

両耳聴ヘッドトルソを用いた音源の方向推定

日大生産工 ○霜山 竜一

1 はじめに

両耳聴によって到来する音の方向を推定できるのは、異なる2箇所の音圧に音源の空間情報が含まれることを意味する。筆者は、両耳聴ヘッドトルソで測定された音響信号からコンピュータで音源位置の2次元分布を推定することにより、脳で聴空間を把握する聴覚情報処理の仕組みの一端を明らかにする研究を行っている。心理物理学実験による知見から、従来、人の聴覚では低い周波数の音を時間差で、高い周波数ではその強度差を用いて音の方向を求めているとされる。しかし音源が遠方にある場合には、音圧の強度差は小さすぎて容易に検出できず。時間差も反響の影響を受け易い。人の聴覚器官は精巧である一方、頑健でもある。聴覚が様々な悪環境でも十分機能する点を考慮すれば、従来の定説は誤りである可能性がある。一方、神経回路に関する生理学実験から Jeffress¹⁾は、人の聴覚中枢で音源の方向が知覚される仕組みについて簡潔で本質的なモデルを提案している。蝸牛からの神経パルスは聴覚中枢にある一種の遅延線上を伝搬し、左右のパルスの出会う空間的な位置は音源の方向に対応するというものである。Konishi ら²⁾は生理学実験を通してメンフクロウが音源の水平方向を検出する仕組みを明らかにした。音の周波数が高い場合に、周期的に発生する神経パルスは、遅延線に沿って複数の異なる一致検出細胞を発火させる。これは周期波である神経パルス

の位相が多義的であることに起因する。メンフクロウの聴覚中枢で位相の多義性を解決して音源の方向を検出するメカニズムについても提案している。

また情報学の観点から、両耳聴の音響信号間のクロススペクトルの特徴と音源位置の関係を直接学習させる試みもある³⁾。これは現象の物理的、数学的側面を無視してデータ間を対応付けるもので、それらの特徴が音源やセンサの位置に著しく依存するため、未学習の環境で音源位置を推定することは難しい⁴⁾。

本報告では、メンフクロウの聴覚中枢で音源の水平方向を検出するメカニズムと類似のアルゴリズム⁵⁾を、両耳聴ヘッドトルソで測定された音響信号に適用し、1個の小型スピーカの方向を検出した。ヘッドトルソの向きを変えて、固定されたスピーカの方向を推定した結果について報告する。

2 音源方向の推定アルゴリズム

ヘッドトルソの両耳部奥にそれぞれマイクロホンが内蔵されており、周囲から到来した音は、ヘッドトルソの頭部や外耳で回折することが予想される。音の回折による効果は音の波長とヘッドトルソの外耳、頭部等の形状、寸法（境界条件）に依存する。人は、自身の頭部、外耳の物理的な形状、寸法を知りようがないため、聴覚からの情報を何らかの方法で補正しているものと考えられる。

本研究では、音の回折を無視した幾何学的計算モデルを採用し、回折の影響を補正する方法について検討した。

図1に示すように、2つのセンサを通る x 軸とそれに直交する y 軸からなる座標系において、原点 O から音源を見た水平方向 θ を算出する。原点は両耳間の中間点に相当する。1対のセンサに直接到来する音の行路差 Δl と周波数 f における位相差 $\Delta\phi$ には次式の関係がある

$$\Delta l = \frac{c\Delta\phi}{2\pi f} \quad (1)$$

但し c は音速である。位相差の多義性を考慮し、整数 m を用いて位相差 $\Delta\phi$ を書き換えれば、

$$\Delta l = \frac{c(\Delta\phi + 2\pi m)}{2\pi f} \quad (2)$$

となる。また、行路差 Δl と音源方向 θ の関係は式(3)で与えられる。

$$\Delta l = \sqrt{(L \cos \theta)^2 + (L \sin \theta + S/2)^2} - \sqrt{(L \cos \theta)^2 + (L \sin \theta - S/2)^2} \quad (3)$$

音源と原点 O 間の距離 L が既知の場合に、式(2)と式(3)が等しい条件で方向 θ を解析的に求めるのは容易ではない。そこで、次の条件を導入し近似解を求める。

$$\varepsilon = |\Delta l - \Delta l'| < \varepsilon_0 \quad (4)$$

ここで ε_0 は閾値である。音源数が1個の場合でも、とり得る整数値 m に応じて複数の音源方向が推定されることになる。

具体的な手順は次の通りである。

- (a) 1対のマイクロホンで検出された2系統の音響信号を $x(t)$, $y(t)$ とするとクロスパワースペクトル $G_{lr}(f)$ は、

$$G_{lr}(f) = X(f)Y^*(f) \quad (5)$$

で与えられる。ここで、 $X(f)$, $Y(f)$ はDFT処理後の各スペクトル、 $*$ は複素共役を

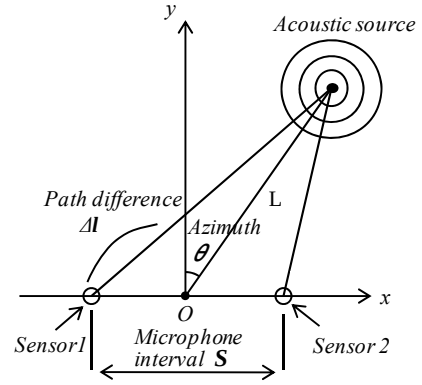


図2 ヘッド音源とセンサの幾何学的関係

示す。 $G_{lr}(f)$ の位相は次式で与えられる。

$$\angle G_{lr}(f) = \angle(X(f)Y^*(f)) \quad (6)$$

- (b) 式(6)で得られた位相差値を用いて式(2)～式(4)より幾何学的に音源方向を求める。

$$\frac{c(\Delta\phi_i + 2\pi m)}{2\pi f_i} \cong \sqrt{(L \cos \theta_{mi})^2 + (L \sin \theta_{mi} + S/2)^2} - \sqrt{(L \cos \theta_{mi})^2 + (L \sin \theta_{mi} - S/2)^2} \quad (7)$$

周波数 f_i で、式(7)が成立する整数 m に対応する複数の方向 θ_{mi} を求める。

- (c) 音源が移動しない場合、その方向は音の周波数に依存せず一定である。この特徴を利用し次式で与えられるハフ変換を用いて複数の音源方向から真の方向を求めた。周波数 f_i における音源方向 θ_{ij} を2次元データ (f_i, θ_{ij}) と考え次式で変換する。

$$K(\theta_j) = \sum_i (-f_i n + \theta_{ij}) \quad (8)$$

周波数に依存しない方向は、式(8)で $n=0$ と置き、度数分布 $K(\theta_j)$ が最大になる方向 θ_j を求める。

3 実験結果

室内の中央で回転台上に固定された、図3に示すヘッドトルソ(4100-D

Bruel&Kjaer 製)の向きを変え,ゴルフボール大のスピーカの方向を推定した(図4)。両耳部のマイクロホンの間隔 S は約 13cm である。24 ビット分解能の A/D 変換ボード(PCI-4474,N.I.製)を装備した PC (GX280,DELL 製)で 2 系統の音響信号を同時収録した。サンプリング周波数は 24kHz でハニング窓を用い,各チャンネルの 1 フレームは 4800 個のデータを含む。図 5 に,様々なヘッドトルソの向きに対する,音源の推定方向と実際の方向の比較を示す。 $-40^{\circ} \sim +40^{\circ}$ の範囲で推定方向は実際の方向と比例する様子が分かる。比例定数は約 1.5 である。音源方向が 50° を越すと繰り返しのよるばらつきが顕著になった。実際の音源方向が -40° である場合の,推定方向の周波数特性を図 6 に示す。式(8)を用いて計算された真の音源方向は -58° であり, 18° の差が生じた。この誤差の原因を探るため,式(2)~式(4)に使われる物理的パラメータの一つ,マイクロホン間隔と推定方向の関係について検討した。図 6 と同じ測定データを用いて,式(3)のマイクロホン間隔 S を仮想的に 17cm に設定して計算された推定方向の周波数特性を図 7 に示す。図 6 と比較して複数ある音源方向の間隔は狭まることが分かる。周波数に依存しない音源方向は -42° で誤差は 2° に減少した。さらにマイクロホン間隔を広げ 21cm に設定すると,推定方向の誤差は増加した。マイクロホン間隔を 17cm に設定した場合の推定方向と実際の方向の関係を図 8 に示す。音源の方向が $-40^{\circ} \sim +40^{\circ}$ の範囲で,推定方向は実際の方向とほぼ一致した。推定値が実際の音源方向と一致するための,等価的なマイクロホンの位置は実際より約 2cm 外側になる。これは音の回折によって音の位相差が大きめに検出されたためと考えられる。音の回



図 3 両耳聴ヘッドトルソの外観
(4100-D, Brüel&Kjaer)

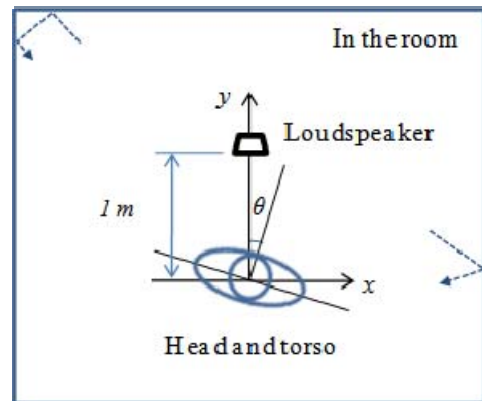


図 4 ヘッドトルソとマイクロホンの配置

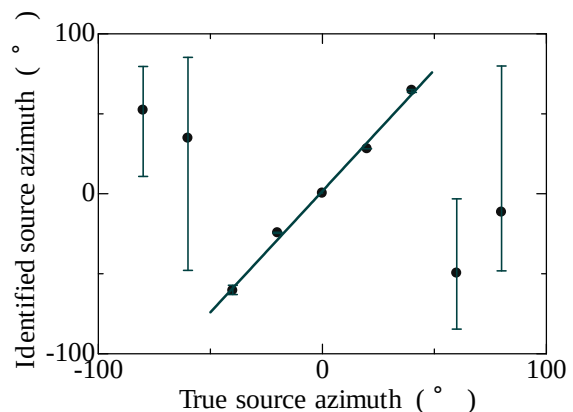


図 5 実際の音源方向と推定方向の比較
($S=13$ cm)

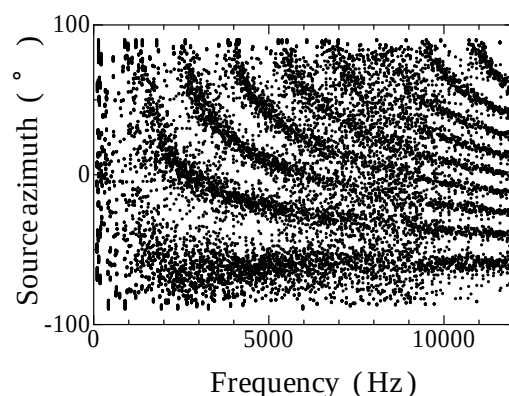


図 6 推定された音源方向の周波数特性
($S=13$ cm)

折を無視した幾何学的計算モデルを用いた場合でも、物理的パラメータの一つであるマイクロホン間隔を調整して、実際の音源方向が比較的正確に推定できることを示した。

4. おわりに

メンフクロウの聴覚中枢で音源の水平方向が検出されるメカニズムと類似のアルゴリズムを、両耳聴ヘッドトルソで検出された音響信号に適用し、1個の小型マイクロホンの方向を推定した。ヘッドトルソの向きを変えて、様々な角度でスピーカ方向を推定した結果、次のような結論が得られた。

- (1) 推定された方向と実際の音源方向に差異がみられた。
- (2) 幾何学計算に用いるマイクロホン間隔を調整すると、 $-40^{\circ} \sim +40^{\circ}$ の範囲で実際の音源方向をほぼ推定できた。

今後は本アルゴリズムをロボット聴覚に適用する予定である。

謝辞 本研究は平成21年度に日本大学学術研究助成(個人研究)を受けた。

[参考文献]

- 1) L. A. Jeffress, "A place theory of sound localization", J. Comp. Physiol. Psychol. 41, 1948, p.p.35-39.
- 2) M. Konishi, "Study of sound localization by owls and its relevance to humans," Comp. Biochem. Physiol. A, 126, 2000, p.p.459-469.

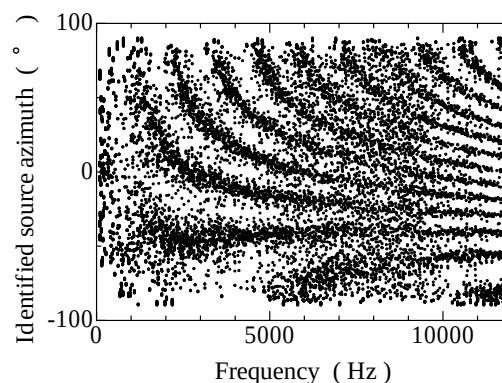


図7 推定された音源方向の周波数特性
(S=17 cm)

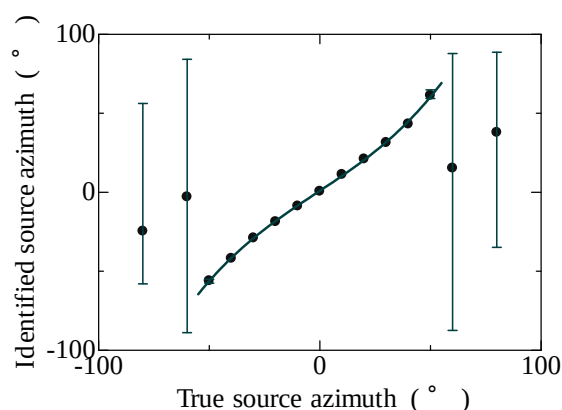


図8 実際の音源方向と推定方向の比較
(S=17 cm)

- 3) P. Smaragdis and P. Boufounos, "Position and trajectory learning for microphone arrays", IEEE Trans, on Audio, Speech and Language Processing, 15-1,2007,p.p.358-368.
- 4) S. Vesa, "Binaural sound source distance learning in rooms", IEEE Trans. on Audio, Speech and Language Processing,17-8,2009,p.p.1498-1507.
- 5) R. Shimoyama and K. Yamazaki, "Multiple acoustic source localization using ambiguous phase differences under reverberative conditions," Acoust. Sci. & Tech., 25, 2004, p.p.446-456.