

VT-LMSを用いたマルチチャネル エコーキャンセラに関する研究

◎園田 哲也 嶋田 浩和 宇佐川 毅 江端 正直
(熊本大学)

1. はじめに

近年、急速にインターネットが普及し、またインターネットの通信費は距離に依存しないという事から、非常に安価になってきている。一方、SOHO(Small Office/Home Office)や Branch Office など、小規模なワークスペースが急激に増加している。そういった状況の中で、離れた場所との会議やコミュニケーションが必要になり、現在、遠隔会議システムが注目されている。通常遠隔会議システムとは、ハンドセットの電話と違い、スピーカとマイクロホンを用いて、複数話者同士の会話を可能にするものである。そのため、効率良く音声を集音するには複数のマイクロホンを用いる事が有効である。しかし、スピーカからの相手側の音声は再びマイクロホンに集音され、エコーとして相手側に伝わってしまうため、複数のマイクロホンに対応した高品位なエコーキャンセラが必要である。図1に実際に伝達関数を求めるために使った電話器、POLYCOM社製の Soundstation Premier を示す。



図 1. Soundstation Premier

2. システム

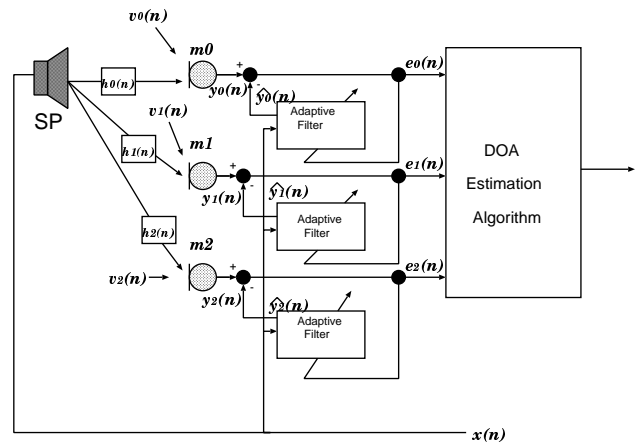


図 2. システムのブロック図

まず、それぞれのマイクロホンにエコー成分とこちら側の音声が入る。本手法では、各マイクロホンごとに適応フィルタを介し、残った成分より方向推定を行い、音声を抽出し、伝えたい信号として相手に伝える。ここで、適応フィルタではこちら側の音声はダブルトークとなり、エコーの収束を悪くしてしまう。故に、ダブルトークが検出されたら、適応動作を中断しなければならない。この処理については次に説明をする。

3. 適応アルゴリズム

本手法で用いた適応アルゴリズムは、これまでに宇佐川らによって有効性が示された、可変タップ長 LMS アルゴリズム [4] である。特長としては、学習初期にはタップ長を短くし、収束の速度をあげ、学習の終わりにはタップ長を長くし、抑制量を高めることができる。

また今回ダブルトークの検出には残留エ

コーのパワの変化を用いた [2]。しかし残留エコーのパワはエコー経路の変動によっても起こるのでその識別方法には、フィルタ出力とエコーの相関を用いた。図 3 に処理の流れを示す。

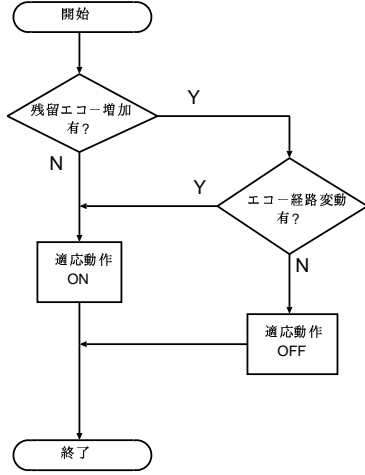


図 3. ダブルトーク検出の流れ

4. 方向推定

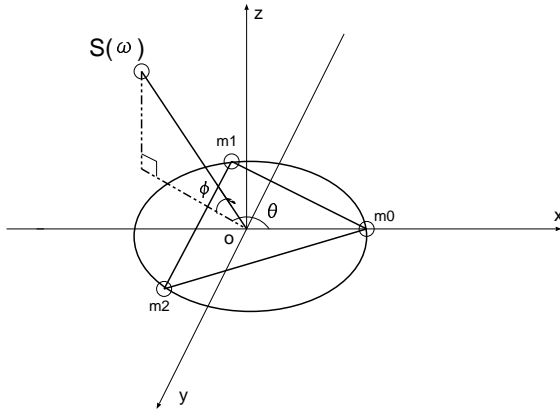


図 4. 信号の到来方向

図 4 に示すように、各マイクロホンまでの距離を一般化すると、

$$d_i = \sqrt{R^2 + r^2 - 2r \cos(\phi) R \cos(\theta - \frac{2\pi i}{3})} \quad (1)$$

基準となるマイクロホン m_0 に入る信号を次式のように示す。 A は振幅とする。

$$M_0(\omega) = AS_{(\omega)}e^{j\omega t} \quad (2)$$

この時、 m_1 , m_2 への遅延は、音速を V とすると、

$$\tau_1 = \frac{d_1 - d_0}{V} \quad (3)$$

$$\tau_2 = \frac{d_2 - d_0}{V} \quad (4)$$

となる。これよりマイクロホン m_1, m_2 に届く信号は次式のように表せる。

$$M_1(\omega) = \left(\frac{d_0}{d_1}\right)^2 AS_{(\omega)}e^{j\omega(t-\tau_1)} \quad (5)$$

$$M_2(\omega) = \left(\frac{d_0}{d_2}\right)^2 AS_{(\omega)}e^{j\omega(t-\tau_2)} \quad (6)$$

また、基準となるマイクロホン m_0 から、 $\frac{1}{3}\pi$ ずつ回転させて、同心円上の 6 点での信号のパワを観測することにする。そして、最もパワの大きい信号を伝えたい音声として抽出する。

ここで m_0 での信号のパワから順に $P_1 \sim P_6$ とすると、各点でのパワは次のように平均をとることで求めることにした。

$$P_1(t) = \alpha |M_0(t)| + (1.0 - \alpha)P_1(t-1) \quad (7)$$

$$P_2(t) = \alpha \left| \frac{M_0(t) + M_1(t)}{2} \right| + (1.0 - \alpha)P_2(t-1) \quad (8)$$

$$P_3(t) = \alpha |M_0(t)| + (1.0 - \alpha)P_3(t-1) \quad (9)$$

$$P_4(t) = \alpha \left| \frac{M_1(t) + M_2(t)}{2} \right| + (1.0 - \alpha)P_4(t-1) \quad (10)$$

$$P_5(t) = \alpha |M_2(t)| + (1.0 - \alpha)P_5(t-1) \quad (11)$$

$$P_6(t) = \alpha \left| \frac{M_2(t) + M_0(t)}{2} \right| + (1.0 - \alpha)P_6(t-1) \quad (12)$$

実際に 60° と 120° という角度で入って来た信号の方向推定を行った結果を図 5,6 に示す。

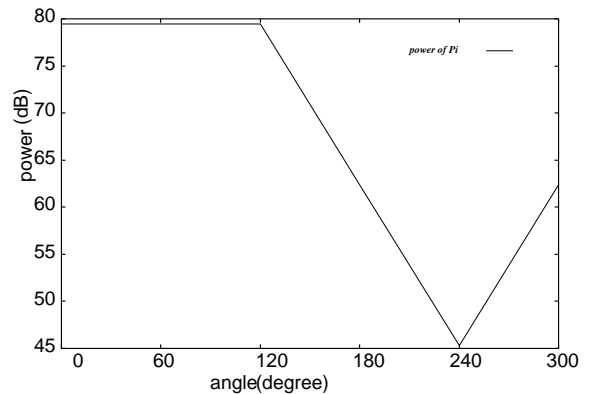


図 5. $\theta=60^\circ$ の方向推定の結果

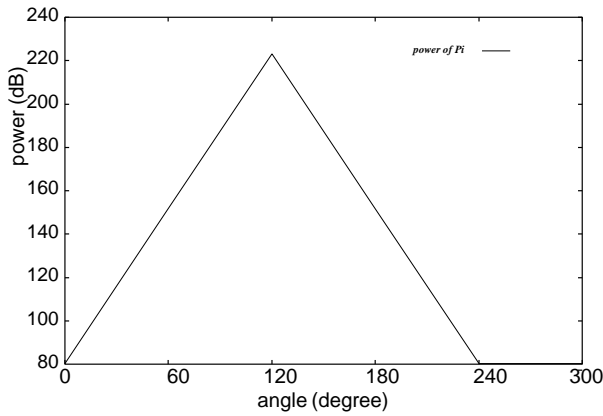


図 6. $\theta=120^\circ$ の方向推定の結果

結果から見て分かるように、それぞれ 60° 、 120° のところでピークが立っていることより、方向推定ができています。また、他の角度においても方向推定ができました。このことより、音声成分が相対的に最も強調された信号を抽出することが可能である。

5. シミュレーション

ここでは実際に図 1 を用いて求めた伝達関数を元に計算機で 3 個のマイクロホンを持つエコーキャンセラを動作させ、その性能を検討した。

5.1 ホワイトノイズ + 音声 (男性)

参照信号ホワイトノイズに対して、 $r = 1.0m$ 、 $R = 0.3m$ 、 $\theta=120^\circ$ の方向から男性の音声を挿入する。まず参照信号と目的信号を示し、方向推定の結果と、抽出するポイントでの入力信号と収束状況を示す。

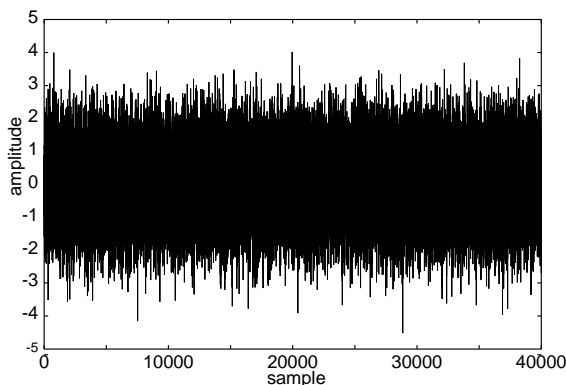


図 7. 参照信号に用いたホワイトノイズ

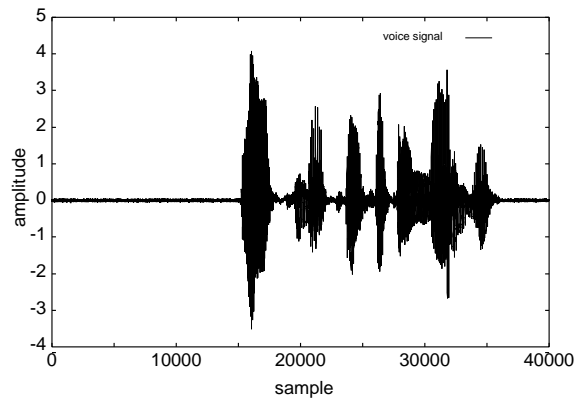


図 8. 目的信号の音声

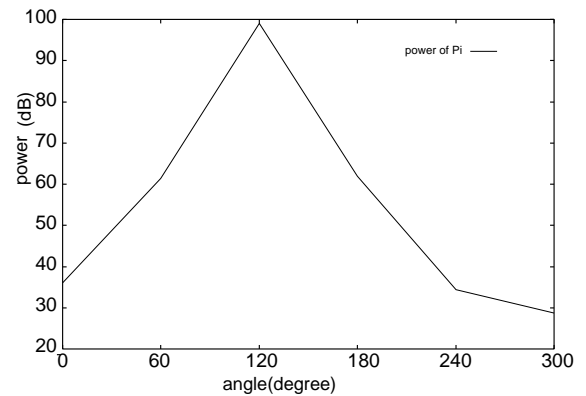


図 9. 方向推定の結果

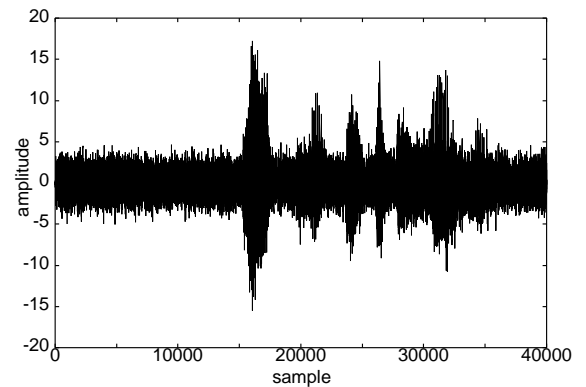


図 10. m1 での入力信号

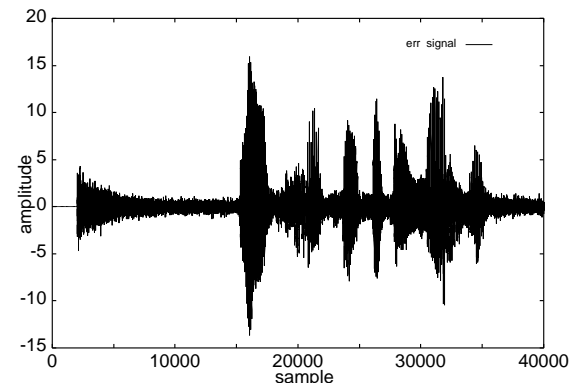


図 11. m1 での収束状況

5.2 音声(女性) + 音声(男性)

次に参照信号に女性の音声を用い、目的信号に男性の音声を付加し、シミュレーションを行った。目的信号の到来の条件は5.1.1と同じである。以下同様に結果を示す。

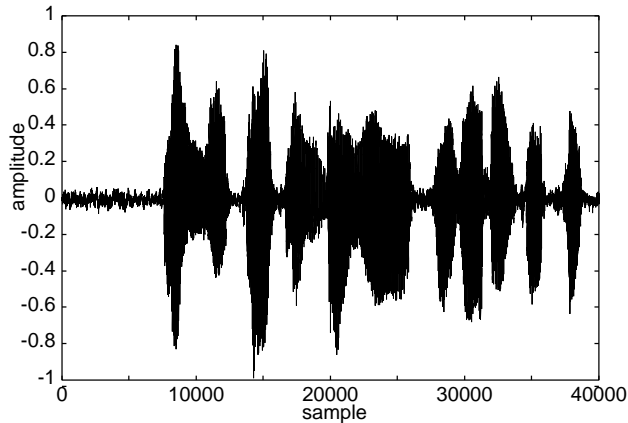


図 12. 参照信号に用いた音声

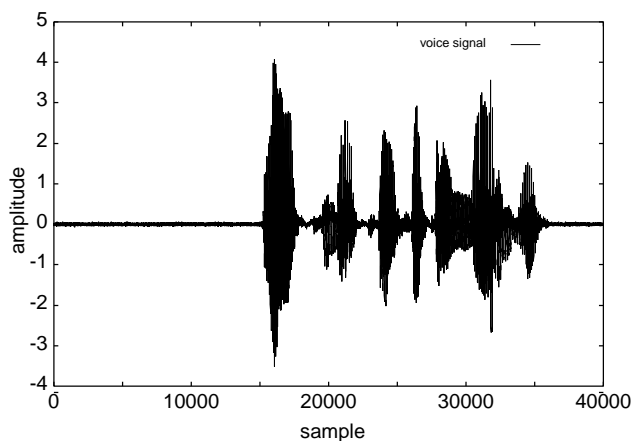


図 13. 目的信号の音声

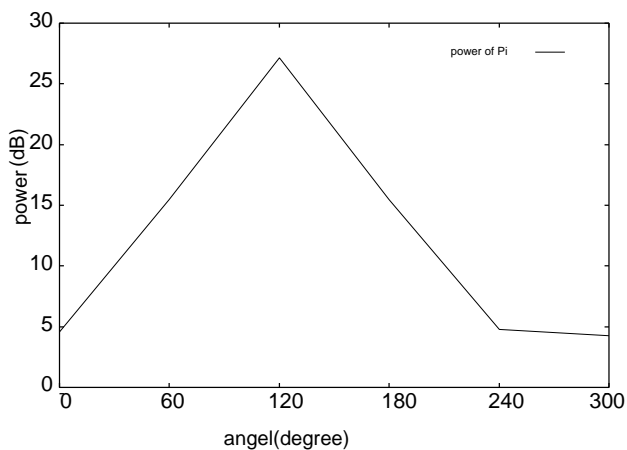


図 14. 方向推定の結果

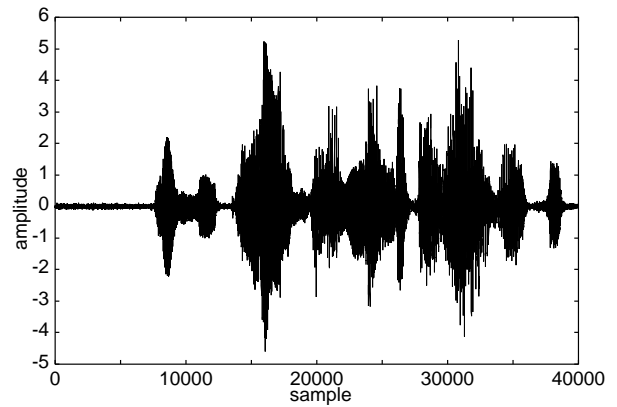


図 15. m1 での入力信号

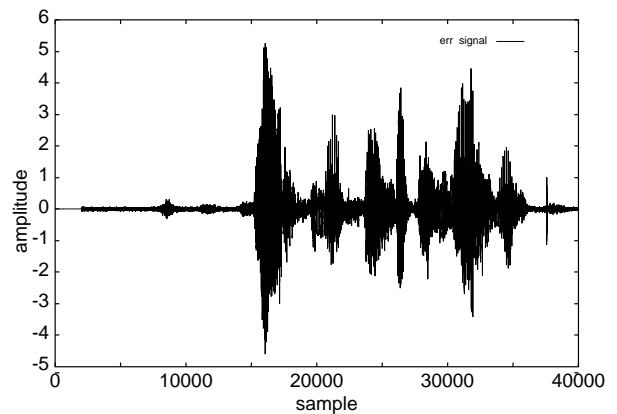


図 16. m1 での収束状況

6. まとめ

ホワイトノイズ+音声, 音声+音声の両方において方向推定が可能であった。しかし、本手法は各チャンネルごとに適応アルゴリズムを動作させているので演算量が多くなってしまう。今後はマイクロホンを有効に使い、演算量を減らすことが重要な課題である。

参考文献

- [1] Tsuyoshi Usagawa and Hideyuki Matsuo and Yuji Morita and Ebata Masanao "A New Adaptive Algorithm Focused on the Convergence Characteristic by Colored input Signal: Variable Tap Length LMS", IEICE TRANS FUNDAMENTALS, Vol. E75-A, pp1493-1499, NO.11 NOVEMBER 1992.
- [2] 森瀬 邦英, "可変タップ長 LMS アルゴリズムにおけるダブルトーク検出に関する研究," 熊本大学修士論文, 1994.
- [3] 嶋田 泰幸, "可変タップ長 LMS アルゴリズムにおけるダブルトーク検出に関する研究," 熊本大学卒業論文, 1993.
- [4] 橋本 あゆみ, "ブラインド処理を用いたディスプレイ組込型マイクロホンアレーシステムの一構成法," IECE EA2000-13, 2000.