

# 耳珠のある能動耳介システムとその動作について

Active pinnae with tragus and their motion

公文誠, 尾堂航, 木元大輔

Makoto KUMON, Wataru ODO, Daisuke KIMOTO

熊本大学大学院自然科学研究科

Graduate School of Science and Technology, Kumamoto University

kumon@gpo.kumamoto-u.ac.jp

## Abstract

本研究では二つのマイクロホンを用いて音源定位を行うロボットシステムとして、マイクロホンの周囲に可動式の反射板である能動耳介を利用するものを考える。耳介の姿勢に伴って伝達特性が変化することを活用することで、音源定位性能の向上を目指す。このためには伝達特性の変化が顕著であることが重要である。動物では耳珠が耳介の効果を強調することが知られているため、本研究では音源定位に向けた能動耳介の効果改善のため耳珠に相当する部位を導入することを提案する。実際に耳珠を備えた能動耳介システムを開発し、音源定位のための動作に反映するための基礎性能を調べたのでこれを報告する。

## 1 はじめに

ロボットが自律的に動作するには、判断のためにロボット周囲の環境を適切に認識することが不可欠である。環境認識の代表的な方法はカメラやレーザ距離計などを用いた光学的な方法が数多く提案されているが、人間の住環境のように壁面など障害物によるオクルージョンの問題や、カメラの視野外は認識できないなど課題もある。一方、音信号を用いることが出来れば、対象が障害物の隠れるような場合でも、これを見つけることが可能であると同時に、自動車のクラクションや電話の呼び出し音のような音記号の理解を通じた環境理解も可能となる。音を通じた環境理解の重要な基礎能力に、音源の位置や方向を推定する音源定位があり、ロボットに向けてはマイクロホンアレイと音響信号処理による方法が種々提案されている。

さて、生物の場合の音源定位を考えると、多くは実用的な範囲を二つの耳で実現している。二つに限られた受聴点の信号から音源定位を実現する生物の能力は興味深く、

このような機能を実現する機序の一つに、頭部など身体を動かすことが指摘されている (Blauert[Blauert 96] など)。さらに、ネコなど一部の動物では、耳介そのものを随意筋によって能動的に動作させることが可能であるが、この耳介動作によっても音源定位能が改善されることが知られている (Populin[Populin 98], Heffner[Heffner 82] など)。

このような生物に倣って、二つあるいは少数のマイクロホンを用いたロボットシステムで高性能の音源定位を実現しようとする試みがある。例えば Kimら [Kim 13] は頭部動作を利用することで音源方向の推定性能が向上することを示している。また、マイクロホンが二つに限られると正中面内の方向が認識できないことが指摘されているが、著者ら [Kumon 05] や Hörnstein[Hörnstein 06] などは耳介を利用することで音源の上下を区別する方法を提案している。あるいは、動く耳介を活用する先行研究として、著者ら [Kumon 11] や 金ら [金 12] の研究が挙げられる。

これらの動作する耳介を用いたロボットシステムでは、耳介の動作に伴って音源からマイクロホンまでの音信号の伝達特性が変化するが、この伝達特性の変化が既知であれば、例えば音信号の到来方向を認識するなどが可能となる。また、著者らの音源の上下を区別する手法 ([Kumon 05]) では、耳介による伝達要素の周波数特性を利用するものだったので、より望ましい特性を呈する耳介の姿勢を求められれば、音源定位能が改善されることが期待される。一方、このことが機能するためには、耳介の姿勢変化によって十分に伝達特性の変化が明瞭であることが必要である。これまでの著者らの結果 [Kumon 11] では、主に高周波数帯域のパワーが姿勢変化に伴ってなだらかに変化することが示されていたが、この変化を顕著にすることが定位性能に貢献することが考えられる。

ところで、動物の外耳には耳介以外にも様々な部位があり、耳珠と呼ばれるものが存在する。耳珠とは動物などの耳に見られる耳孔前方にある小突起で、耳介の効果を強調するとの指摘がある [本田 85]。そこで本研究では能

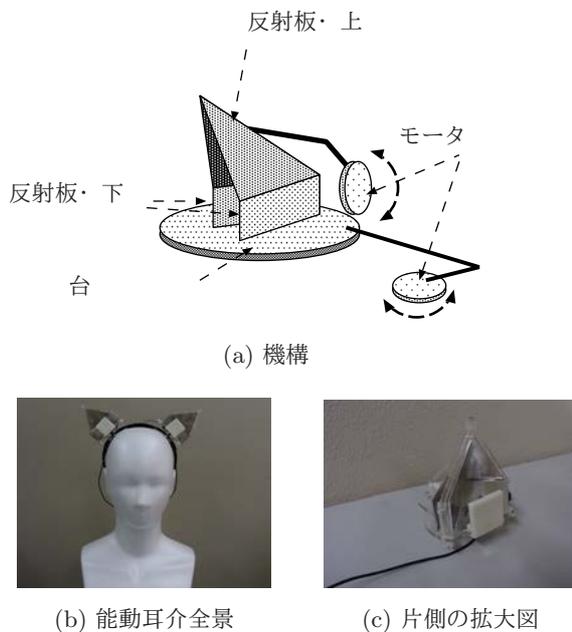


図 1: 耳珠を持つ能動耳介システム

動耳介に合わせて耳珠を耳介開口部に設け、姿勢による周波数特性の変化を強調することを考える。まず、これまでに提案した能動耳介のシステムに耳珠に相当する構造を取り付けた耳介システムを開発し、この特性を調べることとした。また、得られた特性の評価として、いくつかの耳介の姿勢で音収録を行い音源の方向を推定することで行った。

## 2 能動耳介と耳珠

### 2.1 実験装置

まず本研究で用いる能動耳介装置について説明する。図1は装置についてまとめたものである。

図1(a)に示すように耳介は基礎を成す台の上に取り付けられており、アルミ板の反射板を組み合わせたテント型の構造になっている。反射板は上下に分かれており、上部はモータによって姿勢を前後に動かすことが出来、下部は台に固定されている。本研究では用いないが、耳介を取り付けている台もモータによって動かすことが可能で、耳介の左右方向の自由度に対応する。下部の台の中央にはマイクロホンが上向き(耳介の上部方向)に取り付けられている。このような能動耳介を二つ組み合わせ、マネキン頭部に搭載した状態で実験に供した(図1(b))。

本研究の中心となる耳珠[本田 85]. は、図1(c)に示される写真の白い部分であり、上述する耳介の開口部の下半分ほどを覆う正方形の板で模擬した。各耳珠は4mm厚のプラスチック板で、以下では30mm四方のものと40mm四方のもの二種類を検討している。

### 2.2 耳珠の効果: 周波数特性

耳介の姿勢の変化で周波数特性がどのように変化するかを確認するため耳珠を取り付けた場合と取り付けていない場合での伝達特性を検証した。音源は頭部の正面方向1.5mの距離に設置するものとし、頭部の正面高さを基準位置として仰伏角方向に $-20^{\circ} \sim 20^{\circ}$ まで $10^{\circ}$ 刻みで計5点から白色雑音を試験信号として用いた。また、それぞれの音源位置ごとに耳介の姿勢を右耳6パターン、左耳6パターンの計36パターンで収録した。なお、音信号はサンプリング周波数44.1kHzで録音しており、再生装置の制限で呈示される白色信号は16kHzまでの帯域に制限されている。

音源の方位角については、IPDを用いて比較的良好に推定ができると想定できるので、ここでは特に二つのマイクロホンで難しいと考えられる音源の仰角方向について検証することとし、ILDに相当するパワースペクトルに着目する。

図2に収録した白色信号のスペクトログラムを示す。図の列は、左から順に耳珠を取り付けなかった場合、小型(30mm四方)の耳珠を取り付けた場合、大型(40mm四方)の耳珠を取り付けた場合の結果を示しており、図の行は音源の仰角に対応する。各図の横軸は時間、縦軸が周波数を表わし、図中の色によってパワーを示している。

いずれの場合も縞状のパターンが見られ、安定したスペクトル構造が得られている。また、耳珠を取り付けることでパターンが変化していることも確認される。特に耳介の効果は高周波数帯域(10kHz以上)で顕著であり、耳珠も同じ帯域で影響が際立っている(黒枠)。具体的な変化を挙げると、音源方向が $-20^{\circ} \sim -10^{\circ}$ のとき耳珠のない場合では周波数特性が酷似している一方、耳珠を取り付けた場合、13kHz付近のパワーが小さくなっており、これらの別が明瞭であること、また、音源方向が $0^{\circ} \sim 20^{\circ}$ のときには14~16kHzにノッチがはっきりするなどがある。音源方向の推定法では、この周波数特性の違いを手がかりに定位を行うので、このように特性の変化が顕著になることは重要である。この観点からすると、逆に耳珠の大きなものでは、 $-10^{\circ} \sim 0^{\circ}$ の方向での周波数特性の変化が乏しく、耳珠の小さなものに比べ適当とは言い難い。このことから、単に耳珠を取り付ければ良いというわけではなく、適当な大きさのものを選択する必要があることが示唆される。

先の条件に替えて、耳介を前方に傾斜させた姿勢( $\rho_2$ )でも同様の試験を行った。この結果を図3に示すが、先と同様に耳珠の影響が確認される。例えば音源方向が $-20^{\circ}$ および $-10^{\circ} \sim 0^{\circ}$ のとき、それぞれ10~16kHzと1.5~1.6kHzにノッチが見られることが挙げられる。特に重要なのは、同じ音源方向であっても耳介の姿勢が $\rho_1$ から $\rho_2$ に変化することで周波数特性がはっきり変化する

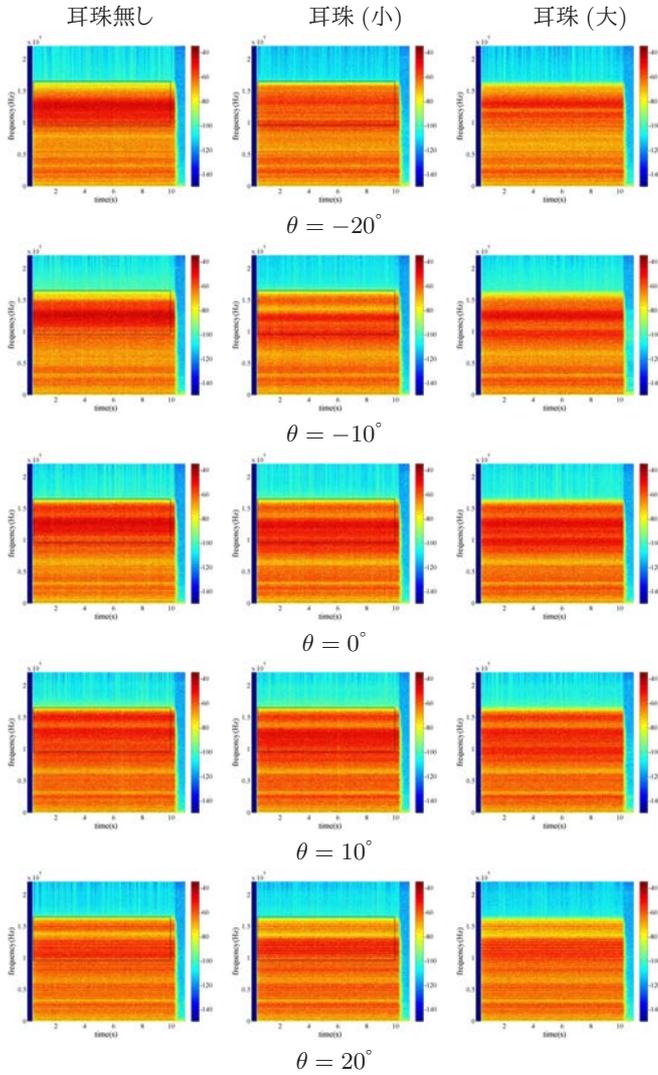


図 2: スペクトログラム (耳珠の影響: 姿勢  $\rho_1$ )

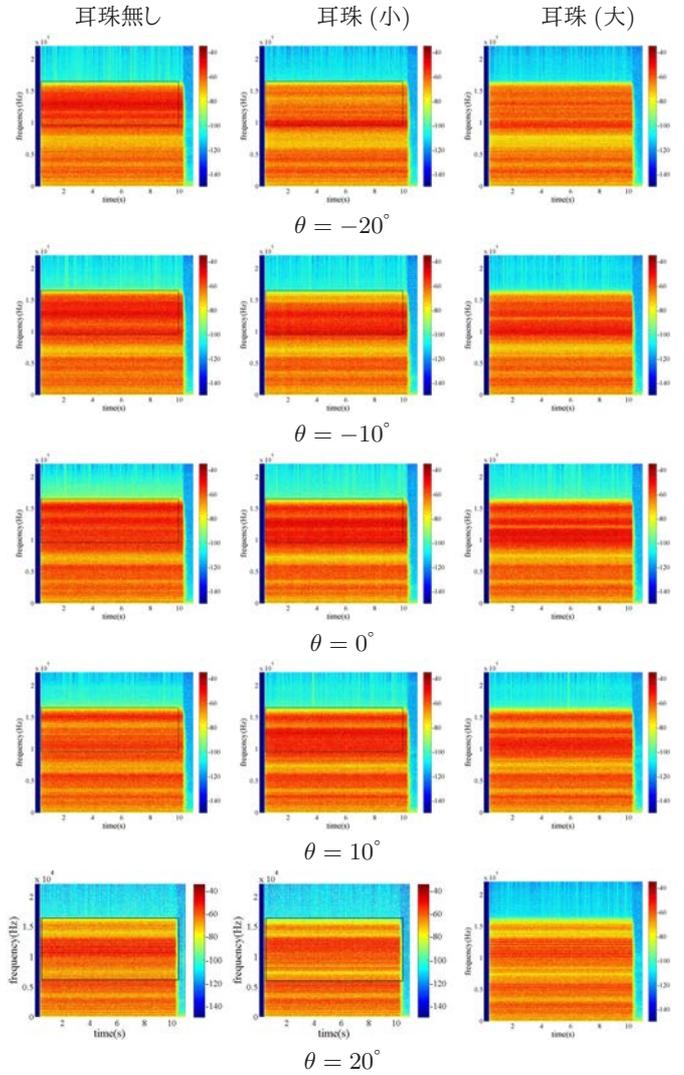


図 3: スペクトログラム (耳珠の影響: 姿勢  $\rho_2$ )

ことである。例えば  $\theta = -20^\circ$  の場合、図 2, 3 を比較すれば、耳珠を取り付けない場合では周波数特性はほとんど変化しないが、耳珠 (小) を取り付けた場合、耳介を傾ける ( $\rho_2$ , 図 3) と 10kHz 付近にピークが出現するという明瞭な変化がある。音源方向を識別する上で耳介の方向を能動的に操作する効果が生じることになるので、能動耳介と耳珠の組み合わせが有用であると言える。

### 3 音源方向の推定

前節に説明した耳珠を有する能動耳介システムを用いて音源方向の推定を行う。本研究では頭部正面を基準とする座標系から見た音源の方向を仰角、方位角で考え、それぞれを  $\theta, \phi$  と表すことにする。また、左右の耳介の姿勢を  $\rho := (\rho_l, \rho_r)$  で表すものとする。

以下ではまず定位の方法を説明し、その後、実際の定位結果を示す。

#### 3.1 推定方法

二つのマイクロホンによる音源方向の推定に用いる特徴量として両耳間レベル差 (Interaural Level Difference; ILD) と位相差 (Interaural Phase Difference; IPD) が良く知られている [Blauert 96]。ここでは、ILD, IPD を周波数領域上の  $N$  個の周波数点上で表現した特徴量ベクトルとして扱うこととし、これらの量を  $z_{ILD}, z_{IPD}$  と表す。

定位を行うために、音源方向と対応づけて、これらの音響特徴量を事前に測定したデータベースを作成しておく。これらの特徴量のことを以下では規範の特徴量と呼び、 $(\theta, \phi)$  方向の規範の ILD, IPD を  $z_{ILD}^d(\theta, \phi, \rho)$ ,  $z_{IPD}^d(\theta, \phi, \rho)$  のように添字  $d$  を付して表すものとする。

耳介の姿勢  $\rho$  にあつて測定された ILD, IPD の特徴量  $z_{ILD}, z_{IPD}$  から音源方向を推定するため、ILD, IPD について次のような尤度を考える。

$$\begin{aligned}
 & l_{ILD}(\theta, \phi | \rho, z_{ILD}) \\
 &= \sum_{i=1}^N \exp \left\{ - (z_{ILD,i} - z_{ILD,i}^d(\theta, \phi, \rho))^2 \right\},
 \end{aligned} \tag{1}$$

$$l_{\text{IPD}}(\theta, \phi | \boldsymbol{\rho}, \mathbf{z}_{\text{IPD}}) = \sum_{i=1}^N \exp \{ \cos(z_{\text{IPD},i} - z_{\text{IPD},i}^d(\theta, \phi, \boldsymbol{\rho})) - 1 \}. \quad (2)$$

ここで  $i$  は各特徴量の  $i$  番目の要素を示す。次に、これらの尤度をまとめて結合尤度を

$$l(\theta, \phi | \boldsymbol{\rho}, \mathbf{z}_{\text{ILD}}, \mathbf{z}_{\text{IPD}}) = l_{\text{ILD}}(\theta, \phi | \boldsymbol{\rho}, \mathbf{z}_{\text{ILD}}) l_{\text{IPD}}(\theta, \phi | \boldsymbol{\rho}, \mathbf{z}_{\text{IPD}}). \quad (3)$$

とする。

今、 $k$  回の観測の後、音源方向について事前分布  $p_k(\theta, \phi)$  が与えられたとし、 $k+1$  回目の観測によって (3) の結合尤度が得られたとすると、繰り返しベイズ推定による音源定位法 [Kumon 13] に従って事後分布を

$$p_{k+1}(\theta, \phi) = \frac{l(\theta, \phi | \boldsymbol{\rho}, \mathbf{z}_{\text{ILD}}, \mathbf{z}_{\text{IPD}}) p_k(\theta, \phi)}{\int l(\theta, \phi | \boldsymbol{\rho}, \mathbf{z}_{\text{ILD}}, \mathbf{z}_{\text{IPD}}) p_k(\theta, \phi) d\theta d\phi} \quad (4)$$

のように更新する。

### 3.2 定位実験

ここでは、音源を仰角方向に 5 通り、方位角方向に 5 通りの計 25 点を対象に音源方向の推定(音源定位)を行うこととし、データベースには先と同じ白色雑音による規範データを作成し、試験信号には図 4 に示す音楽信号を用いることとした。

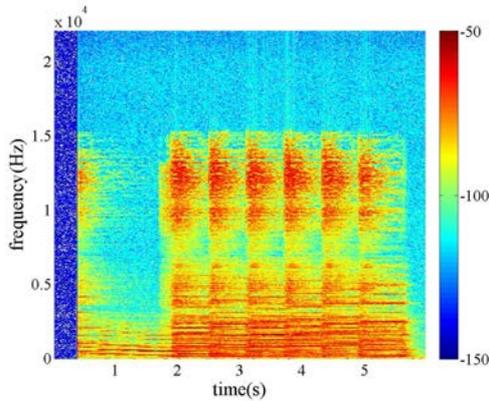


図 4: 定位実験に用いた信号

まず、音源を  $(\theta, \phi) = (-10^\circ, 20^\circ)$  に置いた時、耳介を S1: 両方の耳を直立させた場合 (姿勢  $\boldsymbol{\rho}_1$  とする), S2: 右耳を直立, 左耳を前方に小さく傾斜させた場合 (同  $\boldsymbol{\rho}_2$ ), S3: 右耳を直立, 左耳を前方に大きく傾斜させた場合 (同  $\boldsymbol{\rho}_3$ ), A: 右耳を直立させた状態, 左耳を観測毎に直立させた状態から前方に 6 段階に傾斜させた場合の 4 つの条件で推定を行った。観測の回数は S1, S2, S3, A のいずれの場合も 6 回であり、初期の事前分布には一様分布を用いた。ここで A は片耳が取る全ての場合で一度ずつ音信号を取得するもので、基本的な耳介動作の一つである。

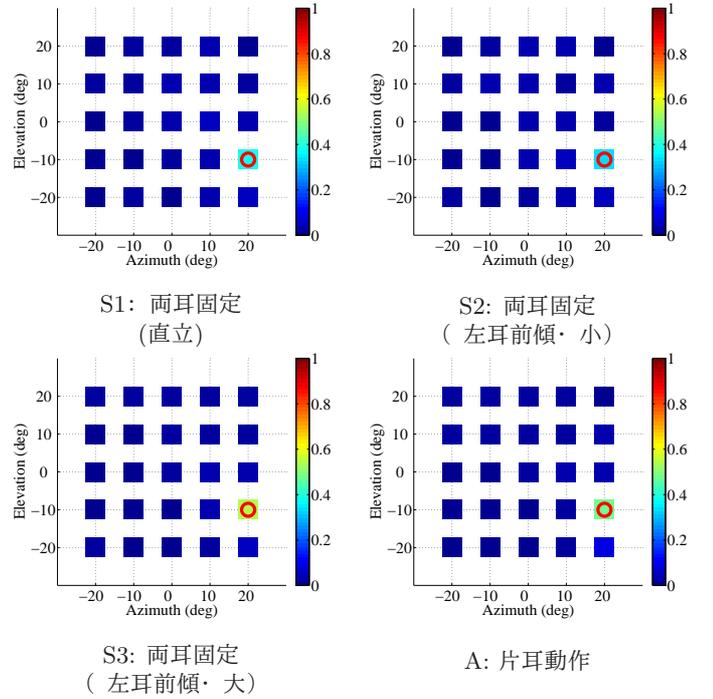


図 5: 音源方向の推定結果

適当な時間区間の信号に対する定位結果として、各条件での 6 回目の事後分布  $p_6$  を図 5 に示す。各図は 25 個のセルからなっており、各セルが仰角・方位角に対応し、そのセルの色が音源の事後分布の値を示している。図中の赤丸が音源の正解方向を示しており、S1, S2, S3, A のいずれの場合も音源の方向に対応するセルの値だけが大きくなっている。この例では耳介の姿勢や動作に依らず、音源方向が正しく定位されているとも見えるが、 $\boldsymbol{\rho}_3$  では推定された値がやや小さいなどの違いがある。紙面の都合で割愛したが、他の音源方向では推定結果にばらつきがあるなど、図 5 のような事後分布だけでは性能を測ることが出来ないため、定位性能を評価する系統的な指標を導入することとした。

収録された音信号に独立な白色雑音を混合し、騒音下での音源定位を模擬した場合を考える。なお、各試行毎に S/N 比 (SNR) が一定となるよう収録音の各時間区間のパワーに合わせて白色信号のパワーを調整している。また、音源定位性能については以下で定義される  $F$  値 [Chinchor 92] を用いることとした。

$$F = \frac{2RP}{R+P}, \quad (5)$$

ここで

$$P = \frac{\# \text{ of estimated sound sources}}{\# \text{ of estimated sound source candidates}}$$

$$R = \frac{\# \text{ of estimated sound sources}}{\text{Total } \# \text{ of sound sources}}.$$

で、 $F$  値は  $[0, 1]$  の範囲を取り、大きいほど定位性能が良いことを示す指標である。

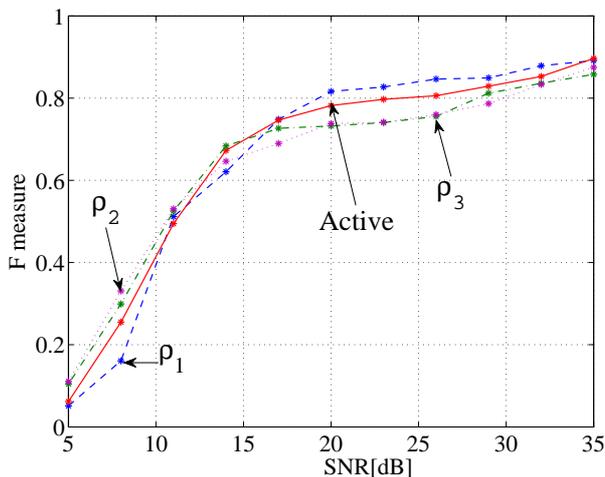


図 6: 音源定位のロバスト性

SNR の変化の下での定位性能を図 6 に示す．図の横軸が SNR，縦軸に  $F$  値を示しており，先の S1 と S2, S3, A の各条件での性能をまとめて表示している．

収録環境が良好な場合 (SNR が大きい場合)，どの場合でも同じ程度の定位性能が得られている．20dB 以上の領域では，両耳を直立させた  $\rho_1$  が最良の結果を示しており，次いで動作させた場合 (A)，片方の耳介を前に傾斜させた場合 ( $\rho_2, \rho_3$ ) の順になっている．これは，音信号が正確に得られる場合は，耳介を立て開口部を広く保つ  $\rho_1$  が有効であるので，自然な結果とも言える．一方，SNR が 15dB 付近の悪条件では状況が異なり， $\rho_1$  は SNR の低下とともに急激に定位性能が劣化し，逆に耳介を伏せた  $\rho_2$  や  $\rho_3$  の方が  $F$  値の点ではやや良い結果となった．ここで，耳介を観測毎に前方へ傾斜させた場合 (A の場合)， $\rho_1$  と  $\rho_2$  や  $\rho_3$  の中間的な性質を示しており，SNR が高いところでは  $\rho_1$  に近い性能を呈するとともに，SNR の低下時は  $\rho_2$  や  $\rho_3$  に近い性能であった．このことは，耳介の姿勢を変化させることで，音源の方向や雑音の程度に応じて伝達特性を変化させることで，それぞれに適した観測を活用し，ロバストな音源定位を実現できる可能性を示唆している．なお，SNR が 12dB 以下になると  $F$  値は 0.5 以下と定位性能は悪く，ここでの  $F$  値の大小を論じることは意味のあることとは言えない．

#### 4 おわりに

本研究では動物などをヒントに耳介を能動的に動作させることで，音源定位性能を改善することを目的とし，耳珠の導入を提案するとともに，実際の音源定位においてその効果を調べた．その結果，耳珠を取り付けることで，単に伝達特性が変化するだけでなく，特に仰角方向の音源方向の推定に重要な耳介ノッチが明瞭になること，また耳介動作による伝達特性の変化が明らかになることを実際の装置によって確認した．また，耳介を観測の度に少し

ずつ前方に傾斜させる動作を考え，この動作の下で音源定位を行った場合の定位性能を調べた所，雑音下でロバスト性を改善出来ることが示され，耳介動作が音源定位性能の向上に繋がる例を挙げる事が出来た．

今回は最も基本的な動作として前方に傾斜し続ける動作を考えたが，実際には定位情報にあたる事後分布を用いて次を取るべき姿勢を計算することが効果的だと考えられる．この音源定位の観点からどのような姿勢を良いとするかは今後の検討が必要である．

また，SNR が著しく低下した場合には定位性能が悪くなるが，耳介姿勢を適切に選ぶことで改善が可能か，可能であればどのようにすれば良いかについても今後の課題である．

#### 参考文献

- [Blauert 96] Blauert, J.: *Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization*, The MIT Press, rev sub edition (1996)
- [Chinchor 92] Chinchor, N.: MUC-4 EVALUATION METRICS, in *Proceedings of Fourth Message Understanding Conference (MUC-4)*, pp. 22–29 (1992)
- [Heffner 82] Heffner, R., Heffner, H., and Stichman, N.: Role of the elephant pinna in sound localization, *Animal Behaviour*, Vol. 30, No. 2, pp. 628–629 (1982)
- [Hörnstein 06] Hörnstein, J., Lopes, M., Santos-Victor, J., and Lacerda, F.: Sound localization for humanoid robots - building audio-motor maps based on the HRTF, in IEEE, ed., *Proceedings of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1170–1176 (2006)
- [Kim 13] Kim, U.-H., Nakadai, K., and Okuno, H. G.: Improved Sound Source Localization and Front-Back Disambiguation for Humanoid Robots with Two Ears, in *IEA/AIE*, pp. 282–291 (2013)
- [Kumon 05] Kumon, M., Shimoda, T., Kohzawa, R., Mizumoto, I., and Iwai, Z.: Audio Servo for Robotic Systems with Pinnae, in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 885–890 (2005)
- [Kumon 11] Kumon, M. and Noda, Y.: Active Soft Pinnae for Robots, in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 112–117 (2011)
- [Kumon 13] Kumon, M., Kimoto, D., Takami, K., and Furukawa, T.: Bayesian Non-Field-of-View Target Estimation Incorporating an Acoustic Sensor, in *Proceedings of 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3425–3432 (2013)
- [Populin 98] Populin, L. and Yin, T.: Pinna movements of the cat during sound localization, *Journal of Neuroscience*, Vol. 18, No. 11, pp. 4233–4243 (1998)
- [金 12] 金 天海, 中臺 一博, 辻野 広司: ウエアラブル人工可動耳介-音追従動作による音源定位能力の向上-, 第 30 回日本ロボット学会学術講演会, 日本ロボット学会 (2012)
- [本田 85] 本田 学: 耳珠のはたらき, *耳鼻咽喉科臨床*, Vol. 78, (1985)