システム/制御/情報, Vol. 46, No. 12, pp. 724-732, 2002

解　説

ステレオエコーキャンセラの課題と解決法

牧野　昭二*・島内　末廣†

1. まえがき

ステレオ音声による拡声通話システムは，臨場感の高い快適な通話手段として期待されている．ステレオシステムが作り出す空間的情報によって，遠端で誰が話しているかを聞き分けることができる．このようなステレオ拡声システムにおいて，双方向同時通話のためには，ステレオエコーキャンセラが必須である．

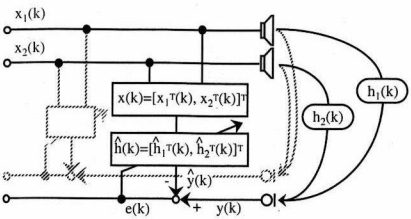
従来のステレオエコーキャンセラの最も大きな問題点は，解が真のエコー経路に収束しないこと，あるいは，収束する場合でも収束速度が非常に遅いことである [1,2]．その結果，従来のステレオエコーキャンセラは，近端のエコー経路変動ばかりでなく，遠端の送話経路変動の影響を受けてしまう．そのため，適応アルゴリズムは，近端のエコー経路変動と遠端の送話経路変動の両方に追随しなければならず，遠端の送話経路が突然変動した場合にも，ステレオエコーキャンセラの性能は劣化してしまう．このため，遠端の送話経路の変動に左右されることのないステレオエコーキャンセラを実現することが重要な課題となっており，この目的のためには，ステレオエコーキャンセラが，素早く真のエコー経路を推定できることが必要である．

聴感上聞こえない程度に無相関雑音を付加したり，左右チャネル間の相互相関に変動を与えることにより，この問題を解決することができる．さらに，無相関雑音や相互相関の変動を強調する適応アルゴリズムにより，収束速度を速めることができる．サブバンド構成もまた有効である．ここでは，ステレオエコーキャンセラの本質的な課題について議論し，最近の解決法について解説する．最後に，インプリメンテーションの一例も示す．

2. ステレオエコーキャンセラの課題

2.1 ステレオエコーキャンセラ

ステレオエコーキャンセラはステレオ信号を線形結合することにより実現できる（第1図）．入力信号ベクトル $\mathbf{x}_1(k)$, $\mathbf{x}_2(k)$ およびフィルタ係数ベクトル $\hat{\mathbf{h}}_1(k)$, $\hat{\mathbf{h}}_2(k)$ は，$\mathbf{x}(k)=[\mathbf{x}_1^{\mathrm{T}}(k),\ \mathbf{x}_2^{\mathrm{T}}(k)]^{\mathrm{T}}$, $\hat{\mathbf{h}}(k)=[\hat{\mathbf{h}}_1^{\mathrm{T}}(k),\ \hat{\mathbf{h}}_2^{\mathrm{T}}(k)]^{\mathrm{T}}$



第1図　ステレオエコーキャンセラの構成

と結合される．結合されたフィルタ係数ベクトル $\hat{\mathbf{h}}(k)$ は適応アルゴリズムによって逐次修正される．このように，ステレオエコーキャンセラは，2個のモノラルエコーキャンセラを線形結合することによって構成できる．もし左右ステレオ信号が共通の音源をフィルタリングすることによって得られたものであれば，入力信号ベクトル $\mathbf{x}_1(k)$, $\mathbf{x}_2(k)$ には，強く固定の相互相関が存在することになる．

2.2 解の誤推定の問題

モノラルエコーキャンセラと異なり，ステレオエコーキャンセラには，解の誤推定という特有の問題がある．すなわち，フィルタ係数は真のエコー経路のインパルス応答に収束しない．

指数重み付き2乗誤差の最小化，

$$J(k)=\sum_{l=1}^{k}\lambda^{k-l}e^2(l) \tag{1}$$

により，正規方程式

$$\mathbf{R}(k)\hat{\mathbf{h}}(k)=\mathbf{r}(k) \tag{2}$$

が得られる．ここで

$$\mathbf{R}(k)=\sum_{l=1}^{k}\lambda^{k-l}\mathbf{x}(l)\mathbf{x}^{\mathrm{T}}(l)=\begin{bmatrix}\mathbf{R}_{11}(k) & \mathbf{R}_{12}(k)\\ \mathbf{R}_{21}(k) & \mathbf{R}_{22}(k)\end{bmatrix} \tag{3}$$

$$\mathbf{r}(k)=\sum_{l=1}^{k}\lambda^{k-l}y(l)\mathbf{x}(l) \tag{4}$$

である．

もし，ステレオ信号 $x_1(k)$, $x_2(k)$ が

* NTTコミュニケーション科学基礎研究所
† NTTサイバースペース研究所

Key Words: echo cancellation, stereo system, adaptive filter, adaptive algorithm, teleconferencing system.

$$x_1(k)=g_1(k)*s(k),\quad x_2(k)=g_2(k)*s(k) \tag{5}$$

と表わされ，雑音がなく，$g_1(k)$，$g_2(k)$が時不変であるならば，入力信号の共分散行列(3)式はフルランクでなくなる．ここで，$g_1(k)$，$g_2(k)$は遠端送話者と2本のマイクロホン間のインパルス応答，$s(k)$は遠端送話音声，$*$は畳込みを表す．そのため，(2)式には，無限の解が存在する．$\hat{\mathbf{h}}(k)$を最小ノルム解として決定すれば，定常解$[\hat{\mathbf{h}}_1^{\mathrm{T}}(k),\hat{\mathbf{h}}_2^{\mathrm{T}}(k)]$は(1)式を最小化することによって，初期値に最も近い部分空間$\boldsymbol{H}_x$内の点に一意に収束する．ここで，$\boldsymbol{H}_x$は$\mathbf{x}_1(k)$と$\mathbf{x}_2(k)$の相互相関によって一意に決定される．しかし，これは$\hat{\mathbf{h}}(k)=\mathbf{h}(k)$ということではない．

モノラルエコーキャンセラでは，自己相関が非常に大きいにもかかわらず，解の誤推定の問題はめったに現れない．その理由は音声の自己相関は単語の中の子音や母音につれて時変であるからである（モノラルエコーキャンセラで解の誤推定の問題は自己相関が時不変となる長い母音区間にのみ起きる．この区間では，たとえ適応フィルタが真のエコー経路に収束せず係数誤差が大きくても，出力誤差を小さくすることができる）．一方，ステレオエコーキャンセラでは，$g_1(k)$，$g_2(k)$が時不変の場合には左右ステレオ信号間の相互相関は固定である．そのため，共分散行列はランク落ちを起こし，ステレオエコーキャンセラにおいて解の誤推定は大きな問題となる．

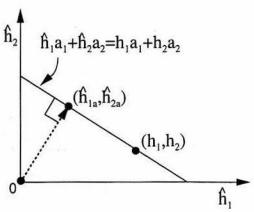
簡単な場合を例に説明する．左右ステレオ信号$\mathbf{x}_1(k)$，$\mathbf{x}_2(k)$が

$$\mathbf{x}_1(k)=a_1\mathbf{s}(k),\quad \mathbf{x}_2(k)=a_2\mathbf{s}(k) \tag{6}$$

と表わされる場合を考える．ここで，a_1，a_2は定数，$\mathbf{s}(k)$は音源信号ベクトルである．初期値$\hat{\mathbf{h}}(0)$を0ベクトルとする．部分空間$\boldsymbol{H}_x$を第2図に示す直線で表せば，フィルタ係数$\hat{\mathbf{h}}_1(k)$，$\hat{\mathbf{h}}_2(k)$は

$$\hat{\mathbf{h}}_{1a}(k)=\frac{a_1^2}{a_1^2+a_2^2}[\mathbf{h}_1(k)+\frac{a_2}{a_1}\mathbf{h}_2(k)]\neq\mathbf{h}_1(k) \tag{7}$$

$$\hat{\mathbf{h}}_{2a}(k)=\frac{a_2^2}{a_1^2+a_2^2}[\frac{a_1}{a_2}\mathbf{h}_1(k)+\mathbf{h}_2(k)]\neq\mathbf{h}_2(k) \tag{8}$$

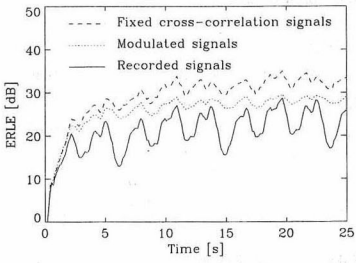


第2図　ステレオ信号の相互相関が時不変の定数である場合の収束点

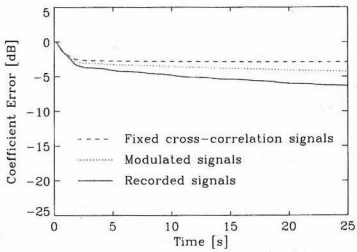
に収束する．

2.3　従来のステレオエコーキャンセラの性能

第3図（破線）に相互相関固定のステレオ入力信号に対するNLMSアルゴリズムを用いた従来のステレオエコーキャンセラの収束特性を示す．これらのステレオ信号は源音声$s(k)$に二つの異なるインパルス応答g_1，g_2を計算機上で畳み込んで作成した．



(a) ERLE



(b) Coefficient error

第3図　従来のステレオNLMSエコーキャンセラの音声入力に対する収束特性

解の誤推定が大きくても，エコー消去は可能である．相互相関固定のステレオ入力信号の場合には，エコー消去量（ERLE）は30[dB]にも達する（第3図(a)）．これに対して，第3図(b)に示すように，係数誤差は数dB程度しか収束しない．これが，解の誤推定の結果である．このような場合には，遠端で$g_1(k)$，$g_2(k)$が変化するとエコー消去量は低下する．

ノイズがなく，入力信号の相互相関が固定のときには，RLSアルゴリズムを含めていかなる適応アルゴリズムを駆使しても真のエコー経路を推定することはできない．

3.　真のエコー経路推定のための手がかり

解の誤推定の問題を解く手がかりは，実際の通信会議の状況の中にある．

第3図（実線）に実際の会議室において2本のマイクロホンで録音されたステレオ音声入力に対する，NLMSアルゴリズムを用いた従来のステレオエコーキャンセラの収束特性を示す．話者は一箇所に留まったままである．第3図(b)において録音された音声に対する係数誤差は少しずつ小さくなっている点に注意していただきたい．

解の誤推定の問題を解く手がかりは，3点ある．

3.1 無相関雑音

現実の世界では，ステレオ信号 $\mathbf{x}_1(k)$，$\mathbf{x}_2(k)$ には騒音や電気的なノイズが含まれる．これらのノイズは左右チャネル間で無相関であり，フィルタ係数を収束させる．

3.2 インパルス応答の打切り効果

遠端室内のインパルス応答 $g_1(k)$，$g_2(k)$ が適応フィルタ $\hat{\mathbf{h}}_1(k)$，$\hat{\mathbf{h}}_2(k)$ より長ければ，遠端室内のインパルス応答の末尾（打ち切られた成分）が無相関なノイズとして作用する．これらのノイズは左右チャネル間で無相関であり，フィルタ係数を収束させる [3]．

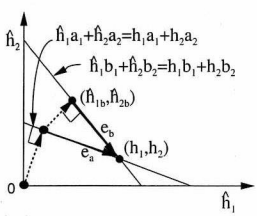
3.3 相互相関の変動

左右ステレオ信号 $\mathbf{x}_1(k)$，$\mathbf{x}_2(k)$ の相互相関は，たとえ会話中に話者が体や頭を動かさなくても，微妙に変動している．

相互相関が変動してもまた誤推定が起こるだけであり，何の解決にもならない，と思われるかもしれない．しかし，幸運にも古い推定結果を初期値として新しい推定が行われる．

左右ステレオ信号 $\mathbf{x}_1(k)$，$\mathbf{x}_2(k)$ の相互相関が変化した場合，たとえば，(6)式の a_1，a_2 が b_1，b_2 に変化した場合を考える（第4図）．最初，$[\hat{\mathbf{h}}_1^{\mathrm{T}}(k), \hat{\mathbf{h}}_2^{\mathrm{T}}(k)]$ は，(7)，(8)式に示すように $[\hat{\mathbf{h}}_{1a}^{\mathrm{T}}(k), \hat{\mathbf{h}}_{2a}^{\mathrm{T}}(k)]$ に収束する．つぎに，“初期値” $[\hat{\mathbf{h}}_{1a}^{\mathrm{T}}(k), \hat{\mathbf{h}}_{2a}^{\mathrm{T}}(k)]$ から最も近い $[\hat{\mathbf{h}}_{1b}^{\mathrm{T}}(k), \hat{\mathbf{h}}_{2b}^{\mathrm{T}}(k)]$ に収束する．その結果，フィルタ係数誤差ベクトルのノルムは $\|\mathbf{e}_b\| < \|\mathbf{e}_a\|$ となる．$\mathbf{h}(k)$ と $\hat{\mathbf{h}}(k)$ の間の係数誤差は，ステレオ信号の相互相関の変動のたびに小さくなるのである．

このようにして，数多くの相互相関の変動の後，ステレオエコーキャンセラは真のエコー経路に収束することができる [4]．

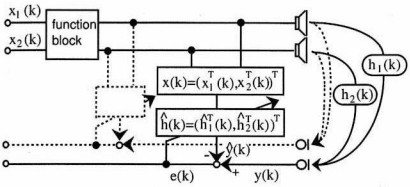


第4図 ステレオ信号の相互相関が変化した場合の収束点

4. 最近の解決法

4.1 無相関雑音と相互相関変動の付加

近年，解の誤推定を解決するためにいくつかの手法が提案されている．第5図中のファンクションブロックは，ステレオ信号に無相関なノイズや相互相関の変動を与えるためのものである [4]．これまでに，ステレオ信号に無相関な成分を作り出す手法 [5–8] や，左右ステレオ信号の相互相関に効果的な変動を付加する手法 [9–11] などが提案されている．これらの手法は，相互相関が固定のステレオ信号，たとえば，ミキサーで生成したステレオ信号などを入力とする場合などには必須である．付加されたノイズが聞こえないように，また，付加された相互相関の変動が音像の定位感を低下させないように注意する必要がある．



第5図 ファンクションブロックを有するステレオエコーキャンセラの構成

人間の聴覚の周波数特性は平坦ではない．また，音声によって歪み成分をマスキングすることができる．そこで，人間の聴覚システムの特徴である同時マスキングや継時マスキングを利用することも提案されている．同時マスキングでは，大きな周波数成分が隣接する小さな周波数成分をマスキングし，継時マスキングでは，大きな音が直前・直後の小さな音をマスキングする．したがって，付加ノイズや相互相関の変動を適切に調節することができれば，歪み成分を聞こえないように，また，ステレオエコーキャンセラの真のエコー経路への収束速度を格段に速めることができる．

4.2 非線形処理

非線形処理を適用して無相関なノイズを生成させる手法 [5] では，

$$\mathbf{x}'(k) = \mathbf{x}(k) + \alpha f[\mathbf{x}(k)] \tag{9}$$

の処理を行う．ここで，f は半波整流などの非線形関数である．この処理によって生成されたノイズはもとの音声と同期して増減するため，もとの音声によってマスキングされやすい．このノイズによって共分散行列の条件数（最大固有値と最小固有値の比）が減少し，係数誤差を小さく収束させることができる．正の定数 α の値を適切に調節すれば，通常の聴取条件で歪みは聞こえず，また，音像の定位に影響はない．筆者らもこの手法を追試して，主観評価実験によりフィルタ係数収束への効果と

音質への影響を確認した．

非線形処理された信号のみを別の適応フィルタに入力することにより，もとのステレオ信号の相互相関の影響を受けることなく真のエコー経路に収束させることもできる [6]．

4.3 ノイズシェイピング

人間の聴覚によるノイズマスキング効果はオーディオ符号化（MPEG オーディオ符号化）ですでに用いられている．

人間の聴覚の同時マスキング効果を取り入れて，ステレオ信号によってマスキングされるように周波数スペクトル形状を加工した無相関ノイズを加える手法が提案されている [7]．適切なマスキング効果を得るために，左右チャネルに付加されるそれぞれの付加ノイズのレベルは，ノイズが聞こえないように，また，音像の定位が変わらないように制御する．可聴しきい値以下に制御されたオーディオ符号化器における時間・周波数の量子化誤差を利用する手法も提案されている [8]．オーディオ符号化器が生成した量子化誤差がマスキングレベル以下であるならば，耳に聞こえない程度までさらにノイズを加えることもできる．

この手法により耳に聞こえないように周波数形状をシェイピングしたノイズ成分によって，真のエコー経路への収束が確保できる．

4.4 相関除去フィルタ

もし，左右ステレオ信号 $\mathbf{x}_1(k)$，$\mathbf{x}_2(k)$ が無相関になれば，真のエコー経路の推定が可能である．そのため，左右ステレオ信号間の相互相関を取り除く相関除去フィルタ

$$\mathbf{x}'_1(k) = \mathbf{x}_1(k) - f_2(k)\mathbf{x}_2(k) \tag{10}$$

$$\mathbf{x}'_2(k) = \mathbf{x}_2(k) - f_1(k)\mathbf{x}_1(k) \tag{11}$$

が試みられているが，過去の検討 [1,12] は失敗に終わっている．相関除去された信号 $\mathbf{x}'_1(k)$，$\mathbf{x}'_2(k)$ もまた同じ音源信号 $s(k)$ をフィルタリングしたものとなっているからである．したがって，ノイズがなく相互相関が固定であれば，完全な相互相関除去の結果 $\mathbf{x}'_1(k)=0$，$\mathbf{x}'_2(k)=0$ となる．このような受話入力信号を用いて，適応フィルタは安定な適応動作を行うことはできない（$f_1(k)$，$f_2(k)$ を計算するためには $g_1(k)$，$g_2(k)$ の逆システムを求める必要があるが，現実の世界では $g_1(k)$，$g_2(k)$ は安定な逆システムを持たない．そのため，現実の世界では完全な相互相関除去を行うことは非常に困難である）．

一方，もし無相関なノイズがあり相互相関が変動していれば，話は別である．(12)，(13) 式のように相関除去された信号 $\mathbf{x}'_1(k)$，$\mathbf{x}'_2(k)$ はそれぞれ $\mathbf{x}_2(k)$，$\mathbf{x}_1(k)$ に直交している．

$$\mathbf{x}'_1(k) = \mathbf{x}_1(k) - \frac{\mathbf{x}_2^{\mathrm{T}}(k)\mathbf{x}_1(k)}{\mathbf{x}_2^{\mathrm{T}}(k)\mathbf{x}_2(k)}\mathbf{x}_2(k) \tag{12}$$

$$\mathbf{x}'_2(k) = \mathbf{x}_2(k) - \frac{\mathbf{x}_1^{\mathrm{T}}(k)\mathbf{x}_2(k)}{\mathbf{x}_1^{\mathrm{T}}(k)\mathbf{x}_1(k)}\mathbf{x}_1(k) \tag{13}$$

相関除去フィルタは収束速度低下の原因となる音声の自己相関を取り除く．そのため，この手法はノイズや相互相関変動を強調することができ，真のエコー経路への収束が確保できる．しかしながら，修正ベクトル $\mathbf{x}'_1(k)$，$\mathbf{x}'_2(k)$ が互いに直交しているわけではなく，この直交化は左右を入れ替えたステレオ信号入力に対して，適応フィルタの誤差が小さいとみなせる場合に意味を持つため，収束の初期段階において不安定になる傾向がある（**5.4** 章に示すように，アルゴリズムに工夫を施すことにより安定動作を確保することができる）．

これに対して，**5.3** 章に示すように，モノラルの場合でもステレオの場合でも，自己相関除去フィルタを適用すれば，修正ベクトル $\mathbf{x}'(k)$ は入力信号ベクトル $\mathbf{x}(k-1)$ に常に直交し，不安定はない．

$$\mathbf{x}'(k) = \mathbf{x}(k) - \frac{\mathbf{x}^{\mathrm{T}}(k-1)\mathbf{x}(k)}{\mathbf{x}^{\mathrm{T}}(k-1)\mathbf{x}(k-1)}\mathbf{x}(k-1) \tag{14}$$

4.5 時変オールパスフィルタ

時変の一次オールパスフィルタを用いてステレオ信号の相互相関に変動を与える手法 [9] が提案されている．時変の量は空間的な知覚（音像の定位）に影響を与えないように，聴覚心理における両耳間時間差の検知限に基づいて制御されている．この手法によって生成される相互相関の変動によって適応フィルタの真のエコー経路への収束が確保できる．

4.6 時変 1 サンプル遅延フィルタ

2 タップの時変フィルタも提案されている [10,11]．このフィルタはステレオ信号の片方を周期的に 1 サンプル遅らせる．時変フィルタによって生じる歪みやクリック音は，パラメータを適切に選ぶことによって聞こえないように調節することができる．相互相関の変動によって，真のエコー経路の推定が可能となる．主観評価実験の結果，中央の聴取位置において音像の定位が最も影響を受けやすいこと，また，この位置においても評価者は処理されたステレオ信号の品質がオリジナルのステレオ信号と比べて許容できることを明らかにしている．筆者らもこの手法を追試して，主観評価によって効果を確認した．

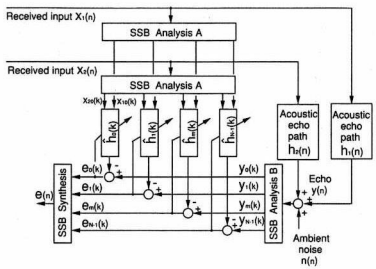
4.7 性能比較

付加ノイズは音声品質を劣化させ，相互相関変動付加は音像の定位をふらつかせる．また，収束速度と主観品質との間にはトレードオフの関係がある．

変動成分の周波数特性は $g_1(k)$，$g_2(k)$ が通過させるすべての周波数成分を含むように設計する必要があることも指摘されている [13]．

4.8 サブバンド処理

サブバンドステレオエコーキャンセラの構成を第6図に示す．サブバンドエコーキャンセラでは，信号はN個の周波数帯域に分割され，間引き率Rで間引かれる．その結果，フルバンドに比べて各サブバンドでサンプリング間隔が広がる．この操作により左右ステレオ信号の相互相関の変動が強調される．その結果，真のエコー経路への収束を速めることができる[14]．



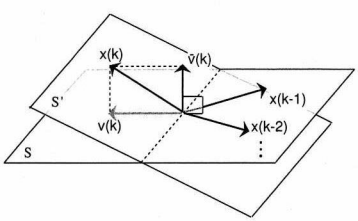
第6図 サブバンドステレオエコーキャンセラの構成

5. 適応アルゴリズム

5.1 無相関雑音と相互相関変動の強調

無相関なノイズや左右ステレオ信号間の相互相関の変動により，真のエコー経路を推定することができることを述べてきた．入力信号の共分散行列はフルランクとなったが，耳に聞こえないという制約のためにまだまだ条件数が大きい．真のエコー経路への収束は相対的に小さな項に頼っている．無相関なノイズや相互相関の変動はフィルタ係数が収束するのを促すが，強い相関成分は収束を阻害する．次のステップは，ノイズや変動を強調し，フィルタ係数が収束するのを加速する方法を見つけ出すことである．

もし相互相関が時刻$k-1$まで一定であれば，第7図に示すように，結合されたステレオ信号ベクトルは，相互相関によって決定される部分空間$\mathbf{x}(k-1)$, $\mathbf{x}(k-2)$,$\ldots \in \boldsymbol{S}$内に存在する．



第7図 相互相関によって決定される部分空間の幾何学的解釈

つぎに時刻kにおいて，相互相関が変化すれば，結合されたステレオ信号ベクトル$\mathbf{x}(k)$はそれまでの部分空間$\boldsymbol{S}$とは異なった部分空間$\mathbf{x}(k) \in \boldsymbol{S}'$内に存在する．そこで，$\mathbf{x}(k)$を二つの直交する成分

$$\mathbf{x}(k)=\mathbf{v}(k)+\bar{\mathbf{v}}(k) \quad [\mathbf{v}(k)\in \boldsymbol{S},\ \bar{\mathbf{v}}(k)\perp \boldsymbol{S}] \tag{15}$$

に分解する．ここで，$\bar{\mathbf{v}}(k)$は収束に有効な新しい成分であり，$\mathbf{v}(k)$は余分な成分である．無相関なノイズ成分は$\bar{\mathbf{v}}(k)$に含まれる．$\mathbf{x}(k)$から$\mathbf{v}(k)$を取り除くことにより，修正ベクトル$\Delta\hat{\mathbf{h}}(k)$の方向が$\bar{\mathbf{v}}(k)$と同じ方向になれば，$\hat{\mathbf{h}}(k)$の修正はむだなく行われる．無相関ノイズと相互相関変動の強調は$\bar{\mathbf{v}}(k)$を求めることにより達成され，フィルタ係数ベクトル$\hat{\mathbf{h}}(k)$は強調されたベクトル$\bar{\mathbf{v}}(k)$の方向に修正される．この議論は相互相関が時刻$k-1$まで少しずつ変動している場合にも成り立つ．この場合には，部分空間$\boldsymbol{S}$は少し広がる．しかし，ベクトル$\bar{\mathbf{v}}(k)$はやはり強調されたベクトルであり，少し広い部分空間$\boldsymbol{S}$に直交している．

$\bar{\mathbf{v}}(k)$は，$\mathbf{x}(k-1)$, $\mathbf{x}(k-2)$,...と相関がある$\mathbf{v}(k)$を，RLSアルゴリズムや射影アルゴリズムを用いて$\mathbf{x}(k)$から除去することによって得られる．RLSアルゴリズムは余分な成分を完全に取り除くが，演算量は非常に多い．これに対して，p次の射影アルゴリズムは，少ない演算量で直近のp個の余分な成分を取り除く[15–19]．

5.2 ステレオRLSアルゴリズム

ステレオRLSアルゴリズムは自己相関を完全に除去するが，演算量は非常に多い．

5.3 ステレオ射影アルゴリズム

ステレオ射影アルゴリズムは，ステレオエコーキャンセラに有効である[4,20]．このアルゴリズムは無相関ノイズやステレオ信号の相互相関変動を強調し，ステレオ信号の自己相関を取り除くことにより，真のエコー経路への収束を速める．

p次の射影アルゴリズムは，フィルタ係数ベクトル$\hat{\mathbf{h}}(k)$を次のように修正する．

$$\hat{\mathbf{h}}(k+1)=\hat{\mathbf{h}}(k)+\mu\Delta\hat{\mathbf{h}}(k) \tag{16}$$

$$\begin{aligned}\Delta\hat{\mathbf{h}}(k)=&\beta_1(k)\mathbf{x}(k)+\beta_2(k)\mathbf{x}(k-1)\\&+\cdots+\beta_p(k)\mathbf{x}(k-p+1)\end{aligned} \tag{17}$$

ここで，μはステップサイズ（$0<\mu<2$）であり，$\beta_1(k)$, $\beta_2(k)$, ..., $\beta_p(k)$は$\mu=1$のときに$\hat{\mathbf{h}}^{\mathrm{T}}(k+1)$が以下の式を満たすように決定される[21]．

$$\begin{aligned}\hat{\mathbf{h}}^{\mathrm{T}}(k+1)\mathbf{x}(k)&=y(k)\\ \hat{\mathbf{h}}^{\mathrm{T}}(k+1)\mathbf{x}(k-1)&=y(k-1)\\ &\vdots\\ \hat{\mathbf{h}}^{\mathrm{T}}(k+1)\mathbf{x}(k-p+1)&=y(k-p+1)\end{aligned} \tag{18}$$

その結果，$\mathbf{x}(k)$から相関成分$\mathbf{x}(k-1)$, $\mathbf{x}(k-2)$,...,$\mathbf{x}(k-p+1)$を取り除くことにより，$\Delta\hat{\mathbf{h}}(k)$は$\mathbf{x}(k)$のうちの

相関除去された成分になる．第7図において，$\bar{\mathbf{v}}(k)$ は $\mathbf{x}(k)$ から相関成分 $\mathbf{x}(k-1)$, $\mathbf{x}(k-2)$,…,$\mathbf{x}(k-p+1)$ を取り除いた成分になっている．すなわち，$\bar{\mathbf{v}}(k)$ は $\mathbf{x}(k-1)$, $\mathbf{x}(k-2)$,…,$\mathbf{x}(k-p+1)$ に直交している．次数 p を適切に選べば，$\mathbf{v}(k)$ のほとんどすべての相関成分が取り除かれ，修正ベクトル $\Delta\hat{\mathbf{h}}(k)$ の方向は $\bar{\mathbf{v}}(k)$ の方向と同じになる．次に説明する Extended LMS アルゴリズムと異なり，ステレオ射影アルゴリズムは相互相関除去を陽に試みていない．しかし，ステレオ信号ベクトル $\mathbf{x}(k)$ の自己相関除去は，音源信号の自己相関除去と同時に，左右ステレオ信号間の相互相関変動を強調する効果がある．ステレオ射影アルゴリズムでは収束過程が不安定になることはない．

5.4 Extended LMS アルゴリズムと Extended 射影アルゴリズム

Extended LMS アルゴリズム [3] と Extended 射影アルゴリズム [20] はアルゴリズムの内部でチャネル間の相互相関を除去する．**4.4** 章で述べたように，強い相互相関を持つ左右ステレオ信号の相互相関を取り除こうとすると，アルゴリズムが不安定になる傾向がある．そこで，Extended LMS アルゴリズムと Extended 射影アルゴリズムは，不安定性を制御するための定数を導入することにより，安定動作を確保している．

左右を入れ替えたステレオ信号入力に対して，適応フィルタの誤差が零（修正不要）という条件を付けることにより，Extended LMS アルゴリズムと Extended 射影アルゴリズムを導出することができる [22]．

6. 収束特性の一例

6.1 ステレオ射影アルゴリズムの収束特性

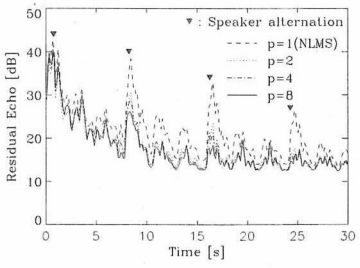
第8図に射影次数 $p=1$, 2, 4, 8 のステレオ射影アルゴリズムの収束特性を示す．$p=1$ は従来の NLMS アルゴリズムに相当する．

ステレオ入力信号は2人の話者が交互に話す音声であり，会議室で2本のマイクロホンにより録音した．2人の話者は部屋の異なる位置におり，体や頭は動かさない．適応フィルタのタップ数はそれぞれのフィルタ $\hat{\mathbf{h}}_1(k)$, $\hat{\mathbf{h}}_2(k)$ で512であり，サンプリング周波数は8 [kHz] である．真のエコー経路のインパルス応答は残響時間150 [ms] の室内で測定した．35 [dB] 一定の SN 比で周囲騒音を加えた．

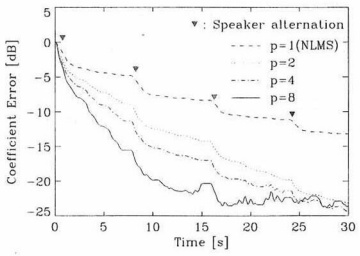
第8図 (a) から，射影次数 p を大きくするほど，話者交代にともなう残留エコーの増加を防止することができることがわかる．また，第8図 (b) から，射影次数 p を大きくするほど，係数誤差がより速く収束することがわかる．

6.2 サブバンド処理の効果

第9図に次数 $p=1$ (NLMS) のサブバンドエコーキャンセラの係数誤差の収束特性を示す．間引き率は $R=N/4$ とした．収束速度は帯域分割数に応じて格段に速くなっている [14]．

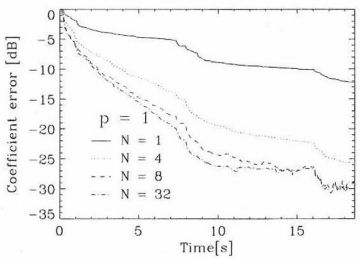


(a) Residual echo



(b) Coefficient error

第8図　ステレオ射影アルゴリズムの音声入力に対する収束特性．$p=1\sim8$



第9図　サブバンドステレオ NLMS アルゴリズムの音声入力に対する収束特性．分割数 $N=1\sim32$．係数誤差

第8図 (b) と第9図の傾向はよく似ている．サブバンド構成には，以下に説明するように，ステレオ射影アルゴリズムと同様の効果がある．

p 次の射影アルゴリズム ($p<L$) は，(18) 式を満足するようにフィルタ係数 $\hat{\mathbf{h}}(k+1)$ を逐次修正する．(18) 式は，入力信号 $\mathbf{x}(k-i+1)$ に対して正しい出力 $y(k-i+1)$ を出力するようなフィルタ係数 $\hat{\mathbf{h}}(k+1)$ を求めることを示している．これに対して，間引き率 R のサブバンド処

理の場合には，$R-1$ サンプルは冗長であり，(18) 式は $i=1, 2, \cdots, R$ に対して自動的に満たされる．そのため，間引き率 R，フィルタ長 L/R，修正回数 $1/R$ のサブバンド処理は，次数 R，オリジナルサンプリング，フィルタ長 L，修正回数 $1/R$ の射影アルゴリズムとほぼ等価となる．

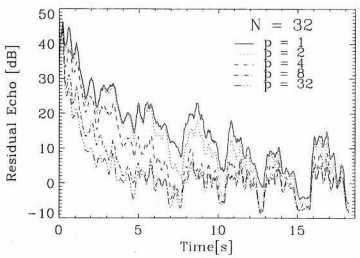
6.3 サブバンドステレオ射影アルゴリズム

サブバンド構成を用いさらにステレオ射影アルゴリズムを用いることにより，収束速度をさらに速めることができる [14]．

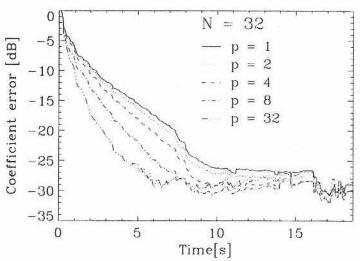
第 10 図 (a) に $N=32$ サブバンドの場合の残留誤差の収束特性を示す．射影次数を大きくするに従い，収束速度が速くなることがわかる．$p=32$ 次の残留誤差は遠端室内における話者交代の影響を受けていないことがわかる．

第 10 図 (b) に $N=32$ サブバンドの場合の係数誤差の収束特性を示す．射影次数を大きくするに従い，収束速度が速くなることがわかる．$p=32$ 次の収束速度は $p=1$ 次（NLMS）の 2 倍以上である．

各サブバンドで必要なタップ数は間引きによって $1/R$ に減る．また，射影アルゴリズムはタップ数の少ない適応フィルタを少ない射影次数で相関除去できる [23]．その



(a) Residual echo



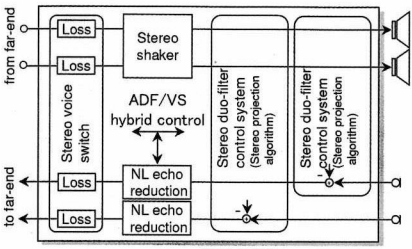
(b) Coefficient error

第 10 図 $N=32$ サブバンドステレオ射影アルゴリズムの音声入力に対する収束特性．射影次数 $p=1\sim32$

ため，フルバンドで演算量の多い高次の射影アルゴリズムを使う代わりに，サブバンドで低次の射影アルゴリズムを使うことができ，演算量を低く抑えることができる．

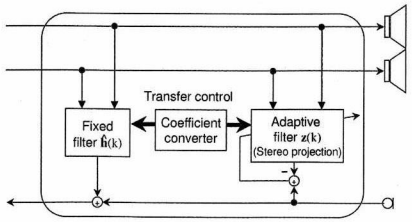
7. インプリメンテーションの一例

第 11 図にステレオエコーキャンセラのインプリメンテーションの一例を示す．DSP を用いて 100 [Hz]–20 [kHz] の周波数帯域を実現している．適応フィルタのタップ数は 0.1–4 [kHz] 帯域で 1200 タップ，4–8 [kHz] 帯域で 800 タップである．8–20 [kHz] 帯域は，ステレオ音声スイッチのみを用いている [24]．



第 11 図 ステレオエコーキャンセラのインプリメンテーションの一例

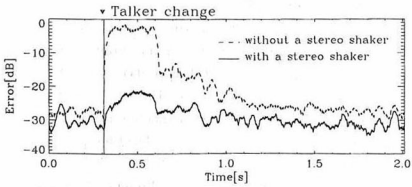
真のエコー経路推定を実現するために，八つの周波数帯域にステレオシェイカー（ファンクションブロック）を導入している．これらは，二つの会議室を結んだ双方向通信会議を行い，耳に聞こえず，かつ音像の定位に影響がないように主観評価実験により調節した．ダブルトーク制御には，適応フィルタと半固定フィルタの二つのフィルタを用いるデュオフィルタ制御方式 [25] を用いている（第 12 図）．適応フィルタには 2 次のステレオ射影アルゴリズムを用い，ステレオ音声スイッチも併用している．



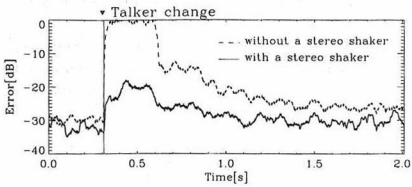
第 12 図 ステレオデュオフィルタ制御方式

7.1 遠端話者変動時の収束特性

DSP で構成したハードウェアを用いて，二つの会議室を接続して，実時間評価実験を行った．それぞれの部屋は容積約 150 [m^3]，残響時間約 300 [ms] である．周囲騒



(a) right channel S_{out}



(b) left channel S_{out}

第 13 図　(a) 右チャネル S_{out} および (b) 左チャネル S_{out} 誤差の収束特性

音の SN 比は約 25 [dB] であった．射影アルゴリズムのステップサイズ μ は 0.5 に設定した．

第 13 図に (a) 右チャネル S_{out} および (b) 左チャネル S_{out} の誤差を示す．音声スイッチは動作しておらず，エコーキャンセラだけが動作している．遠端からのステレオ信号は相互相関固定の白色雑音である．遠端の右位置の話者が 90 秒話した後，右位置から左位置へ移動した．ステレオシェイカーがないときには，エコー経路は正しく推定されていないため，誤差は 20 [dB] も増加した．これに対して，ステレオシェイカーを使えば，エコー経路が正しく推定されているため，遠端の話者位置が変化しても誤差の増加は少ない．誤差の増加分が 20 [dB] ということは，エコーキャンセラの係数誤差も約 20 [dB] であることを示唆している．

このステレオエコーキャンセラは通信会議システムの中で日常的に使われており，従来のシステムに比べてスピーカ音量を 10 [dB] 程度増加させることができた．ハウリングとエコーがなくなり，音声品質が向上した．2.4 [m] (W)×1.3 [m] (H) の大型スクリーンと組み合わせることで，音像の定位により通信会議に臨場感をもたらした．これまでに 500 人以上に使われている [24]．

8. あとがき

ステレオエコーキャンセラの本質的な課題について議論し，最近の解決策について解説した．1 年以上にわたってハードウェアを日常的に使った経験からいえば，現実の世界においては無相関なノイズや相互相関の変動は非常に多く存在する（ときには，ノイズ除去が必要なほどノイズが多すぎることさえある．現実の世界のノイズや変動の方が，人工的に付加したものよりも大きい場合も多い）．遠端室内の空間的に離れた位置に二つ以上の無相関な音源があれば，解の誤推定の問題は本質的に消滅してしまう．ステレオエコーキャンセラにおける次のステップは，ダブルトーク，音声スイッチ，エコーリダクションを総合した制御手法の開拓である．

謝　辞

ご討論いただいた金田豊博士，羽田陽一博士，田中雅史氏に感謝する．また，細部にわたってディスカッションいただいた Jacob Benesty 博士に感謝する．

(2002 年 7 月 9 日受付)

参 考 文 献

[1] M. Sondhi, D. Morgan and J. Hall: Stereophonic acoustic echo cancellation — an overview of the fundamental problem; *IEEE Signal Processing Letters*, Vol. 2, No. 8, pp. 148–151 (1995)

[2] T. Gansler and J. Benesty: Multichannel acoustic echo cancellation: what's new?; *Proc. IWAENC2001*, pp. 14–18 (2001)

[3] J. Benesty, F. Amand, A. Gilloire and Y. Grenier: Adaptive filtering algorithm for stereophonic acoustic echo cancellation; *Proc. ICASSP95*, pp. 3099–3102 (1995)

[4] S. Shimauchi and S. Makino: Stereo projection echo canceller with true echo path estimation; *Proc. ICASSP95*, pp. 3059–3062 (1995)

[5] J. Benesty, D. Morgan and M. Sondhi: A better understanding and an improved solution to the problems of stereophonic acoustic echo cancellation; *Proc. ICASSP97*, pp. 303–306 (1997)

[6] S. Shimauchi, Y. Haneda, S. Makino and Y. Kaneda: New configuration for a stereo echo canceller with nonlinear pre-processing; *Proc. ICASSP98*, pp. 3685–3688 (1998)

[7] A. Gilloire and V. Turbin: Using auditory properties to improve the behaviour of stereophonic acoustic echo cancellers; *Proc. ICASSP98*, pp. 3681–3684 (1998)

[8] T. Gansler and P. Eneroth: Influence of audio coding on stereophonic acoustic echo cancellation; *Proc. ICASSP98*, pp. 3649–3652 (1998)

[9] M. Ali: Stereophonic acoustic echo cancellation system using time-varying all-pass filtering for signal decorrelation; *Proc. ICASSP98*, pp. 3689–3692 (1998)

[10] Y. Joncour and A. Sugiyama: A Stereo echo canceller with pre-processing for correct echo-path identification; *Proc. ICASSP98*, pp. 3677–3680 (1998)

[11] A. Sugiyama, Y. Joncour and A. Hirano: A Stereo echo canceller with correct echo-path identification

based on an input-sliding technique; *IEEE Trans. Signal Processing*, Vol. 49, No. 11, pp. 2577–2587 (2001)

[12] Y. Mahieux, A. Gilloire and F. Khalil: Annulation D'echo en teleconference stereophonique; *Proc. Quatorzieme Colloque GRETSI*, pp. 515–518 (1993) (in French)

[13] A. Hirano, K. Nakayama and K. Watanabe: Convergence analysis of stereophonic echo canceller with pre-processing — relation between pre-processing and convergence —; *Proc. ICASSP99*, pp. 861–864 (1999)

[14] S. Makino, K. Strauss, S. Shimauchi, Y. Haneda and A. Nakagawa: Subband stereo echo canceller using the projection algorithm with fast convergence to the true echo path; *Proc. ICASSP97*, pp. 299–302 (1997)

[15] K. Ozeki and T. Umeda: An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties; *Trans. IEICE Japan*, Vol. J67-A, pp. 126–132 (1984) (in Japanese)

[16] M. Tanaka, Y. Kaneda, S. Makino and J. Kojima: Fast projection algorithm and its step size control; *Proc. ICASSP95*, pp. 945–948 (1995)

[17] S. Gay and S. Tavathia: The fast affine projection algorithm; *Proc. ICASSP95*, pp. 3023–3026 (1995)

[18] M. Tanaka, Y. Kaneda, S. Makino and J. Kojima: A fast projection algorithm for adaptive filtering; *Trans. IEICE Japan*, Vol. E78-A, No. 10, pp. 1355–1361 (1995)

[19] F. Amand, J. Benesty, A. Gilloire and Y. Grenier: A fast two-channel projection algorithm for stereophonic acoustic echo cancellation; *Proc. ICASSP96*, pp. 949–952 (1996)

[20] J. Benesty, P. Duhamel and Y. Grenier: A multichannel affine projection algorithm with applications to multichannel acoustic echo cancellation; *IEEE Signal Processing Letters*, Vol. 3, No. 2, pp. 35–37 (1996)

[21] S. Makino and Y. Kaneda: Exponentially weighted stepsize projection algorithm for acoustic echo cancellers; *Trans. IEICE Japan*, Vol. E75-A, No. 11, pp. 1500–1508 (1992)

[22] S. Shimauchi and S. Makino: Stereo echo cancellation algorithm using imaginary input-output relationships; *Proc. ICASSP96*, pp. 941–944 (1996)

[23] S. Makino, J. Noebauer, Y. Haneda and A. Nakagawa: SSB subband echo canceller using low-order projection algorithm; *Proc. ICASSP96*, pp. 945–948 (1996)

[24] S. Shimauchi, S. Makino, Y. Haneda, A. Nakagawa and S. Sakauchi: A stereo echo canceller implemented using a stereo shaker and a duo-filter control system; *Proc. ICASSP99*, pp. 857–860 (1999)

[25] Y. Haneda, S. Makino, J. Kojima and S. Shimauchi: Implementation and evaluation of an acoustic echo canceller using duo-filter control system; *Proc. EUSIPCO96*, pp. 1115–1118 (1996)

著者略歴

牧野 昭二

1956年6月4日生．1981年3月東北大学大学院工学研究科機械工学専攻修士課程修了．同年4月日本電信電話公社（現NTT）入社．2000年4月NTTコミュニケーション科学基礎研究所 知能情報研究部 信号処理研究グループリーダとなり現在に至る．音響エコーキャンセラ，ブラインド音源分離などの音響信号処理の研究に従事．IEEE Senior Member．日本音響学会，電子情報通信学会などの会員．

島内 末廣

1967年5月16日生．1993年3月東京工業大学大学院総合理工学研究科物理情報工学専攻博士前期課程修了．同年4月日本電信電話（株）（NTT）入社，現在に至る．音響エコーキャンセラ，適応フィルタなどの音響信号処理の研究開発に従事．IEEE，日本音響学会，電子情報通信学会などの会員．