

F-13

指向性マイクを利用した移動音源分離

Moving source separation from their mixture using directional microphones

伊藤 雅紀† 竹内義則† 松本哲也†
Masanori Ito Yoshinori Takeuchi Tetsuya Matsumoto工藤博章‡ 向井利春* 大西昇†
Hiroaki Kudo Toshiharu Mukai Noboru Ohnishi

1. まえがき

音源分離の分野でさまざまな音源分離手法が提案されている。しかし、実環境で分離を行っているものは少ない。それらの中には、時間領域で分離を行うもの[1]や、周波数領域で分離を行うもの[5]がある。しかし、音源が移動しているときを扱っているものは少ない[4]。移動音源を扱う場合は、混合行列が一定でなく時変となるために、短時間に信号を分割し、その分割された信号ごとに分離フィルタを計算しなければならない。例えば、移動速度が速い場合、ごく短く分割しなければならない、しかしごく短い信号を分割するときにはデータ数が不足して分離がうまく行えないことが起こりうる。本論文では、マイクに指向性マイクを使用し、その指向特性を利用することで位置などに対し、混合行列の変化を小さくし、移動音源に対しても分離を可能にする手法を述べる。なお、分離には、河本らによる音源の非定常性を用いた2次統計量を利用したアルゴリズムを用いた[1]。

2. 混合モデルと分離方法

2.1 混合モデル

簡単のため音源数を2として、音源1を固定、音源2を移動したときのみを考える。

源信号を $s(t)$ 、混合行列を $A(z)$ 、混合信号を $x(t)$ とすると、次の関係が成り立つ

$$x(t) = A(z)s(t) \quad (1)$$

ここで $A(z)$ はシステム関数を要素に持つ行列で、

$$A(z) = \begin{pmatrix} a_{11}(z) & a_{12}(z, t) \\ a_{21}(z) & a_{22}(z, t) \end{pmatrix} \quad (2)$$

である。音源の移動によって、伝達関数が時間とともに変動するものと、時間によらず一定のものが存在する。推定する分離行列を B はアルゴリズムの性質上対角成分は1であり、

$$B(z) = \begin{pmatrix} 1 & b_{12}(z) \\ b_{21}(z) & 1 \end{pmatrix} \quad (3)$$

と表すことができる。分離で順序の変化がおきないと仮定すれば、 BA が対角行列になれば分離ができているので、 BA の各成分について次の等式が成立する。

$$b_{12}(z)a_{12}(z, t) + a_{22}(z, t) = 0 \quad (4)$$

$$b_{21}(z)a_{11}(z) + a_{21}(z) = 0 \quad (5)$$

(4)から b_{12} は時間とともに変化する分離パラメータであり、



図1 マイク近接配置

(5)から b_{21} は時間によらず一定であるということがわかる。音源が一つ静止であるということがわかっている場合は、分離行列の一つのみが変化するという性質を利用することで、うまく分離処理を行うことが考えられる。

2.2 分離方法

音源が移動する場合、混合信号を分割してその分割された信号で分離処理を行えばよいと考える。ただし、その分割された信号では音源が静止していると仮定できるような分割でなければならない。すなわち、分割が大きすぎるとある時点とある時点の分離パラメータが異なってしまう、同一の分離パラメータで分割信号を分離することができないと考える。

分離パラメータは音源からマイクに到達するまでの時間差に依存した値に収束することがわかっている[2]。したがって、音源からそれぞれのマイクの距離を等しくすれば、いつでも同じような収束点に分離パラメータが収束すると考える。そこで、指向性マイクを、そのマイクの指向性を異なった方向に向けて分離する手法を提案する[3] (図1参照)。このように配置すればマイクの受音部を同じ個所に配置して、指向性により、複数の信号の入力レベルで差をつけている。この手法では、直接波のみが存在し、周波数などによらず指向特性が一定であるとすれば瞬時値混合であるとみなすこともできる。ただし、実際には瞬時値混合にはならず、畳み込み混合を考えなくてはならない。

また分割された信号を独立に分離処理を行うので、順列、大きさの違いが発生するので解決しなければならない。

3. 分離パラメータの変化

図2のように指向性マイクを離して配置した場合と、図3のように指向性マイクを近接させて配置した場合で、手で音源を10cmずつ移動させて混合信号を観測する。その観測信号をそれぞれ用いて分離をして、得られる分離パラメータ(200tap)の変化を観測した。

図2のようにマイクを離して配置した場合は、分離パラメータの変化は移動する音源とマイクとの距離と密接した関係で変化し、10cmといったわずかな移動距離で移動前と移動後の分離パラメータの相関係数は0.1程度になった。

図3のようにマイクを近接して配置し、指向性を別の方向に向けた場合、30cm程度移動しても前後の分離パラメータの相関係数は0.6以上を保っていた。したがって、図3のように指向性マイクを配置した場合は、分離信号を分割して分離処理を行うとき、混合信号をおおまかに分離し

†名古屋大学大学院工学研究科

‡名古屋大学情報メディア教育センター

*理化学研究所バイオミメティックコントロール研究センター

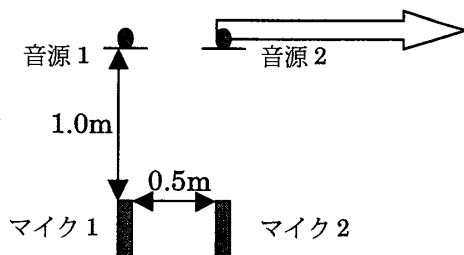


図2 マイクを離れた実験配置

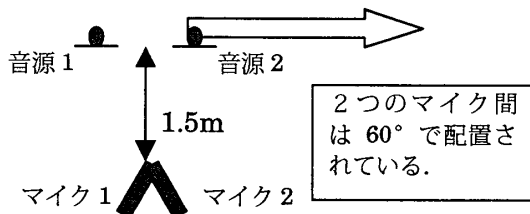


図3 マイクを近接した実験配置

でも同一の位置にいると仮定して、一つの分離パラメータで処理をすることができると考える。また、この手法を用いることで分離のタップ数の低減を考慮することができ、処理の高速化も望める[3]。

4. 分離実験

図3と同じ配置で、音源1を固定、音源2を移動させて混合信号を観測した。このとき、音源はスピーカを用いて、スピーカをロボット(ActivMedia, Pioneer2)に固定してロボットを移動させた。

実験環境は、残響時間が0.81秒の実験室を用いて、指向性マイクにはSONYのECM-670を用いた。サンプリング周波数は16kHz、ロボットは毎秒10cmの速度で2mの距離を等速直線運動させて20秒間信号を観測した。この信号を2秒ごとに分割し(ただし前の分割と1秒のオーバーラップをもうけている)、それぞれを分離処理をして結合した結果が図4である。このとき、単純に信号をつなぎ合わせたものであり、順序の入れ替えはこのときは発生しなかったが、発生したら対処方法を考えねばならない。

この図4から源信号と分離信号が似通っていることから分離処理がうまく行えていることがわかる。

5. 考察

指向性マイクの指向特性を異なった方向に向けることで、理想的には瞬時値混合をつくりだし、移動による分離パラメータの変化が小さいと考えることで、移動していてもある一点に静止していると仮定できる距離を大きいと考えることができる。すなわち、移動に対して、分離パラメータの変化は小さいと考えることができる。その点からマイクを近接させた手法を移動音源に用いるのは効果的であると考える。

音源を一つのみを移動させた場合、分離パラメータは4つのうち1つしか変化しない。この性質を利用すれば分割された信号を結合するとき順列やスケールングの問題についても解決できるのではないかと考える。また1秒のオーバーラップ部分を利用して、順列の問題を解決する手法も考えられる。

6. まとめ

指向性マイクの指向特性を利用した移動音源分離の手法を提案した。

今後の課題としては、現在これの分離性能について評価していないので、性能評価が課題である。また、移動速度が遅いときしか行っておらず、さらに速く移動した場合の実験も必要である。また、画像処理などを用いて指向性を音源に向けなければならない。

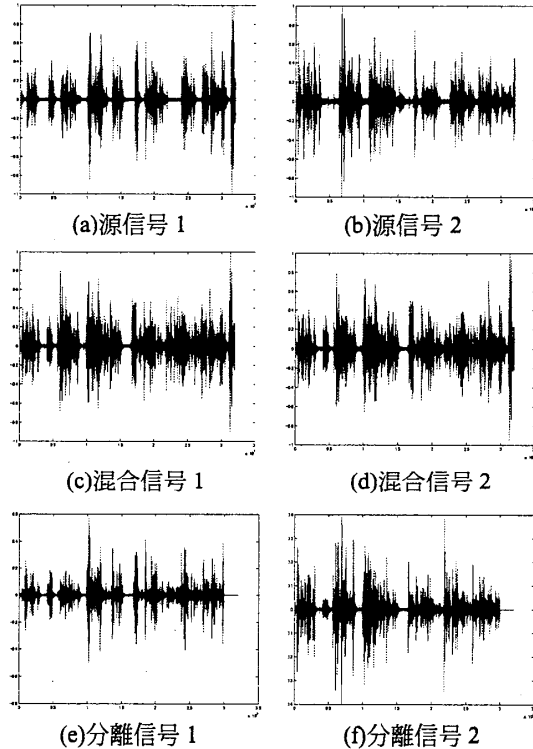


図4 信号の分離

参考文献

- [1] 河本満, A.K. Barros, A. Mansour, 松岡清利, 大西昇, "時空間的に混合した非定常信号のブラインド信号分離," 信学論(A), Vol.J82-A, No.8. pp.1320-1328, Aug.1999.
- [2] 丸吉政博, 河本満, 竹内義則, 松本哲也, 工藤博章, 大西昇, "分離パラメータ収束点推定を用いた源信号分離の収束性能の向上," 信学技法 EA-2001-122.
- [3] 片山雄介, 竹内義則, 松本哲也, 工藤博章, 大西昇, 向井利春, "実環境音源分離におけるタップ数の低減," 2002年電子情報通信学会総合大会, A-4-47, p.157, 2002.3.
- [4] 澤井克之, 猿渡洋, 鹿野清宏, "音源方位情報を利用した独立成分分析によるセミブラインド移動音源分離," 信学技法 EA-2001-34.
- [5] S. Ikeda, N. Murata, "A method of ICA in time-frequency domain," Proc. ICA'99, pp.365-371, 1999.