結晶型マイクロフォンアレイを用いた 残響環境下における楽器音分離

伊藤信貴^{†1,†2}北野佑^{†1} 小野順貴^{†1}嵯峨山茂樹^{†1}

本稿では、残響のある室内でマイクロフォンアレイを用いて録音された音楽信号を 各楽器音に分離する問題を扱う。これを実現するために、我々がこれまで拡散性雑音 の抑圧のために研究してきた、結晶型マイクロフォンアレイと呼ぶ対称なマイクロフォ ン配置を用いる信号処理手法を応用する。提案法を残響環境下での楽器音分離に適用 する実験結果を示す。

Instrument Separation in Reverberant Environments Using Crystal Microphone Arrays

NOBUTAKA ITO ,^{$\dagger 1,\dagger 2$} YU KITANO,^{$\dagger 1$} NOBUTAKA ONO^{$\dagger 1$} and Shigeki SAGAYAMA^{$\dagger 1$}

This paper deals with instrument separation from a music signal recorded with a microphone array in a reverberant environment. To this end, we apply our signal processing method for suppressing diffuse noise, which utilizes symmetrical microphone arrangements called crystal microphone arrays. We show some results of experiments where the proposed method is applied to instrument separation in a reverberant environment.

1. はじめに

音響信号処理の分野では、マイクロフォンアレイと呼ばれる複数のマイクロフォンを用い

†1 東京大学大学院情報理工学系研究科

Graduate School of Information Science and Technology, The University of Tokyo †2 METISS group, IRISA-INRIA, France た信号処理により、様々な音が存在する環境で所望の方向の音を強調する技術が研究されて きた¹⁾。そのような技術は音楽信号に対しても様々な応用が期待される。例えば、クラシッ クコンサートをライブ録音する際に雑音を除去したり、アマチュアバンドや合唱団の練習を 各楽器音/歌声に分離して個別に確認したり、ポータブルメディアプレーヤのためのハミン グ検索²⁾において性能低下の原因となりうる雑音を除去したり、といった応用が期待され る。本稿では、マイクロフォンアレイ信号処理技術の音楽信号への応用の一例として、残響 のある室内で録音された音楽信号を各楽器音に分離する問題を扱う。

上述のようなマイクロフォンアレイ信号処理技術として最も基本的な遅延和ビームフォー マと呼ばれる方法では、各マイクロフォンによる観測音を、目的音成分が同期するように 時間シフトした後に加算することにより、目的音方向に指向性を形成する。この方法では、 鋭い指向性を実現するためにはアレイサイズを大きくする必要がある。一方、適応ビーム フォーマと呼ばれる手法では、観測音に対し環境に適応したフィルタを適用することによ り、マイクロフォン数-1個以下の方向に指向性の零点を向けることができるため、原理的 にはアレイサイズに関わらず少数方向から到来する雑音を除去することができる。しかしな がら、残響のある室内では壁における反射波が等価的な雑音源となって様々な方向から雑音 が到来する。このようないわゆる拡散性の雑音環境では、全ての雑音方向に指向性の零点を 向けることはできないため、適応ビームフォーマによる雑音抑圧には限界がある。

一方、拡散性雑音の抑圧に適する枠組みとして、上述のような従来型ビームフォーマによる処理の後で、更に時間周波数マスクと呼ばれる時間周波数平面上のフィルタを乗算する、いわゆるポストフィルタリングが提案されている³⁾。特に、最小分散ビームフォーマと呼ばれる適応ビームフォーマの出力に対し Wiener ポストフィルタと呼ばれる時間周波数マスクを適用する処理が、平均二乗誤差規範で最適であることが示されている⁴⁾。この Wiener ポストフィルタを設計するためには、雑音を含む観測音を用いて目的音のパワースペクトログラムを精確に推定することが重要である。その際、アレイサイズが波長に対して小さい場合のチャネル間の雑音相関が従来の推定法の誤差要因であった^{3),5),6)}が、我々は結晶型マイクロフォンアレイと呼ぶ対称なマイクロフォン配置を用いて雑音を無相関化することにより、アレイサイズに関わらず目的音のパワースペクトログラムを精確に推定し、拡散性雑音を効果的に抑圧することが可能な手法を提案した⁷⁾。

本稿では、この雑音抑圧法を用いた残響環境下における音楽信号の楽器音への分離につい て検討する。本稿の構成は以下のようになっている。まず、2節で結晶型マイクロフォンア レイを用いた拡散性雑音抑圧法を紹介する。次に、3節で本手法による残響環境下での楽器 情報処理学会研究報告 IPSJ SIG Technical Report

音分離実験について述べ、最後に4節で結論を述べる。

2. 結晶型マイクロフォンアレイを用いた拡散性雑音抑圧法

2.1 記 法

本稿では以下の記法を用いる。上付き記号*、^T、^Hはそれぞれ複素共役、転置、Hermite 転置を表す。信号は時間周波数領域で表現し、フレーム番号を $\tau \in \mathbb{Z}$ 、角周波数を $\omega \in \mathbb{R}$ で表す。スカラーの信号 $\alpha(\tau, \omega)$ 、 $\beta(\tau, \omega)$ のクロススペクトログラムを

 $\phi_{\alpha\beta}(\tau,\omega) \triangleq \mathcal{E}[\alpha(\tau,\omega)\beta^*(\tau,\omega)] \tag{1}$

で表し、零平均の確率ベクトル $m{\gamma}(au,\omega)$ の共分散行列を

 $\boldsymbol{\Phi}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}(\tau,\omega) \triangleq \mathcal{E}[\boldsymbol{\gamma}(\tau,\omega)\boldsymbol{\gamma}^{\mathsf{H}}(\tau,\omega)] \tag{2}$

で表す。ここで、 *E*[·] は期待値を表す。

2.2 観測モデル

本稿では、拡散性雑音の存在下で、既知の方向から到来する目的音をマイクロフォンアレ イにより観測する状況を考える。 $s(\tau, \omega)$ を基準点(マイクロフォン配置の重心など)におけ る目的音、 $x_m(\tau, \omega)$ および $v_m(\tau, \omega)$ をm番目のマイクロフォンにおける観測音および拡散 性雑音成分とする。目的音が平面波として伝播するという仮定の下、 $s(\tau, \omega)$ から $x_m(\tau, \omega)$ への伝達関数 $d_m(\omega)$ は次式で近似できる。

$$d_m(\omega) = \mathrm{e}^{-\mathrm{j}\omega\delta_m} \tag{3}$$

ここで、 δ_m は基準点から *m* 番目のマイクロフォンまでの目的音の到来時間遅れであり、目 的音の到来方向とマイクロフォン位置(いずれも既知と仮定)から計算することができる。 以上から、観測モデルは次式で与えられる。

 $x_m(\tau,\omega) = s(\tau,\omega)d_m(\omega) + v_m(\tau,\omega)$ (4)

上式は次のようにベクトル形式で書き直せる。

 $\boldsymbol{x}(\tau,\omega) = \boldsymbol{s}(\tau,\omega)\boldsymbol{d}(\omega) + \boldsymbol{v}(\tau,\omega)$ (5)

ここで、各ベクトルは以下のように定義される。

$$\boldsymbol{x}(\tau,\omega) \triangleq [x_1(\tau,\omega) \ x_2(\tau,\omega) \ \dots \ x_M(\tau,\omega)]^{\mathsf{T}}$$
(6)

$$\boldsymbol{d}(\omega) \triangleq [d_1(\omega) \ d_2(\omega) \ \dots \ d_M(\omega)]^\mathsf{T}$$
(7)

$$\boldsymbol{v}(\tau,\omega) \triangleq \left[v_1(\tau,\omega) \ v_2(\tau,\omega) \ \dots \ v_M(\tau,\omega) \right]^\mathsf{T}$$
(8)

ここで M はマイクロフォンの個数である。 $s(\tau, \omega)$ と $v(\tau, \omega)$ は零平均、無相関の確率変数 として扱う。

2.3 Wiener ポストフィルタ

目的音の線形推定量

$$\hat{s}(\tau,\omega) \triangleq \boldsymbol{w}^{\mathsf{H}}(\tau,\omega)\boldsymbol{x}(\tau,\omega) \tag{9}$$

のうち、平均二乗誤差

$$\mathcal{E}[|\hat{s}(\tau,\omega) - s(\tau,\omega)|^2]$$
(10)
を最小化するものは次式で与えられる^{4),8)}。

$$\hat{s}_{o}(\tau,\omega) \triangleq \mathcal{E}[s(\tau,\omega)\boldsymbol{x}^{\mathsf{H}}(\tau,\omega)]\mathcal{E}[\boldsymbol{x}(\tau,\omega)\boldsymbol{x}^{\mathsf{H}}(\tau,\omega)]^{-1}\boldsymbol{x}(\tau,\omega)$$
(11)

$$=\phi_{ss}(\tau,\omega)\boldsymbol{d}^{\mathsf{H}}(\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{x}(\tau,\omega)$$
(12)

ここで、式 (1)、(2) における定義から、 $\phi_{ss}(au,\omega)$ は目的音のパワースペクトログラム

$$\phi_{ss}(\tau,\omega) \triangleq \mathcal{E}\left[|s(\tau,\omega)|^2\right] \tag{13}$$

を表し、 $\Phi_{xx}(\tau,\omega)$ は観測信号の共分散行列

$$\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}(\tau,\omega) \triangleq \mathcal{E}[\boldsymbol{x}(\tau,\omega)\boldsymbol{x}^{\mathsf{H}}(\tau,\omega)]$$
(14)

を表すことに注意する。

次に、式 (12) の推定量 $\hat{s}_{o}(\tau,\omega)$ と最小分散ビームフォーマの出力

$$y(\tau,\omega) \triangleq \frac{\boldsymbol{d}^{\mathsf{H}}(\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{x}(\tau,\omega)}{\boldsymbol{d}^{\mathsf{H}}(\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{d}(\omega)}$$
(15)

の関係を示す。上式より $y(\tau, \omega)$ のパワースペクトログラムは次のように計算できる。

$$\phi_{yy}(\tau,\omega) = \frac{\boldsymbol{d}^{\mathsf{H}}(\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}(\tau,\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{d}(\omega)}{[\boldsymbol{d}_{1}^{\mathsf{H}}(\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{d}(\omega)]^{2}}$$
(16)

$$= \frac{1}{\boldsymbol{d}^{\mathsf{H}}(\omega)\boldsymbol{\Phi}_{\boldsymbol{x}\boldsymbol{x}}^{-1}(\tau,\omega)\boldsymbol{d}(\omega)}$$
(17)

式 (12)、(17) より、 $\hat{s}_{o}(\tau,\omega)$ は次のように書き直せる^{4),8)}。

$$\hat{s}_{o}(\tau,\omega) = \underbrace{\frac{\phi_{ss}(\tau,\omega)}{\phi_{yy}(\tau,\omega)}}_{\triangleq p(\tau,\omega)} \cdot \underbrace{\frac{d^{\mathsf{H}}(\omega)\Phi_{xx}^{-1}(\tau,\omega)x(\tau,\omega)}{d^{\mathsf{H}}(\omega)\Phi_{xx}^{-1}(\tau,\omega)d(\omega)}}_{= y(\tau,\omega)}$$
(18)

即ち $\hat{s}_{o}(\tau,\omega)$ は、最小分散ビームフォーマの出力 $y(\tau,\omega)$ を、Wiener ポストフィルタと呼ばれる時間周波数マスク $p(\tau,\omega)$ により処理すれば得られる。

2.4 結晶型マイクロフォンアレイを用いた $\phi_{ss}(\tau,\omega)$ 推定法

Wiener ポストフィルタ $p(\tau, \omega)$ の設計においては、分子の目的音のパワースペクトログ ラム $\phi_{ss}(\tau, \omega)$ を、雑音を含む観測音 $x(\tau, \omega)$ から精確に推定することが重要である。本小 節では、結晶型マイクロフォンアレイと呼ぶ対称なマイクロフォンアレイを用いた推定法⁷⁾ 情報処理学会研究報告 IPSJ SIG Technical Report

を紹介する。

 $s(\tau,\omega) \geq \boldsymbol{v}(\tau,\omega)$ の無相関性より、

 $\Phi_{xx}(\tau,\omega) = \phi_{ss}(\tau,\omega)d(\omega)d^{\mathsf{H}}(\omega) + \Phi_{vv}(\tau,\omega)$ (19) が成り立つ。ここで、 $\Phi_{xx}(\tau,\omega)$ は観測音から計算可能な既知の行列、 $\phi_{ss}(\tau,\omega)$ は推定す べき目的音のパワースペクトログラム(スカラ)、 $d(\omega)d^{\mathsf{H}}(\omega)$ は既知の $d(\omega)$ を用いて計 算できる既知の行列、 $\Phi_{vv}(\tau,\omega)$ は未知の行列である。我々の着眼点は、雑音共分散行列 $\Phi_{vv}(\tau,\omega)$ を対角化することができれば、観測信号の共分散行列 $\Phi_{xx}(\tau,\omega)$ の非対角成分は ノイズフリーとなるため、それらを用いて $\phi_{ss}(\tau,\omega)$ を精確に推定することができるはずで ある、というものである。ここで、被対角化行列 $\Phi_{vv}(\tau,\omega)$ は未知であるから、 $\Phi_{vv}(\tau,\omega)$ の値に依らない「ブラインドな」対角化が必要である。

そのために我々は、拡散性雑音場の空間的特性に着目し、次の条件を満たす等方性雑音としてモデル化する。

(a) 拡散性雑音成分 $v_m(\tau, \omega)$ のパワースペクトログラム $\phi_{v_m v_m}(\tau, \omega)$ は、全てのマイクロ フォンにおいて等しい。即ち、

 $\phi_{v_1v_1}(\tau,\omega) = \phi_{v_2v_2}(\tau,\omega) = \dots = \phi_{v_Mv_M}(\tau,\omega)$

が成り立つ。

(b) 拡散性雑音成分 $v_m(\tau, \omega)$ のチャネル間クロススペクトログラム $\phi_{v_m v_n}(\tau, \omega)$ は、距離 が等しいマイクロフォン対に対しては等しい。即ち、 r_{mn} を m 番目と n 番目のマイク ロフォン間の距離とすると、

$$r_{mn} = r_{kl} \Rightarrow \phi_{v_m v_n}(\tau, \omega) = \phi_{v_k v_l}(\tau, \omega)$$
(21)

が成り立つ。

式(1)、(2)における定義から、 $\phi_{v_mv_n}(\tau,\omega)$ は $\Phi_{vv}(\tau,\omega)$ の(m,n)要素である。従って、 仮定(a)は $\Phi_{vv}(\tau,\omega)$ の対角成分が全て等しいことを意味する。更に、仮定(b)より、距離の 等しいマイクロフォン対を多数含む対称性の高いマイクロフォン配置を用いれば、 $\Phi_{vv}(\tau,\omega)$ の非対角成分の多くが等しくなり、 $\Phi_{vv}(\tau,\omega)$ は少数のパラメタで表されることが分かる。 我々はこの点に着目した考察の結果、図1に示す「結晶型マイクロフォンアレイ」を用い れば、 $\Phi_{vv}(\tau,\omega)$ の値に依らない定ユニタリ行列による対角化が可能であることを発見し た^{9),10)}。

以下では、図1中の長方形配置を例にとって、このことが可能であることを示す。まず、



図1 結晶型マイクロフォンアレイ。正多角形配置、長方形配置、正多面体配置、正多角柱配置、直方体配置が含ま れる。

仮定 (a) より、

(20)

$$\phi_{v_1v_1}(\tau,\omega) = \phi_{v_2v_2}(\tau,\omega) = \phi_{v_3v_3}(\tau,\omega) = \phi_{v_4v_4}(\tau,\omega) =: \alpha(\tau,\omega)$$
(22)
である。また、マイクロフォン間の距離について、

 $r_{12} = r_{21} = r_{34} = r_{43} \tag{23}$

$$r_{13} = r_{31} = r_{24} = r_{42} \tag{24}$$

$$r_{44} = r_{41} = r_{23} = r_{32} \tag{25}$$

であるから、仮定(b)より、対応するクロススペクトログラムは以下のように等しい。

$$\phi_{v_1 v_2}(\tau, \omega) = \phi_{v_2 v_1}(\tau, \omega) = \phi_{v_3 v_4}(\tau, \omega) = \phi_{v_4 v_3}(\tau, \omega) =: \beta(\tau, \omega)$$
(26)

$$\phi_{v_1 v_3}(\tau, \omega) = \phi_{v_3 v_1}(\tau, \omega) = \phi_{v_2 v_4}(\tau, \omega) = \phi_{v_4 v_2}(\tau, \omega) =: \gamma(\tau, \omega)$$
(27)

$$\phi_{v_1v_4}(\tau,\omega) = \phi_{v_4v_1}(\tau,\omega) = \phi_{v_2v_3}(\tau,\omega) = \phi_{v_3v_2}(\tau,\omega) =: \delta(\tau,\omega)$$
(28)

情報処理学会研究報告 IPSJ SIG Technical Report

式 (22)、(26)、(27)、(28) より、**Φ**_{vv}(τ,ω) は次の構造を持つ。

$$\boldsymbol{\Phi}_{\boldsymbol{v}\boldsymbol{v}}(\tau,\omega) = \begin{bmatrix} \alpha(\tau,\omega) & \beta(\tau,\omega) & \gamma(\tau,\omega) & \delta(\tau,\omega) \\ \beta(\tau,\omega) & \alpha(\tau,\omega) & \delta(\tau,\omega) & \gamma(\tau,\omega) \\ \gamma(\tau,\omega) & \delta(\tau,\omega) & \alpha(\tau,\omega) & \beta(\tau,\omega) \\ \delta(\tau,\omega) & \gamma(\tau,\omega) & \beta(\tau,\omega) & \alpha(\tau,\omega) \end{bmatrix}$$
(29)

この行列が、 $\alpha(\tau, \omega)$ 、 $\beta(\tau, \omega)$ 、 $\gamma(\tau, \omega)$ 、 $\delta(\tau, \omega)$ の値に関わらず、下記の定ユニタリ行列により対角化されることは容易に確認できる。

図 1 中の他の配置に対する対角化行列については小野ら¹⁰⁾を参照されたい。 $\Phi_{vv}(\tau,\omega)$ の値に依らない対角化手法であること、共分散行列の対角化は無相関化を意味することから、上述の対角化手法をブラインド雑音無相関化と呼ぶ。

ブラインド雑音無相関化に基づき、目的音のパワースペクトログラムは以下のようにして 推定される。ブラインド雑音無相関化における対角化行列 U を用いて、式 (19) は次のよう に変換される。

 $U^{\mathsf{H}} \Phi_{xx}(\tau,\omega) U = \phi_{ss}(\tau,\omega) U^{\mathsf{H}} d(\omega) d^{\mathsf{H}}(\omega) U + U^{\mathsf{H}} \Phi_{vv}(\tau,\omega) U$ (31) ここで、 $U^{\mathsf{H}} \Phi_{vv}(\tau,\omega) U$ は対角行列であると仮定すると、 $U^{\mathsf{H}} \Phi_{xx}(\tau,\omega) U$ の非対角成分は 次のようにノイズフリーである。

 $u_m^{\mathsf{H}} \Phi_{xx}(\tau, \omega) u_n = \phi_{ss}(\tau, \omega) u_m^{\mathsf{H}} d(\omega) d^{\mathsf{H}}(\omega) u_n \ (m \neq n)$ (32) ここで、 u_m は U の第 m 列を表す。上式において、 $u_m^{\mathsf{H}} \Phi_{xx}(\tau, \omega) u_n, u_m^{\mathsf{H}} d(\omega) d^{\mathsf{H}}(\omega) u_n$ は既知であるから、式 (32) は $\phi_{ss}(\tau, \omega)$ に関する過決定の連立方程式とみなせる。そこで本 手法では、最小二乗法に基づき $\phi_{ss}(\tau, \omega)$ を次式で推定する。

$$\hat{\phi}_{ss}(\tau,\omega) \triangleq \frac{\sum_{m \neq n} \left[\boldsymbol{u}_m^{\mathsf{H}} \boldsymbol{d}(\omega) \boldsymbol{d}^{\mathsf{H}}(\omega) \boldsymbol{u}_n \right]^* \boldsymbol{u}_m^{\mathsf{H}} \boldsymbol{\Phi}_{\boldsymbol{xx}}(\tau,\omega) \boldsymbol{u}_n}{\sum_{m \neq n} \left| \boldsymbol{u}_m^{\mathsf{H}} \boldsymbol{d}(\omega) \boldsymbol{d}^{\mathsf{H}}(\omega) \boldsymbol{u}_n \right|^2}$$
(33)

2.5 結晶型マイクロフォンアレイを用いた Wiener ポストフィルタ設計

式 (33) の $\phi_{ss}(\tau, \omega)$ の推定値を用いて、提案する Wiener ポストフィルタ設計は次式で与

えられる。

$$\frac{\hat{\phi}_{ss}(\tau,\omega)}{\hat{\phi}_{m}(\tau,\omega)} \tag{34}$$

ここで、分母の $\hat{\phi}_{yy}(\tau,\omega)$ は Zelinski³⁾ による $\phi_{yy}(\tau,\omega)$ の推定値

 $\hat{p}(\tau,\omega) \triangleq$

$$\hat{\phi}_{yy}(\tau,\omega) \triangleq \frac{1}{M} \sum_{m=1}^{M} \phi_{x_m x_m}(\tau,\omega)$$
(35)

である。ここで、最小分散ビームフォーマの出力 $y(\tau, \omega)$ は観測可能であるから、それを用 いてパワースペクトログラム $\phi_{yy}(\tau, \omega)$ を直接計算することも可能であるが、我々の実験に よると式 (35) を用いた場合の方が雑音抑圧性能が若干高くなる傾向があるため、式 (35) を 用いている。

なお、 $p(\tau, \omega)$ は理論上 $0 \le p(\tau, \omega) \le 1$ を満たすから、 $\hat{p}(\tau, \omega)$ もこれを満たすように次 の簡単な後処理を行う。

$$\begin{cases} \hat{p}(\tau,\omega) \leftarrow 0, \text{ if } \hat{p}(\tau,\omega) < 0\\ \hat{p}(\tau,\omega) \leftarrow 1, \text{ if } \hat{p}(\tau,\omega) > 1 \end{cases}$$
(36)

3.実験

提案法により残響環境下において音楽信号を楽器音に分離するシミュレーションを行った。実験条件は以下の通りである。図 2 に音源等の配置を示す。マイクロフォンアレイによる観測信号は、鏡像法¹¹⁾により生成した各音源から各マイクロフォンまでのインパルス応答を用いて計算した。残響時間 RT₆₀ は 270 ms であった。各楽器音は SiSEC¹²⁾ のデータベースから取った。データ長は約 6 s であり、サンプリング周波数は 16 kHz とした。

楽器音分離は各音源に対して計算した最小分散ビームフォーマ (15) とポストフィルタ (34) を用いて行った。短時間 Fourier 変換のフレーム長は 512 点、フレームシフトは 16 点とし、 Hamming 窓を用いた。最小分散ビームフォーマ (15) における Φ_{xx} は全フレームに亙って $x(\tau,\omega)x^{H}(\tau,\omega)$ を時間平均することにより求めた。一方、ポストフィルタ (34) を計算する 際に必要となる、式 (33) における $\Phi_{xx}(\tau,\omega)$ および式 (35) における $\phi_{x_mx_m}(\tau,\omega)$ は、短 時間定常性を仮定して 32 フレーム毎に $x(\tau,\omega)x^{H}(\tau,\omega)$ および $x_m(\tau,\omega)x_n^*(\tau,\omega)$ を時間平 均することにより求めた。最小分散ビームフォーマ (15) と式 (33) の $d(\omega)$ は音源方向を既 知と仮定して計算した。

観測信号、最小分散ビームフォーマの出力、提案法の出力の SN 比⁴⁾ および LSD¹³⁾ をそ れぞれ表 1、表 2 に示す。ここで、SN 比は雑音抑圧性能の尺度であり、値が大きいほど雑



図2 実験における音源等の配置。

表1 各音源に対する観測信号(@mic1)、最小分散ビームフォーマの出力、提案法の出力の SN 比 (dB)。

	音源 1	音源 2	音源 3
観測信号 (@mic 1)	-4.4	-8.9	1.2
最小分散ビームフォーマ	4.1	1.1	4.6
提案法	10.8	3.9	12.5

音抑圧性能が高いことを示す。一方、LSD は各手法の出力のスペクトログラムと目的音の スペクトログラムの間の距離尺度であり、値が小さいほど各手法の出力の目的音に対する歪 みが小さいことを示す。表より、提案法は LSD を大きく増加させることなく、大幅に SN 比を増加させたことが分かる。

図3は、音源3、観測信号(@mic1)、音源3に対する最小分散ビームフォーマおよび提案法の出力のスペクトログラムである。(c)の最小分散ビームフォーマの出力では他の音源の音が十分に除去されずに残っている(特に低音部)のに対し、(d)の提案法の出力では音源3が効果的に強調されていることが分かる。

表2 各音源に対する観測信号(@mic 1)、最小分散ビームフォーマの出力、提案法の出力の音源信号に対する LSD(dB)。

	音源 1	音源 2	音源 3
観測信号 (@mic 1)	10.1	9.2	7.8
最小分散ビームフォーマ	4.8	4.8	6.2
提案法	4.6	6.1	6.5



図 3 音源 3 に対するスペクトログラム。(a) 音源 3、(b) 観測信号(@mic 1)、(c) 最小分散ビームフォーマの出力、(d) 提案法の出力。

4. 結 論

本稿では、結晶型マイクロフォンアレイを用いた拡散性雑音抑圧法の音楽信号への応用例 として、残響環境下で録音された音楽信号を各楽器音に分離する問題を扱った。シミュレー ションでは提案法は最小分散ビームフォーマに比べて 3~8 dB 高い SN 比を与えた。

本稿では音源方向は既知と仮定したが、現在我々は提案手法を音源方向が未知の場合にも 適用可能とするために、結晶型アレイを用いて拡散性雑音環境下で音源方向を推定する手法 の研究を進めている。

謝 辞

本研究は科学技術振興機構 CREST の補助を受けて行なわれた。

参考文献

- M. Brandstein and D. Ward, Microphone Arrays: Signal Processing Techniques and Applications. Berlin: Springer-Verlag, 2001.
- 2) http://www.music-ir.org/mirex/2009/index.php
- R. Zelinski, "A microphone array with adaptive post-filtering for noise reduction in reverberant rooms," in *Proc. ICASSP '88*, New York, Apr. 1988, pp. 2578–2581.
- 4) K. U. Simmer, J. Bitzer, and C. Marro, "Post-filtering techniques," in *Microphone Arrays: Signal Processing Techniques and Applications*, M. Brandstein and D. Ward, Eds. Berlin: Springer-Verlag, 2001, ch. 3, pp. 39–60.
- I. A. McCowan and H. Bourlard, "Microphone array post-filter based on noise field coherence," *IEEE Trans. Speech Audio Process.*, vol. 11, no.6, pp. 709–716, Nov. 2003.
- 6) S. Lefkimmiatis and P. Maragos, "A generalized estimation approach for linear and nonlinear microphone array post-filters," *Speech Commun.*, vol.49, no. 7–8, pp. 657–666, July–Aug. 2007.
- 7) N. Ito, N. Ono, and S. Sagayama, "A blind noise decorrelation approach with crystal arrays on designing post-filters for diffuse noise suppression," in *Proc. ICASSP* 2008, Las Vegas, Apr. 2008, pp. 317–320.
- 8) H. L. Van Trees, Optimum Array Processing. New York: John Wiley & Sons, 2002.
- 9) H. Shimizu, N. Ono, K. Matsumoto, and S. Sagayama, "Isotropic noise suppression in the power spectrum domain by symmetric microphone arrays," in *Proc.* WASPAA, New Paltz, NY, Oct. 2007, pp. 54–57.
- 10) N. Ono, N. Ito, and S. Sagayama, "Five classes of crystal arrays for blind decorre-

lation of diffuse noise," in Proc. SAM, Darmstadt, Germany, Jul. 2008, pp. 151–154.

- J. A. Allen, D. A. Berkeley, "Image method for efficiently simulating small-room acoustics," JASA, vol. 65, no. 4, pp. 943–950, Apr. 1979.
- 12) http://sisec.wiki.irisa.fr/tiki-index.php
- I. Cohen, "Multichannel Post-Filtering in Nonstationary Noise Environments," IEEE Trans. Signal Process., vol. 52, no. 5, pp. 1149–1160, May 2004.