Diffraction Signal-Based Human Recognition in Non-Line-of-Sight (NLOS) Situation for Millimeter Wave Radar

Jianghaomiao He, Shota Terashima, Hideyuki Yamada, and Shouhei Kidera, *Member, IEEE*

***Abstract*—In driver assistance or self-driving systems, millimeter-wave radar is an indispensable sensing tool because of its applicability to all weather conditions or non-line-of-sight (NLOS) sensing.This study focuses on a human recognition issue in the NLOS scenario by applying the support vector machine (SVM)-based machine learning approach to a diffraction signal.We show that there is a significant difference in diffraction signals between man-made objects (e.g., metallic cylinder and human body) even without motion.Hence, by exploiting such difference, an SVM achieves a high recognition rate, even in deeply NLOS situations.The experimental investigation, using a 24-GHz millimeter-wave radar in an anechoic chamber demonstrates that a diffraction signal-based recognition accurately classifies the real human and human mimicking man-made object, even in the NLOS scenario shielded by the parking vehicle.**

***Index Terms*—Automotive radar, diffraction effect, millimeter-wave (MMW) radar, non-line-of-sight (NLOS) detection, pedestrian detection, radar beamforming, supervised machine learning.**

I. Introduction

SELF-DRIVING automotive technology has been considerably advanced in recent years. The main environmental optical sensors, such as a camera or a laser range finder, are not applicable to bad weather conditions (e.g., rain, fog, and other smoggy situations) that limit driver's visibility and may cause an accident especially at night. Millimeter-wave (MMW) radar is an essential module of self-driving technology or advanced driver assistance systems (ADAS) because it can provide a reliable and environmentally robust sensing of surroundings, such as pedestrians in the vicinity of the vehicle, stationary objects (e.g., guardrails and barriers), and nearby vehicles, even in severe weather conditions. Particularly, in the pedestrian scenario, each motion of the human body (e.g., of arms and legs) exhibits a unique and periodic characteristic in Doppler velocity, which is often referred to as the micro-Doppler signature. Several studies have demonstrated that micro-Doppler features extracted from temporal frequency analysis can serve as an important basis and method of human–vehicle classification [1]–[9]. Xu *et al.* improved the range and Doppler velocity resolution assuming the micro-Doppler analysis in the line-of-sight (LOS) scenario [10]. The study [11] explored differences between Doppler signatures resulting from the motion of two different observation objects, the human body, and the vehicle, with a 24-GHz radar. In [12], pedestrian body movements and bicycle rotation-related micro-Doppler frequency shifts with significant variations were used to distinguish pedestrians and bicyclists from other objects on the road. Khomchuk *et al.* introduced a motion direction estimator for pedestrians in automotive radar [13]. In addition, MMW radar has been used to detect the human gait level [14]–[16] and limb motion [17] in a short-range and stationary environments or through-the-wall situations [18]–[21].

The aforementioned studies have focused on the machine learning-based human recognition or classification by an MMW radar. A previous study [22] has reported that by extracting features from distance and obtaining Doppler profiles, greater than 90% accuracy can be achieved when distinguishing between cars and pedestrians using a support vector machine (SVM). The literature [23] reported used an SVM to classify the pedestrians, cyclists, and vehicles using radar crosssection data, and another study [24] introduced a deep neural network-based gait level recognition system using micro-Doppler components. However, these studies assumed completely LOS scenarios, and few studies have focused on target identification issues in non-LOS (NLOS) situations, which would be important in real-world collision avoidance scenarios, such as children suddenly running out behind parked vehicles, that cannot be observed by a driver.

As the studies dealing with the NLOS case, the study [25] exploited multiple reflections in a complicated indoor situation and [26] discussed the limitations of detecting a human behind a wall. However, there are no studies that directly compared a diffraction signal between a human body and man-made objects, such as metallic cylinder or vehicles, which should be considered in the actual collision warning system to detect walking or running children that are invisible because of obstacles, such

Manuscript received February 17, 2021; revised March 31, 2021; accepted April 13, 2021. Date of publication April 15, 2021; date of current version May 6, 2021. This work was supported in part by PRESTO Program, Japan Science and Technology Agency, Japan, under Grant JPMJPR1771, and in part by the National Institute of Information and Communications Technology (NICT). *(Corresponding author: Shouhei Kidera.)*

Jianghaomiao He is with the Graduate School of Informatics and Engineering, University of Electro-Communications, 182-8585 Tokyo, Japan (e-mail: he.jianghaomiao@ems.cei.uec.ac.jp).

Shota Terashima and Hideyuki Yamada are with the Technical Research Center, Mazda Motor Corp., 221-0022 Yokohama, Japan (e-mail: terashima.s@mazda.co.jp; yamada.hidey@mazda.co.jp).

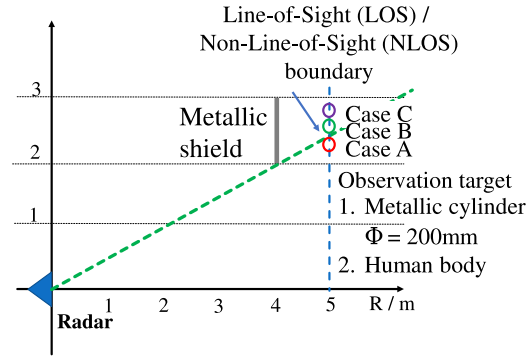
Shouhei Kidera is with the Graduate School of Informatics and Engineering, University of Electro-Communications, 182-8585 Tokyo, Japan, and also with the Japan Science and Technology Agency (JST), 332-0012 Kawaguchi, Japan (e-mail: kidera@ee.uec.ac.jp).

Digital Object Identifier 10.1109/JSTARS.2021.3073678

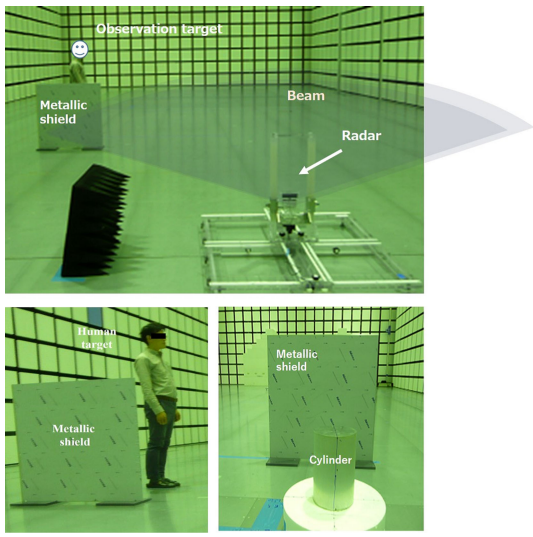
as parked vehicles. Hence, this study focuses on human body recognition from artificial objects using raw radar data-based machine learning recognition. Of note, an SVM is used in our proposed scheme, which has been developed and applied to various types of recognition schemes, by exploiting the so-called kernel trick for the nonlinear classification problem.

Specifically, we focus on the recognition scheme by exploiting a diffraction complex time series data for an SVM nonlinear classifier, where a unique characteristic of a human body even at static motion would be recognized from other artificial objects such as a metallic cylinder. Of note, the challenging issue in this scenario is that the strength of the diffraction signal is considerably lower [i.e., extremely low signal-to-noise ratio (SNR)], especially in the NLOS case. Therefore, we introduce a noise reduction scheme and design several types of feature vectors, where their time-variant structure is effectively implemented, and that its effectiveness has been demonstrated using a simple metallic plate obstacle in [27].

In the experimental validation, we use a 24-GHz band frequency-modulated continuous wave (FMCW) radar and assume two shielding objects, e.g., a metallic plate and a real vehicle such as SUV, where a real human body and a human-mimicking metallic cylinder are the targeted objects for the recognition located at both LOS and NLOS areas. The significance of this article is that we conducted a number of experimental tests, and the results demonstrate that the diffraction signal from a static human body differs considerably from that of a man-made object (i.e., a metallic cylinder), which contributes significantly to the high recognition rate in NLOS human detection and collision avoidance systems with an automotive radar sensor. Additionally, SVM-based classifiers can provide a much higher recognition rate even in NLOS situations by exploiting the time-variant complex reflection or diffraction signals. Furthermore, as a significant update to the literature [27], this study also demonstrates a high recognition performance between upright standing and stepping motions with real humans, which could be more informative in terms of providing reliable driver assessments in realistic collision avoidance scenarios.



(a)



(b)

Fig. 1. (a) Observation geometry. Green broken line is the boundary divided by the LOS and NLOS area by the metallic shield. (b) Optical image including the radar site and metallic shield in an anechoic chamber. Lower figure shows human body and metallic cylinder shadowed by the metallic shield.

II. Method

A. Observation Model

In this study, a pulse Doppler radar or the FMCW radar system is assumed; let τ be the slow time, which has a discrete form sampled by the pulse repetition interval (PRI). The recorded complex signal is denoted as $s(R,\tau)$, where $R = ct/2$, and t and c denote the fast time and speed of c of a radio wave in air, respectively.

Assuming NLOS or LOS situations, it is necessary to determine whether the recorded signal has a human originating feature. Here, because the radar can accurately measure the range from the sensor to the target by analyzing the measured data $s(R,\tau)$, e.g., local peak search of $|s(R,\tau)|$, we focus on a specific range $\tilde{R}$, where the target exists and analyze time variance along the slow-time τ direction. Note that, $\tilde{R}$ can be extracted from the local maxima of $s(R,\tau)$ along the R-axis. First, we introduce a simple feature vector $\boldsymbol{x}$ as

$$\boldsymbol{x} \equiv \left(s(R',\tau_0), s(R',\tau_1), \ldots, s(R',\tau_N)\right) \tag{1}$$

B. Support Vector Machine (SVM)

There are various types of recognition or classification methods such as linear, quadratic, or nonlinear discrimination analyses, specifically, logistic regression, random forest, k-means, hidden Markov model, neural network, and SVM. As a powerful and promising recognition approach, this study introduces an SVM-based nonlinear classifier. An SVM is one of the most promising and effective classifiers, which has been already applied to various types of classification issues [28]. Here, we briefly explain the basis of the SVM as follows. First, an SVM exploits a linear or nonlinear conversion from original data samples to those in the hyperspace to separate data samples by maximizing the margin to the separation boundary in the hyperspace. The aforementioned problem can be expressed as

$$\begin{aligned} &\text{minimize } ||\boldsymbol{w}|| \\ &\text{subject to : } y_i(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}_i - b) \geq 1, \text{ for } i = 1, ..., n \end{aligned} \tag{2}$$

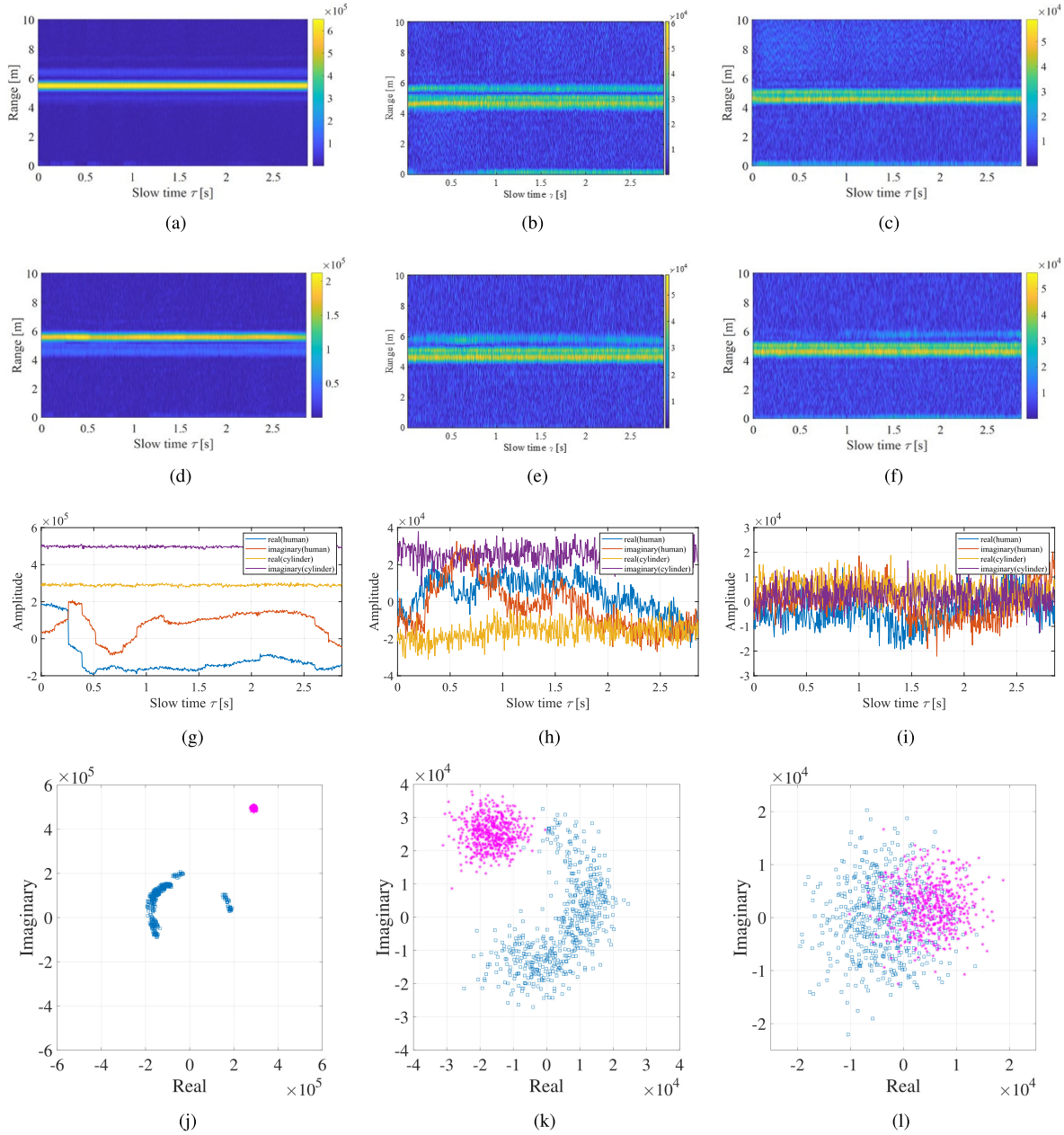


Fig. 2. Observation data at each target and Case at the metallic plate shield. First line: Radar raw data along slow time τ at the metallic cylinder target. Second line: Radar raw data along slow time τ at the static human body. Third line: Complex value reflection signal at $R = 5587$ mm, along slow time τ for both target case. (Blue and red lines denote the real and imaginary part of signal from human body. Orange and purple lines denote the real and imaginary part of the signal from the metallic cylinder.) Fourth line: Scattered plot of complex signal data for each case at $R = 5587$ mm. (Magenta and blue dots denote data of metallic cylinder and human body, respectively.). (a) Case A. (b) Case B. (c) Case C. (d) Case A. (e) Case B. (f) Case C. (j) Case A. (k) Case B. (l) Case C.

where $y_i \in \pm 1$ denotes the output of the hyperspace conversion and $\boldsymbol{\omega}$ is the normal weight vector to the hyperplane.

In most cases of classification problems, the training data could not be separated in the linear plane; a nonlinear data mapping is indefensible to achieve a higher classification performance. However, its computational complexity will greatly increase because of several nonlinear mapping operations. To achieve a nonlinear hyperspace conversion with less complexity, an SVM introduces a kernel trick, which can avoid the explicit mapping of data sample using only the inner product on the hyperspace. Note that, the kernel function transforms the training data so that a nonlinear surface could be transformed to a linear equation in nonlinear hyperspaces, and it makes role of the inner product between two points in the assumed hyperspace. Let $\boldsymbol{\phi}(\boldsymbol{x})$ be a nonlinear mapping operator; the kernel operator is defined as

$$K(\boldsymbol{x}_1, \boldsymbol{x}_2) = \boldsymbol{\phi}(\boldsymbol{x}_1)^{\mathrm{T}} \boldsymbol{\phi}(\boldsymbol{x}_2). \tag{3}$$

Of note, the aforementioned inner products in the hyperplane can be directly calculated without a nonlinear mapping operator, which greatly reduces the computational cost. Here, we introduce the most promising kernel function as the Gaussian

TABLE I
CLASSIFICATION RESULTS AND SNRS IN EACH CASE AND FEATURE EXTRACTION SCHEME AT THE METALLIC PLATE SHIELD

Method / Case	Accuracy				SNR [dB]	
	Feature 1 (Raw)	Feature 2 (Time derivative)	Feature 3 (Time shift)	Feature 4 (STFT)	Human	Cylinder
Case A (LOS)	100.0 %	99.0 %	100.0 %	100.0 %	27.73	35.43
Case B (Partially NLOS)	100.0 %	99.0 %	100.0 %	93.0%	8.35	9.33
Case C (Complete NLOS)	60.0 %	64.0 %	83.0 %	81.0%	1.51	-0.30

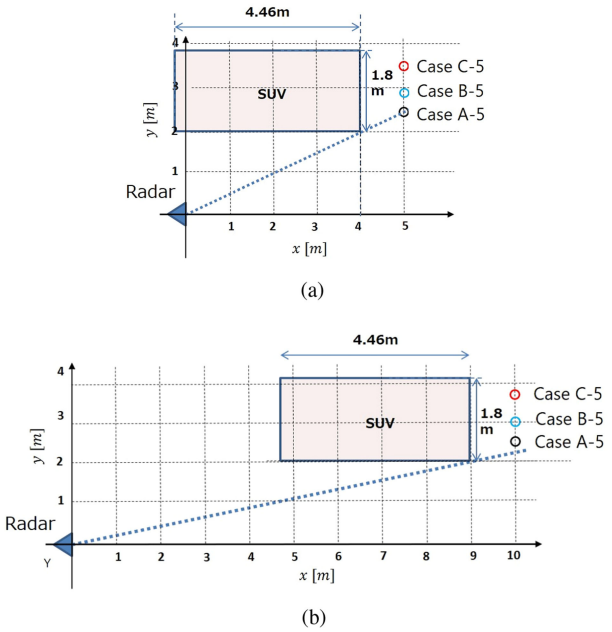


Fig. 3. Experimental geometries in the case of the real vehicle obstacle at different distance to the target. Blue broken line denotes the boundary divided by the LOS and NLOS area by SUV. (a) R = 5 m. (b) R = 10 m.

kernel

$$K(\boldsymbol{x}_1, \boldsymbol{x}_2) = \exp\left(-\frac{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^2}{2\sigma^2}\right) \tag{4}$$

where σ is the parameter of the Gaussian function and it is empirically determined. Gaussian kernel is one of the most useful kernels, where the prior knowledge of the data sample is not available.

C. Feature Extractions

The most important and critical part of the SVM is the feature extraction, and we introduce several different feature vectors for the SVM training using the real and imaginary parts of the discrete-time signal $\boldsymbol{x}$. To discriminate between an actual pedestrian and other objects (e.g., utility pole and guardrail), we introduce feature vectors by considering a temporal structure along slow time; it is based on the fact that even in static motion, the human body produces motion by breathing or posture control, which causes a slight displacement of the surface of human skin.

Hence, in this study, the following four types of features are introduced.

Feature 1:$s(R, \tau)$.
$\boldsymbol{X}_1 \equiv \boldsymbol{x}_1$, where

$$x_1[i] = (x[i]). \tag{5}$$

Feature 2: $s(R, \tau)$ and $\frac{\partial s(R,\tau)}{\partial \tau}$.
$\boldsymbol{X}_2 \equiv (\boldsymbol{x}_2[1], \boldsymbol{x}_2[2], \ldots, \boldsymbol{x}_2[N]$, where

$$\boldsymbol{x}_2[i] = \left(x[i], \frac{x[i+1] - x[i-1]}{t[i+1] - t[i-1]}\right)^T. \tag{6}$$

Feature 3: A number of time shifted data of $s(R, \tau)$ along τ direction.
$\boldsymbol{X}_3 \equiv (\boldsymbol{x}_3[1], \boldsymbol{x}_3[2], \ldots, \boldsymbol{x}_3[N])$, where

$$\boldsymbol{x}_3[i] = (x[i], x[i+1], \ldots x[i+M])^T. \tag{7}$$

where M denotes the length in the time shift.

Feature 4: Short time Fourier transform (STFT) of $s(R, \tau)$ in the τ direction.
$\boldsymbol{X}_4 \equiv (\boldsymbol{x}_4[1], \boldsymbol{x}_4[2], \ldots, \boldsymbol{x}_4[N])$, where

$$\boldsymbol{x}_4[i] = \mathcal{F}^{\mathrm{DFT}}\left[(x[i], x[i+1], \cdots x[i+M])^T\right] \tag{8}$$

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

It should be noted that all features are extracted at the specific range R', which can be extracted from the local peak of the average signal of $|s(R, \tau)|$.

III. EXPERIMENTAL TEST

A. Experimental Setup

This section describes the recognition performance by assuming pedestrian detection in LOS or NLOS situations. Fig. 1 shows the scene and geometrical setup for the experimental measurement in an anechoic chamber, provided by the Applied Electro-Magnetic Research Institute, National Institute of Information and Communications Technology (NICT). We use the FMCM multiple-input multiple-output radar produced by Sakura Tech Corp., which has a 24-GHz center frequency and a 700-MHz bandwidth. The sweep interval in the FMCW sequence is 5.2 ms, which corresponds to the PRI as the sampling interval of the slow time τ. The horizontal beamwidth is wide $\pm$ 45°, and the vertical beamwidth is narrowed to $\pm$ 6.5°. The optical image in Fig. 1 visualizes the upper part of the human body; the radar cannot receive a direct reflection because of a much narrow horizontal beam, i.e., the NLOS situation is recognized. We assume two different types of targets. One is a metallic cylinder with a 200-mm diameter and 300-mm height,

TABLE II
CLASSIFICATION RESULTS AND SNR WITH 0.35-M HEIGHT OF THE RADAR IN THE CASE OF THE REAL VEHICLE OBSTACLE

		Accuracy				SNR [dB]	
Distance	Case	Feature 1 (Raw)	Feature 2 (Time derivative)	Feature 3 (Time shift)	Feature 4 (STFT)	Human	Cylinder
5 m	Case A (LOS)	100.0 %	100.0 %	100.0 %	100.0 %	-9.5	3.7
	Case B (Partially NLOS)	100.0 %	100.0 %	100.0 %	100.0%	-27.0	-18.1
	Case C (Complete NLOS)	73.0 %	72.5 %	84.0 %	84.5%	-29.3	-29.8
10 m	Case A (LOS)	100.0 %	100.0 %	100.0 %	100.0 %	-9.0	-1.2
	Case B (Partially NLOS)	100.0 %	100.0 %	100.0 %	100.0%	-12.9	-15.0
	Case C (Complete NLOS)	100.0 %	100.0 %	100.0 %	100.0%	-23.3	-18.1

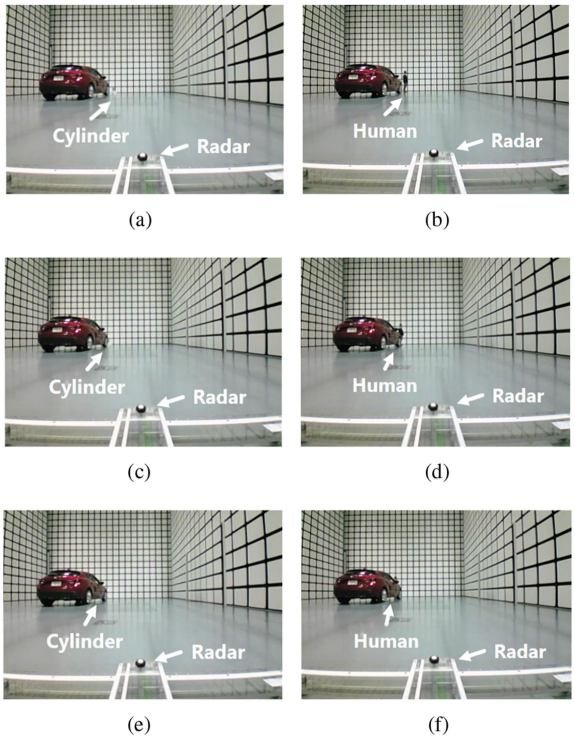


Fig. 4. Experimental scenes for each target and case with a real vehicle obstacle. (a) A, Cylinder. (b) A, Human. (c) B, Cylinder. (d) B, Human. (e) C, Cylinder. (f) C, Human.

which is put on the styrofoam table with a 200-mm height, that is the same height level as the radar. The other target is a real human with a 1700-mm height and wearing light clothes, who stands still during the measurement without motion. In the following evaluations, all reflection data were subtracted by the background data, which have been measured at a situation without any targets and shielding object and includes only direct or multipath reflection of the anechoic chamber.

B. Shield: Metallic Plate

The metallic shielding plate with 1000-mm width and 1000-mm height is positioned on the left-hand side from the radar, which generates the NLOS area. We put the metallic cylinder or the human body behind the metallic shield and generate LOS and NLOS situations by changing the location of these two objects along the y-axis, as shown in Fig. 1. Although the height of the shielding plate is lower than that of a human, the radar has a much narrower vertical beam compared to that a horizontal beam, and the NLOS case can be generated using this shield. Here, the radar receives a reflection signal from the object or obstacles at nearly the same height as the radar, e.g., 0.35 or 0.5 m. Then, even if the upper side of the human body is visible in the optical camera (see Fig. 1), the radar cannot receive a direct reflection echo from the NLOS area. We assume three cases, i.e., Case A (the target is in the LOS area), Case B [partially NLOS (target is on the boundary between LOS and NLOS areas)], and Case C [complete NLOS (target is completely within the NLOS area)], as shown in Fig. 1.

Fig. 2(a)–2(f) shows the reflection responses $s(R, \tau)$ for each case, and we clearly recognize two distinguished echoes at $R \simeq 4.5$ and $R \simeq 5.7$ m, which correspond to the reflections from the metallic plate and objects (cylinder or human body), respectively. The aforementioned figures also indicate that the signal strength from the object becomes weaker in the NLOS case, compared with that in the LOS case. Fig. 2(g)–2(i) illustrates the real and imaginary parts of reflection responses versus slow time τ and Fig. 2(j)–2(l) also show the Gaussian plane distributions of the complex reflection data $s(R, \tau)$ at the fixed distance of $R = 5587$ mm, where each target exists in each case and the number of samples along the slow time direction is 550; then, the total observation time is 2.86 s. By extracting the range-τ data at a specific range in which that target exists, an unnecessary reflection, such as shielding plate or multipath scattering in the chamber, could be excluded from the test data, and then, we could focus on only the diffraction or reflection signal from the targets (e.g., a cylinder or the human body). The results for Case A show that there are clearly unique characteristics in the pedestrian reflection response in its phase change, which are not observed for the metallic cylinder. Of note, because a human would hold the same posture without motion, this phase change is caused by breathing or slight displacement because of human posture control. Note that, we have also confirm that these phase changes could be measured from human without breathing, and this displacement is mainly caused by posture control in considering the height of radar beam (lower than human breast). Nevertheless, it is expected that the aforementioned difference between a metallic object and a static human is promising for the classification issue. In contrast, in case C (i.e., complete NLOS), the signal strength becomes considerably lower compared to the direct reflection signal in the LOS case; thus, it suffers from

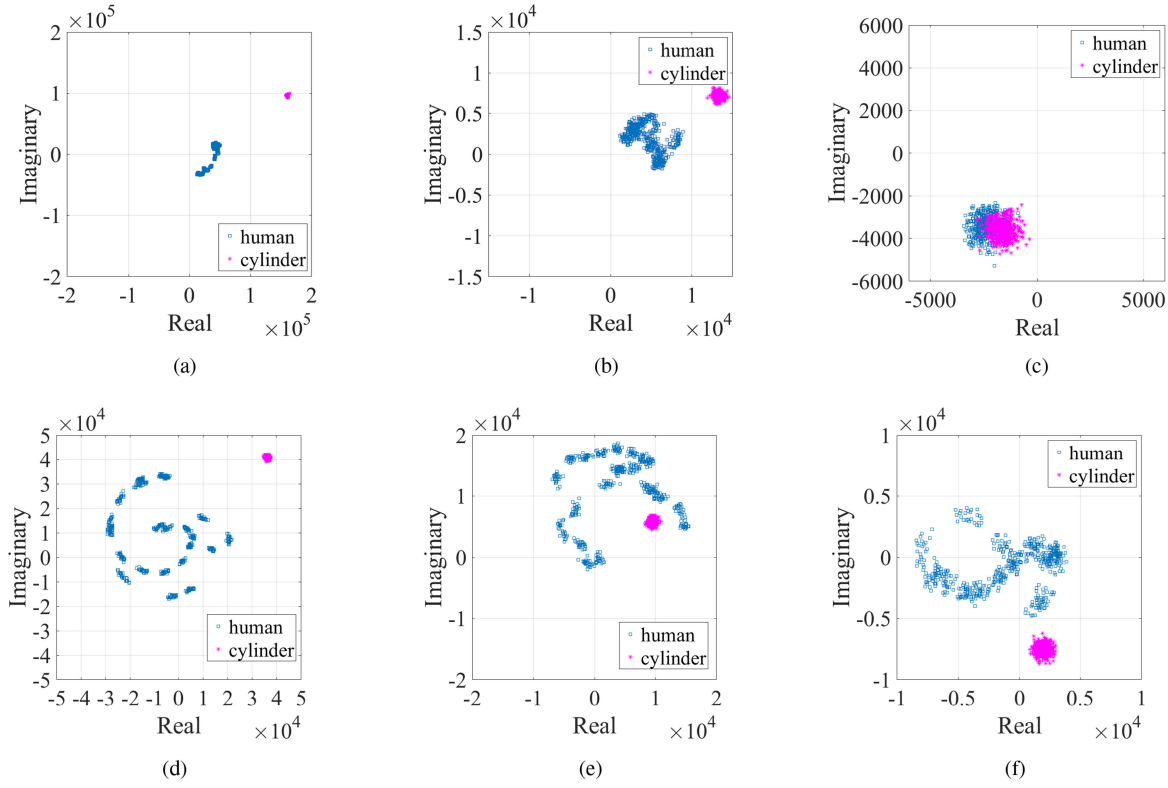


Fig. 5. Scattered plot on the Gaussian plane of reflection signals from human body without motion and metallic cylinder, where the height of the radar is set to 0.35 m, in the case of the real vehicle obstacle. Blue and magenta dots denote the response from the human body without motion and metallic cylinder, respectively. (a) A, $R = 5$m. (b) B, $R = 5$m. (c) C, $R = 5$m. (d) A, $R = 10$m. (e) B, $R = 10$m. (f) C, $R = 10$m.

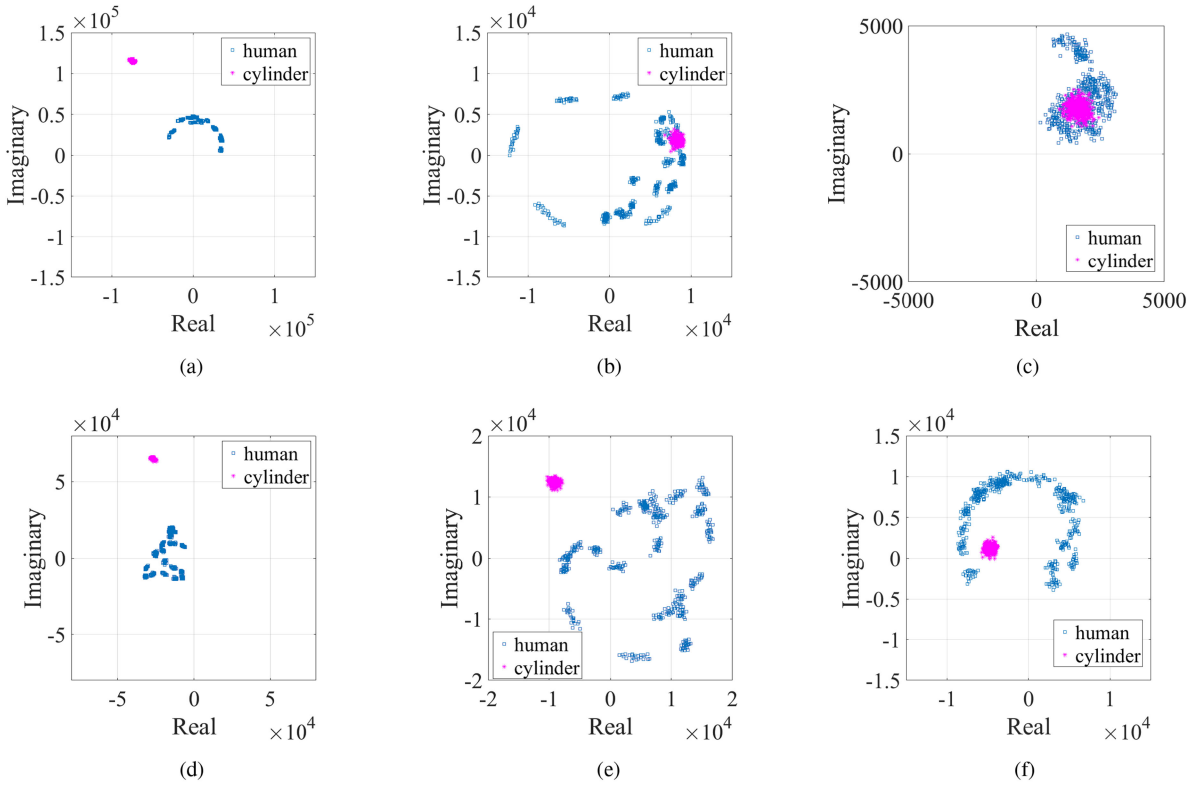


Fig. 6. Scattered plot on the Gaussian plane of reflection signals from the human body without motion and metallic cylinder, where the height of the radar is set to 0.5 m, in the case of the real vehicle obstacle. Blue and magenta dots denote the response from the human body without motion and metallic cylinder, respectively. (a) A, $R = 5$m. (b) B, $R = 5$m. (c) C, $R = 5$m. (d) A, $R = 10$m. (e) B, $R = 10$m. (f) C, $R = 10$m.

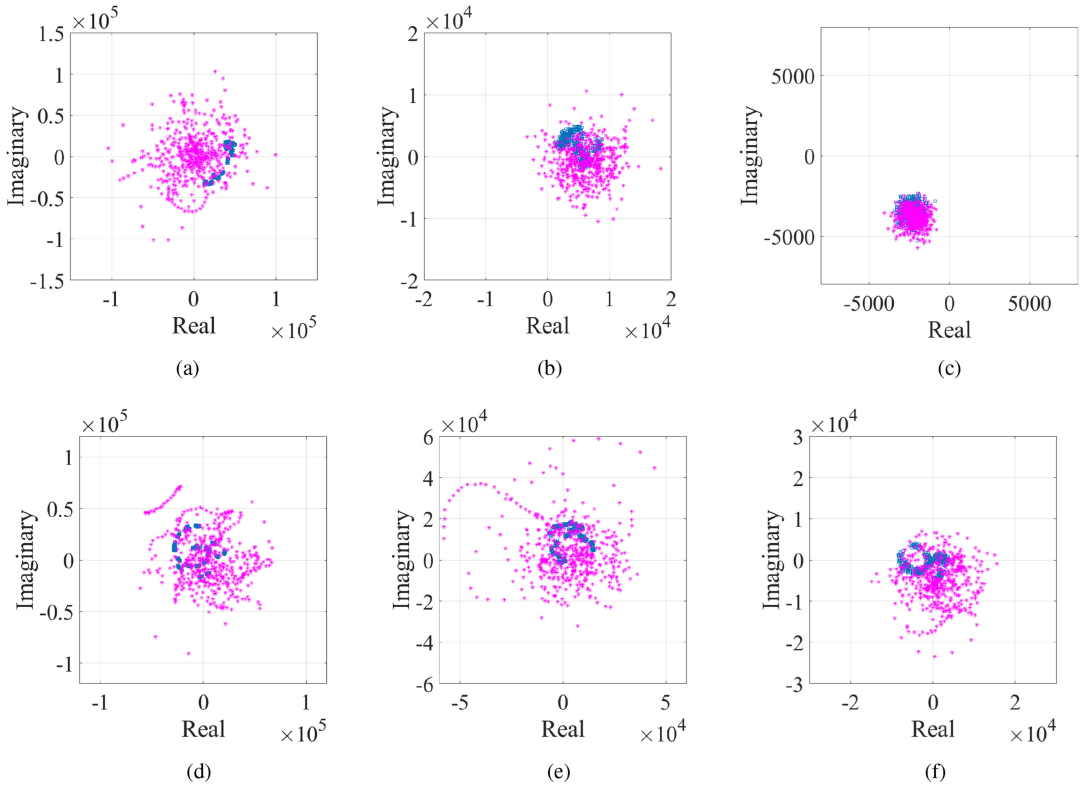


Fig. 7. Scattered plot on the Gaussian plane of each reflection signal from the human body with and without stepping motion, where the height of the radar is set to 0.35 m. Blue and magenta dots denote the response from human without and with stepping motion, respectively. (a) A, $R = 5$m. (b) B, $R = 5$m. (c) C, $R = 5$m. (d) A, $R = 10$m. (e) B, $R = 10$m. (f) C, $R = 10$m.

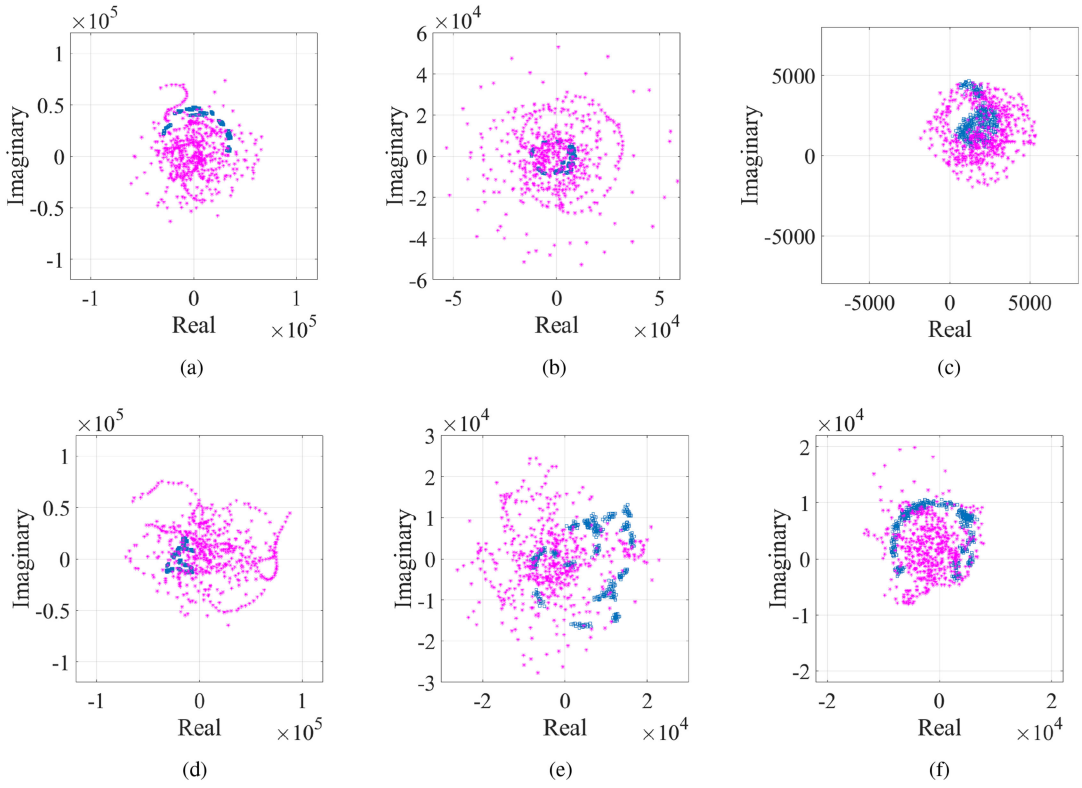


Fig. 8. Scattered plot on the Gaussian plane of each reflection signal from the human body with and without stepping motion, where the height of the radar is set to 0.5 m. Blue and magenta dots denote the response from human without and with stepping motion, respectively. (a) A, $R = 5$m. (b) B, $R = 5$m. (c) C, $R = 5$m. (d) A, $R = 10$m. (e) B, $R = 10$m. (f) C, $R = 10$m.

TABLE III
CLASSIFICATION RESULTS AND SNR WITH 0.5-M HEIGHT OF THE RADAR IN THE CASE OF THE REAL VEHICLE OBSTACLE

		Accuracy				SNR [dB]	
Distance	Case	Feature 1 (Raw)	Feature 2 (Time derivative)	Feature 3 (Time shift)	Feature 4 (STFT)	Human	Cylinder
5 m	Case A (LOS)	100.0 %	100.0 %	100.0 %	100.0 %	-8.4	2.2
	Case B (Partially NLOS)	100.0 %	100.0 %	100.0 %	100.0%	-22.4	22.0
	Case C (Complete NLOS)	96.5 %	95.0 %	98.5 %	98.0%	-30.8	-32.8
10 m	Case A (LOS)	100.0 %	100.0 %	100.0 %	100.0 %	-13.3	-3.2
	Case B (Partially NLOS)	100.0 %	100.0 %	100.0 %	100.0%	-18.3	-16.4
	Case C (Complete NLOS)	95.0 %	93.5 %	98.5 %	96.5 %	-21.4	-26.3

TABLE IV
PARAMETER DEPENDENCE OF M IN THE CASE C IN THE REAL VEHICLE OBSTACLE

			Accuracy	
Height	Distance	M	Feature 3 (Time shift)	Feature 4 (STFT)
0.35 m	5 m	2	81.0 %	83.5 %
		3	84.0 %	84.5 %
		5	75.5 %	68.0 %
	10 m	2	100.0 %	100.0 %
		3	100.0 %	100.0 %
		5	96.5 %	95.0 %
0.5 m	5 m	2	95.5 %	95.0 %
		3	100.0 %	100.0 %
		5	100.0 %	100.0 %
	10 m	2	96.0 %	95.5 %
		3	98.5 %	96.5 %
		5	93.5 %	92.0 %

a much lower SNR value. In addition, there is a significant difference between LOS and NLOS cases due to the contamination of additive random noise, which masks the characteristic feature inherent to human motion. Thus, a noise reduction scheme or an appropriate feature extraction is required by considering a unique temporal structure by human body motion. In addition, diffraction would occur at the corner of the shielding plate. Here, the vertical beam is quite narrow ($\pm 6.5\,\mathrm{deg}$); thus, a transmitting signal could not directly reach targets located at the NLOS area, i.e, case C, and the received signal in case C primarily contains the diffraction signal from the corner of the shield.

Next, the classification results are validated by an SVM using different feature vectors as described in Section II-C. We set the number of training and test datasets to 500 and 50, respectively, which is common for both pedestrian and cylindrical targets. Particularly, three discrete samples along slow time are used in Features 3 and 4. Table I quantitatively validates the effectiveness of our approach that the features, including temporal variance, have a certain level of superiority to that obtained using the raw data (Feature 1), especially for Features 3 and 4 with 80 % higher recognition rates. Of note, we tested a different number of sequences in Features 3 and 4 and confirmed that the chosen number of sequences slightly exceeds the recognition rate than other numbers; and then, there was no severe sensitivity for the chosen number of data sequences.

C. Shield: Real Vehicle

Next, we perform the test using another obstacle (i.e., a real vehicle), which simulates human body detection behind a parking vehicle. Figs. 3 and 4 show the observation geometries with two different distances, and the actual scenes for each case, where the SUV vehicle is located to the left side from the radar. Similar to the approach described in the previous subsection, we use a metallic cylinder object as a nonhuman target and a real human target, the dimension of which is the same as in the previous section. LOS (Case A), partially NLOS (Case B), and completely LOS (Case C) are assumed, as shown in Figs. 3 and 4. We investigated two different scenarios, where the distances to the target are selected as 5 or 10 m. Here, the height of the radar is set to 0.35 m.

Fig. 5 shows the Gaussian distribution of the complex-value signal for each object. As shown in Fig. 5, especially for the Cases A and B, we could recognize a substantial phase change in human body reflection or diffraction signal, which was not recognized in those from metallic cylinder. Table II shows the quantitative analysis for the recognition rate between the cylinder and the human body for each target's distance and radar height, and demonstrates that Features 3 or 4 provide a recognition rate higher than 80% even in the completely NLOS case at both 5- and 10-m distance, which is similar to that obtained in Table I using metallic plate shielding. Particularly, for the 10-m-distance case, the recognition rate will be upgraded to 100 % using any feature even in a complete NLOS situation. While the vertical beamwidth of the radar is designed to be narrow as $\pm 6.5°$, beam spread along the vertical direction is nonnegligible in the far field case (10-m distance), compared to that of the near field case (5-m distance). A part of the direct scattering echo from the lower or upper part of the human body, such as foot or head, can be received at the radar site, which would enhance the equivalent SNR compared to the near field case, as shown in Table II. Thus, the far-side case obtains a higher recognition rate than that near-side case. This is beneficial and convenient for the actual collision warning system in a self-driving system because the driver could timely notice a pedestrian shadowed by a parked vehicle.

D. Dependency of Radar Height

Additionally, we conducted a similar investigation using different radar heights. In this case, the radar height is set to

TABLE V
CLASSIFICATION RESULTS AND SNR WITH 0.35-M HEIGHT OF THE RADAR BETWEEN HUMAN BODY WITHOUT AND WITH STEPPING MOTION

		Accuracy				SNR [dB]	
Distance	Case	Feature 1 (Raw)	Feature 2 (Time derivative)	Feature 3 (Time shift)	Feature 4 (STFT)	Human w/o motion	Human w motion
5 m	Case A (LOS)	58.5 %	79.0 %	96.0 %	99.5 %	-9.5	-9.6
	Case B (Partially NLOS)	83.0 %	87.0 %	93.0 %	95.5 %	-27.0	-24.0
	Case C (Complete NLOS)	43.5 %	46.5 %	45.5 %	43.0 %	-29.3	-28.7
10 m	Case A (LOS)	52.5 %	44.5 %	57.5 %	70.0 %	-9.0	-4.6
	Case B (Partially NLOS)	67.0 %	84.0 %	100.0 %	100.0 %	-12.9	-9.2
	Case C (Complete NLOS)	52.5 %	50.5 %	59.5 %	84.0 %	-23.3	-17.0

TABLE VI
CLASSIFICATION RESULTS AND SNR WITH 0.5-M HEIGHT OF THE RADAR BETWEEN HUMAN BODY WITHOUT AND WITH STEPPING MOTION

		Accuracy				SNR [dB]	
Distance	Case	Feature 1 (Raw)	Feature 2 (Time derivative)	Feature 3 (Time shift)	Feature 4 (STFT)	Human w/o motion	Human w motion
5 m	Case A (LOS)	97.0 %	84.0 %	97.0 %	93.5 %	-8.4	-10.5
	Case B (Partially NLOS)	83.0 %	83.0 %	88.5 %	100 %	-22.4	-13.5
	Case C (Complete NLOS)	78.0 %	73.0 %	69.5 %	78.0 %	-30.8	-30.3
10 m	Case A (LOS)	47.5 %	54.5 %	52.5 %	81.0 %	-13.3	-7.4
	Case B (Partially NLOS)	57.5 %	57.5 %	81.5 %	90.0 %	-18.2	-17.9
	Case C (Complete NLOS)	88.5 %	79.5 %	85.0 %	82.0 %	-21.4	-23.5

0.5 m, which is 0.15 m greater than that of the previous setting. Fig. 6 shows the complex value reflection signal at each distance from the metallic cylinder and static human. Table III shows the recognition rates with different features and SNRs in this case. These figures and table also demonstrate that the temporal feature of the radar signal directly offers a discriminating capability, even at different radar heights. Of note, because the radar site is slightly higher than that of a previous setting, the reflection signal includes more components of human breathing, i.e., the reflection from the upper part of a human. Note that a received signal includes a reflection component along various paths, including 3-D spreading, e.g., a floor, a ceiling, or the backside wall of the anechoic chamber (or multiple reflections between them). However, the training or test data for object recognition are extracted by the local peak of the range profile by focusing on the specific range where a target exists. Then, unnecessary reflections from other paths with late or early arrivals could be excluded from the test data.

In addition, we investigate the dependence of the parameter M, namely, the data length in the Features 3 and 4. Table IV shows the recognition results at the case C, and validates that the recognition rates depend on the selected M, and the results at $M = 3$ provides a better performance in this case. We consider that if M is too small, the time-dependent feature could not be considered in the recognition, and if M is too large, the recognition problem would be complex due to much higher dimension input to the SVM.

E. Effect of Human Stepping Motion

This section describes the recognition performance for a human with stepping motion. We assume two different motions of the human body, i.e., standing upright without any motion and standing upright with stepping motion at the same location. This comparison allows us to discriminate the human motion in the NLOS case, which allows the driver to better determine whether a pedestrian would or would not jump out on the road. Figs. 7 and 8 show the scattered plot on the Gaussian plane of reflection responses, where the target is located 5 or 10 m from the radar site, and the radar height is set to 0.35 or 0.5 m. These figures indicate that there are significant discrepancies in responses between with and without stepping motion, and a larger phase change and amplitude variance are confirmed for any case with stepping motion because the motion of each part of the human body fluctuates the amplitude or phase of the backscattered signal. However, in the case of the near side (5 m) and with a lower height of radar (0.35 m), the responses are much close between with and without motion, which is considerably contaminated with a random noise. Furthermore, Tables V and VI summarize the recognition rates for each case of a human body without and with stepping motion and the SNRs. The aforementioned table shows that a high recognition rate (e.g., higher than 75 %) is achieved even for the complete NLOS case, except the case of 5-m distance and 0.35-m height, using Feature 4 (STFT), and this is because the Doppler velocity components are more significant to recognize the motion of the human body. Note that, we can see a certain level of the drop of the recognition rate, compared with that of discrimination between the metallic cylinder and the static human body, as shown in Tables II and III, we consider that this occurs because both complex signals of the human body without and with stepping motion include a common component due to human breathing or posture control, which does not occur in a metallic cylinder target. These results show that we can recognize a situation, whether a human has a motion or not, at a certain level of accuracy, which helps drivers to make better judgments in realistic situations.

IV. Conclusion

This article proposes the raw data-based radar signal recognition method that is based on an SVM classifier for the MMW radar pedestrian detection in NLOS situations. First, we determined that there are significant differences in complex reflection responses between a metallic cylinder and a human body with or without motion, where the phase change due to posture control or breathing can be recognized in the responses of the human body. By exploiting these results, we introduce four types of a feature vector, by considering the temporal variance along slow time, into the SVM-based nonlinear classifier.

In experimental examinations, we introduce two shielding cases, i.e., metallic plate and real vehicle. In the first test using a metallic shield, we clearly demonstrated that the responses from the metallic cylinder (i.e., an artificial object) have a very stable response in the Gaussian plane, whereas those of the human body have a considerable amount of phase change even during static motion. Although the aforementioned difference would not be recognizable in a perfect NLOS situation, the proposed time-frequency feature in an SVM largely improves the recognition rate to approximately 80%. Additionally, in the real vehicle shielding case, we also confirm the dependence of the radar height and distance to the object. Of note, the recognition rate was further improved at a longer distance to the object because of the smaller diffraction angle, which is beneficial to the application of the ADAS system or self-driving system and provides extra time to judge or inform the driver about a hidden pedestrian behind a parked vehicle. This study investigated the recognition issue for human motion with stepping motion, and the number of studied cases demonstrated that it accurately recognized whether the human had a walking motion or not, even in complete NLOS situations. Of note, although there are various machine learning approaches, this study introduces a simpler and faster classifier SVM. However, there should be further investigation using other classifiers or features, which is our important future work. In realistic situations, a number of objects, e.g., the human body, guard rails, utility poles, or trees, may exist in the observation area. If each object is separated over a range resolution, these components could be decomposed by extracting the data at a specific range, which could be extracted by the local peak of a range profile, as in the process of this article. In contrast, if numerous objects are located in the range resolution, its separation would become difficult, and another decomposition scheme, e.g., wavenumber (angular) space conversion or Doppler velocity space, is required. We plan to investigate such extensions in future work.

References

[1] Y. Kim, S. Ha, and 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.

[2] X. Shi, X. Yao, X. Bai, F. Zhou, Y. Li, and L. Liu, "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.

[3] 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.

[4] G. Li and P. K. Varshney, "Micro-Doppler parameter estimation via parametric sparse representation and pruned orthogonal matching pursuit," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 12, pp. 4937–4948, Dec. 2014.

[5] L. Du, B. Wang, P. Wang, Y. Ma, and H. Liu, " Noise reduction method based on principal component analysis with beta process for micro-Doppler radar signatures," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 4028–4040, Aug. 2015.

[6] J. Le Kernec and C. Li, "Continuous human motion recognition with a dynamic range-Doppler trajectory method based on FMCW radar," *IEEE Trans. Geosci. Remote Sens.* vol. 57, no. 9, pp. 6821–6831, Sep. 2019.

[7] M. Li, T. Chen, and H. Du, "Human behavior recognition using range-velocity-time points," *IEEE Access*, vol. 8, pp. 37914–37925, 2020.

[8] C. Clemente and J.J. Soraghan, "Vibrating target micro-Doppler signature in bistatic SAR with a fixed receiver," *IEEE Trans. Geosci. Remote Sens.* vol. 50, no. 8, pp. 3219–3227, Aug. 2012.

[9] M. Gustafsson, A. Andersson, T. Johansson, S. Nilsson, A. Sume, and A. Örbom, "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.

[10] Z. Xu, C. J. Baker, and S. Pooni, "Range and Doppler cell migration in wideband automotive radar," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5527–5536, Jun. 2019.

[11] A. Rizik, A. Randazzo, R. Vio, A. Delucchi, H. Chible, and D. D. Caviglia, "Feature extraction for human-vehicle classification in FMCW radar", in *Proc. 26th IEEE Int. Conf. Electron., Circuits Syst.*, Nov. 2019, pp. 131–132.

[12] D. Belgiovane and C.-C. Chen, "Micro-Doppler characteristics of pedestrians and bicycles for automotive radar sensors at 77 GHz," in *Proc. 11th Eur. Conf. Antennas Propag.*, Mar. 2017, pp. 2912–2916.

[13] 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, Jun. 2016.

[14] K. Zhang, S. C. Lan, and G. Y. Zhang, "Mining spatio-temporal features from mmW radar echoes for hand gesture recognition," in *Proc. IEEE Asia-Pacific Microw. Conf.*, Dec. 2019, pp. 93–95.

[15] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1709–1723, Aug. 2018.

[16] M. S. Seyfioğlu 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.

[17] Y. Wang and Y. Zheng, "An FMCW radar transceiver chip for object positioning and human limb motion detection," *IEEE Sensors J.*, vol. 17, no. 2, pp. 236–237, Jan. 2017.

[18] S. Guo, G. Cui, L. Kong, Y. Song, and X. Yang, "Multipath analysis and exploitation for MIMO through-the-wall imaging radar," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3721–3731, Oct. 2018

[19] G. Gennarelli, R. Solimene, F. Soldovieri, and M. G. Amin, "Three-dimensional through-wall sensing of moving targets using passive multistatic radars," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 1, pp. 141–148, Jan. 2016

[20] Y. Song, J. Hu, T. Jin, Z. Li, N. Chu, and Z. Zhou, "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, Sep. 2019,

[21] H. Li, G. Cui, L. Kong, G. Chen, M. Wang, and S. Guo, "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, Apr. 2019.

[22] S. Heuel and H. Rohling, "Two-stage pedestrian classification in automotive radar systems," in *Proc. 12th Int. Radar Symp.*, Sep. 2011, pp. 477–484.

[23] X. Tang, X. Wu, S. B. Yeap, R. Luo, T. Dai, and L. Huang, "Experimental results of target classification using mmWave corner radar sensors," in *Proc. Asia-Pacific Microw. Conf.*, Nov. 2018, pp. 842–844.

[24] S. Lee, Y.-J. Yoon, J.-E. Lee, and S.-C. Kim, "Human-vehicle classification using feature-based SVM in 77-GHz automotive FMCW radar," *IET Radar Sonar Navigat.*, vol. 11, no. 10, pp. 1589–1596, 2017.

[25] S. Guo, Q. Zhao, G. Cui, S. Li, L. Kong, and X. Yang, "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, Jan. 2020, doi: 10.1109/JSTARS.2020.2963924.

[26] S. Li, G. Cui, S. Guo, C. Jia, L. Kong, and X. Yang, "On the electromagnetic diffraction propagation model and applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 884–895, Feb. 2020, doi: 10.1109/JSTARS.2020.2974529.
[27] J. He, S. Terashima, H. Yamada, and S. Kidera, "Human body recognition method using diffraction signal in NLOS scenario for millimeter wave radar," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2020, pp. 766–769.
[28] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Feb. 1995.

Jianghaomiao He received the B.E. degree in electrical and electronic engineering from the Baoji University of Arts and Sciences, Baoji, China, in 2017, and the M.E. degree in 2021 from the Graduate School of Informatics and Engineering, University of Electro-Communications, Tokyo, Japan, where he is currently working toward the Ph.D. degree.

His research interests include signal processing for millimeter and microwave radar as well as its applications.

Shota Terashima received the B.E. degree in electrical and electronic engineering from Seikei University, Tokyo, Japan, in 2005.

He joined Mazda Motor Corporation, Hiroshima, Japan, in 2005, and has been involved in research on automotive electronic control units. He is currently involved in research work for automotive radar propagation.

Hideyuki Yamada received the B.S. degree in physics from Yamaguchi University, Yamaguchi, Japan, in 1982.

He was with Hitachi Computer Engineering Corporation, Kanagawa, Japan, from 1982 to 1993. In 1993, he joined Mazda Motor Corporation, Yokohama, Japan. From 1996 to 2011, he was involved in research on the intelligent transportation system. Since 2012, he has been involved in research on vehicle-to-vehicle radio wave propagation and automotive radar propagation.

Shouhei Kidera (Member, IEEE) received the B.E. degree in electrical and electronic engineering and the M.I. and Ph.D. degrees in informatics, all from Kyoto University, Kyoto, Japan, in 2003, 2005, and 2007, respectively.

He has been with the Graduate School of Informatics and Engineering, University of Electro-Communications, Tokyo, Japan, since 2009, and is currently an Associate Professor. His current research interests include advanced radar signal processing or electromagnetic inverse scattering issue for ultra wideband 3-D sensor or biomedical applications. He was with the Cross-Disciplinary Electromagnetics Laboratory, University of Wisconsin Madison, as the Visiting Researcher in 2016. He was a Principal Investigator of the PRESTO Program with the Japan Science and Technology Agency from 2017 to 2021.

Dr. Kidera was the recipient of the 2012 Ando Incentive Prize for the Study of Electronics, the 2013 Young Scientist's Prize by the Japanese Minister of Education, Culture, Sports, Science and Technology, and the 2014 Funai Achievement Award. He is a Senior Member of the Institute of Electronics, Information, and Communication Engineers of Japan, and a Member of the Institute of Electrical Engineering of Japan, and the Japan Society of Applied Physics.