

位相情報を用いたサブバンドピークホールド処理による音源方向推定の検討*

◎佐藤耕平, 金田豊 (東京電機大・工)

1 はじめに

室内音源方向推定では雑音や反射音が悪影響し推定精度が低下する問題がある. これまで我々はサブバンドピークホールド(SBPH)処理により反射音耐性の向上を実現してきたが[1], 室内雑音に対する対策は未検討であった. 本稿では従来の SBPH 処理と位相情報を組み合わせることで, 雑音耐性の向上を確認したので報告する.

2 音源方向推定とピークホールド処理

本稿では2マイクによる音源方向推定を考える. 図1(a)は2つのマイクの受信信号(ここではパルス音を仮定) $x_1(t)$ と $x_2(t)$ を表している. 受信信号には直接音に加えて反射音が含まれている. この時, 直接音の時間差 τ_s が音の到来方向を表すが, 反射音の影響により相互相関関数 $\phi_{12}(\tau)$ が τ_s 以外にピークを持ち(図1(b))推定誤差の原因となる.

そこで, 先行する振幅値を保持する処理を行い後続の反射音をマスクする(図1(c)). この信号の時間差分をとることで信号の立ち上がり部分が抽出でき(図1(d)), 相互相関関数を求めると直接音の時間差 τ_s が明確になる.

3 サブバンドピークホールド処理(SBPH)

音声のように帯域毎に立ち上がり時刻が異なる信号に対応するため, 受信信号を帯域分割し, 帯域毎にピークホールド処理を行うことが有効である. 処理の流れを図2に従って説明する. まず①短時間フーリエ変換(STFT)を行い, ②時変振幅スペクトル $|X_i(\omega, t)|$ を出力し, ③ピークホールド処理(PH)を行う. さらに後続反射音を抑圧するため④対数操作(log)を行い, ⑤時間差分(Diff)を取ることで帯域毎に音声信号の立ち上がりを取り出す. これらの信号の⑥相互相関関数(Cor)を計算して, ⑦全ての帯域での相互相関関数を平均化し時間差 τ_s の推定を行う.

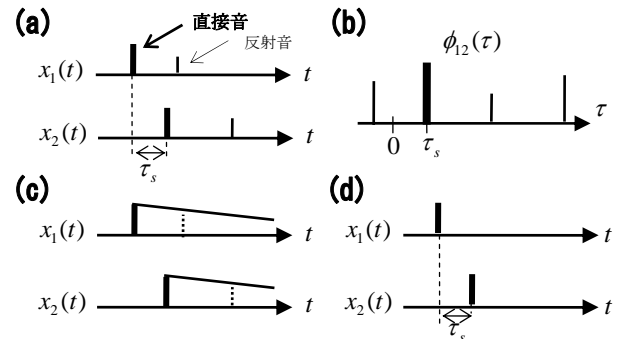


図1 ピークホールド処理のモデル図

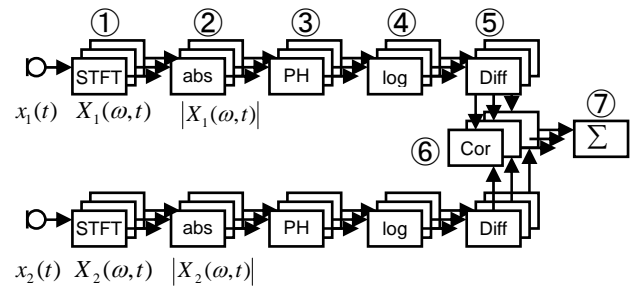


図2 SBPH処理のブロック図

4 提案法

4.1 従来法の問題点

従来法において, 雑音の影響で方向推定誤差が発生した場合についてシミュレーションにより検討する. 図3(a)は, 0° 方向から目的音(白色雑音)が到来している場合の各サブバンドの相関関数(図2⑥の出力)をカラーで表示したものである. 図において横軸は相関関数の時間軸を推定方向に変換して表示した. 一方, 縦軸はサブバンドの周波数を表している. 図より全サブバンドにおいて相関値は 0° 方向で最大になっていることがわかる.

一方, 図(b)は各マイクロホンに妨害雑音として独立な白色雑音を付加した結果である. 図より, 雑音の影響で相関関数の山(赤色の相関値の大きな部分)の位置がサブバンドごとに左右に大きく変動していることがわかる. 図(c)は, 各サブバンドの相関の総和をとった結果(図2⑦の出力)を表したもので, 最大値は正解方向である 0° からずれたものとなっている.

* A study of sound source direction estimation using Sub-band Peak Hold processing with phase information, by SATO, Kohei and KANEDA, Yutaka (Tokyo Denki University).

4.2 位相情報の利用(提案法)

図4に提案法のブロック図を示す。提案法は図2の処理に加えて、⑧でSTFTの位相角 $\angle X_1(\omega, t)$ を求めて、この位相角を⑥で相関関数を計算する前に付加してやる。この時、短時間フーリエ変換を行うデータシフト長を1 sample とすると、各サブバンド信号は、対応する周波数成分の信号となる。

この位相情報を付加した場合の妨害雑音なしの場合の相関関数(図4⑥の出力)を図3(d)に示す。図より高域サブバンドほど山の幅が鋭くなっていること、またサイドローブの山の位置がサブバンド毎に少しずつ変化していることがわかる。

次に、4.1項と同様に雑音を付加した相関関数を図3(e)に示す。雑音によって相関関数の最大値は変動しているが、細かい山の位置の変動は大きくないことがわかる。図3(f)は提案法の各サブバンドの相関の総和をとった結果(図4⑦の出力)を表し、雑音が付加された場合でも正解方向で最大値を得ることができた。

5 評価実験

実験は部屋の壁際に2つのマイクを設置して音源方向検出を行った。雑音による推定精度の影響を確認するため、測定した音声信号に別途録音した室内の雑音を付加し、相対雑音レベルを30~60 dBの範囲で変化させた。音源には人間の発話で日本語の単音を使用し(71単音, 3回発話, 7方向, 全1491個)、マイク間距離 d は0.6 m、マイクから音源までの距離は2 mとし、許容誤差 $\pm 5 \text{ deg}$ での正答率を従来法と提案法で比較を行った。

実験結果を図5に示す。図は横軸に雑音レベル縦軸に許容誤差を $\pm 5 \text{ deg}$ とした場合の方向推定正答率を示す。赤線が提案法、青破線が従来法の結果である。図より提案法は相対雑音レベルが高くなって従来法より高い正答率が得られ、正答率が90%に低下する雑音レベルは約6dB上昇しており雑音耐性の向上が確認された。

6 むすび

従来のSBPH処理に位相情報を用いた手法を提案し、行った実音場実験の結果、提案法は従来法と比べて約6dB雑音耐性が向上していることがわかった。

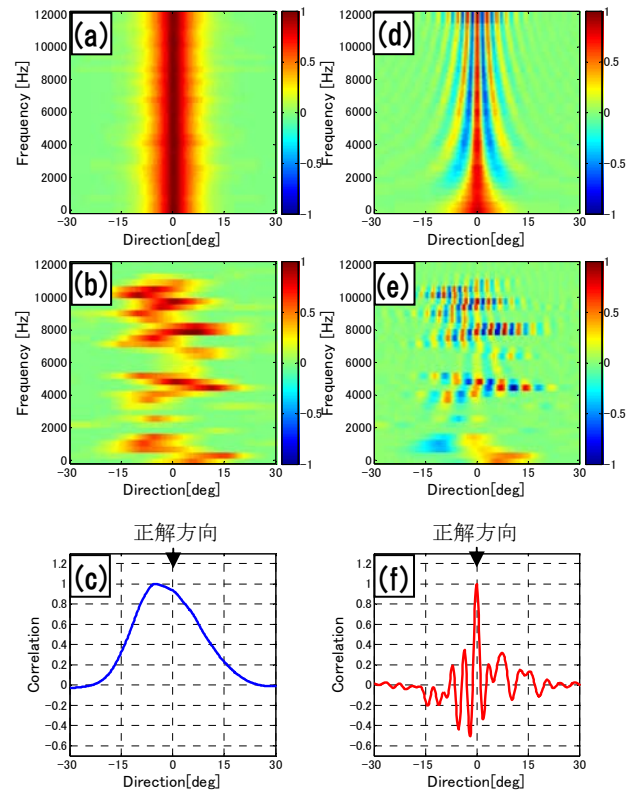


図3 従来法(左)と提案法(右)の相関関数

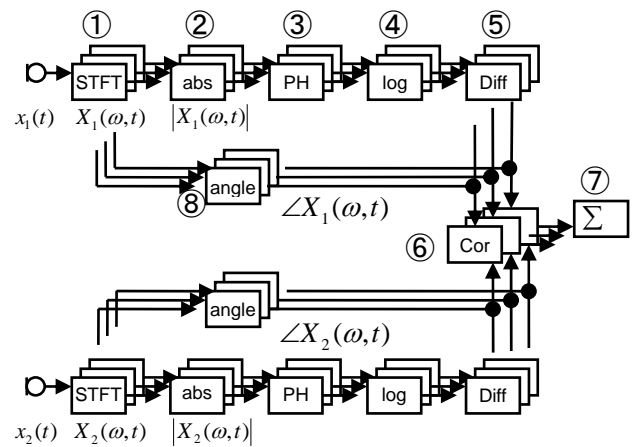


図4 提案法のブロック図

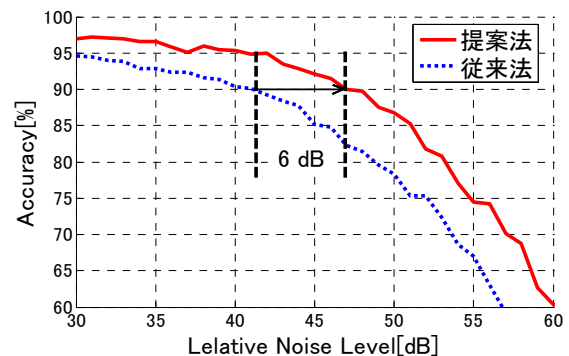


図5 評価実験の結果

参考文献

[1] 鈴木, 金田, 信学技報 EA2008-113(2008-12).