

複数反響路を有する音響エコーキャンセラの構成法

Acoustic Echo Canceller with Multiple Echo Paths

小泉宣夫 牧野昭二 及川弘一
 Nobuo KOIZUMI Shoji MAKINO Hiroshi OIKAWA

NTTヒューマンインタフェース研究所（*現NTT通信機器事業部）
 NTT Human Interface Laboratories

1. まえがき

会議室において、多くの臨席者のもとで通信会議を行なう場合、複数のマイクロホンを用いた拡声通話方式が用いられる。この拡声通話系において生じるハウリングとエコーの抑圧には従来音声スイッチが用いられてきたが、最近では話頭切断のないエコーキャンセラ方式がデジタル信号処理技術の発達とともに採用されつつある。しかし、従来のエコーキャンセラを用いた方式では1つの帰還路にしかうまく対応できず、マイクロホンが1本かあるいは複数本を直接混合した系にのみ用いられてきた。本稿では、ゲインの切り替わる複数のマイクロホンを用いた系においても安定に動作する帰還補償系の構成法について述べる。

2. 複数マイクロホン系の構成

音声会議におけるマイクロホン系の構成形態を図1に示す。1つの側に複数の列席者がいて音声会議を行なう場合、なるべくマイクロホンを発言者の近傍において直接音の成分を多く含んだほうが品質が良い。したがって、マイクロホンは列席者の数だけ、あるいは少なくとも2人に1本位を目安に設置することが望ましい。今これらの複数のマイクロホンの出力を直接加算して送話入力信号を合成すると、室内の背景雑音が加算され音声信号のSN比の劣化、残響成分の加算による歪みによる劣化が生じる。また背景雑音と同じようにまわり込み信号レベルも増加し、音響結合が増え、系が不安定となる。マイクロホンの本数をN本とするとS/Nの劣化量あるいは音響結合の増加量は $10 \log_{10} N$ (dB)となる。また、残響時間が1秒の部屋では、6本加算接続すること

で明瞭度の客観評価尺度であるSTI値は0.8程度低下する[1]。一方、マイクロホンの使用形態としては、通常発言者は一人であり、時々複数の発言が割り込むかたちが多い。

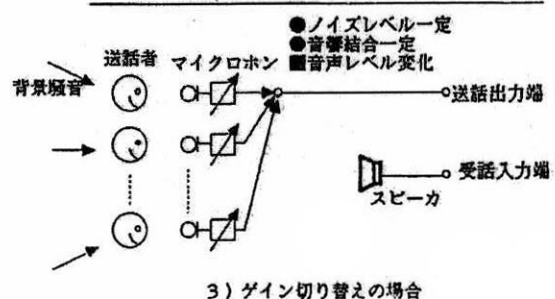
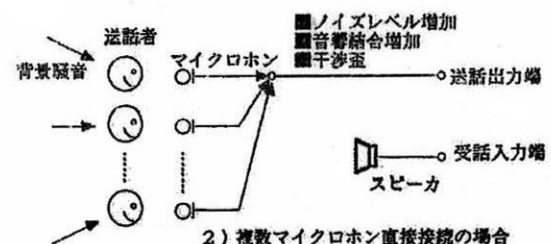
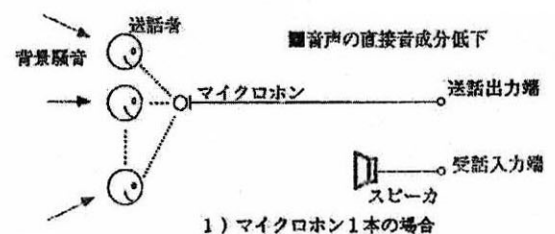


図1 マイクロホン系の使用形態

このような観点から、発言者に対応するマイクロホンあるいはその近傍のマイクロホンのみ感度を上げ、その他は感度を下げようとするゲイン切換えを動的に行なうことが望ましい。この場合、背景雑音あるいは音響結合量が一定になるように切り替え、かつ切り替え時に不自然な音声の切断を生じないようにゲインを設定する必要がある[2]。このようにすれば、一人の送話信号のレベルは複数が同時に発言する場合に下がることになるが、そのような状態は長くは続かないので実際上問題にはならない。今感度を上げた状態をオン、下げた状態をオフと呼ぶことにし、オフマイクロホンのゲインを1に正規化し、オンマイクロホンのゲインを a'_k で表すことにする。

ただし k は同時にオンとなるマイクロホンの個数であり、その個数に応じて一定の値に設定されるものとする。それぞれのマイクロホンに印加される信号は無相関であると仮定すると、背景雑音レベルあるいは搬送信号レベルすなわち音響結合が一定となる1本がオンの場合と k 本がオンの場合のゲインの関係は

$$a'^2_{j2+(N-1)} = k a'^2_{k2+(N-k)}$$

あるいは

$$a'_k = \sqrt{\frac{a'^2_{j2+(N-1)} + k - 1}{k}} \quad (1)$$

で表せられる。この条件をゲイン切り替え時の理想的な配分条件と考えることにする。

3. エコーキャンセラの構成

スピーカから放射される受信信号の内、動作中のすべてのマイクロホンにまわり込んで再び送話信号となる成分を除去するために、エコーキャンセラは、スピーカから個々のマイクロホンに至る音響経路と入力の総和に至る部分からなる反響路に対して、疑似反響路フィルターを構成することになる。しかし、疑似反響路フィルターに用いるタップ係数を複数組記憶する機能のない従来の構成では、マイクロホンゲインの切換えによる反響路の変化に対してはすぐに対応できず、疑似反響路の修正は同定誤差に基づく適応動作に頼ることになり、回復には時間を要する。そこで、ここではそれぞれのマイクロホンに応じた複数の反響路のメモリを有し、必要に応じてこのメモリからの呼び出しにより全体の反響路を合成することにより、マイクロホンのゲイン切り替えに瞬時に適応できる構成を提案する。構成の概要を図2に示す。記憶する反響路の個数は接続されたマイクロホンの個数に等しいものとするが、記憶されるインパルスレスポンスは必ずしも1つのマイクロホンに至る反響路そのものとは限らない。

この構成では、オフマイクロホンのゲインが完全に0の場合は、任意のゲイン設定に対して疑似反響路を個々のマイクロホンに至る反響路の線形和によって合成することができる。しかし、実際にオフマイクロホンについてもゲインを持つ場合には、ゲイン設定に対して拘束条件を設けないと全体の反響路を線形加算によって得ることはできない。その条件を次に求めることにする。

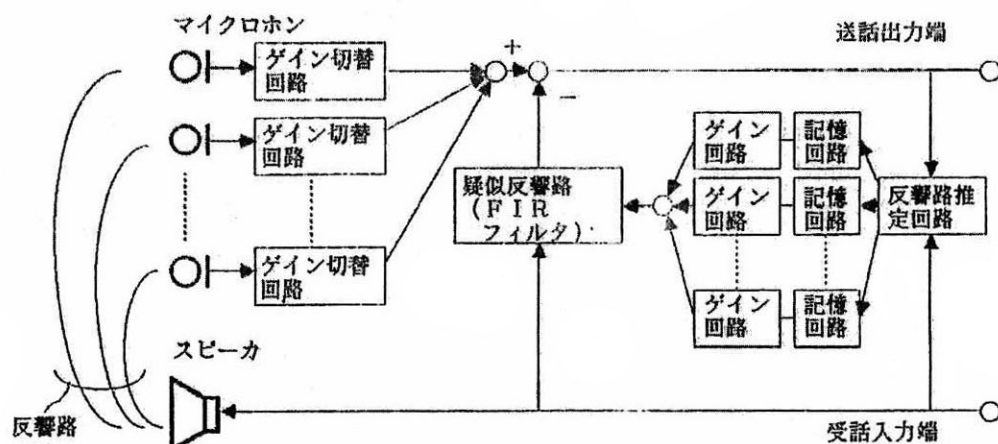


図2 複数反響路エコーキャンセラの構成

今、接続されたマイクロホンの総数を N 本とし、スピーカから l 番目のマイクロホンに至る音響経路のインパルスレスポンスを H_l とする。使用状態ではオフマイクロホンも生きているので、この H_l は動作中に観測あるいは同定することはできない。さて、 k 本がオン状態のときの全体の反響路のインパルスレスポンス H は

$$H = (a_k - 1) \sum_{j: \text{on}} H_j + \sum_{i=1}^N H_i \quad (2)$$

で与えられる。ただし、右辺第1項は k 本のオン状態についての和を表すものとする。一方、 l 番目のマイクロホンが1本だけオン状態にあるときの、全体の反響路のインパルスレスポンス H^* は次式で与えられる。

$$H^*_j = (a_1 - 1) H_j + \sum_{i=1}^N H_i \quad (3)$$

オン状態のすべてに関する総和をとると、

$$\sum_{j: \text{on}} H^*_j = (a_1 - 1) \sum_{j: \text{on}} H_j + k \sum_{i=1}^N H_i \quad (4)$$

が得られる。したがって、 k 本がオンのときのゲイン a_k と1本がオンのときのゲイン a_1 とが

$$a_k - 1 = \frac{a_1 - 1}{k} \quad (5)$$

なる関係にあれば、全体の反響路のインパルスレスポンスは次式のように1本がオンのときの線形和によって表すことができる。

$$H = \frac{1}{k} \sum_{j: \text{on}} H^*_j \quad (6)$$

1本がオンのときのインパルスレスポンス H^*_j は動作中に観測あるいは同定することが容易であり、2本以上がオンの場合はこの式により即座に疑似反響路を合成することができる。

式(5)の条件は帰還量一定の条件である式(1)とは一致しない。図3に式(1)と式(5)との比較を示す。1つがオンのときのゲイン a_1 を2(6dB)および3(9.5dB)とした場合の2本以上のときの値を示してあるが、 a_k と a'_k の差は1dB以内にあるので、式(5)を本構成におけるゲインの設定条件として用いても、音響結合はほぼ一定に保たれるものと考えられる。

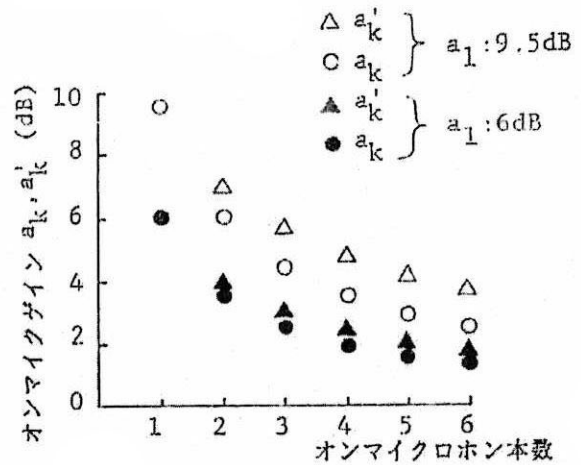


図3 オンマイクロホンゲインの設定

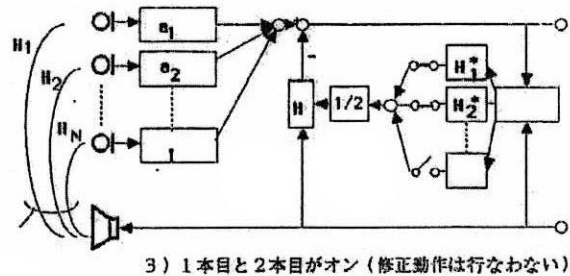
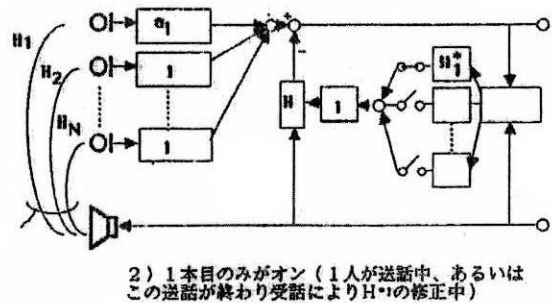
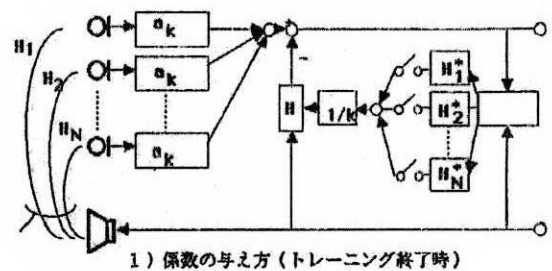


図4 動作概念

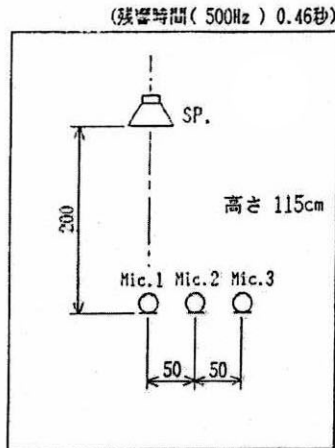


図5 音響系の配置

本構成における動作概念を図4に示す。まず、動作開始に当たっては、1本をオン状態になるように順に切り替え、トレーニング信号により、 H_i ($i=1, N$) の初期値を同定し、メモリに格納する。さて会議が始まると、発言者が1人のとき、すなわち1本がオンでは格納した内のどれかが疑似反響路に呼び出され、2人以上の発言で複数がオンになれば、式(6)に従って疑似反響路が合成される。次に、発言が中断し、受話状態になると誤差信号に基づく反響路の修正が始まる。この場合、最後まで送話のあったマイクロホンのオン状態を保持するかたちが自然である。送話状態が1本がオンの状態で終了し、オン状態の頻度の高いマイクロホンと受話状態になる寸前のオンマイクロホンとは対応すると仮定すれ

ば、この方法で反響路の使用頻度に応じた修正を実行することができる。

本提案によるハードウェアは反響路のメモリの増加程度ですみ、演算手順も簡単であり、容易に実現可能な構成であるといえる。

4. シミュレーション

実際の部屋でのインパルスレスポンスの測定値を用いて、時間特性のシミュレーションを行なった。図5に示すように、約0.5秒の部屋の中に3本のマイクロホンを50cm間隔に配列し、それぞれのマイクロホンに至るインパルスレスポンスを測定した。これのタップ長は4000、8kHzサンプリングで遅延時間長は0.5秒である。

このデータをもとに、従来の1つの反響路でエコーキャンセラを構成したときのマイク切換えによるエコー消去量の変化を図6に示す。切り替えは等間隔とし、前半は1本の切り替え、後半は2本オンの組み合わせによる切り替えを示した。横軸はサンプリングの繰り返し数、縦軸はエコー消去量： $10 \log$ (まわり込み信号のパワー/誤差信号のパワー) (dB) を表す。適応動作には学習同定法を適用し、信号には白色雑音を用いた。図から明らかなように、マイクロホンの位置が替っただけで消去量は大きく落ち込むことが分かる。この場合、性能は適応修正アルゴリズムに大きく依存するが、現状の実時間処理技術では学習同定法程度のものが限界であり、実際の音声信号を用いれば、収束はさらに遅くなり、回復に時間を要することになる。

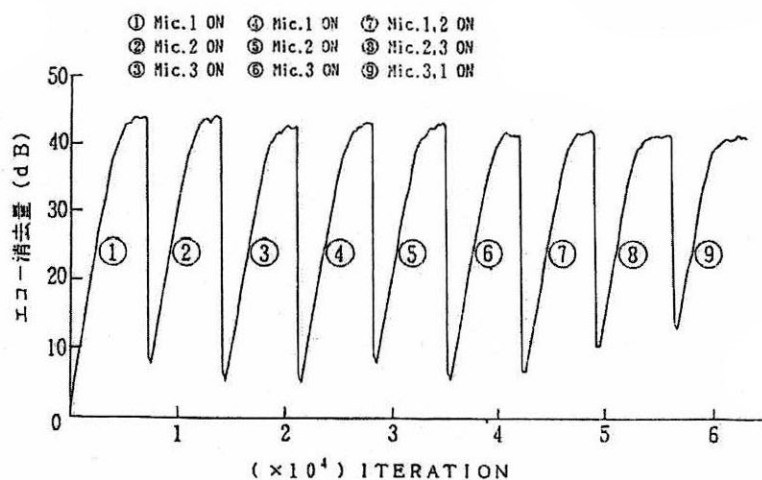


図6 タップ係数を置き換えない場合の消去特性の変化

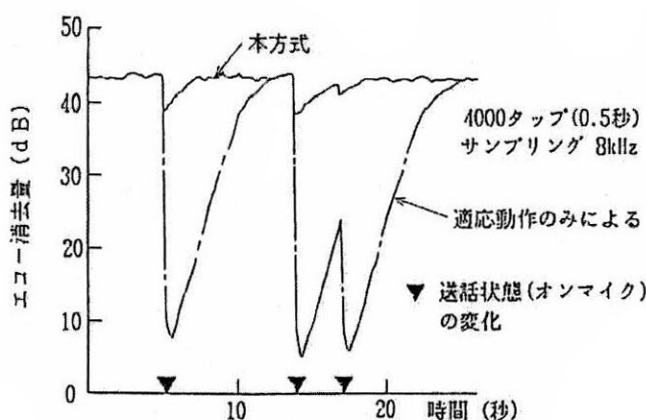


図7 送話状態変化時のエコー消去特性

一方、提案した構成を用いればこれらの低下は回避され、わずかに音場の変動分だけの収束を待てばよいことになる。この特徴を示したのが図7である。ここでは使用条件として帯域4kHz(8kHzサンプリング)を仮定し、横軸を時間で表した。実際には発言者すなわちオンマイクロホンが替った後は適応動作を禁止し、受話状態になってから適応が再開されるが、図では送話直後に適応動作を開始する形で示してある。頻繁に送話状態が替る系では、適応動作だけでは十分な消去性能を回復できず、本手法が有効であることがわかる。

5. 考察

本構成の過渡的特性についてシミュレーションにより示したが、次に、定常的な性能に対してどのような効果があるかを考察する。実際の室内における消去性能は、残響時間に対して十分な遅延処理長があるとすれば、部屋の背景騒音に依存する。このことを、実際にDSPチップを用いて試作したハードウェアによる測定結果[3]である図8、図9を用いて説明する。この装置の帯域は4kHz、遅延処理時間は144msecである。また、1つの反響路の記憶回路のみを持つ従来構成であるので、1本のマイクロホンによる結果を示している。測定に用いた部屋の残響時間は遅延処理時間内の0.1秒である。図8は受話入力レベルに対するまわり込み信号の送話入力とキャンセラを経由した後の送話出力のレベルを、背景騒音のレベルを変えて測定した例である。送話入力は、当然ながら受話入力に比例する。しかし、誤差信号に相当する送話出力は受話入力によらず一定である。図9はこの場合の送話入力から送話出力を引いた消去量を示しているが、消去量は受話

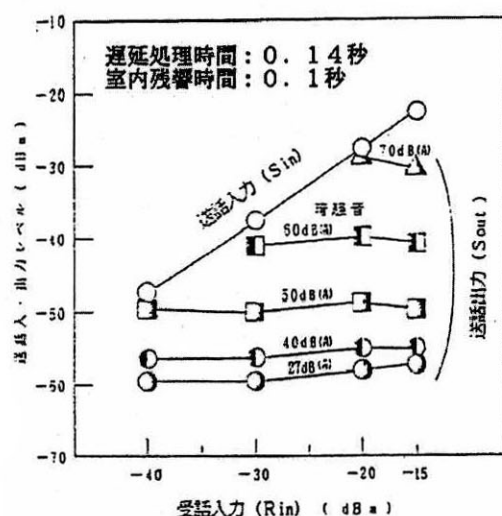


図8 背景騒音レベルを変化させたときの受話入力レベルと送話入/出力レベル

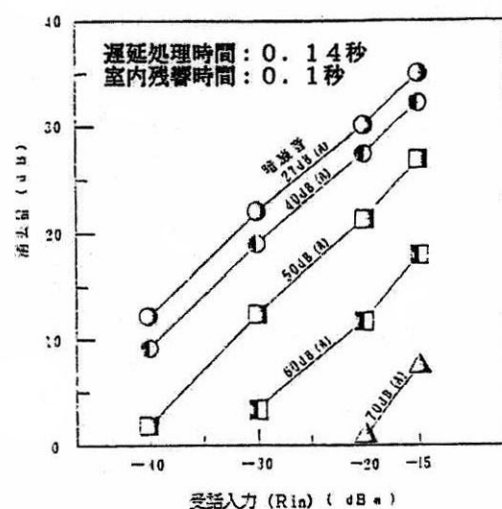


図9 背景騒音レベルを変化させたときの受話入力レベルと消去量

入力に依存しており、まわり込みの送話入力と、背景騒音のレベル比で決まる。以上のことから推察した、従来の複数のマイクロホン出力を直接混合した場合と、本切り替え方式の場合との比較を図10に示す。図10でのレベルの絶対値は前述のハードウェアに準拠して示してある。また、比較の例として、背景騒音のレベルは40dB、接続本数は6本、1つがオンのときのゲイン差を3(9.5dB)とした。この場合、本方式を用いるとまわり込み信号の加算後のレベルは約6dB下げられる。しかし、背景騒音

の加算後のレベルも同じだけ下がるので、消去性能は低下せず、受話入力レベルに対する消去量は直接接続の場合に等しい。したがって、結果としてまわり込みレベルの低下分だけまわり込みの送話出力、すなわち残留エコーレベルを低減させることができる。また、この低減量はマイクロホンの接続本数とオン/オフのゲイン差によって決まることがわかる。

5. あとがき

複数のマイクロホンがあり、ゲインが切り替わることにより複数の反響路が存在する場合のエコーキャンセラの構成法を提案した。この構成ではマイクロホンの本数分の反響路の記憶回路が必要であり、ゲインの設定はほぼ帰還量一定の条件を満たすものである。本手法の特徴は以下のように要約される。

(1) 適応アルゴリズムだけにより修正する方法に比べ、発言者の交替により送話マイクロホンが切り替わっても、消去量の劣化は経時変動分だけですむので、正常な特性を保持できる。

(2) 定常特性としては、複数マイクロホン直接接続時に比べ、音響結合量を低減できるので、残留エコー成分を低減させることができる。

本手法は従来の構成と比べて、記憶回路の増大だけですむので、ハードウェア的にも十分実現できるものである。

謝辞

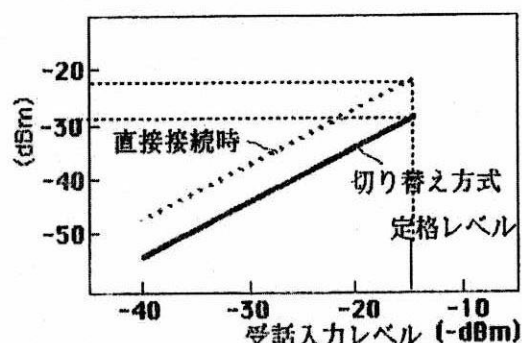
日頃ご指導いただく杉山精部長ならびに川嶋功部長に感謝いたします。

文献

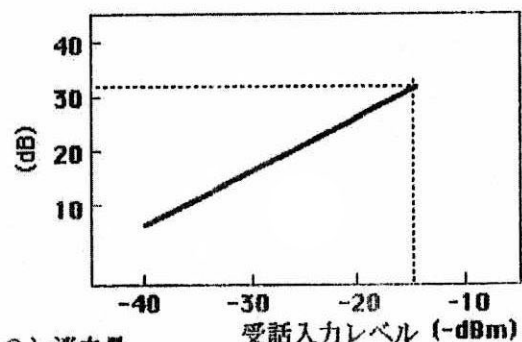
[1]小泉、牧野、青木：MTFによる音声会議でのマイクロホン配置の評価について、音響学会講義集、p631, 昭61-10.

[2]西野、飛田、野村：音声会議における複数マイクロホン制御系の構成法、信学会情報システム部大1-301, 1987.

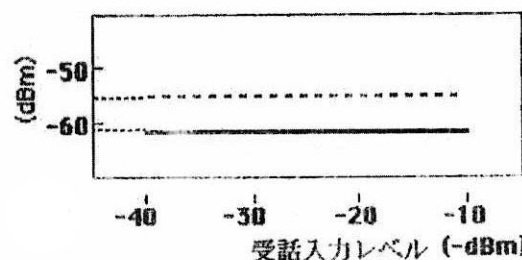
[3]牧野、小泉：音響エコーキャンセラの室内環境における消去特性について、信学技報E A 87-43, 1987.



1) まわり込みの送話入力レベル



2) 消去量



3) まわり込みの送話出力(残留エコー)レベル

図10 定常消去特性の改善