EMPLOYING MOMENTS OF MULTIPLE HIGH ORDERS FOR HIGH-RESOLUTION UNDERDETERMINED DOA ESTIMATION BASED ON *MUSIC*

Yuya Sugimoto[1], *Shigeki Miyabe*[1], *Takeshi Yamada*[1], *Shoji Makino*[1] *and Fred Juang*[2]

[1]University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577 Japan
[2]Georgia Institute of Technology, 777 Atlantic Dr. NW. Atlanta, GA 30332-0250 USA
Email: sugimoto@mmlab.cs.tsukuba.ac.jp, {miyabe, makino}@tara.tsukuba.ac.jp,
takeshi@cs.tsukuba.ac.jp, juang@ece.gatech.edu

ABSTRACT

Several extensions of the MUltiple SIgnal Classification (MUSIC) algorithm exploiting high order statistics were proposed to estimate directions of arrival (DOAs) with high resolution in underdetermined conditions. However, these methods entail a trade-off between two performance goals, namely, robustness and resolution, in the choice of orders because use of high-ordered statistics increases not only the resolution but also the statistical bias. To overcome this problem, this paper proposes a new extension of MUSIC using a nonlinear high-dimensional map, which corresponds to the joint analysis of moments of multiple orders and helps to realize the both advantages of robustness and high resolution of low-ordered and high-ordered statistics. Experimental results show that the proposed method can estimate DOAs more accurately than the conventional MUSIC extensions exploiting moments of a single high order.

Index Terms— MUSIC, Higher order statistics, Microphone array, Underdetermined DOA estimation, Subspace analysis

1. INTRODUCTION

In microphone array signal processing, direction of arrival (DOA) estimation is essential in various applications such as speaker tracking and preprocessing for signal enhancement. MUltiple SIgnal Classification (MUSIC) [1] is a representative high-resolution DOA estimation method based on the subspace analysis of the observed signals. However, the effectiveness of MUSIC is conditioned on the dimensionality of the covariance matrix because it requires the estimation of the noise subspace that is orthogonal to the transfer function vectors of the observed signals. With N sound sources, M $(> N)$ sensors are required and the DOA estimation performance degrades significantly as N approaches M.

For high-resolution DOA estimation in underdetermined conditions where the number of sources exceeds the number of microphones, several extensions of MUSIC were developed. In the early 90's, extended MUSIC exploiting fourth-order cumulants was proposed [2], and recently this approach was generalized to cumulants analysis of arbitrary even order, referred to as $2q$-MUSIC [3]. These methods improve the resolution of DOA estimation and can also estimate DOAs in underdetermined conditions by increasing the dimensionality of observation with higher-order cross cumulants. As an alternative extension, we proposed *the mapped MUSIC* [4] using a high-dimensional nonlinear map of the observation, which is inspired by the nonlinear beamformer [5, 6, 7] for speech enhancement in a similar underdetermined condition. We showed that the mapped MUSIC corresponds to the analysis of arbitrary even order cross moments. The mapped MUSIC achieves a performance in DOA estimation comparable to that of conventional extensions of MUSIC based on cumulant analysis, but offers an advantage in computational complexity due to its efficient algorithm in dimensionality expansion.

In this paper, we propose a method for joint analysis of moments of multiple orders to further improve the resolution of DOA estimation in the MUSIC framework. Although use of higher order statistics yields a higher resolution in the DOA estimate, the result is less robust due to the greater bias in the higher-order statistics, as examined in [4]. Thus the question of a strategy for trade-off between resolution and robustness arises. To explore the possibility of realizing both advantages of the high and low orders of statistics, we evaluate moments of multiple orders simultaneously. This goal was found to be possible and can be accomplished by properly combining low-and-high-ordered moments in a nonlinearly expanded space, as inspired by the mapped MUSIC method. Experimental results of speech DOA estimation demonstrate that the mapped MUSIC with the proposed map achieves a higher accuracy than other MUSIC extensions.

2. PROBLEM STATEMENT

Throughout this paper, signals are expressed as complex amplitudes in the time-frequency domain. Observed signals can be modeled as

$$\mathbf{x}(\omega,t) = [x_1(\omega,t),\cdots,x_M(\omega,t)]^{\mathrm{T}} = \sum_{i=1}^{N}\mathbf{a}_i(\omega)s_i(\omega,t)+\mathbf{n}(\omega,t), \quad (1)$$

$$\mathbf{n}(\omega,t) = [n_1(\omega,t),\cdots,n_M(\omega,t)]^{\mathrm{T}}, \quad (2)$$

$$\mathbf{a}_i(\omega) = [a_{1,i}(\omega),\cdots,a_{M,i}(\omega)]^{\mathrm{T}}, \quad (3)$$

where $t = 1,\cdots,L$ is a time frame index, ω is the angular frequency, M is the number of sensors, N is the number of sound sources, $[\cdot]^{\mathrm{T}}$ denotes transpose, $s_i(\omega,t)$ is the complex amplitude of the i-th sound source, $x_j(\omega,t)$ is the complex amplitude of the signal observed with the j-th sensor, $n_j(\omega,t)$ is the complex amplitude of the noise observed with the j-th sensor, and $a_{j,i}(\omega)$ denotes the transfer function from the i-th source to the j-th sensor.

In the signal model expressed by (1), each sensor observes a mixture of source signals and noise. The problem in this paper is to estimate the DOAs of the source signals $s_1(\omega,t),\cdots,s_N(\omega,t)$ by analyzing these observed signals and finding steering vec-

This work was supported by KAKENHI 23760325.

978-1-4799-0972-8/13/$31.00 ©2013IEEE

tors $\mathbf{b}(\omega;\theta_1),\cdots,\mathbf{b}(\omega;\theta_N)$ similar to the transfer function vectors $\mathbf{a}_1(\omega),\cdots,\mathbf{a}_N(\omega)$ from a set of steering vectors $\mathbf{b}(\omega;\theta)$ ($||\mathbf{b}(\omega;\theta)||_2=1$) of arbitrary directions θ.

$$\mathbf{b}(\omega;\theta)=\frac{1}{\sqrt{M}}\left[\exp(-j\omega\tau_1),\cdots,\exp(-j\omega\tau_M)\right], \quad (4)$$

where τ_i $(i=1,\cdots,M)$ are time delays from a reference point at each sensor, given by a signal from direction θ.

3. MAPPED MUSIC

3.1. DOA estimation algorithm of mapped MUSIC

This section briefly reviews the DOA estimation with mapped MUSIC [4]. Mapped MUSIC maps the M-dimensional observed vector $\boldsymbol{x}(\omega,t)$ onto M'-dimensional Euclidian space $(M'\geq M)$ with a nonlinear map $\boldsymbol{\phi}:\mathbb{C}^M\rightarrow\mathbb{C}^{M'}$ and conducts a similar analysis to MUSIC with mapped steering vectors $\boldsymbol{\phi}(\mathbf{b}(\omega;\theta))$. To estimate DOAs properly with mapped MUSIC, the information about the correlations between each microphone signal must be retained after mapping. This requirement for the map $\boldsymbol{\phi}$ is given by the following three conditions:

1. The magnitude relation of the norm is retained.
$$||\mathbf{x}||_2\geq||\mathbf{y}||_2\rightarrow||\boldsymbol{\phi}(\mathbf{x})||_2\geq||\boldsymbol{\phi}(\mathbf{y})||_2 \quad (5)$$
2. The origin remains intact.
$$\mathbf{x}=0\rightarrow\boldsymbol{\phi}(\mathbf{x})=0 \quad (6)$$
3. The orthogonality between vectors is preserved.
$$\mathbf{x}^{\mathrm{H}}\mathbf{y}=0\rightarrow\boldsymbol{\phi}^{\mathrm{H}}(\mathbf{x})\boldsymbol{\phi}(\mathbf{y})=0 \quad (7)$$

We describe the DOA estimation algorithm of mapped MUSIC with the map $\boldsymbol{\phi}$ satisfying (5)–(7). The covariance matrix of $\boldsymbol{\phi}(\mathbf{x}(\omega,t))$ is expressed as

$$\mathbf{R}(\omega)=E[\boldsymbol{\phi}(\mathbf{x}(\omega,t))\boldsymbol{\phi}^{\mathrm{H}}(\mathbf{x}(\omega,t))], \quad (8)$$

where $E[\cdot]$ denotes the expectation of the argument, and $[\cdot]^{\mathrm{H}}$ denotes a complex conjugate transpose. The following equations describe the eigen decomposition of covariance matrix $\mathbf{R}(\omega)$:

$$\mathbf{R}(\omega)=\mathbf{V}(\omega)\mathbf{E}(\omega)\mathbf{V}^{\mathrm{H}}(\omega), \quad (9)$$
$$\mathbf{V}(\omega)=[\mathbf{v}_1(\omega),\cdots,\mathbf{v}_{M'}(\omega)],\ \mathbf{V}^{\mathrm{H}}(\omega)\mathbf{V}(\omega)=\mathbf{I}_{M'}, \quad (10)$$
$$\mathbf{E}(\omega)=\mathrm{diag}[e_1(\omega),\cdots,e_{M'}(\omega)],$$
$$e_1(\omega)\geq\cdots\geq e_{M'}(\omega), \quad (11)$$
$$M'=\dim[\boldsymbol{\phi}(\mathbf{x}(\omega,t))], \quad (12)$$

where $\mathbf{v}_i(\omega)$ $(i=1,\cdots,M')$ are the eigenvectors associated with each eigenvalue $e_i(\omega)$ $(i=1,\cdots,M')$ respectively, $\mathbf{I}_i$ denotes the i-dimensional identity matrix, $\mathrm{diag}[\cdot]$ is a diagonal matrix consisting of elements within the argument vector, and $\dim[\cdot]$ is the dimensionality of the argument. With N' denoting the number of the eigenvectors corresponding to the mapped directional source signals;

$$\mathrm{span}[\boldsymbol{\phi}(\mathbf{a}_i(\omega))|i=1,\cdots,N]\simeq\mathrm{span}[\mathbf{v}_j(\omega)|j=1,\cdots,N'], \quad (13)$$

where $\mathrm{span}[\cdot]$ denotes the space spanned by argument vectors, we can define the noise subspace $\mathcal{N}(\omega)$ in the mapped space;

$$\mathcal{N}(\omega)\triangleq\mathrm{span}[\mathbf{v}_k(\omega)|k=N'+1,\cdots,M']. \quad (14)$$

Note that the value of N' depends on the property of the map $\boldsymbol{\phi}$ and the number of the source signal N in the original observation. Although the orthogonality between the mapped transfer functions $\boldsymbol{\phi}(\mathbf{a}_i(\omega))$ and the mapped noise subspace $\mathcal{N}(\omega)$ is retained by the conditions (5)–(7), the dimensionality N' of the mapped signal subspace is increased by the higher-dimensional map or the greater number of sources N. Thus we must choose N' appropriately to satisfy the following condition:

$$|\boldsymbol{\phi}^{\mathrm{H}}(\mathbf{a}_i(\omega))\mathbf{v}_k(\omega)|^2\ll 1$$
$$(i=1,\cdots,N\quad k=N'+1,\cdots,M'). \quad (15)$$

Since we assume $\mathbf{b}(\omega;\theta_i)\cong\mathbf{a}_i(\omega)$ $(i=1,\cdots,N)$, we can find the true sound source directions by searching the orthogonal projection of mapped steering vectors $\boldsymbol{\phi}(\mathbf{b}(\omega;\theta))$ onto the mapped noise subspace. Here, we define the following *MUSIC score* of mapped MUSIC $f(\omega;\theta)$, which is locally maximized at the direction θ_i:

$$f(\omega;\theta)\triangleq\frac{1}{\sum_{k=N'+1}^{M'}|\boldsymbol{\phi}^{\mathrm{H}}(\mathbf{b}(\omega;\theta))\mathbf{v}_{\mathbf{k}}(\omega)|^2}. \quad (16)$$

Since the MUSIC score $f(\omega;\theta)$ is defined for each narrowband, it is necessary to integrate the information from all frequency bins in order to obtain a broadband DOA estimation. Among various average operations to integrate the narrowband MUSIC scores [8], we use the following geometric mean in this paper;

$$\overline{f(\theta)}\triangleq\left[\prod_{\omega}f(\omega;\theta)\right]^{\frac{1}{K}}, \quad (17)$$

where K denotes the number of averaged frequency bins. Estimated DOAs are given by finding peaks of $\overline{f(\theta)}$.

3.2. Map for analysis of arbitrary even order moments

In [4], we proposed a map $\boldsymbol{\phi}_d:\mathbb{C}^M\rightarrow\mathbb{C}^{M^d}$ for an analysis of cross moments of $2d$-th order:

$$\boldsymbol{\phi}_d(\mathbf{x}(\omega,t))\triangleq\left[\prod_{l=1}^{d}x_{c_{1l}}^{l\otimes}(\omega,t),\cdots,\prod_{l=1}^{d}x_{c_{M^d l}}^{l\otimes}(\omega,t)\right]^{\mathrm{T}}, \quad (18)$$

$$x^{l\otimes}\triangleq\begin{cases}x & (\text{if } l \text{ is odd})\\ x^* & (\text{if } l \text{ is even})\end{cases}, \quad (19)$$

where $[\cdot]^*$ denotes a complex conjugate. c_{kl} is an index specifying the element index of vector $\mathbf{x}$ for calculating $\boldsymbol{\phi}_d$, and this is the (k,l) element of the $M^d\times d$ index matrix $\mathbf{C}$ that arranges d-repeated-permutations of M in an arbitrary order as its row vectors:

$$\mathbf{C}\triangleq[c_{kl}]_{kl}=\begin{bmatrix}1 & 1 & \cdots & 1\\ 1 & 1 & \cdots & 2\\ & & \vdots & \\ M & M & \cdots & M\end{bmatrix}, \quad (20)$$

where $[\cdot]_{ij}$ denotes a matrix consisting of the argument as its (i,j) element. We showed in our previous work [4] that the mapped MUSIC with the map $\boldsymbol{\phi}_d$ is as accurate as the conventional MUSIC extensions based on cumulants analysis with less computational complexity.

4. PROPOSED METHOD

To utilize the advantages of high-resolution estimation with high-ordered moments and robust estimation with low-ordered moments, we propose the new map for joint analysis of cross moments of multiple orders. The proposed map $\boldsymbol{\phi}_{d_1,\cdots,d_m}$ is given by a direct sum of maps $\boldsymbol{\phi}_{d_1},\cdots,\boldsymbol{\phi}_{d_m}$ as

$$\boldsymbol{\phi}_{d_1,\cdots,d_m} \triangleq \boldsymbol{\phi}_{d_1} \oplus \cdots \oplus \boldsymbol{\phi}_{d_m}, \tag{21}$$

$$\{\boldsymbol{\phi}_{d_1},\cdots,\boldsymbol{\phi}_{d_m}\} = \{\boldsymbol{\phi}_d \mid d = d_1,\cdots,d_m\}$$
$$d_{i+1} > d_i \quad (i = 1,\cdots,m-1), \tag{22}$$

where $\oplus$ denotes a direct sum. Mapped MUSIC with the map $\boldsymbol{\phi}_{d_1,\cdots,d_m}$ gives the joint analysis of multiple moments because the covariance matrix $\mathbf{R}^{\oplus}_{d_1,\cdots,d_m}(\omega)$ of $\boldsymbol{\phi}_{d_1,\cdots,d_m}$ contains the moments of multiple orders as

$$\mathbf{R}^{\oplus}_{d_1,\cdots,d_m}(\omega) = E\left[\boldsymbol{\phi}_{d_1,\cdots,d_m}(\mathbf{x}'(\omega,t))\boldsymbol{\phi}^{\mathrm{H}}_{d_1,\cdots,d_m}(\mathbf{x}'(\omega,t))\right]$$
$$= \begin{bmatrix} \mathbf{R}_{d_1,d_1}(\omega) & \mathbf{R}_{d_1,d_2}(\omega) & \cdots & \mathbf{R}_{d_1,d_m}(\omega) \\ \mathbf{R}_{d_2,d_1}(\omega) & \mathbf{R}_{d_2,d_2}(\omega) & \cdots & \mathbf{R}_{d_2,d_m}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{d_m,d_1}(\omega) & \mathbf{R}_{d_m,d_2}(\omega) & \cdots & \mathbf{R}_{d_m,d_m}(\omega) \end{bmatrix}, \tag{23}$$

$$\mathbf{R}_{d_i,d_j}(\omega) \triangleq E\left[\boldsymbol{\phi}_{d_i}(\mathbf{x}'(\omega,t))\boldsymbol{\phi}^{\mathrm{H}}_{d_j}(\mathbf{x}'(\omega,t))\right] = [r_{pq}]_{pq}$$
$$(i,j = 1,\cdots,m \quad p = 1,\cdots,M^{d_i}$$
$$q = 1,\cdots,M^{d_j}), \tag{24}$$

$$r_{pq} = E\left[\left(\prod_{l=1}^{d_i} x^{l\otimes}_{c_{pl}}\right)\left(\prod_{l=1}^{d_j} x^{l\otimes}_{c_{ql}}\right)^*\right]. \tag{25}$$

where $\mathbf{x}'(\omega,t)$ is an observed signal normalized properly (normalization rule is described later), and submatrices $\mathbf{R}_{d_i,d_j}(\omega)$ gives $(d_i + d_j)$-th order cross moments. The dimensionality of $\mathbf{R}^{\oplus}_{d_1,\cdots,d_m}(\omega)$ is $\sum_{i=1}^{m} M^{d_i}$, which is still larger than that of the highest-dimensional single cross moment matrix $\mathbf{R}_{d_m,d_m}(\omega)$. In addition, by analyzing the low-ordered moments together with the high-ordered moments, the significance of the statistical bias in the analysis of high-ordered moments is relaxed. With the increased dimensionality and reduced significance of statistical bias, the mapped MUSIC with the proposed map $\boldsymbol{\phi}_{d_1,\cdots,d_m}$ achieves further improvement of DOA estimation performance.

Note that we must take care of the scaling of the observed signal for the sufficient joint analysis of the moments of multiple orders. As we find in (23)–(25), $\mathbf{R}^{\oplus}_{d_1,\cdots,d_m}(\omega)$ contains moments of various orders, and it is obvious that the absolute values of the elements in the high-ordered moment matrices becomes significantly larger than those of the low-ordered moment matrices when the norm $||\mathbf{x}(\omega,t)||$ of the observation is large. In contrast, the absolute values of the elements in the high-ordered moments becomes considerably small when the norm $||\mathbf{x}(\omega,t)||$ is small. Since the eigen decomposition (9)–(12) takes the moments of the orders with large values into account more significantly, all the orders of moments must be weighted appropriately for the sufficient joint analysis of moments of multiple orders. Since such the proportion is defined by the norm, we normalize the observed signal with the L_{2d_m}-norm based on the maximum degree d_m within the proposed map $\boldsymbol{\phi}_{d_1,\cdots,d_m}$ in each frequency bins as

$$\mathbf{x}'(\omega,t) = \frac{w\mathbf{x}(\omega,t)}{\sqrt[2d_m]{E\left[\frac{1}{M}\sum_{i=1}^{M}|x_i(\omega,t)|^{2d_m}\right]}}, \tag{26}$$

where w is a parameter to adjust the L_{2d_m}-norm of the observed signal. With large value of w, the proposed method analyzes the high-ordered moments more significantly, and vice versa.

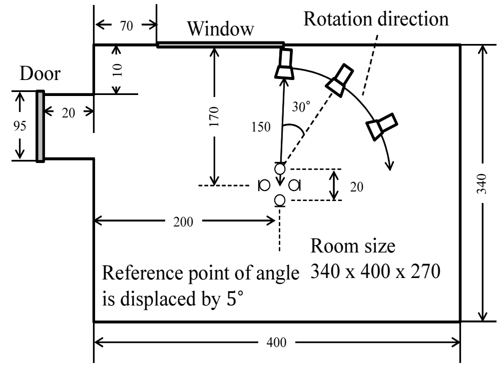

Figure 1: Recording environment.

5. EXPERIMENT

This section describes an experiment to verify the effectiveness of the proposed method. We evaluated mapped MUSIC with the proposed map $\boldsymbol{\phi}_{1,2,3}$, MUSIC, mapped MUSIC ($d = 2, 3$) and $2q$-MUSIC ($q = 2, 3$) are implemented for comparison.

5.1. Experimental condition

We conducted an evaluation of the DOA estimation of each method by using mixed signals that were created as convolutive mixtures of the Japanese speeches and the impulse responses, measured at the environment shown in Fig.1. To emulate noisy observation, diffused pink noises [9] was superimposed to the observed signals. To verify estimation performance for both overdetermined and underdetermined environments, we also evaluated the performance with two different number of sources. Table 1 shows the other experimental conditions.

Table 1: Experimental conditions

Sensor array configuration		Circular array with radius of 0.1 m	
Sound sources		Speakers 1.5 m apart from array	
# of sound sources	3, 5	Sampling frequency	16 kHz
# of sensors	4	SNR	20 dB
Room size		3.4 m × 4 m × 2.7 m	
Reverberation		T_{60} of 0.3 s	
Speech duration		1 s	
Frame length		512 samples	
Frame shift length		256 samples	
Window function		Hamming window	

For the evaluation, we employed the root mean squared error (RMSE) between the estimated direction and the true sound direction. We evaluated 100 combinations of the positions of three or five sources selected randomly from the directions $\{0^\circ, 30^\circ \cdots, 330^\circ\}$. When the estimation score had fewer peaks than sources, we added a penalty equal to the average error for all directions. This paper employed $N' = 15$ in (13)–(16) as the number of source is three, and $N' = 30$ for five sources. We also employed normalization parameter $w = 2$ in (26) for both of the number of sources. DOA estimation with MUSIC when $M \leq N$ is performed by regarding the one-dimensional subspace associate with the minimum eigenvalue as the noise subspace in every frequency bin.

5.2. Experimental results

Figures 2–3 show the experimental results obtained under different number of sources. Throughout all the conditions, the proposed method performs the best, and mapped MUSIC performs slightly better than $2q$-MUSIC. When we compare the results of mapped MUSIC $(d = 2)$ and that of $(d = 3)$, the former shows the better performance, and this tendency also appears as to $2q$-MUSIC. As examined in [3], these results suggest that the latter is biased greater than the former because of the shortage of snapshots, and the observation of one second is not enough to reduce the influence of statistical bias derived from the high-ordered statistics. While, the comparison between the proposed method and mapped MUSIC $(d = 2)$ denotes the effectiveness of joint analysis of moments of multiple orders. Although the high-ordered moments is inferior as for the sole use, we can confirm its usefulness to support the low-ordered moments. Furthermore, the proposed method can be expected to keep showing the best performance even if the time length becomes larger because the information of the high-ordered becomes more robust. From these discussion, the effectiveness of the proposed method utilizing the advantage of low-ordered and high-ordered moments simultaneously is verified.

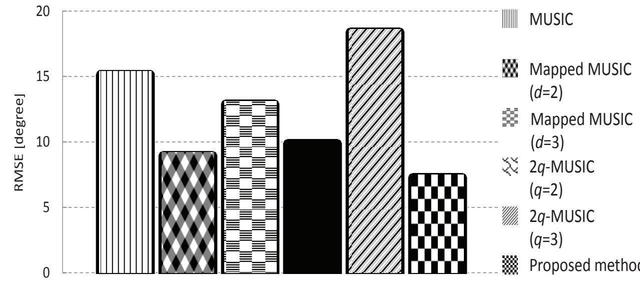

Figure 2: Experimental results for 3 sources within 1 second.

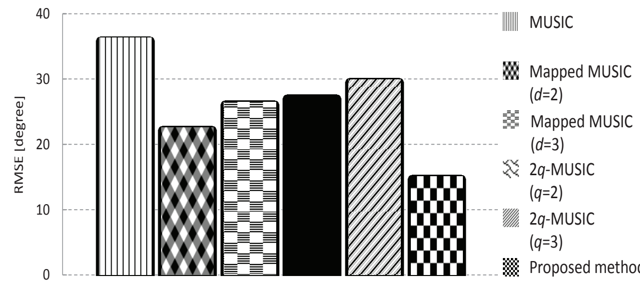

Figure 3: Experimental results for 5 sources within 1 second.

6. CONCLUSION

This paper proposed extended mapped MUSIC realizes to estimate DOAs with higher resolution than other conventional MUSIC extensions. We reviewed mapped MUSIC with the map for moments analysis of single arbitrary even order. We proposed new map as a direct sum of maps of multiple degrees, and showed that the proposed method corresponds to joint moments analysis of multiple orders taking advantage of both resolution with high-ordered moments and robustness with low-ordered moments. The experiment comparing with the other MUSIC extensions based on single high order statistics clarified the effectiveness of the joint analysis and excellent DOA estimation performance by the proposed method.

7. REFERENCES

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

[2] B. Porat and B. Friedlander, "Direction finding algorithms based on higher order statistics," *IEEE Trans. Signal Process.*, vol.39, no.9, pp.2016–2024, 1991.

[3] P. Chevalier, A. Ferreol and L. Albera, "High resolution direction finding from higher order statistics: The $2q$-MUSIC algorithm," *IEEE Trans. Signal Process.*, vol.54, no.8, pp.2986–2997, 2006.

[4] Y. Sugimoto, *et al.*, "Underdetermined DOA estimation by the non-linear MUSIC exploiting higher-order moments," *Proc. IWAENC*, 2012.

[5] H. Saruwatari, *et al.*, "Speech Enhancement Using Nonlinear Microphone Array Based on Complementary Beamforming," *IEICE Trans. Fundam.*, vol.E82-A, no.8, pp.1501–1510, 1999.

[6] H. Kamiyanagida, *et al.*, "Direction of Arrival Estimation Using Nonlinear Microphone Array," *IEICE Trans. Fundam.*, vol.E84-A, no.4, pp.999-1010, 2001.

[7] S. Miyabe, *et al.*, "Kernel-based nonlinear independent component analysis for underdetermined blind source separation," *Proc. ICASSP*, pp.1641–1644, 2009.

[8] M. Wax, T.-J. Shan and T.Kailath, "Spatio-temporal analysis by eigenstructure methods," *IEEE Trans. Acoust. Speech Signal Process.*, vol.32, pp.817–827, 1984.

[9] I. A. McCowan and H. Bourlard, "Microphone array postfilter based on noise field coherence," *IEEE Trans. Speech Audio Process.*, vol.11, no.6, pp.709–716, 2003.