

Title	雑音抑圧処理を付加した音源方向推定に関する研究
Author(s)	小林, 利秋
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1472
Rights	
Description	Supervisor: 赤木 正人, 情報科学研究科, 修士

雑音抑圧処理を付加した音源方向推定に関する研究

小林 利秋

北陸先端科学技術大学院大学 情報科学研究科

2001 年 2 月 15 日

キーワード: 音源方向推定、マイクロホンアレイ、雑音抑圧.

1 はじめに

複数のマイクロホンを用いた音源方向推定は、TV 会議や遠隔会議、ロボットのマンマシンインタフェース、音声認識の前処理等の幅広い分野で求められているが、雑音、反射音の存在する環境下では推定精度が低下するという問題点がある [1]。

本研究では、音声を目的音とする音源方向推定を扱う。そして雑音に対応するために、音源方向推定システムに目的音声中有声音の基本周期を手掛かりとした時間 - 周波数フィルタによる雑音抑圧処理を組み合わせることを提案する。

さらにこの雑音抑圧波形に対して、反射音に対する不応期制御を含んだ西田らの提案法 [2] を適用することで反射音に対応することにより、システム全体として、雑音と反射音に対し頑健な音源方向推定のアルゴリズム構築を目指す。

2 マイクロホンアレイ

本提案法では、最少のマイクロホン数で 2 次元平面全方位に対する方向推定が可能なアレイとして、西田らの推定法と同様、正三角形の頂点に 1 つずつマイクロホンを配置したマイクロホンアレイを使用する。このようなアレイを用いることにより、従来多く用いられていた直線配置のマイクロホンアレイで生じていた前後の判断の誤りなどを解消することができる。

3 音源方向推定のアルゴリズム

提案法の音源方向推定は、図 1 に示すフローチャートに沿って行われる。

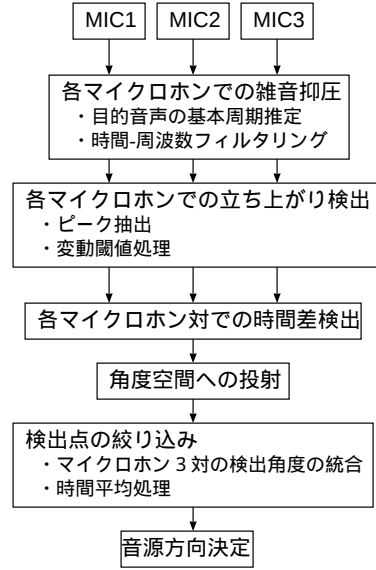


図 1: 音源方向推定フローチャート

この図の雑音抑圧で雑音に対応し、続く立ち上がり検出の反射音に対する不応期処理により、反射音に対応する。この立ち上がり検出処理以降は、西田らの提案法の流用となる。

4 雑音抑圧処理

雑音に対応するため、提案法では先ず各マイクロホン入力に対し雑音抑圧を行う。後段で時間差に基づく方向推定を行えるように、ここでは目的音声のもつ時間情報を保存しつつ雑音抑圧を実現する手段として、目的音声の基本周期を用いた時間-周波数フィルタ [3] を使用する。

4.1 時間-周波数フィルタ

目的音 $s(t)$ の音声は、基本周期 T (基本周波数 ω_0) の調波複合音で代表されると仮定すると、マイクロホンに入力する混合音は、次のように表せる。

$$x(t) = s(t) + n(t) \quad (1)$$

$$= \sum_n a_n e^{j(n\omega_0 t + \theta_n)} + \sum_k b_k e^{j(n\omega_k t + \theta_k)} \quad (2)$$

キャンセレーション法では、まず混合音 $x(t)$ より目的音 $s(t)$ をキャンセルし、雑音成分 $n(t)$ を推定する。そのためにまず下記の如く $g(t)$ を定義する。

$$g(t) = \frac{2x(t) - x(t - T) - x(t + T)}{4} \quad (3)$$

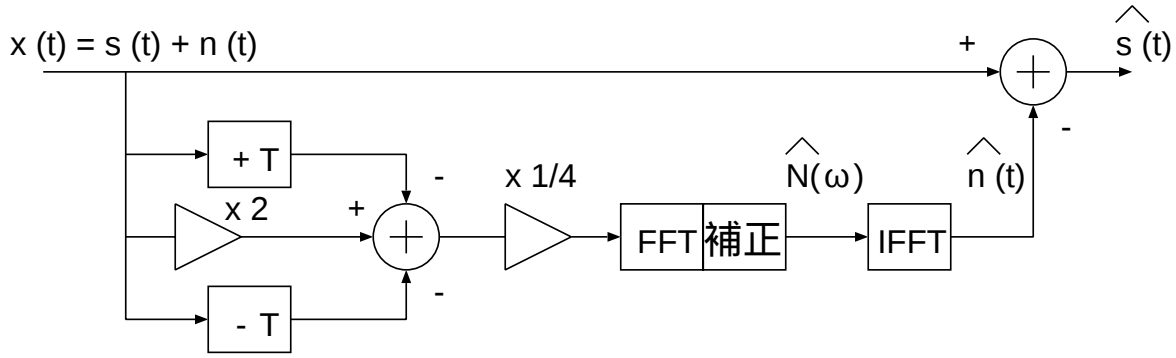


図 2: 時間 - 周波数フィルタ ブロック図

$$= \sum_k b_k e^{j(n\omega_k t + \theta_k)} \sin^2 \frac{\omega_k}{\omega_0} \pi \quad (4)$$

上式には目的音の成分が、振幅、位相項ともに全く含まれていないことが留意せよ。そして次に上式をフーリエ変換する。 $n(t)$ のフーリエ変換を $N(\omega_k)$ 、 $g(t)$ のフーリエ変換を $G(\omega_k)$ とすれば、

$$G(\omega_k) = N(\omega_k) \sin^2 \frac{\omega_k}{\omega_0} \pi \quad (5)$$

よって雑音スペクトルは

$$N(\omega_k) = G(\omega_k) / \sin^2 \frac{\omega_k}{\omega_0} \pi \quad (6)$$

となる。

上式では $\omega_k / \omega_0 = \text{整数}$ の時、分母がゼロとなり、計算処理に不都合が生じる。よって上式を、微小量 ε を設定し、以下のように場合分けする。

$$N(\omega_k) = \begin{cases} G(\omega_k) / \sin^2 \frac{\omega_k}{\omega_0} \pi & \text{for } |\sin \frac{\omega_k}{\omega_0} \pi| \geq \varepsilon \\ 0 & \text{for } |\sin \frac{\omega_k}{\omega_0} \pi| < \varepsilon \end{cases} \quad (7)$$

この $N(\omega_k)$ を逆フーリエ変換した雑音 $n(t)$ を、もとの混合音から引き去ることにより目的音 $s(t)$ を推定としての雑音抑圧信号を得る。参考として、以上の処理を表したブロック図を図 4.1 に示す。以上の処理により、目的音の調波構造に合わせた楕円フィルタ特性を実現している。

4.2 立ち上がり検出

立ち上がり検出では、反射音に対応するため、人間の音源定位に有効な先行音効果を工学的にモデル化し、以下の通り実現される。

先ず雑音抑圧波形を全波整流した後、ピーク抽出を行う（図 3(a)）。なだらかに変化する波形のピークのみを残すことで、時間差検出に不必要な情報を取り除く。次にこのピーク抽出波形に対して、波形に適応する形で変動閾値を設定し、振幅が急激に変化するポイントの検出（立ち上がり検出）を行う。この立ち上がり検出の結果を、一例として図 3(b) に示す。

またこの変動閾値の値は、次の立ち上がり検出点に至るまで、時間とともに指数減衰させている。一般に反射音は直接音より遅れて到来し、その振幅は直接音より小さいことを考慮すれば、この変動閾値処理により、反射音に対する不応期を実現していることが分かる。

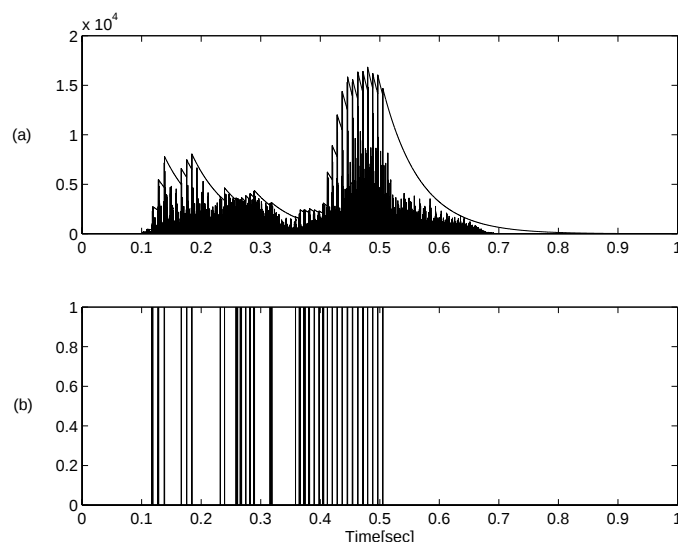


図 3: dynamic threshold & onset detect

5 実験結果

本研究で構築した音源方向推定法の有効性を確認するために、様々な条件を設定し、計算機シミュレーションを行なった。その中で、西田らの推定法、白色化相互相関 (Crosspower-Spectrum Phase : CPSP)[1] との比較実験を行なった結果、他の手法に比べ、方向性の雑音に対してロバストであることが認められた。その結果を図 4 に、そのときのシミュレーションパラメータを表 1 に示す。

ここで誤り率とは、音源方向に対し ± 5 度以内に入らなかった異常検出点の数を音源方向推定によって得られた総検出点数で割った値を百分率で示したものである。図 4 を見ると、明らかに提案法の優位性がわかる。この結果は提案法の雑音抑圧処理が有効であることを示している。

表 1: simulation parameters

target signal	word speech of ATR database, 18 samples arriving angle : setting at random
noise	white noise arriving angle : 60° SNR: 10dB ~ 30dB
1st echo	echo-to-signal ratio : -4dB arrival delay: 2ms
2nd echo	echo-to-signal ratio : -8dB arrival delay: 10 ~ 15ms

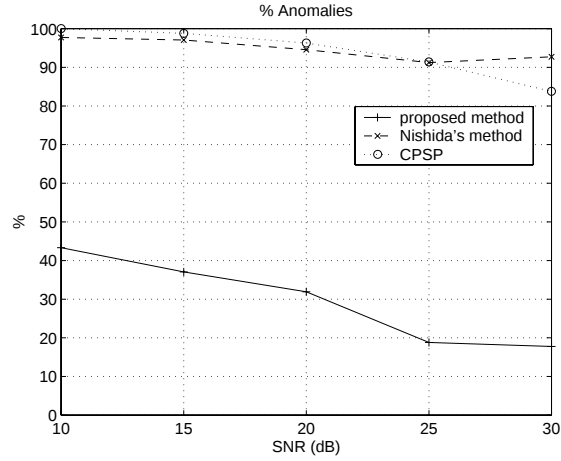


图 4: simulated result in the presence of arriving noise & echoes

次に室内音響シミュレーション [4] により、残響に対する評価を行った結果について図 6 に示す。本評価では、図 5 に示す室内空間を仮定している。

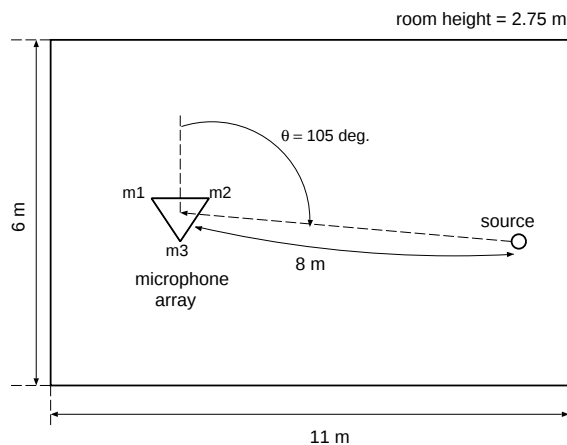


図 5: simulated room

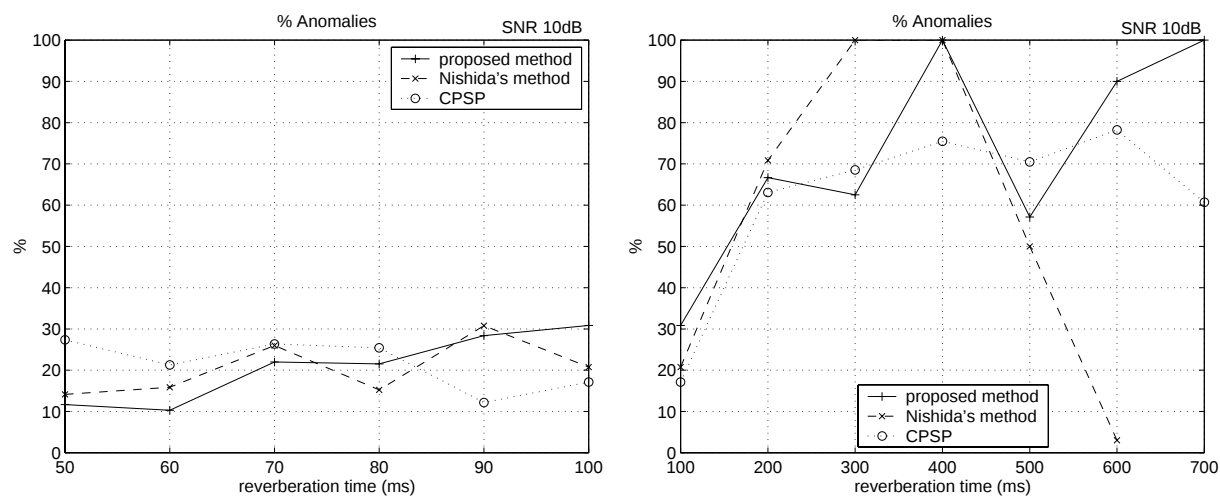


図 6: experimental results for simulating room acoustics : SNR 10dB

この結果をみると、残響時間 70ms 以下では、他の手法に比べ提案法が優位だが、残響時間 200ms を越えると、従来法と同様、提案法の性能は大きく劣化してしまう。このことは、提案法、西田らの手法の反射者に対する不応期制御が、反射音と直接音のレベル差に期待した処理であるため、反射音が集合した残響のレベルが直接音のレベルを上回る環境では正常に機能せず、方向推定が行えなかったためと考えられる。

6 結論

雑音抑圧処理を付加した音源方向推定を提案し、特に提案法が方向性の雑音に有効であることをシミュレーションにより確認した。しかし一方、室内音響シミュレーションにより行った残響に対する評価では、残響時間が 200ms を越えると性能が大幅に劣化することが確認された。これに対しては、現行の反射音に対する不応期制御とは別に、残響に対応する処理を、今後検討していく必要がある。

参考文献

- [1] Maurizio Omologo, "Acoustic Event Localization Using A Crosspower-Spectrum Phase Based Technique", Proceedings of ICASSP 94, vol.2, pp. 273-276, Adelaide, 1994.
- [2] T.Nishida, "Research on the estimation of sound source direction using multi microphone," JAIST master thesis, 1999.
- [3] M.Akagi et.al., "Speech Enhancement and Segregation based on Human Auditory Mechanisms," Proc, Int. Conf. Information Society in the 21st Century(IS2000), pp.102-126, Aizu-Wakamatsu, November 2000.
- [4] Jont B. Allen, David A. Berkley, "Image Method for Efficiently Simulating Small-Room Acoustics," J.Acoust.Soc.Am. 65(4), pp. 954-950, April 1979.