

相原 洋平 岩城 護 木 竜 徹

新潟大学大学院自然科学研究科 〒950-2181 新潟市五十二風2の町 8050 番地

1. はじめに

人間には、選択的に音を受聴する能力があり、無意識のうちに音を聞き分けることが出来る。このような選択的受聴能力は、音を用いたヒューマンインターフェイスを、うるさくない高効率な聴覚補助とするための鍵のひとつである。著者らは、音源の選択的受聴を目的とした方位選択性フィルタリングシステム<sup>[1]</sup>を拡張し、方位ごとに分割された信号の振幅包絡の類似性による音源の方位推定手法を提案した<sup>[2]</sup>。そこで、本研究ではこのシステムの音源方位推定性能の検証を行った。

2. 提案手法

本研究の提案手法の概要を図1に示す。

2.1 位相差, 強度差による信号分割

指向性マイクロフォン2本が、正面から±45度開いて設置されている。指向性はカージオイド曲線で近似する。左右の受音信号は、定Q帯域幅のGammatone Filterbankによって周波数分解され、時間遅延を伴う対向伝送線路(Delayline)に入力される。左右のマイクロフォンで観測された受音信号の時間差に対応する線路上の位置において、左右の信号の位相差が一致する。また Delayline で抽出された各成分に対してマイクロフォンの指向性の強度差補正を行うことによって強度差も一致する。これらが一致する位置  $l_m$  は音源方位  $\theta_m$  によって定まるので、この一致情報は特定方位からの音を選別するために利用できる。以降では  $XR(f, l, t)$ ,  $XL(f, l, t)$  はそれぞれ左右のマイクロフォンで観測された、方位  $l(l=l_0, \dots, l_N)$ , 周波数帯域  $f(f=f_0, \dots, f_M)$  における信号を表すものとする。

2.2 強調信号作成

各 Delayline 上で観測された左右の信号の位相差と強度差の類似度を算出し、強調信号を作成する。ある音源方位  $l$  の強度差の類似度  $A_{sim}(f, l, t)$ , 位相差の類似度  $P_{sim}(f, l, t)$  を以下の(1),(2)式で定義する。

$$A_{sim}(f, l, t) = \frac{\min\{|c_l(l)XL(f, l, t)|, |c_r(l)XR(f, l, t)|\}}{\max\{|c_l(l)XL(f, l, t)|, |c_r(l)XR(f, l, t)|\}} \quad (1)$$

$$P_{sim}(f, l, t) = \max\{\cos(\angle XL(f, l, t) - \angle XR(f, l, t)), 0\} \quad (2)$$

点  $l$  での強度差を算出する際、その方位に相当する指向性の逆数  $c_l(l)$ ,  $c_r(l)$  により伝播信号を補正している。この補正により点  $l$  相当の方位音であれば、補正した伝播波形の強度が一致するため類似度は大きくなるが、他の方位音の伝播成分では小さくなる。類似度を delayline の全ての点  $l$  で計算し、図2で示す写像関数により変換された重みに基づき、伝播信号  $XL(f, l, t)$ ,  $XR(f, l, t)$  を減衰させ出力する。処理後の左右音の中から、合成信号  $X(f, l, t)$  を算出する。

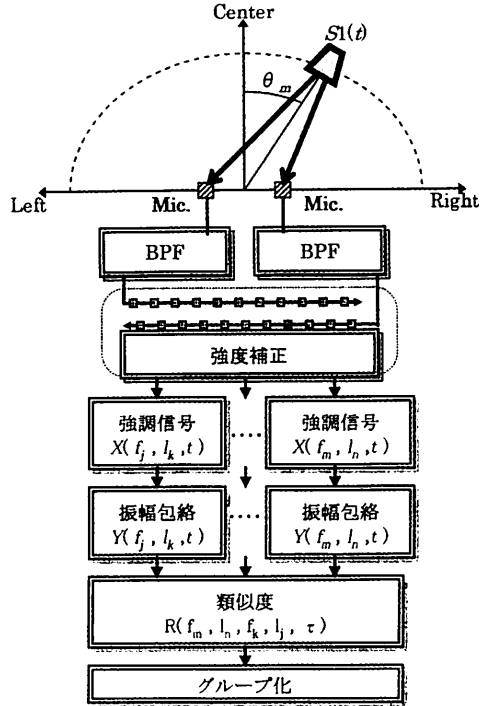


図1 提案手法概要

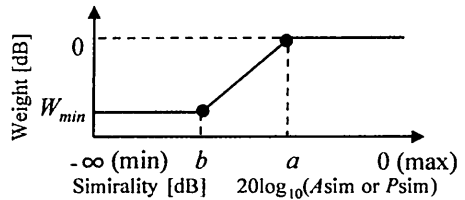


図2 写像関数

2.3 振幅包絡とそれらの類似度の算出

信号  $X(f, l, t)$  の絶対値を短時間平均より平滑化し、振幅包絡の時系列  $Y(f, l, t)$  を算出した。この  $Y(f, l, t)$  から信号の類似度  $R(f_m, l_n, f_k, l_j, t)$  を算出した。類似度は各帯域、方位  $(f_m, l_n)$  ( $m=0, 1, \dots, M-1, M; n=0, 1, \dots, N-1, N$ ) での振幅包絡  $Y(f_m, l_n, t)$  に対して、 $l_n$  の近傍方位  $l_k$  と全ての周波数帯域  $f_j$  との組み合わせ  $(f_j, l_k)$  ( $j=0, 1, \dots, M-1, M; k=n-1, n, n+1$ ) の振幅包絡との間での相関によって計算する。

2.4 グループ化

算出された類似度に対し、グループ化を行う。得られた類似度  $R(f_m, l_n, f_k, l_j, t)$  の中で最も類似度の値が高い組み合わせ  $\{(f_m, l_n), (f_j, l_k)\}$  を一つのグループに統合する。順次グループ化を行い、適度なグループ数になるまで同様のグループ化を繰り返した。

3 シミュレーション

3.1 シミュレーション設定値

本研究では強調信号作成時の重み関数が音源方位推定性能に与える影響を評価するために、写像関数での類似性の閾値を変化させた際の方位推定の精度の評価を行った。最適閾値を調べるために図2のa, bは同一の値とした。シミュレーションの設定値を表1に示す。シミュレーションは重み計算での閾値  $a(=b)$  を0~-10まで変化させた場合の音源方位推定をそれぞれ行い、精度の評価を行った。音源数は2つとし、それぞれ左90度と正面に設置した場合、右20度と左20度に設置した場合について行った。

3.2 方位推定結果

音源入力方位が /左90度, 正面/ の場合について、方位推定結果より上位2グループの重心をそれぞれ図3に示す。また、各方位推定グループの分散を図4に示す。同様に音源入力方位が /左20度, 右20度/ の場合についての結果を図5, 6にそれぞれ示す。これらの図より、抽出された各グループは音源方位に確認することができた。図3では写像関数の閾値が0と-8のときに、有効なグループが1つだけとなった。ここで、有効なグループとは、方位・周波数の分解数(17×19)に対して5%以上の大きさを持つものとした。また、図5では写像関数の閾値が-1, -2の時にのみ有効グループが2つ確認された。また、図4, 6でのグループの分散結果より、閾値を大きくするに当たってグループが大きくなり1つのグループに統合される事が確認できた。これは図2の強調信号のための減衰重みがゆるくなったために、近傍方位同士の信号が混入したまま強調されてしまい、以後のグループ化処理で同一のグループと判断されるためであると考えられる。またどちらの結果でも閾値が0の場合は性能が悪化しており、類似性に対する基準がきつすぎると思われる。

4 まとめ

シミュレーションにより強調信号算出時の写像関数の閾値を検討した。その結果、方位推定結果に有効なグループが確認されたのは閾値が-1~-2付近の時であった。今後の課題として、方位推定精度の向上、様々な音を入力した際の定位、音源種類の特定などがあげられる。

表1. 設定したパラメータ

マイクロフォン間隔	80 [mm]
マイク間の最大時間差	450 [ $\mu$ s]
指向性の最大減衰量	-5 [dB]
中心周波数	60 [Hz]~6 [kHz]
周波数分割数( $l$ )	17Ch
方位解析範囲( $\theta$ )	19Ch(180deg)
最大減衰値 $W_{min}$	50 [dB]
標準化周波数	20 [kHz]
入力信号長	1[sec]
包絡推定信号長	1[sec]
入力信号 A	[すいません]
入力信号 B	[はじめまして]

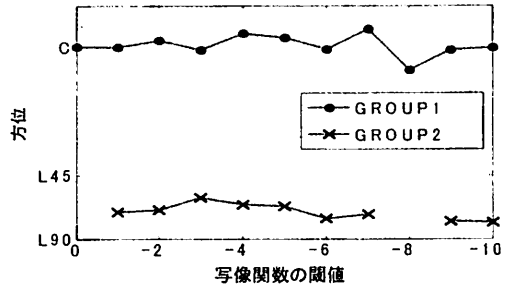


図3 入力方位左90度, 正面でのグループ重心

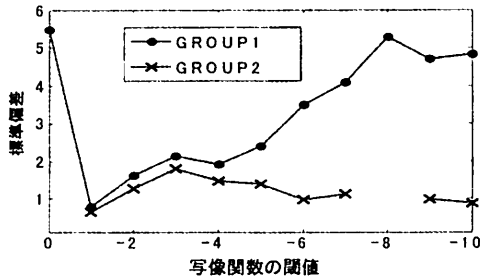


図4 入力方位左90度, 正面でのグループの分散

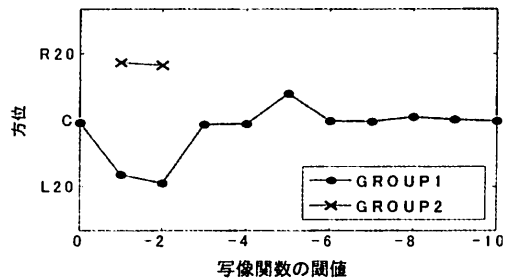


図5 入力方位左20度, 右20度でのグループ重心

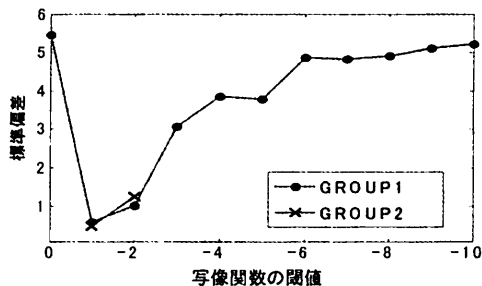


図6 入力方位左20度, 右20度でのグループの分散

参考文献

- [1]手塚智章, 岩城護, "音源の選択的受聴を目的とする方位選択性フィルタリングシステムの検討", 日本音響学会秋期大会, 3-Q-14, 2004-9.
- [2]相原洋平, 岩城護, 木竜徹"振幅包絡時系列を利用した音源到来方位推定手法の検討", 信学技報, SP2005-36, pp 7-12.