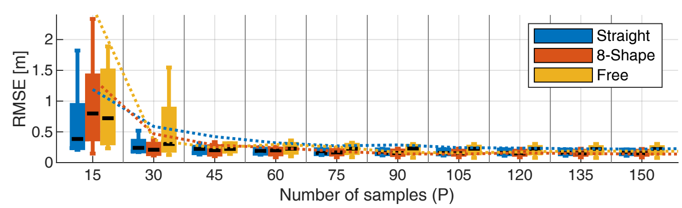


Fig. 5. Statistical dispersion of trajectory RMSE when calibrating with P points from the trajectory in various 2R-1T setups. The dotted lines represent the median of the radar position using P samples.

II. Position and orientation errors. To evaluate mmSCALE's accuracy in estimating the positions and orientations of multiple radar devices, we first consider four 2R-1T and four 2R-2T setups (Fig. 2a and Fig. 2d), applying our method to calibrate radar 2 with respect to radar 1. For every setup and every trajectory shape (see Section V-A), we collected two 40 s-long sequences, for a total of 32 sequences. However, since we observed that a time window as short as 6 s is sufficient for our algorithm to converge (see Section V-B-III) we only used the first 6 s of the tracks for all the calibrations.

We define the *orientation error* as the absolute value of the difference between the true orientation angle of the radar device and the estimated one, which are derived, after calibration, from the rotation matrices, as explained in Section IV-B. The *position error* is defined as the Euclidean distance between the estimated position of the radar device and its true location.

The calibration performance is summarized in Fig. 4, which shows the distribution of the orientation and localization errors for the three trajectory types in a single-target setting (1T), and for the free movement case in a 2 targets setting (2T). Regarding the orientation error, we observe that its median, for all the three trajectories, is about 3°, without a great impact of the shape of the trajectory. On the other hand, the latter influences the radar position estimation, where the best results are provided by the 8-shape trajectory, reaching errors as low as 11 cm. The worst case is represented by the free movement, giving a median error of 25 cm.

With 2 concurrently (and freely) moving subjects, mmSCALE obtains similar orientation angle estimation accuracy, with a slightly increased variance, whereas the positioning error significantly improves with respect to the single free moving target case (see also Tab. 1). This is a distinctive advantage of the track association process, which can simultaneously exploit all trajectories seen by the two radars. As a result, the number of measurements available for the calibration process increases over the same observation period. With 2 subjects, we also empirically assessed suitable values for the threshold parameter A_{thr}. Our results showed that spurious tracks lead to association costs closer to 0 in all the observed cases, whereas complete human trajectories have costs strictly lesser than -2. Hence, we set $A_{\text{thr}} = -2$.

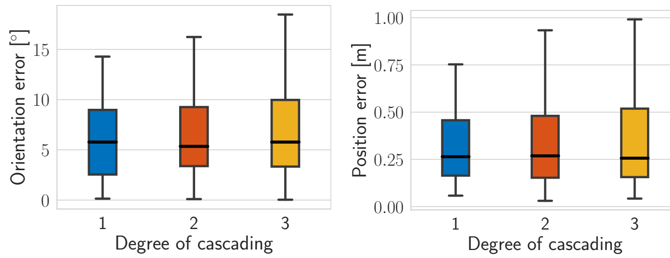
III. Calibration quality assessment. We now examine the number of samples mmSCALE needs in order to achieve a sufficiently low residual calibration error, namely the value of P in Eq. (2), and propose a practical method to assess the calibration quality at run time. As a measure of the calibration quality, we use the residual error after transforming the trajectories of radar i with respect to radar 1 (the reference); if tracks from the two radars match well, we can assume the calibration has reached a low positioning and orientation error. This is demonstrated in Fig. 5, which shows the distributions of the RMSEs obtained by comparing the transformed track from radar i with the one from the reference radar 1. To study the impact of the number of samples, we only used the first P trajectory points to compute the rigid transformation and then transformed the entire sequence. P is varied from 15 to 150, which corresponds to a time duration from 1 to 10 seconds. With the dotted curves instead, we show the median error on the estimated position of radar i with respect

TABLE 1
SUMMARY OF THE MEDIAN ERRORS FOR DIFFERENT SETUPS

	2 Radar setup				4 Radar setup
Median error	**1T Straight**	**1T 8-Shape**	**1T Free**	**2T Free**	**1T Free**
Orientation [°]	3.1	2.78	2.66	2.86	5.6
Position [m]	0.21	0.11	0.24	0.18	0.26

to the GT. In practice, the GT is not available, so only the transformed trajectory RMSE can be computed. We observe that the median RMSE of all the trajectories falls below 0.5 m with just 30 samples, i.e., 2 seconds of data. The RMSE distributions for the straight and 8-Shape trajectory reach a plateau with 45 samples (3 seconds), whereas, for the free trajectory, it takes about 75 samples (5 seconds). The estimation error on the radar position also reaches low values within 90 samples (6 seconds), showing an identical trend to the trajectory RMSE. Therefore, 6 seconds are sufficient to assess the quality of the calibration. If needed, a re-calibration can be performed by collecting another set of trajectories and re-applying mmSCALE.

IV. Cascaded calibration. mmSCALE works by calibrating pairs of radars, taking one of them as the reference. In practical indoor radar networks scenarios, the aim is to have a single global reference system and calibrate multiple radars with respect to it. We expect that not all the deployed radars have a FoV that partially overlaps with the reference radar, i.e., the origin of the global reference system. As a result, direct one-to-one calibration is infeasible. In these cases, mmSCALE can be applied in *cascade*, calibrating each radar with respect to another radar with which it partially shares the FoV, e.g., in Fig. 2e, radar 2 with respect to radar 1, radar 3 with respect to radar 2, and so on. We define a *cascade sequence* as the ordered list of the radar numbers according to the calibration order ($[1, 2, 3, 4]$ in the previous example from Fig. 2e). Furthermore, we refer to the *degree* of a radar in the cascade sequence as its position in the list, starting from 0, which is the degree of the reference radar. The final calibration with respect to the reference radar can be obtained by composing the obtained rigid transformations sequentially. Note that cascaded calibration can occur in any order, as long as adjacent radars in the cascade sequence have overlapping FoVs. We applied this technique to the 4R-1T setups shown in Fig. 2e-2g. We consider all feasible cascading permutations: e.g., in Fig. 2g all permutations involving the calibration of radar 4 with respect to 3 (or vice-versa) are infeasible as their 120° FoVs do not overlap. Fig. 6 shows the calibration error distributions for radars with different degree in the cascade. The median errors are almost unaffected by cascaded calibration, while the variance is slightly higher for higher degrees in the cascade because the errors in θ and $\mathbf{t}$ keep adding up linearly as the degree of the radar increases. The "4 Radar setup" section of Tab. 1 reports the overall median errors across all the degrees of cascade. The above results show that mmSCALE can keep the cascaded calibration error within reasonable levels in realistic scenarios, where a room would include a network of typically up to 3-4 radars.



(a) Orientation error. (b) Position error.

Fig. 6. Statistical distributions of all the orientation errors and position errors of the radars for different degrees of cascade in the 4R-1T setups.

VI. CONCLUSIONS

In this paper, we proposed mmSCALE, a practical and fully autonomous self-calibration method to estimate the relative position and orientation of multiple mmWave radars, by exploiting the tracking of people moving in the environment. We evaluated the system in several realistic conditions, including up to 4 radars and 2 subjects, achieving median position and orientation errors of 0.18 m and 2.86°, while still allowing targets to move freely during the calibration process. Moreover, our system requires a calibration window as short as 6 s, handles multiple targets, and is robust to occlusion events. These results suggest that mmSCALE can be a viable candidate to calibrate mmWave radar networks for real-time applications. Future work will incorporate our method for real-time multi-radar people tracking, identification, and activity recognition in large indoor spaces.

REFERENCES

[1] P. Zhao *et al.*, "mID: Tracking and Identifying People with Millimeter Wave Radar," in *Proc. DCOSS*, May 2019.

[2] J. Pegoraro and M. Rossi, "Real-time people tracking and identification from sparse mm-wave radar point-clouds," *IEEE Access*, May 2021.

[3] P. Vaishnav and A. Santra, "Continuous Human Activity Classification With Unscented Kalman Filter Tracking Using FMCW Radar," *IEEE Sensors Lett.*, Apr 2020.

[4] R. G. Guendel *et al.*, "Continuous human activity recognition for arbitrary directions with distributed radars," in *Proc. IEEE RadarConf21*, May 2021.

[5] S. A. Shah and F. Fioranelli, "RF sensing technologies for assisted daily living in healthcare: A comprehensive review," *IEEE Aerosp. Electron. Syst. Mag.*, Nov 2019.

[6] S. Li *et al.*, "Pedestrian trajectory based calibration for multi-radar network," in *Proc. IEEE INFOCOM WKSHPS*, May 2021.

[7] S. Iwata *et al.*, "Multiradar data fusion for respiratory measurement of multiple people," *arXiv preprint arXiv:2107.11525*, 2021.

[8] R. E. Kalman, "A new approach to linear filtering and prediction problems," *ASME Transactions, Journal of Basic Engineering*, 1960.

[9] S. M. Patole *et al.*, "Automotive radars: A review of signal processing techniques," *IEEE Signal Process. Mag.*, Mar 2017.

[10] N. Knudde *et al.*, "Indoor tracking of multiple persons with a 77 GHz MIMO FMCW radar," in *Proc. EURAD*, Oct 2017.

[11] Z. Meng *et al.*, "Gait Recognition for Co-Existing Multiple People Using Millimeter Wave Sensing," in *Proc. AAAI*, Feb 2020.

[12] M. Ester *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *2nd International Conference on Knowledge Discovery and Data Mining*, Aug 1996.

[13] T. Wagner *et al.*, "Radar signal processing for jointly estimating tracks and micro-Doppler signatures," *IEEE Access*, Feb 2017.

[14] O. Sorkine-Hornung and M. Rabinovich, "Least-squares rigid motion using SVD," *Computing*, 2017.

[15] Kuhn, Harold W, "The Hungarian method for the assignment problem," *Naval research logistics quarterly*, 1955.