

ブラインド処理を用いたディスプレイ組込型 マイクロホンアレーシステムの一構成法

◎橋本 あゆみ 宇佐川 毅 江端 正直
(熊本大学)

1. はじめに

近年、急速にコンピュータの普及が進み、それとともにキーボード入力のみならず、マイクロホンを用いた音声入力技術も開発されてきている。現在は、発話者のすぐ近くにマイクロホンを設置し音声入力する手法が利用されている。しかし、この手法では、発話者がコンピュータとつながれた装置を装着しなければならない、または、その装置のすぐそばにいないといけないという負担がかかることとなる。そこで、マイクロホンアレーを用いた信号分離システムを使用し、発話者に負担のかからない集音方法が要求される。本報告では、コンピューターディスプレイに設置できるような、ディスプレイ組込型マイクロホンアレーを用いたビームフォーミング法を提案する。図 1 に示すような、平面状のディスプレイ組込型マイクロホンアレーを用いて、音源の情報が無い状況で、目的信号の方向および雑音の到来方向を推定し、ANF を用いたビームフォーミングにより目的信号の抽出を実現する。最後にその性能を検討する。

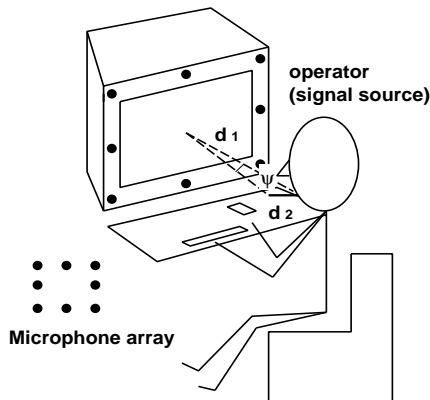


図 1. ディスプレイ組込型マイクロホンアレー

2. ディスプレイ組込型マイクロホンアレーを用いたビームフォーミング

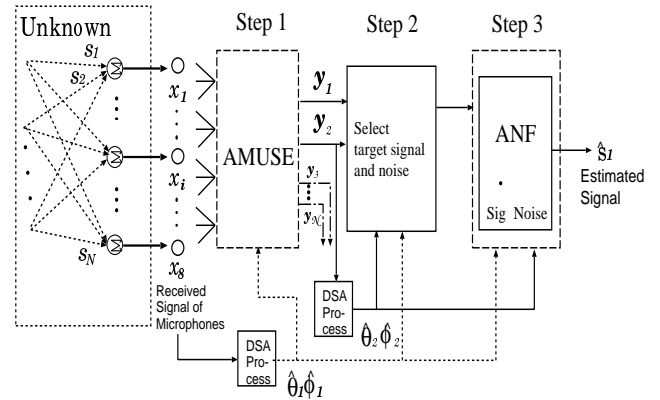


図 2. Beamforming のブロック図

本手法はディスプレイ組込型マイクロホンアレーを用いてビームフォーミングを行うシステムである。図 2 に本手法のブロック図を示す。図 2 にあるように、本手法は大きく 3 つの Step に分けられるが、全ての Step はマイクロホン観測信号をいくつかのフレームに分割し、各フレームごとに行う。

2.1 マイクロホンアレー出力信号の定義

図 3 に示すように、 $m_1 \sim m_8$ の 8 個のマイクロホンを d_1, d_2 の間隔でディスプレイ型に配置し、方向の異なる複数の音源を観測するマイクロホンアレーシステムを想定する。ここで、各音源信号は無限遠点に存在し、マイクロホン受信信号は平面波であり、反射は起こらないものと仮定する。また、1 フレームの長さを K とし、離散時間 k は各フレームごとに 0 から始まるものとする。

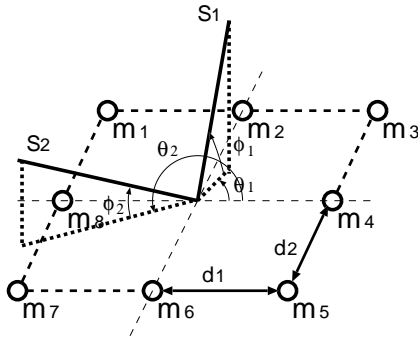


図 3. 平面状マイクロホンアレー

マイクロホン m_1 で受信される N 個の音源信号のうち最大パワを有する音源 (第 1 音源) を $s_1(k\Delta T)$, 他の音源を $s_n(k\Delta T) (n = 2, \dots, N)$ とし, 各音源方向は $(\theta_1, \phi_1), \dots, (\theta_N, \phi_N)$ とする. ただし, マイクロホン m_i における付加雑音を $n_i(k\Delta T)$ とする. その際, ΔT はサンプリング間隔を表す. また, マイクロホン m_i の出力信号 $x_i(k\Delta T)$ は, 次式で表現できる.

$$\begin{aligned}
 x_i(k\Delta T) = & s_1(k\Delta T - \tau_{(\theta_1, \phi_1), i}) \\
 & + s_2(k\Delta T - \tau_{(\theta_2, \phi_2), i}) \\
 & + \dots + s_N(k\Delta T - \tau_{(\theta_N, \phi_N), i}) \\
 & + n_i(k\Delta T)
 \end{aligned} \quad (1)$$

ただし, $\tau_{(\theta, \phi), i}$ はマイクロホン m_1 を基準としたときの方向角 (θ, ϕ) に対応するマイクロホン m_i の時間差を表す.

2.2 Blind Deconvolution による二音源信号の分離 (Step1)

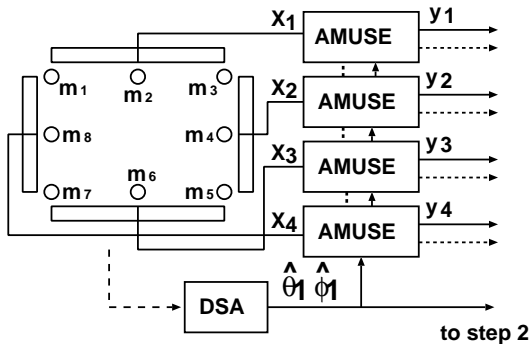


図 4. Step1 のブロック図

図 6に Step1 のブロック図を示す. Step1 では, 以下の処理を行う.

- DSA 法により, 最大パワを持つ音源 (第 1

音源) の DOA(Direction Of Arrival) を推定する.

- 受信信号を複数の sub-array に分解し, AMUSE 法により各音源信号の強調を行なう.
- 各 AMUSE による出力信号から, 第 1 音源と 2 番目に大きなパワを持つ音源の強調信号のみを取り出し, Step2 への入力信号とする.

2.3 目的信号と主雑音の決定 (Step2)

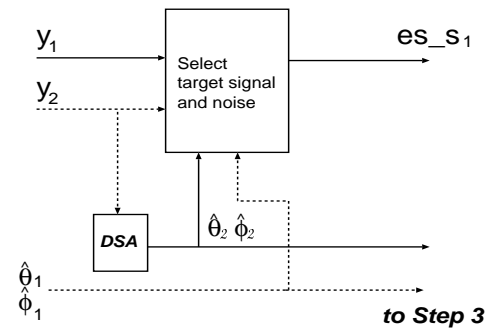


図 5. Step2 のブロック図

Step2 では, まず, 4 個の sub-array 出力信号より第 2 音源の DOA を推定し, 次に, 第 1 音源, 第 2 音源の強調信号から目的信号を選択する.

2.3.1 第 2 音源の DOA 推定

AMUSE 法により分離されたパワの小さい方の出力信号を用い, 第 2 音源の DOA $(\hat{\theta}_2, \hat{\phi}_2)$ を以下の手順で推定する.

- AMUSE の 4 つの出力信号は, それぞれ, m_2, m_4, m_6, m_8 に同期化された信号である. そこで, m_2, m_6 を pair 1, m_4, m_8 を pair 2 とする.
- pair 1, pair 2 それぞれで DSA 法を用いて α, β を推定する.
- 推定した α, β により, 第 2 音源の DOA が第 1 象現, 第 2 象現, 第 3 象現, 第 4 象現のどの象現に存在するか判定し, 次表の式を用いて第 2 音源の DOA を算出する.

第 1 象限	$\hat{\theta}_2 = \tan^{-1} \frac{\cos \alpha}{\cos \beta}$ $\hat{\phi}_2 = \cos^{-1} \frac{\cos \alpha}{\sin \beta}$
第 2 象限	$\hat{\theta}_2 = \pi - \tan^{-1} \frac{\cos \alpha}{-\cos \beta}$ $\hat{\phi}_2 = \cos^{-1} \frac{\cos \alpha}{\sin \beta}$
第 3 象限	$\hat{\theta}_2 = \pi + \tan^{-1} \frac{\cos \alpha}{\cos \beta}$ $\hat{\phi}_2 = \cos^{-1} \frac{\cos \alpha}{\sin \beta}$
第 4 象限	$\hat{\theta}_2 = 2\pi - \tan^{-1} \frac{-\cos \alpha}{\cos \beta}$ $\hat{\phi}_2 = \cos^{-1} \frac{\cos \alpha}{\sin \beta}$

2.3.2 目的信号の決定

各フレーム内での第 1 音源および第 2 音源の決定は、フレーム毎で異なり、2 音源信号間のパワー比に依存する。よって、第 1 音源、第 2 音源の置換が起こる可能性がある。そこで、この置換を防ぐため、前フレームにおいて推定した第 1 音源信号の DOA 情報を基に、目的信号を決定を行う。

2.4 ANF を用いたビームフォーミングと合成処理 (Step3)

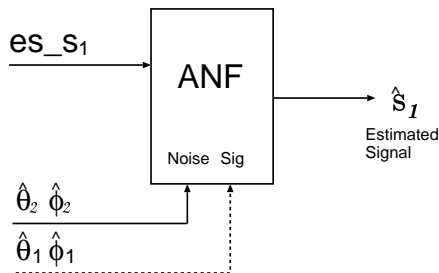


図 6. Step3 のブロック図

Step3 では、以下のような処理を行う。

- AMUSE 法の出力信号と各音源信号の DOA をもとに構成した ANF を用いて目的信号を推定し抽出する。
- ハニング窓を用いて、各フレームの出力信号の合成処理を行う。

ここで、ANF を構成する場合、逆フィルタを求める必要があるが、これは複数の零点を持ち、直接的な手法では安定な逆フィルタを求めることが困難である。この問題を解決する方法として、幸田らにより提案されている反復型エコー除去法 [3] を適用する。

3. シミュレーション

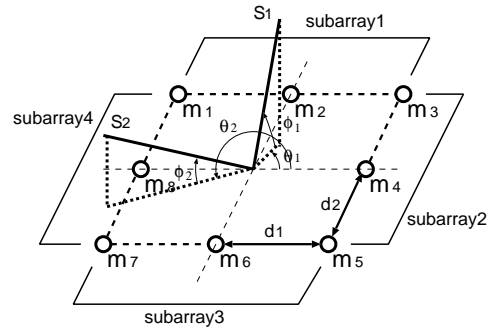


図 7. マイクロホンの配置図

図 7 に示したように、8 個のマイクロホンを平面上に $d1 = 0.15\text{m}$, $d2 = 0.10\text{m}$ で並べたディスプレイ型マイクロホンアレーを用いてシミュレーションを行った。第 1 音源として音声 /大きな交差点があります/

を $\theta_1 = 30^\circ$, $\phi_1 = 80^\circ$ 、第 2 音源として第 1 音源に対して -5dB のパワーを持つ音声

/会場はどちらですか/

を $\theta_2 = 240^\circ$, $\phi_2 = 40^\circ$ の方向に配置した。サブアレイ数は 4 個、各サブアレイのマイクロホン数は 3 個となる。サンプリング周波数は 15kHz 、受信信号は平面波を仮定し、反射は起こらないものとする。

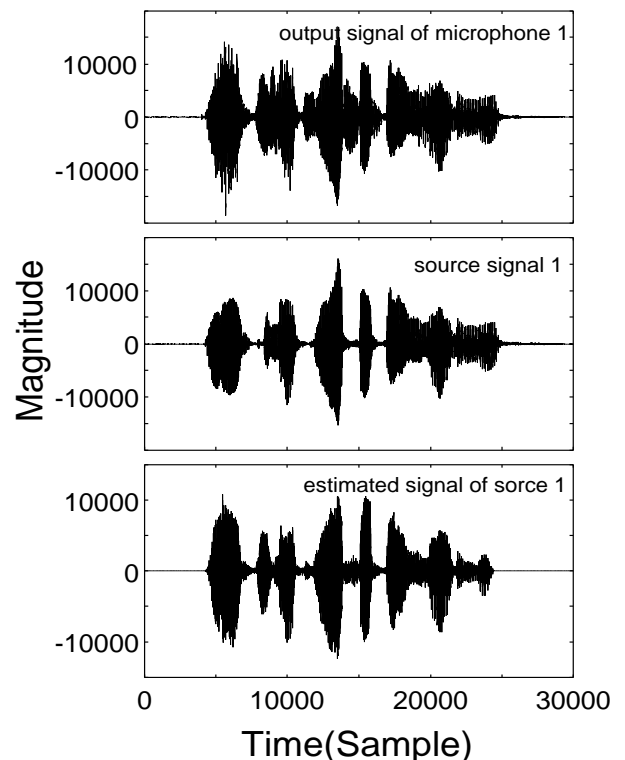
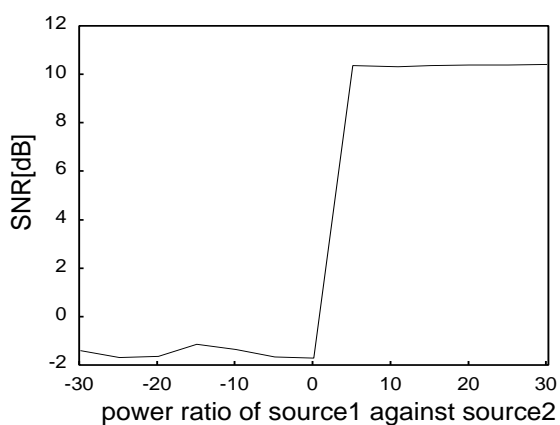


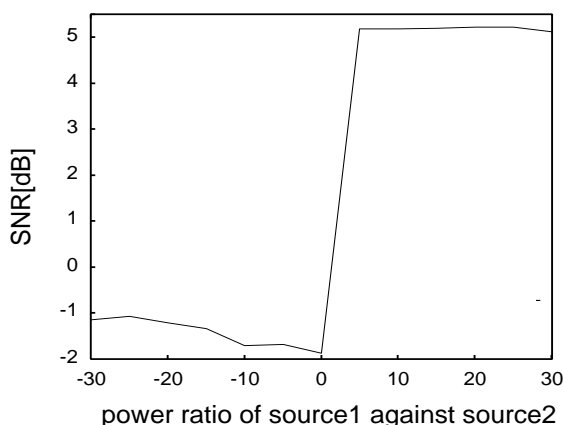
図 8. 目的信号の推定波形

図 8は、上から、マイクロホン 1 の出力信号、第 1 音源の原信号、第 1 音源の出力信号の時間特性である。

本手法の分離精度を確認するために、第 1 音源 (音声/a/) に対して第 2 音源 (ホワイトノイズ) のパワ比を $-30dB$ から $30dB$ まで変化させた場合の、出力信号の $300Hz$ から $3000Hz$ の周波数帯域における SNR を図 9 に示す。図 9 の (a) は、付加雑音を加えない場合、図 9 の (b) は、付加雑音としてマイクロホン 1 の出力信号に対し、 $-30dB$ のパワを持つホワイトノイズを加えた場合である。次に、第 1 音源 (音声/a/) の DOA を $\theta_1 = 60^\circ, \phi_1 = 80^\circ$ と固定し、第 2 音源としてホワイトノイズを半球面から到来させた場合の出力信号の SNR を図 10 に示す。原点からの方向が第 2 音源の方向、それぞれの方向の原点からの距離が SNR である。



(a) 観測雑音なし



(b) 観測雑音あり

図 9. 出力信号の SNR

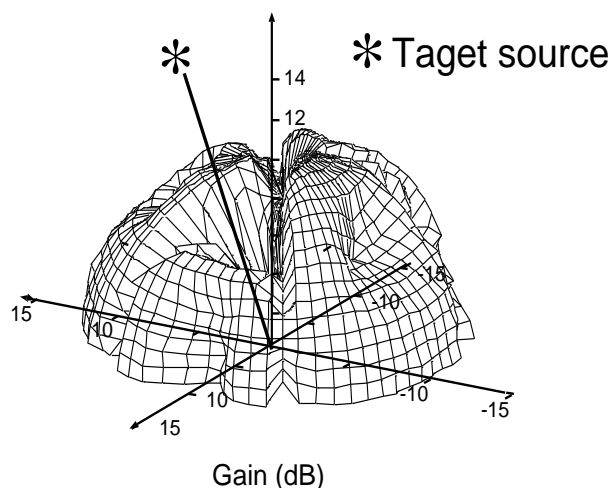


図 10. 出力信号の SNR

4. まとめ

今回、コンピューターディスプレイに設置できるような平面状のマイクロホンアレーを用いてビームフォーミングを行う方法を提案した。シミュレーションの結果、雑音除去できていることがわかった。しかし、付加雑音による影響がみられた。また、ほぼ全方向からの雑音に対応できることがわかった。今後、方向推定精度の向上、計算時間の短縮などに取り組み、認識実験を行う。

参考文献

- [1] Jin-Nam Park, Tsuyoshi Usagawa and Masanao Ebata, "A Signal Enhancement Method Using the Iterative Blind Deconvolution for Microphone Array System," IEICE E-A, Vol.E82-A, No.4, pp.611-618, 1999.4.
- [2] 澤田 慶昭, "反復型エコー除去法を応用した音源分離", 熊本大学修士学位論文, 1998.
- [3] 幸田 晃, 宇佐川 毅, 江端 正直, "系の零点を考慮した反復型エコー除去法," 日本音響学会誌, 53 巻 4 号, pp.268-276, 1997.
- [4] 本多 崇, 朴真南, 宇佐川 毅, 江端 正直, "マイクロホンアレーを用いたブラインド処理による音源分離", IECE EA99-97, 1999.
- [5] 橋本 あゆみ, 宇佐川 毅, 江端 正直, "ブラインド処理を用いたディスプレイ組込型マイクロホンアレーシステムの一構成法", IECE EA2000-13, 2000.