

IEICE **TRANSACTIONS**

on Communications

DOI:10.23919/transcom.2025EBP3011

This advance publication article will be replaced by the finalized version after proofreading.

A PUBLICATION OF THE COMMUNICATIONS SOCIETY



The Institute of Electronics, Information and Communication Engineers
Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3chome, Minato-ku, TOKYO, 105-0011 JAPAN

PAPER

Non-line-of-sight Pedestrian Recognition Approach using Multiple Reflections for Millimeter Wave Radar Sensing Applications

Junya OKAZAKI[†], *Nonmember*, Takeru MATSUOKA[†], Jianghaomiao HE[†], *Student Members*,
and Shouhei KIDERA^{†a)}, *Senior Member*

SUMMARY We introduce a target recognition scheme based on multiple signal reflections assuming a non-line-of-sight (NLOS) scenario for a millimeter wave (MMW) radar used in automotive sensor applications. For a pedestrian or a cyclist jumping out from the rear of an obstacle, such as a parked vehicle, a target recognition scheme in a NLOS region is highly required for collision prevention, especially in driving support and autonomous driving systems. Since the diffraction signal obtained from a NLOS region is too weak for identifying a specific target type, particularly at high frequencies and under full NLOS conditions, we introduce a target recognition approach based on multiple reflections. Two machine learning approaches, based on the support vector machine (SVM) and long-short-term memory (LSTM) algorithms are introduced with only single transmitter and receiver data. The experiments conducted using a 24-GHz band MMW radar in an actual road scenario demonstrated that the proposed approach can accurately recognize a real human body from an artificial object by exploiting the unique characteristics of reflected signals.

key words: Millimeter-wave (MMW) radar, Automotive radar, Non-line-of-sight (NLOS) detection, Pedestrian detection, Multiple reflection, Supervised machine learning, Long-short-term memory (LSTM)

1. Introduction

In the emerging automated driving and advanced driver-assistance systems, there is an urgent demand for establishing a reliable sensing technology. Millimeter wave (MMW) sensors, namely, radar sensors, have become essential in optically challenging situations such as adverse weather, darkness, or strong back light, and there are a number of studies focusing on the radar imaging scheme [1–6]. However, the MMW radar exhibits an inherent drawback in terms of spatial resolution, which is of the order of cm, even when the high MMW frequency (60 or 79 GHz) bands are employed [7, 8]. Specifically, in far-range situations, *i.e.*, > 10 m from the radar site, the detection or recognition of a human body from other artificial objects, such as guardrails and barriers, becomes particularly challenging, because of the lower cross-range resolution than that in near-range situations. Several studies on pedestrian detection using micro-Doppler signal-based recognition, have been reported [9–12], where each part of the human body (*e.g.*, arms and legs) exhibits a different and periodical motion, leading to a Doppler velocity variation during slow-time measurements. Furthermore,

the gait level recognition of a human body has been investigated using a MMW radar in short-range scenarios [13–16] or through-the-wall situations [17–20]. Several studies have shown that a micro-Doppler signature can provide a significant feature for human-vehicle classification [21–25]. However, these studies focused on line-of-sight (LOS) targets. Recently, there has been a strong demand for non-line-of-sight (NLOS) target recognition or detection, because a child or a cyclist may jump out suddenly from the rear of a parked vehicle, and the driver or the automated braking system may not respond in time to prevent the accident.

Recent studies on MMW radar-based NLOS sensing scenarios, such as [26–28], have been reported. In these studies, imaging or localization schemes were employed for a human body hidden in the NLOS region. Specifically, a radar imaging method exploiting multiple reflections in a complicated indoor environment was proposed in [29–31] and an imaging or localization scheme for a human body behind a wall was introduced in [32]. In addition, the study in [33] assumes a scenario wherein a vehicle suddenly emerges from an NLOS area shadowed by the buildings. In this context, multipath scattering waves are induced by strategically placed reflectors. The study experimentally investigates the relationship between the reflector angle and the signal-to-noise ratio (SNR). However, these studies did not deal with the target identification or recognition problem. Conversely, our previous study [34] demonstrated that there is a significant difference in the reflection responses of a complex-valued radar backscattered signal along slow-time measurements. Specifically, the scattered signal response obtained from a static human body includes a phase rotation due to a mm-scale displacement caused by respiration, heart-beat, and posture control. In [34] a direct target recognition scheme employing the abovementioned signal reflection responses was presented. In this scheme, several feature vectors, such time-series or short-time Fourier transform data, are inserted into the support vector machine (SVM). Notably, this study also demonstrated a high recognition rate, even in the NLOS scenario case, *i.e.*, the diffracted signal could be used for target recognition at 10 m away from the radar. However, this study focused only the diffracted signal, and the experimental validation was performed in an ideal environment, *i.e.*, an anechoic chamber. Therefore, the detection of a diffraction signal would be much smaller in a real road scenario because the radar data include a number of clutter components such as roads, trees, buildings, or

[†]The authors are with Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo, Japan.

a) E-mail: kidera@uec.ac.jp*

other surrounding objects.

In a real-case scenario, a child or a cyclist may suddenly appear from spaces between parked vehicles. In this case, multiple scattered signals between the pedestrian and the parked vehicles can be distinctly obtained, particularly in the NLOS case. Thus, to address the above-mentioned issues in [34], this study introduces a target recognition scheme based on multiple scattering components that enhances the recognition accuracy of pedestrians hidden in NLOS areas, where the use of multiple reflections considerably enhances the signal-to-noise ratio (SNR) compared with that obtained using the diffraction signals in [34]. Additionally, we introduce the long-term-short-memory (LSTM) algorithms [35], one of the recurrent neural network (RNN) scheme, to identify a target using time-series data of multiple reflection signals.

The main contributions of this paper are the following:

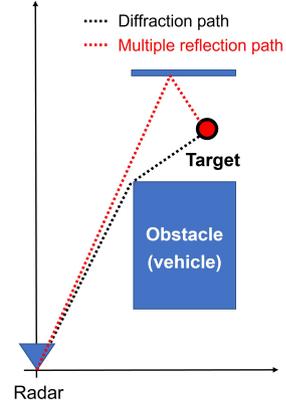
1. Multiple reflection signals are initially used to distinguish a pedestrian (target) from an artificial object in cases where the target is in the NLOS region between obstacles (*e.g.*, parked vehicles).
2. This approach can recognize a pedestrian in an NLOS area more accurately than the use of diffraction signals only. In particular, the SNR is considerably improved by observing multiple reflection signals.
3. Phase rotation due to a pedestrian's respiration or posture control can be clearly recognized from man-made objects using complex time-series data and the SVM or LSTM based learning scheme successfully identifies a pedestrian target even in deep NLOS area.

The experiments conducted using a 24-GHz band MMW frequency-modulated continuous-wave (FMCW) radar and assuming a real road environment. The findings demonstrate that our scheme, which uses multiple reflection signals, significantly improves the recognition accuracy between a real pedestrian and a dummy object in an NLOS area.

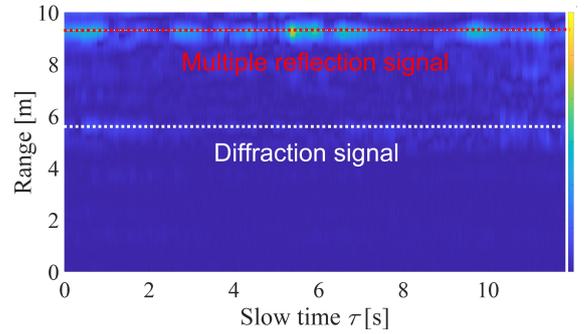
2. Method

2.1 Extractions on Diffraction Signals

Our previous study [34], introduced diffraction-signal-based NLOS target identification, whose methodology is briefly described as follows. In this paper, we employ a FMCW radar; hence, the repeatedly transmitted signal is recorded at the receiver, with a constant pulse repetition interval (PRI). The complex-valued received signal is $s(R, \tau)$, where τ denotes the measurement slow time discretized by the PRI. $R = ct/2$ where c is the speed of light and t is the fast time. We can obtain the target range R' by extracting the local maxima of $s(R, \tau)$ as to R . Notably, the study in [34] demonstrated that there is a distinct signature of a static human body in the complex-valued signal reflection responses. These signatures primarily result from human body displacements owing to breathing or posture control. Because they are characterized in temporal variations in the phase or



(a) Geometry



(b) Scattered data

Fig. 1: (a): Conceptual illustration of the proposed method for diffraction and multiple reflection paths in typical NLOS scenario. (b): Example of range- τ profile as $s(R, \tau)$, where the time-delay in fast time direction, namely, R of the multiple reflection signal can be observed at the certain range as R_M .

magnitude of the complex-valued scattered signals, we introduce multiple feature vectors, particularly for extracting a temporal or frequency variations, to facilitate target identification between a human body and an artificial object, such as a cylinder. At first, we introduce the following complex-valued feature vector \mathbf{x} as the time variance in the τ direction:

$$\mathbf{x}_S \equiv (s(\tilde{R}, \tau_0), s(\tilde{R}, \tau_1), \dots, s(\tilde{R}, \tau_N)) \quad (1)$$

In our previous study [34], five types of features are introduced to characterize the temporal variation of $s(R, \tau)$ along τ . In the following, i denotes the index number of the i -th slow-time sample as τ_i .

Feature 1: $s(R, \tau)$.

$$\mathbf{X}_1[i] \equiv (x[i]). \quad (2)$$

where $x[i]$ denotes the i -th components of the vector \mathbf{x} .

Feature 2: $s(R, \tau)$ and $\frac{\partial s(R, \tau)}{\partial \tau}$.

$$\mathbf{X}_2[i] \equiv \left(x[i], \frac{x[i+1] - x[i-1]}{\tau[i+1] - \tau[i-1]} \right)^T. \quad (3)$$

where T denotes the transpose operator.

Feature 3: A number of time shifted data of $s(R, \tau)$ along τ direction.

$$\mathbf{X}_3[i] \equiv (x[i], x[i+1], \dots, x[i+M])^T. \quad (4)$$

where M denotes the length in the time shift.

Feature 4: Short time Fourier transform (STFT) of $s(R, \tau)$ in the τ direction.

$$\mathbf{X}_4[i] \equiv \mathcal{F}^{DTFT} \left[(x[i], x[i+1], \dots, x[i+M])^T \right], \quad (5)$$

where $\mathcal{F}^{DTFT}[\ast]$ denotes the discrete Fourier transform operator, and M denotes the length of the DTFT.

Feature 5: Instantaneous frequency and spectrum entropy response of $s(R, \tau)$ in the τ direction.

$$\mathbf{X}_5[i] \equiv \begin{bmatrix} F[i], & F[i+1], & \dots, & F[i+M-N-1] \\ G[i], & G[i+1], & \dots, & G[i+M-N-1] \end{bmatrix} \quad (6)$$

where $F[i]$ and $G[i]$ denote the instantaneous frequency and the spectral entropy [36], respectively, that are mostly used to extract the feature a time-frequency data as:

$$F[i] \equiv \frac{\sum_{j=1}^N j P[i, j]}{\sum_{j=1}^N P[i, j]} \quad (7)$$

$$G[i] \equiv - \sum_{j=1}^N P[i, j] \log_2 P[i, j] \quad (8)$$

$$P[i, j] \equiv \frac{S[i, j]}{\sum_{j=1}^N S[i, j]} \quad (9)$$

holds. Here, $S[i, j]$ denotes the short time DFT of $x[i]$ with the length of M , and N is the data length, which determines the temporal resolution of $F[i]$ and $G[i]$.

The rationale behind selecting these features is as follows. Feature 1 has a minimal input dimension; however, it does not include temporal variation along τ in each input vector. The other features (Features 2, 3, 4, and 5) account for the temporal variation of responses to extract a unique pedestrian features caused by breathing, posture control, and similar movements, which differ from those of artificial objects. In particular, Feature 2 includes a temporal differential along τ , corresponding to the instantaneous frequency; however, it is essentially sensitive to random noise. Feature 3 includes a number of temporal data, thus retaining some temporal variation, in phase or amplitude. Feature 4 (DTFT) denotes data using STFT along τ , thus directly characterizing a unique feature in each target in the frequency domain. Feature 5 denote the instantaneous frequency and spectral entropy, which extracts the first moment of the power spectrogram, and power distribution in temporarily varying frequency spectrum, respectively. The recognition performance in Features 3, 4, and 5 depends on the selected data length M , which should be determined considering the balance between the number of input dimensions and the necessary frequency resolution.

2.2 Extractions on Multiple Scattered Signals

Our previous study [34] demonstrated that high recognition accuracy could be achieved, even under full NLOS conditions. However, in the above study, only the diffraction signal from the object in the NLOS case was considered. In assuming an actual road traffic scenario, where a diffraction signal is buried under other interference signals originating from other moving objects, *e.g.*, swaying of the leaves of the trees. In particular, in high-frequency radar systems, such as 79 GHz band, the diffraction effect is more alleviated, thus increasing the difficulty of NLOS target identification.

To address these issues, in this study, we initially focus on the multiple reflections between an object and other obstacles, such as parked vehicles, telephone poles, guardrails, walls, and other surrounding objects. Stronger multiple reflection components with a longer delay time are expected than the corresponding diffraction signal components and can be extracted from the received signal $s(R, \tau)$. Figure 1 shows a conceptual illustration of the proposed method. Here, because the radar can accurately measure the sensor–target distance by extracting the peaks of $|s(R, \tau)|$, we can also determine a specific range \tilde{R}_M corresponding to that derived from multiple reflection signals as

$$\tilde{R}_M = \arg \max_R \frac{1}{T} \left| \int_0^T s(R, \tau) d\tau \right| \quad (10)$$

where T denotes the total observation time along τ . Then, the input vector in the proposed method is defined as:

$$\mathbf{x}_M \equiv (s(\tilde{R}_M, \tau_0), s(\tilde{R}_M, \tau_1), \dots, s(\tilde{R}_M, \tau_N)) \quad (11)$$

In post-machine learning, the proposed method also uses the five feature vectors (Eqs. (2), (3), (4), (5), and (6)). Figure 3 shows the processing flow of the proposed scheme.

2.3 Machine Learning Based Object Identification

This section describes the machine learning schemes to identify the object in the NLOS area, using the diffraction or multiple reflection signals. As a promising learning scheme, we introduce the two approaches, as SVM and LSTM, the methodology of which are briefly described as follows:

2.3.1 SVM

As in [34], the above features are inserted into post-SVM learning for target identification. SVMs are promising classification schemes that achieve nonlinear classification with low complexity [37]. For fast nonlinear classification in a hyperspace, an SVM uses a well-known technique where the pairwise similarity between input data is directly measured using a kernel function. Here, we use the Gaussian kernel detailed in [34] as:

$$K(\mathbf{x}_1, \mathbf{x}_2) = \exp \left(- \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{2\sigma^2} \right) \quad (12)$$

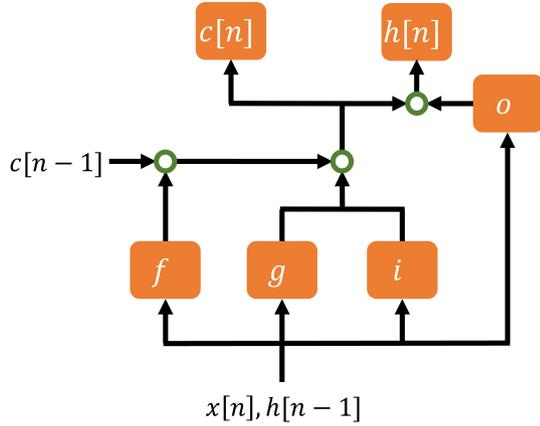


Fig. 2: Structure and processing flow of the LSTM learning scheme.

where σ expresses the standard deviation of the Gaussian function. This scheme has been already introduced in the literature [34], using the diffraction signal based feature vector as \mathbf{x}_S , while the case using the multiple reflection signal \mathbf{x}_M has not been validated.

2.3.2 LSTM

The LSTM is one of the most efficient deep learning scheme for time-serious data, in the recurrent neural network (RNN) framework. Notably, while traditional RNNs inherently struggle to retain long-term memory owing to the problem of exploding and vanishing gradients, LSTM networks are considered as a potential solution. Through the introduction of additional gates to regulate the information in a hidden state that is to be exported as output to the next hidden state, LSTM networks improve the learning of long-term dependencies. A LSTM structure is made of a number of cells, as input, output, forget gates, as shown in Fig. 2, which controls the flow of information from input to output cells. In particular, forget gate determines to discard or retain the cell information of the previous state, which enables a long-term control of input information.

The processing flow of the LSTM is summarized as:

Step 1): Each gate and unit is fed with the input $x[n]$ at time-index n and the output of the memory unit $h[n-1]$.

Step 2): Cell state vector $c[n-1]$ is multiplied by the output of the forget gate $f[n]$ (to determine how much of the past input is stored).

Step 3): The input unit $i[n]$ is multiplied by the output of the input gate $g[n]$ (to determine how much of the input value to put in).

Step 4): The outputs of Step 2 ($f[n] \circ c[n-1]$) and Step 3 ($i[n] \circ g[n]$) are added together to update the state of the memory cell $c[n]$.

Step 5): The state of the memory cell is passed to the

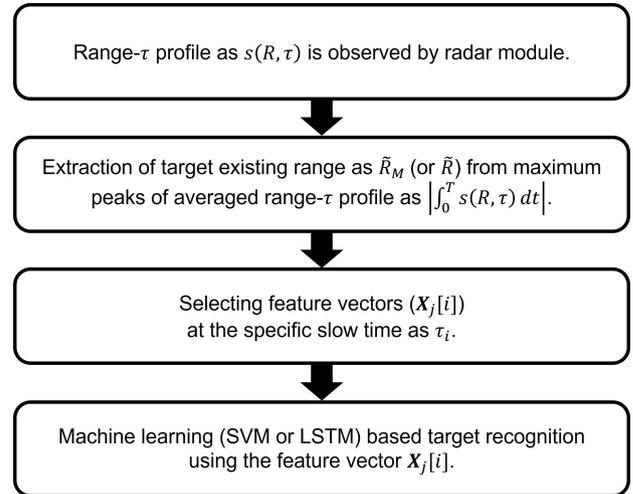


Fig. 3: Processing flow of the proposed target identification scheme, by exploiting the multiple reflection signals.

output unit $h[n]$.

Step 6): The output of the output gate $o[n]$ is multiplied by the state of the memory cell $c[n]$ to obtain the output from the memory unit $h[n]$.

3. Result

3.1 Experimental Setup

This section describes the target recognition performance using the proposed scheme and assuming a real university road environment. A multiple-input multiple-output (MIMO) FCM radar produced by Sakura Tech Corp. operating at a 24-GHz center frequency and 0.2 GHz bandwidth is used in the experiment. Two transmitters and four receivers forms the 1D array along the horizontal axis, where the total module size is 104 mm in width, 76 mm in height, and 6 mm in thickness; the maximum output power is 8 dBm. Figure 4 shows the experimental setup, which includes the radar and a vehicle, which is considered as an obstacle. The horizontal and vertical antenna beamwidths are ± 45 and ± 6.5 degrees, respectively, denoting a quasi two-dimensional sensing scenario. Here, target signals acquired at the same radar height can be observed in a wide range of azimuth directions, and clutter signals from the ground or other surroundings at different heights are suppressed by the narrower vertical beamwidth. The radar height was 0.5 m. The PRI was 2.1 ms, and the total pulse hits were 2125.

We considered two different target types. One was a dummy doll (child) with a 1.23 m height, which was produced under the European New Car Assessment Program (Euro NCAP) standards. The other was a male human target with a 1.70 m height wearing light clothes and standing still during the measurement process. These two targets were placed at three different positions (A, B, and C). A was located in front of the radar (namely, the LOS region

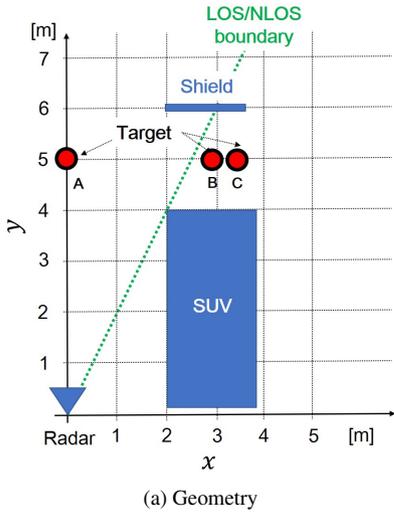


Fig. 4: Experimental setup.

within the center of the radar beam), where the radar–target distance was 5.0 m. B and C were located in the NLOS region 2.96 m and 3.45 m far from A along x axis, and an 1.79 m height SUV vehicle obstacle was located at the right side of the radar. To generate multiple scattering effects, an aluminum shield was placed at 2.0 m behind the obstacle vehicle (Fig. 5), referred to as rear shield. During the recognition pre-processing, all reflection data were subtracted from the background data, which were obtained without considering any targets and obstacle vehicles. Table 1 summarizes the conditions for each case.

3.2 Recognition Performance

3.2.1 Scattered Data in LOS Scenario

Initially, we investigated the responses of each target under LOS conditions, namely, Cases 1, 2, and 3. Figure 6 shows the reflection responses $s(R, \tau)$ and the scattered plots for a specific range \hat{R} of targets in the LOS case, where each target is located at positions A, B, and C. To suppress the clutter generated by the static objects, such as building walls, roadside strips, and other obstacles, including parked vehicles in the NLOS region, the zero-Doppler components were eliminated in each case by subtracting the average value signals along the τ direction. As shown in this figure, the responses from the front side of the vehicle obstacle of a reflection from the ground or a building could be suppressed

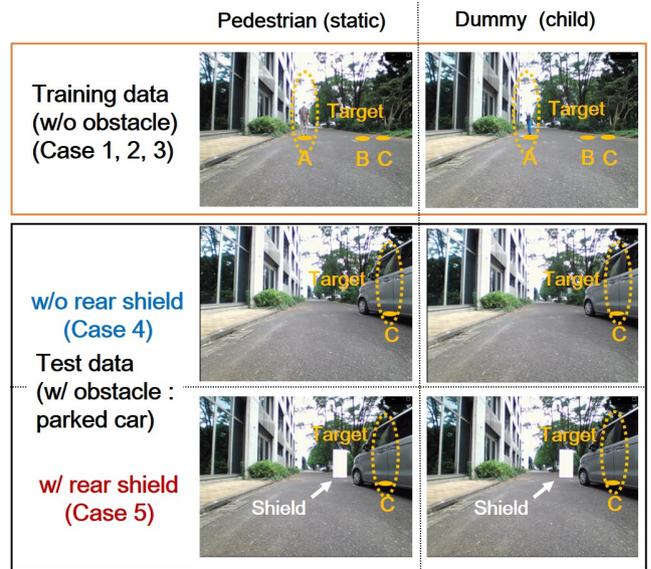


Fig. 5: Experimental model, and conditions for training and test data acquisition.

Table 1: Definition for each case. N/A: Not applied S: Single scattering. M: Multiple scattering. Case 4 corresponds to the situation in [34]. Case 5 corresponds to the situation in the proposed method.

Case	Abbreviation	Target location	Vehicle obstacle	Rear shield
1	A-LOS-S	A	N/A	N/A
2	B-LOS-S	B	N/A	N/A
3	C-LOS-S	C	N/A	N/A
4	C-NLOS-S	C	Applied	N/A
5	C-NLOS-M	C	Applied	Applied

considerably by eliminating the zero-Doppler components, which allows us to extract a nonstationary response caused by a pedestrian’s respiration or posture control. In Fig. 6, we observe some strong responses at the range \hat{R} of each target. We can also identify time-variant components in the pedestrian response along the measurement slow time, even when the pedestrian is static. These components are regarded as mm-order displacements of the human body due to breathing or position control, which have been identified in [34]. In particular, the response in Case 1, *i.e.*, at the radar center beamwidth, is more distinctive than that in the other cases because the main radar beam is in the direction at the front of the radar. Nonetheless, in each case, there are significant discrepancies between the dummy and the static pedestrian responses. Cases 1, 2, and 3 were used for the training sequences of the SVM or LSTM learning approach to identify the targets in the NLOS scenario, as described in the next section.

3.2.2 Scattered Data in NLOS Scenario

Next, we investigated the target identification performance

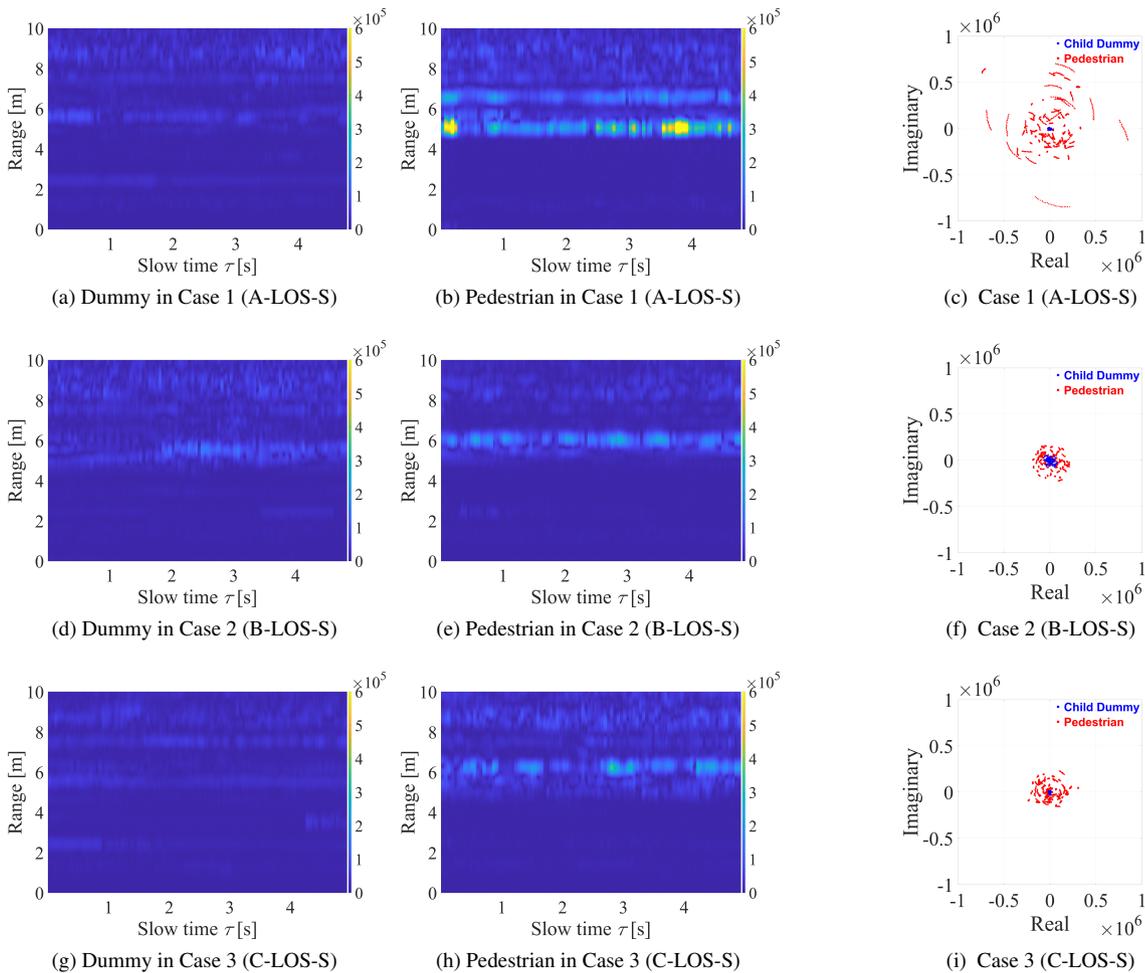


Fig. 6: (a), (b), (d), (e), (g), and (h) express range- τ profiles as $|s(R, \tau)|$ in Case 1, 2, and 3. (c), (f), and (i) denote the scattered plots of each response at the specific extracted range as $\tilde{R} = 5.1$ m, $\tilde{R} = 5.9$ m, and $\tilde{R} = 6.2$ m, respectively. Blue and red dots denote responses from the child-dummy and the actual pedestrian, respectively.

in the NLOS region with and without multiple scattering effects. The SUV vehicle was located between the radar and the target position, as shown in Fig. 4. Then, positions B and C were included in the full NLOS region. Figure 7 shows the reflection responses $s(R, \tau)$ at position C in this NLOS scenario, where a metallic rectangular shield, which generates multiple scattered signals between the shield and target, either does not exist, or it is located at 2 m behind the vehicle, as shown in Fig. 4. The results in Case 4 (C-NLOS-S), namely, the NLOS case without multiple scattering components, corresponding to the situation in [34], show that the diffraction signal strength is considerably lower than those obtained in the LOS scenario, as in Case 3 (C-LOS-S). Also, we can hardly identify any distinctive features in the scattered signal plots of the responses between the dummy and the pedestrian. In contrast, in Case 5 (C-NLOS-M), *i.e.*, in the situation assumed in the proposed scheme, the scattered data exhibit distinctive responses obtained from the pedestrian at the far range ($\tilde{R}_M = 9.35$ m) and are considered as multiple scattered signals between the shield and the targets.

Table 2: Parameters used in the LSTM.

Parameter	Value
Number of epochs	3
Mini-batch size	25
Initial learning rate	0.01
Sequence length	100

The scattered signal plots extracted for this range exhibit a significant discrepancy between the dummy and the pedestrian, which is similar to Case 3.

3.2.3 Classification Results

Next, we investigated the two-class classification performance (dummy v.s. pedestrian) using the SVM and the LSTM for Cases 4 and 5. To evaluate the applicability of this scheme, we introduced three different training datasets in the LOS case, which were extracted from Cases 1, 2, and 3, as shown in Fig. 5. To train the SVM and LSTM, we introduce five-fold cross validation scheme, where the pro-

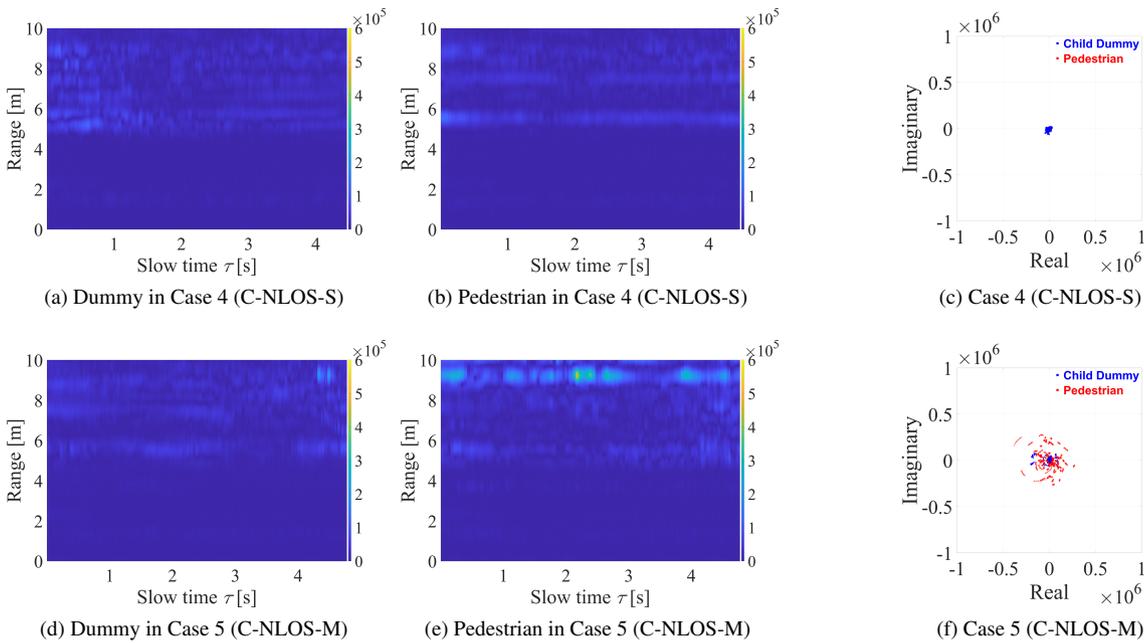


Fig. 7: Case 4 corresponds to the situation in [34]. Case 5 corresponds to the situation in the proposed method. (a), (b), (d), and (e) express range- τ profiles as $|s(R, \tau)|$ in Case 4 and 5. (c) and (f) denote the scattered plots of each response at the specific extracted range as $\tilde{R} = 6.60$ m and $\tilde{R} = 9.15$ m, respectively. Blue and red dots denote responses from the child-dummy and the actual pedestrian, respectively.

Table 3: Recognition accuracy by the SVM scheme in Case 4 (C-NLOS-S) *i.e.*, assuming the scenario in, [34], in using each training data the situation A-LOS-S (A), B-LOS-S (B), and C-LOS-S (C). In Feature 5, the notation in Data length parameter denotes $M(N)$ in Eq. (6).

Feature	Data length	A	B	C
1	1	47.0 %	49.9 %	37.7 %
2	2	49.5 %	49.9 %	40.0 %
3	10	47.1 %	49.7 %	37.1 %
	20	47.0 %	49.7 %	37.1 %
	50	47.4 %	50.0 %	37.5 %
	100	49.4%	50.0 %	41.7 %
4	10	49.3 %	50.0 %	44.2 %
	20	49.8 %	50.0 %	45.7 %
	50	50.0 %	50.0 %	45.8 %
	100	50.0 %	50.0 %	48.9 %
5	10(6)	37.7 %	37.9 %	37.8 %
	20(16)	36.4 %	37.4 %	36.9 %
	50(46)	35.1 %	38.7 %	35.1 %
	100(96)	35.1 %	40.9 %	35.0 %

portion of the training and validation data is defined as 80 % (1280 samples) and 20 % (320 samples). We extract the average accuracy for these five training patterns, where the unknown test data (200 samples) are input. In the LSTM, the bi-directional LSTM [38] with two hidden layers are used, where the adaptive moment estimation (Adam) solver is used. Table 2 also summarizes the parameters used in the LSTM. At first, Tables 3 and 4 show the classification accuracy in Case 4 (C-NLOS-S), where only the diffraction signals are available in scattered data as in Fig. 6-(a), (b), and

Table 4: Recognition accuracy by the LSTM scheme in Case 4 (C-NLOS-S) *i.e.*, assuming the scenario in, [34], in using each training data the situation A-LOS-S (A), B-LOS-S (B), and C-LOS-S (C). In Feature 5, the notation in Data length parameter denotes $M(N)$ in Eq. (6).

Feature	Data length	A	B	C
1	1	54.7 %	57.0 %	49.5 %
2	2	49.8 %	50.0 %	50.0 %
3	10	40.3 %	36.7 %	45.6 %
	20	56.9 %	48.8 %	51.4 %
	50	60.2 %	45.8 %	47.4 %
	100	31.4 %	45.8 %	35.7 %
4	10	52.0 %	46.3 %	51.5 %
	20	43.7 %	50.3 %	59.8 %
	50	57.9 %	50.2 %	60.4 %
	100	48.6 %	45.8 %	46.7 %
5	10 (6)	49.8 %	37.8 %	44.4 %
	20 (16)	40.7 %	44.3 %	41.3 %
	50 (46)	38.1 %	41.8 %	40.3 %
	100 (96)	34.2 %	44.1 %	44.6 %

(c). Here the accuracy is defined as the average value of the true positive and true negative rates. These tables indicate that a high accuracy (over 50%) could not be obtained in Case 4 in the both learning schemes SVM and LSTM, even when using various types of training datasets or features, because the strength of each diffraction signal was not sufficient to identify the features of targets due to the obtained low SNR values, and the reflection responses from both the dummy and the pedestrian are dominated by random noise (Fig. 7 (c)); thus, all test data are identified as either target

Table 5: Recognition accuracy by the SVM scheme in Case 5 (C-NLOS-M), *i.e.*, assuming the proposed scenario, in using each training data the situation A-LOS-S (A), B-LOS-S (B), and C-LOS-S (C). In Feature 5, the notation in Data length parameter denotes $M(N)$ in Eq. (6).

Feature	Data length	A	B	C
1	1	90.8 %	90.9 %	90.8 %
2	2	85.5 %	89.6 %	93.9 %
3	10	90.1 %	90.9 %	90.3 %
	20	89.6 %	90.3 %	90.6 %
	50	87.5 %	91.9 %	94.1 %
	100	87.1 %	91.2 %	96.4 %
4	10	80.7 %	84.8 %	89.8 %
	20	81.2 %	83.8 %	87.6 %
	50	88.8 %	85.5 %	91.2 %
	100	92.0 %	90.3 %	93.3 %
5	10 (6)	76.7 %	77.6 %	76.7 %
	20 (16)	79.1 %	79.4 %	78.7 %
	50 (46)	82.0 %	77.4 %	82.4 %
	100 (96)	87.3 %	74.3 %	88.7 %

Table 6: Recognition accuracy by the LSTM scheme in Case 5 (C-NLOS-M), *i.e.*, assuming the proposed scenario, in using each training data the situation A-LOS-S (A), B-LOS-S (B), and C-LOS-S (C). In Feature 5, the notation in Data length parameter denotes $M(N)$ in Eq. (6).

Feature	Data length	A	B	C
1	1	59.2 %	50.7 %	50.3 %
2	2	50.0 %	50.0 %	49.9 %
3	10	47.7 %	54.1 %	49.9 %
	20	50.6 %	45.9 %	55.1 %
	50	45.1 %	53.0 %	45.7 %
	100	42.9 %	54.4 %	53.9 %
4	10	52.2 %	55.1 %	55.8 %
	20	42.3 %	59.6 %	43.5 %
	50	54.9 %	59.8 %	43.8 %
	100	49.6 %	40.2 %	60.6 %
5	10 (6)	77.7 %	74.3 %	76.3 %
	20 (16)	81.1 %	76.0 %	77.4 %
	50 (46)	85.7 %	79.4 %	82.5 %
	100 (96)	81.4 %	72.1 %	83.5 %

(the dummy in this case). Conversely, Table 5 and 6 summarize the accuracy in assuming Case 5 (C-NLOS-M), where the multiple reflection signals are included in the reflection data. These results demonstrate that the multiple scattering components greatly enhance the identification accuracy due to the obtained high SNR values, compared with those obtained in Case 4, especially in using the SVM with any feature or data length. Notably, in this case, while the LSTM could not retain a sufficient accuracy in using Feature 1, 2, 3, and 4, the Feature 5 remarkably enhance the accuracy, particularly in using longer data length (*e.g.*, 100). Thus, the accuracy of the LSTM highly depends on the selected feature vector, unlike the SVM results, Furthermore, in any case, the accuracy do not much depend on the location of the training data as A, B, or C, which demonstrates the applicability of the proposed scheme in terms of robustness to training dataset.

Table 7: SNRs in each case.

Case	Dummy	Pedestrian
1	11.4 dB	40.0 dB
2	22.8 dB	30.3 dB
3	11.1 dB	31.8 dB
4 ([34])	18.3 dB	9.6 dB
5 (Proposed)	27.6 dB	33.7 dB

The above discussions are also validated in the view point of the available SNR. Table 7 shows the SNR values in each case, where the $\text{SNR} \equiv P_S/P_N$ is defined where signal power P_S is determined as.

$$P_S = \max_{R, \tau} |s(R, \tau)|^2 \quad (13)$$

Conversely, the noise power P_N is defined as the average of lower 10 % data of $|s(R, \tau)|^2$, which approximately estimates a variance of additive white noise. This table shows that in Case 5 (the proposed scheme), 10 to 20 dB higher SNR values were obtained by extracting the multiple scattering echo signals, compared to those in Case 4, which contributes more accurate recognition for pedestrian existing. This experimental test assumes an ideal metallic shield with a smooth surface to generate multiple scattered signals. However, in real-world cases, such shielding obstacles do not necessarily have smooth surfaces, such as cylindrical poles and walls with corners. In particular, if the rear obstacle is another parked vehicle, then its front or rear surface will be more complex, and its surface reflection strength should be reduced considerably relative to that of the flat shield. In these cases, the expected SNR will be lower than that assumed in this experiment, and some noise reduction scheme should be implemented, such as coherent radar imaging processing along multiple elements. Nevertheless, the SNR will be enhanced relative to the case involving only diffraction signals. Furthermore, although this study adopts SVM or LSTM based classification, many machine learning schemes are available, such as random forests, k-nearest neighbor models, and logistic regression. Among these algorithms, SVMs are prevalent, easily implemented nonlinear classifiers needing less parameter selection and less sensitive to the feature vector, compared with the LSTM based RNN approaches. According to the above results, with such a widely used classifier, the use of multiple scattered signals can significantly help enhance recognition accuracy compared with that obtained using the strategy in [34].

3.3 Further Discussions

This section discusses the advantages and limitations of the proposed method based on a realistic scenario. First, the proposed method does not require target position estimation for classification; instead, it only extracts the temporal variations of scattered signals at a specific range, determined by the maximum response along fast time. Depending on the arrangement of the target and reflector, dominant propagation paths, such as direct, diffracted, or multipath scatter-

ing waves, may vary, resulting in a possible change in their relative intensities. However, as the method selects the signal with the maximum amplitude response, it need not distinguish whether the received signal originates from direct, diffracted, or multipath scattering waves. Moreover, it does not require any knowledge of the target or reflector location. In addition, focusing on Case C in the NLOS area, the propagation length (delay) is equivalent to a scenario wherein a target is located in the LOS area at $y = 7$ m with the same x position as Case C. Thus, the recognition accuracy could potentially be improved via the inclusion of training dataset corresponding to this target position. However, even when using only the training dataset from Cases A, B, or C in the LOS area at $y = 5$ m, the proposed scheme maintains sufficient recognition accuracy by effectively extracting multiple scattering signals, particularly through the use of SVM, as shown in Table 5. This is another advantage of the proposed method.

Furthermore, assuming that the distance between reflectors (e.g., parked vehicles) is approximately 1 to 2 m, similar to parallel parking scenarios, and that pedestrians are likely positioned near the midpoint, we set the target and reflector positions accordingly, with the reflector located 1 m behind the target, as illustrated in Fig. 4-(a). Notably, even when the target-reflector distances along the y -axis were set to 0.5 m or 1.5 m, multipath scattered signals would be observable according to shifting the reflection points of the reflectors. Although these distance variations introduce slight differences in propagation length, they do not significantly impact classification accuracy. This is because classification accuracy, whether using training datasets A or B (which involve different propagation lengths), is preserved by utilizing multiple scattering signals, as presented Table 5. Further changes in target position along the y -axis could alter the dominant propagation path among direct, diffracted, and multipath scattered waves. However, these variations are expected to be analogous to those observed when shifting the position along the x -axis.

4. Conclusion

In this study, we introduced a multiple scattering signature-based pedestrian recognition scheme under NLOS conditions using an MMW collision avoidance radar. The proposed scheme extracts a multiple signal scattering signature from the slow-time range profile and recognizes a static pedestrian from an artificial dummy by exploiting a unique feature of the human body displacements due to breathing or posture control without using imaging processing. The experimental tests using a 24-GHz MWW radar and assuming a real road environment demonstrated that the multiple reflections have a great potential to enhance the SNR and target identification accuracy by employing an SVM or LSTM based learning process, where the SVM has less sensitivity to selecting the feature vector, compared with the LSTM.

Notably, this study assumes a static radar scenario with a moving vehicle-mounted radar. Driving-induced vibration

will affect the phase fluctuation in the target response. In this case, the phase variations derived from only the breathing or posture control of the pedestrian should be extracted using the characteristic signature (periodical phase rotation due to respiration or walking motion). In addition, we can suppress the clutter from other obstacles, such as parked vehicles, walls, and other traffic surroundings, by considering the relative velocity to the moving radar site, which can be estimated using the time-varying responses in $s(R, \tau)$, using the Hough transform or other tracking approaches. Additionally, while this study assumes extraction from the range- τ profile; it can be expanded to obtain a complex-valued radar image via MIMO array-based beam-forming process. By extracting a phase variation in these radar images, it is promising to recognize a pedestrian in NLOS areas, while their images would be located at mirror position to the wall or parked vehicle. Although we need further investigations to clarify the above point in the future, the proposed multipath scheme is a promising solution that expands the application range of NLOS target recognition using vehicle-mounted MMW radars.

Finally, 79 GHz and other high-frequency MMW radars have appeared in recent years, outperforming 24 GHz ones in terms of spatial resolution or wider bandwidth. Since diffraction effects are further abated in higher-frequency systems, the use of multiple signals will be indispensable in NLOS sensing, and our ongoing work aims to extend this scheme to 79 GHz MMW radar systems.

References

- [1] D. A. Andrews, S. W. Harmer, N. J. Bowring, N. D. Rezgui, and M. J. Southgate, "Active Millimeter Wave Sensor for Standoff Concealed Threat Detection", *IEEE Sensors Journal*, Vol. 13, No. 12, pp. 4948–4955, Dec., 2013.
- [2] H. Sun, C. Gao, Z. Zhang, X. Liao, X. Wang and J. Yang, "High-Resolution Anisotropic Prestack Kirchhoff Dynamic Focused Beam Migration," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 11753–11760, 15 Oct., 2020
- [3] D. Xu, W. Yu, B. Wang and S. Liu, "Millimeter Wave Real-Time Tracking and Imaging of Moving Objects Based on Virtual MIMO Array and State Vector Prediction," *IEEE Access*, vol. 11, pp. 115854–115864, 2023.
- [4] J. W. Smith and M. Torlak, "Efficient 3-D Near-Field MIMO-SAR Imaging for Irregular Scanning Geometries," *IEEE Access*, vol. 10, pp. 10283–10294, 2022.
- [5] T. Kishigami, T. Morita, H. Mukai, M. Otani, Y. Nakagawa, "Advanced Millimeter-Wave Radar System to Detect Pedestrians and Vehicles by Using Coded Pulse Compression and Adaptive Array", *IEICE Trans. Commun.*, Vol.E96-B, No.9, pp.2313–2322, Sep., 2013.
- [6] H. Kamoda, T. Derham, T. Iwasaki, and T. Kuki, "Millimeter-Wave Imaging System Using Simultaneous Frequency-Encoding Technique", *IEICE Trans. Commun.*, Vol.E94-C No.2 pp.206–214, Feb., 2011.
- [7] P. J. B. Morris and K. V. S. Hari, "Detection and Localization of Unmanned Aircraft Systems Using Millimeter-Wave Automotive Radar Sensors," *IEEE Sensors Letters*, vol. 5, no. 6, pp. 1–4, June, 2021.
- [8] J. Song, C. Cui, S. Kim, B. Kim and S. Nam, "A Low-Phase-Noise 77-GHz FMCW Radar Transmitter With a 12.8-GHz PLL and a $\times 6$ Frequency Multiplier," *IEEE Microwave and Wireless Components*

- Letters*, vol. 26, no. 7, pp. 540-542, July, 2016.
- [9] Y. Kim, S. Ha, J. Kwon, "Human Detection Using Doppler Radar Based on Physical Characteristics of Targets", *IEEE Geosci. Remote Sens. Lett.* vol. 12, no. 2, pp. 289 - 293, Feb., 2015.
- [10] X. Shi, X. Yao, X. Bai, *et al.*, "Radar Echoes Simulation of Human Movements Based on MOCAP Data and EM Calculation", *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 859 - 863, Jun., 2019.
- [11] Y. Wang and Y. Zheng, "An FMCW Radar Transceiver Chip for Object Positioning and Human Limb Motion Detection", *IEEE Sensors Journal*, vol. 17, no. 2, pp. 236-237, Jan., 2017.
- [12] T. Wagner, R. Feger and A. Stelzer, "Radar Signal Processing for Jointly Estimating Tracks and Micro-Doppler Signatures," *IEEE Access*, vol. 5, pp. 1220-1238, 2017.
- [13] M. Li, T. Chen and H. Du, "Human Behavior Recognition Using Range-Velocity-Time Points," *IEEE Access*, vol. 8, pp. 37914-37925, 2020
- [14] K. Zhang, S. C. Lan, G. Y. Zhang, "Mining Spatio-Temporal Features from mmW Radar echoes for Hand Gesture Recognition," *2019 IEEE Asia-Pacific Microwave Conference (APMC)*, Dec., 2019.
- [15] M. S. Seyfioglu and S. Z. Gürbüz, "Deep neural network initialization methods for micro-Doppler classification with low training sample support", *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2462-2466, Dec., 2017.
- [16] T. Ohmori and S. Kidera, "Doppler Velocity Enhanced Range Migration Algorithm for High Resolution and Noise-Robust Three-Dimensional Radar Imaging," *IEEE Sensors Journal*, vol. 21, no. 18, pp. 20616-20628, 15 Sept., 2021.
- [17] Y. Song, J. Hu, T. Jin, *et al.*, "Estimation and Mitigation of Time-Variant RFI Based on Iterative Dual Sparse Recovery in Ultra-Wide Band Through-Wall Radar," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3398-3411, Sept., 2019.
- [18] H. Li, G. Cui, L. Kong, *et al.*, "Robust Human Targets Tracking for MIMO Through-Wall Radar via Multi-Algorithm Fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1154-1164, April, 2019.
- [19] M. Wang, G. Cui, L. Kong and X. Yang, "First-Order Rear-Wall Multipath Positioning and Suppression for Through-Wall Imaging Radar," *IEEE Sensors Journal*, vol. 18, no. 20, pp. 8261-8274, Oct., 2018.
- [20] M. Becquaert, E. Cristofani, B. Lauwens, M. Vandewal, J. H. Stiens and N. Deligiannis, "Online Sequential Compressed Sensing With Multiple Information for Through-the-Wall Radar Imaging," *IEEE Sensors Journal*, vol. 19, no. 11, pp. 4138-4148, June, 2019.
- [21] X. Li, Y. He, F. Fioranelli, X. Jing, A. Yarovoy and Y. Yang, "Human Motion Recognition With Limited Radar Micro-Doppler Signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6586-6599, Aug. 2021.
- [22] P. Khomchuk, I. Stainvas and I. Bilik, "Pedestrian motion direction estimation using simulated automotive MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 3, pp. 1132-1145, June, 2016.
- [23] S. Zhu, R. G. Guendel, A. Yarovoy and F. Fioranelli, "Continuous Human Activity Recognition With Distributed Radar Sensor Networks and CNN-RNN Architectures," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1-15, 2022.
- [24] Y. Kim, I. Alnujaim and D. Oh, "Human Activity Classification Based on Point Clouds Measured by Millimeter Wave MIMO Radar With Deep Recurrent Neural Networks," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13522-13529, June, 2021.
- [25] Z. Zhao *et al.*, "Point Cloud Features-Based Kernel SVM for Human-Vehicle Classification in Millimeter Wave Radar," *IEEE Access*, vol. 8, pp. 26012-26021, 2020.
- [26] M. Gustafsson, Å. Andersson, T. Johansson, *et al.*, "Extraction of Human Micro-Doppler Signature in an Urban Environment Using a "Sensing-Behind-the-Corner" Radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 187-191, Feb., 2016.
- [27] Q. Wang, Z. Li, H. Zhang, Y. Yang and X. Meng, "An Indoor UWB NLOS Correction Positioning Method Based on Anchor LOS/NLOS Map," *IEEE Sensors Journal*, vol. 23, no. 24, pp. 30739-30750, Dec., 2023.
- [28] Y. Xiang, S. Guo, S. Xia, X. Gu, J. Chen and G. Cui, "NLOS Target Positioning Method Based on UAV Millimeter-Wave Radar," *IEEE Sensors Journal*, vol. 24, no. 2, pp. 1975-1987, Jan., 2024.
- [29] S. Guo, Q. Zhao, G. Cui, *et al.*, "Behind Corner Targets Location Using Small Aperture Millimeter Wave Radar in NLOS Urban Environment," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 460-470, 2020,
- [30] Q. Tang, J. Li, L. Wang, Y. Jia and G. Cui, "Multipath Imaging for NLOS Targets Behind an L-Shaped Corner With Single-Channel UWB Radar," *IEEE Sensors Journal*, vol. 22, no. 2, pp. 1531-1540, Jan., 2022.
- [31] X. Yang, S. Fan, S. Guo, S. Li, G. Cui and W. Zhang, "NLOS Target Localization Behind an L-Shaped Corner With an L-Band UWB Radar," *IEEE Access*, vol. 8, pp. 31270-31286, 2020
- [32] S. Li, G. Cui, S. Guo, *et al.*, "On the Electromagnetic Diffraction Propagation Model and Applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 884-895, 2020.
- [33] D. Solomitskii, C. B. Barneto, M. Turunen, M. Allen, Y. Koucheryavy and M. Valkama, "Millimeter-Wave Automotive Radar Scheme with Passive Reflector for Blind Corner Conditions," *2020 14th European Conference on Antennas and Propagation (EuCAP)*, Copenhagen, Denmark, 2020, pp. 1-5.
- [34] J. He, S. Terashima, H. Yamada and S. Kidera, "Diffraction Signal-Based Human Recognition in Non-Line-of-Sight (NLOS) Situation for Millimeter Wave Radar," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4370-4380, 2021.
- [35] S. Hochreiter; J. Schmidhuber, "Long short-term memory", *Neural Computation*. 9 (8): 1735–1780, 1997.
- [36] Pan, Y. N., J. Chen, and X. L. Li, "Spectral Entropy: A Complementary Index for Rolling Element Bearing Performance Degradation Assessment." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*. Vol. 223, Issue 5, 2009, pp. 1223–1231.
- [37] C. Cortes and V. Vapnik, "Support-vector networks", *Machine learning*, vol. 20, no. 3, pp. 273–297, Feb. 1995.
- [38] Kim, Jinah and Namme Moon. "BiLSTM model based on multivariate time series data in multiple field for forecasting trading area." *Journal of Ambient Intelligence and Humanized Computing*, (2019): n. pag.

Junya Okazaki received the B.E. degree in electrical and electronic engineering from Tokyo University of Science, Japan, in 2021, and received the M.E degree with the Graduate School of Informatics and Engineering, the University of Electro-Communications, Tokyo, Japan in 2023. His research interests are signal processing for millimeter radar signal processing as well as its applications.

Takeru Matsuoka received the B.E. degree in communication engineering and informatics from the University of Electro-Communications, Tokyo, Japan, in 2024. He is currently pursuing the M.E. degree with the Graduate School of Informatics and

Engineering in the University of Electro-Communications, Tokyo, Japan. His research interest includes signal processing for millimeter radar signal processing as well as its applications.



Jianghaomiao He received the B.E. degree in electrical and electronic engineering from Baoji University of Arts and Sciences, China, in 2017, and received the M.E degree with the Graduate School of Informatics and Engineering, the University of Electro-Communications, Tokyo, Japan in 2021. and is currently pursuing the Ph.D.degree with the Graduate School of Informatics and Engineering, the University of Electro-Communications, Tokyo, Japan. His research interests are signal processing for millimeter and microwave radar as well as its applications.

limeter and microwave radar as well as its applications.



Shouhei Kidera received his B.E. degree in Electrical and Electronic Engineering from Kyoto University in 2003 and M.I. and Ph.D. degrees in Informatics from Kyoto University, Kyoto, Japan, in 2005 and 2007, respectively. In 2009, he joined as an Assistant Professor with the University of Electro-Communications, Tokyo, Japan, where he is currently a full Professor with Graduate School of Informatics and Engineering in the University of Electro-Communications, Tokyo, Japan.

His current research interest is in advanced radar signal processing or electromagnetic inverse scattering issue for ultra wideband (UWB) three-dimensional sensor or bio-medical applications. He has been stayed at the Cross-Disciplinary Electromagnetics Laboratory in the University of Wisconsin Madison as the visiting researcher in 2016. He has been a Principal Investigator of the PRESTO Program of Japan Science and Technology Agency (JST) from 2017 to 2021. He was a recipient of the 2012 Ando Incentive Prize for the Study of Electronics, 2013 Young Scientist's Prize by the Japanese Minister of Education, Culture, Sports, Science and Technology (MEXT), and 2014 Funai Achievement Award, 2022 KDDI Foundation Award, Contribution Award, and 2023 RIEC Award. He is a senior member of the Institute of Electrical and Electronics Engineering (IEEE), and a member of the Institute of Electrical Engineering of Japan (IEEJ), and the Japan Society of Applied Physics (JSAP).