# Model-Based Vehicle Position Estimation Using Millimeter Wave Radar

Yoshihiro Suzuki, Yosuke Sugiura, Tetsuya Shimamura, Osamu Isaji, Kazuaki Hamada, and Kazuhiko Shite

*Abstract*— **In this paper, we propose a method to accurately estimate the position of the vehicle by millimeter wave radar (MWR). In recent years, many techniques of autonomous driving have been developed actively. In order to realize a safety autonomous driving system, MWR plays an important role for its inherent robustness against external circumstances. However, MWR has low spatial resolution. To achieve highly accurate estimation, we propose a model-based matching approach to the point cloud observed from MWR. Simulation experiments show that accurate estimation of the position of a moving vehicle can be obtained.**

*Index Terms*—**Autonomous driving, object detection, Millimeter Wave Radar (MWR), point cloud, least-squares method.**

## I. Introduction

In recent years, many techniques of Advanced Driving Assistance System (ADAS) have been developed actively. Examples of these are Adaptive Cruise Control (ACC), Parking Aid, and so on. For these techniques, various types of sensors are used, which are camera, LiDAR (Light Detection and Ranging), millimeter wave radar (MWR), and so on. The mainstream of vehicle detection is based on image processing using a camera. Although the camera is easy to deploy, the image-based method provides low accuracy to measure the distance to the target vehicle and it is seriously affected by external factors such as light source and weather. LiDAR has been also used in recent years due to its high resolution. However, the detection performance for long distance measurement by LiDAR is insufficient, and it is susceptible to external factors in the same way as the camera. On the other hand, MWR can measure the distance to the target object with high accuracy, and it is robust to external factors. However, MWR possesses the disadvantage of low spatial resolution. Owing to that, it is difficult to accurately estimate the shape and position of the target object. For these reasons, the sensors mentioned above are often used in a combined form. Actually, MWR often assists the camera and LiDAR. However, in order to improve the overall sensing performance, it is necessary to improve the performance of each sensor.

Radar is a sensor that measures the distance and angle to the target object using the radio wave reflected from the target object. The reflection strength which represents the power of the radio wave reflected from the object can be obtained and utilized in radar systems. The point cloud is obtained from the radar points extracted by threshold processing. This is because the radar points reflected from the target object have high reflection strength. This feature is utilized for object detection. MWR which uses the millimeter wave radio can be, especially, used as a car-equipped radar for detecting obstacles. MWR is a small device but has detection ability of long-distance measurement. On the other hand, the spatial resolution of MWR is lower due to the wavelength. In fact, the number of radar points that can be utilized for object detection is extremely small, and thus it is very difficult to estimate the shape and position of the object. This is a different feature when compared to LiDAR. Therefore, signal processing to improve the low spatial resolution of MWR has been tried as front-end processing [1]-[3].

However, accurate estimation of the object position by improving only the spatial resolution of MWR may be difficult. This is because the point cloud obtained by MWR often includes outliers due to multipath and clutter phenomena. The point cloud obtained by LiDAR is sufficient with high resolution, and object recognition is often conducted by converting the point cloud data to depth map data [4]. The point cloud expressing the target object in LiDAR is of high density. Hence, it is possible to eliminate outliers in advance by pre-processing. On the other hand, the point cloud obtained by MWR is insufficient, and denoising is difficult. This is because it is difficult for MWR to distinguish the outliers. Therefore, a matching processing technique considering the outliers would be required.

In this paper, we propose a method to accurately estimate the position of the vehicle even in the above conditions. We utilize a model which has reference points (denoted as evaluation point in this paper). These points are used to evaluate a distance between the model and point cloud. We accomplish robust estimation for the outliers by introducing a new projection function and by embedding constraint conditions. In simulation experiments, position estimation is set out in a situation where a single vehicle is the target object. The estimation error is evaluated and discussed.

After this Introduction, Section 2 explains the model of the observation data of the vehicle, and Section 3 describes the related work for the object detection and point cloud fitting. Section 4 derives the proposed method and Section 5 shows simulation experiments. In Section 6, we conclude this paper.

Manuscript received January 29, 2019; revised July 28, 2019.

Yoshihiro Suzuki, Yosuke Sugiura, and Tetsuya Shimamura are with the Graduate School of Science and Engineering, Saitama University, Saitama 338-0823, Japan (e-mail: {suzuki, sugiura, shima}@sie.ics.saitama-u.ac.jp).

Osamu Isaji, Kazuaki Hamada, and Kazuhiko Shite are with DENSO TEN Limited, Kobe 652-8510, Japan (e-mail: {osamu.isaji, kazuaki.hamada, kazuhiko.shite}@denso-ten.com).

doi: 10.18178/ijfcc.2019.8.3.547

## II. Vehicle Data Modeling

By radar, we can obtain three types of information; direction, distance, and reflection strength. The former two can be easily transformed into a 2-D coordinate. The reflection strength is affected by the material of the target object. It is easy to detect a metal object such as vehicles, while it is difficult to detect an object with less reflection, which is, for example, a pedestrian.

In this paper, referring to [5], a vehicle is recognized as a 2-D model with four corners and four wheel covers of the vehicle. The four corners and four wheel covers are considered as the reflection points of the radio wave. Fig. 1 shows the vehicle model used in this paper. The size of this vehicle is 1.8 m in width and 4.9 m in height. The wheel covers are set to a smaller size by 1.0 m from the front and rear of the vehicle. As shown in Fig. 1, the vehicle model has totally eight reflection points, which are illustrated as circles.

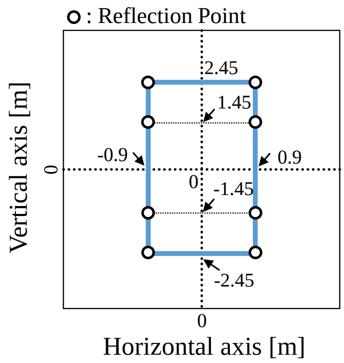



Fig. 1. Arrangement of the reflection point on the vehicle model.

In MWR, the position of radar points is often scattered around the true one due to the existence of noise and due to the low spatial resolution of MWR. In this paper, we model the radar points as a spread that obeys the Gaussian distribution around the true reflection points. In addition, the point cloud obtained in real situation contains many outliers such as reflections from the ground, other obstacles, and so on. However, it is difficult for MWR to distinguish the target object from the outliers since the point cloud is poor. Therefore, it is not easy to remove the outliers by front-end processing. In simulation experiments to be followed, we assume that the outliers occur at random positions and cannot be removed by front-end processing.

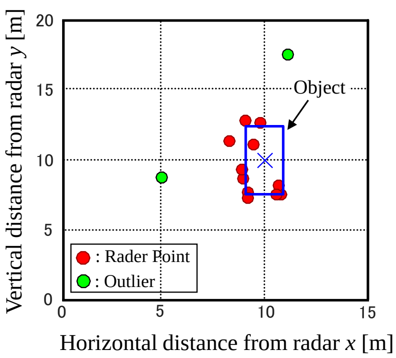



Fig. 2. An example of target vehicle and point cloud obtained by MWR.

An example of simulation data is shown in Fig. 2. In Fig. 2, the centroid of the target vehicle is located on $(x, y) = (10, 10)$. The blue bounding box indicates the target vehicle position, the red points indicate the point cloud, and the green points indicate the outliers. As can be seen in Fig. 2, the radar points are distributed along the boundary of the target vehicle.

## III. Related Work

For LiDAR-based object detection, the detection process requires two steps; region proposal and classification [4]. In the region proposal step, multi-scale anchors are generated as candidates of the region occupied by the object. In the classification step, the type of the object is detected, and the detailed information such as size and orientation are estimated. Most of such state-of-the-art approaches utilize convolutional neural networks (CNN). In the method shown in [4], the point cloud is projected onto a depth map as input data. PointNet [6] utilizes the point cloud directly as the input, and VoxelNet [7] divides the point cloud into 3D voxels. However, for MWR-based object detection, it is invalid to utilize such methods because the point cloud obtained by MWR is of low density and with outliers. Therefore, it is necessary to set the model of the object and then match the model with the point cloud because a small number of points represent the shape of the object.

For fitting between the point cloud and the model, a plane [8], [9] or a voxel [10] is designed as the model. These approaches are usually based on the principle of the least-squares method. However, the squared norm is affected by outliers, and a large estimation error occurs. When considering the model as the point cloud, ICP (Iterative Closest Point) [11], which is an algorithm for fitting two point-clouds, is often utilized. However, this algorithm leads a local optimal solution unless the two point-clouds are set to the appropriate initial position. Also, Generalized-ICP [12], which searches for the global solution, has been proposed. It would be difficult to match the model with the point cloud obtained by MWR due to the characteristics of ICP [13] that means those of weakness against ununiform samples and outliers. Therefore, it is not suitable to utilize ICP directly for MWR-based position estimation.

## IV. Proposed Approach

### A. Matching with the Template Model

In MWR-based position estimation of the object, we should firstly design a template model to match the set of radar points reflected from the surface of the object. The template model is constructed from $M$ evaluation points ($M$ is a positive integer) for the simple calculation instead of a line object. The evaluation points are arranged so as to match to the reflection points on the vehicle model shown in Fig. 1.

The objective function to minimize is defined by the sum of the distance between each evaluation point of the template model and each radar point. Letting $\boldsymbol{T}_i = (T_x^{(i)}, T_y^{(i)})$ and $\boldsymbol{S}_j = (S_x^{(j)}, S_y^{(j)})$ be the $i$-th evaluation point and the $j$-th radar point, respectively. The objective function, $E$, can be described as follows;

$$E = \sum_i^M \sum_j^N d_{ij} \quad (1)$$

$$d_{ij} = \|T_i - S_j\|_2^2 \quad (2)$$

### B. Weighting for the Evaluation Point

Next, in order to improve the matching accuracy between the point cloud and the template model, we utilize a weight for each evaluation point. The weight of the *i*-th evaluation point, $w_i$, is set so that the area where the radar points gather is largely weighted. Using the weight $w_i$, the objective function $E$ can be redefined as follows;

$$E = \sum_i^M w_i \sum_j^N d_{ij} \quad (3)$$

The weight $w_i$ can be determined uniquely from the relative position between the radar device and the template model. The reflection is strong at the evaluation point close to the radar device, which indicates high reliability to represent the target object. On the other hand, the reflection is weak at the evaluation point far from the radar device. This means that weak reflection points may correspond to outliers. Therefore, $w_i$ can be defined as follows;

$$w_i = \cos\theta_i \quad (4)$$

$$\theta_i = \tan^{-1}\frac{T_y^{(i)}}{T_x^{(i)}} \quad (5)$$

where $\theta_i$ is the angle between the position of the radar device and the *i*-th evaluation point, which is limited to $-\pi/2 \leqq \theta_i \leqq \pi/2$. According to (4), $w_i$ becomes a larger value when the evaluation point is closer to the radar device. Also, due to the characteristics of the radar, it is impossible to observe the back of the object where the radar points disappear. Therefore, we set the value of $w_i$ to 0 in this case. Fig. 3 shows an example of actual weights. In Fig. 3, the centroid of the vehicle is located on $(x, y) = (-5, 10)$.

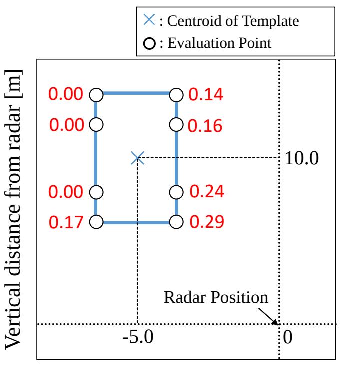



Fig. 3. An example of the weights of the evaluation points.

### C. Data Formatting by Threshold

Generally, the radar points reflected from the target object have high reflection strength. On the other hand, the radar points which have low reflection strength are likely outliers. That is, the radar points with low reflection strength may cause the error of object position estimation. Therefore, the radar points with lower reflection strength than a threshold value, $\beta$, are removed from the objective function. Therefore, the objective function $E$ is further redefined as follows;

$$E = \sum_i^M w_i \sum_j^N \phi_j d_{ij} \quad (6)$$

$$\phi_j = \begin{cases} 1 & (r_j > \beta) \\ 0 & (\text{otherwise}) \end{cases} \quad (7)$$

where $r_j$ is the reflection strength of the *j*-th radar point.

### D. Suppression of the Influence of Outliers

Outliers may cause the error of object position estimation because the distance to outliers tends to become long. In order to suppress the effect of the outliers, we propose the following projection;

$$\tilde{d}_{ij} = f(d_{ij}) \quad (8)$$

where $f(\cdot)$ is a monotonically increasing function, which should be designed so as to compress the large values which corresponds to outliers. In this paper, $f(x)$ is defined as follows;

$$f(x) = \begin{cases} x & (x < \alpha) \\ \log(x - \alpha + 1) + \alpha & (\text{otherwise}) \end{cases} \quad (9)$$

where $\alpha$ is a positive constant. Fig. 4 shows the shape of $f(x)$. As shown in Fig. 4, $f(x)$ plots a linear line for $x < \alpha$ and a logarithmic curve for $x \geqq \alpha$.

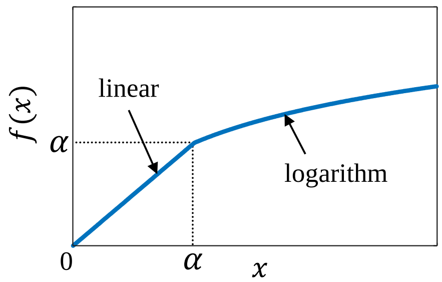



Fig. 4. Shape of the projection f(x).

### E. Constraint Conditions for Improving Tracking Accuracy

The coordinate of the position estimated at the current time, $\Theta_t$, can be expressed as follows;

$$\Theta_t = \left(\frac{1}{M}\sum_i^M T_x^{(i)}, \frac{1}{M}\sum_i^M T_y^{(i)}\right) \quad (10)$$

When the object moves continuously, it is assumed that the position at the current time, $\Theta_t$, exists near the position estimated at the previous time, $\Theta_{t-1}$. Then, the problem can be expressed by the following equation:

$$\text{minimize} \quad \sum_i^M w_i \sum_j^N \phi_j \tilde{d}_{ij} \quad (11)$$

$$\text{subject to} \quad \|\Theta_t - \Theta_{t-1}\|_2 < K \quad (12)$$

In (12), $K$ is a positive constant, which limits the search range. The above problem can be rewritten to the following formula by use of Lagrange multiplier in inequality constraint:

$$E = \sum_i^M s_i \sum_j^N \phi_j \tilde{d}_{ij} + \lambda \|\Theta_t - \Theta_{t-1}\|_2 \tag{13}$$

where $\lambda$ is a positive constant. In this paper, we obtain the value of $\Theta_t$ by raster scanning of the template in the search area. In order to ensure matching accuracy, the scanning resolution should be set to a small value. From this point of view, we set the scanning resolution to 0.2 m.

## V. Experiments

### A. Parameter Tuning

In this section, we simulated a scene where one vehicle equipped with an MWR device (which is referred to as the ego vehicle) tracks a vehicle passing through in a right area of the ego vehicle. When we show the location of the ego vehicle at the origin $(0, 0)$ in the x-y coordinate, the moving vehicle is found in the front right direction of the ego vehicle. The moving vehicle is assumed to be the target vehicle. The search area for position estimation is, in this case, restricted to $0\,[m] < x < 20\,[m]$, $5\ [m] < y < 25\,[m]$. For our simulation, we assume that the ego vehicle goes forward at a speed of 0.2 m per observation time. The target vehicle passes through at a faster speed than that the ego vehicle. Under the above conditions, simulation data for 100 observation times are prepared and utilized for evaluating the performance of the proposed method.

As the evaluation metric, the error between the centroid of the estimated position and the true position of the target vehicle is utilized. The evaluation value is calculated by averaging the estimation error, which is considered according to the following three groups;

- Long distance $(20\,[m] < \overline{\Theta}_t^{(y)} < 25\,[m])$
- Medium distance $(10\,[m] < \overline{\Theta}_t^{(y)} < 20\,[m])$
- Short distance $(5\,[m] < \overline{\Theta}_t^{(y)} < 10\,[m])$,

where $\overline{\Theta}_t = (\overline{\Theta}_t^{(x)}, \overline{\Theta}_t^{(y)})$ is the true position of the target vehicle.

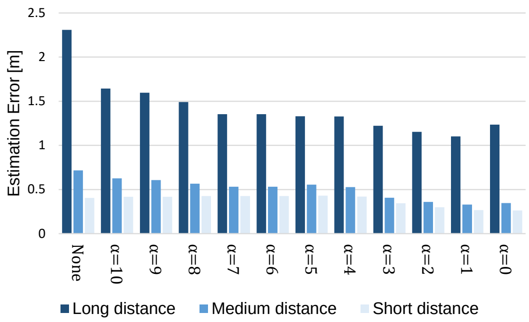



Fig. 5. Parameter tuning results with changing α.

Fig. 5 shows the estimation error versus the constant value $\alpha$. Improvement in estimation accuracy can be confirmed as the value $\alpha$ decreases. From the results in Fig. 5, $\alpha = 1$ is used in subsequent evaluation experiment. Fig. 6 shows the estimation error versus the constant value $\lambda$. Improvement in estimation accuracy can be confirmed as the value $\lambda$ increases. From the results in Fig. 6, $\lambda = 1.0$ is used in subsequent evaluation experiment.

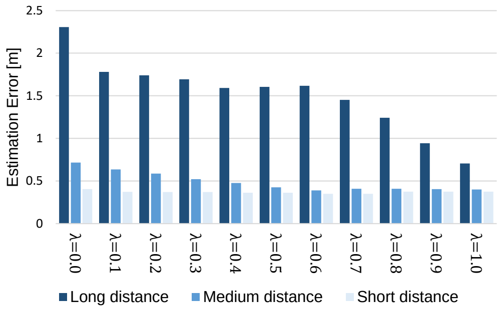


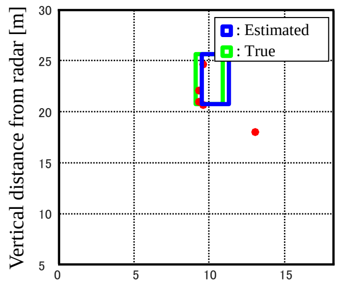

Fig. 6. Parameter tuning results with changing λ.

### B. Evaluation Experiment

We tried to additionally implement two methods to confirm the performance of the proposed method. The first one is the least-squares method to minimize (6), which is referred to as “LSM”. The second one is the LSM using the projection shown in (9), which is referred to as “Projection”. Now the proposed method, which minimizes (13), is referred to as “Proposed”. The experimental conditions and evaluation metric are the same as those in the parameter tuning case.

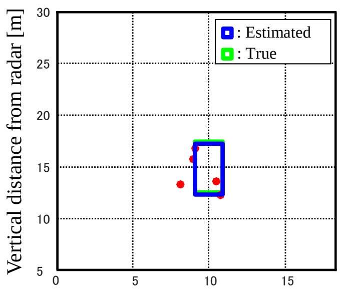
TABLE I: Estimation Results for Each Method

| Method | Estimation error [m] | | |
|---|---|---|---|
| | Long distance | Middle distance | Short distance |
| LSM | 2.31 | 0.72 | 0.41 |
| Projection | 1.10 | 0.33 | 0.27 |
| Proposed | 0.40 | 0.25 | 0.23 |




Horizontal distance from radar [m]

Fig. 7. Position estimation in long distance case (blue box: estimated position, green box: true position).




Horizontal distance from radar [m]

Fig. 8. Position estimation in medium distance case (blue box: estimated position, green box: true position).

The results are summarized in Table I. Experiments show that the proposed method achieves improvements compared with "LSM" and "Projection". Especially for the long distance case, where the radar points are insufficient and affected by outliers, the proposed method achieves large improvements compared with the two methods. Fig. 7-Fig. 9 show states of actual estimation results by the proposed method in each group.

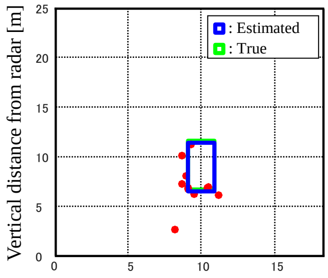



Fig. 9. Position estimation in short distance case (blue box: estimated position, green box: true position).

## VI. Conclusion

In this paper, considering the characteristics of MWR, we have derived a new objective function to accurately estimate the position of the vehicle. Simulation experiments have shown that accurate estimation of the position can be obtained. In future, we will aim at further estimation improvement by developing a method to estimate the size of the object.

## References

[1] Y. Akiyoshi and I. Matsunami, "Vehicle 3D shape estimation using MUSIC algorithm for millimeter wave 2D-MIMO radar," in *Proc. International Workshop on Nonlinear Circuits, Communications and Signal Processing*, March 2018.

[2] S. Ogaki, K. Jimi, and I. Matsunami, "High spatial resolution SFM-MIMO radar using Khatri-Rao product extension processing," in *Proc. International Workshop on Nonlinear Circuits, Communications and Signal Processing*, March 2018.

[3] H. Yamada, T. Kobayashi, Y. Yamaguchi, and Y. Sugiyama, "High-resolution 2D SAR imaging by the millimeter-wave automobile radar," in *Proc. IEEE Conference on Antenna Measurements & Applications*, Dec. 2017, pp. 149-150.

[4] Y. Zeng, Y. Hu, S. Liu, J. Ye, Y. Han, X. Li, and N. Sun, "RT3D: Real-time 3D vehicle detection in LiDAR point cloud for autonomous driving," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3434-3440, Oct. 2018.

[5] M. Bühren and B. Yang, "Automotive Radar Target List Simulation based on Reflection Center Representation of Objects," in *Proc. Workshop on Intelligent Transportation*, March 2006.

[6] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, July 2017, pp. 77-85.

[7] Y. Zhou and O. Tuzel, "Voxelnet: End-to-end learning for point cloud based 3D object detection", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, June 2018, pp. 4490-4499.

[8] M. Alexa, J. Behr, D. Cohen-Or, S. Fleishman, D. Levin, and C. T. Silva, "Point set surfaces," in *Proceedings Visualization, 2001. VIS '01*, Oct. 2001, pp. 21-29.

[9] F. Qian, W. Kan, C. Lailiang, and A. Jianfeng, "The application of point cloud data plane fitting in the Guishan Han Tomb modeling," in *Proc. International Conference on Multimedia Technology*, July 2011, pp. 1781-1784.

[10] T. Funatomi, I. Moro, S. Mizuta, and M. Minoh, "Surface reconstruction from point cloud of human body by clustering," *Systems and Computers in Japan*, vol. 37, pp. 44-56, Oct. 2006.

[11] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239-256, Feb. 1992.

[12] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP." *Proc. of Robotics: Science and Systems*, vol. 2, no. 4, p. 435, June 2009.

[13] H. Q. Luong, M. Vlaminck, W. Goeman, and W. Philips, "Consistent ICP for the registration of sparse and inhomogeneous point clouds," in *Proc. of IEEE Sixth International Conference on Communications and Electronics*, pp. 262-267, July 2016.

**Yoshihiro Suzuki** was born in Tochigi, Japan, in 1994. He received the B.E. degree from Saitama University, Saitama, Japan, in 2017. He researched speech information processing. He is now with Graduate School of Science and Engineering, Saitama University, Japan. He belongs to the Shimamura Laboratory and researches signal processing of Millimeter-Wave Radar.

**Yosuke Sugiura** was born in Hiroshima, Japan, in 1986. He received his B.E., M.E., Ph.D. degrees from Osaka University, Japan, in 2009, 2011, and 2013, respectively. In 2013, he joined Tokyo University of Science, Japan. Since 2015, he has been with the Graduate School of Science and Engineering, Saitama University, Japan, where he is currently an Assistant Professor. His research interests include digital signal processing, adaptive filter theory, and speech information processing.

**Tetsuya Shimamura** received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1986, 1988, and 1991, respectively. In 1991, he joined Saitama University, Saitama, Japan, where he is currently a Professor. His interests are in digital signal processing and its applications to speech, image, and communication systems. He is a member of IEEE, IEICE, and IEEJ.

**Osamu Isaji** was born in Shiga, Japan, in 1963. He received his B.E. from Kyoto Institute of Technology, Japan in 1987. He joined FUJITSU TEN Limited in 1987. Since 2017, he has been with DENSO TEN Limited, where he is engaged in research of the Millimeter-Wave Radar. His research interests include adaptive array antenna, digital signal processing, and Millimeter-Wave devices.

**Kazuaki Hamada** was born in Tokyo, Japan, in 1976. He received his B.E. and M.E. from University of Electro-communications, Japan. He joined Fujitsu Limited in 2001. Since 2017, he has been with DENSO TEN Limited, where he is currently a manager. His research interests include signal processing of Millimeter-Wave Radar.

**Kazuhiko Shite** was born in Oita, Japan, in 1970. He received his B.E., M.E. degrees from Kyushu Institute of Technology, Japan, in 1992 and 1994, respectively. He joined Fujitsu Limited. in 1998, and has been with DENSO TEN Limited since 2017, where he is engaged in research of the Millimeter-Wave Radar. His research interests include signal processing of Millimeter-Wave Radar.