ガウス回帰に基づく両耳間レベル差の補間

Interpolation of interaural level difference based on Gaussian regression

木元 大輔，尾堂　航，公文　誠
Daisuke KIMOTO，Wataru ODO，Makoto KUMON
熊本大学
Kumamoto University
d.kimoto@ick.mech.kumamoto-u.ac.jp

Abstract

著者らはバイノーラル聴覚ロボットでの規範データを用いた音源定位手法として，音響特徴量の不確かさを考慮した音源定位手法を提案している．この際，規範データ取得に工数がかかるため，これを軽減する方法として疎な収録データを補間して規範データを擬似的に取得する方法が考えられる．このような方法として本報告では補間された点での不確かさを適当なモデルの下で推定するガウス回帰に基づいた補間方法に着目した．この方法より得られる音響特徴量とその不確かさを利用して，不確かさを考慮した音源定位手法を用いた場合，一定条件の下で，定位性能が向上することが確認できた．

1　はじめに

音源定位は，音の情報を利用して周辺環境を認識する上で基本的な聴覚機能であり，ロボットにおいても，ロボット周辺の環境認識を行う方法の 1 つとして有用である．これはロボット聴覚として盛んに研究されている[奥乃, 2010]．その中でも，人間や動物は 2 つの耳で音源位置の情報を得ていることから，2 つのマイクロホンを持つロボットを用いたバイノーラル聴覚ロボットにおける研究が行われている[奥乃, 2002]．バイノーラル聴覚ロボットを用いた音源定位手法の 1 つとしては，事前に学習した音響特徴量を規範データとし，それとマッチングを行うことで，仰伏角及び方位角方向を推定する方法が提案されている[章, 2008]．

人間や動物には，耳に耳介と呼ばれる音の反射・集音を果たす器官が存在している．この耳介の影響により，音の到来方向に応じて音響特性が変化することが知られている[Shaw, 1968]．このことから著者らのグループは，到来方向により音響特性が変化することで音の到来方向が推定しやすくなると考え，マイクロホン近辺に動物の耳介の形状に類似した反射板を取り付けた装置[野田, 2012]を提案してきた．しかしながら，頭部形状による影響は複雑であり，正確にこれをモデル化することは困難で，常に一定の不確かさを考慮する必要がある．著者らはこのことを踏まえ，音響特徴量の不確かさを考慮した音源定位法を提案した[木元, 2013]．この方法を用いることで従来の不確かさを考慮していない音源定位手法に比べ，定位性能が向上することを確認している．

規範データを使用する音源定位手法では，環境ごとに規範データを作成しなければならず，多くの規範データを観測より得ることは困難である．この問題を解決する方法として伝達特性を線形補間する方法が提案されている[中村, 2012]．しかし，この方法では補間より得られる音響特徴量の不確かさは考慮しておらず著者らが提案している不確かさを考慮した音源定位手法を使用することができない．そこで，本研究では音響特徴量の補間方法としてベイズ回帰に基づく方法を提案し，補間より作成した規範データの不確かさも同時に推定し，その検証を行う．

2　バイノーラル聴覚ロボットでの音源定位に用いる特徴量

ロボットに搭載されているマイクロホンに受聴される音信号について考える．環境やロボット自身の影響により，ロボットに搭載されている 2 つのマイクロホンに収録される音は原信号 $s_o(\omega)$ とは異なったものとなる．今，ロボットを基準とした音源位置を $\boldsymbol{x}$ とすると，ある周波数 ω の音源から 2 つのマイクロホンへの伝達関数は $H_l(\omega, \boldsymbol{x})$，$H_r(\omega, \boldsymbol{x})$ と表わすことができる．これより，ロボットに搭載されている 2 つのマイクロホンが受聴する音信号 $s_l(\omega, \boldsymbol{x})$，$s_r(\omega, \boldsymbol{x})$ は Figure 1 に示すような

$$
\begin{aligned}
s_l(\omega, \boldsymbol{x}) &= H_l(\omega, \boldsymbol{x}) s_o(\omega) \\
s_r(\omega, \boldsymbol{x}) &= H_r(\omega, \boldsymbol{x}) s_o(\omega)
\end{aligned} \tag{1}
$$

の関係がある．

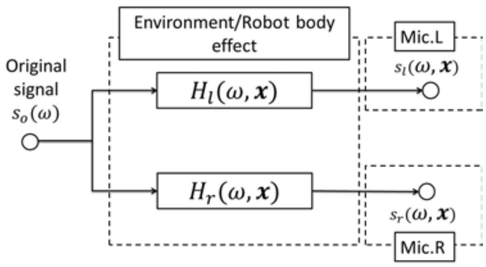


Figure 1: 左右のマイクロホンに受聴される音信号

両耳間レベル差 (ILD) は 2 つのマイクロホン間での音の大きさの違いに相当する．ロボットに搭載されているマイクロホンに受聴される音信号が式 (1) で与えられるとすると，周波数 ω における ILD の値を $Z(\omega, \boldsymbol{x})$ と表すと，

$$
\begin{aligned}
Z(\omega, \boldsymbol{x}) &\equiv 20 \log |s_l(\omega, \boldsymbol{x})| - 20 \log |s_r(\omega, \boldsymbol{x})| \\
&= 20 \log |H_l(\omega, \boldsymbol{x}))| - 20 \log |H_r(\omega, \boldsymbol{x})|
\end{aligned} \tag{2}
$$

で表わすことができる．原信号 $s_o(\omega)$ の影響が消去され ILD に影響がないことがわかる．ILD が $\boldsymbol{x}$ ごとに異なっていれば，予め測定した規範データと観測量を比べることで，原信号 $s_o(\omega)$ によらず音源定位に利用できる．

3　両耳間レベル差における不確かさ

式 (2) の ILD の実際の観測量には不確かさが含まれる．著者らの提案する不確かさを考慮した音源定位手法[木元, 2013]ではこの不確かさを適当なガウス分布に従うと考えてきたが，ここではこの仮説の妥当性について検証する．このため本研究ではコルモゴロフースミルノフ検定 (KS 検定) を用いて，観測から得られた ILD がガウス分布に従うと言えるのかを確認する．

まず，帰無仮説を標本が確率密度関数 $F(z)$ から発生するとする．標本が $z_1, z_2, \cdots z_n$ で与えられたとすると，この標本の経験分布は

$$
\begin{aligned}
F_n(z) &= \frac{1}{n} \sum_{i=1}^{n} R_i(z) \\
R_i &= \begin{cases} 1(z_i \leq z) \\ 0(z_i > z) \end{cases}
\end{aligned} \tag{3}
$$

となる．これより KS 検定統計量は

$$
D = \sup_z |F_n(z) - F(z)| \tag{4}
$$

で与えられる．ここで，ガウス分布の場合 $F(z)$ は

$$
F(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \tag{5}
$$

で与えられる．

KS 検定統計量 D の有意確率は

$$
\Pr(D\sqrt{n} > \lambda) = 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2j^2\lambda^2} \tag{6}
$$

で与えられる．これより，有意水準 5% で与えられた場合には λ　1.36 となり，$D\sqrt{n}$ が 1.36 以上の場合，帰無仮説を棄却し，標本 $z_1, z_2, \cdots z_n$ は確率密度関数 $F(z)$ と一致しないという結論となる．

実際に KS 検定を行った．まず，本研究では Figure 2 に示すマイクロホン近傍に耳介を取り付けたバイノーラル聴覚ロボットを使用している．検定に使用する音は，Figure 3 に示すような環境で収録を行い，Figure 4 に示すような位置にスピーカを設置し収録を行った．音源位置は $x = -0.5$m, $y = 2.0$m を 1，$x = -0.4$m, $y = 2.0$m を 2，...，$x = 0.5$m, $y = 1.0$m を 121 と番号を割り当てている．対象音としては，白色雑音を使用し 32000Hz でサンプリングを行い，FFT 長 1024 点で処理を行った．また，ILD に周波数方向のフィルターをかけ，平滑化を行った結果を用いている．

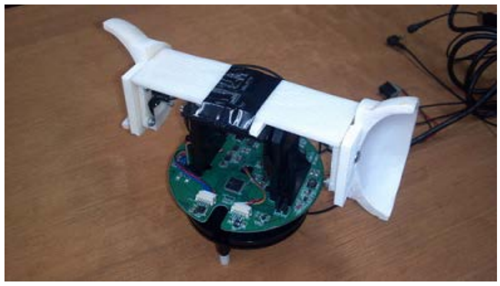


Figure 2: バイノーラル聴覚ロボット

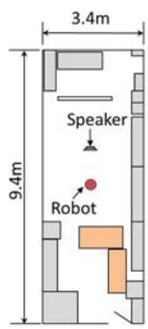


Figure 3: 収録環境

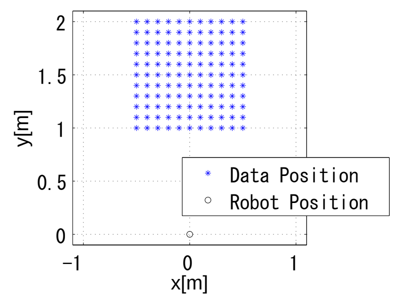


Figure 4: 収録位置

検定を行った結果を Figure 5 に示す．横軸が音源位置を表す番号，縦軸が周波数，色が検定結果を表しており，

黒色となっている部分が帰無仮説を棄却し，ガウス分布ではないとされた部分となる．

本来，帰無仮説が棄却されなかったことから，帰無仮説を採択することはできないが，一点のみでなく 98.26% の音源位置及び周波数帯域でガウス分布ではないとされていないことから，本研究では，ILD はガウス分布に従うとした．

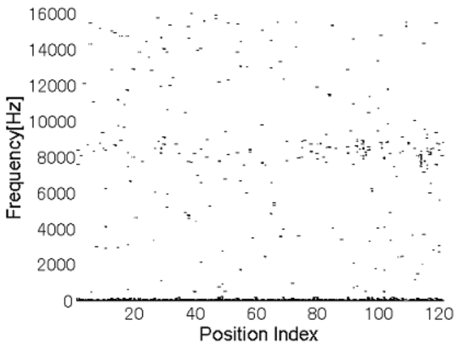


Figure 5: 検定結果

4 両耳間レベル差 (ILD) の補間

この節は，参考文献[ビショップ, 2009]に基づき ILD の補間方法について説明する．

ある周波数での観測される ILD z_n を

$$z_n = Z_n + w_n \tag{7}$$

とする．ここで，$Z_n = Z(\omega, \boldsymbol{x}_n)$ であり位置 $\boldsymbol{x}_n$ で観測される理想的な ILD を表し，$\boldsymbol{x}_n$ は n 番目の観測位置，w_n は n 番目の観測値に含まれるノイズで，独立同分布であると考える．ここで，ノイズはガウス分布に従い

$$p(z_n|Z_n) = \mathcal{N}(z_n|Z_n, \beta^{-1}) \tag{8}$$

であるものとする．また β はノイズの精度を表す超パラメータである．ノイズは各データに対して独立に決まるため，$\boldsymbol{Z}_{1:N} = (Z_1, ..., Z_N)^T$ が与えられた下での ILD $\boldsymbol{z}_{1:N} = (z_1, ..., z_N)^T$ の同時分布は以下の等方的なガウス分布に従う．

$$p(\boldsymbol{z}_{1:N}|\boldsymbol{Z}_{1:N}) = \mathcal{N}(\boldsymbol{z}_{1:N}|\boldsymbol{Z}_{1:N}, \beta^{-1}\boldsymbol{I}_N) \tag{9}$$

ここで，$\boldsymbol{I}_N$ は $N \times N$ の単位行列とする．ガウス過程のモデルとしてカーネル関数を式 (10) と考えるとき

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \theta_0 \exp\left(-\frac{\theta_1}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2\right) \tag{10}$$

周辺分布 $p(\boldsymbol{Z}_{1:N})$ は，平均が $\boldsymbol{0}$ で共分散が $K_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ を i, j 要素に持つグラム行列 $\boldsymbol{K}$ で与えられるガウス分布となる．

$$p(\boldsymbol{Z}_{1:N}) = \mathcal{N}(\boldsymbol{Z}_{1:N}|\boldsymbol{0}, \boldsymbol{K}) \tag{11}$$

位置 $\boldsymbol{x}_1, ...\boldsymbol{x}_N$ で条件づけられたときの周辺分布 $p(\boldsymbol{z}_{1:N})$ を求めるためには，$\boldsymbol{Z}_{1:N}$ についての積分が必要であるが，それは以下のように求まる．

$$\begin{aligned} p(\boldsymbol{z}_{1:N}) &= \int p(\boldsymbol{z}_{1:N}|\boldsymbol{Z}_{1:N})p(\boldsymbol{Z}_{1:N})d\boldsymbol{Z}_{1:N} \\ &= \mathcal{N}(\boldsymbol{z}_{1:N}|\boldsymbol{0}, \boldsymbol{C}_{1:N}) \end{aligned} \tag{12}$$

ここで，共分散行列 $\boldsymbol{C}_{1:N}$ は要素

$$C_{1:N}(\boldsymbol{x}_i, \boldsymbol{x}_j) = k(\boldsymbol{x}_i, \boldsymbol{x}_j) + \beta^{-1}\delta_{ij} \tag{13}$$

を持つ．

訓練集合として，位置 $\boldsymbol{x}_1, ..., \boldsymbol{x}_N$ と対応する $\hat{\boldsymbol{z}}_{1:N} = (\hat{z}_1, ..., \hat{z}_N)^T$ が与えられているときに，新しい位置 $\boldsymbol{x}_{N+1}$ に対する ILD z_{N+1} を予測したいものとする．そのために，予測分布 $p(z_{N+1}|\hat{\boldsymbol{z}}_{1:N})$ を求める必要がある．

条件付き分布 $p(z_{N+1}|\boldsymbol{z}_{1:N})$ を求めるためには，同時分布 $p(\boldsymbol{z}_{1:N+1})$ を書き下す必要がある．ここで，$\boldsymbol{z}_{1:N+1}$ はベクトル $(z_1, ..., z_N, z_{N+1})^T$ を表し，式 (12) から同時分布は

$$p(\boldsymbol{z}_{1:N+1}) = \mathcal{N}(\boldsymbol{z}_{1:N+1}|\boldsymbol{0}, \boldsymbol{C}_{1:N+1}) \tag{14}$$

で与えられる．ここで，$\boldsymbol{C}_{1:N+1}$ は，$(N+1) \times (N+1)$ の共分散行列であり，その要素は式 (13) で与えられる．この同時分布はガウス分布なので，条件付きガウス分布が得られる．これを行うために，次のように共分散行列の分割を行う．

$$\boldsymbol{C}_{1:N+1} = \begin{pmatrix} \boldsymbol{C}_{1:N} & \boldsymbol{k} \\ \boldsymbol{k}^T & c \end{pmatrix} \tag{15}$$

ここで，$\boldsymbol{C}_{1:N}$ は要素が式 (13)($n, m = 1, ..., N$ に対する) であるような $N \times N$ の共分散行列，$\boldsymbol{k}$は要素 $k(\boldsymbol{x}_n, \boldsymbol{x}_{N+1})(n = 1, ..., N)$ を持つベクトルであるとする．また，スカラー $c = k(\boldsymbol{x}_{N+1}, \boldsymbol{x}_{N+1}) + \beta^{-1}$ とする．これらを用いると，条件付き分布 $p(z_{N+1}|\hat{\boldsymbol{z}}_{1:N})$ は，次に示すような平均と共分散を持つようなガウス分布になることが知られている．

$$\begin{aligned} \mu(\boldsymbol{x}_{N+1}) &= \boldsymbol{k}^T\boldsymbol{C}_{1:N}^{-1}\hat{\boldsymbol{z}}_{1:N} \\ \sigma^2(\boldsymbol{x}_{N+1}) &= c - \boldsymbol{k}^T\boldsymbol{C}_{1:N}^{-1}\boldsymbol{k} \end{aligned} \tag{16}$$

式 (16) を用いることで，位置 $\boldsymbol{x}_{N+1}$ の ILD $\tilde{z}(\omega, \boldsymbol{x}_{N+1})$ とその分散 $\tilde{\sigma}^2(\omega, \boldsymbol{x}_{N+1})$ の補間が可能である．

5 不確かさを考慮した音源定位法

不確かさを考慮した音源定位では，予め各音源位置での ILD の平均，分散を取得しておく必要がある．

ある位置 $\boldsymbol{x}$ より得られる周波数 ω の ILD $\tilde{z}(\omega, \boldsymbol{x})$ と分散 $\tilde{\sigma}^2(\omega, \boldsymbol{x})$ が式 (16) より得られる．今，観測により ILD

$z(\omega)$ が得られたとすると，この時，音源が位置 $\boldsymbol{x}$ にある尤度 $l(\omega,\boldsymbol{x})$ を

$$l(\omega,\boldsymbol{x}) = \exp\left[-\frac{\{z(\omega)-\tilde{z}(\omega,\boldsymbol{x})\}^2}{\tilde{\sigma}^2(\omega,\boldsymbol{x})}\right] \tag{17}$$

とする．本研究では，1 つ以上の周波数点において音源が存在すると考えられれば高い尤度を与えるとの考えから，式 (17) の尤度の否定に相当する音源が位置 $\boldsymbol{x}$ にはない尤度 $\overline{l(\omega,\boldsymbol{x})}$ として

$$\overline{l(\omega,\boldsymbol{x})} = 1 - l(\omega,\boldsymbol{x}) \tag{18}$$

を考える．これを用いて，音源定位で使用する周波数帯域全体にわたって音源が位置 $\boldsymbol{x}$ にはない結合尤度 $\overline{l(\boldsymbol{x})}$ を考え，

$$\ln\ \overline{l(\boldsymbol{x})} = \frac{1}{M}\sum_{i=1}^{M}\ln\overline{l(\omega_i,\boldsymbol{x})} \tag{19}$$

とする．ここで，M は音源定位に使用する周波数点数である．これより，最終的に位置 $\boldsymbol{x}$ に音源がある尤度 $L(\boldsymbol{x})$ を

$$L(\boldsymbol{x}) = \eta\left[1-\exp\left\{\ln\overline{l(\boldsymbol{x})}\right\}\right] \tag{20}$$

と定めることとする．ここで η は正規化項である．

式 (20) より得られる各音源位置の尤度に閾値 ϵ を考え，$\epsilon \leq L(\boldsymbol{x})$ となった位置 $\boldsymbol{x}$ を音源位置と見做す．

6 検証

6.1 ILD の補間の検証

上述した方法を用いて，実際に ILD の補間を行った，検証用の音は KS 検定で使用した音データと同一のデータを使用した．訓練集合として Figure 4 の位置から 0.2m 四方の位置で取り出した Figure 6 に示す位置のデータを使用した．また本検証では，超パラメータである θ_0，θ_1 は適当な値を与え，β は，EM 法[ビショップ, 2009]を用いて推定を行った．

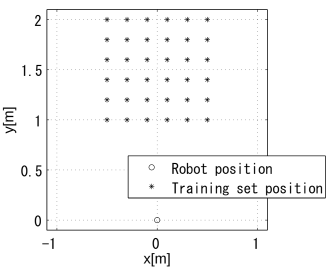


Figure 6: 訓練集合位置

ガウス回帰による補間 (GR 補間) の妥当性について検証を行うために，中村らの方法[中村, 2012]に基づいて ILD を線形補間によって求めたものとの比較を行う．線形補間は，補間したい ILD $\tilde{z}(\omega,\boldsymbol{x}_{N+1})$ が Figure 7 のような位置 $\boldsymbol{x}_{N+1}$ である場合，それを囲む 4 点の ILD を用いて

$$\begin{aligned}\tilde{z}(\omega,\boldsymbol{x}_{N+1}) = &(1-\zeta)(1-\xi)z(\omega,\boldsymbol{x}_1) + \zeta(1-\xi)z(\omega,\boldsymbol{x}_2)\\ &+(1-\zeta)\xi z(\omega,\boldsymbol{x}_3) + \zeta\xi z(\omega,\boldsymbol{x}_4) \\ &(0 \leq \xi,\zeta \leq 1)\end{aligned} \tag{21}$$

のように行う．

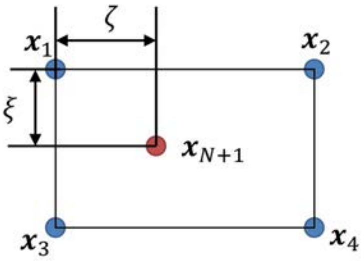


Figure 7: 線形補間概要

実際に補間を行った結果を Figure 8，9 に示す．Figure 8，9 はそれぞれ 500Hz，5000Hz の各位置での ILD を示している．

まず，500Hz での補間を見ると，GR 補間，線形補間ともに実際の ILD の分布と同様な分布が得られていることから，補間が可能であるということがわかる．5000Hz の補間結果を見ると，GR 補間，線形補間は同様な ILD の分布となっているが，実際の分布と比較すると，位置によって細かく変化している ILD の補間は必ずしも完全ではないが，おおよその傾向は再現できていることがわかる．

GR 補間と線形補間を比較するために，各位置で補間した ILD $\tilde{\boldsymbol{z}}(\boldsymbol{x}_i)$ と実際の ILD $\boldsymbol{z}(\boldsymbol{x}_i)$ を式 (22) のように内積を行った．

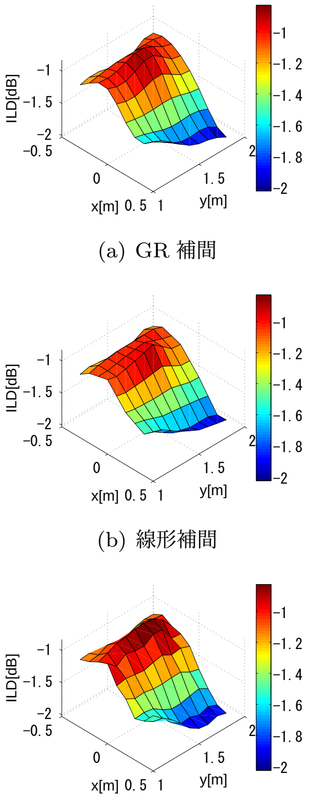
$$\mathrm{IP}_i = \frac{<\tilde{\boldsymbol{z}}(\boldsymbol{x}_i),\boldsymbol{z}(\boldsymbol{x}_i)>}{|\tilde{\boldsymbol{z}}(\boldsymbol{x}_i)||\boldsymbol{z}(\boldsymbol{x}_i)|} \tag{22}$$

その結果，各位置で得られる内積結果の平均が GR 補間で 0.9203，線形補間で 0.9301 となり，どちらの手法を用いても補間の性能にほとんど差がないことがわかった．

6.2 音源定位結果

補間より得られた ILD を規範データとし実際に音源定位を行った．音源定位には，GR 補間より得られた結果を用い，音源定位は不確かさを考慮した方法 (提案手法) と考慮していない方法で行った．音源定位の対象音としては白色雑音を使用した．また，音源定位には，500〜5000Hz の帯域を使用した．Figure 10 に代表的な定位結果を示す．$*$ が正解の位置を示しており，色で尤度を表す．

Figure 10(a)，10(b) また 10(c)，10(d)，10(e)，10(f) はそれぞれ提案手法の方がうまく定位が出来ている点，双方の手法で定位が出来ている点，双方の手法で定位が出

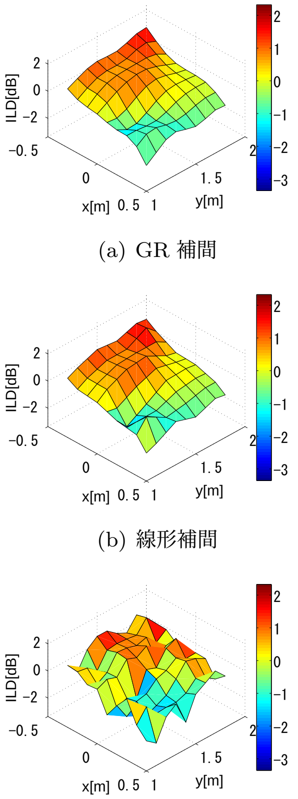


(a) GR 補間

(b) 線形補間

(c) 実際の ILD

Figure 8: ILD の補間結果 (500Hz)



(a) GR 補間

(b) 線形補間

(c) 実際の ILD

Figure 9: ILD の補間結果 (5000Hz)

来ていない点という組合せとなっている．まず，すべての結果を見ると不確かさを考慮した提案手法の方が全体の尤度が高くなっており，これは不確かさを考慮したことによる影響だと考えられる．Figure 10(a)，10(b) を見ると，不確かさを考慮していない手法に比べ，不確かさを考慮した手法の方が正解位置のピークが際立っている形となっている．Figure 10(c)，10(d) を見ると，どちらの手法を用いても正解位置の尤度が高くなっており，音源定位が出来ている結果となっている．Figure 10(e)，10(f) を見るとどちらの手法を用いても正解位置よりも x の位置が-側の尤度が高くなっており，正確な音源定位が出来ていない．補間より得られた ILD と実際の ILD の内積を求めると，Figure 10(a)，10(b) の位置での補間では 0.950，Figure 10(d)，10(e) の位置での補間では 0.925，Figure 10(e)，10(f) の位置での補間では 0.898 となっていることから，Figure 10(e)，10(f) の位置で音源定位が出来なかった原因として ILD の補間自体がうまく出来ていなかったためだと考えられる．

6.3 音源定位性能の評価

定位結果の評価方法として，ROC 曲線[James, 1989] における検出 / 誤り率に基づいた指標を用いる．

ROC 曲線では，尤度に適当な閾値 ϵ を設け，その下で二値判別を行った時の False positive の割合 (FP)，True positive の割合 (TP) を考えるので，音源位置 $\boldsymbol{x}$ での実験データに対して，曲線上の点 $(FP, TP)^T$ は

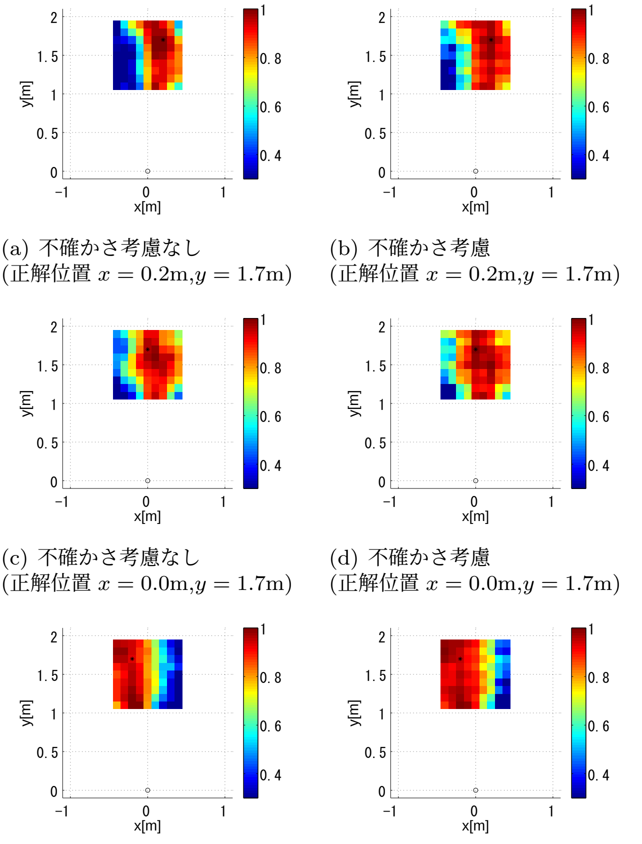
$$(FP, TP)^T = \mathrm{ROC}(\epsilon, \boldsymbol{x}) \tag{23}$$

と表される．$(FP, TP)^T = (0, 1)^T$ が理想的な定位を実現している状態に対応していることから ROC 曲線の値が $(FP, TP)^T = (0, 1)^T$ に近いほど定位性能が良いと考える．

閾値 ϵ を 0.3 から 0.9 まで変化させ ROC 曲線を求めた結果を Figure 11 に示す．Figure 11 の ROC 曲線は，全音源位置で求めた True positive の割合と False positive の割合の平均を使用して求めている．

Figure 11 を見ると，不確かさを考慮した音源定位の方が ROC 曲線の値が $(FP, TP)^T = (0, 1)^T$ に近く，また $\epsilon = 0.9$ の点を比較すると提案手法の方が良好な性能を示していることがわかる．

次に，True positive の割合，False positive の割合が最も悪い場合の結果を用いて ROC 曲線を求めた結果を Figure 12 に示す．この結果を見ると，不確かさを考慮した音源定位の方が考慮していないものに比べ $(FP, TP)^T = (0, 1)^T$ に近いことから，不確かさを考慮することで定位性能の低下が抑えられることがわかる．このことから，不確かさを考慮することで定位性能の改善が行えることがわかった．



(a) 不確かさ考慮なし (正解位置 $x = 0.2$m,$y = 1.7$m)

(b) 不確かさ考慮 (正解位置 $x = 0.2$m,$y = 1.7$m)

(c) 不確かさ考慮なし (正解位置 $x = 0.0$m,$y = 1.7$m)

(d) 不確かさ考慮 (正解位置 $x = 0.0$m,$y = 1.7$m)

(e) 不確かさ考慮なし (正解位置 $x = -0.2$m,$y = 1.7$m)

(f) 不確かさ考慮 (正解位置 $x = -0.2$m,$y = 1.7$m)

Figure 10: 音源定位結果

7 まとめ

音響特徴量の補間の手法として，ガウス回帰に基づく手法を使用し，補間点の不確かさを考慮した．規範データとして，補間より得られた音響特徴量を使用し，対象音を白色雑音とし音源定位を行った場合，音源定位性能が向上することを確認することができた．

音楽，音声での音源定位は今後の課題である．

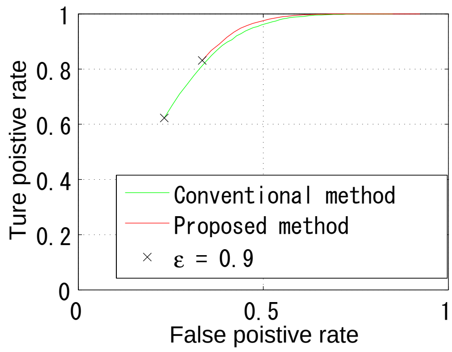


Figure 11: ROC 曲線

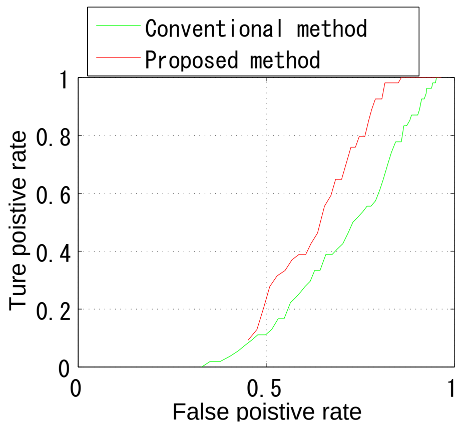


Figure 12: ROC 曲線 (最悪値)

参考文献

[奥乃, 2010] 奥乃　博: ロボット聴覚の現状と展望, 日本ロボット学会誌, vol.28, no.1, pp.2-5, 2010.

[奥乃, 2002] 奥乃　博，中臺　一博: ロボットの耳は二つで十分か (< 小特集 > なぜ耳は二つあるか?), 日本音響学会誌, vol.58, no.3, pp.205-210, 2002.

[章, 2008] 章忠，井和章，三宅哲夫，今村孝，堀畑聡：バイノーラルモデルを用いた音源方向推定， 日本機械学会論文集 C 編， Vol.74-739, pp. 642-649, 2008.

[Shaw, 1968] Shaw E.A.G., Teranishi R.: Sound pressure generated in an external-ear replica and real human ears by a nearby point source, J.Acoust.Soc.Am., vol.44, pp.240-249, 1968.

[野田, 2012] 野田　佳孝，公文　誠: 2 つの能動耳介による正中面内の音源方向推定, 第 13 回システムインテグレーション部門講演会 (SI2012), pp.1643-1646, 2012.

[木元, 2013] 木元　大輔，尾堂　航，公文　誠: 観測データの不確かさを考慮したバイノーラル聴覚ロボットでの音源定位手法, 第 31 回日本ロボット学会学術講演会, RSJ2013AC3D3-05, 2013.

[中村, 2012] 中村　圭介，中臺　一博：ハイブリッド伝達関数補間法とそのロボット聴覚システムへの応用, 日本ロボット学会第 30 回記念学術講演会， 3D1-5, 2012.

[ビショップ, 2009] C.M. ビショップ: パターン認識と機械学習 (上, 下) ベイズ理論による統計的予測， シュプリンガー・ジャパン， 2009

[James, 1989] James A.Hanley: Receiver operating characteristic(ROC) methodology:The state of the art, Crit.Rev.Diagn.Imaging, vol.29, Issue3, pp.307-335, 1989.