

# 位相差スペクトルイメージによる音源までの距離推定 —音源が正面に位置しない場合— Distance Estimation to a Sound Source Off-Centered Using Phase Difference Spectrum Images

霜山 竜一<sup>†</sup>  
R. Shimoyama

## 1. まえがき

視覚障害者でも、聴覚情報だけで音源までのおよその距離を推定できることが知られている。Bronkhorst ら[1]は距離推定における直接音と反射音のエネルギー比の役割について報告している。Nakadai ら[2]と Berglund ら[3]は“Active audition”という手法を提案している。複数のマイクロホン装着した、ロボットの頭部を回転させて受音点を移動させ、回転前後の音響信号を用いて音源の位置を推定する。この手法では回転駆動に伴う内部雑音が発生し、回転中の音源の移動は無視される。本研究は、人の聴覚にみられる、瞬時に距離を推定する機能の模擬を目的とし、静止した1対のマイクロホンで音源までの距離を推定する。筆者は先に、室内で1対の無指向性マイクロホンの真正面に設置した音源から連続的に広帯域雑音を発生させ、マイクロホンと音源間の距離と音圧の位相差スペクトル分布の関係を調べた[4]。位相差スペクトルのばらつきを標準偏差で表すと、標準偏差の平均値はマイクロホンと音源間の距離にほぼ比例することを報告した。一方、音源がマイクロホンの正面にない場合は、マイクロホンと音源のなす角度に応じて位相差スペクトルのばらつきが著しく変化する。本報告では、マイクロホンと音源のなす角度に応じて標準偏差の上昇分を補正して音源までの距離を推定した結果について報告する。

## 2. 推定アルゴリズム

室内の中央で1対の無指向性マイクロホンから、距離  $L$  だけ離して1台の小型スピーカを置く。スピーカから広帯域雑音を連続的に発生させると、測定される音圧の位相差の周波数特性には連続性がほとんどみられず、位相差値は不規則にばらつく。Fig.1 で角度  $\theta$  をゼロとすれば  $Y$  軸上に音源がある限り、スピーカからマイクロホンに直接到来する音の位相差は周波数に拘らず理論的にはゼロとなる。そうならない理由は、マイクロホンが直接音と反射音の合成された音を検出するため、位相差値の不規則な変動は壁や床面による反射音に起因する。位相差のばらつきを標準偏差値で定量的に評価する。周波数帯域幅  $\Delta f$  に含まれる  $2m+1$  個の位相差値の標準偏差をその中心周波数  $f_i$  を変えながら式(1), (2)から求める。周波数  $f_i$  における標準偏差  $\rho_i$  は、

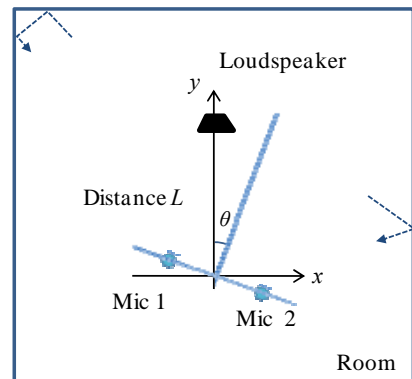


Fig. 1 Configuration of sensors and a source (a plane figure).

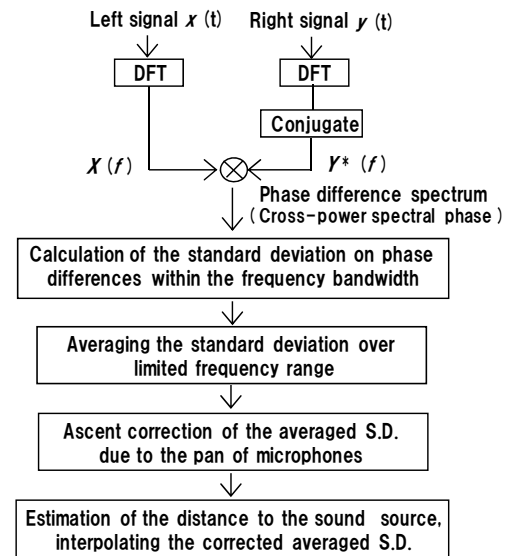


Fig. 2 Proposed algorithm.

$$\rho_i = \sqrt{\frac{1}{2m+1} \sum_{j=i-m}^{i+m} (\varphi_j - \bar{\varphi}_i)^2} \quad (1)$$

$$\bar{\varphi}_i = \frac{\sum_{j=i-m}^{i+m} \varphi_j}{2m+1} \quad (2)$$

<sup>†</sup> 日本大学生産工学部電気電子工学科  
〒275-8575 習志野市泉町 1-2-1

で与えられる。ここで  $\varphi_j$  は周波数  $f_j$  における位相差、 $\bar{\varphi}_i$  は帯域幅  $\Delta f$  内の位相差の平均値である。

距離を推定するためのアルゴリズムを Fig.2 に示す。2 系統の音響信号を同時に計測し、それぞれ DFT 処理後にクロススペクトルの位相 (位相差スペクトル) を計算する。式(1),(2)から標準偏差の周波数特性を計算し、特定の周波数帯域で求めたその平均値を、音源までの距離に依存する特徴量とみなす。Fig.1 の角度  $\theta$  がゼロの場合は、距離と標準偏差値の平均値の関係を実測で求め、位相差スペクトルの標準偏差値を線形補間すれば距離が推定できる。一方、スピーカに対するマイクロホンの角度  $\theta$  を変えると、角度が大きいくほど標準偏差値は増加する傾向があり、角度の影響を分離する必要が生じる。そこで、マイクロホンの角度に応じて標準偏差の平均値の上昇分を予め実測して求め、この上昇分を補正して距離を推定する。

### 3. 結果

24 ビット分解能の A/D 変換ボード(NI,PCI-4474)を装備した PC (Dell,GX280) で 2 系統の音圧波形を同時計測した。マイクロホン(Type 4190 B&K)の間隔は 20(cm)である。  $\Delta f$  は 0.8kHz ( $m = 160$ )である。サンプリング周波数 24kHz, 1 フレームは 9600 データでハンギング窓を用いた。標準偏差の平均値は、距離による違いの顕著な周波数範囲(4kHz~11kHz)で求めた。スピーカをマイクロホンの正面に設置して ( $\theta = 0$ )、マイクロホンとスピーカの距離と位相差の標準偏差の平均値の関係を測定した。最少二乗法による近似直線は  $\delta = +17.7L + 4.1$  で表わされ、距離の増加に伴う標準偏差値の上昇率は約  $+17.7$  ( $^\circ$ /距離)となった。次にスピーカまでの距離  $L$  をそれぞれ 1(m), 2(m), 3(m) に設定し、マイクロホンの角度  $\theta$  を  $-60^\circ$  から  $+60^\circ$  まで順次変えて測定した位相差の標準偏差の平均値を Fig.3 に示す。角度の正負に拘わらずマイクロホンの角度が増すほど標準偏差値の上昇する様子が分かる。標準偏差値の上昇率の平均はほぼ  $+0.8$  ( $^\circ$ /角度)となった。マイクロホンとスピーカ間の実際の距離と推定された距離の関係を Fig.4 に示す。黒丸は補正前の、白丸はマイクロホンの角度に応じて標準偏差の平均値の上昇分を補正した場合の推定距離を示す。補正しない場合はマイクロホンの角度が約  $60^\circ$  で距離は 2 (m) 程度長めに推定された。補正するとほぼ 20~30(cm)の誤差で推定できることが分かる。

### 4. あとがき

室内でスピーカから連続的に発生させた広帯域雑音を 1 対の無指向性マイクロホンで測定し、マイクロホンの角度を変えて音圧の位相差の標準偏差の平均値からスピーカまでの距離を推定した。マイクロホンと音源のなす角度に応じて標準偏差の上昇分を補正して音源までの距離を推定した結果、(1) マイクロホンのなす角度による標準偏差値の上昇率は平均するとほぼ  $+0.8$  ( $^\circ$ /角度) である、(2) 角度による標準偏差の上昇分を補正した場合にはほぼ 20~30(cm)の誤差で推定できる、ことを示した。マイクロホンに近く、あるいは遠ざかるという音源の移動を検出できそうである。本手法では反響音を利用して音源までの距離を推定しており、推定距離は室内の反響の状況に依存する。

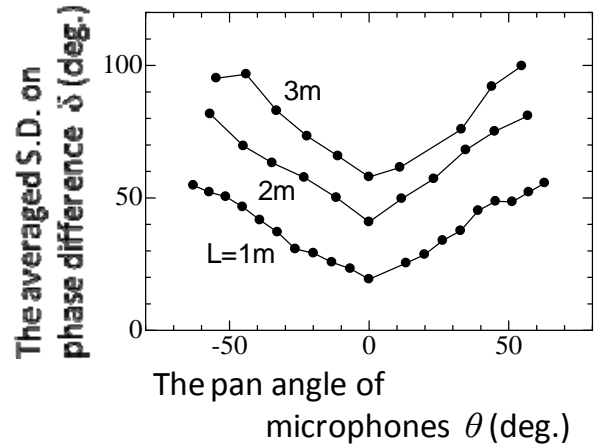


Fig. 3 Averaged standard deviation on phase difference corresponding to the pan angle of microphones.

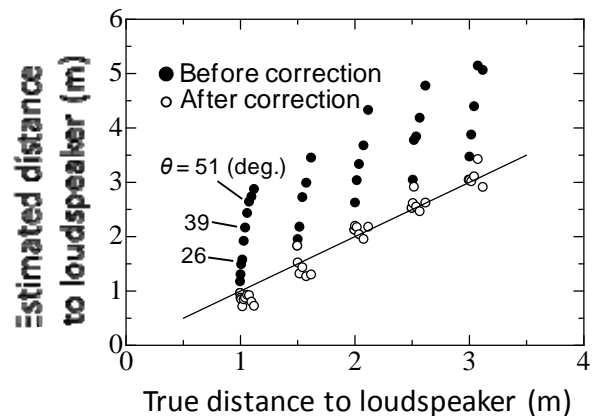


Fig. 4 Estimated distances to loudspeaker before and after the correction

今後は、音源の方向定位と距離推定を同時に行い、マイクロホン周囲の音源の空間的な分布を可視化する予定である。上述した結果から予想されるマイクロホンの単位角度あたりの推定距離の誤差は約  $0.045$  (m/角度) である。例えば約  $5$  ( $^\circ$ ) の精度で音源の方向が検出できる場合には  $30$  (cm) 以下の誤差で距離が推定されることになる。

### 参考文献

- [1] A. W. Bronkhorst and T. Houtgast, "Auditory distance perception in rooms", *Nature*, 397 (6719), 1999, p.p. 517-520.
- [2] K. Nakadai, T. Lourens, H. Okuno, H. Kitano, "Active audition for humanoid", *AAAI-2000*, 2000, p.p. 832-839.
- [3] E. Berglund, J. Sitte and G. Wyeth, "Active audition using the parameter-less self-organising map" *Auton Robot*, 24, 2008, p.p.401-417.
- [4] 霜山, 山崎, "位相差スペクトルイメージによるマイクロホンと音源間の距離推定", *FIT2008*, H-059, 2008, p.p.191-192.