High-resolution Imaging and Separation of Multiple Pedestrians using UWB Doppler Radar Interferometry with Adaptive Beamforming Technique

Motoshi Anabuki[1], Shigeaki Okumura[1], Takuya Sakamoto[1,2], Kenshi Saho[3], Toru Sato[1],
Mototaka Yoshioka[4], Kenichi Inoue[4], Takeshi Fukuda[4], and Hiroyuki Sakai[4]
[1]Graduate School of Informatics, Kyoto University, Kyoto, Japan, manabuki@sato-lab.0t0.jp.
[2]Graduate School of Engineering, University of Hyogo, Hyogo, Japan.
[3]College of Science and Engineering, Ritsumeikan University, Shiga, Japan
[4]Advanced Research Division, Panasonic Corporation, Kyoto, Japan

***Abstract*—Ultra-wideband (UWB) radar imaging has attracted attention for use in security and intelligent transportation system (ITS) applications. Conventional UWB Doppler interferometry is an effective way to obtain high-resolution images while using a simple radar system. However, this method produces ghost images when multiple closely-spaced human targets are present. To resolve this problem, we propose a new technique that combines UWB Doppler interferometry with an adaptive beamforming method called estimation of signal parameters via rotational invariance techniques (ESPRIT). We also propose a tracking and separation algorithm that uses the k-nearest neighbor method. Through a combination of numerical simulations and measurements, we demonstrate the remarkable performance improvement that can be achieved using our proposed method. The proposed method can separate multiple humans with a root-mean-square error of 5.2 cm, which makes its accuracy 1.9 times higher than that of the conventional method.**

***Index Terms*—UWB radar, high-resolution imaging, UWB Doppler radar interferometry, adaptive array processing, k-nearest neighbor.**

I. Introduction

Human body imaging using UWB radar is an effective technique for use in both security and intelligent transportation system (ITS) applications. Optical cameras are commonly used in these applications, but these conventional systems cannot monitor pedestrians when the view is blocked by heavy rain, dense fog, smoke or atmospheric dust particles. In contrast, UWB radar can work effectively in these low visibility situations.

For these applications, a simple but high-accuracy system is required. Recently, high-resolution imaging systems that use simple Doppler radar arrays have been proposed [1], [2], [3]. These systems successfully obtained high-resolution three-dimensional images of single pedestrians in experiments. However, when more than one human target is present in the same range bin simultaneously, the imaging accuracy is seriously reduced.

To resolve this problem, our group have proposed a high-accuracy imaging technique that uses radar antenna arrays and an adaptive beamforming technique [4]. The technique separates the targets that are present in the same range bin based not only on their Doppler velocities but also on their directions of arrival (DOAs).

To generate a three-dimensional image of a specific target, the DOA of and distance to each of the scattering points is required. Adaptive beamforming techniques such as the Capon method [8] and MUSIC [9], however, require peak searching in the horizontal and vertical angles to estimate the DOAs of the scattering points. Two-dimensional peak searching requires high computational complexity.

In this paper, to reduce the computational complexity, we combine UWB Doppler interferometry with the unitary estimation of signal parameters via rotational invariance techniques (ESPRIT) method [10], [11]. The ESPRIT method is an adaptive beamforming technique that estimates the DOA numerically without peak searching. The proposed method requires only a small number of array elements and offers low computational complexity.

First, we apply UWB Doppler radar interferometry and separate the targets based on their Doppler velocities and their range bins. Second, we use the ESPRIT method to separate the targets depending on their DOAs.

Additionally, automatic identification of multiple pedestrians is important for security applications. In this paper, we propose a new identification method that uses the k-nearest neighbor method, which is a clustering technique. The previous step of UWB Doppler radar interferometry combined with the ESPRIT technique estimates the distance, DOA, and intensity of each target. We therefore modify the k-nearest neighbor method and track the targets using these three estimated quantities.

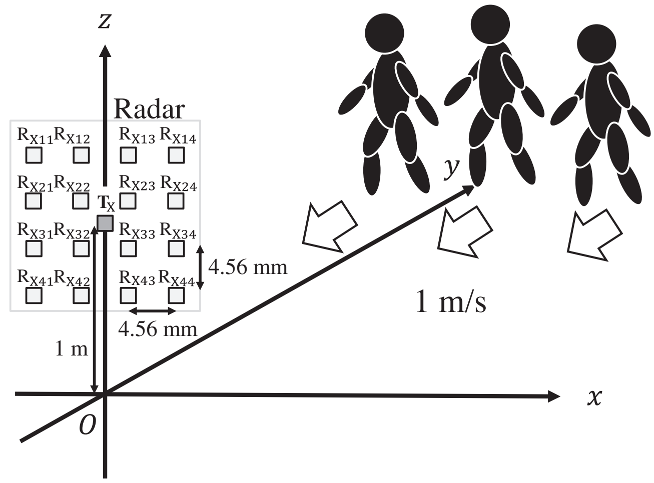


Fig. 1. System model.

II. System Model

We assume the use of an antenna array and three pedestrian targets, as shown in Fig. 1. The pedestrians are walking side by side towards radar antennas at approximately 1 m/s. The pedestrians are walking close to each other and the lateral distance between them is approximately 10 cm. The pedestrian models are formed from actual motion capture data taken from the tracking of markers that were attached to a participant's body.

The signal transmitted by the radar system is a UWB signal with a center frequency $f_0 = 60.5$ GHz. The wavelength λ at the center frequency is 4.96 mm. The UWB signal is assumed to be a Gaussian-shaped pulse with a 10 dB bandwidth of 1.25 GHz, which corresponds to downrange resolution of $\Delta r = c/2W = 12.0$ cm, where c is the speed of light. One transmitting antenna, $\mathrm{T_X}$, and 16 receiving antennas, $\mathrm{R_{X11}}$ to $\mathrm{R_{X44}}$ are set up on the $x-z$ plane. The spacing between the receiving antennas d is 4.56 mm. $\mathrm{T_X}$ is located at the central position $(0, 0, z_c)$, where $z_c = 1$ m. The pulse repetition interval ΔT is 0.457 ms.

III. Materials and Methods

A. UWB Doppler Radar Interferometry

The UWB Doppler radar imaging algorithm can separate the multiple scattering points in the Doppler velocity and range domain [5], [6], [7]. In this algorithm, the time-frequency distribution $S_i(t, v)$ is obtained using a short-time Fourier transform (STFT) of the signal received at the i-th antenna $s_i(t)$. $S_i(t, v)$ is calculated using:

$$S_i(t, v) = \int s_i(\tau) h(\tau - t) \mathrm{e}^{-j2kv\tau} d\tau, \quad (1)$$

where $h(t)$ is the Hamming window function, and k is the wavenumber of the center frequency. The significant peaks in $S_i(t, v)$ represent the target Doppler velocities and the range. Echoes from targets with the same Doppler velocity and the same range bin show the same time-frequency distributions. The current method cannot separate the targets in such a situation.

B. DOA Estimation Using 2D-Unitary ESPRIT Method

To separate the targets that have similar Doppler velocities and that are in the same range bin, we use the difference in the DOA of each target. We apply the 2D-unitary ESPRIT method [10], [11] to the signal after the STFT. This method is a high-resolution DOA and elevation-azimuth angle estimation technique. The ESPRIT method requires two adjacent antenna sub-arrays with the same form. In this study, we assume the use of two types of sub-arrays, which are composed of 4×3 and 3×4 elements, respectively.

First, we transform the received signal vector $\boldsymbol{S}_m(t, v)$ into the real-valued vector $\boldsymbol{Y}_m(t, v)$ using the unitary transform method as follows:

$$\boldsymbol{X}_m(t, v) = U^{\mathrm{H}} \boldsymbol{S}_m(t, v), \quad (2)$$

$$\begin{aligned} \boldsymbol{S}_1(t, v) &= [S_{11}(t, v_n)\ S_{12}(t, v) \cdots \\ &\quad S_{43}(t, v_n)\ S_{44}(t, v)]^{\mathrm{T}}, \quad (3) \\ \boldsymbol{S}_2(t, v) &= [S_{11}(t, v_n)\ S_{21}(t, v) \cdots \\ &\quad S_{34}(t, v_n)\ S_{44}(t, v)]^{\mathrm{T}}, \quad (4) \end{aligned}$$

where n is the target number, and U is the unitary matrix. The covariance matrix, R_l, is then given by:

$$R_m = \mathrm{E}[\boldsymbol{X}_m(t, v_n) \boldsymbol{X}_m(t, v_n)^{\mathrm{H}}] \quad (5)$$

Next, we solve the problem of the ESPRIT method as follows:

$$J_1 \boldsymbol{A}_m \Phi_m = J_2 \boldsymbol{A}_m, \quad (6)$$

$$\Phi_m = diag[\exp(-j\Phi_m(1)) \cdots \exp(-j\Phi_m(N))], \quad (7)$$

$$\Phi_1(n) = kd \cos\phi_n \sin\theta_n, \quad (8)$$

$$\Phi_2(n) = kd \sin\phi_n \sin\theta_n, \quad (9)$$

where J_1 and J_2 are the selection matrices that are used to select the sub-arrays, $\boldsymbol{A}_m$ is the steering vector, N is the number of echo signals present in the received signal, and (ϕ_n, θ_n) is the DOA of the n-th target. The 2D-unitary ESPRIT method estimates (ϕ_n, θ_n) using the paired eigenvalues of R_1 and R_2. Note that the ESPRIT technique estimates both the target DOA and the target intensity.

The outputs of the Doppler interferometry with ESPRIT method are $(x_n(t), y_n(t))$ and $P_{\mathrm{out}}(x_n(t), y_n(t))$, where $(x_n(t), y_n(t))$ represents the estimated position of the n-th target, and P_{out} is the estimated intensity.

C. Separation and Tracking of Multiple Targets

The received signal includes echoes from multiple parts of each body. The scattering points are largely distributed around the torso [5]. Therefore, we estimate the torso position first to separate and then track the targets. We assume that the positions of the scattering points of the targets represent their torso positions. We determine these positions using the estimated power for each target.

We smooth the echo signal power data on the $x-y$ plane using a 2D Gaussian filter [12]. The signal power in the $t-x-y$ space, $P(t,x,y)$, is given by:

$$P(t,x,y) = \sum_n [P_{\mathrm{out}}(x_n(t), y_n(t)) \\ (G(x,y) * \delta(x - x_n(t), y - y_n(t)))], \quad (10)$$

where $G(x,y)$ is the Gaussian distribution in the $x-y$ plane, δ is the Dirac delta function, and $*$ denotes a convolution. After the smoothing process, we then find the significant peak positions $\boldsymbol{p}_u(t)$ of the signal power $P(t,x,y)$.

Next, we track the targets. First, we estimate the candidate moving distances using the following equation:

$$\boldsymbol{\epsilon}_{uw}(t) = \boldsymbol{p}_u(t+\Delta T) - \boldsymbol{p}_w(t), \quad (11)$$

where u and w are the target numbers. Because we assume that the targets are human, the moving distance during the pulse repetition time should be short. We regard the target at position $\boldsymbol{p}_u(t+\Delta T)$ as being the same target that is at position $\boldsymbol{p}_w(t)$ if $\boldsymbol{p}_u(t+\Delta T)$ satisfies the following:

$$\frac{|\boldsymbol{\epsilon}_{uw}(t)|}{\Delta T} < v_{\max}, \quad (12)$$

$$|\boldsymbol{\epsilon}_{uw}(t)| = \min_h |\boldsymbol{\epsilon}_{uh}(t)|, \quad (13)$$

where $v_{\max}$ is the assumed maximum target speed.

D. Clustering Using the k-Nearest Neighbor Method

In this step, we classify the scattering points that arise from the same targets into the same group using the tracked torso positions. To cluster the scattering points, we apply the 2-NN method [13] on the $x-y$ plane. First, we acquire the training data. We consider the estimated scattering center at position $\boldsymbol{q}(t)$ as the training data if $\boldsymbol{q}(t)$ satisfies the following:

$$|\boldsymbol{p}(t) - \boldsymbol{q}(t)| < \eta, \quad (14)$$

where η is the radius that is centered at $\boldsymbol{p}(t)$.

Next, we separate and classify the scattering points using the 2-NN method. The k-nearest neighbor (k-NN) method is a simple separation method that is based on the closest training samples located in the feature space. In this method, we calculate the distance D_{ab} between the a-th unlabeled point at $\boldsymbol{r}_a(t)$ and the b-th training data point at $\boldsymbol{q}_b(t)$ as follows:

$$D_{ab} = |\boldsymbol{r}_a(t) - \boldsymbol{q}_b(t)| \quad (15)$$

D_{ab} is calculated for all training data. We then classify $\boldsymbol{z}_a(t)$ with respect to the label of the nearest training data point.

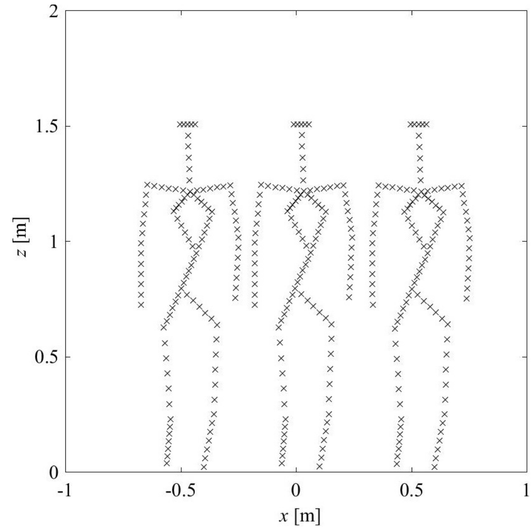


Fig. 2. Scattering points of pedestrians used in simulations.

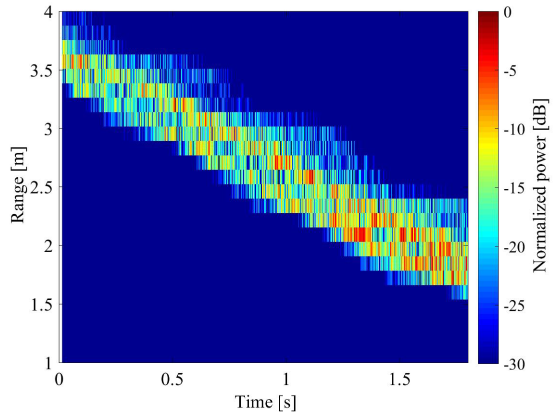


Fig. 3. Received signal calculated using ray-tracing methods.

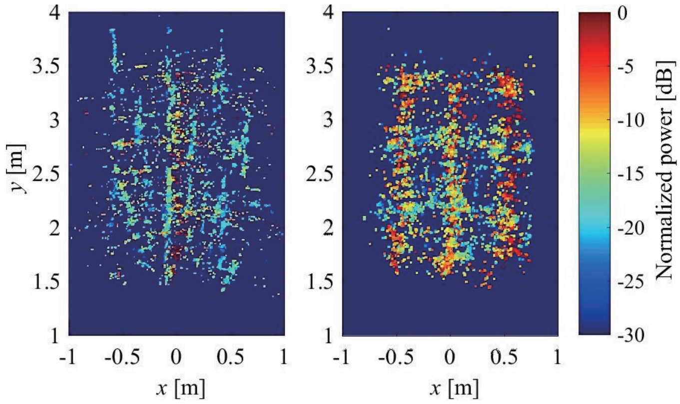
IV. Simulation Results

We used numerical simulations to examine the performance of the proposed algorithm. In the simulations, we assumed that three pedestrians were walking side by side at approximately 1 m/s, as shown in Fig 1. Fig 2 shows the scattering points of the pedestrians. Fig 3 shows the received signals. These received signals are calculated using ray-tracing methods and the scattering points that were taken from the action data of the pedestrians, which were acquired by motion capture.

For the conventional technique, we use the UWB interferometry technique with three receiving antennas: $\mathrm{R}_{\mathrm{X}22}$, $\mathrm{R}_{\mathrm{X}32}$, and $\mathrm{R}_{\mathrm{X}33}$[5], [6], [7].

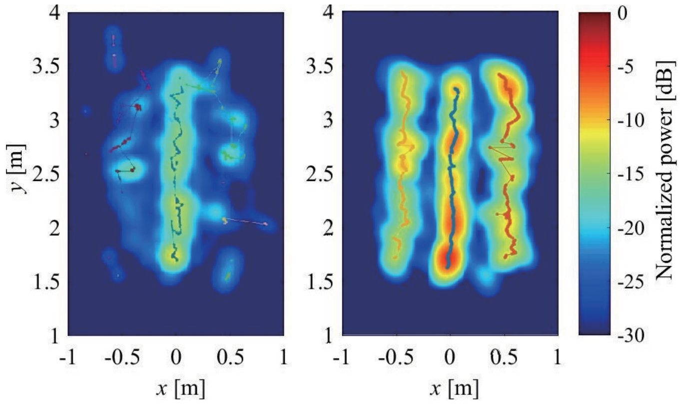
To evaluate the imaging accuracy, we used the root-mean-square error (RMSE) between the actual scattering points and the estimated scattering points.

Fig. 4 shows the top view of the scattering points that are estimated using both the conventional method and the proposed method. Fig. 5 shows smoothed versions of the images from Fig. 4 and the peaks in their power. The colors



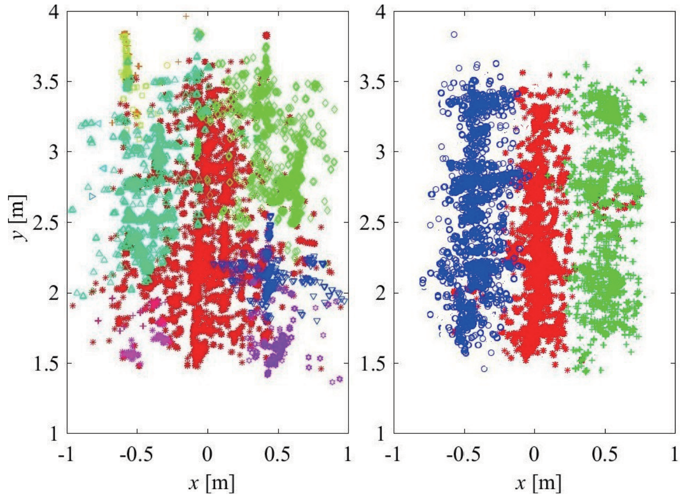
(a) Conventional method (a) Proposed method

Fig. 4. Top view of the estimated images.



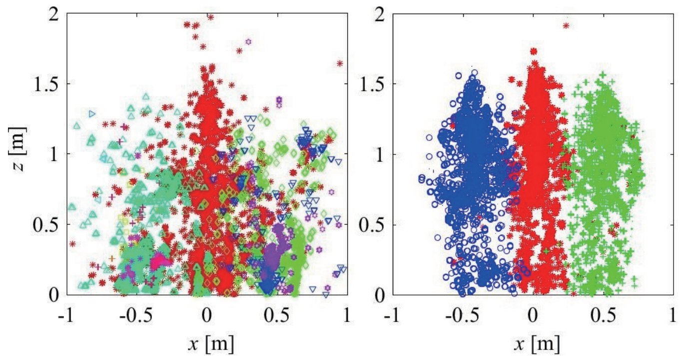
(a) Conventional method (a) Proposed method

Fig. 5. Top view of smoothed power data taken from the estimated image shown in Fig. 4 and the peaks of the power.



(a) Conventional method (a) Proposed method

Fig. 6. Top view of images separated using the 2-NN method.



(a) Conventional method (a) Proposed method

Fig. 7. Frontal view of the estimated image after separation.

of these peak points show the labels that are attached by the tracking algorithm.

Figs. 6 and 7 show the top view and the front view of the estimated points after application of the 2-NN method, respectively. The colors of the peak points show the labels that were attached by the 2-NN method.

As shown in Figs. 5, 6, and 7, the conventional method failed to identify and track the pedestrians. The RMSE for this method was 10 cm, and the number of estimated scattering points was 16141.

In contrast, the proposed method successfully identified the three pedestrians. The RMSE was 5.2 cm, and the number of estimated scattering points was 25436. When compared with the image that was acquired using the conventional method, the imaging accuracy when using the proposed method showed an improvement of 1.9 times.

V. Conclusion

We have proposed a new UWB radar imaging and separation algorithm that combines the UWB Doppler radar interferometry technique, the 2D-unitary ESPRIT method, and the k-NN method. We evaluated the imaging accuracy of the proposed method using numerical simulations. The proposed method created high-resolution images with a typical RMSE of 5.2 cm, and accurately separated three pedestrians, whereas the RMSE of the conventional method was 10 cm. These results demonstrate that the proposed method is effective for the imaging of multiple pedestrians using UWB radar.

Acknowledgment

This research was partially supported by the Supporting Program for Interaction-based Initiative Team Studies (SPIRITS) as part of the Japan-Netherlands joint development of sleep monitoring technology using ultra-wideband radar; by the Center of Innovation (COI) Program The Last 5X Innovation R&D Center for a Smart, Happy, and Resilient Society; by JSPS KAKENHI under Grant Nos. 25249057, 15K18077, and 15J05687; and by the R&D project for the expansion of radio spectrum resources or more efficient use of frequency resources for the future, which is supported by The Ministry of Internal Affairs and Communications, Japan.

REFERENCES

[1] A. Lin and H. Ling, "Doppler and direction-of-arrival (DDOA) radar for multiple-mover sensing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1496–1509, 2007.

[2] A. Lin and H. Ling, "Frontal imaging of human using three-element Doppler and direction-of-arrival radar," *Electron. Lett.*, vol. 42, no. 11, pp. 660–661, 2006.

[3] A. Lin and H. Ling, "Three-dimensional tracking of humans using very low-complexity radar," *Electron. Lett.*, vol. 42, no. 18, pp. 1062–1063, 2006.

[4] M. Anabuki, S. Okumura, T. Sakamoto, K. Saho, T. Sato, M. Yoshioka, K. Inoue, T. Fukuda, and H. Sakai, "High resolution imaging and identification of multiple pedestrians using UWB Doppler radar interferometry and adaptive array processing," in *ISAP2016*, accepted.

[5] K. Saho, T. Sakamoto, T. Sato, K. Inoue, and T. Fukuda, "Pedestrian imaging using UWB Doppler radar interferometry," *IEICE Trans. Commun.*, vol. E96–B, no. 2, pp. 613–623, 2013.

[6] K. Saho, T. Sakamoto, T. Sato, K. Inoue, and T. Fukuda, "Accurate and real-time pedestrian classification based on UWB Doppler radar images and their radial velocity features," *IEICE Trans. Commun.*, vol. E96–B, no. 10, pp. 2563–2572, 2013.

[7] K. Saho, H. Homma, T. Sakamoto, T. Sato, K. Inoue, and T. Fukuda, "Accurate image separation method for two closely spaced pedestrians using UWB Doppler imaging radar and supervised learning," *IEICE Trans. Commun.*, vol. E97–B, no. 6, pp. 1223–1233, 2014.

[8] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[9] R.O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas. Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[10] M. Haardt and J. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1232–1242, 1995.

[11] M.D. Zoltowski, M. Haardt, and C.P. Mathews, "Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 316–28, 1996.

[12] R.A. Haddad and A.N. Akansu, "A class of fast Gaussian binomial filters for speech and image processing," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 723–727, 1991.

[13] T.M. Cover and P.E. Hart, "Nearest neighbor pattern classification", *Proc. IEEE Trans. Inform. Theory*, vol. 13, no. 1, pp. 21–27, 1967.