# Modified Sound Field Interpolation Method for Rotation-robust Beamforming with Unequally Spaced Circular Microphone Array


Shuming Luan
*Graduate School of Information Science*
*Nagoya University*
Nagoya, Japan
shuming.luan@g.sp.m.is.nagoya-u.ac.jp

Yukoh Wakabayashi
*Department of Computer Science*
*Tokyo Metropolitan University*
Tokyo, Japan
wakayuko@tmu.ac.jp

Tomoki Toda
*Information Technology Center*
*Nagoya University*
Nagoya, Japan
tomoki@icts.nagoya-u.ac.jp



***Abstract*—In this paper, we propose an improved method for interpolating a sound field to make it possible to use an unequally spaced circular microphone array (CMA) in rotation-robust beamforming. The previous method of sound field interpolation for rotation-robust beamforming required a regular CMA. This constraint causes degeneration when the microphones are non-uniformly spaced and makes the previous method difficult to utilize in practical applications. To address this issue, the proposed method first estimates the sound signal at equally spaced positions from that observed with the unequally spaced CMA after rotation and then performs a variant of the previous method to obtain the signal before rotation. A series of simulated experiments are conducted. The experimental results indicate that the proposed method avoids the adverse effects of unequal placement and shows a greater improvement in estimating the signal before rotation under different circumstances. Also, the proposed method significantly improves the performance of beamforming compared with the previous method.**



***Index Terms*—Sound field interpolation, rotation-robust beamforming, unequally spaced circular microphone array**


## I. Introduction

Array signal processing methods are still major challenges in some research topics, such as source separation and source enhancement. The common advanced source separation methods based on, for example, beamforming [1], [2], independent component analysis [3], [4], and nonnegative matrix factorization [5], [6], accomplish an impressive performance by modifying the calculation methodology, the spatial model, or the source model. However, a time-invariant acoustic transfer system (ATS) is usually required by most state-of-the-art methods to maintain the performance. The utilization of a time-invariant ATS in these methods means that not only the sound source but also the microphones' positions cannot move when performing source separation and enhancement.

In this paper, we assume a situation where a user or a humanoid robot wears a circular microphone array (CMA) on the head and rotates the head to listen to the target sound source in a noisy environment. The rotation of the CMA represents the variation of the ATS, which requires the re-estimation of the spatial filter. Statistical information such as the covariance matrix is usually the key to estimating the spatial model in most array signal processing methods, which has a long calculation time. Therefore, the rotation of the CMA in our assumed situation makes real-time processing difficult.

A new framework of beamforming [7] robust to one of the ATS movements, the rotation of a CMA, was proposed to address this bottleneck in online processing. This new technique employs sound field interpolation based on a non-integer sample shift theorem, which utilizes the periodicity of the sound field on the circumference of a circle and the relationship between sensing the sound field with a CMA and discretizing the sound field. By applying sound field interpolation before beamforming, we don't need to update the beamformer's filter and can directly use the previous one in arbitrary sound fields when the CMA rotates. It was demonstrated that this method enables the robust estimation of the lower band spectrum and achieves a high performance when applied to an existing beamformer, even when the CMA rotates. At the same time, this framework requires an equally spaced CMA to sense the sound field so that the discretized sound field is a periodic function. This means that an error of any microphone's position on the CMA changes the equally spaced CMA into an unequally spaced one, markedly reducing the efficacy of the method. In most practical applications, it is highly likely that the microphones on a CMA are not set at uniform positions. Hence, an equally spaced CMA rarely exists but an unequally spaced one is much more universal, which causes difficulties in the application of this approach in a real environment.

In this study, we focus on both the rotation of a CMA and the errors of the microphones' positions, and present a new method on the basis of our previous work [7]. We develop a new method for beamforming that is robust to the array's rotation using an unequally spaced CMA. The conceptual diagram of this proposed method is illustrated in Fig. 1. This newly proposed method makes it possible for both the previously proposed scheme and the conventional beamforming method to work well when an unequally spaced CMA rotates because, by applying our new technique, the time-variant ATS on an unequally spaced CMA can be regarded as a time-invariant one

ISBN: 978-1-6654-6798-8

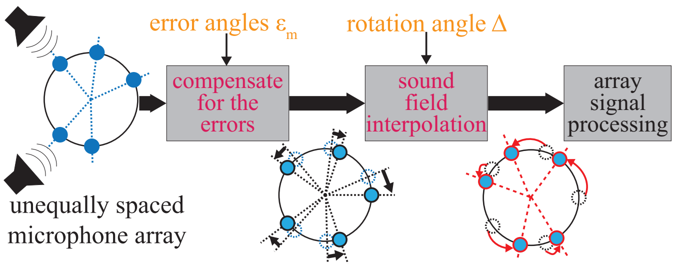


Fig. 1. Conceptual diagram of the proposed system.

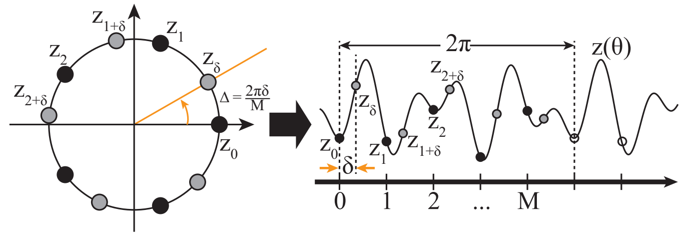


Fig. 2. Continuous sound field on a circle's circumference and the discretized sound field function with a $\delta$ sample shift.

on an equally spaced CMA virtually. The proposed framework is still based on sound field interpolation and the noninteger sample shift theorem, and we utilize a variant of previous formulations to compensate for the errors of the microphones' positions on an unequally spaced CMA. The high performance of this method is confirmed by simulated experiments on an existing beamformer.

## II. Conventional Sound Field Interpolation Method with Regular CMA

In this section, we explain our previous research [7] on sound field interpolation based on a noninteger sample shift theorem for a time-variant ATS using a regular CMA.

Let $z(\theta)$ be a continuous sound field function on a circle's circumference, as shown in Fig. 2. Obviously, $z(\theta)$ is a periodic function with $2\pi$ as the period, where $\theta \in [0, 2\pi)$ is the spatial angle. Because observing the sound field with a CMA corresponds to discretizing the sound field function $z(\theta)$, to make the discretized sound field function also a periodic function, we must set the microphones on a CMA at even intervals. Therefore, $z(\theta)$ is discretized by an M-channel CMA with interval $2\pi/M$ so that the observed signal in the $m$th channel is represented as

$$z_m = z\left(2\pi\frac{m}{M}\right), \quad m = 0, ..., M-1. \tag{1}$$

The continuous sound field function $z(\theta)$ can be reconstructed from the discretized sound signal $z_m$ provided that the sampling theorem is satisfied. Hence, sound field interpolation is possible using the noninteger sample theorem in the Fourier domain. A $\Delta$-rad-rotated sound field $z(2\pi m/M + \Delta)$ observed by a CMA is consistent with a $\delta$-sample-shifted discretized sound signal $z_{m+\delta}$, where $\delta = M\Delta/2\pi$. Using the sample shift theorem in the DFT, $z_{m+\delta}$ can be represented by $z_0$, $z_1$, ..., $z_{M-1}$ as

$$z_{m+\delta} = \sum_{n=0}^{M-1} z_n U_{m,n,\delta}. \tag{2}$$

$U_{m,n,\delta}$ is the coefficient of sound field interpolation, which is calculated utilizing the $sinc$ function as

$$U_{m,n,\delta} = \begin{cases} \frac{\left(1-e^{jL\pi}\right)}{M} + \frac{\mathrm{sinc}\left(\frac{L}{2}\right)\cos\left(\frac{M+2}{2M}L\pi\right)}{\mathrm{sinc}\left(\frac{L}{M}\right)}, & M \text{ is even}, \\ \frac{1}{M} + \frac{M-1}{M}\frac{\mathrm{sinc}\left(\frac{L(M-1)}{2M}\right)\cos\left(\frac{M+1}{2M}L\pi\right)}{\mathrm{sinc}\left(\frac{L}{M}\right)}, & M \text{ is odd}, \end{cases} \tag{3}$$

where $L = (n - m - \delta)/2M$ and $j = \sqrt{-1}$. In matrix representation, (2) can also be defined as

$$\begin{bmatrix} z_{0+\delta} \\ \vdots \\ z_{M-1+\delta} \end{bmatrix} = \begin{bmatrix} U_{0,0,\delta} & \cdots & U_{0,M-1,\delta} \\ \vdots & \ddots & \vdots \\ U_{M-1,0,\delta} & \cdots & U_{M-1,M-1,\delta} \end{bmatrix} \begin{bmatrix} z_0 \\ \vdots \\ z_{M-1} \end{bmatrix} = \boldsymbol{U}_\delta \boldsymbol{z}_0, \tag{4}$$

where $\boldsymbol{U}_\delta$ is the rotation transform matrix and independent of the frequency.

## III. Proposed Sound Field Interpolation Method with Unequally Spaced CMA

Two problems are addressed in this paper: the time-variant ATS and the unequally spaced CMA. We propose a modified sound field interpolation method capable of addressing these two problems.

### A. Overview

In our proposed method, we assume that the error angle of each microphone, $\epsilon_m$, is already known beforehand. As shown in Fig. 1, the proposed method is divided into two steps. In the first step, we use $\epsilon_m$ and the observation recorded by the unequally spaced CMA to estimate the signal on a virtual equally spaced CMA. Then, utilizing the rotation angle $\Delta$, we generate the sound signal before rotation by a variant of the previous sound field interpolation method with a regular CMA. Finally, the estimated result is applied to other array signal processing methods.

### B. Formulation

In this subsection, we introduce the formulation of the proposed method. The error vector is represented as $\boldsymbol{\epsilon} = [\epsilon_0, \epsilon_1, ..., \epsilon_{M-1}]^\mathsf{T}$. Accordingly, the sound signal observed by an unequally spaced CMA is expressed as $\boldsymbol{z}_{\boldsymbol{\epsilon}} = \left[z_{0+\epsilon_0}, z_{1+\epsilon_1}, ..., z_{M-1+\epsilon_{M-1}}\right]^\mathsf{T}$.

From (2), for the $m$th channel, $\epsilon_m$ can be regarded as the sample shift $\delta$, and the corresponding shifted sound signal $z_{m+\epsilon_m}$ can be computed from the coefficient $U_{m,n,\epsilon_m}$. From matrix formulation (4), it can be inferred that if we obtain the $m$th shifted signal $z_{m+\delta}$, only the $(m+1)$th row of the rotation matrix $\boldsymbol{U}_\delta$ is necessary. Similarly, the relationship between the $m$th channel signal of the unequally spaced CMA, $z_{m+\epsilon_m}$, and the pseudo-observation recorded by an equally spaced CMA, $\boldsymbol{z}_0$, can be defined as

$$z_{m+\epsilon_m} = \boldsymbol{U}_{\epsilon_m}(m+1,:)\,\boldsymbol{z}_0, \tag{5}$$

where $\boldsymbol{U}_{\epsilon_m}(m+1,:)$ is the $(m+1)$th row of the rotation matrix $\boldsymbol{U}_{\epsilon_m}$. Referring to (5), the observation recorded by an unequally spaced CMA can be represented as

$$\boldsymbol{z}_{\boldsymbol{\epsilon}} = \boldsymbol{U}_{\boldsymbol{\epsilon}}\boldsymbol{z}_0, \tag{6}$$

where $\boldsymbol{U}_{\boldsymbol{\epsilon}}$ is defined as

$$\boldsymbol{U}_{\boldsymbol{\epsilon}} = \begin{bmatrix} \boldsymbol{U}_{\epsilon_0}(1,:) \\ \boldsymbol{U}_{\epsilon_1}(2,:) \\ \vdots \\ \boldsymbol{U}_{\epsilon_{M-1}}(M,:) \end{bmatrix}. \tag{7}$$

From (6), $\boldsymbol{z}_0$ can be calculated as

$$\boldsymbol{z}_0 = \boldsymbol{U}_{\boldsymbol{\epsilon}}{}^{-1}\boldsymbol{z}_{\boldsymbol{\epsilon}}. \tag{8}$$

Note that $\boldsymbol{U}_{\boldsymbol{\epsilon}}{}^{-1}$ does not always exist because $\boldsymbol{U}_{\boldsymbol{\epsilon}}$ may be a nearly singular matrix. In such a situation, we apply singular value decomposition (SVD) to $\boldsymbol{U}_{\boldsymbol{\epsilon}}$ and discard the singular values and singular vectors corresponding to the abnormally small condition number. Then, the factorized results of this truncated SVD are employed to compute the inverse matrix of $\boldsymbol{U}_{\boldsymbol{\epsilon}}$.

By using (8), we compensate for the errors and obtain the sound signal of a virtual equally spaced CMA from the observation of an unequally spaced CMA. In other words, the unequally spaced CMA is transformed to an equally spaced one virtually using the inverse matrix of $\boldsymbol{U}_{\boldsymbol{\epsilon}}$.

For $\boldsymbol{z}_{\boldsymbol{\epsilon}}$, the $\delta$-sample-shifted result is $\boldsymbol{z}_{\boldsymbol{\epsilon}+\delta}$, which can be expressed as $\left[z_{0+\epsilon_0+\delta}, z_{1+\epsilon_1+\delta}, ..., z_{M-1+\epsilon_{M-1}+\delta}\right]^\mathsf{T}$. According to (6) and (7), the relationship between $\boldsymbol{z}_0$ and $\boldsymbol{z}_{\boldsymbol{\epsilon}+\delta}$ is

$$\boldsymbol{z}_{\boldsymbol{\epsilon}+\delta} = \boldsymbol{U}_{\boldsymbol{\epsilon}+\delta}\boldsymbol{z}_0 \tag{9}$$

$$\boldsymbol{U}_{\boldsymbol{\epsilon}+\delta} = \begin{bmatrix} \boldsymbol{U}_{\epsilon_0+\delta}(1,:) \\ \boldsymbol{U}_{\epsilon_1+\delta}(2,:) \\ \vdots \\ \boldsymbol{U}_{\epsilon_{M-1}+\delta}(M,:) \end{bmatrix}. \tag{10}$$

$\boldsymbol{z}_0$ in (9) can be substituted with (8) so that $\boldsymbol{z}_{\boldsymbol{\epsilon}+\delta}$ can be directly calculated utilizing $\boldsymbol{z}_{\boldsymbol{\epsilon}}$:

$$\boldsymbol{z}_{\boldsymbol{\epsilon}+\delta} = \boldsymbol{U}_{\boldsymbol{\epsilon}+\delta}\boldsymbol{U}_{\boldsymbol{\epsilon}}{}^{-1}\boldsymbol{z}_{\boldsymbol{\epsilon}}. \tag{11}$$

Therefore, sound field interpolation for a time-variant ATS on an unequally spaced CMA is achieved.

## IV. Experimental evaluation

### A. Setup

Simulated experiments were conducted to assess the performance of the proposed method and verify its robustness to the rotation of an unequally spaced CMA. The SiSEC database [8] was adopted, each utterance of which was sampled at 16 kHz. Eight samples (four female and four male) were selected from the database as the sound sources in different directions, as displayed in Fig. 3. The sound signals were convolved with room impulse responses (RIRs) simulated by an RIR generator [9] based on the image method [10], and the reverberation time was roughly 100 ms. The signals were recorded by an unequally spaced $M$-channel CMA with a radius of 0.05 m in a noise-free room. The error of each microphone's position, $\epsilon_i$ (deg), $i \in \{0, ..., M-1\}$ had a Gaussian distribution with zero mean and variance $\sigma^2 \in \{0, 10, 20, ..., 490, 500\}$, and all the errors were independently identically distributed. For each Gaussian distribution with a certain variance, there were a total of 100 samples. We also recorded the $\Delta$-rad-rotated sound field with the CMA. Then, we utilized the $\Delta$-rad-rotated sound signal to estimate the signals before rotation. Here, the rotational angle $\phi = \Delta\pi/180$ deg was known. To conduct the sound field interpolation in the time-frequency domain, we performed the short-time Fourier transform (STFT) using a 1/8-shifted Blackman window with a length of 64 ms.

In the first experiment, the performance in the case of only one source was evaluated in terms of the signal-to-error ratio (SER), defined as

$$\mathrm{SER}_{m,k} = 10\log_{10}\left(\frac{\sum_t |x_{m,t,k}|^2}{\sum_t |\hat{x}_{m,t,k} - x_{m,t,k}|^2}\right), \tag{12}$$

where $x_{m,t,k}$ is the time-frequency domain signal and $\hat{x}_{m,t,k}$ is its estimate. $m$, $t$, and $k$ denote the channel, time frame, and frequency bin, respectively. $M$ and $\phi$ were set as 4–6 and 10, 20, and 30 deg, respectively.

In the second experiment, we utilized the minimum variance distortionless response (MVDR) beamformer [11]–[13], where the power of the output signal is minimized subject to a single constraint assuring an undistorted response for the target sound source, to compare the source enhancement performance with different interpolation methods using source to distortion ratio (SDR) and source to interferences ratio (SIR) [14]. We used the RIR from the target to each microphone to calculate the relative transfer function (RTF) [15] in order to estimate the beamformer's filter as follows:

$$\mathbf{w}_{\mathrm{MVDR}} = \frac{\boldsymbol{\Phi}_{nn}^{-1}\tilde{\mathbf{h}}_0}{\tilde{\mathbf{h}}_0^H\boldsymbol{\Phi}_{nn}^{-1}\tilde{\mathbf{h}}_0}, \tag{13}$$

where $\boldsymbol{\Phi}_{nn}$ and $\tilde{\mathbf{h}}_0$ denote the covariance matrix of the interference signal and RTF respectively. Two sources were randomly chosen and mixed into the observation so that 12 environments (two patterns at each of six angles, 30, 60, ..., 180 deg) were simulated. Here, $M$ and $\phi$ were set as 5 and 10, 20, and 30 deg, respectively.

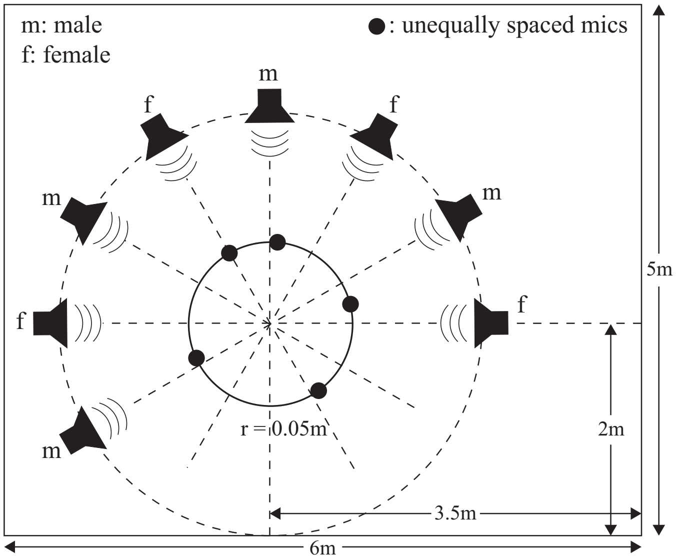


Fig. 3. Simulated environment in the experiments.

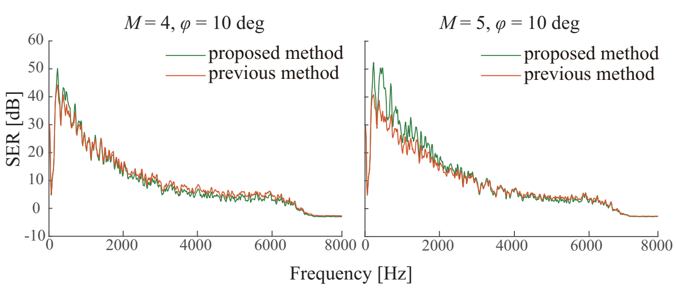


Fig. 4. Examples of SERs as a function of frequency.

## B. Comparative results of proposed and previous methods

First, only the sound source in the direction of 0 deg is considered. For such an environment, some examples of SER results when $\phi$ is 10 deg and $M$ is 4 or 5 are illustrated in Fig. 4. We display the mean SER among $M$ channels for a certain variance (400) of the error angle $\epsilon_i$. As demonstrated by the figure, the proposed method is generally able to estimate the spectrum almost as well as the previous method [7] for $M = 4$. This new scheme performs significantly better when estimating the lower-frequency component for $M = 5$.

In Fig. 5, we present the relationship between the variance of the error angle and the SER improvement in the frequency range of up to 1 kHz relative to the cases without sound field interpolation when $M$ is 5 and $\phi$ is 10 or 20 deg. We demonstrate the mean SER improvement over 0–1 kHz for five channels for each sample, thus, there are 100 data points in each box. It is clear that with increasing variance, the SER improvement of the previous method markedly degrades, whereas our proposed method achieves much steadier and better results. Namely, directly applying ordinary sound field interpolation to an unequally spaced CMA is not sensible, and it is advantageous to utilize our new technique.

Fig. 6 depicts SER improvement with different $M$ and $\phi$

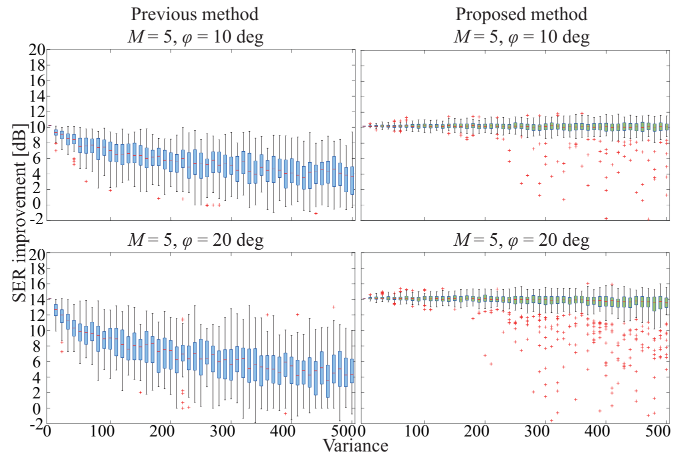


Fig. 5. Boxplots of the relationship between the variance of the error angle and the SER improvement in the frequency range of up to 1 kHz relative to the cases without interpolation.

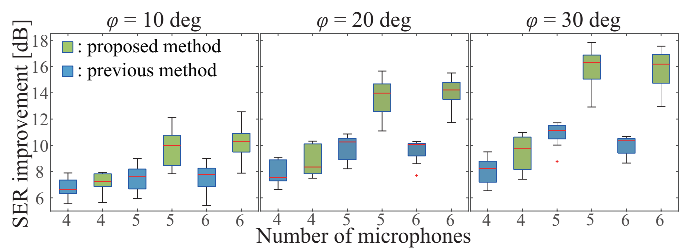


Fig. 6. Boxplots of mean SER improvement over the frequency range of up to 1 kHz for $M$ channels in eight situations relative to the cases without interpolation.

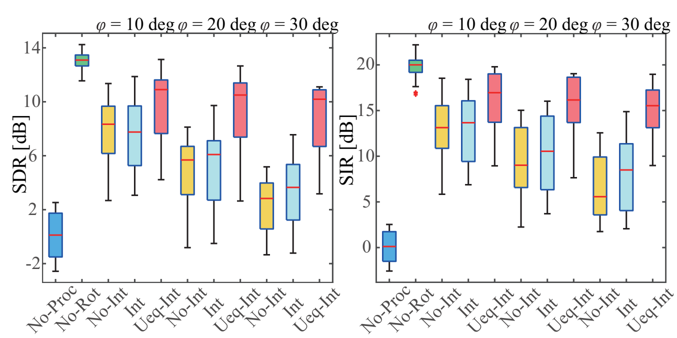


Fig. 7. Boxplots of SDR and SIR obtained by MVDR beamformer in five situations: unprocessed (**No-Proc**), no rotation of the CMA (**No-Rot**), without interpolation when the CMA rotates (**No-Int**), with ordinary interpolation when the CMA rotates (**Int**), and unequally spaced interpolation when the CMA rotates (**Ueq-Int**).

values for a variance of 100. We calculate the mean SER improvement relative to the cases without interpolation over the frequency range of up to 1 kHz for $M$ channels. In each box of Fig. 6, there are eight samples, corresponding to the mean SER improvement of eight different sound sources. These results illustrate that in most situations, the proposed method shows a greater improvement than the previous technique.

### C. Evaluation results of source enhancement

In this experiment, we examine the source enhancement performance using the MVDR beamformer. First, the MVDR beamformer's filter weight $\boldsymbol{w}$ is produced utilizing the RTF and the multichannel STFT spectrogram obtained from the unequally spaced CMA at the position before rotation. Then, $\boldsymbol{w}$ is directly applied to this spectrogram without rotating; we denote this simulation as **No-Rot**. We also estimate the spectrograms without any interpolation after the unequally spaced CMA rotates (**No-Int**), the spectrograms with ordinary interpolation when the CMA rotates (**Int**), and the spectrograms with unequally spaced interpolation on the time-variant CMA (**Ueq-Int**). Similarly, these spectrograms are also post-processed by $\boldsymbol{w}$ so as to generate the estimated target sound signal. The unprocessed case (**No-Proc**) and **No-Rot** are employed as baselines for comparison with the results in other cases.

The SDR and SIR results for a variance of 100 are depicted in Fig. 7. As expected, **No-Proc** has the lowest SDR and SIR values, whereas **No-Rot** has the greatest source enhancement performance of the MVDR beamformer when the ATS does not rotate. Interestingly, **Int** does not perform as well as expected. In most environments, **Int**'s SDR and SIR values show almost no difference from those of **No-Int**, and under some circumstances, **Int**'s performance is even lower than **No-Int**'s. This demonstrates that if the CMA is unequally spaced, the previous method of sound field interpolation cannot work when the CMA rotates. In source enhancement using the MVDR beamformer, ordinary interpolation is unnecessary because of the slight difference between **Int** and **No-Int**. Moreover, it is likely that even without interpolation, we can achieve better source enhancement than with interpolation. The proposed method (**Ueq-Int**) outperforms the cases without interpolation (**No-Int**) and the cases with the previous interpolation technique (**Int**), and attains the closest results to the best performance (**No-Rot**) regardless of the type of simulated environment used. Our unequally spaced interpolation method is robust to a non-uniform distribution of the microphones on the CMA to some extent and significantly improves the performance of array signal processing.

## V. Conclusion

In this paper, we presented a novel framework of beamforming robust to the rotation of an unequally spaced microphone array on the basis of our previous work. We extended simple sound field interpolation to an unequally spaced sound field interpolation method by analyzing the principle and process of interpolation. This new framework virtually regards the time-variant ATS on an unequally spaced CMA as a time-invariant one on an equally spaced CMA by utilizing the noninteger sample shift theorem. By simulated experiments, it was demonstrated that our newly presented system is not affected by a non-uniform distribution and compensates for the errors of the microphones' positions. This method also assists array signal processing, even when the CMA rotates. However, our proposed method still requires improvement, especially when there are a small number of microphones on the CMA. Also, the estimation of the high-frequency component is still challenging. In this study, we assumed that the errors of each microphone's position were already known but, in reality, these errors will not be known under most conditions. Thus, without these known variables, performing sound field interpolation on an unequally spaced CMA is an interesting and worthwhile research direction. These problems will be investigated in our future work.

## Acknowledgment

This work was partly supported by JST CREST Grant Number JPMJCR19A3, Japan.

## References

[1] K. Yamaoka, N. Ono, S. Makino, and T. Yamada, "Time-frequency-bin-wise switching of minimum variance distortionless response beamformer for underdetermined situations," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 7908–7912.

[2] Y. Kubo, T. Nakatani, M. Delcroix, K. Kinoshita, and S. Araki, "Mask-based MVDR beamformer for noisy multisource environments: Introduction of time-varying spatial covariance model," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 6855–6859.

[3] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 9, pp. 1626–1641, 2016.

[4] N. Makishima, S. Mogami, N. Takamune, D. Kitamura, H. Sumino, S. Takamichi, H. Saruwatari, and N. Ono, "Independent deeply learned matrix analysis for determined audio source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 10, pp. 1601–1615, 2019.

[5] A. Ozerov and C. Févotte, "Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 3, pp. 550–563, 2009.

[6] R. Scheibler and N. Ono, "Fast and stable blind source separation with rank-1 updates," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 236–240.

[7] Y. Wakabayashi, K. Yamaoka, and N. Ono, "Rotation-Robust Beamforming Based on Sound Field Interpolation with Regularly Circular Microphone Array," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 771–775.

[8] S. Araki, F. Nesta, E. Vincent, Z. Koldovskỳ, G. Nolte, A. Ziehe, and A. Benichoux, "The 2011 Signal Separation Evaluation Campaign (SiSEC2011):-audio source separation," in *International Conference on Latent Variable Analysis and Signal Separation*, 2012, pp. 414–422.

[9] E. A. Habets, "Room impulse response generator," *Technische Universiteit Eindhoven, Tech. Rep*, vol. 2, no. 2.4, p. 1, 2006.

[10] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

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

[12] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[13] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[14] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1462–1469, 2006.

[15] S. Doclo, W. Kellermann, S. Makino, and S. E. Nordholm, "Multichannel signal enhancement algorithms for assisted listening devices: Exploiting spatial diversity using multiple microphones," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 18–30, 2015.