

# 複素スペクトル円心 (CSCC) 法と雑音音源方向推定を 組み合わせた雑音抑圧\*

井上和士 西本卓也 嵯峨山茂樹 (東大・情報理工)

## 1 はじめに

近年、マイクロホンアレイによる雑音除去技術の研究が盛んである。音声認識では雑音による認識性能低下が大きな問題であり、雑音の低減により性能向上が期待される [1]。またハンズフリー電話などでは目的音の SNR の向上にマイクロホンアレイ技術が有用と考えられる。

我々は事前の学習なしで非線形なフィルタを構成する「複素スペクトル円心 (Complex Spectrum Circle Centroid: CSCC) 法」[3] を提案した。本稿では CSCC 法の枠組で雑音の到来方向の推定を行うことにより、円心の推定を向上させる方法を報告する。

## 2 CSCC 法

異なる方向の目的音源及び単一の雑音源から平面波の音響信号が  $K$  個のマイクロホンに入力すると仮定する。各マイクロホンで観測した信号にそれぞれ適当な遅延を加え目的音源からの信号の時間差を補正すれば、 $i$  番目のマイクロホンの観測信号  $m_i(t)$  は以下のように表せる。

$$m_i(t) = s(t) + n(t - \tau_i), \quad i = 1, 2, \dots, K \quad (1)$$

ここで  $s(t)$  と  $n(t)$  はそれぞれ時刻  $t$  における目的信号と雑音信号、 $\tau_i$  は  $i$  番目のマイクロホンでの目的信号と雑音の間の到達時間差である。これをフーリエ変換すれば、周波数  $\omega$  の成分は

$$M_i(\omega) = S(\omega) + N(\omega)e^{-j\omega\tau_i}, \quad i = 1, 2, \dots, n \quad (2)$$

となる。式 (2) を幾何学的に解釈すると、図 1 に示すように、複素平面上で全ての  $M_i(\omega)$  は  $S(\omega)$  を中心とした半径  $\|N(\omega)\|$  の円上に位置する。したがってこれらの円周上の点が 3 個以上あれば、円の中心が求められ、目的信号のスペクトル  $S(\omega)$  が得られる。この原理に基づく信号処理手法を CSCC 法と呼ぶ。

推定された目的信号スペクトル  $\tilde{S}(\omega)$  を逆フーリエ変換し、フレーム毎にずらしながら重ね合わせることで雑音を抑圧した信号を得ることができる。

## 3 雑音到来方向推定を用いた円心推定

CSCC 法では原理的に雑音到来方向が周波数毎に異なってもよいが [3]、フレーム内で雑音到来方向が全帯域で同一である場合は、そのような条件下での CSCC アルゴリズムを設計すれば雑音除去性能を向上できる可能性がある。

実際に得られる各  $M_i(\omega)$  は複素平面上で上記の円上からずれて位置する。その理由として、同相化の誤差、マイクロホン特性のばらつき、一つの帯域への複数方向からの雑音、残響の要因等が考えられる。

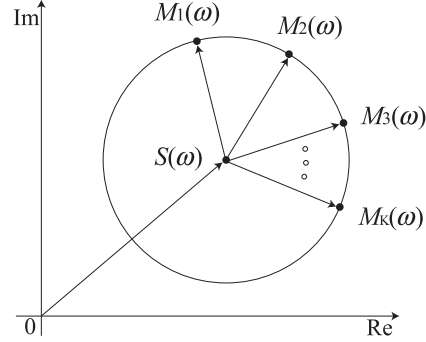


図 1: 各マイクロホンで受信した信号の周波数  $\omega$  における複素平面上での配置。

これら誤差要因を個別に議論せずに、誤差は複素平面上において半径方向及び角度方向にそれぞれ正規分布すると仮定し、最尤法により円心  $S(\omega)$  を推定する。

半径方向の誤差は円の半径の条件から、 $M_i(\omega)$  と  $S(\omega)$  の距離  $\|M_i(\omega) - S(\omega)\|$  の誤差を評価し、対数尤度:

$$L_r(M(\omega); S(\omega), R) = -\frac{K}{2} \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^K (\|M_i(\omega) - S(\omega)\| - R)^2 \quad (3)$$

を得る。 $M(\omega) = (M_1(\omega), M_2(\omega), \dots, M_K(\omega))^T$ ,  $R$  は円の半径、 $\sigma^2$  は正規分布の分散を表す。

角度方向の誤差は、雑音到来方向を推定し評価する。複素平面上において円心  $S(\omega)$  から見た観測点  $M_{i+1}(\omega)$  と  $M_i(\omega)$  の間の角度  $\phi_i(\omega)$  (以後これをマイクロホン間角度と呼ぶ) は、マイクロホン配置が直線的であり、その間隔  $d$  が既知かつ一定の場合、マイクロホンによらず  $\phi_i(\omega) = d \sin \theta \omega / c$  となる。ここで  $\theta$  は目的信号の到来方向から見た雑音の到来方向、 $c$  は音速である。つまりマイクロホン間角度は周波数  $\omega$  に比例し、その傾きは雑音の到来方向に当たる。傾き  $a$  は観測点から得られる  $\phi_i(\omega)$  及び推定目的信号  $\tilde{S}(\omega)$  から回帰分析を用いて、

$$a = \frac{\sum_{\omega} \sum_{i=1}^{K-1} \phi_i(\omega) \omega}{(K-1) \sum_{\omega} \omega^2} \quad (4)$$

により推定できる。ここから各  $\omega$  におけるマイクロホン間角度の推定角度  $\tilde{\Phi}(\omega) = a\omega$  が得られ、尤度は

$$L_a(M(\omega); S(\omega)) = -\frac{K-1}{2} \log 2\pi\zeta^2 - \frac{1}{2\zeta^2} \sum_{i=1}^{K-1} (\phi_i(\omega) - \tilde{\Phi}(\omega))^2 \quad (5)$$

により推定できる。ここで  $\zeta^2$  は角度方向誤差分布の分散である。

\* “Complex Spectrum Circle Centroid method combined with estimation of noise direction for Microphone-Array-Based noise reduction” by Kazuhito INOUE, Takuya NISHIMOTO and Shigeki SAGAYAMA (The University of Tokyo).

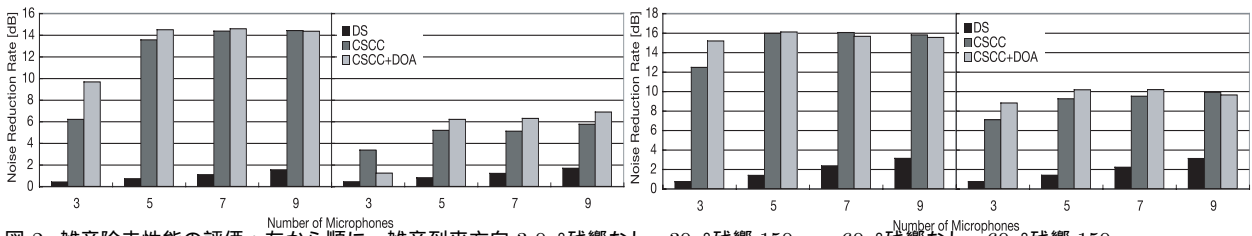


図 2: 雑音除去性能の評価: 左から順に, 雑音到来方向 30° 残響なし, 30° 残響 150ms, 60° 残響なし, 60° 残響 150ms.

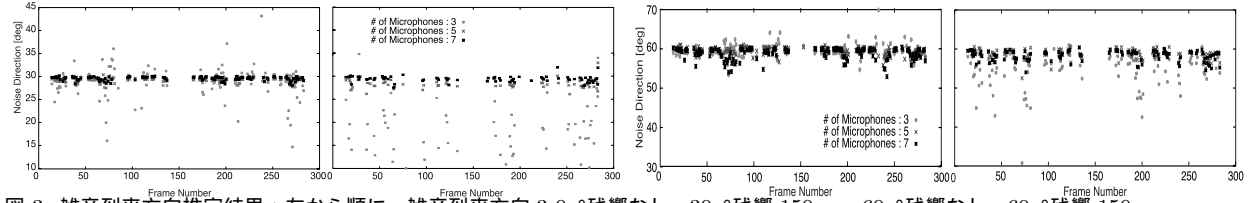


図 3: 雑音到来方向推定結果: 左から順に, 雑音到来方向 30° 残響なし, 30° 残響 150ms, 60° 残響なし, 60° 残響 150ms.

したがって, 半径方向及び角度方向を組合せた対数尤度は, それぞれは直交することから

$$L_{ra}(M(\omega); S(\omega), R) = L_r + L_a \quad (6)$$

により得ることができる. ただし, 半径方向の誤差は雑音のパワーの大きさにより変化し, 角度方向の誤差は周波数によって変化すると考えられる. したがって式 (6) における確率変数は半径  $R$  に対する比, 推定マイクロホン間角度  $\tilde{\Phi}(\omega)$  に対する比と仮定し, 各分布の平均を 1 に正規化した.

式 (6) の最大化問題は解析的に解くことはできない. 本稿では勾配法を用いて最大化を行った.

以上の手順を以下に示す.

- 従来の CSCC 法 [3] で初期推定目的信号スペクトル  $\tilde{S}(\omega)$  を求める.
- マイクロホン間角度  $\phi_i$  を求め, 回帰分析により雑音の到来方向を推定する.
- 式 (6) の対数尤度を最大化する  $\tilde{S}(\omega)$  を求める.
- 雑音到来方向が収束するまで繰り返す.

#### 4 シミュレーション実験による評価

実験条件 イメージ法 [2] を用いて計算機上で閉空間の音響特性を作成し, シミュレーション実験を行った. 部屋の大きさは  $6.00 \times 5.00 \times 3.00 \text{m}^3$  とし, 残響時間は 0ms (反射無し), 150ms とした. マイクロホンアレーは直線型とし, 素子間距離を 4cm とし, 素子数は 3, 5, 7, 9 とした. 目的音源はアレー正面から見て 0° 方向, 妨害音源は 30° あるいは 60° から到来するとし, マイクロホンアレーからの距離は 1.50m とした. 目的音声, 妨害音声は日本音響学会の新聞記事読み上げ音声コーパスから選び, 8 通りの組み合わせで実験を行った. 音源のサンプリング周波数は 8kHz, 分析フレーム幅を 37.5ms, フレームシフトを 18.75ms, 窓関数をハミング窓とした. 性能評価尺度としては Noise Reduction Rate (NRR = 出力 SNR - 入力 SNR) を用いた. 評価の対象として, 雑音到来方向推定を用いた CSCC 法 (CSCC+DOA), 従来法として CSCC 法, Delay-and-Sum (DS) を選んだ. CSCC+DOA で用いるパラメータ  $\sigma^2$  及び  $\zeta^2$  はそれぞれ 0.01 とした. また雑音として音声を扱っているため, 評価音声には雑音がない区間も存在する. 回帰分析による直線当てはめの精度により雑音到来方向情報の利用を閾値処理で決定した.

実験結果と考察 雑音除去結果及び CSCC+DOA による雑音到来方向推定結果をそれぞれ図 2 図 3 に示す. 無響条件及び残響 150ms の条件とも CSCC+DOA は DS に比べ, マイクロホンの数, 雑音到来方向によらず, 雑音の低減を行うことができた. 無響条件では最大で約 14dB, 150ms の残響下では最大約 9dB の SN 比の改善が見られた.

残響により約 6 から 8dB の性能低下が見られた. これは, 残響環境では目的信号と雑音がともに様々な方向から到来し, 複数雑音を扱う問題に相当するためであると考えられる. 現実的な対策としては円の当てはめの精度にしたがって Delay-and-Sum 法などの他の手法との組合せを考える必要がある.

CSCC+DOA は CSCC に比べ, 残響 150ms 下では平均で約 1dB の性能の向上を得た. 無響下では特にマイクロホン 3 本の条件で大きな性能向上を示した. 図 3 より, 雑音到来方向推定は円心の推定に対し有益な情報を与えていたことが分かる. しかし, 雑音到来方向 30°, マイクロホン数 3 の条件において CSCC+DOA の性能が低下した. この原因については図 3 から分かるように雑音到来方向の推定がうまくいかなかったためであると考えられ, より頑健な雑音方向推定のアルゴリズムが要望される.

#### 5 結論

本稿では, CSCC 法の原理に基づいて雑音除去を行うアルゴリズムを検討した. フレーム内で雑音が 1 つである場合, 複素平面上において推定円心と観測点でつくる角度を評価することにより, 雑音の到来方向を CSCC 法の枠組において推定し, 雑音除去性能を向上できる. 雑音到来方向推定を用いる CSCC 法は従来の CSCC 法に比べ残響環境下では平均で約 1dB の性能改善を示し, 残響なしでは特にマイクロホン 3 本の条件で大きな性能向上を示した. 今後は本手法を実環境において評価する予定である.

#### 参考文献

- [1] J. Bitzer, K.U. Simmer, K.-D. Kammeyer, "Multi-microphone noise reduction techniques as front-end devices for speech recognition," *Speech Communication*, vol. 34, pp. 3-12, 2001.
- [2] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 62, no. 4, pp. 943-950, 1979.
- [3] 岡嶋, 鎌本, 西本, 嵯峨山, "マイクロホンアレー入力の周波数領域での幾何学的処理による雑音環境下での音声認識," *音講論集*, vol.1, 3-10-1, pp. 583-584, 2004-Sep.