

パラボラ反射板による音響伝達特性の変化を用いたシングルチャネル音源方向推定*

高島遼一, 滝口哲也, 有木康雄 (神戸大)

1 はじめに

これまでに提案されてきた音源方向の推定方法は、マイクロフォンアレーにおける各観測信号の時間差を用いた手法が多く、複数のマイクロフォンが必要であった。単一マイクロフォンによる音源方向推定を可能にするため、我々はアクティブマイクロフォンというものを作成し、アクティブマイクロフォンが各角度で収録した音声のパワーを比較することで音源方向を推定するという手法を提案した [1]。しかし音声のような非定常信号では角度毎の入力信号のパワーが異なるため、正しく音源方向が推定されないことがある。

本稿ではパワーの代わりに観測信号の音響伝達特性を推定し、それらを比較することで音源方向推定を行う手法を提案し、パワーの比較による手法と比較して提案手法の有効性を示す。

2 アクティブマイクロフォン

本研究で用いたアクティブマイクロフォンは、Fig. 1(a) のように放物面型の反射板の焦点位置に無指向性マイクが取り付けられている。反射板からの位相のずれた反射波は観測信号のパワースペクトルに影響を与えるため、マイクロフォンアレーを用いない音源方向推定法として反射板が用いられることがある [2, 3]。本手法では反射板とマイクの向きを手動で変えながら各方向毎に音声を収録し、角度毎の音声を比較することで音源方向を検出する。

2.1 パラボラ反射板

平面波を仮定した場合、放物面の軸と平行に入射した波はすべて放物面の焦点に向かって反射される。一方別の角度から入射した波が反射によって焦点に到達することはない。音源が放物面の正面に位置している場合の観測信号を Fig. 1(b) に示す。 $x_0(t)$ は焦点 O へ直接到達する信号、 $x_m(t) (m = 1, \dots, M)$ は放物面を反射して O へ到達する信号である。放物線の性質より $x_0(t)$ と $x_m(t)$ の焦点までの時間差は m の値に関わらず一定であるので、 $x_0(t)$ 、 $x_m(t)$ の伝達関数をそれぞれ $h_0(t)$ 、 $h_m(t)$ 、クリーン音声を $s(t)$ と

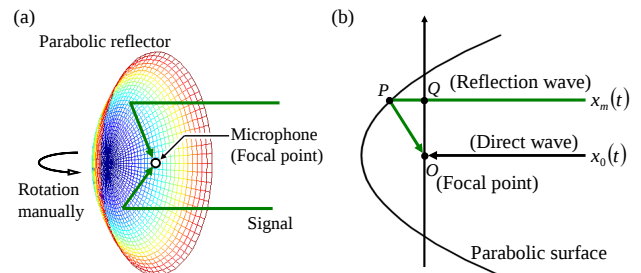


Fig. 1 (a) アクティブマイクロフォン (b) 放物面が音源方向を向いているときの焦点における観測信号。

すると O における観測信号は

$$\begin{aligned} o(t) &= x_0(t) + \sum_{m=1}^M x_m(t) \\ &= s(t) * h_0(t) + \sum_{m=1}^M s(t - \tau) * h_m(t) \quad (1) \end{aligned}$$

フレーム n 毎に短時間フーリエ変換を行うと

$$\begin{aligned} O(\omega; n) &\approx S(\omega; n) \cdot (H_0(\omega; n) + e^{-j2\pi\omega\tau} \cdot \sum_{m=1}^M H_m(\omega; n)) \\ &= S(\omega; n) \cdot (H_p(\omega; n) + H_r(\omega; n)) \quad (2) \end{aligned}$$

が得られ、反射板の影響を受けない伝達関数 H_p と反射板により追加される伝達関数 H_r を用いて表すことができる。一方音源が放物面の正面に位置していない場合、 O へ到来する信号は $x_0(t)$ のみであるので、 O における観測信号は

$$O(\omega; n) \approx S(\omega; n) \cdot H_0(\omega; n) = S(\omega; n) \cdot H_p(\omega; n) \quad (3)$$

となる。

2.2 音響伝達特性による音源方向の推定

(2), (3) 式より、アクティブマイクロフォンの向きを θ 、音源方向を $\hat{\theta}$ とすると伝達特性 $H_\theta(\omega; n)$ は

$$H_\theta(\omega; n) = \begin{cases} H_p(\omega; n) + H_r(\omega; n) & (\theta = \hat{\theta}) \\ H_p(\omega; n) & (\theta \neq \hat{\theta}) \end{cases} \quad (4)$$

となる。ここでアクティブマイクロフォンの角度が変化しても H_p の値がほぼ一定であれば、 H_θ はマイクが音源方向を向いていない状態ではほぼ同じ空間上に分布し、音源方向を向いたときのみそれらとは離

* Monaural sound-source-direction estimation using change of acoustic transfer function by parabolic reflector, by Ryoichi Takashima, Tetsuya Takiguchi and Yasuo Ariki (Kobe Univ.)

れた位置に分布することになる．そこで (5) 式を用いて各角度で得られる H_θ の中で最も離れた位置に分布する $\hat{H}_\theta$ を見つけ、それに対応する角度を音源方向として出力する．

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_{\theta'} (\bar{H}_\theta - \bar{H}_{\theta'})^2 \quad (5)$$

ここで $\bar{H}$ は H のフレーム平均を表す．

2.3 音響伝達特性の推定

本節では前節で述べた H の求め方について述べる．観測信号 O はケプストラム領域では次のように表される．

$$O_{cep}(d; n) \approx S_{cep}(d; n) + H_{cep}(d; n) \quad (6)$$

d はケプストラムの次元を表す．(6) 式より、 O と S を観測することができれば H を求めることができるが、実際には S を観測することはできない．そこで S の代わりにあらかじめクリーン音声の GMM (Gaussian mixture model) を作成しておき、 O に対して、その GMM の尤度が最大となるように H を EM アルゴリズムにより求める [4]．

$$\hat{H} = \underset{H}{\operatorname{argmax}} \Pr(O|\lambda_S, H) \quad (7)$$

λ_S はクリーン音声のモデルパラメータであり、混合要素 k 毎に重み w_k 、平均 $\mu_k^{(S)}$ 、分散 $\sigma_k^{(S)^2}$ によって表される．このとき、 Q 関数は次式のように導出される．

$$Q(\hat{H}|H) = - \sum_k \sum_n \gamma_k(n) \sum_{d=1}^D \left\{ \frac{1}{2} \log(2\pi)^D \sigma_{k,d}^{(S)^2} + \frac{(O(d;n) - \mu_{k,d}^{(S)} - \hat{H}(d;n))^2}{2\sigma_{k,d}^{(S)^2}} \right\} \quad (8)$$

$$\gamma_k(n) = \Pr(O(n), k|\lambda_S) \quad (9)$$

この Q 関数を最大にする $\hat{H}$ は、 $\hat{H}$ について偏微分して解くことにより求めることができる．

$$\hat{H}(d;n) = \frac{\sum_k \gamma_k(n) \frac{O(d;n) - \mu_{k,d}^{(S)}}{\sigma_{k,d}^{(S)^2}}}{\sum_k \frac{\gamma_k(n)}{\sigma_{k,d}^{(S)^2}}} \quad (10)$$

3 実環境下における実験

実環境下で提案手法とパワーを用いた手法の比較実験を行った．パワーを用いた手法は各角度で収録される音声のパワーのフレーム平均値が最も大きい角度を音源方向とするものである．マイクの角度は 30° から 150° までの 7 方向、音源距離は 2 m、音源方向は 90° 、反射板は直径 24 cm、焦点距離 9 cm のもの

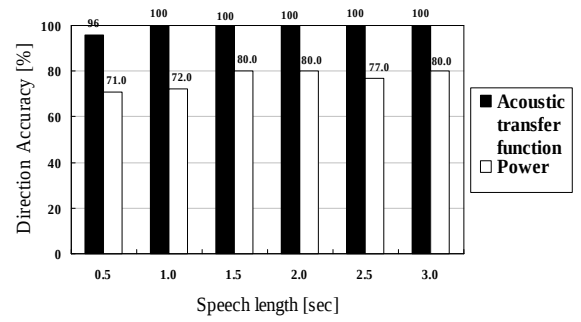


Fig. 2 角度毎の入力音声と同じ場合における音源方向の正解率

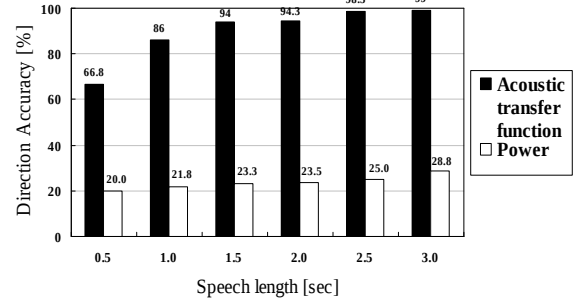


Fig. 3 角度毎の入力音声が異なる場合における音源方向の正解率

を用いている．また、サンプリング周波数は 12 kHz、提案手法の特徴量は MFCC 2 次元を用いた．

角度毎の入力音声と同じ場合と異なる場合のそれぞれについて、1 方向における収録時間毎の正解率を Fig. 2, Fig. 3 に示す．角度毎の入力音声と同じ場合、提案手法では 1 方向での収録時間が 1 秒以上であれば 100% の正解率が得られた．一方、角度毎の入力音声異なる場合、パワーを用いた手法では入力音声と同じ場合に比べて正解率が大きく低下しているのに対して、提案手法では収録時間が短くなるにつれて正解率は低下しているが、パワーを用いた手法と比較すると低下率は抑えられていた．

4 おわりに

本稿ではアクティブマイクروفोनによる音響伝達特性を用いて音源方向推定を行う手法を提案し、以前のパワーを用いた手法との比較実験によりその優位性を示した．今後は反射の影響を強く受ける周波数帯域などについて調査し、さらに細かく伝達特性を分析する予定である．

参考文献

- [1] 高島他, 音講論 (春), 765-766, 2008.
- [2] 市川他, 音講論 (春), 687-688, 2003.
- [3] 小野他, 信学技報, 25-32, 2000.
- [4] 住田他, 音講論 (春), 771-772, 2008.