

実時間信号処理システムを用いた音声信号の適応型ブラインド分離

吉岡 慎二* 河野 清尊*

* 米子工業高等専門学校

Adaptive Blind Source Separation for Audio Signals Using the Real Time Signal Processing System

Shinji Yoshioka* and Kiyotaka Khono*

*Yonago National College of Technology

Abstract: In this paper, we consider that the proposed adaptive AMUSE algorithms are effective for blind source separation of audio signals through computer simulation. Then we propose a new adaptive algorithm and clarify problems for implementing the adaptive algorithm in the real time signal processing system.

1. はじめに

空間的混合 MIMO システムにおける音声信号のブラインド分離アルゴリズムとして、AMUSE が提案されている¹⁾。

本研究では、AMUSE の実時間処理への拡張として、まず、既に提案した適応型 AMUSE (シリアル型 AMUSE およびパラレル型 AMUSE)²⁾ について、そのアルゴリズムの有効性を計算機シミュレーションにより確認するとともに、問題点を明らかにする。

次に、明らかになった問題点を解決するために、新たな適応型 AMUSE (逐次型 AMUSE) を提案し、その有効性を計算機シミュレーションにより検証する。

2. AMUSE アルゴリズム

AMUSE (Algorithm for Multiple Unknown Signals Extraction)¹⁾ は原信号 $s(t)$ の独立成分が時間的に相関のある有色信号であるとき、自己共分散を用いて分離システム W を推定するブラインド分離アルゴリズムである。観測信号 $y(t)$ の自己共分散 $C_y(t)$ は

$$C_y(t) = E[y(t)y(t)^T] \quad (1)$$

これを固有値分解すると、

$$C_y(t) = U\Lambda U^T \quad (2)$$

となり観測信号を無相関化する行列 V が求まる。

$$V = \sqrt{\Lambda}^{-1} U^T \quad (3)$$

観測信号を行列 V で無相関化する。

$$x(t) = Vy(t) \quad (4)$$

$x(t)$ に対する任意の時間差 τ の時間差共分散は

$$C_x(\tau) = E[x(t)x(t-\tau)^T] \quad (5)$$

となり、これを対称化して、

$$\bar{C}_x(\tau) = \frac{1}{2}[C_x(\tau) + C_x(\tau)^T] \quad (6)$$

$$= W^T \bar{C}_s(\tau) W \quad (7)$$

すなわち、時間差共分散分離 $\bar{C}_x(\tau)$ を求め、固有値分解することで、分離システム W を推定することができる。

3. 適応型 AMUSE

オリジナル AMUSE は、ある長さ N の観測信号列を用いてバッチ処理で分離システム W を推定するものであり、再推定を行わない限りは分離システム W が変更されることはない。従って、このようなバッチ処理型の AMUSE では、混合システム H が時間的に変化するような場合には対応できない。そこで、混合システム H が事変の場合にも対応できる適応型 (オンライン処理) のアルゴリズムが必要となる²⁾。

3.1 既存の適応型 AMUSE

このような適応型 AMUSE として、シリアル型 AMUSE およびパラレル型 AMUSE を提案した²⁾。

シリアル型の適応型 AMUSE は、観測信号 $y(t)$ を一定長 N ずつ確保し、 N 時間ごとに分離システム W を推定し更新するアルゴリズムである。

一方、パラレル型の適応型 AMUSE は、複数のシリアル型を並行して実行するもので、段数 l に対して、 $M (= N/l)$ 時間ずらしながら観測信号 $y(t)$ を一定長 N ずつ確保し、 M 時間ごとに分離システム W を推定し更新するアルゴリズムである。

3.2 シミュレーション結果

原信号 $s(t)$ には、入力数 $n = 2$ 、サンプリング周波数 44.1kHz、データ長 296×10^4 (約 1 分) の 2 人の音声信号を用い、 $t = 147 \times 10^4$ において混合システム H が $H1$ から $H2$ に変化するものとする。分離システム W を推定するために必要なデータ長 N は $N = 2 \times 10^4$ [個] であった。

$$H1 = \begin{bmatrix} 0.913 & 0.249 \\ 0.183 & 0.831 \end{bmatrix}, H2 = \begin{bmatrix} 0.913 & 0.249 \\ 0.438 & 0.700 \end{bmatrix}$$

分離精度の評価尺度としては多チャンネル符号間干渉 M_{ISI} [dB] を用いる。

Fig:1, Fig:2, Fig:3 に、時変の場合のオリジナル AMUSE (非適応型)、シリアル型 AMUSE、パラレル型 AMUSE のシミュレーション結果をそれぞれ示す。

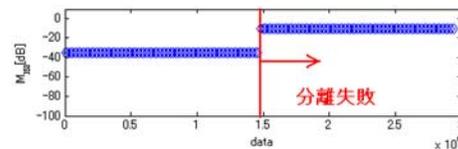


Fig.1 Original AMUSE

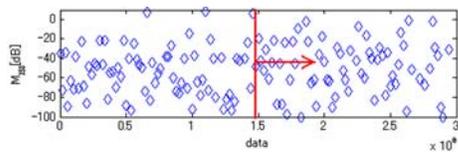


Fig.2 Adaptive AMUSE(Serial type)

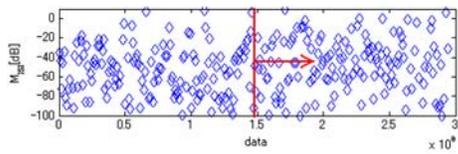


Fig.3 Adaptive AMUSE(Parallel type)

Fig:1 より, 非適応型 AMUSE では, 混合システム H が時間的に変化すると分離に失敗するが, Fig:2, Fig:3 より, シリアル型とパラレル型 AMUSE では, 時変の場合でも分離できることが分かった. しかし, 分離システム W の精度にばらつきがあること, また, 分離した音声を実際に聞いてみると, 分離システム W の再推定時に, 分離信号の順番が入れ替わる問題があることが判明した.

前者の問題は, $N = 2 \times 10^4$ [個] ごとの異なる観測信号 $y(t)$ を用いて W を再推定していることが原因であると考えられる. また, 後者の問題は, 復元信号の順番に関する任意性の問題であり, AMUSE の分離アルゴリズムでは解決できない問題であると考えられる³⁾.

4. 新しい適応型 AMUSE (逐次型)

既存の適応型 AMUSE に存在する分離精度のばらつきの問題を解決するために, 逐次型 AMUSE を提案する.

4.1 逐次型 AMUSE のアルゴリズム

このアルゴリズムは, 各時刻ごとの観測信号ベクトル y' , および無相関化した観測信号ベクトル $x'(t)$ を用いて, 自己共分散行列 $C_y(t)$ および時間差 の時間差共分散行列 $C_x(t)$ を次のようにその時刻までの移動平均で求める.

$$C_y(t) = (1 - \alpha)C_y(t-1) + \alpha y'y'^T \quad (8)$$

$$C_x(t) = (1 - \beta)C_x(t-1) + \beta x'(t)x'^T(t-1) \quad (9)$$

後は AMUSE のアルゴリズムにより各時刻ごとに分離システム W を推定し更新するアルゴリズムである. ただし, $0 < \alpha < 1, 0 < \beta < 1$ とする.

4.2 シミュレーション結果

3.2 と同様の条件で計算機シミュレーションを行った. ただし, 原信号は最初の 10×10^4 [個] を用い, 時変の場合には $t = 5 \times 10^4$ において混合システム H が $H1$ から $H2$ に変化するものとする. また, 初期値として以下のような値を用いた.

$$C_y(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, C_x(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= 0.0001, \quad = 0.0001,$$

Fig:4 に混合システム H が時不変の場合の, Fig:5 に時変の場合のシミュレーション結果を示す.

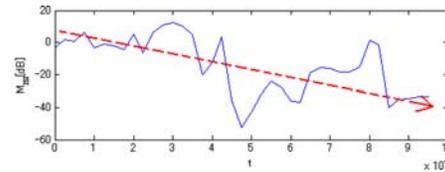


Fig.4 One-by-one AMUSE (Time invariant)

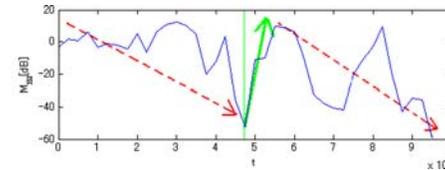


Fig.5 One-by-one AMUSE (Time variant)

Fig:4 および Fig:5 より, 時間が経過するごとに分離精度が向上していることがわかる (破線矢印). また, Fig:5 より, 時変の場合には分離精度が一時的に悪くなる (実線矢印) がすぐに適応していることがわかる (破線矢印).

また, 分離システム W の精度がばらつく現象であるが, Fig:4, Fig:5 のように時間変化によって緩やかに変化はしていても瞬間的に大きく変化はしていないことが分かる. これにより分離システム W がばらつく現象が解決できたと思われる. しかし, M_{ISI} が平均的には十分な値にならず, 分離が完全に成功しているとは言いがたい. これは, α, β を調整することで解決すると考えられる.

5. おわりに

- 今後の課題としては以下のものが挙げられる.
- ・逐次型 AMUSE のアルゴリズムの改良
- ・オリジナル AMUSE の実時間信号処理システムへの実装
- ・適応型 AMUSE の実時間信号処理システムへの実装と, 実環境でのブラインド音声分離の実現

参考文献

- 1) Aapo Hyvarinen, Juha Karhunen, Erkki Oja 著, 根元幾, 川勝真喜訳: 詳解 独立成分分析 信号解析の新しい世界, pp. 42-45, pp. 386-387, 東京電機大学出版局, (2006).
- 2) 荒木貴行, 吉岡慎二, 河野清尊: "実時間信号処理システムを用いた音声信号のブラインド分離", 第 17 回 計測自動制御学会中国支部学術講演会論文集, pp. 64-65(2008).
- 3) 村田昇: 入門 独立成分分析, pp. 1-60, 東京電機大学出版局, (2004).