

# 音源方向検出法の室内残響下での性能評価について\*

○上杉信敏, 金田豊 (東京電機大・工)

## 1 はじめに

信号到来方向(DOA)の検出はこれまで多くの手法が提案されている。しかしこれらの音源方向探査技術を室内で利用する場合には、室内反射音や周囲雑音を無視する事は出来ない。本研究では最も基本的な 2ch マイクロホン信号を利用する手法を対象とし、雑音や反射音の影響などを想定し性能評価を行い、各手法の特徴を明らかにする事を目的とする。

## 2 DOA 検出法

### 2.1 一般化相互相関関数

DOA を検出するための標準的な手法は、2ch マイクロホン出力信号  $x_1(t)$  と  $x_2(t)$  間の遅延時間を次式で定義される相互相関関数(CC: Cross Correlation)  $\phi_{12}(\tau)$  の最大値から推定する方法である。

$$\phi_{12}(\tau) = E[x_1(t)x_2(t+\tau)] \quad (1)$$

ここで式(1)を  $x_1(t)$  と  $x_2(t)$  のクロススペクトル  $G_{12}(\omega)$  を用いると、次式で表される。

$$\phi_{12}(\tau) = \int G_{12}(\omega) e^{j\omega\tau} d\omega \quad (2)$$

一方、雑音や反射音などの環境下での性能を確保するため  $G_{12}(\omega)$  に様々な周波数重み  $\psi(\omega)$  を付ける事が提案されてきた[1-2]。これを一般的に表したものが一般化相互相関関数  $R_{GCC}(\tau)$  で、次式で表される。

$$R_{GCC}(\tau) = \int \psi(\omega) G_{12}(\omega) e^{j\omega\tau} d\omega \quad (3)$$

### 2.2 平滑化コヒーレンス変換

平滑化コヒーレンス変換(SCOT: Smoothed Coherence Transform)はコヒーレンス関数  $\gamma_{12}(\omega)$  を用いる手法である。SCOT において周波数重み  $\psi_{SCOT}(\omega)$  は次式で表される。

$$\psi_{SCOT}(\omega) = \frac{1}{\sqrt{G_{11}(\omega)G_{22}(\omega)}} \quad (4)$$

式(3)に代入して得られる被積分項  $\gamma_{12}(\omega)$

$$\gamma_{12}(\omega) = \frac{G_{12}(\omega)}{\sqrt{G_{11}(\omega)G_{22}(\omega)}} \quad (5)$$

は複素コヒーレンス関数を表す。

### 2.3 位相変換

位相変換(PHAT: Phase Transform)は白色化相互相関、CSP などとも呼ばれる手法で周波数重み  $\psi_{PHAT}(\omega)$  は次式で表される。

$$\psi_{PHAT}(\omega) = \frac{1}{|G_{12}(\omega)|} \quad (6)$$

式(3)に代入で得られる被積分項は、パワースペクトルの特性を打ち消し位相項のみを表す。

### 2.4 最尤推定法に基づく手法

雑音が独立な白色雑音であると仮定して最尤推定法(ML: Maximum Likelihood)に基づく手法である。この周波数重み  $\psi_{ML}(\omega)$  は次式で表される。

$$\psi_{ML}(\omega) = \frac{1}{|G_{12}(\omega)|} \cdot \frac{|\gamma_{12}(\omega)|^2}{1 - |\gamma_{12}(\omega)|^2} \quad (7)$$

## 3 雑音耐性の検討

前節 2.1~2.4 で示した 4 つの手法に対する雑音の影響を評価するために計算機シミュレーションを行なった。条件はマイクロホン間距離 0.6 m, サンプリング周波数 12kHz, 相関長 250 ms とし、目的音は男性音声を到来方向が  $-90^\circ \sim 90^\circ$  範囲で 40 方向に変化させ、方向検出結果の正解率を評価した。

雑音は多くの研究で行われているようにマイク信号に無相関なガウス性白色雑音を加え、SN 比は  $-15 \sim 15$  dB で変化させた。許容誤差を  $\pm 2^\circ$  とし、各 SN 比における各手法の方向検出正解率を図 1 に示す。図より SN 比によらず ML の正解率が最も高く、PHAT の性能が悪いことが分かる。ML は無相関白色雑音に対して最適な手法として導出された手法であること、PHAT は SN 比の低い高周波を持ち上げてしまうので誤差が大きくなることが原因として考えられる。この結果は過去の研究[3-4]と同様の結果である。

しかし無相関な白色雑音は実際の騒音環境を反映しているとは言えない。そこで標準的な室内騒音スペクトルを持った Hoth 雑音が

\* Performance of sound source direction estimation methods under reverberant conditions, by Nobutoshi Uesugi and Yutaka Kaneda (Tokyo Denki Univ.).

周囲一様から到来している状態（拡散音）をシミュレートして同様な実験を行った。その結果は図2に示すように、PHATの性能は良好だがMLが劣化するという無相関白色雑音の結果（図1）とは逆の結果が得られた。これは、Hoth雑音と音声のスペクトルが類似して全帯域でほぼSN比が等しいので、白色化は悪影響を受けないこと、ML導出の理論条件が満たされていないことなどがその理由として考えられる。

一方、図3は許容誤差を $\pm 10^\circ$ とした場合の結果を示している。この場合にはSCOTの性能が向上している。SCOTの相関関数はPHATと比べると鋭いピークをとらないので、許容誤差が小さいと正解率は低いが、許容誤差が大きい場合には、雑音に対してロバストに推定ができることを示している。

#### 4 反射音環境での検討

残響時間が約400msの部屋の中央で同様の実験を行った。目的音は男性が $-60^\circ \sim 60^\circ$ 範囲で8方向、距離は1.5mとして発声した。雑音は拡散性Hoth雑音を計算機上で加算し、マイクロホンの条件等は前項と同様である。

許容誤差 $\pm 10^\circ$ とした結果を図4に示す。図よりSN比によらずPHATの正解率が最も高いことがわかるが、一般性の高い評価を得るためには、より多くの条件での検討が必要と考える。

#### 5 むすび

雑音のある場合のDOA検出法として、現実の音場を反映した拡散性Hoth雑音に対して、許容誤差が小さい場合にはPHATが優れており、許容誤差が $10^\circ$ 程度と大きい場合にはSCOTの性能が良いことが分かった。

また反射音に対しては今後多くの条件で検討を進める必要がある。

#### 参考文献

- [1] 金井浩, "音・振動のスペクトル解析", コロナ社, (1999).
- [2] C.H.Knapp and G.C.Carter, "The Generalized Correlation Method for Estimation of Time Delay", IEEE Trans.On Acoust., Speech and Signal Proc., ASSP-24, 4, pp.320-327, (1976).
- [3] M.S.Brandstein, "Time-delay estimation of reverberated speech exploiting harmonic structure", J.Acoust.Soc.Am., 105, 5, pp.2914-2929(1999).
- [4] V.C.Rayar et al, "Speaker Localization Using Excitation Source Information in Speech", IEEE Trans.On Speech and Audio Proc., 13, 5, pt.2, pp.751-761, (2005).

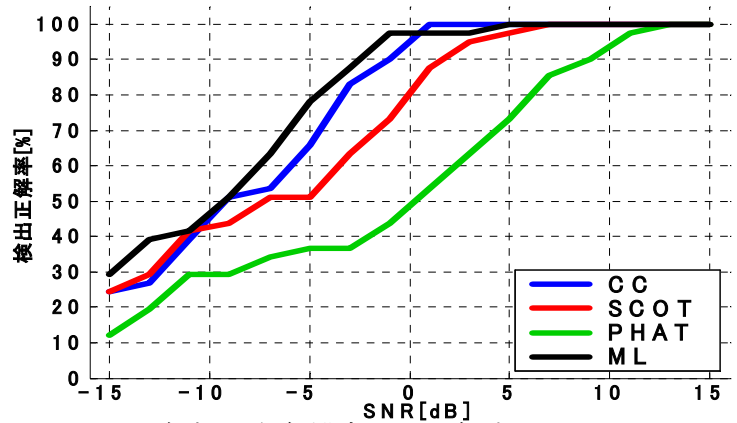


Fig. 1 無相関白色雑音での正解率( $\pm 2^\circ$ )

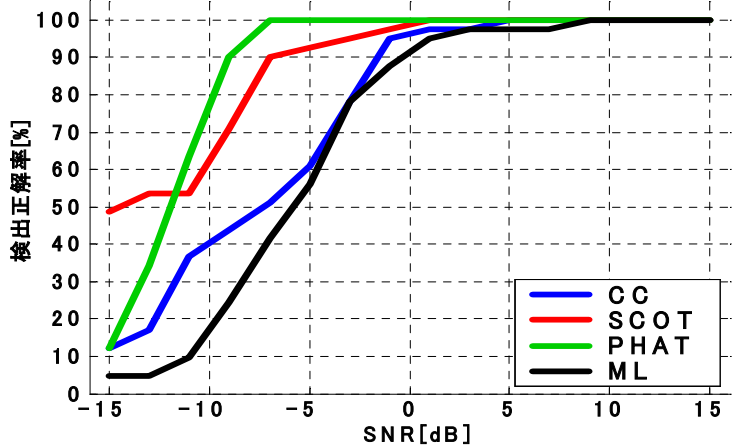


Fig. 2 拡散性 Hoth 雑音での正解率( $\pm 2^\circ$ )

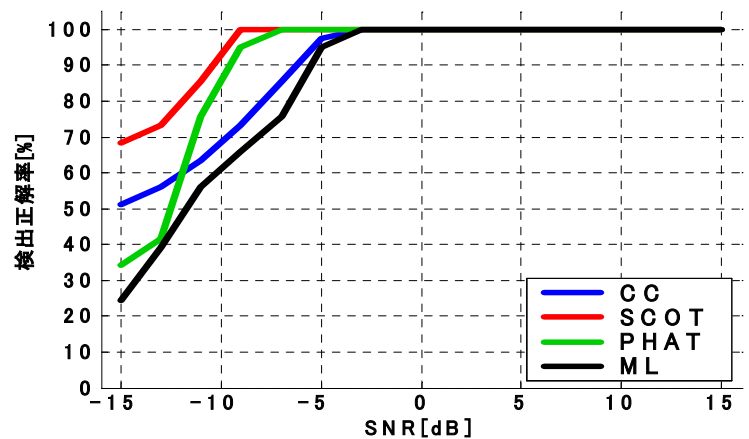


Fig. 3 拡散性 Hoth 雑音での正解率( $\pm 10^\circ$ )

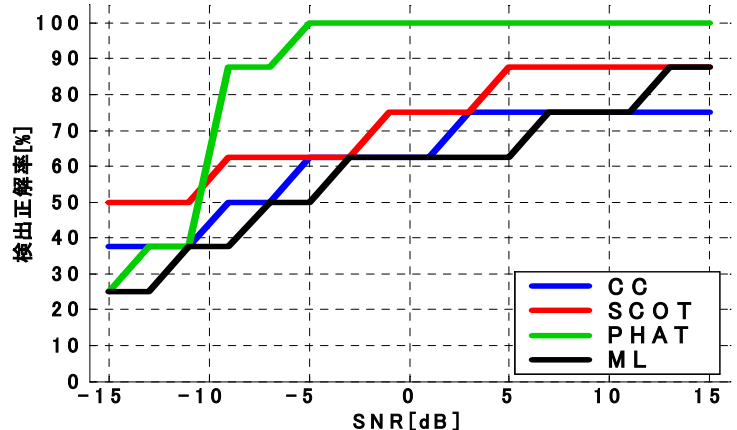


Fig. 4 残響下での正解率( $\pm 10^\circ$ )