PAPER

Fundamental Trial on DOA Estimation with Deep Learning*

Yuya KASE[†a)], ***Student Member***, **Toshihiko NISHIMURA**[†], ***Senior Member***, **Takeo OHGANE**[†b)], **Yasutaka OGAWA**[†], ***Fellows***, **Daisuke KITAYAMA**[††], ***Member***, ***and*** **Yoshihisa KISHIYAMA**[††], ***Senior Member***

SUMMARY Direction of arrival (DOA) estimation of wireless signals has a long history but is still being investigated to improve the estimation accuracy. Non-linear algorithms such as compressed sensing are now applied to DOA estimation and achieve very high performance. If the large computational loads of compressed sensing algorithms are acceptable, it may be possible to apply a deep neural network (DNN) to DOA estimation. In this paper, we verify on-grid DOA estimation capability of the DNN under a simple estimation situation and discuss the effect of training data on DNN design. Simulations show that SNR of the training data strongly affects the performance and that the random SNR data is suitable for configuring the general-purpose DNN. The obtained DNN provides reasonably high performance, and it is shown that the DNN trained using the training data restricted to close DOA situations provides very high performance for the close DOA cases.

key words: *DOA estimation, deep learning, machine learning*

1. Introduction

Direction of arrival (DOA) estimation of radio signal sources and/or scatters is frequently required not only for radar applications but also for user localization which is used in services for augmented reality, pushing sales information, etc. DOA estimation has a long history starting from a beamformer (matched filter), and now many algorithms including sensor-network-based one are available [2]. Among them, traditional high-resolution algorithms such as a Capon technique [3], a multiple signal classification (MUSIC) algorithm, and estimation of signal parameters via rotation invariance (ESPRIT) are widely known [4].

In recent years, methods based on compressed sensing have also been proposed [5]. Compressed sensing is a technique mainly used in the fields of signal processing and data compression. The technique reconstructs an original signal, which is given by a sparse vector, from fewer observations than the dimension of the signal vector [6]. In an example of DOA estimation using a compressed sensing solver called half-quadratic regularization (HQR), higher accuracy has been reported compared with MUSIC [7].

Nonlinear algorithms such as compressed sensing require a heavy computational load in general. Actually, HQR needs an iterative calculation of an inverse matrix of which dimension is the same as that of the original signal vector. Along with the development of computer technology, however, very heavy computation becomes feasible now and thus deep learning is attracting attention rapidly.

Deep learning is a machine learning technique using multiple layers of neural network, i.e., a deep neural network (DNN). The technique has been extensively studied in the fields of images, sounds, languages, etc., starting with the research by Hinton et al. [8], [9]. During the learning process, the DNN requires a large amount of calculations. However, estimation using the learned network can be performed with simple calculations which mainly consists of multiplication of a matrix and a vector. This computational load in estimation phase is lighter than the ones of MUSIC (eigenvalue decomposition) and HQR (iterative calculation of an inverse matrix). In addition, the fact that we can design the DNN for specific scenarios where estimation is difficult with conventional algorithms is unique to machine-learning-based estimation. Thus, it can be said that using the DNN for DOA estimation is an attractive option when the heavy computational load of the learning process is acceptable.

Although DOA estimation using deep learning was first applied in the speech source localization field as in [10], several papers on the radio resource localization can be seen recently [1], [11], [12]. In [11], the DOA estimation results are used for massive MIMO channel estimation, and estimation performance related to the number of hidden layers was precisely measured. The DNN proposed in [12] is unique. It consists of two parts, a spatial filtering part for subregion decomposition and a spatial spectrum estimation part, and succeeds in obtaining accurate spectra. These papers discuss application-oriented performance, and the DNN configurations are not general in terms of the array size [11] or the use of subregion decomposition.

In this paper, we apply a simple DNN to DOA estimation and discuss the estimation performance under a scenario where two equal-power and uncorrelated signals are incident on a uniform linear array in order to check the fundamental applicability of deep learning to the DOA estimation. In other fields, it is known that the performance of DNN changes

Manuscript received December 24, 2019.
Manuscript publicized April 21, 2020.

[†]The authors are with the Graduate School of Information Science and Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.

[††]The authors are with NTT DOCOMO, INC., Yokosuka-shi, 239-8536 Japan.

*A part of this paper was presented at IEEE Workshop on Positioning, Navigation and Communications (WPNC) 2018 [1].

a) E-mail: kase@m-icl.ist.hokudai.ac.jp
b) E-mail: ohgane@ist.hokudai.ac.jp

DOI: 10.1587/transcom.2019EBP3260

depending on the type of training data, so the same can be expected with DNN in DOA estimation. Therefore, we decided to investigate the relationship between SNR and DOA performance. In addition to discussion on types of training data, we consider designing a DNN for a specific scenario because DNNs are able to be specialized for various purposes. In general, it is a difficult problem to estimate DOAs of closely incident signals accurately. Therefore, we design a DNN using training data sets assuming two waves with close DOAs, evaluate the estimation accuracy, and propose an a parallel use of multiple DNNs designed for different scenarios. In the rest of paper, the array structure and signal arrival model, the configuration and training process of DNN, and performance evaluation by computer simulation are described.

2. Preparation for DOA Estimation

DOA is estimated using received signal information. In general, the information is categorized into two types: raw waveforms and statistical data when multiple observations are allowed. MUSIC and ESPRIT use a correlation matrix of the received signal since those DOA estimation techniques are based on the eigen structure of the correlation matrix. Therefore, DOA can be estimated without waveform information such as pilot signals. In this paper, we use the correlation matrix for the DNN input as in [12].

Although estimation of two equal-power waves is discussed later, we formulate the following equations under a general case where K plane waves with a wavelength λ and a complex amplitude $s_k(t)$ are incident at an angle θ_k and a time t on an L-element uniform linear array with an element spacing d antenna as shown in Fig. 1. The received signal at the lth antenna is expressed as

$$x_l(t) = \sum_{k=1}^{K} s_k(t) e^{-j\frac{2\pi}{\lambda}(l-1)d\sin\theta_k} + n_l(t), \tag{1}$$

where $n_l(t)$ is an additive white noise at the lth antenna.

The received signals at all antennas can be expressed in a vector form as

$$\begin{aligned}\boldsymbol{x}(t) &= [x_1(t),\, x_2(t), \ldots,\, x_L(t)]^T \\ &= \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t),\end{aligned} \tag{2}$$

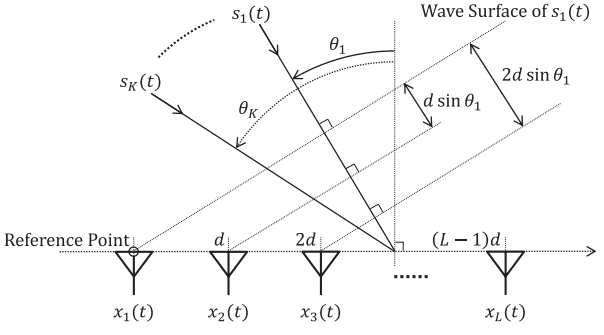


Fig. 1 An L-element uniform linear array antenna and incident waves.

where $[\cdot]^T$ denotes the transpose and

$$\boldsymbol{s}(t) = [s_1(t),\, s_2(t), \ldots,\, s_K(t)]^T \tag{3}$$

$$\boldsymbol{n}(t) = [n_1(t),\, n_2(t), \ldots,\, n_L(t)]^T \tag{4}$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j\frac{2\pi}{\lambda}d\sin\theta_1} & \cdots & e^{-j\frac{2\pi}{\lambda}d\sin\theta_K} \\ \vdots & \ddots & \vdots \\ e^{-j\frac{2\pi}{\lambda}(L-1)d\sin\theta_1} & \cdots & e^{-j\frac{2\pi}{\lambda}(L-1)d\sin\theta_K} \end{bmatrix}. \tag{5}$$

The $L \times K$ matrix $\boldsymbol{A}$ is called a mode matrix.

The $L \times L$ correlation matrix of received signal vector $\boldsymbol{x}(t)$ is expressed as

$$\boldsymbol{R}_{xx} = E\left[\boldsymbol{x}(t)\,\boldsymbol{x}^H(t)\right] = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \boldsymbol{R}_N, \tag{6}$$

where $E[\cdot]$ and $[\cdot]^H$ denote the ensemble average and the conjugate transpose, respectively. $\boldsymbol{S}$ and $\boldsymbol{R}_N$ are the $K \times K$ signal correlation matrix and the $L \times L$ noise correlation one, respectively, and are given by

$$\boldsymbol{S} = E\left[\boldsymbol{s}(t)\,\boldsymbol{s}^H(t)\right] \tag{7}$$

$$\boldsymbol{R}_N = E\left[\boldsymbol{n}(t)\,\boldsymbol{n}^H(t)\right]. \tag{8}$$

We assume that all the noise components are mutually uncorrelated and have the same power σ^2. Thus, we have

$$\boldsymbol{R}_N = \sigma^2\boldsymbol{I}, \tag{9}$$

where $\boldsymbol{I}$ is the L-dimensional identity matrix. Then, we can rewrite (6) as

$$\boldsymbol{R}_{xx} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma^2\boldsymbol{I}. \tag{10}$$

Note that this receive correlation matrix is a Hermitian matrix.

3. DOA Estimation with Deep Learning

3.1 Formulation of DNN

In general, a single layer dense neural network of J units with I inputs and J outputs can be illustrated as Fig. 2. The output of the jth unit z_j can be expressed as

$$z_j = f(u_j) \tag{11}$$

$$u_j = \sum_{i=1}^{I} w_{j,i}\, y_i + b_j, \tag{12}$$

where f, y_i, $w_{j,i}$, and b_j are all real-valued and represent an activation function, the ith input, the weight for the jth unit multiplied by y_i, and a constant bias, respectively. Assuming the 0th input $y_0 = 1$ and the corresponding weight $w_{j,0} = b_j$, we can modify (12) as

$$u_j = \sum_{i=0}^{I} w_{j,i}\, y_i. \tag{13}$$

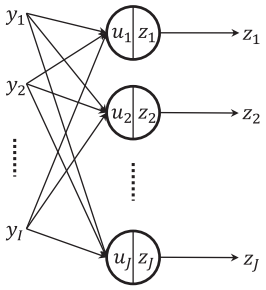


Fig. 2 A factor graph for a single layer neural network.

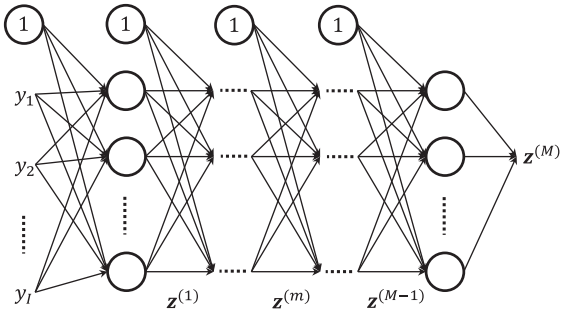


Fig. 3 Factor graph expression of a deep neural network (DNN).

These relationships can be expressed in a vector-matrix form as

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_J \end{bmatrix} = \begin{bmatrix} f(u_1) \\ f(u_2) \\ \vdots \\ f(u_J) \end{bmatrix} \tag{14}$$

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_J \end{bmatrix} = \begin{bmatrix} w_{1,0} & w_{1,1} & \cdots & w_{1,I} \\ w_{2,0} & w_{2,1} & \cdots & w_{j,I} \\ \vdots & \vdots & & \vdots \\ w_{J,0} & w_{J,1} & \cdots & w_{J,I} \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ \vdots \\ y_I \end{bmatrix}, \tag{15}$$

and simplified as

$$\boldsymbol{z} = f(\boldsymbol{u}) \tag{16}$$

$$\boldsymbol{u} = \boldsymbol{W} \begin{bmatrix} 1 \\ \boldsymbol{y} \end{bmatrix}. \tag{17}$$

DNN has a multilayer structure with hidden layers as shown in Fig. 3 where M single layer dense neural networks are stacked. The output of the mth layer can be expressed as

$$\boldsymbol{z}^{(m)} = f^{(m)}\left(\boldsymbol{u}^{(m)}\right), \tag{18}$$

where $[\cdot]^{(m)}$ denotes the mth layer index and

$$\boldsymbol{u}^{(m)} = \boldsymbol{W}^{(m)} \begin{bmatrix} 1 \\ \boldsymbol{z}^{(m-1)} \end{bmatrix}. \tag{19}$$

Note that $\boldsymbol{z}^{(M)} = f^{(M)}\left(\boldsymbol{u}^{(M)}\right)$ and $\boldsymbol{z}^{(0)} = \boldsymbol{y}$ correspond to the DNN output and input, respectively.

3.2 DNN Input

In the paper, we use the $L \times L$ correlation matrix $\boldsymbol{R}_{xx}$ for an input to the DNN as described in Sect. 2. Since $\boldsymbol{R}_{xx}$ is a Hermitian matrix where the off-diagonal elements are the complex conjugates of the diagonally-opposite entries, only the lower triangular part is used for the input of the DNN. Equations (18) and (19) are defined in real space. Therefore, we need to decompose the complex value of each entry of the lower triangular part of $\boldsymbol{R}_{xx}$ into two real values except for the diagonal elements. Then, the input vector to the DNN is written as

$$\begin{aligned} \boldsymbol{y} = [&r_{1,1},\ r_{2,2}, \ldots,\ r_{L,L},\ \Re(r_{2,1}),\ \Im(r_{2,1}), \\ &\Re(r_{3,1}),\ \Im(r_{3,1}),\ \Re(r_{3,2}), \ldots,\ \Im(r_{L,L-1})]^T, \end{aligned} \tag{20}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts, respectively. The dimension of the column vector $\boldsymbol{z}^{(0)} = \boldsymbol{y}$, i.e., the number of input units becomes L^2.

3.3 DNN Output

Definition of the DNN output directly determines the DOA estimation accuracy. Clearly, the DNN provides a discrete output. In that sense, the DOA estimation using deep learning is generally classified as the "on-grid" type similar to the conventional compressed sensing algorithms. In this paper, we apply the on-grid method straightforwardly and thus assign each entry of $\boldsymbol{z}^{(M)}$ to a certain discrete angle from a predefined grid set[†].

The number of output units depends on the size of the predefined grid set, i.e., a required angle resolution and search range. When the angle resolution is $\Delta\theta$ and the search range is between $\theta_{\min}$ and $\theta_{\max}$, the number of output units becomes $(\theta_{\max} - \theta_{\min})/\Delta\theta + 1$. For each output value, we define the probability [%] that a plane wave is incident at the corresponding angle. Thus, in the training phase, each output value is set as

$$z_j^{(M)} = \begin{cases} 100 & \text{when a plane wave is incident} \\ 0 & \text{otherwise} \end{cases}. \tag{21}$$

4. Evaluation of Estimation Accuracy

4.1 Simulation Conditions

We evaluated the DOA estimation performance of the DNN by computer simulations. The followings are parameters used here. To simplify performance analysis, we considered a basic estimation problem where the number of targets was two. In addition, it was assumed that both signals from the

[†]Even in the on-grid case, the off-grid angle can be estimated in some ways. For example, interpolation is used in [12].

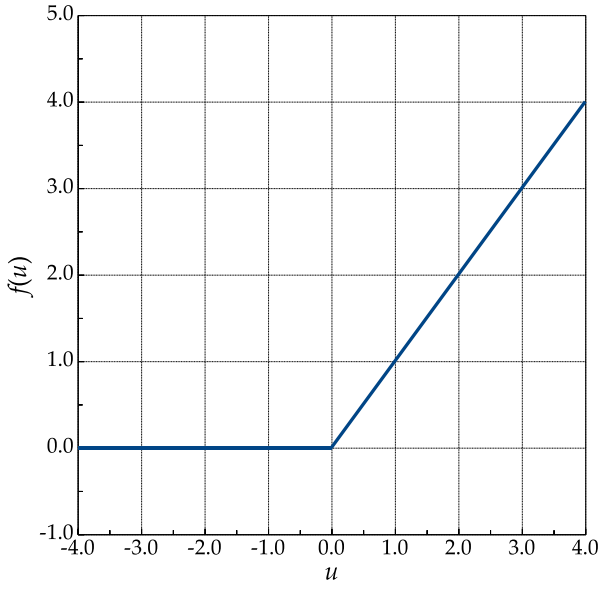


Fig. 4 A ramp function.

targets were narrowband for (1) to hold, had equal power, and were uncorrelated.

The DOA of each plane wave ranged from $-60°$ to $60°$ with a $1°$ step. This integer DOA assumption is not general but makes the performance analysis very clear. Although each integer DOA was set randomly per estimation, we coordinated two DOAs not to be coincident. The number of array elements was five, and the element spacing was half-wavelength of the center frequency of the narrowband signals. We used 100 snapshots of received signals to calculate the correlation matrix in each estimation.

According to the number of antennas, the dimension of input vector $\boldsymbol{z}^{(0)} = \boldsymbol{y}$ connected to the DNN are 25. We assumed that required angle resolution was $1°$ and that the search range of DOA was between $-60°$ and $60°$. Thus, 121 output units compose the DOA grid corresponding to integer angles from $-60°$ to $60°$ which covers the possible DOA range appropriately. In the estimation phase, two signals' DOAs were determined by the angles of two output units having the largest and the second largest probabilities without peak detection.

The number of hidden layers is four where each layer has 150 units. Note that these parameters are not optimized but empirically determined with a limited number of tests. The activation function of all units except for the output layer was set to a ramp function shown in Fig. 4, and we have

$$f(u) = \max(u, 0) = \begin{cases} u & \text{for } u \geq 0 \\ 0 & \text{elsewhere} \end{cases}. \tag{22}$$

For the activation function of the output layer, the identity function is used.

Online learning based on back propagation was applied to the DNN using 3,000,000 training data where each training data sample was generated by setting two different random integer DOAs within the search range and white Gaussian noise[†]. In online learning, the weights of the DNN are updated at every training sample which is input one by one. Even though batch learning is more common, we applied online learning because we thought that it is simple and gives straightforward observation on the training progress. Adaptive moment estimation (Adam) [13] was used to determine the learning rate. Adam determines the learning rate adaptively and automatically from the past gradient. Thus, there is no need to adjust the learning rate manually. In addition, it is known that the DNN learned with Adam provides better performance than DNNs learned with other techniques such as AdaGrad [14] and AdaDelta [15].

During the learning process, the DNN performance is validated every 100 learning cycles based on the estimation success rate using 100,000 validation data where not only incident angles but also SNR are randomly set. The range of SNR in this validation phase was from 0 to 30 dB. The estimation success is defined as the case where both estimated DOAs are equal to the true value. Finally, the DNN providing the highest success rate is selected from the validation as the best one.

In the evaluation phase, the estimation performance is evaluated using three measures: estimation success rate, RMSE, and the absolute value of error. RMSE is expressed as

$$\text{RMSE} = \sqrt{\frac{1}{KN}\sum_{k=1}^{K}\sum_{n=1}^{N}\left(\hat{\theta}_k^{(n)} - \theta_k^{(n)}\right)^2}, \tag{23}$$

where $[\cdot]^{(n)}$ and N denotes the nth test index and the number of tests, respectively. The evaluation is performed using 100,000 test data ($N = 100,000$) at each SNR from 0 dB to 30 dB in 5 dB step where the incident angles are randomly set per test. Thus, the number of total evaluation data becomes 700,000.

4.2 Training Data and Behavior

For the training data, six different sets where the SNR was changed in predetermined patterns were prepared as follows.

(i) 0 dB constant ("0 dB")
(ii) 30 dB constant ("30 dB")
(iii) increased linearly from 0 dB to 30 dB ("increase")
(iv) increased in 5 dB step from 0 dB to 30 dB ("stepwise")
(v) decreased linearly from 30 dB to 0 dB ("decrease")
(vi) random in the range between 0–30 dB ("random")

SNR transition in each set is illustrated in Fig. 5.

[†]For implementation of such a DOA estimator based on the deep learning, preparing a large amount of training data samples is one of the most difficult problems. It might be a reasonable idea that batch training with computer-generated samples is done before the operation and that online training using observation data samples for known targets is done to update the network. It is too difficult for us to discuss system-level-training at this moment. Thus, only the computer-simulation-based training and evaluation are shown in this paper.

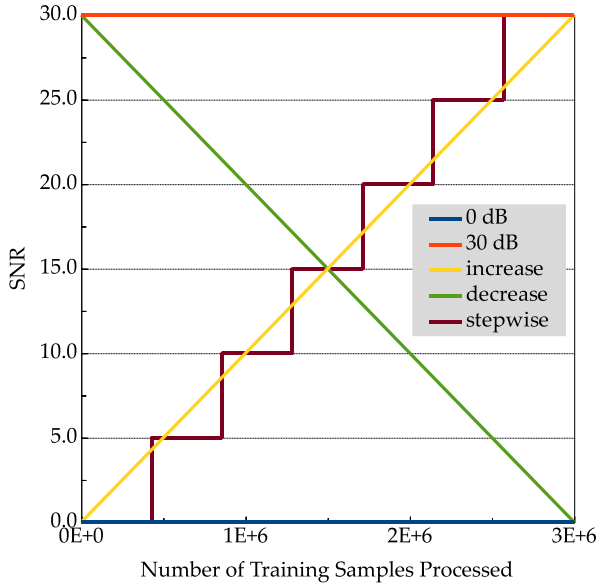


Fig. 5 SNR transition in each training pattern. (The random case is not shown.)

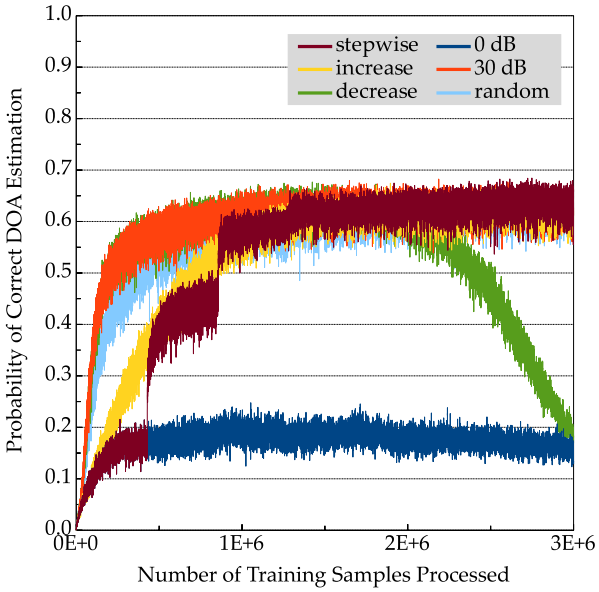


Fig. 6 Learning progress in terms of success rate. In the graph, success rate is plotted at every 100 training data samples.

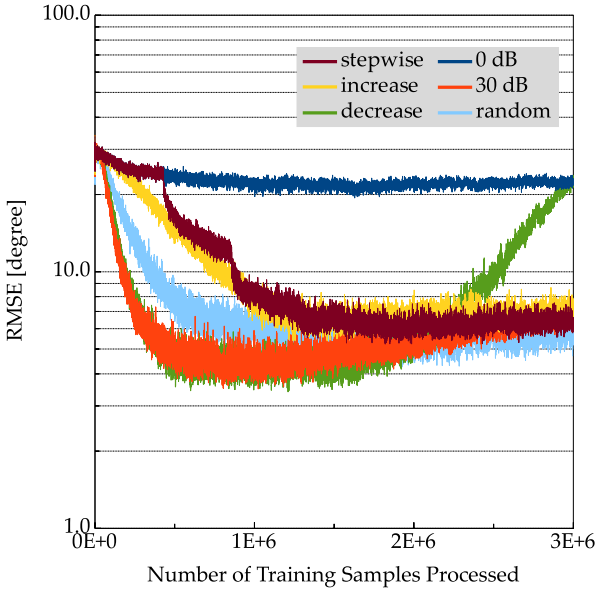


Fig. 7 Learning progress in terms of RMSE. In the graph, RMSE value is plotted at every 100 training data samples.

Estimation success rates observed at every 100 learning cycles are shown in Fig. 6. The number of learning cycles where the success rate of the test was the maximum is 1,052,900 for "0 dB", 2,186,700 for "30 dB", 1,895,200 for "increase", 2,748,600 for "stepwise", 1,424,000 for "decrease", and 2,995,500 for "random."[†] Note that the highest success rates are obtained not at the end of learning except for the "stepwise" and "random" data sets. It implies the possibility of overfitting as discussed later. The training data set "0 dB" could not provide good success rate at any learning cycles. It is supposed that featureless correlation matrices generated by the low SNR data confuse the learning process. Except for the "0 dB" and "decrease" cases, the success rates seem to be improved with the number of leaning cycles although the performance improvement after about 1,500,000 learning cycles is very small. Therefore, it may be difficult to find the best timing to terminate learning.

Figure 7 shows RMSEs observed at every 100 learning cycles. The RMSE using the training data set "30 dB" becomes visibly worse beyond the best learning cycles. This suggests that over fitting occurred. Since such an over fitting phenomenon is not observed in the success rate performance in Fig. 6, it is expected that over fitting makes DOA estimation errors larger when the estimation fails. However, we selected the DNN providing the highest success rate for the later performance evaluation because this paper focuses on success rate performance.

4.3 Estimation Performance vs. Training Data Set

Estimation success rates and RMSEs of DNNs trained according to the SNR patterns (i) to (vi) are shown in Figs. 8 and 9. Note that the abscissa represents SNR for the estimation phase, not the learning. As a reference, the success rate and RMSE of Root MUSIC [16] are also shown. Since Root MUSIC outputs real angles of the peaks, each angle value is rounded to the nearest integer and then regarded as an estimated DOA. Note that the success rate and RMSE in these figures are different form ones in Figs. 6 and 7 because the SNR of validation data in these figures was randomly set from about 0 dB to 30 dB.

In Fig. 8, both the "30 dB" and "stepwise" cases show the highest success rate. Although the "stepwise" case provides slightly better success rate, the RMSE is visibly worse

[†]In the estimation phase, we use the DNNs obtained at these learning cycles.

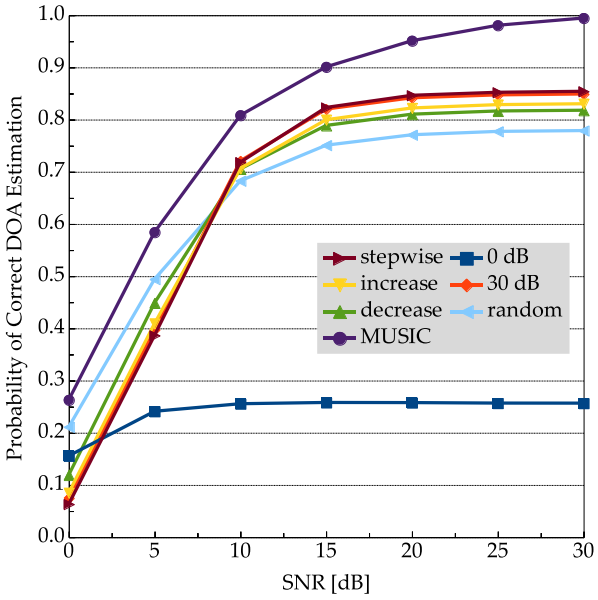


Fig. 8 Success rate performance for each training data set.

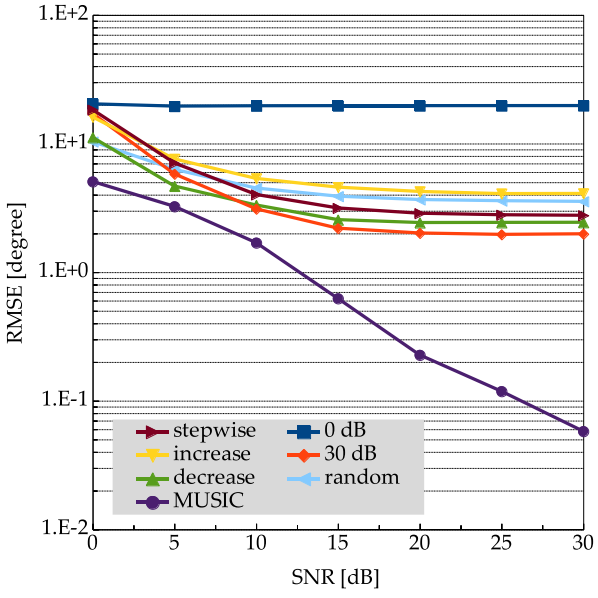


Fig. 9 RMSE for each training data set.

in comparison to the "30 dB" case. Thus, it can be said that the DNN trained using high SNR data ("30 dB") provides higher performance in terms of not only the success rate but also the RMSE in the SNR range from 10 dB to 30 dB. Note that neither success rates nor RMSEs of DNNs reach ones of Root MUSIC. This performance difference clearly indicates that our DNNs are not trained well or the network configuration is not optimized yet. Thus, modifications on training data sets and network configuration are urgent issues for us. However, even in the current DNN configuration, an advantage can be obtained as will be shown later in 4.5.

Although the DNN trained using random SNR data achieves a lower success rate compared to the "30 dB", "stepwise", "increase", and "decrease" cases in the SNR range from about 10 dB to 30 dB, its success rate in the SNR range from 0 dB to about 10 dB is higher than that of the others. It can be said that training using random SNR data produces a DNN suitable for a wide SNR range. In the following discussions, we use the DNN trained using random SNR data (iv).

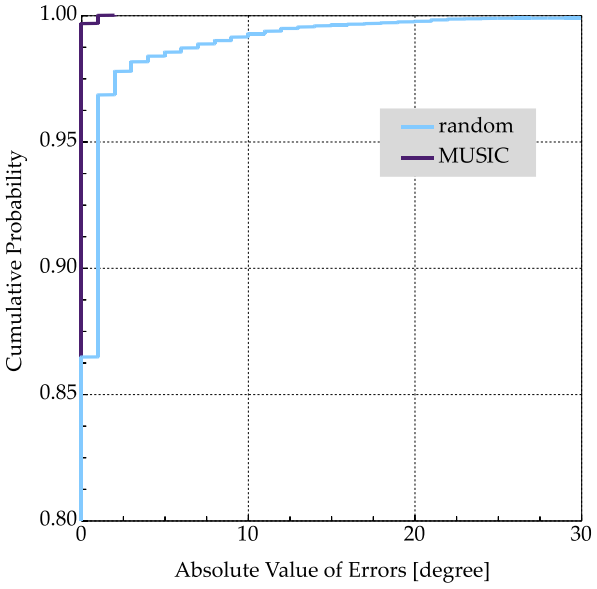


Fig. 10 Cumulative probability of absolute value of errors.

4.4 Absolute Value of Errors

Figure 10 shows the cumulative probability of the absolute value of errors at the SNR of 30 dB for the case using the DNN trained with random SNR data. The one for the Root MUSIC case is also shown as a reference. Let us recall that the estimation status is flagged as "success" only when both DOAs are correctly estimated. Specifically, Fig. 8 shows that "random" DNN failed about 22% of estimations (100,000 in total). However, Fig. 10 indicates that zero error occurs in about 86% of waves (200,000 in total) and that the probability that the error falls within 1° exceeds 97%. Although the probability of success of the DNN is less than that of Root MUSIC (near 100%), it is very attractive that the obtained DNN achieves high accuracy as in Fig. 10 with a reasonably low complexity in the estimation phase as long as we consider an on-grid problem.

4.5 DNN for Close DOA Scenario

In the previous discussion, we found that the estimation performance depended on the training data. For example, the DNN trained using random SNR data works well for the wide range of SNR. This implies that it may be possible for us to design the special purpose DNN. Here, we try to construct a DNN suitable to the close DOA case. For this purpose, we trained the DNN using 300,000 training data which were

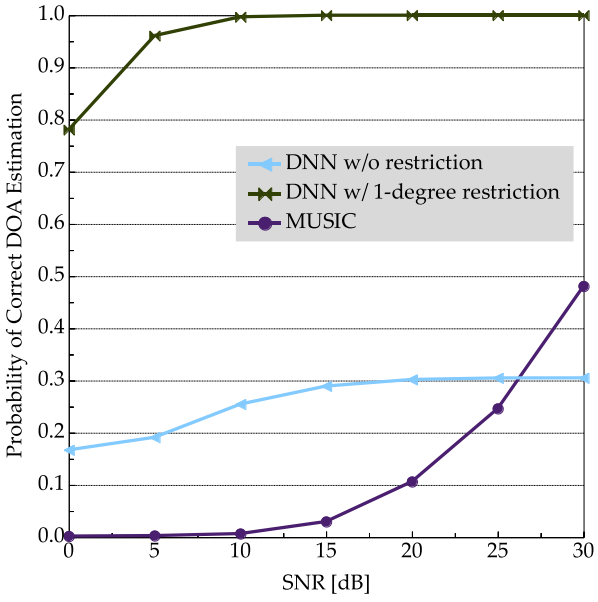


Fig. 11 Success rate performance when $|\theta_1 - \theta_2| = 1°$.

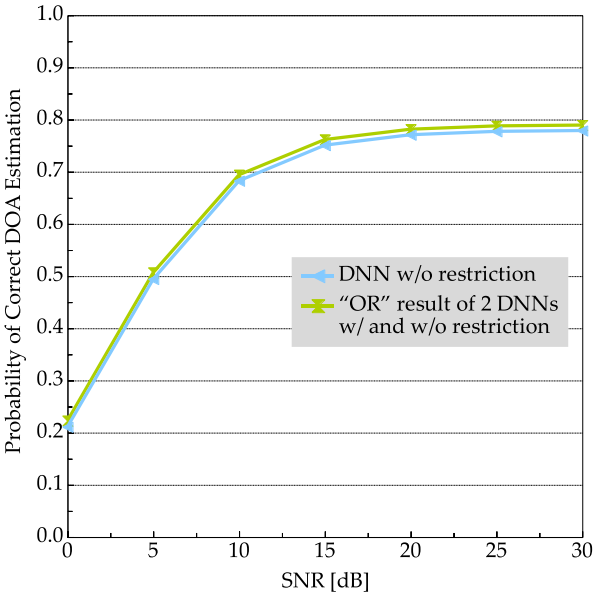


Fig. 12 Success rate performance for parallel use of DNNs.

produced as to be $|\theta_1 - \theta_2| = 1°$. SNR of the training data was randomly set. The other conditions were the same as in Sect. 4.1.

The estimation success rate using this DNN is shown in Fig. 11. Note that the estimation data is restricted to $|\theta_1 - \theta_2| = 1°$ as well as the training phase. As a reference, the success rates of Root MUSIC and the DNN trained using random SNR and DOA data without restriction on DOAs, i.e., the same DNN as in Fig. 10, are also shown. In general, it is difficult to estimate DOAs of which difference is 1°. Actually, the success rate of the DNN trained using random DOA data is lower than about 30%. Furthermore, the success rate of MUSIC is still lower than about 50%. In contrast, the DNN trained using special training data with the restriction of $|\theta_1 - \theta_2| = 1°$ achieves the success rate of 100% at the SNR of 20 dB or higher. The fact that we can design the DNN for specific scenarios is unique to machine-learning-based estimation and shows potential capability of complementary use of such DNNs with other techniques.

4.6 Parallel Use of General-Purpose and Specially-Designed DNNs

As described above, the specially-designed DNN for the close DOA scenario (1° DOA difference) can be used jointly with other DOA estimators. Here, we use two DNNs trained using random DOA data and 1°-restricted data in parallel. We simply assumed that the estimation status was flagged as "success" when at least one of the two DNNs estimated correct DOAs, i.e. the "OR" result of two DNNs' success flags although the method selecting the correct DNN has not been developed yet.

Figure 12 shows the success rates for the parallel use case and for the single use case of the "general-purpose" DNN trained with random DOA data. Since the probability that 1° DOA difference occurs is about 1.6%, preparation for such a case may not be so important. However, the parallel use certainly improves the success rate by about 1.6% compared with the single use of the general-purpose DNN. This property implies the possibility of performance improvement by combination of specially-trained DNNs.

5. Conclusions

In this paper, we have evaluated an on-grid DOA estimation method using deep learning with simple DNN configuration for a simple estimation problem where only two narrowband signals of equal power are incident from integer angles in degrees. Six training data sets are created by setting different SNR transitions. Among them, the data set having the constant SNR of 30 dB was found to provide the best overall estimation performance when the SNR of estimation tests was 10 dB or higher. It has also been shown that the random SNR training data can generate a DNN that works well over a wide SNR range.

Although the success rates of the trained DNNs at the SNR of 30 dB are still lower than 90%, the absolute value of errors equal to or less than 1° is achieved for about 95% signals. It can be said that the DNN has potential as a high-resolution DOA estimator.

In addition, it has been indicated that the DNN designed for a specific scenario involving two incident signals where these DOAs differ only by 1° achieves the very high success rate in the estimation of the same scenario. This implies the integrated use of such specialized DNNs improves the DOA estimation performance.

Unfortunately, except the case of 1° DOA difference, the performance of Root MUSIC outperforms that of DNN. Thus, the DNN parameters such as the number of layers and units, the activation function for each unit, and training data

sets should be definitely refined. In addition, evaluations on off-grid estimation performance and the development of multiple DNN integration are urgent and interesting topics for us.

References

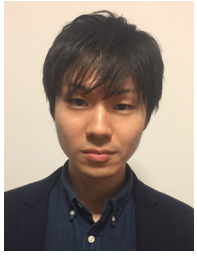
[1] Y. Kase, T. Nishimura, T. Ohgane, Y. Ogawa, D. Kitayama, and Y. Kishiyama, "DOA estimation of two targets with deep learning," Proc. IEEE WPNC 2018, Oct. 2018.
[2] J.C. Chen, K. Yao, and R.E. Hudson, "Source localization and beamforming," IEEE Signal Process. Mag., vol.19, no.2, pp.30–39, March 2002.
[3] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," Proc. IEEE, vol.57, no.8, pp.1408–1418, Aug. 1969.
[4] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," IEEE Sig. Process. Mag., vol.13, no.4, pp.67–94, July 1996.
[5] A. Massa, P. Rocca, and G. Oliveri, "Compressive sensing in electromagnetics — A review," IEEE Antennas Propag. Mag., vol.57, no.1, pp.224–238, Feb. 2015.
[6] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," IEICE Trans. Commun., vol.E96-B, no.3, pp.685–712, March 2013.
[7] T. Terada, T. Nishimura, Y. Ogawa, T. Ohgane, and H. Yamada, "DOA estimation for multi-band signal sources using compressed sensing techniques with Khatri-Rao processing," IEICE Trans. Commun., vol.E97-B, no.10, pp.2110–2117, Oct. 2014.
[8] G.E. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," Neural Comput., vol.18, no.7, pp.1527–1544, July 2006.
[9] G.E. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol.313, no.5786, pp.504–507, July 2006.
[10] R. Takeda and K. Komatani, "Sound source localization based on deep neural networks with directional activate function exploiting phase information," Proc. IEEE ICASSP 2016, pp.405–409, March 2016.
[11] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," IEEE Trans. Veh. Technol., vol.67, no.9, pp.8549–8560, Sept. 2018.
[12] Z. Liu, C. Zhang, and P.S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," IEEE Trans. Antennas Propag., vol.66, no.12, pp.7315–7327, Dec. 2018.
[13] D.P. Kingma and J.L. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980v9, Jan. 2017.
[14] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," J. Machine Learning Research, vol.12, pp.2121–2159, July 2011.
[15] M.D. Zeiler, "ADADELTA: An adaptive learning rate method," arXiv:1212.5701v1, Dec. 2012.
[16] B.D. Rao and K.V.S. Hari, "Performance analysis of Root-MUSIC," IEEE Trans. Acoust., Speech, Signal Process., vol.37, no.12, pp.1939–1949, Dec. 1989.

Yuya Kase received a B.E. degree in electrical and electronic engineering from National Institute of Technology, Asahikawa College, Asahikawa, Japan, in 2017. He is currently a Ph.D. student at the master course student at the Graduate School of Information Science and Technology, Hokkaido University. His interests are DOA estimation and mobile communication.

Toshihiko Nishimura received the B.S. and M.S. degrees in physics and Ph.D. degree in electronics engineering from Hokkaido University, Sapporo, Japan, in 1992, 1994, and 1997, respectively. Since 1998, he has been with Hokkaido University, where he is currently an Associate Professor. His current research interests are in MIMO systems using smart antenna techniques. He received the Young Researchers' Award of IEICE in 2000, the Best Paper Award from IEICE in 2007, and TELECOM System Technology Award from the Telecommunications Advancement Foundation of Japan in 2008, the best magazine paper award from IEICE Communications Society in 2011, and the Best Tutorial Paper (in Japanese) Award from the IEICE Communications Society in 2018. He is a member of the IEEE.

Takeo Ohgane received the B.E., M.E., and Ph.D. degrees in electronics engineering from Hokkaido University, Sapporo, Japan, in 1984, 1986, and 1994, respectively. From 1986 to 1992, he was with Communications Research Laboratory, Ministry of Posts and Telecommunications. From 1992 to 1995, he was on assignment at ATR Optical and Radio Communications Research Laboratory. Since 1995, he has been with Hokkaido University, where he is currently a Professor. During 2005–2006, he was at Centre for Communications Research, University of Bristol, U.K., as a Visiting Fellow. His research interests are in MIMO signal processing for wireless communications. He received the IEEE AP-S Tokyo Chapter Young Engineer Award in 1993, the Young Researchers' Award of IEICE in 1990, the Best Paper Award from IEICE in 2007, TELECOM System Technology Award from the Telecommunications Advancement Foundation of Japan in 2008, the Best Magazine Paper Award from IEICE Communications Society in 2011, and the Best Tutorial Paper (in Japanese) Award from IEICE Communications Society in 2018. He is a member of the IEEE.

Yasutaka Ogawa received the B.E., M.E., and Ph.D. degrees from Hokkaido University, Sapporo, Japan, in 1973, 1975, and 1978, respectively. Since 1979, he has been with Hokkaido University, where he is currently a Professor Emeritus. During 1992–1993, he was with ElectroScience Laboratory, the Ohio State University, as a Visiting Scholar, on leave from Hokkaido University. His professional expertise encompasses super-resolution estimation techniques, applications of adaptive antennas for mobile communication, multiple-input multiple-output (MIMO) techniques, and measurement techniques. He proposed a basic and important technique for time-domain super-resolution estimation for electromagnetic wave measurement such as antenna gain measurement, scattering/diffraction measurement, and radar imaging. Also, his expertise and commitment to advancing the development of adaptive antennas contributed to the realization of space division multiple accesses (SDMA) in the Personal Handy-phone System (PHS). He received the Yasujiro Niwa Outstanding Paper Award in 1978, the Young Researchers' Award of IEICE in 1982, the Best Paper Award from IEICE in 2007, TELECOM system technology award from the Telecommunications Advancement Foundation of Japan in 2008, the Best Magazine Paper Award from IEICE Communications Society in 2011, the Achievement Award from IEICE in 2014, and the Best Tutorial Paper (in Japanese) Award from IEICE Communications Society in 2018. He also received the Hokkaido University Commendation for excellent teaching in 2012. He is a Fellow of the IEEE.



Daisuke Kitayama received the B.S. degree in electrical and electronic engineering, the M.S. degree in electronics and applied physics, all from Tokyo Institute of Technology, Tokyo, Japan, in 2010 and 2012, respectively. He joined NTT Device Technology Laboratories, NTT Corporation in 2012. He has been engaged in research on metamaterials for controlling terahertz waves. Since October 2017, he has been working as a researcher in 5G Laboratories, NTT DOCOMO, INC. and is involved in research of fifth generation mobile communications system. He received the Young Researchers' Award from IEICE in 2016. He is a member of IEICE.

Yoshihisa Kishiyama received his B.E., M.E., and Ph.D. degrees from Hokkaido University, Sapporo, Japan in 1998, 2000, and 2010, respectively. He is now a Senior Research Engineer of 5G Laboratory in NTT DOCOMO, INC. Since he joined NTT DOCOMO in 2000, he has been involved in research and development on 4G/5G radio access technologies and physical layer standardization in 3GPP. In 2012, he received the International Telecommunication Union Association of Japan (ITU-AJ) Award for contributions to LTE.