

# 非対称な球面配置 2ch マイクロホンによる音源定位\*

深町 聡一郎 (東大工), 小野 順貴, 西本 卓也, 嵯峨山 茂樹 (東大院情報理工)

## 1 はじめに

一般に音源定位はマイクロホン数が多いほど高い性能が実現できるが, 2ch アレイは, 多くの計算機がもつ標準的な音響入力で処理可能であり, 音源定位や, 定位に基づく音源分離の適用を容易にする利点をもつ. 通常の 2ch アレイではその対称性から, 360 度の方位, もしくは 3 次元的な方位を定位することはできないが, 人間や生物の聴覚系を模倣することにより, これを可能にする研究が様々行なわれている [1, 2, 3]. これらの研究では耳介や頭部の周波数特性が定位の大きな手がかりとなっているが, 一般に耳介状の凹凸が生じる反射・回折は複雑なため, 周波数特性の実測が不可欠となり, 形状の最適化も困難である.

一方, 単純な形状である球は, 回折の解析解が古くから知られており, 球面上多 ch アレイを利用した音源定位なども提案されている [4]. 本研究の目的は, 球にマイクロホンを非対称配置することにより, 単純な形状で, 360 度定位可能な 2ch 音源定位システムを実現することである. Handzel らも, マイクロホンを装着した球の微小回転を利用した定位システム [5] の枠組の中で非対称配置に触れているが, 白色雑音を用いた簡単な実験に留まっていた. 本稿では, 非対称配置による音源定位の理論と全方位定位可能性の検討, 音声を対象としたシミュレーション実験, また, マイクロホンの最適配置について論じる.

## 2 非対称な球面配置 2ch アレイによる全方位定位の理論

### 2.1 球面配置 2ch アレイの問題設定

自由空間に配置された 2 個のマイクロホンでは, その対称性から前後方向の区別を行うことができない. そこで本研究では, Fig.1 のように 2 個のマイクロホンを球面の赤道上の  $\pm\phi$  の角度に非対称に配置し, 同一水平面内の  $\theta$  方向から到来する音波を定位することを考える. (ただし  $\phi = 90^\circ$  の場合は対称配置となる.) また以下では, 左右 2 個のマイクロホンで観測された信号の Fourier 変換をそれぞれ  $M_L(\omega)$ ,  $M_R(\omega)$  で表す. このときこれらの観測信号は

$$M(\omega) = S(\omega)b(\omega, \theta) + N(\omega) \quad (1)$$

のようなモデルで表される. ここで  $M(\omega) = (M_L(\omega), M_R(\omega))^T$ ,  $S(\omega)$  は音源信号,  $N$  は観測雑音を表す. また,  $b(\omega, \theta)$  はステアリングベクトルと呼ばれる音源方向に依存した複素ベクトルであり, 音源を任意性なく定位するためには,  $b(\omega, \theta)$  が音源方向と 1 対 1 の関係にあることが必要となる.

### 2.2 球による音波の回折

平面波入射に対する球の回折については理論式が知られており [6], 球上の点における音圧  $p(\omega, \psi)$  は, Legendre 関数  $P_n(x)$  と球 Hankel 関数  $h_n(x)$  を用いて

$$p(\omega, \psi) = D(\omega, \psi) p_0(\omega) \quad (2)$$

$$D(\omega, \psi) = \frac{1}{ka} \sum_{n=0}^{\infty} \frac{j^{(2n+1)} P_n(\cos \psi)}{ka h_{n+1}^{(2)}(ka) - nh_n^{(2)}(ka)} \quad (3)$$

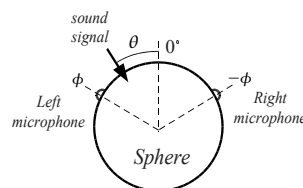


Fig. 1 マイクロホンの配置と座標系. 正面を 0 とし, マイクロホンを球体赤道上の  $\pm\phi$  に配置, 音源は赤道面内に存在し, その方向を  $\theta$  とする.

と表される. ここで,  $p_0$  は球が無いと仮定した場合の球中心にあたる点での音圧,  $a$  は球の半径,  $k = \omega/c$  で  $c$  は音速,  $\psi$  は球中心からみた球上の点と音源方向との間の角度である. また,  $D$  を回折係数と呼ぶ.

球面上の 2ch マイクロホンに対するステアリングベクトルは, この回折係数  $D$  を用いて次のように表される.

$$b(\omega, \theta) = \begin{pmatrix} b_L(\omega, \theta) \\ b_R(\omega, \theta) \end{pmatrix} = \begin{pmatrix} D(\omega, \theta - \phi) \\ D(\omega, \theta + \phi) \end{pmatrix} \quad (4)$$

### 2.3 球回折を利用した 2ch 全方位定位可能性の検証

球の回折を利用した場合に, 2ch で全方位が定位可能かどうか確認するために, 音源方向  $\theta$  に対するステアリングベクトル  $b(\omega, \theta)$  の変化を調べた. 定位に関係するのは  $b(\omega, \theta)$  の方向であるから, 絶対値の任意性を排除するために, 成分比である

$$ITF(\omega, \theta) = \frac{D(\omega, \theta - \phi)}{D(\omega, \theta + \phi)} \quad (5)$$

をチャンネル間伝達関数 (Inter-channel Transfer Function; ITF) と定め, この変化の例を Fig. 2 に示す.

非対称配置の場合は,  $\theta$  と ITF がほぼ 1 対 1 の関係をもち, 観測信号から, 前後判断を含めた  $360^\circ$  の音源方向が定位できそうであることがわかる.

## 3 最尤法に基づく定位アルゴリズム

2ch アレイにおける音源定位では位相差や強度差などを特徴量とすることが多いが, 本稿では, 今後, 雑音の性質を導入しやすいよう, 一般のマイクロホンアレイで用いられる最尤法の枠組を用いる. 観測モデルである式 (1) において, 観測雑音  $N$  が分散  $\sigma^2$  の複素ガウス分布に従うと仮定し, 未知の音源信号  $S$  に最尤値を仮定すると, 音源方向  $\theta$  に対する対数尤度は,

$$\log p(M; b) = -\log(2\pi\sigma^2) - \frac{1}{2\sigma^2} |M|^2 + \frac{1}{2\sigma^2} \cdot \frac{|b^H M|^2}{|b|^2} \quad (6)$$

と求まる. この尤度を最大とする  $\theta = \theta_0$  を音源方向推定値として用いる.

## 4 音源定位シミュレーション実験

### 4.1 時間周波数成分毎の定位

予備的な検討として, 時間周波数成分毎の尤度最大方向が真の音源方向に一致することを確認するた

\*Sound Source Localization by Asymmetrically Arrayed 2ch Microphones on Sphere by FUKAMACHI, Soichiro, NISHIMOTO, Takuya, ONO, Nobutaka and SAGAYAMA, Shigeki (The University of Tokyo)

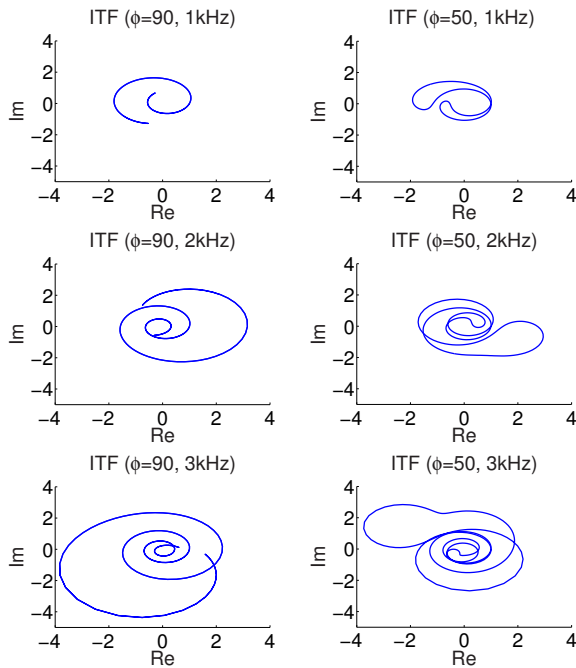


Fig. 2 球半径 85mm, 1kHz, 2kHz, 3kHz における, 音源方向を変化させた時の複素平面での ITF の軌跡. マイクロホンを両極に配置した場合 (左列) は, 前後対称な位置において信号比が同一であるが, 前後非対称な配置 (右列) ではこれがほとんどの点で異なる.

めのシミュレーション実験を行った. サンプル周波数 16kHz の音声信号を音源信号とし, フレーム長を 1024 サンプル, シフト幅 512 サンプルの短時間 Fourier 変換して用いた. この条件は以降の実験でも同様である. 球の直径を 85mm, マイクロホンの配置は  $\phi = 50^\circ$  とし, 音源方向を  $\theta = 60^\circ$  としたときに, 各フレームに対して周波数ごとに対数尤度最大を与える  $\theta_0$  を求め, プロットしたものを Fig. 3 に示す. ただし意味のない結果を除くため, 信号パワー ( $|M_L|^2 + |M_R|^2$ ) がある閾値より小さい点を除外してある.

多くの場合で推定音源方向が正しい方向の周りに集まっているが, 高い周波数帯域で定位に失敗している点も見られる. これは Fig. 2 でもわかる通り, 高い周波数帯域においては, ITF が重なる場合があり,  $\theta$  が一意に定まっていなかったためであると考えられる.

#### 4.2 時間フレーム毎の定位

前述のような時間周波数成分毎の定位に生じる曖昧性を除くため, 1 フレーム内で音源は 1 つと仮定し, 毎フレームでの対数尤度を周波数方向に加算した評価量を用いた定位について検討した. 結果を Fig. 4 に示す. ただしこの実験では, 左右独立な 10dB の白色雑音を加えてあるが, 全周波数の情報を統合したことにより, ほとんどの方向を正しく定位できていることがわかる. 音源方向が  $0^\circ$  と  $180^\circ$  に近い点で前後の混乱が生じているが, これは, 周波数にかかわらずこの 2 点とともに ITF=1 となるからである.

#### 4.3 非対称配置の最適化実験

Fig. 2 より, 左右マイクロホン間の信号比の軌跡はマイクロホンの配置角度  $\phi$  によって変化することが示された. マイクロホンの配置によってこの軌跡の重なり具合が変化し, 前後判断の正確さに影響を及ぼすことが予想されるため, 最適な  $\phi$  を求めることが重要になる. ここでは実験的に音声信号を音源信号とし,  $\phi$  の変化による性能を評価するための指標として, 左右

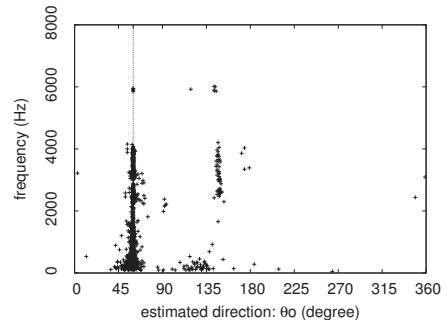


Fig. 3  $60^\circ$  方向の音源に対する尤度最大点の散布図. 30 フレーム分に対する結果を重ねがきしてあり, 信号強度が小さい点は除外してある.

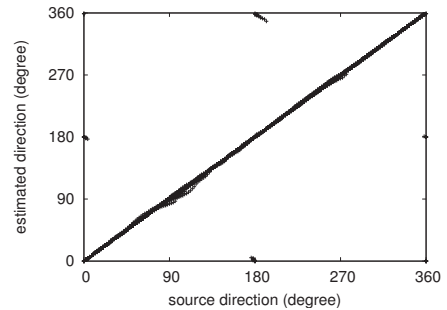


Fig. 4 SN 比が 10[dB] の場合の定位結果

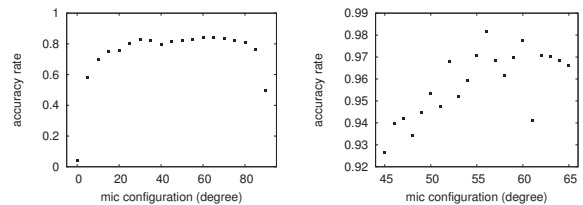


Fig. 5 ノイズを加えた際の, マイクロホンの配置角度による定位正解率の変化. (左図) における最大値付近である  $45^\circ \sim 65^\circ$  の範囲を拡大したものが (右図).

独立な白色雑音を観測信号に付加した際の定位正解率を用いた. 球の半径が 85mm, 雑音 SNR=10, 0 の場合にそれぞれ 4.2 節の方法で  $\theta = 0^\circ \sim 360^\circ$  でシミュレーションを行い, 正解率の平均をとった. 定位方向が音源方向から  $\pm 5^\circ$  以内であれば定位に正解したとみなした. 本実験でも前節と同様, 観測信号のパワーが小さい点を除いてある. 結果を Fig. 5 に示す. 最適値として  $\phi = 56^\circ$  という値を得た.

現在, この最適化に基づく実機を試作中であり, 今後, 実環境での定位実験と評価を行なう予定である.

#### 参考文献

- [1] N. Ono, et al., Proc. Transducers'01, Vol. 2, pp.1654-1657, Munich, Jun. 2001.
- [2] H. Nakashima, et al., Acoust. Sci & Tech., Vol. 24, No. 4, pp. 172-178, 2003.
- [3] 西野隆典他, 日本音響学会誌, 63 巻, pp. 3-12, 2007.
- [4] 福留公利, 日本音響学会誌, 44 巻, pp. 272-281, 1988.
- [5] A. A. Handzel et al., IEEE Sensors J., Vol. 2, No. 6, pp. 607-616, 2002.
- [6] 早坂壽雄他, 音響振動論, 丸善, 1974.