

SBPH 音源方向推定法の帯域制限の有効性の検討*

☆沼田 亮, 金田 豊 (東京電機大)

1 はじめに

一般の室内で音源方向推定を行う際、反射音の影響で誤推定が起こる. この問題に対し、反射音耐性に優れたサブバンドピークホールド(SBPH: Sub-Band Peak Hold)処理[1]が提案され、従来の代表的手法である PHAT 法(または CSP 法)[2]などに対する有用性が確認された. 本研究では SBPH 処理に位相情報を付加した SBPHP 法[3]に周波数帯域制限を加える事で、さらなる推定精度向上の検討を行った.

2 SBPHP-CC 法

2.1 相互相関関数(CC)による音源方向推定

図 1 に示すよう、2ch マイクロホンによる受信信号 $x_1(t)$, $x_2(t)$ には時間差 τ が生じる. 音源方向 θ は、音速 c とマイク間距離 d から

$$\theta = \sin^{-1}\left(\frac{c\tau}{d}\right)$$

と表すことができ、 c と d は固定値であるので、時間差 τ がわかれば音源方向 θ を求められる. 時間差 τ は $x_1(t)$ と $x_2(t)$ との相互相関関数の最大値を取る時刻 τ として求められる.

2.2 ピークホールド処理

受信した音声には、音源方向情報を含んだ直接音とそれに後続する反射音が存在する(図 2(a)). ピークホールド処理により、先行する直接音の振幅値に残響相当の減衰を持たせ、反射音をマスクし(図 2(b)), 時間差分をとることで、直接音だけを取り出す(図 2(c)).

2.3 サブバンド処理

音源方向推定において直接音、つまり信号の立ち上がりが重要であるが、音声信号の立ち上がりの発生時刻は帯域毎に異なる. そこで、信号をサブバンド化(帯域分割)した後に、ピークホールド処理することにより帯域毎の立ち上がり部分の取出しを可能とする.

2.4 SBPHP-CC 法

処理の流れを図 3 のブロック図に示す. 2ch マイクロホンで録音した音声を、短時間フーリエ変換(STFT)をする事で信号を帯域分割し、絶対値(abs)を取る. 次に、各帯域において PH

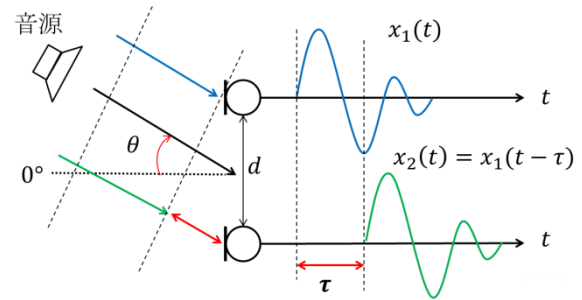


図 1 2ch マイクロホンを用いた音源方向推定

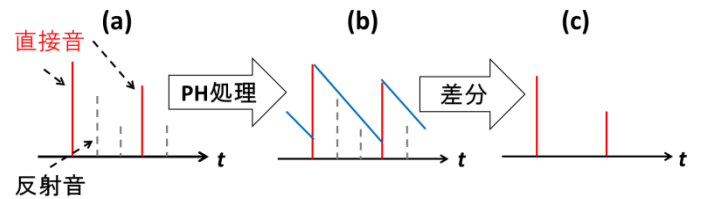


図 2 ピークホールド処理のモデル図

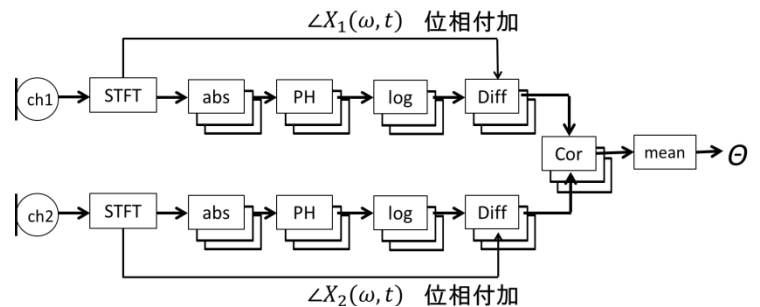


図 3 SBPHP-CC 法のブロック図

処理により反射音を除去する. そして対数操作(log)を行い、時間差分(Diff)を取る. これにより、周波数帯域毎の音声信号の立ち上がりを得ることができる.

次に、絶対値をとる時に損失した各帯域の位相成分を付加する. 最後に各周波数帯域における信号の相互相関関数(Cor)を計算し、平均化(mean)した相関値より方向推定を行う.

3 有効帯域の評価実験

相互相関関数の周波数平均化を行うにあたり、本報告では方向推定に効果のある帯域を実験により評価し、有効な帯域のみを利用して方向推定をして、推定性能の向上を試みた.

実験は反射音の影響の大きい部屋の隅に 2 つのマイクを 0.35 m 間隔で設置し、音声を録音して方向推定を行った. 測定室の寸法は 9.1 x 3.5 x 2.8 m, 残響時間は 0.4 秒であった. サンプルング周波数は 48 kHz, 帯域分割には

* A study of effectiveness of the band-limitation in sub-band peak hold DOA estimation, by Ryo NUMATA, and Yutaka KANEDA(Tokyo Denki University).

ハニング窓を使用し、帯域分割数は 9, 17, 33 とした。2名の成人男性が距離 1m, 2m, 角度 $0^\circ, \pm 30^\circ, \pm 60^\circ$ (1m のみ) で 10 種類の単語を発話し、総発話回数は 160 回であった。

4 実験結果

SBPHP-CC 法の有効な帯域を調査するため、短時間フーリエ変換された一つの帯域の相互相関関数のみから方向推定し、その推定結果を図 4 に示す。横軸は使用した周波数帯域、縦軸は誤推定数(許容誤差 $\pm 5^\circ$)を示す。

図より、どの帯域分割数においても、約 15kHz 以上の高周波数帯域において誤推定数が大きい。これは、音声の成分が高周波数帯域では小さなものになるためだと考えられる。

一方、低周波数帯域においても誤推定数が増大するが、誤推定が増大する周波数は帯域分割数によって異なり、共通して 2 バンド(第 2 帯域)目までの推定性能が低いことが確認できる。1 バンド目は方向推定に不適な直流成分の影響と考えられる。また、今回帯域分割に利用されたハニング窓は周波数分解能が低いため、2 バンド目も直流成分の影響を受けていると考えた。

図 5 は、2 種の窓関数で推定性能の比較した結果を示す。図中の緑丸より、周波数分解能が高い方形窓を利用した場合の誤推定数の上昇は 1 バンド目だけであったことから、図 4 の 2 バンド目の誤推定要因はハニング窓の影響によるものだと確認できた。

これらの結果から、SBPHP-CC 法の有効帯域は、低周波数帯域は 3 バンド目以降、高周波数帯域は 15kHz 以下と判断した。

なお、図 4 では有効帯域内であっても誤推定数が総発話の 1/3 程度となっている。これら誤推定した帯域は、音声の周波数成分を含まない帯域と考えられる。ただし、相関関数の値は小さいため、平均化する際、推定結果に悪影響は及ぼさない。

以上の帯域制限の有効性を確認するため、従来の代表的手法である PHAT 法を含めて、比較実験を行った。実験は SBPHP-CC 法および PHAT 法それぞれに対して、帯域制限の有無の場合の誤推定数を評価した。PHAT 法の帯域制限は多数の帯域制限の条件を試行して最良のものを使用した。実験結果を表 1 に示す。表より、今回示した帯域制限により

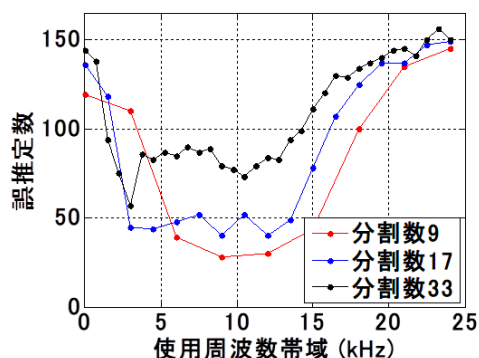


図 4 帯域分割数と各周波数帯域毎の推定性能

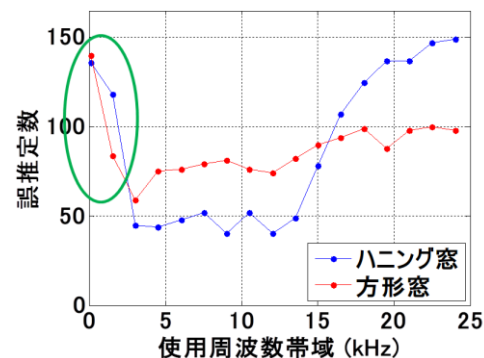


図 5 ハニング窓と方形窓の比較

表 1 帯域制限の有無による誤推定数

	帯域制限無し	帯域制限有り
SBPHP-CC	20	5
PHAT	31	17

SBPHP-CC 法の推定誤差は 1/4 に減少し、PHAT 法に対する優位性も示されている。

5 まとめ

音源方向推定において、帯域毎の推定性能を調査し、有効帯域を特定することで、推定性能の向上を試みた。低周波数帯域は直流成分、およびハニング窓による影響により 2 バンド目まで、高周波数帯域は音声の成分が小さいため 15kHz 以上の周波数帯域が方向推定に不適であることが確認できた。上記の周波数帯域を制限することで、推定性能が大きく向上することを確認できた。

謝辞

本研究の一部は JSPS 科研費 15H02728 の助成を受けたものです。

参考文献

- [1] 鈴木 他, 音講論集(秋), 751-752, (2007).
- [2] C. Knapp 他, IEEE Trans. on ASSP vol. 24, 4, pp. 320-327, (1976).
- [3] 佐藤 他, 音講論集(秋), 783-784, (2009).