COHERENCE-ADJUSTED MONOPOLE DICTIONARY AND CONVEX CLUSTERING FOR 3D LOCALIZATION OF MIXED NEAR-FIELD AND FAR-FIELD SOURCES

Tomoya Tachikawa, Kohei Yatabe, and Yasuhiro Oikawa

Department of Intermedia Art and Science, Waseda University, Tokyo, Japan

ABSTRACT

In this paper, 3D sound source localization method for simultaneously estimating both direction-of-arrival (DOA) and distance from the microphone array is proposed. For estimating distance, the off-grid problem must be overcome because the range of distance to be considered is quite broad and even not bounded. The proposed method estimates positions based on an extension of the convex clustering method combined with sparse coefficients estimation. A method for constructing a suitable monopole dictionary based on coherence is also proposed so that the convex clustering based method appropriately estimate distance of sound sources. Numerical experiments of distance estimation and 3D localization show possibility of the proposed method.

Index Terms— Sparsity, direction-of-arrival (DOA), distance estimation, convex optimization, primal-dual splitting.

1. INTRODUCTION

Sound source localization techniques using a microphone array [1,2] has been interested in many applications. Many of them assume far-field sources, and therefore plane waves are used as a dictionary for estimating the direction-of-arrival (DOA) of sound sources [3–10].

For estimation, DOA is usually discretized into a grid so that only several directions must be considered. However, such discretization can be a source of estimation error because only a direction which coincides with the grid is represented by the dictionary perfectly. This discretization issue is called the off-grid problem, and many methods have been proposed to overcome it [10–13]. It can be regarded as a trade-off between accuracy and computational cost: a fine grid provides better accuracy with more computational complexity.

In contrast to DOA estimation which only considers direction of sound sources, there are less research on 3D source localization which estimates both direction and distance from a microphone array [14–18]. For estimating distance, both near- and far-field sources must be taken into account simultaneously. Then, a natural dictionary for such distance estimation includes both monopoles and plane waves. Nevertheless, distance of plane waves may be regarded as infinity, and thus it is only possible to estimate distance of sources in the near-field by the dictionary. Therefore, a monopole-only dictionary is necessary to treat wide range of distance. However, the above mentioned off-grid problem becomes a quite serious problem for such a monopole dictionary because the number of 3D grid points grow extremely fast if wide range of distance is considered. It can be impossible to solve the large problem when one requires, say, one-meter accuracy of estimation within fifteen meter range. Although a greedy strategy as in [19] may be a choice for reducing the computational cost [14], local minima may prevent accurate estimation for some situations.

In this paper, a method for constructing an efficient monopole-only dictionary is proposed. By combining the proposed dictionary with a convex optimization formulation, it allows a simultaneous estimation of direction and distance without concerning local minima traps. Moreover, an extension of the convex clustering method [20–22] with a coordinate transformation is proposed in order to estimate source positions from many candidates of the positions. A convex optimization algorithm is also proposed to solve the clustering problem effectively.

2. SPARSE LOCALIZATION USING MONOPOLES

A sound signal at position $\mathbf{q} \in \Omega$ in a region $\Omega \subset \mathbb{R}^3$ can be approximated by linear combination as

$$y(\mathbf{q}, \omega) \simeq \sum_i \varphi_i(\mathbf{q}, \omega)\, x_i, \tag{1}$$

where $\varphi_i(\mathbf{q}, \omega)$ is an element of an overcomplete dictionary, ω is the angular frequency, and $x_i \in \mathbb{C}$ is a coefficient. For the elements of the dictionary, monopoles,

$$\varphi_i(\mathbf{q}, \omega) = \exp(jk \left\|\mathbf{q} - \mathbf{p}_i\right\|_2) \,/\, 4\pi \left\|\mathbf{q} - \mathbf{p}_i\right\|_2, \tag{2}$$

are considered here, where $j = \sqrt{-1}$, $k\ (= \omega/c)$ is the wave number, c is the speed of sound, $\mathbf{p}_i$ is position of a monopole, $\|\cdot\|_p$ is the ℓ_p-norm. Then, a observed signal of M microphones placed at $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_M\}$ can be written as

$$\mathbf{y}(\omega_l) = \mathbf{\Phi}(\omega_l)\mathbf{x}_l, \tag{3}$$

where $\mathbf{y}(\omega_l) = [\, y(\mathbf{q}_1, \omega_l),\, y(\mathbf{q}_2, \omega_l),\, \cdots,\, y(\mathbf{q}_M, \omega_l)\,]^{\mathrm{T}}$, $\mathbf{x}_l = [x_1, x_2, \cdots, x_N]^{\mathrm{T}}$, $\mathbf{x}^{\mathrm{T}}$ is transpose of $\mathbf{x}$, and $\mathbf{\Phi}(\omega_l) \in \mathbb{C}^{M\times N}$ is a matrix whose (j,i)-th element is $[\mathbf{\Phi}(\omega_l)]_{ji} = \varphi_i(\mathbf{q}_j, \omega_l)$. Sparsity based methods estimate source locations via sparse optimization [6–13, 23]. Although the proposed dictionary can be integrated into many other formulations, a group sparse formulation [23–27],

$$\min_{\mathbf{x}} \quad \left\|\mathbf{\Phi}\mathbf{x} - \mathbf{y}\right\|_2^2 + \lambda \left\|\mathbf{x}\right\|_{2,1}, \tag{4}$$

is adopted for this paper as an example. Here, $\mathbf{x}$ and $\mathbf{y}$ are vertically concatenated vectors of $\{\mathbf{x}_l\}$ and $\{\mathbf{y}(\omega_l)\}$, respectively, $\mathbf{\Phi}$ is the normalized block diagonal matrix whose diagonal block consists of $\{\tilde{\mathbf{\Phi}}(\omega_l)\}$, $\tilde{\mathbf{\Phi}}(\omega_l)$ is the column-wise normalized version of $\mathbf{\Phi}(\omega_l)$, $\lambda > 0$ is a regularization parameter, and $\left\|\mathbf{x}\right\|_{2,1} = \sum_{i=1}^{N} \left\|[(x_i)_1,\, (x_i)_2,\, \cdots,\, (x_i)_L]\right\|_2$.

After solving this convex optimization problem, positions corresponding to non-zero elements of estimated $\hat{\mathbf{x}}$ are treated as candidates of sound source locations. When the position of a source and one of the monopoles coincides, it can be directly estimated by choosing appropriate λ. However, such condition cannot be met in general, and thus estimation of off-grid position from non-zero coefficients is necessary.

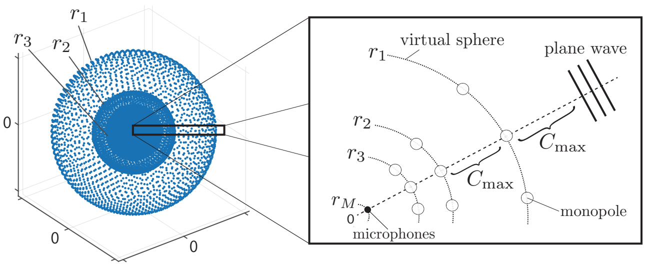

Fig. 1. Schematic diagram of coherence-adjusted monopoles where each monopole lies on a sphere whose radius is denoted by r_n.

3. PROPOSED METHOD

For estimating source location from its candidates, a simple extension of convex clustering is proposed. A primal-dual algorithm for efficiently computing it is also proposed. In addition, a method for constructing a coherence-adjusted dictionary, which is essential for distance estimation, is proposed.

3.1. Generating coherence-adjusted monopole dictionary

In the previous section, positions of monopoles $\{\mathbf{p}_i\}$ can be chosen arbitrarily. Then, it is natural to ask which choice is better than the others. Here, a method for obtaining better positions in terms of coherence is proposed.

Coherence C_{ij} between two functions φ_i and φ_j is a measure of similarity which takes a value in $[0, 1]$:

$$C_{ij} = \frac{|\langle \varphi_i, \varphi_j \rangle|}{\|\varphi_i\|_2 \|\varphi_j\|_2}, \tag{5}$$

where $|\cdot|$ is absolute value, and $\langle \cdot, \cdot \rangle$ is the standard inner product which is calculated from the vectors obtained by the microphones. It is known that a dictionary with smaller worst-case coherence,

$$C_{\max} = \max_{i \neq j} C_{ij}, \tag{6}$$

is more preferable for sparse estimation [27–29]. Therefore, monopoles should be distributed so that $C_{\max}$ is minimized among dictionaries having the same number of monopoles. However, it should be a quite difficult problem to solve because such position determination is highly non-convex. In addition, the far-field, whose distance from origin may be considered as infinity, must be treated specially.

In order to obtain a better placement of monopoles, a greedy method is proposed. For making the problem simple, several spheres, on which each set of monopoles lies, are considered as in Fig. 1. The number of monopoles and their distribution on a sphere are supposed to be given by the user. Then, only radii of the spheres $\{r_n\}$ have to be determined.

As $C_{\max}$ within a sphere can be calculated from given information, our aim is to adjust $\{r_n\}$ so that $C_{\max}$ between adjacent spheres becomes similar to that. For notational convenience, let $\mathcal{I}_{r_n}$ be an index set of monopoles lying on the sphere of radius r_n. Then, the proposed method adjusts the mean value of $C_{\max}$ between adjacent spheres (r_n and r_{n-1}), which can be written as

$$\frac{1}{|\mathcal{I}_{r_n}|} \sum_{i \in \mathcal{I}_{r_n}} \max_{j \in \mathcal{I}_{r_{n-1}}} C_{ij} = \frac{1}{|\mathcal{I}_{r_n}|} \sum_{i \in \mathcal{I}_{r_n}} \max_{j \in \mathcal{I}_{r_{n-1}}} \frac{|\langle \varphi_i, \varphi_j \rangle|}{\|\varphi_i\|_2 \|\varphi_j\|_2},$$

to have a similar value to $C_{\max}$ within a sphere. In the adjustment, the radial component r_n of monopoles on that sphere $\{\varphi_i\}_{i \in \mathcal{I}_{r_n}}$ is the variable while the direction (azimuth and polar angles) is constant.

Algorithm 1 Proposed method for adjusting monopoles

Input: $\{\mathbf{p}_i\}_{i \in \mathcal{I}_\infty}, r_M$
Output: $\{r_1, r_2, \cdots, r_{n-1}\}$
Calculate $C_{\max}^{\infty} = \max_{i,j \in \mathcal{I}_\infty, i \neq j} C_{ij}$.
Find r_1 $(<\infty)$ such that $\frac{1}{|\mathcal{I}_{r_1}|} \sum_{i \in \mathcal{I}_{r_1}} \max_{j \in \mathcal{I}_\infty} C_{i,j} = C_{\max}^{\infty}$.
Set $n = 1$.
while $r_n > r_M$ **do**
  Set $n = n + 1$
  Find r_n $(< r_{n-1})$ such that $\frac{1}{|\mathcal{I}_{r_n}|} \sum_{i \in \mathcal{I}_{r_n}} \max_{j \in \mathcal{I}_{r_{n-1}}} C_{i,j} = C_{\max}^{\infty}$.
end while

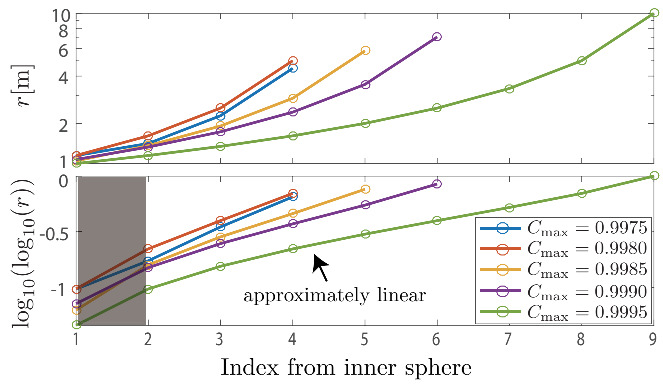

Fig. 2. Examples of obtained $\{r_n\}$ by Algorithm 1. Note that the horizontal axis represents index from the inner sphere for visibility which is opposite to those of Fig. 1 and Algorithm 1.

The reason of taking the mean is to defuse dependency of coherence on the relative position between monopoles and microphones.

For taking the far-field into account, plane waves are considered first, where the propagating directions are denoted by $\{\mathbf{p}_i\}_{i \in \mathcal{I}_\infty}$. Then, spheres with radii $\{r_n\}$ are repeatedly added from outer to inner until the smallest sphere reaches to the microphone array. The proposed algorithm is summarized in Algorithm 1, where r_M is the radius of enclosing sphere of the array. Note that finding r_n can be achieved easily by a simple line search because it is single-variable optimization and coherence varies smoothly.

After $\{r_n\}$ is obtained, the plane waves are replaced by monopoles because plane waves are not suitable for distance estimation. Figure 2 shows obtained $\{r_n\}$ by Algorithm 1 for several settings. It can be seen that $\{r_n\}$ can be regarded as sampled points of an approximately linear function in the doubly logarithmic scale. Therefore, the plane waves ($r = \infty$) are approximated by monopoles lying on a sphere of radius r_0,

$$r_0 = 10^{[\log_{10}(r_1)]^2 / \log_{10}(r_2)}, \tag{7}$$

which is linear extrapolation in double logarithmic domain [1].

An example of obtained radii which will be used for the simulation is $R_{\mathrm{co}} = \{1.34, 1.78, 2.68, 5.37, 17.44\}$ [m] which is illustrated by the red line in Fig. 2. Since the interval of outer spheres is notably large, the off-grid problem arises as a serious issue for distance estimation.

[1] One may try to find r_0 $(> r_1)$ such that their maximal coherence coincide with $C_{\max}^{\infty}$, which can be regarded as applying Algorithm 1 reversely. However, we empirically found such strategy can easily fail for some situations where the condition cannot be met. In contrast, the extrapolation can always be performed without concerning any condition.

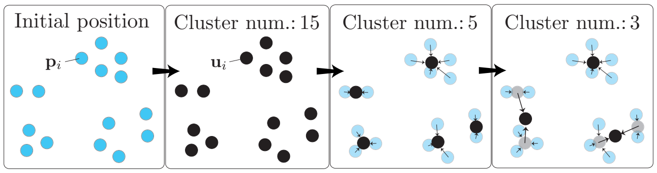

Fig. 3. Schematic explanation of convex clustering. Blue circles represent $\{\mathbf{p}_i\}$, while black circles are $\{\mathbf{u}_i\}$. The second block is for $\gamma = 0$, while the right two blocks correspond to larger γ.

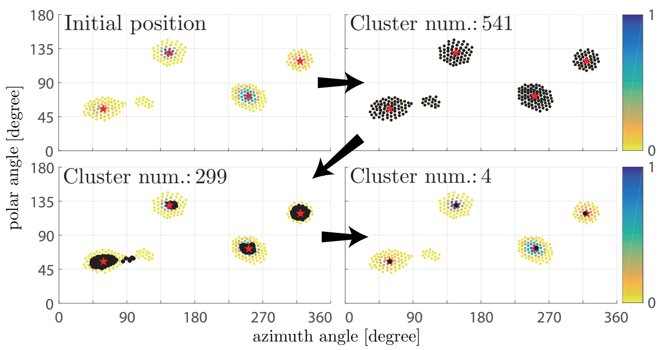

Fig. 4. An example of DOA using the extended convex clustering. The magnitude of monopoles correspond to non-zero coefficients are represented by the color. The red stars are true positions of the sound sources, while the black circles are estimated positions which are listed in Table 2.

3.2. Extension of convex clustering for 3D localization

For estimating off-grid source location, an extension of convex clustering is proposed. Before the estimation, the following coordinate transformation is necessary:

$$\mathbf{p}_i = \frac{\min\{d_1, d_2\}}{\max\{d_1, d_2\}} \frac{\mathbf{p}_i}{\|\mathbf{p}_i\|_2} \Big[\log_{10}\big(\log_{10}(\|\mathbf{p}_i\|_2)\big) + r_M\Big], \quad (8)$$

where d_1 and d_2 are respectively minimum distance of monopoles within a sphere and between the spheres. This transformation ensures the directional components and the radial components in the doubly logarithmic domain have similar order so that the following method equally treat them.

Then, the proposed method is formulated as the following convex optimization problem

$$\min_{\{\mathbf{u}_i\}} \frac{1}{2} \sum_{i \in \mathcal{I}_{\hat{x}\neq 0}} w_i \|\mathbf{p}_i - \mathbf{u}_i\|_2^2 + \gamma \sum_{(n,m)\in\mathcal{N}} \chi_{(n,m)} \|\mathbf{u}_n - \mathbf{u}_m\|_2, \quad (9)$$

where $\mathcal{I}_{\hat{x}\neq 0}$ is the index set of non-zero elements of estimated coefficient $\hat{\mathbf{x}}$ in Eq. (4), $\{\mathbf{p}_i\}_{i\in\mathcal{I}_{\hat{x}\neq 0}}$ is the set of positions of monopoles corresponding to the non-zero elements, $\{\mathbf{u}_i\}$ is a copy of $\{\mathbf{p}_i\}$ to be optimized, $\gamma > 0$ is a regularization parameter, $\mathcal{N}$ is an index set of adjacency, and w_i and $\chi_{(n,m)}$ are positive weights which play important roles.

This formulation is an extension of the convex clustering method [20,21] which does not have $\{w_i\}$ in the first term. As in the original convex clustering, k-nearest-neighbor is considered in the weight as

$$\chi_{(n,m)} = \kappa^k_{(n,m)} \exp\big(-\psi\|\mathbf{p}_n - \mathbf{p}_m\|_2^2\big), \quad (10)$$

where ψ is positive constant, and $\kappa^k_{(n,m)}$ is 1 if $\mathbf{p}_n$ is k-nearest-neighbor of $\mathbf{p}_m$ and 0 otherwise. Note that this weights can be regarded as elements of the adjacency matrix of the monopoles: in that case, monopoles can be considered as a graph signal [30]. In the proposed method, the weight $\{w_i\}$ is also considered in order to take the estimation result of Eq. (4) into account. Here, a weighting rule,

$$w_i = r_i\,|\hat{x}_i|\,, \quad (11)$$

is proposed, where $\hat{x}_i$ is the non-zero element of estimated coefficient in Eq. (4), and r_i is the radius of the sphere on which the corresponding monopole lies. Figure 3 schematically illustrates process of the proposed method. The second term of Eq. (9) tries to sparsify the distance of monopoles, while the first term keeps the positions close to the original ones. These effects result in concentrated points which will be treated as the estimated locations. An example of DOA estimation for four sources, which is the same condition as the simulation in Table 2, by the proposed method is shown in Fig. 4. It can be seen that the proposed method can correctly localize the directions.

Algorithm 2 Primal-dual splitting method for solving (9)

Input: $\{w_i\}_{i\in\mathcal{I}_{\hat{x}\neq 0}}$, $\{\mathbf{p}_i\}_{i\in\mathcal{I}_{\hat{x}\neq 0}}$, $\mathbf{L}$, μ_1, μ_2
Output: $\mathbf{u}^{[t+1]}$
Set $\mathbf{u}_i^{[1]} = \mathbf{p}_i$, $\mathbf{v}^{[1]} = \mathbf{0}$
for $t = 1, 2, 3, \cdots$ **do**
  $\boldsymbol{\xi}^{[t]} = \mathbf{u}^{[t]} - \mu_1 \mathbf{L}^{\mathrm{T}} \mathbf{v}^{[t]}$
  for $i = 1, 2, 3, \cdots$ **do**
    $\mathbf{u}_i^{[t+1]} = \frac{1+\mu_1 w_i}{\mu_1}(w_i \mathbf{p}_i + \frac{1}{\mu_1}\boldsymbol{\xi}_i^{[t]})$
  end for
  $\boldsymbol{\zeta}^{[t]} = \mathbf{v}^{[t]} - \mu_2 \mathbf{L}(2\mathbf{u}^{[t+1]} - \mathbf{u}^{[t]})$
  for $j = 1, 2, 3, \cdots$ **do**
    $\mathbf{v}_j^{[t+1]} = \min\{1, \gamma/\|\boldsymbol{\zeta}_j\|_2\}\, \boldsymbol{\zeta}_j$
  end for
end for

To obtain an estimated result as in Fig. 4, one needs to solve the above convex optimization problem. In this paper, the recent primal-dual algorithm is adopted for solving it [31,32]. The primal-dual splitting, which is an iterative procedure of the following form

$$\boldsymbol{\xi}^{[t+1]} = \mathrm{prox}_{\mu_1 f}\left[\, \boldsymbol{\xi}^{[t]} - \mu_1 \mathbf{L}^{\mathrm{T}} \boldsymbol{\zeta}^{[t]} \,\right] \quad (12)$$

$$\boldsymbol{\zeta}^{[t+1]} = \mathrm{prox}_{\mu_2 g^*}\left[\, \boldsymbol{\zeta}^{[t]} - \mu_2 \mathbf{L}(2\boldsymbol{\xi}^{[t+1]} - \boldsymbol{\xi}^{[t]}) \,\right], \quad (13)$$

can solve a convex optimization problem of the following form:

$$\min_{\boldsymbol{\xi}}\ f(\boldsymbol{\xi}) + g(\mathbf{L}\boldsymbol{\xi}), \quad (14)$$

where f and g are (proper and lower-semicontinuous) convex functions, t is the iteration index, $\mathbf{L}$ is a linear operator, μ_1 and μ_2 are the constants satisfying $\mu_1\mu_2\|\mathbf{L}\|_{\mathrm{op}} \leq 1$ ($\|\cdot\|_{\mathrm{op}}$ stands for the operator norm). $\mathrm{prox}_f[\cdot]$ denotes the proximity operator of f [33],

$$\mathrm{prox}_f[\,\mathbf{y}\,] = \arg\min_{\mathbf{x}} \Big[\, f(\mathbf{x}) + \frac{1}{2}\,\|\mathbf{y} - \mathbf{x}\|_2^2 \,\Big], \quad (15)$$

and g^* is the convex conjugate of g, whose proximity operator can be computed by the following Moreau decomposition formula:

$$\mathrm{prox}_{\mu g^*}[\,\mathbf{y}\,] = \mathbf{y} - \mu\, \mathrm{prox}_{\frac{1}{\mu} g}\Big[\, \frac{1}{\mu}\mathbf{y} \,\Big]. \quad (16)$$

Let $\mathbf{u}$ is the vector obtained by vertically concatenating all $\mathbf{u}_i$, and $\mathbf{L}$ be the linear operator mapping $\mathbf{u} \mapsto \mathbf{v}$, where $\mathbf{v}$ is the vector obtained by vertically concatenating $\{\chi_{(n,m)}(\mathbf{u}_n - \mathbf{u}_m)\}_{(n,m)\in\mathcal{N}}$. Then, Eq. (9) can be viewed as a convex optimization problem in the form of Eq. (14). Therefore, the primal-dual splitting algorithm can be applied to obtain an iterative procedure for solving Eq. (9). The proposed primal-dual algorithm is summarized in Algorithm 2.

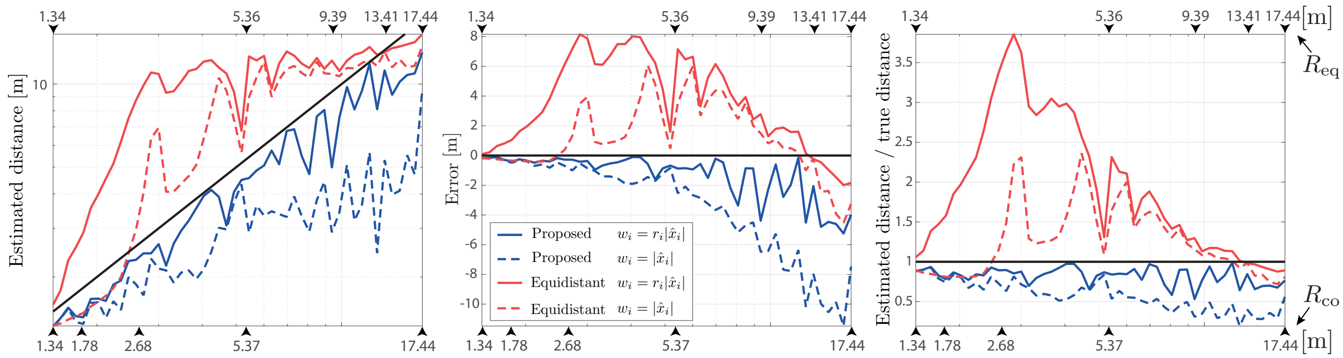

Fig. 5. Distance estimation for a single source. A coherence-adjusted monopole dictionary was obtained by the proposed method whose radii of the spheres are $R_{\text{co}} = \{1.34, 1.78, 2.68, 5.37, 17.44\}$ [m]. For comparison, equally spaced set $R_{\text{eq}} = \{1.34, 5.36, 9.39, 13.41, 17.44\}$ [m] (4.025 m interval) was also used. The horizontal axes represent true position of the source. The vertical axis of the left figure shows estimated position. The figure in the center shows error of the estimation, while the right one is relative error normalized by the true value.

Table 1. The condition of simulation.

Number of microphones	$M = 8$
Number of monopoles on each sphere	2401
Calculated coherence	$C^{\infty}_{\text{max}} = 0.998$
Regularization parameter in Eq. (4)	$\lambda = 0.005$
Number of nearest neighbor	$k = 100$
Size parameter in Eq. (10)	$\psi = 0.5$

Table 2. 3D localization result for four sound sources. Each position is represented by the spherical coordinate, where radial distance r is in meter while azimuth and polar angles are in degrees. Left columns are the true positions, and right are the estimated ones.

	r_{true}	azimuth	polar	r_{est}	azimuth	polar
1	7.07	59.03	55.55	9.17	59.37	55.00
2	4.69	146.31	129.76	4.23	147.53	129.87
3	3.31	251.56	72.45	2.78	252.87	72.93
4	10.48	319.40	118.47	9.43	318.03	118.97

4. NUMERICAL EXPERIMENTS

Performance of the proposed method was investigated numerically. Simulation condition for both results shown below is listed in Table 1. The eight microphones were located at $(0.15, 0.15, 0.21)$, $(-0.15, 0.15, 0.21)$, $(-0.15, -0.15, 0.21)$, $(0.15, -0.15, 0.21)$, $(0.25, 0.25, -0.35)$, $(-0.25, 0.25, -0.35)$, $(-0.25, -0.25, -0.35)$, and $(0.25, -0.25, -0.35)$.

Firstly, appropriateness of the proposed coherence-adjusted monopole dictionary was examined by comparison with non-adjusted dictionary. Result of distance estimation for a single source is shown in Fig. 5, where γ was chosen automatically so that the number of sources becomes one. From the figure, it is confirmed that only combination of the proposed dictionary with the proposed weighting rule can properly estimate distance of the source.

Next, 3D positions of four sound sources are estimated. The true and estimated positions with $\gamma = 0.012$ are summarized in Table 2, which are illustrated in Figs. 4 and 6. It can be seen that the proposed method can reasonably estimate positions of multiple sound sources simultaneously. Although γ in this experiment was chosen so that correct number of sources was obtained, the proposed method has possibility of providing number of sources automatically. Since the convex optimization problem in Eq. (9) is strongly convex, the solution for chosen γ is unique. Therefore, the optimization procedure can be regarded as mapping form γ to number of sound sources. As

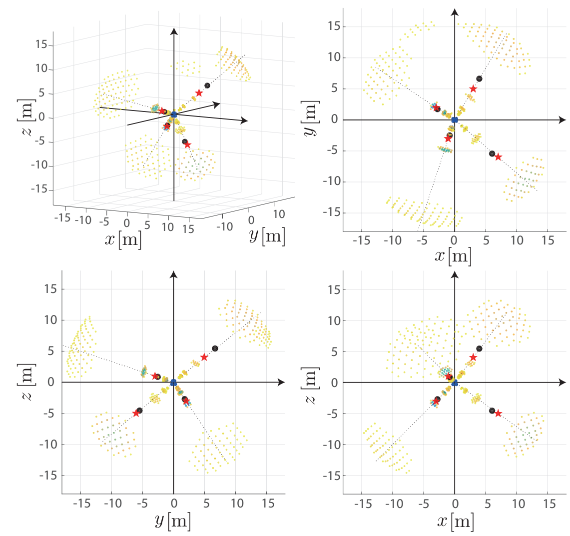

Fig. 6. 3D localization result for four sources as in Table 1. True positions are depicted by red stars whose direction are shown by black dotted lines. Estimated positions are shown by black circles.

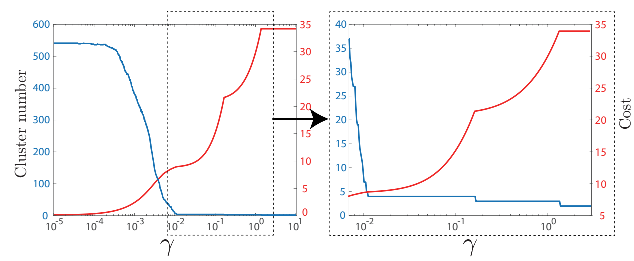

Fig. 7. Number of clusters and the least squares cost [the first term of Eq. (9)] with respect to γ for the 3D localization of four sources.

in Fig. 7, the least squares cost [first term in Eq. (9)] rapidly increases around true number of sources. This property should be useful for automatic detection of number of sources.

5. CONCLUSIONS

In this paper, for 3D source localization, a method for constructing a coherence-adjusted monopole dictionary and a method of estimating position based on convex clustering are proposed. For the future works, automatic detection of number of sources should be investigated as discussed in the last section.

6. REFERENCES

[1] Y. Koyano, K. Yatabe, Y. Ikeda, and Y. Oikawa, "Physical-model based efficient data representation for many-channel microphone array," in *2016 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar 2016, pp. 370–374.

[2] Y. Koyano, K. Yatabe, Y. Ikeda, and Y. Oikawa, "Infinite-dimensional SVD for analyzing microphone array," in *2017 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar 2017.

[3] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul 1996.

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

[5] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul 1989.

[6] I.F. Gorodnitsky and B.D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar 1997.

[7] J. Yin and T. Chen, "Direction-of-arrival estimation using a sparse representation of array covariance vectors," *IEEE Trans. on Signal Process.*, vol. 59, no. 9, pp. 4489–4493, Sep 2011.

[8] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *J. Acoust. Soc. Am*, vol. 136, no. 1, pp. 260–271, 2014.

[9] M.M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4646–4655, Sep 2010.

[10] D. Malioutov, M. Cetin, and A.S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug 2005.

[11] M. Ibrahim, F. Römer, R. Alieiev, G. Del Galdo, and R.S. Thomä, "On the estimation of grid offsets in CS-based direction-of-arrival estimation," in *2014 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 6776–6780.

[12] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan 2013.

[13] H. Jamali-Rad and G. Leus, "Sparsity-aware multi-source TDOA localization," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4874–4887, Oct 2013.

[14] L. Kumar, K. Singhal, and R.M. Hegde, "Near-field source localization using spherical microphone array," in *Hands-free Speech Commun. Microphone Arrays (HSCMA), 2014 4th Jt. Workshop*, May 2014, pp. 82–86.

[15] Q. Huang and T. Wang, "Acoustic source localization in mixed field using spherical microphone arrays," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–16, 2014.

[16] J. Liang and D. Liu, "Passive localization of mixed near-field and far-field sources using two-stage MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 108–120, Jan 2010.

[17] S. I. Adalbjörnsson, T. Kronvall, S. Burgess, K. Âström, and A. Jakobsson, "Sparse localization of harmonic audio sources," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 1, pp. 117–129, Jan 2016.

[18] J.J. Jiang, F.J. Duan, J. Chen, Y.C. Li, and X.N. Hua, "Mixed near-field and far-field sources localization using the uniform linear sensor array," *IEEE Sens. J.*, vol. 13, no. 8, pp. 3136–3143, Aug 2013.

[19] Y. Oikawa and Y. Yamasaki, "Direction of arrival estimation using matching pursuit under reverberant condition," *Acoust. Sci. Tech.*, vol. 26, no. 4, pp. 365–367, 2005.

[20] T.D. Hocking, A. Joulin, F. Bach, and J.-P. Vert, "Clusterpath an algorithm for clustering using convex fusion penalties," in *28th int. conf. mach. learn.*, United States, June 2011, p. 1.

[21] E.C. Chi and K. Lange, "Splitting methods for convex clustering," *J. Comput. Graph. Stat.*, vol. 24, no. 4, pp. 994–1013, 2015.

[22] K.M. Tan and D. Witten, "Statistical properties of convex clustering," *Electron. J. Stat.*, vol. 9, no. 2, pp. 2324–2347, 2015.

[23] Z. Liu, X. Wang, G. Zhao, G. Shi, J. Lin, and Z. Gao, "Wideband DOA estimation based on sparse representation — an extension of ℓ_1-SVD in wideband cases," in *Signal Process., Commun. Comput. (ICSPCC), 2013 IEEE Int. Conf.*, Aug 2013, pp. 1–4.

[24] L. Meier, S. van de Geer, and P. Bühlmann, "The group lasso for logistic regression," *J. Roy. Stat. Soc. B*, vol. 70, no. 1, pp. 53–71, 2008.

[25] K. Yatabe and Y. Oikawa, "Optically visualized sound field reconstruction using Kirchhoff-Helmholtz equation," *Acoust. Sci. Tech.*, vol. 36, no. 4, pp. 351–354, 2015.

[26] K. Yatabe and Y. Oikawa, "Optically visualized sound field reconstruction based on sparse selection of point sound sources," in *2015 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr 2015, pp. 504–508.

[27] X. Lv, G. Bi, and C. Wan, "The group lasso for stable recovery of block-sparse signal representations," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1371–1382, Apr 2011.

[28] Y.C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun 2010.

[29] D.L. Donoho, M. Elad, and V.N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan 2006.

[30] A. Sandryhaila and J.M.F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr 2013.

[31] N. Komodakis and J.C. Pesquet, "Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 31–54, Nov 2015.

[32] R.I. Bot, E.R. Csetnek, A. Heinrich, and C. Hendrich, "On the convergence rate improvement of a primal-dual splitting algorithm for solving monotone inclusion problems," *Mathematical Programming*, vol. 150, no. 2, pp. 251–279, 2015.

[33] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Opt.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.