

Matching Pursuit による音源分離法の性能改善

学籍番号：90163021 飯國研究室 大久保智広

1. はじめに

近年、複数の音声が入力された信号から個々の音声を分離するブラインド音源分離が盛んに研究されている。音源分離手法として、時間周波数解析法である Matching Pursuit (MP) を用いたステレオ MP が提案されている。この方法はステレオ入力から 3 音声以上の分離が可能である。MP の基底選択方法は、1 回の基底選択毎に残差の 2 乗ノルムを最小化する選択法であるが、必ずしもこの選択基準が音声分離に適しているとは限らない。そこで、MP による音源分離の性能改善を目的として、MP の基底選択法を検討する。

2. 最大基底選択ステレオ MP

MP は信号のサンプル数を越える個数の基底を持つ基底系から、逐次的に基底を選択して信号を近似していく信号分解法である。この MP をステレオ信号に適用するために提案され、従来の音源分離で用いられているステレオ MP では、基底選択基準を左右の内積の 2 乗和としているために信号全体の近似性能は高い。これに対し本研究では、基底選択基準を個々の内積の絶対値とすることで個々の信号の特徴を示す基底を選択できる、最大基底選択ステレオ MP を提案する。ここで、左右のチャンネルの信号が I 個の音源から混合され、

$$(f_l(t), f_r(t)) = \sum_{i=1}^I (\cos \theta_i x_i(t), \sin \theta_i x_i(t))$$

と受信されているとする。 θ_i は両チャンネルに現れる信号の振幅比を決めるパラメータであり、パンポットパラメータと呼ばれる。はじめに、左右のチャンネルの初期残差信号に $Rf_l^{(0)}(t) = f_l(t), Rf_r^{(0)}(t) = f_r(t)$ を与えて、次の手順を反復して行う。

1. D に含まれる全ての基底に対して左右の残差信号との内積 $|\langle Rf_l^{(m)}, g \rangle|, |\langle Rf_r^{(m)}, g \rangle|$ を計算する。

2. 残差信号に対して最良近似を実現する基底

$$g_{\gamma_m} := \arg \max_{g \in D} |\langle Rf_c^{(m)}, g \rangle|, \quad c = l, r$$

を選択する。また、 g_{γ_m} に対応する内積の最大値 $p_{l,m}, p_{r,m}$ を係数として記録する。

3. 左右の残差信号を次のように更新する。

$$Rf_c^{(m+1)}(t) := Rf_c^{(m)}(t) - p_{c,m} g_{\gamma_m}(t), \quad c = l, r$$

また、分解後に左右の内積値 $p_{l,m}, p_{r,m}$ の絶対値から

$$\theta_m = \arg(|p_{l,m}|, |p_{r,m}|)$$

としてパンポットパラメータを算出し、 θ_m についてのヒストグラムを作成して、参照し基底を分類することで 2 つ以上の音声を分離することができる。

3. ヒストグラムを考慮した基底選択

音源分離性能を向上させるには、基底の分類を適切に行うことができればよいことが予想される。また反復を最後まで行わなくとも、反復の初期にはヒストグラムの概形が決まることが予想される。これらのことから、過去に選択された基底のヒストグラムを考慮し重み付けをして、個々の音声の持つ θ_i に近い θ_m を持つ基底を優先的に選択することで分離能力の向上を図る。

4. シミュレーション

従来法であるステレオ MP の結果と、本研究での提案法である最大基底ステレオ MP とヒストグラムを考慮した基底選択を組み合わせた結果の比較を行う。8000kHz サンプリング、3 秒間混合音声について 2000 個の基底に分解した。基底系にはサンプル数に対して 992 倍の冗長度を持たせた Gabor 関数の基底群を用いた。図 1 には、提案法での 4 音声のヒストグラムを、表 1 には SN 比による結果を示す。SN 比の向上から、提案法で分離性能が向上できることを確認できた。

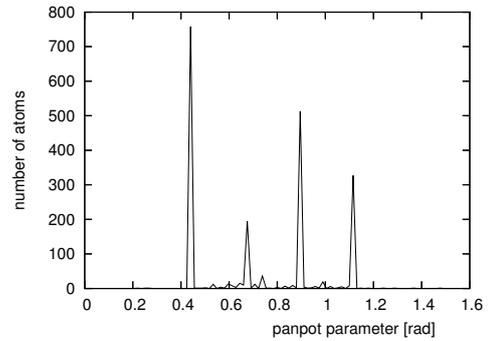


図 1: 4 音声のヒストグラム (提案法)

表 1: SN 比による実験結果 (上: 3 音声, 下: 4 音声)

	分解前	従来法	提案法
女声 1	-3.0	7.4	9.3
女声 2	-2.9	4.8	6.8
女声 3	-3.0	6.8	8.9
女声 1	-4.9	4.8	5.9
男声 1	-4.8	2.4	2.8
女声 2	-4.8	2.7	3.1
男声 2	-4.7	4.5	4.7

5. 結論

本研究では、提案法により音源分離の分離性能を改善できることを示した。また、基底選択の基準について大きな自由度があることがわかった。今後、基底選択法を改善することで、さらなる分離能力の向上が期待できる。