

## 通信会議における遠方音声収録のための残響抑圧方式の研究

古家, 賢一

<https://doi.org/10.15017/459055>

---

出版情報 : Kyushu University, 2005, 博士 (芸術工学), 論文博士  
バージョン :  
権利関係 :

## 第4章

# 入力相関行列と音声の白色化フィルタを用いたSemi-blind残響抑圧

### 4.1 まえがき

第3章において、入力相関行列を用いた二入力残響抑圧法を提案し、その有効性について報告した。低次の非最小位相伝達系の場合にはほぼ完璧に逆フィルタが計算でき残響回復できることが確認された。しかし、実際の数千次におよぶ非最小位相伝達系に対しては、十分な効果が得られなかった。その原因は、最小固有ベクトルを計算しているため雑音や計算誤差に弱く精度よく計算することが困難であること、帯域分割した場合、各帯域に不定乗数が残る、その決定が困難であることなどが挙げられる。

本章では、二入力残響抑圧法のように入力相関行列の最小固有ベクトルを用いインパルス応答を推定し逆フィルタを計算するのではなく、直接、入力相関行列から方程式を立て、逆フィルタを計算するSemi-blind残響抑圧法を提案する。この方法は、音源に最も近いマイクロホンのみを既知として、MINT法の多チャンネル逆フィルタを入力相関行列を用いて近似的に計算する。完全なBlind処理ではなく、音源に最も近いマイクロホンがどれであるかという事前情報を利用しており、この意味でSemi-blind処理となっているが、事前のインパルス応答測定が不要であり、実用上MINT法などに比べ大幅に使用条件を緩和できる。

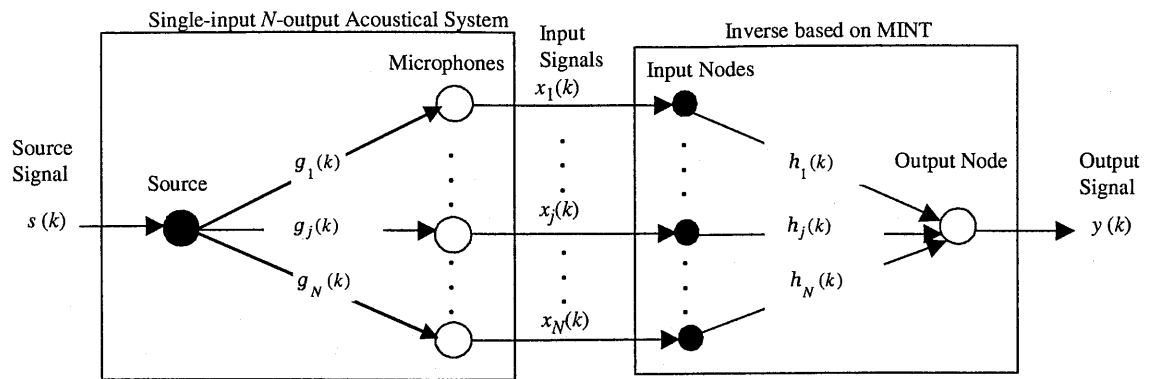


図 4.1: 1 入力  $N$  出力音響システムに対する MINT 法による残響抑圧

第 4.2 節 次節では Semi-blind 残響抑圧法の原理および音声の平均スペクトルを平坦化する白色フィルタについて述べる. 第 4.3 節で実験を通じて, 本手法での仮定により実際の音場でどの程度の残響抑圧効果があるのか検討を行い, 第 4.4 節でそれらの結果をまとめる.

## 4.2 Semi-blind 残響抑圧法

本節では, MINT 法の逆フィルタ  $\mathbf{H}$  を室内インパルス応答行列  $\mathbf{G}$  を用いることなく近似的に計算する Semi-blind 残響抑圧法の原理について述べる.

図 4.1 に示す 1 入力  $N$  出力の音響システムを考える. ここで  $N$  はマイクロホン数で  $N > 1$  である.  $s(k)$  は音源信号,  $x_j(k) (j = 1, \dots, N)$  は  $j$  番目のマイクロホンでの受信信号,  $k$  は離散時間を表す. さらに  $y(k)$  は逆フィルタ出力であり,  $s(k)$  の回復信号を表す.  $g_j(k)$  は音源から  $j$  番目のマイクロホンまでの室内インパルス応答,  $h_j(k)$  は逆フィルタを構成する  $j$  番目のマイクロホン出力に繋がる FIR フィルタである.

逆フィルタ  $\mathbf{H}$  を室内インパルス応答を用いず計算するため, 音源および音場について次の仮定を行う.

- (i) 音源は白色信号であり，その自己相関関数がデルタ関数となる．
- (ii) 音源に最も近いマイクロホンが何番目の入力チャンネルに接続されているかが既知である．

以上の仮定が成り立つとすると，次節以降で説明するように室内インパルス応答行列の代わりに入力チャンネル間の相関行列を用いて近似的に逆フィルタを計算することができる．入力チャンネル間の相関行列は，入力信号のみで計算できるため，事前に室内インパルス応答を測定する必要はない．第4.2.1項では上記(i)の仮定を用いた相関関数とインパルス応答の関係を説明し，第4.2.2項では(ii)の仮定を用いた相関行列からの逆フィルタ計算式を導出する．第4.2.3項では，仮定(i)が成り立たない音声のような有色性の音源に対して，平均スペクトル特性を平坦にする白色化フィルタを畳み込むことにより，近似的に白色信号とすることを提案する．

#### 4.2.1 入力チャンネル間相関行列と室内インパルス応答の関係式

$N$ チャンネルの入力信号  $\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_N]$  に対して入力相関行列  $\mathbf{R}$  を次式のように定義する．

$$\begin{aligned} \mathbf{R} &= E\{\mathbf{X}^T \mathbf{X}\} \\ &= \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1N} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{N1} & \mathbf{R}_{N2} & \cdots & \mathbf{R}_{NN} \end{bmatrix} \end{aligned} \quad (4.1)$$

但し，

$$\mathbf{X}_j = [x_j(k) \ x_j(k-1) \ \cdots \ x_j(k-(L-1))],$$

$$\mathbf{R}_{ij} = \begin{bmatrix} r_{ij}(0) & r_{ij}(-1) & \cdots & r_{ij}(-L+1) \\ r_{ij}(1) & r_{ij}(0) & \cdots & r_{ij}(-L+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{ij}(L-1) & r_{ij}(L-2) & \cdots & r_{ij}(0) \end{bmatrix},$$

$r_{ij}(n)$  は  $i$  番目入力信号  $x_i(k)$  と  $j$  番目入力信号  $x_j(k)$  の相関関数を表し,  $r_{ij}(n) = E\{x_i(k)x_j(k+n)\}$  ( $i = 1, \dots, N, j = 1, \dots, N$ ) であり,  $E\{\cdot\}$  は期待値を表す.  $L$  は逆フィルタ長であり, したがって  $\mathbf{R}$  は  $NL \times NL$  行列である. さらに入力信号  $\mathbf{X}$  は, 音源信号  $\mathbf{S}$  と室内インパルス応答  $\mathbf{G}$  を用いて次式で表せる.

$$\mathbf{X} = \mathbf{S}\mathbf{G} \quad (4.2)$$

但し,  $\mathbf{S} = [s(k) \ s(k-1) \ \cdots \ s(k-(K+L-2))]$ ,

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \cdots & \mathbf{G}_N \end{bmatrix},$$

$$\mathbf{G}_j = \begin{bmatrix} g_j(0) & 0 & \cdots & 0 \\ g_j(1) & g_j(0) & \cdots & 0 \\ \vdots & g_j(1) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_j(K-1) & \cdots & \cdots & g_j(K-L) \\ 0 & g_j(K-1) & \cdots & g_j(K-L+1) \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & g_j(K-1) \end{bmatrix}$$

である. 式 (4.1) に式 (4.2) を代入すると,

$$\begin{aligned} \mathbf{R} &= E\{\mathbf{G}^T \mathbf{S}^T \mathbf{S} \mathbf{G}\} \\ &= \mathbf{G}^T E\{\mathbf{S}^T \mathbf{S}\} \mathbf{G} \end{aligned} \quad (4.3)$$

である。ここで、入力信号を白色であると仮定する。つまり、

$$E\{s(k)s(k+n)\} = \delta(n) \quad (4.4)$$

とすると、

$$E\{\mathbf{S}^T \mathbf{S}\} = \mathbf{I} \quad (4.5)$$

となる。ただし、 $\delta(n)$  はデルタ関数、 $\mathbf{I}$  は単位行列である。この時、入力チャンネル間相関行列と室内インパルス応答行列  $\mathbf{G}$  には次式の関係が成り立つ。

$$\mathbf{R} = \mathbf{G}^T \mathbf{G} \quad (4.6)$$

この関係を用いて、次節で逆フィルタの計算式を導く。

#### 4.2.2 入力チャンネル間相関行列を用いた逆フィルタ計算式

式(2.19)と式(4.6)を用いて、逆フィルタ  $\mathbf{H}$  を近似的に計算する方法を述べる。

式(2.19)の両辺に  $\mathbf{G}^T$  を掛け、式(4.6)を用いると、

$$\begin{aligned} \mathbf{G}^T \mathbf{B} &= \mathbf{G}^T \mathbf{G} \mathbf{H} \\ &= \mathbf{R} \mathbf{H} \end{aligned} \quad (4.7)$$

となる。ここで仮定条件(ii)から、音源が  $j$  番目のマイクロホンに最も近いとすると、

$$g_j(0) = \begin{cases} g_1(0) \neq 0 & j = 1 \\ 0 & j \neq 1 \end{cases} \quad (4.8)$$

と表現でき、最初に到達する  $g_1(0)$  以外の  $g_j(0)$  は0となる。つまり、このとき式(4.7)左辺の  $\mathbf{G}^T$  の最左列は室内インパルス応答の先頭  $g_1(0)$  以外は0となり、次式で表せる。

$$\mathbf{G}^T \mathbf{B} = g_1(0) \mathbf{B}' \quad (4.9)$$

ここで,

$$\mathbf{B}' = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} : NL \text{ 行の列ベクトル}$$

であり, 音源からの音が1番目のマイクロホンに最も最初に到達していることを表す. 式(4.7)と式(4.9)から逆フィルタ  $\mathbf{H}$  は,

$$\mathbf{H} = g_1(0)\mathbf{R}^{-1}\mathbf{B}' \quad (4.10)$$

と計算できる. 式(4.10)から逆フィルタ  $\mathbf{H}$  は, 入力相関関数行列  $\mathbf{R}$  が既知であればスケールファクタ  $g_1(0)$  だけ大きさが異なるが, 計算可能である.

### 4.2.3 白色化フィルタを用いた音声平均スペクトルの補正

Semi-blind 残響抑圧法では, 音源信号の白色性を仮定して逆フィルタ  $\mathbf{H}$  を計算している. しかし, 一般には音声は白色信号ではなく, 長時間観測した場合, 音声特有の平均スペクトルを有する. このような音声など白色性が低い音源信号では, 逆フィルタの計算において近似誤差が大きくなり逆フィルタの精度が悪く, 残響抑圧性能も低下する. そこで, 音声の平均スペクトルを白色化フィルタで補正し, 擬似的に音源信号の白色性を高める方法を提案する.

図4.2に, 白色信号と音声(男性音声と女性音声)との10s間観測した自己相関関数と平均スペクトルを示す. 理想的には, 音源信号は白色信号であり, 自己相関関数はデルタ関数, 平均スペクトルは平坦となるのが良いが, 実際の音源は音声信号であり自己相関関数はデルタ関数と大きく異なり, 平均スペクトルも平坦ではない. これを白色化フィルタで補正する.

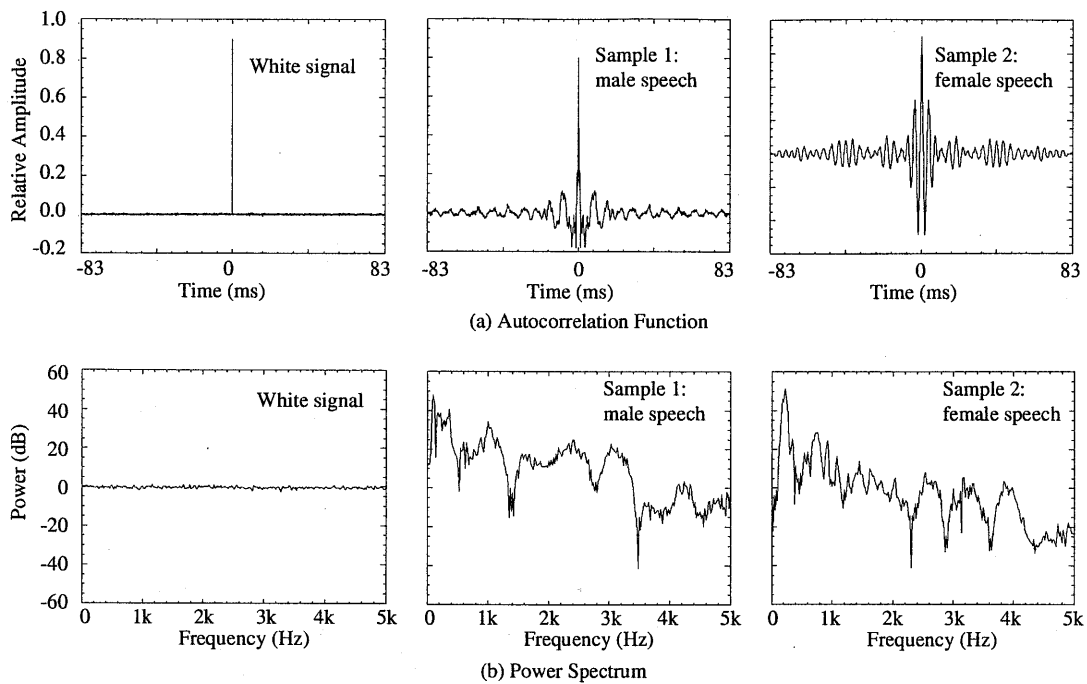


図 4.2: 白色信号と音声信号の違い (a) 自己相関関数 (b) 平均スペクトル

まず、音声信号を白色信号と音声の平均スペクトル特性フィルタで図 4.3 のようにモデル化する。これは、一般的な音声分析で用いられている短時間（数十ミリ秒）の線形予測分析のように音素に対応したホルマントと残差信号に分解するのではなく、室内インパルス応答より十分長い時間（十数秒）観測したときの平均スペクトル特性のフィルタと白色音源信号でモデル化している。これは、MINT 法や Semi-blind 残響抑圧法では線形時不変システムの逆フィルタを求める方法であり、一般的な室内インパルス応答長より変化の早いフィルタでモデル化した場合、システムの時不変性が満たされなくなり時不変な逆フィルタが存在しなくなるため、ここでは平均スペクトル特性のフィルタでモデル化した。

実際の音声では、必ずしも音声の平均スペクトル特性と室内残響特性を完全に分離して推定することは容易ではないが、ここでは直接音からある時刻  $T$  ms までの相関は音声の影響が大きく、それ以上離れた信号間の相関は室内残響に



## Model of Speech Signal

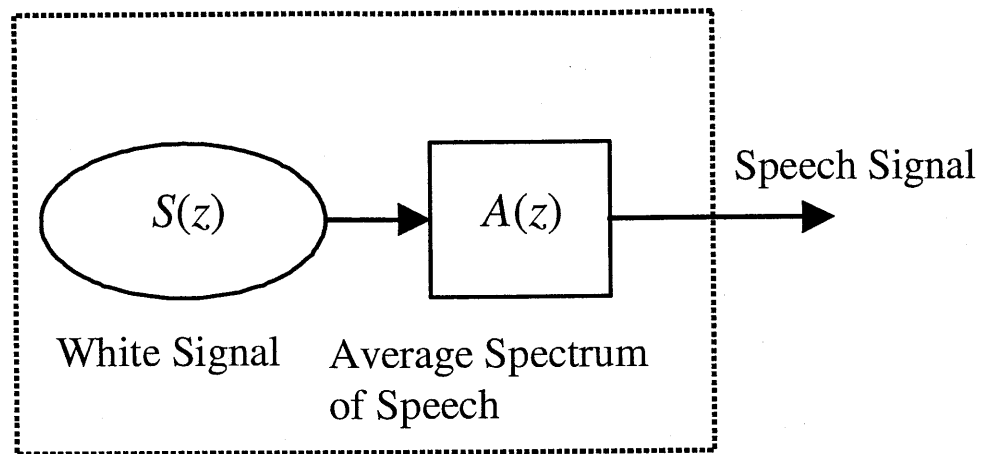


図 4.3: 音声信号のモデル化

よると考える。したがって、直接音から  $T$  ms までの信号の自己相関関数を観測し、それを平均してそのスペクトルを音声の平均スペクトルとする。 $T$  の値は、あまり小さいと音声のスペクトル特性を表しきれず、音声のスペクトルも残響の一部として抑圧されてしまう。一方、 $T$  の値が大きすぎると残響特性が平均スペクトル特性に含まれてしまい、その分残響が抑圧できなくなる。また、この  $T$  の値は白色化フィルタのタップ長  $Q_w$  に対応しており、サンプリング周波数が  $SF$  Hz のとき、 $Q_w = \frac{T \times SF}{1000}$  の関係がある。白色化フィルタの適切な  $Q_w$  の値については、第 4.3.4 節で実験的に検討する。

白色化フィルタによる平均スペクトル補正の結果を図 4.4 に示す。白色化フィルタ長 500 タップとした。図 4.4(a) は、入力信号 (Input signal) と白色化された入力信号 (Decorrelated signal) の自己相関関数である。図 4.4(b) は、入力信号 (Input signal) と白色化された入力信号 (Decorrelated signal) の平均スペクトルである。入力信号は、音声の特性により白色信号からは大きく異なるが、白色化を行うことにより自己相関関数の概形はデルタ関数に近づき、平均スペクトルの包絡は平坦になることが分かる。

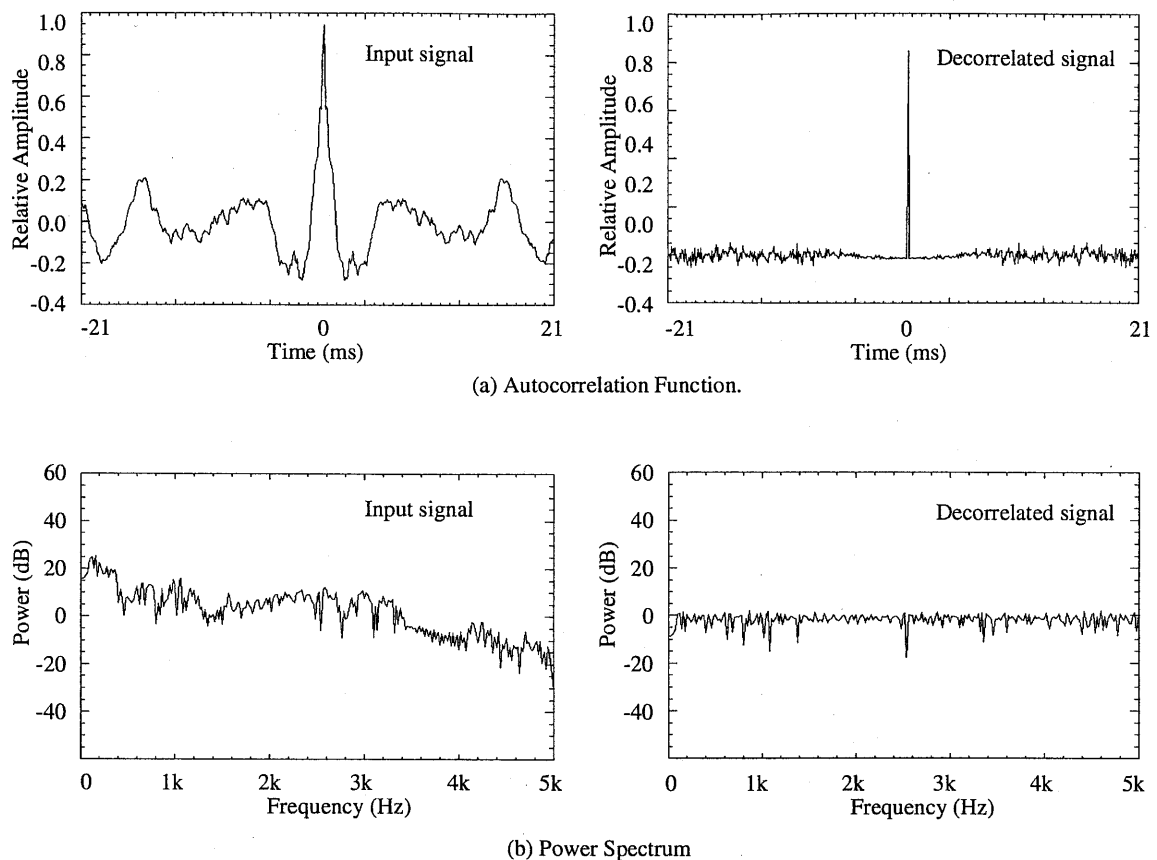


図 4.4: 音声信号の白色化. (a) 自己相関関数 (b) 平均パワースペクトル. 左側が  
入力信号 (Input signal), 右側が白色化された入力信号 (Decorrelated signal).

#### 4.2.4 提案手法の処理の流れ

提案手法における処理の流れを図 4.5 に示す. モデル化された音声信号は, まず白色信号  $S(z)$  に音声の平均スペクトル特性  $A(z)$  が畳み込まれ生成される. その音声は室内に放射され室内インパルス応答が畳み込まれ, マイクロホンで受音される. 受音信号は白色化フィルタ  $1/A(z)$  により白色化され, その入力相関行列から逆フィルタ  $H(z)$  が計算される. マイクロホン入力信号に逆フィルタ  $H(z)$  を畳み込むことにより残響抑圧した回復音声を得られる.

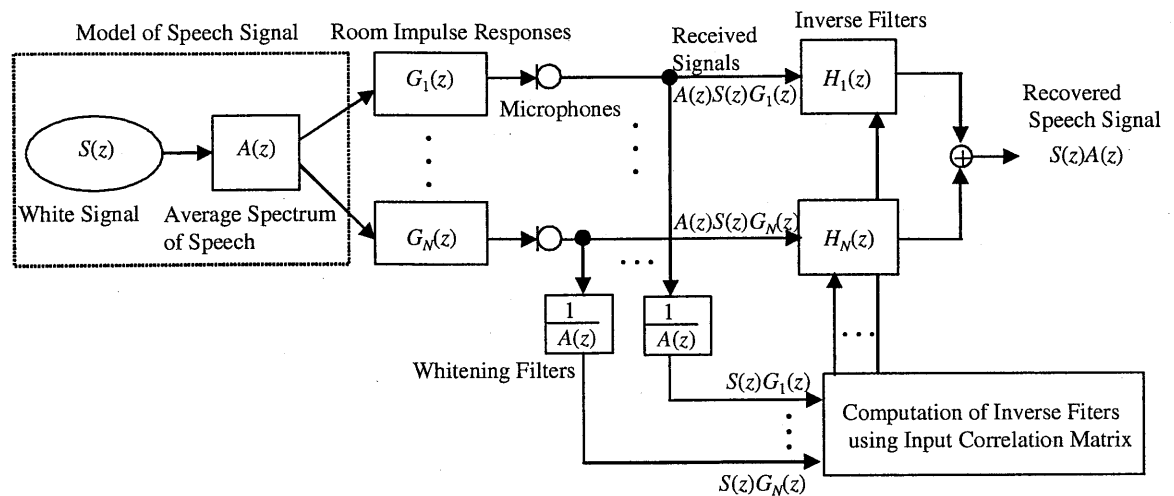


図 4.5: 提案手法の処理の流れ

ここで、音声の平均スペクトル特性  $A(z)$  は、図 4.5 に示されるようにすべてのチャンネル信号に共通に含まれているため、音声の平均スペクトル特性を補正する白色化フィルタ  $1/A(z)$  は各チャンネル共通のフィルタ特性を持たせる。これにより、チャンネル間の位相関係や振幅関係を損なわないまま、入力信号の白色化が図れる。各マイクロホンでの受信信号に白色化フィルタ  $1/A(z)$  をかけることにより、音源を白色信号で近似でき、これを用いて Semi-blind 残響抑圧法で逆フィルタを計算すれば精度の良いものが得られると期待できる。

## 4.3 提案手法の性能評価

### 4.3.1 実験条件

提案手法の効果を確認するため、実際の部屋において実験をおこなった。目的音源は発声内容の異なる男性・女性音声の 2 種類であり、無指向性マイクロホン 8 本、サンプリング周波数 12kHz を用い、部屋の大きさは  $6.6\text{m} \times 4.6\text{m} \times 3.1\text{m}$ 、部屋の残響時間は約 0.55s (インパルス応答長 6,600 タップ) である。マ

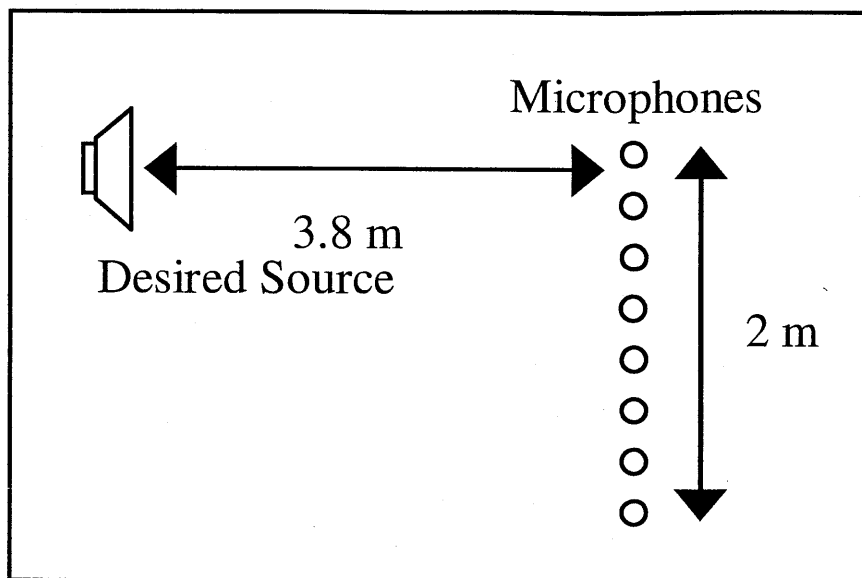


図 4.6: 実験配置図

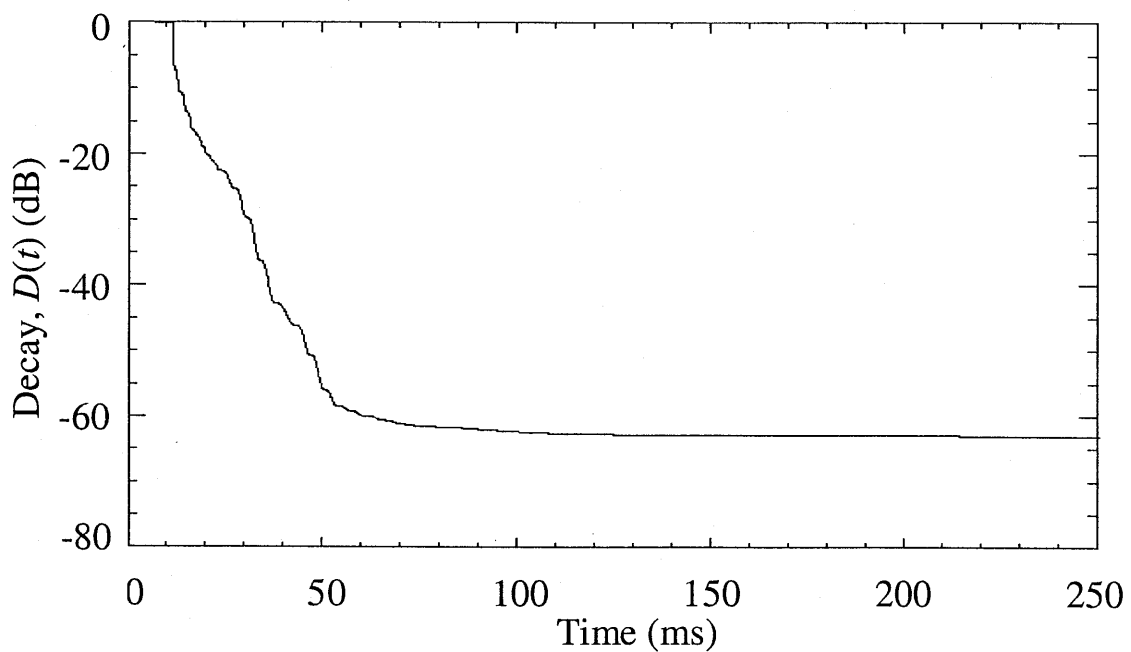


図 4.7: 理想条件での性能

イクロホン及び音源の配置を図 4.6 に示す。残響抑圧処理はインパルス応答を用いずに行ったが、性能比較のため、実測したインパルス応答を用いて残響曲線、残響抑圧改善量も算出した。

### 4.3.2 理想条件での性能

仮定条件 (i),(ii) (式で示すと式 (4.4), 式 (4.8)) が完全に満たされた理想的な状態、つまり、自己相関関数がデルタ関数になるような理想的な白色信号の場合の効果をみる。実際の 6,600 タップの室内インパルス応答を用いて、音源の自己相関関数としては完全なデルタ関数を与え、式 (4.6) と式 (4.10) から計算した。図 4.7 に、提案手法による処理結果の残響曲線を示す。これは、提案手法の性能限界を示すものである。直接音から 50ms 経過するまでに減衰する残響成分パワー  $LD_{50}$  は、60 dB 以上の減衰があり、原理的には提案手法により直接音から 50ms 以内に残響を 60 dB 以上抑圧できることを示している。しかし、実際には音源はこのような白色音源でなく、また、相関関数の平均時間の影響や白色化フィルタ長の影響を受ける。

### 4.3.3 相関関数の平均時間の影響

式 (4.4) で仮定しているように、理想的には受信信号について無限長時間平均すれば音源信号の相関関数がデルタ関数となることを利用しているため、平均時間が短いと分散が大きくなり相関関数がデルタ関数で近似できず、残響抑圧性能は劣化する。平均時間  $T_m$  を 1s から 10s まで変化させた場合の残響曲線を図 4.8 に示す。音源は音声を用いた。図より平均時間が長くなるにつれ、ほぼ単調に残響抑圧効果が大きくなるのがわかる。しかし、実際の使用条件では、話者はその位置に固定したまま、ある時間しゃべり続けなければならない。通常

は平均時間が長くなればなるほど話者が移動する可能性が大きくなり平均の意味が無くなる可能性が大きくなる。

#### 4.3.4 白色化フィルタ長の影響

白色化フィルタの長さが残響抑圧性能に与える影響を実験的に検討した。白色化フィルタの長さ  $Q_w$  を、10タップから1000タップまで変化させた場合の残響曲線を図4.9に示す。実験では、相関関数の平均時間を10sに固定した。図4.9(a)が音源が男声の場合、図4.9(b)が女声の場合である。これらの図からタップ長が500と1000タップの場合は、男声女声どちらもほとんど同じ減衰量である。しかし、10タップから200タップまでは、男声と女声で性能が異なる。これは、平均時間が10sであるため、音源特性によるバラツキの差が出ていると考えられる。つまり、女声はタップ長が短い場合には、白色化フィルタによる白色化の効果が低く、十分近似ができていないと考えられる。この程度の平均時間では、音源特性のバラツキを吸収するためにも少なくとも500タップ程度の長さが必要であることが分かる。

#### 4.3.5 逆フィルタ長の影響

逆フィルタ長は、理論的にはインパルス応答長を式(2.20)に代入して計算されるMINT逆フィルタ長に設定されるが、実際の室内インパルス応答長は連続的であり、ある値で打ち切って離散的に決めなければならない。したがって、打ち切り誤差が生じ、測定誤差などにより室内インパルス応答長が正確には推定できない場合がある。そのため、逆フィルタ長がMINT逆フィルタ長とは厳密には異なる場合でも、残響抑圧効果がそれほど低下しないことが望まれる。

最小固有ベクトルを用いた残響抑圧では、インパルス応答長の推定値から計

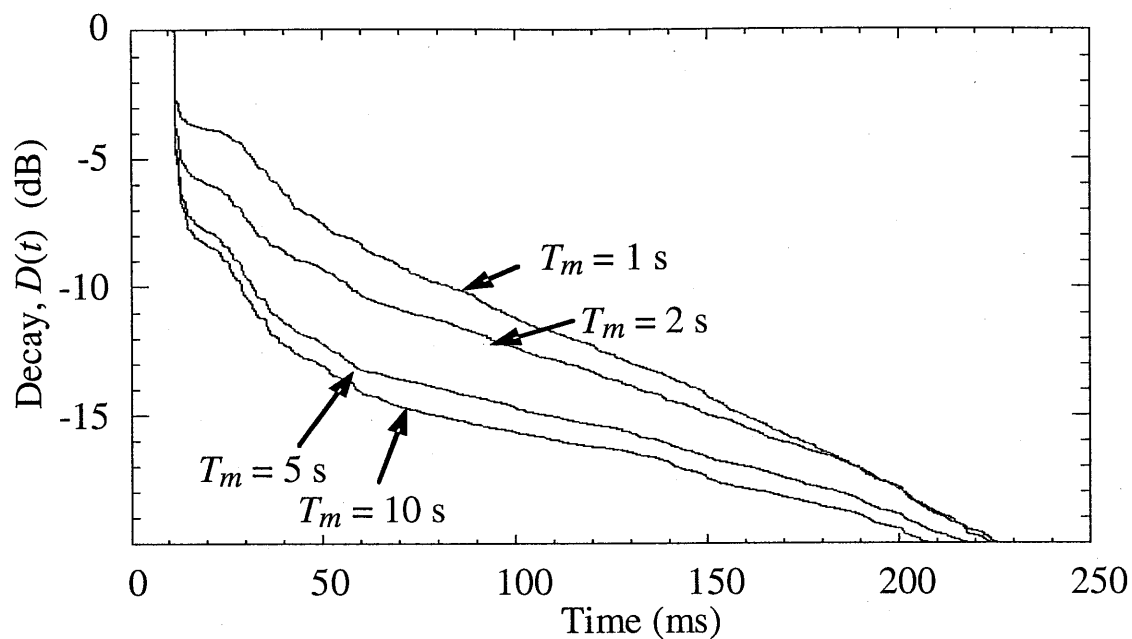


図 4.8: 自己相関関数の平均時間  $T_m$  の影響

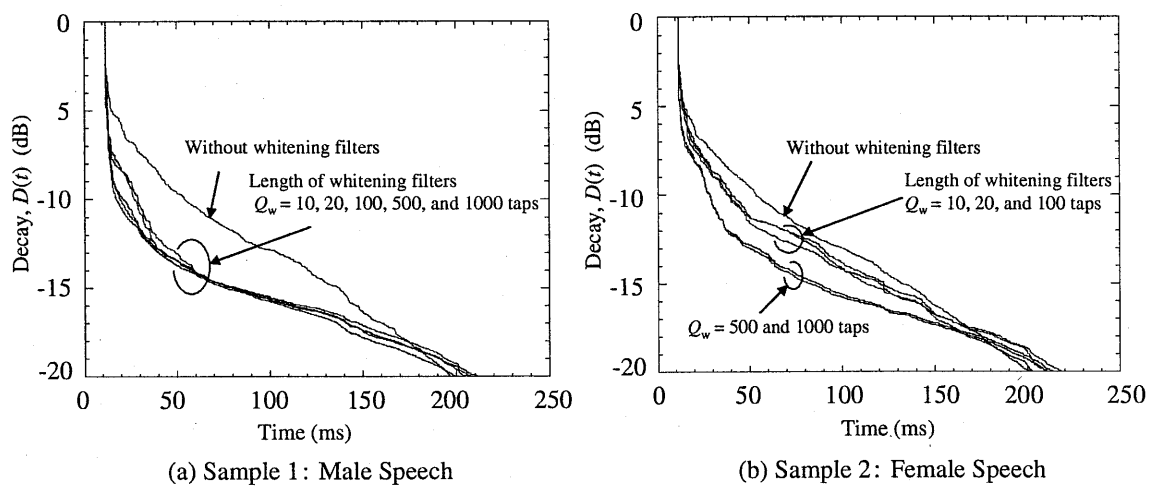


図 4.9: 残響抑圧における白色化フィルタの長さの影響

算した逆フィルタ長と MINT 逆フィルタ長が同じでない場合や、測定誤差により真の最小固有値が誤差により埋もれてしまう場合などには、逆フィルタの計算に大きな誤差が生じ、残響抑圧することが困難であった。その一例として、図 4.10 に、目的音源として白色信号を用い、チャンネル数  $N = 2$ 、インパルス応答長  $K = 30$  タップと比較的短い場合の二入力残響抑圧法と提案手法の逆フィルタ精度のシミュレーション結果を示す。また、測定雑音を模擬するために、各チャンネルに無相関な白色雑音を SN 比 40 dB で加えた場合も行った。図 4.10 の横軸は逆フィルタ長、縦軸はインパルス応答に逆フィルタを畳み込んだ後の処理結果とデルタ関数との差を直接音のパワーで正規化し dB 表示したもので、逆フィルタの精度が高ければ縦軸の誤差は小さく、低ければ誤差は大きくなる。ここで、

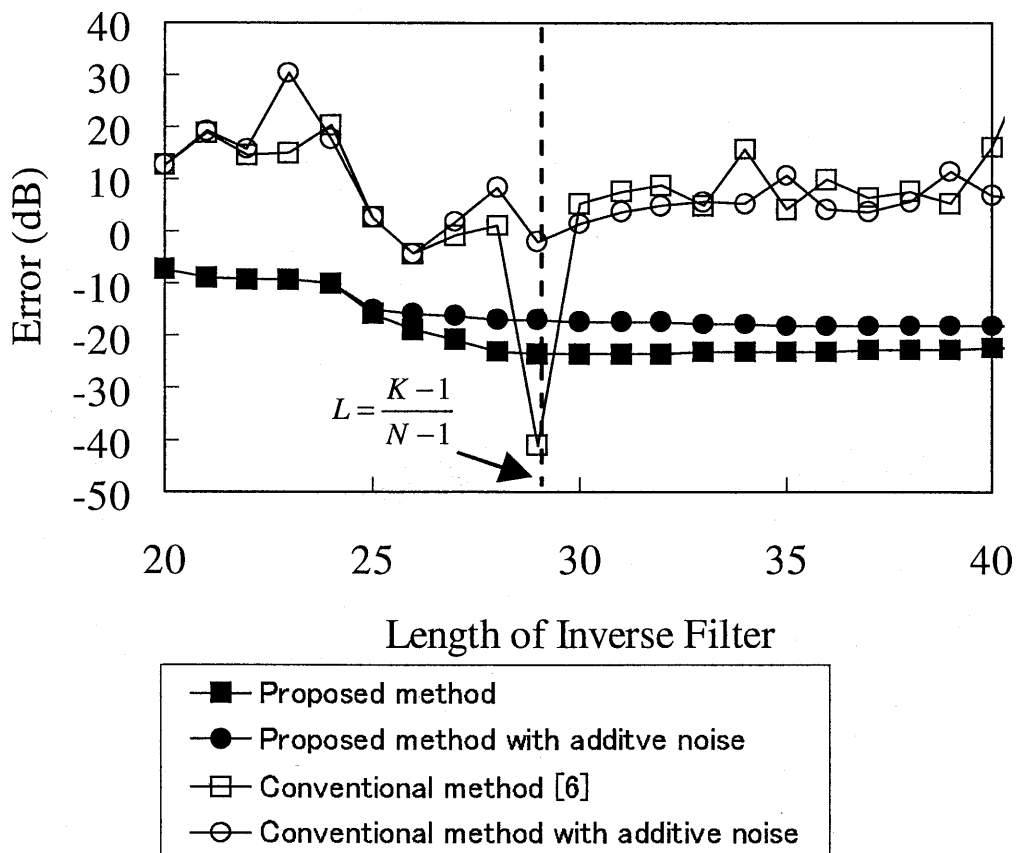


図 4.10: 逆フィルタの長さ精度の関係



実際のインパルス応答長は30タップなので、式(2.20)から計算されるMINT逆フィルタ長は29タップと計算され、これが必要十分な逆フィルタの長さとなる。図4.10から分かるように、従来法では、測定に雑音が無い場合は、逆フィルタ長がMINT逆フィルタ長と同じであれば精度の高い逆フィルタが得られるが、逆フィルタ長に大きく影響を受け、1タップでも異なったり測定に雑音が加わると精度の良い逆フィルタは得られなくなる。一方、提案手法では、逆フィルタ長がMINT逆フィルタ長以上の長さがあれば、逆フィルタ長が変わってもほぼ性能が変わらない安定した精度の逆フィルタが得られている。また、測定に誤差が含まれる場合でも精度は若干劣化するが逆フィルタ長が変化しても安定した精度の逆フィルタが得られている。逆フィルタの精度は、少なくとも逆フィルタ処理前と処理後で誤差が小さくなっていなければ、逆フィルタを処理する効果がない。図4.10の場合、処理前の誤差(残響成分)のパワーと直接音パワーの比は3dBであったので、それ以下である必要がある。提案手法は、条件によらず安定して3dB以下の低い誤差を示しており、一方、従来法は、条件がMINT逆フィルタ長で雑音が無いとき以外は、ほとんど3dB以上となっている。

従来法ではMINT逆フィルタ長を正確に推定する必要があるが、インパルス応答長が数千タップに及ぶ残響時間の長い室内インパルス応答に対しては、打ち切り誤差や測定誤差、さらには大きいサイズの行列の最小固有値計算に伴う誤差によりますますインパルス応答長を厳密に推定することは困難となり、精度のよい逆フィルタを計算することは不可能に近くなる。一方、提案手法では、インパルス応答長が数千タップに及ぶ場合でも大雑把にMINT逆フィルタ長以上に逆フィルタ長を設定することにより安定した精度の逆フィルタを計算することができると考えられる。図4.6に示した実験条件では残響時間は0.55sでインパルス応答長はおよそ6600タップである。この条件での提案手法の逆フィルタ長に対する残響抑圧改善量RRR (Reverberation Reduction Rate) を図4.11

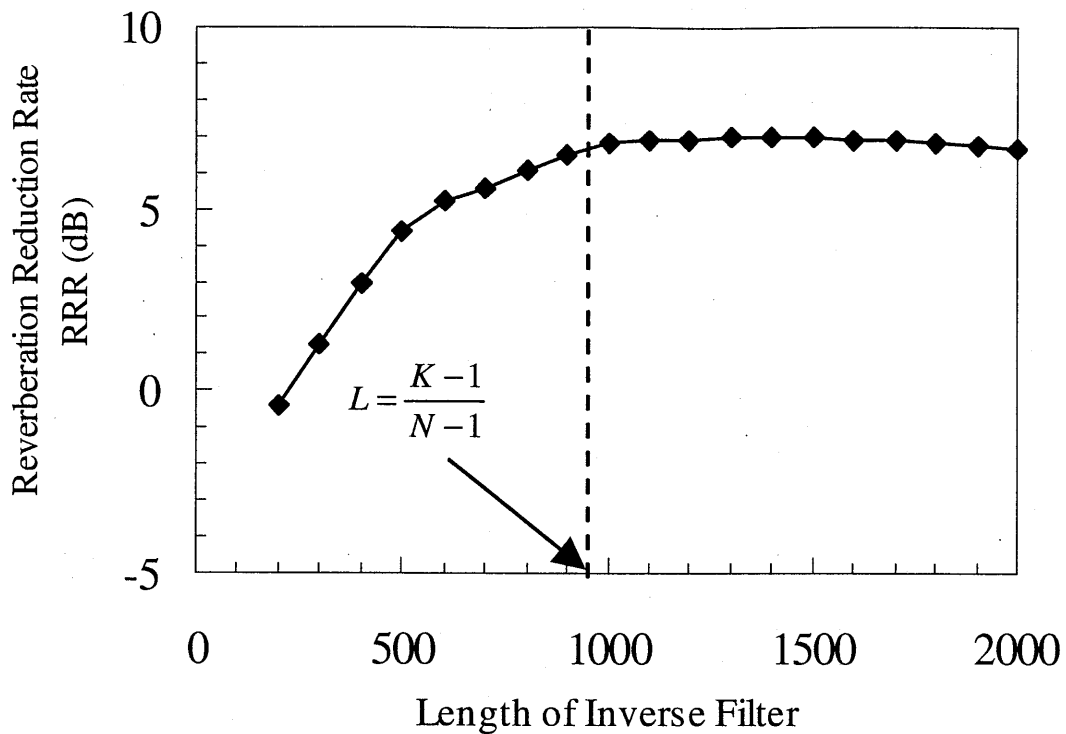


図 4.11: 逆フィルタの長さとお響抑圧改善量 RRR

に示す。この図から、提案手法では逆フィルタ長が MINT 逆フィルタ長以上であれば、安定して約 7 dB の抑圧効果があり、正確に MINT 逆フィルタ長つまりインパルス応答長を推定する必要がないことが分かる。

#### 4.3.6 提案手法の性能

提案手法の性能を評価するため、(i) 白色化フィルタを用いない場合、(ii) 白色化フィルタを用いた本提案手法、(iii) 目的音源として音声でなく白色信号を用いた場合の残響曲線を図 4.12 に示す。処理前の残響曲線には音源に最も近いマイクロホン信号を用いた。提案手法により処理前に比べ 7dB 程度残響エネルギーを抑圧できている。また、白色化なしの場合に比べて、白色化フィルタを導入することにより 4dB の改善が得られている。平均時間 10s での提案手法の限界

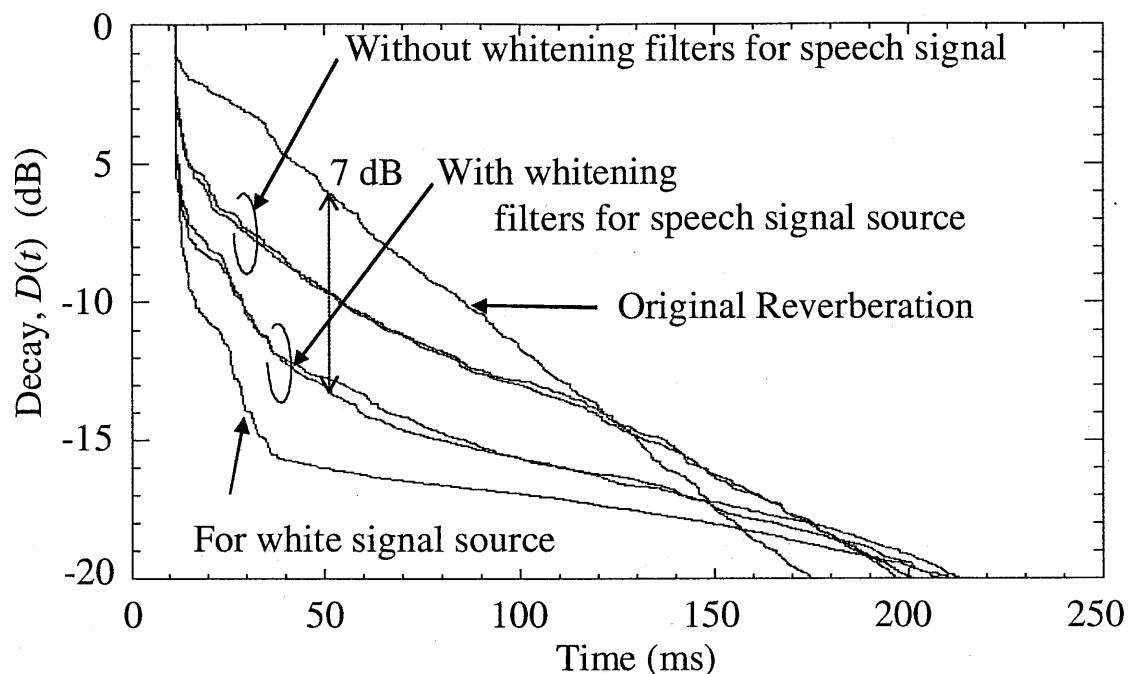


図 4.12: 残響抑圧における白色化の効果.

は白色信号入力の場合であるが、白色化ありの提案手法と比較すると、今回の白色化フィルタでは完全な白色化が行われておらず、改善の余地があることがわかる。原因としては、室内インパルス応答が畳み込まれた音声信号から原音声の平均スペクトルを推定しているため、推定精度が悪いことが考えられる。

#### 4.3.7 遅延和アレーとの比較

代表的な残響抑圧技術として遅延和アレー法 [24] と本提案手法を比較した。遅延和アレー法では、各チャンネルに単純な遅延を接続して、音源からの遅延を補正し、音源からの信号が各チャンネルで同じ位相となるように調整した。マイクロホン数は同じ 8 本である。図 4.13 にその結果を示す。本手法は、約 15 dB の残響エネルギーの減衰まではすばやく残響を抑圧している。減衰量が 15 dB を越えた時点から本手法では減衰量が緩やかになる。一方、遅延和アレー法は、処

処理前の残響曲線をほぼ 3 dB そのまま平行移動した形になっている。表 4.1 に残響抑圧改善量 RRR を示す。表から、従来技術である遅延和アレー法に比べ本手法は抑圧量が 4 dB 大きい。遅延和アレー法による改善量がマイクロホン 8 本で 3dB と小さい理由として、処理前の残響曲線が音源に最も近いマイクロホンが用いられている点、マイクロホン間の残響成分が互いに相関を持っている点が考えられる。

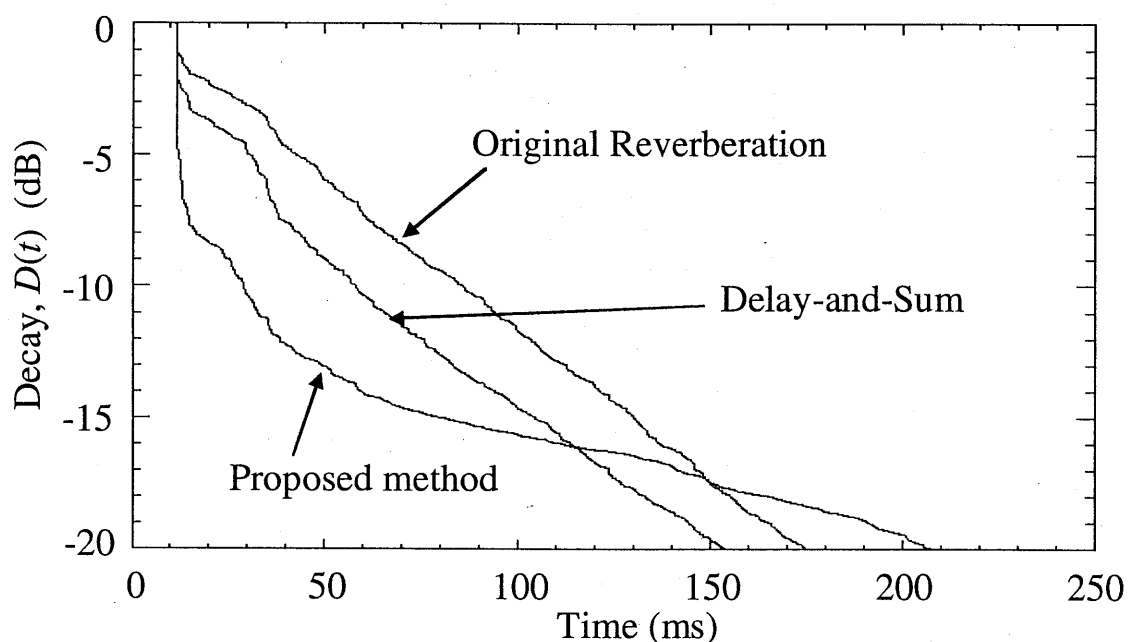


図 4.13: 遅延和アレー法との比較

#### 4.3.8 波形の回復

音声信号を音源とした場合の残響抑圧後の時間波形とインパルス応答を示す。平均時間  $T_m$  は 10s, 白色化フィルタ長  $Q_w$  は 500 タップである。図 4.14 に、残響抑圧処理前後の音声波形を示す。処理後の音声は処理前に比べ、音声のピッチ成分が明確になり残響が抑圧できていることが分かる。さらにインパルス応答

表 4.1: 残響抑圧改善量 RRR

	Reverberation reduction rate: RRR
Delay-and-Sum	3.0dB
Proposed method for speech signal source	7.0 dB
Proposed method for white signal source	10.0 dB

の残響抑圧処理前と処理後の比較を図 4.15 に示す。処理前波形は残響のある室内インパルス応答であるが、処理後インパルス応答の直接音以降の部分と比較すると残響抑圧処理によってインパルス応答の残響成分が十分抑圧されていることがわかる。

#### 4.4 むすび

本章では、二入力残響抑圧法のように入力相関行列の最小固有ベクトルを用いインパルス応答を推定し逆フィルタを計算するのではなく、直接、入力相関行列から方程式を立て、逆フィルタを計算する Semi-blind 残響抑圧法を提案した。音源が有色信号である音声のため性能が低下する点を、提案手法では収録された音に音声の平均スペクトルの逆特性を持つ白色化フィルタを用いることにより改善を図った。実験の結果、提案手法では 7dB の残響抑圧効果があり、白色化フィルタが無い場合に比べ 4dB の改善が図れた。また、一般的な遅延和アレー法による残響抑圧に比べても 4dB の改善効果があった。

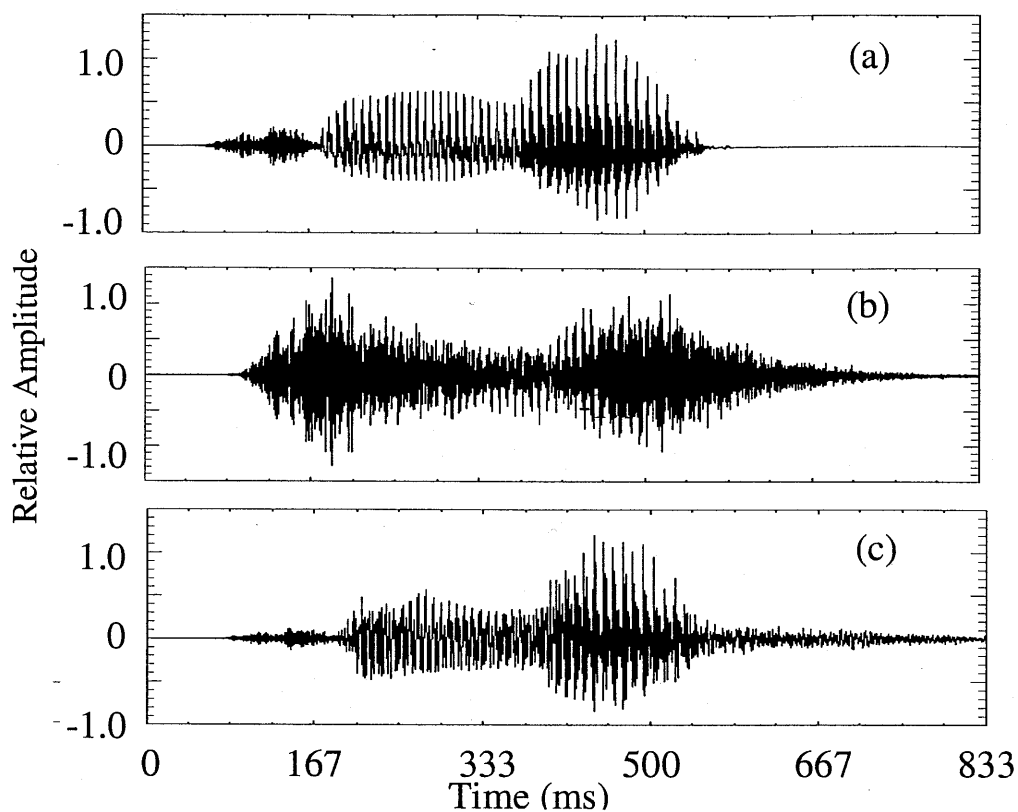


図 4.14: 音声波形の残響回復 (a) 原音声 (b) 残響音声 (c) 残響抑圧処理後の回復音声

今回、音声の平均スペクトル特性を室内残響特性と分離して白色化するために、実験的に必要な白色化フィルタ長を決定し、直接音から約 50ms までの信号間の相関を音声の平均スペクトルによるものであり、それ以上の時間間隔の空いた信号間の相関は室内残響によるものであると単純化している。しかし、実際には信号間の時間間隔だけで明確に分離できるものではなく、音声の平均スペクトル特性と室内残響特性が重なっている部分が存在し、白色化の効果を低下させている。表 4.1 において白色信号を音源とした場合に比べ、提案手法はまだ 3dB の差がある。これは、白色化を最適化すればまだ 3dB 程改善できることを示す。

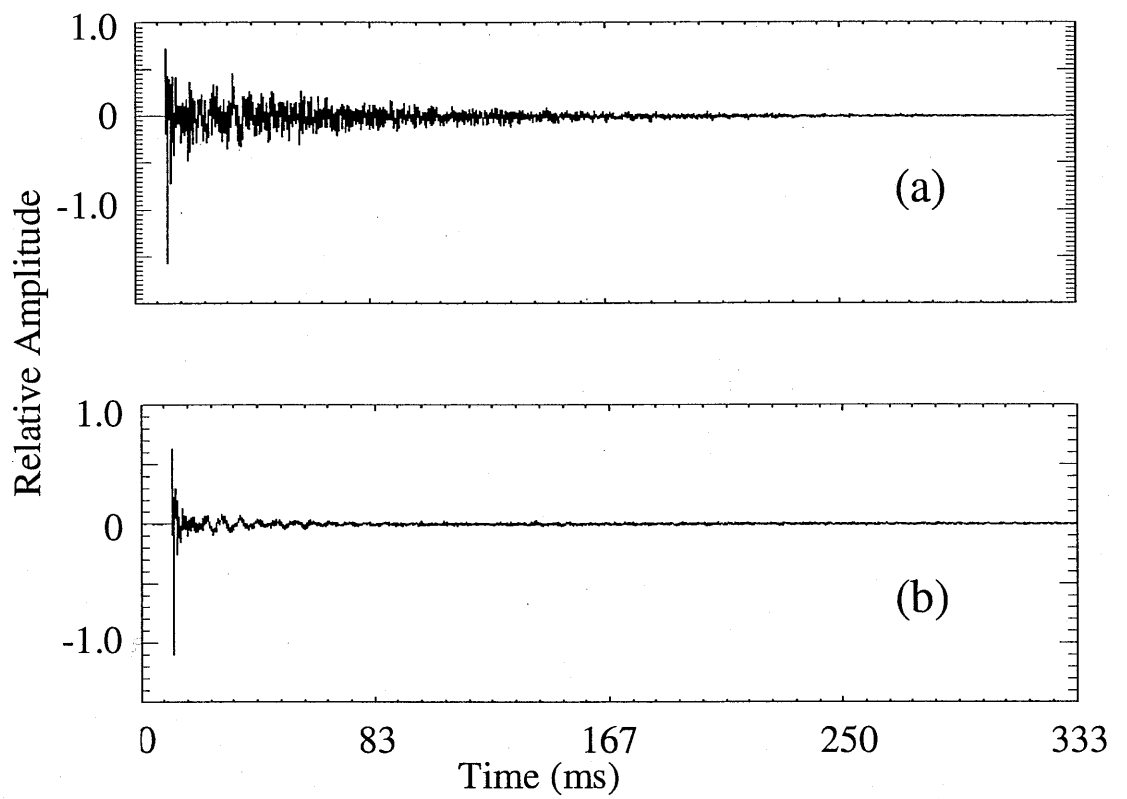


図 4.15: 提案手法による残響抑圧結果 (a) 室内インパルス応答 (b) 処理後の  
 応答