Distributed Radar Fusion for Extended Target Location and Velocity Reconstruction

Kruse, Nicolas; Guendel, Ronny; Fioranelli, Francesco; Yarovoy, Alexander

DOI
10.1109/RadarConf2458775.2024.10548772

Publication date
2024

Document Version
Final published version

Published in
RadarConf 2024 - 2024 IEEE Radar Conference, Proceedings

Distributed Radar Fusion for Extended Target Location and Velocity Reconstruction

Nicolas Kruse
Microwave Sensing, Systems, and Signals
Delft University of Technology
Delft, the Netherlands
n.c.kruse@tudelft.nl

Ronny Guendel
Microwave Sensing, Systems, and Signals
Delft University of Technology
Delft, the Netherlands
r.guendel@tudelft.nl

Francesco Fioranelli
Microwave Sensing, Systems, and Signals
Delft University of Technology
Delft, the Netherlands
f.fioranelli@tudelft.nl

Alexander Yarovoy
Microwave Sensing, Systems, and Signals
Delft University of Technology
Delft, the Netherlands
a.yarovoy@tudelft.nl

2024 IEEE Radar Conference (RadarConf24) | 979-8-3503-2920-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/RADARCONF2458775.2024.10548772

***Abstract*—The application of distributed radar to human motion monitoring is considered. A novel sensor fusion method has been proposed that yields a two-dimensional map of reflection intensity and a vector field of reconstructed velocities in lieu of conventional Doppler spectrograms or radial velocity components. The method has been verified using experimental datasets in two case studies involving fall detection in sequences of activities, and arm motion discrimination for in-place activities. A true positive rate and precision of respectively $99.3\,\%$ and $93.0\,\%$ have been demonstrated for the fall detection task, and the output of the proposed method for arm motion characterisation indicates suitability for classification in future research.**

***Index Terms*—Distributed Radar, Sensor Fusion, Human Monitoring, Fall Detection, Activity Classification**

I. Introduction

Human monitoring has important applications in such fields as security and healthcare support. As a sensor, radar presents a set of properties that make it an advantageous modality for these purposes. These include sensing capabilities in adverse lighting conditions, as well as the non-contact, privacy-preserving nature of radar imaging that may be advantageous for user compliance and acceptance. Specifically in the context of healthcare support, significant research efforts using contactless radar sensing have been directed towards fall detection to aid the prolonging of independent living for the older and more vulnerable population [1], [2]. Additionally, as part of the more general study of contactless human activity classification, radar utilisation is also investigated for such applications as gesture and sign language recognition [3]–[7].

In radar sensing and beyond, fusion of information from multiple sensors or multiple data representations is widely considered to be an effective means of increasing detection and classification performance. Specifically, sensor fusion is employed to improve, among others, gesture recognition [8], [9], sign language recognition [10], classifications of human activities [11], [12], and classification of gait [13]. Often in this research context, sensor fusion occurs in either an intermediate feature domain that is not easily interpretable by humans [14], or at a later stage in the form of e.g., decision fusion [13].

In this work, a novel signal-level sensor fusion method is proposed that takes inputs from a network of distributed radar sensors. The method outputs a reconstructed 2D map of reflection intensity in the horizontal plane, as well as a vector field of reconstructed absolute velocities, instead of conventional micro-Doppler spectrograms representing only radial velocity components. The proposed sensor fusion yields a data representation that remains intelligible to humans and which can be utilised for further processing to perform e.g. classification of motions featuring distinct velocity components in multiple directions. The utility of the method is demonstrated in two case studies with experimental data collected using a network of five radars. In the first study, the intensity map is computed for experimental sequences of human activities and employed to perform fall detection based on alterations in the shape of an extended target in the horizontal plane. In the second study, several arm motions are captured and the feasibility of discriminating between them is demonstrated using the computed velocity vector field.

The remainder of this work is organised as follows. In Section II, the proposed sensor fusion method and its outputs are described, followed by a description of the two case studies and experimental setups in Section III. The results of the case studies are shown in Section IV with conclusions and recommendations for future work given in Section V.

II. Proposed Method

Serving as input to the proposed processing pipeline are N complex-valued range-time matrices from N spatially distributed radar sensors. For computing the intensity map at time t_0, the interval $[t_0, t_0+CPI]$ is selected from all N range-time

This research is funded in part by the Dutch Research Council (NWO) through the project *RAD-ART* (Radar-aware Activity Recognition with Innovative Temporal Networks).

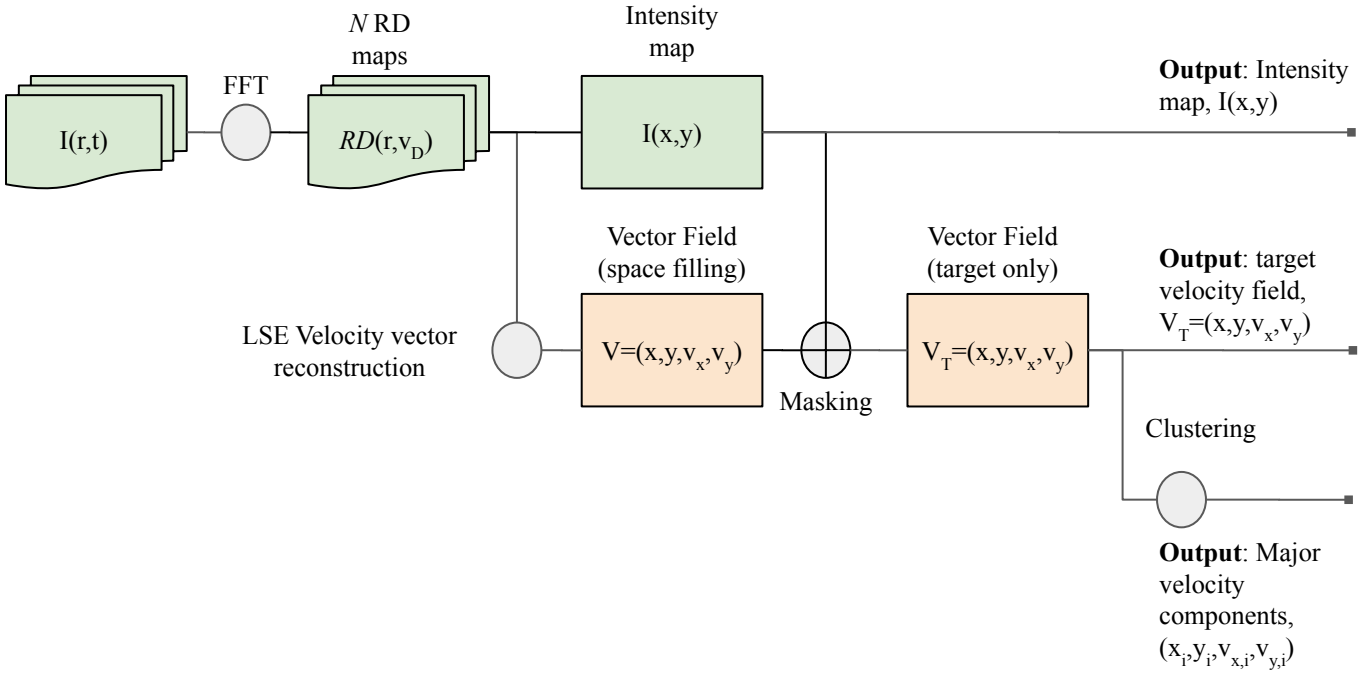

Fig. 1. Processing pipeline for the proposed sensor fusion method using a network of distributed radar sensors. The method outputs are organised in three horizontal branches.

matrices, where CPI indicates a chosen coherent processing interval.

Through the application of a Fast Fourier Transform (FFT) along the slow-time dimension, N complex range-Doppler maps $\mathcal{RD}_n$ are acquired. With the range-Doppler maps and the known positions of the radar sensors (x_n, y_n), the measured signal intensity I at an arbitrary position (x, y) can be determined by first computing the range to all N sensors and subsequently evaluating the respective range-Doppler maps at the computed range, as:

$$r_n(x,y) = \sqrt{(x-x_n)^2 + (y-y_n)^2} \tag{1}$$

$$I(x,y) = \sum_{n=1}^{N} \sum_{v_D} \left| \mathcal{RD}_n(r_n, v_D) \right|, \tag{2}$$

where v_D denotes Doppler or radial velocity which is removed through summation over the Doppler dimension. Note that the final intensity I at (x, y) is computed by summing the amplitude contributions from all N radar sensors, and that the vertical dimension z is not considered separately here.

Aside from the 2D intensity map, an estimated spatial distribution of velocities in the form of a vector field can be reconstructed using measurements from the N radar sensors. For a position (x, y), the computed ranges from (2) are again utilised to evaluate the range-Doppler maps at the corresponding range, but this time selecting the Doppler component with maximum intensity as:

$$v_n(x,y) = \mathrm{argmax}_{v_D} \left| \mathcal{RD}_n(r_n, v_D) \right|. \tag{3}$$

The N velocities v_n are all projections on the respective axes connecting the position (x, y) to the sensor locations (x_n, y_n). To retrieve the true velocity $\mathbf{v}(x, y)$ that gives rise to the projections $\mathbf{v_n}$ we establish from Figure 2 that the lines constructed from the orthogonal unit vectors $\hat{\mathbf{n}}_\mathbf{n}$ intersect at $\mathbf{v}$ in the case of ideal projection. For N non-ideal projection vectors, we can extend this approach by finding the best fitting intersection point of the N lines constructed from the unit vectors $\hat{\mathbf{n}}_\mathbf{n}$.

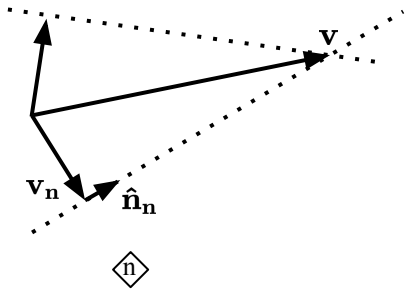

Fig. 2. Schematic representation of the reconstruction of the true velocity vector $\mathbf{v}$ from a set of N projected velocity vectors $\mathbf{v_n}$ and their respective orthogonal unit vectors $\hat{\mathbf{n}}_\mathbf{n}$. The location of a single radar sensor n is indicated with a diamond.

This problem can be approached as a least-square-error optimisation, where the sum of squares of distances from a point $\mathbf{v}'$ to the lines constructed from $\hat{\mathbf{n}}_\mathbf{n}$ is minimised as:

$$\mathbf{v} = \mathrm{argmin}_{\mathbf{v}'} \sum_n \mathcal{D}(\mathbf{v}'; \hat{\mathbf{n}}_\mathbf{n})^2, \tag{4}$$

with $\mathcal{D}$ a distance operator yielding the shortest Euclidean distance between a point and a line. A full derivation of the optimisation can be found at e.g., [15] and results in the following closed form expression for the true velocity $\mathbf{v}$:

$$\mathbf{v} = \left(\sum_n \hat{\mathbf{n}}_\mathbf{n} \hat{\mathbf{n}}_\mathbf{n}^T \right)^{-1} \left(\sum_n \hat{\mathbf{n}}_\mathbf{n} \hat{\mathbf{n}}_\mathbf{n}^T \mathbf{v}_n \right). \tag{5}$$

By utilising (5) a velocity vector field $\mathcal{V} = (x, y, v_x, v_y)$ can be computed, which contains the components (v_x, v_y) of the reconstructed absolute velocity vector at each position (x, y). With regards to the accuracy of the reconstructed field $\mathcal{V}$, two important assumptions are made:

1) Sensor range resolution is sufficiently high that a dominant velocity exists within a single range cell.
2) The Doppler component with the highest intensity measured by one sensor corresponds to those with highest intensity measured by the other sensors, allowing recombination of the projections.

The vector field $\mathcal{V}$ can be computed for every location in (x, y)-space, but is only physically interpretable at those locations where there is a target. For this reason, the intensity map $I(x, y)$ is utilised as a guide to mask $\mathcal{V}$ with both a geometrical mask and an intensity threshold. For the geometrical mask, the location of maximum intensity is first determined and then a square area of $2\,\mathrm{m}$ by $2\,\mathrm{m}$ around this point is selected, with the assumption that in this case humans are the objects of interest. For the intensity threshold, only those areas of $\mathcal{V}$ are kept where the intensity map exceeds a user-defined cutoff intensity. What remains after masking and thresholding is a vector field $\mathcal{V}_\mathbf{T}(x, y, v_x, v_y)$ describing the motions of the extended target as a function of location.

To improve interpretability of the vector field, a clustering algorithm can be applied to identify regions of homogeneous motion corresponding to e.g., an individual limb. Clustering is applied in (x, y, v_x, v_y) space to detect not only spatial proximity, but also comparable velocity vectors. For this work, a *k-means clustering algorithm* [16] with *k-means++* cluster seeding [17] is utilised due to the consistency of cluster sizes and locations between time steps.

III. Case Studies & Experimental Setups

Two case studies are conducted in order to demonstrate the capabilities of the proposed method. Specifically, these include an approach to fall detection using the intensity map and a study on distinguishing arm motions using the masked velocity vector field $\mathcal{V}_\mathbf{T}$.

In both case studies, the radar sensors utilised for the experimental data capture are Humatics PulsON P410 pulsed Ultra-Wideband(UWB) Single Input Single Output (SISO) radars with an antenna pattern that is approximately symmetric in azimuth. The sensors, one of which is shown in Figure 3, feature a centre frequency of 4.3 GHz, a bandwidth of 2.2 GHz, and a PRF of 122 Hz, resulting in a maximum unambiguous velocity of ±2.2 $\mathrm{m\,s^{-1}}$ and a range resolution of approximately $6.8\,\mathrm{cm}$. It should be noted that the sensors utilized provide only range and Doppler information, and no azimuth or elevation.

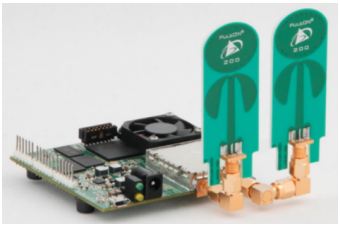

Fig. 3. A single Humatics PulsON P410 sensor with antennae, as used in the distributed networks in both case studies.

The preprocessing of data from the sensors is also identical for both case studies, with each sensor outputting a real-valued vector representation of the backscattered signal. The quadrature component of this signal is obtained through a Hilbert transform and a complex-valued fast-time/slow-time matrix can subsequently be constructed.

A. Fall Detection

To perform fall detection, the intensity map $I(x, y)$ is computed at intervals of 64 slow time steps ($0.52\,\mathrm{s}$) with a CPI of 32 slow time samples ($0.26\,\mathrm{s}$) for continuous sequences of human activities featuring falls. The sequences are part of a publicly available dataset [18] where sensors have been placed in a semicircular baseline of diameter $6.38\,\mathrm{m}$ at regular intervals. An image of the experimental setup is shown in Figure 5. The selected sequences comprise four distinct types for all 14 participants, with a different number of instances for each sequence type:

1) Walking around along arbitrary trajectories and falling at random intervals ($4\times$);
2) Falling from a stationary position at random locations and facing random directions ($4\times$);
3) A mix of various activities including sitting down and bending, performed at predetermined locations in the room, but with unconstrained duration ($1\times$);
4) The same mix of activities performed at random locations ($1\times$).

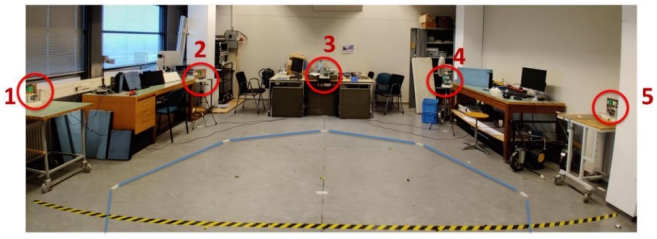

Fig. 4. The experimental setup for the fall detection case study. Five sensors are arranged in a semicircle at 45° intervals.

Since the intensity map can give insight in the approximate shape and size of the extended target, it can be utilised to determine when the target rapidly transitions from a state that

is localised in the (x, y)-plane (associated with an upright posture), to a state that is less localised (associated with a prone position during/following a fall event). To this end, the following steps are performed:

1) Determine the coordinates (x_c, y_c) of maximum intensity in the map I.
2) Apply a $2\,\mathrm{m}$ by $2\,\mathrm{m}$ geometrical mask around (x_c, y_c), with these dimensions assumed to contain the signature of a human participant regardless of their posture.
3) Compute the standard deviation σ of the Euclidean distances of all locations in the unmasked area with respect to (x_c, y_c), weighted by the intensity at the respective locations.

With the variable σ established as the measure of the spatial localisation of the extended target, a simple biLSTM (Long-Short Term Memory) neural network with 5 hidden units is used as a binary classifier. The network is trained to detect fall events based on the time evolution of the quantity σ. Due to the imbalance in the dataset between fall and non-fall events, a cross-entropy loss function is employed that weights the two classes according to their frequency in the dataset.

A leave-one-person-out (L1PO) scheme is adopted for performance evaluation, where the sequences of a single participant are used for testing the network which is trained on the sequences of the remaining 13 participants. This process is repeated to obtain a test result for all participants.

B. Arm Motion Classification

Utilising the vector field of reconstructed velocities enables a representation of human motions that combines information from multiple sensors whilst remaining physically interpretable. To demonstrate the potential of the method as a classification tool, several motions are recorded that feature limbs moving in separate directions while the target remains overall in a stationary location. For data capture, five sensors are arranged on the corners of a regular pentagon with sides of approximately $2.85\,\mathrm{m}$, as shown in Figure 5. Each radar sensor has a circular dead zone with a radius of $1\,\mathrm{m}$ within which no data is captured, and the target is located in the middle of the pentagon facing one of the sensors. Two motion types are initially captured for feasibility and shown in Figure 6: raising and lowering both arms laterally, and raising and lowering one arm forward, and one to the side simultaneously.

IV. Results

First the results for fall detection are discussed. As described in Section III-A a L1PO validation scheme has been employed and the performance metrics are shown on an individual participant basis in Table I. Out of a total of 973 fall events in 280 minutes of data, 966 have been correctly identified, yielding a True Positive Rate (TPR) of $99.3\,\%$. The amount of false alarms is 73 which corresponds to a precision of $93.0\,\%$. Inspection of the network predictions reveals that the precision is lower in the mixed sequences (i.e., sequences containing several types of activities mixed together), which

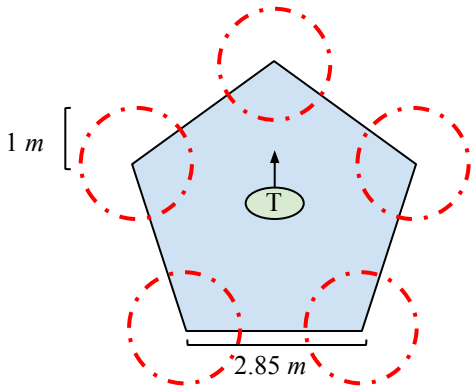

Fig. 5. The experimental setup with a network of five distributed radars utilised for the case study involving arm motion discrimination in Section III-B. Dashed red circles indicate radar sensor dead zones where no data is captured, and the extended target T is in the middle facing one of the corners of the regular pentagon.

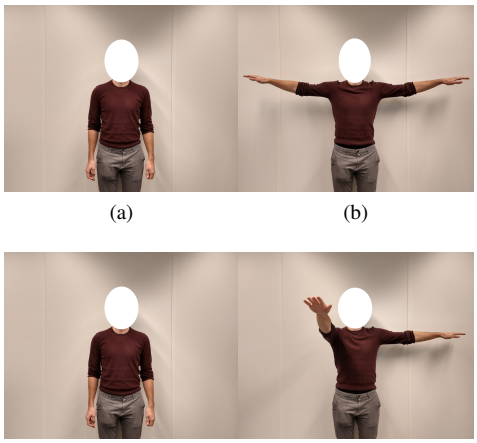

Fig. 6. Arm motions captured for the case study described in Section III-B. (a)(b): Raising and lowering both arms laterally. (c)(d): Raising and lowering one arm forward, one arm laterally simultaneously.

can be attributed to both their comparative difficulty, and their underrepresentation in the imbalanced dataset.

The results for the arm motion case study are displayed in Figure 7. The subfigures incorporate intensity maps $I(x, y)$ and masked velocity vector fields $\mathcal{V}_{\mathbf{T}}(x, y, v_x, v_y)$. Clustering has also been applied to aid in visual discrimination of the motions, and cluster centroids are indicated with exaggerated vectors to indicate gross limb movement. The classes of motion (i.e., 7a: Raising arms laterally; 7b: Lowering arms laterally; 7c: Raising left arm laterally & right arm frontally) can be visually discerned in the sub-figures, suggesting that a classification method can be developed based on the data representation generated by the proposed method.

V. Conclusion

In this work, a novel signal-level radar sensor fusion method is proposed that utilises a distributed network of radar sensors to reconstruct a 2D map of reflection intensity, as well as

TABLE I

TABLE SHOWING TRUE POSITIVE RATE (TPR) AND PRECISION FOR ALL PARTICIPANTS (L1PO TEST) IN THE FALL DETECTION CASE STUDY. AVERAGES OF THE METRICS AS WELL AS STANDARD DEVIATION ARE ALSO REPORTED IN THE LAST TWO COLUMNS.

Metric	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Average	StDev
TPR	98.7%	97.4%	100.0%	100.0%	97.6%	100.0%	100.0%	100.0%	100.0%	100.0%	96.8%	100.0%	100.0%	100.0%	99.3%	1.2%
Precision	89.3%	96.2%	97.2%	93.1%	95.2%	97.5%	95.7%	92.6%	95.5%	93.1%	85.7%	83.1%	96.7%	90.8%	93.0%	4.4%

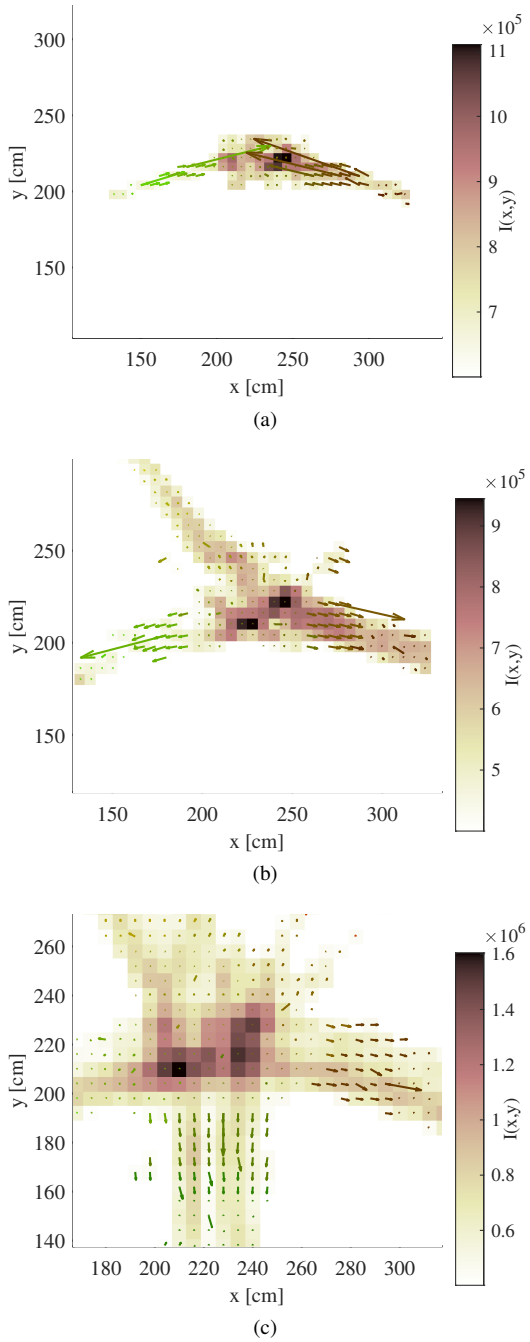

Fig. 7. Reflection Intensity maps & velocity vector fields for various arm motions. Larger vectors represent cluster centroids and are scaled for visibility. In all sub-figures, the target is facing in the negative y direction. (a) Raising arms laterally. (b) Lowering arms laterally. (c) Raising left arm laterally & right arm frontally.

a vector field of absolute velocities. Aided by a clustering algorithm to extract gross limb motion, these outputs are used to demonstrate two use cases of the method, namely: fall detection using only the intensity map, and distinguishing limb movements. In the case of fall detection, a TPR and precision of respectively $99.3\,\%$ and $93.0\,\%$ are achieved on a publicly available dataset of continuous human activities. Additionally, for the second use case sample snapshots of arm motions are presented that can be empirically classified, indicating that an automated classification method can be developed based on this novel data representation.

In future work, deep learning approaches will be investigated to perform activity classification on the reconstructed velocity field, expanding the motion classes of interest to include a wider variety of activities of daily living. Such a study will benchmark the performance of the proposed method with respect to reference works in literature. Furthermore, studies will be performed on the influence of the geometry of the sensor network in terms of sensor quantity and placement and their effects on reconstruction accuracy. Finally, the reconstruction performance of the method in a multi-target scenario shall be gauged experimentally.

REFERENCES

[1] S. A. Shah, A. Tahir, J. Le Kernec, A. Zoha, and F. Fioranelli, "Data portability for activities of daily living and fall detection in different environments using radar micro-doppler," *Neural Computing and Applications*, vol. 34, no. 10, pp. 7933–7953, may 2022.

[2] Y. Yao, C. Liu, H. Zhang, B. Yan, P. Jian, P. Wang, L. Du, X. Chen, B. Han, and Z. Fang, "Fall Detection System Using Millimeter Wave Radar Based on Neural Network and Information Fusion," *IEEE Internet of Things Journal*, pp. 1–1, 2022.

[3] E. Kurtoglu, A. C. Gurbuz, E. A. Malaia, D. J. Griffin, C. S. Crawford, and S. Z. Gürbüz, "Sequential Classification of ASL Signs in the Context of Daily Living Using RF Sensing," in *2021 IEEE Radar Conference (RadarConf21)*, vol. 2021-May. IEEE, may 2021, pp. 1–6.

[4] C. Wang, X. Zhao, and Z. Li, "DCS-CTN: Subtle Gesture Recognition based on TD-CNN-Transformer via Millimeter Wave Radar," *IEEE Internet of Things Journal*, pp. 1–1, 2023.

[5] E. Klinefelter and J. A. Nanzer, "Hand Gesture Recognition Using a Dual Axis Millimeter-Wave Interferometric-Doppler Radar and Convolutional Neural Networks," in *2021 18th European Radar Conference*, no. April, London, 2022, pp. 297–300.

[6] G. Li, S. Zhang, F. Fioranelli, and H. Griffiths, "Effect of sparsity-aware time-frequency analysis on dynamic hand gesture classification with radar micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 12, no. 8, pp. 815–820, 2018.

[7] L. Wang, Z. Cui, Y. Pi, C. Cao, and Z. Cao, "Adaptive framework towards radar-based diversity gesture recognition with range-Doppler signatures," *IET Radar, Sonar & Navigation*, vol. 16, no. 9, pp. 1538–1553, sep 2022.

[8] M. Arsalan, T. Zheng, A. Santra, and V. Issakov, "Contactless Low Power Air-Writing Based on FMCW Radar Networks Using Spiking Neural Networks," in *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, dec 2022, pp. 931–935.

[9] N. Kern, M. Steiner, R. Lorenzin, and C. Waldschmidt, "Robust Doppler-Based Gesture Recognition with Incoherent Automotive Radar Sensor Networks," *IEEE Sensors Letters*, vol. 4, no. 11, pp. 1–4, 2020.
[10] S. Z. Gürbüz, M. M. Rahman, E. Kurtoglu, E. A. Malaia, A. C. Gurbuz, D. J. Griffin, and C. S. Crawford, "Multi-Frequency RF Sensor Fusion for Word-Level Fluent ASL Recognition," *IEEE Sensors Journal*, pp. 1–1, 2021.
[11] R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Distributed radar fusion and recurrent networks for classification of continuous human activities," *IET Radar, Sonar & Navigation*, apr 2022.
[12] S. Z. Gurbuz, M. M. Rahman, E. Kurtoglu, T. Macks, and F. Fioranelli, "Cross-frequency training with adversarial learning for radar micro-Doppler signature classification (Rising Researcher)," in *Radar Sensor Technology XXIV*, K. I. Ranney and A. M. Raynal, Eds., vol. 11408, International Society for Optics and Photonics. SPIE, 2020, p. 114080A.
[13] H. Li, A. Mehul, J. Le Kernec, S. Z. Gürbüz, and F. Fioranelli, "Sequential Human Gait Classification with Distributed Radar Sensor Fusion," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7590–7603, 2021.
[14] S. Zhu, R. G. Guendel, A. Yarovoy, and F. Fioranelli, "Continuous Human Activity Recognition With Distributed Radar Sensor Networks and CNN–RNN Architectures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.
[15] J. Traa, "Least-squares intersection of lines," Lecture Notes, 2013.
[16] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.
[17] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," vol. 8, 01 2007, pp. 1027–1035.
[18] R. G. Guendel, M. Unterhorst, F. Fioranelli, and A. Yarovoy, "Dataset of continuous human activities performed in arbitrary directions collected with a distributed radar network of five nodes," Nov 2021. [Online]. Available: https://data.4tu.nl/articles/dataset/Dataset_of_continuous_human_activities_performed_in_arbitrary_directions_collected_with_a_distributed_radar_network_of_five_nodes/16691500