

## 残響音場における相互相関関数を用いた音源包絡の回復

広林 茂樹<sup>†a)</sup> 林 智也<sup>†</sup> 山淵 龍夫<sup>†</sup> 吉沢 寿夫<sup>†</sup>

Signal Envelope Recovery Using a Cross-Correlation Function in Reverberation Sound Field

Shigeki HIROBAYASHI<sup>†a)</sup>, Tomonari HAYASHI<sup>†</sup>, Tatsuo YAMABUCHI<sup>†</sup>, and Toshio YOSHIZAWA<sup>†</sup>

あらまし 残響のある室内で観測された信号から音源波形を回復する研究は、機械故障診断や拡声系会議通信のめいりょう度向上にとって重要である。この研究は残響音場伝達関数の逆フィルタ処理を施すことなく、観測信号のみから回復する方法を目的としている。残響は過去に発生した複数のエコー成分によって形成されている。本論文では観測信号をそれぞれ短い時間区間（フレーム）で細分化し過去のフレームと比較してエコーの影響を軽減することを試みた。対象フレームに対し、過去にさかのぼって、それぞれのフレームとの相関関数を求めた。各相関関数の最大値を重み係数として、過去のフレームを対象フレームより差し引くことによって残響信号を回復した。更に回復効果を向上させるため、各フレームで帯域分割し回復過程を狭域化した。その結果、残響音声に対し回復音声の包絡線ひずみが最大 8 dB、平均 4.6 dB 回復することが分かった。

キーワード デコンボリューション、相関係数、信号処理、音源回復、残響抑圧

## 1. ま え が き

残響は、室内壁面に繰り返し反射された音が音源からの直接音に積み重なったものである。この反射音は音に広がりや奥行を与え、高臨場感や仮想現実感 (Virtual Reality) を創出するシステムにおいては人工的に付加することが求められている。しかし、音声認識や拡声系会議通話においては、室内反射音は音源波形にひずみを与え特徴を埋もれさせる原因となっている。実世界における音場において音源信号と観測信号は、室内伝達系を介してほぼ線形システムで近似できることがよく知られている。このような線形システムにおける逆問題、すなわち観測信号から音源信号を復元する逆問題では伝達系の推定が必要である。伝達特性は一般に時変系であるため、常に適応的な推定が必要であり、逆フィルタ処理による音源波形復元は実用的に難しい。S.T. Neely と J. B. Allen らは、単一マイクロホンで受音された信号から室内伝達特性の最小位相成分のみを取り除いて、回復信号を求める最小位

相逆フィルタ法 [1] を提唱している。音場伝達特性が最小位相特性を有する場合には効果的に残響を抑圧できる。しかし、残響時間の増加に伴い非最小位相成分も一般的に増加するため、伝達系に与える影響が大きくなる。すなわち、非最小位相成分の最小位相化による位相ひずみの増加によって、回復精度が低下する。これに対し、三好、金田らは複数のマイクロホンを用いた音場逆フィルタ理論 (MINT) [2] を提唱している。音源の数に対しマイクロホンを一つ以上多く配置することによって、非最小位相成分の影響を強く受けた伝達特性でも、音源信号を正確に復元できる。しかし、音源から各々のマイクロホンに至る伝達特性の零点が重複する場合や接近する場合では解が発散する。また、タップ長の増加に伴い、解を求めるための逆行列演算に膨大な時間を要する。更に継承的な演算の増加に伴い計算誤差が大きくなる。また、この MINT 法を発展させた中島ら [3] や伊勢ら [4] によって最小ノルム解に基づく逆フィルタ理論も提案されているが、残響時間の増加に対しては対策が論じられていない。このように、長い残響時間を有する室内環境では実際にフィルタの解を求めることができず、時変系では伝達特性の適応的な測定あるいは推定が一般的に求められる。そこで、筆者らは既に伝達特性の測定を必要としな

<sup>†</sup> 富山大学工学部、富山市  
Faculty of Engineering, Toyama University, Toyama-shi,  
930-8555 Japan

a) E-mail: hiro@iis.toyama-u.ac.jp

いブラインド解法としてパワーエンベロープ逆フィルタ処理 [5], [6] を提案した. これは MTF (Modulation Transfer Function) の理論 [7] に基づき, 音源信号と伝達系を信号の包絡線処理によって回復する手法である. コンサートホールなどに代表される拡散音場的な伝達特性を有し, 指数関数的に減衰する場合には高い改善効果があるものの, 残響時間が短く, その減衰過程が変則的な伝達特性では高い改善効果が得られない. 例えば, 室内壁面材料の特性に伴う帯域間における残響時間の変化がある場合などでは回復精度が低下する. このため, 帯域ごとの残響時間の推定が必要である [8]~[10]. また, 残響時間の推定では, 過去に発生した信号と相関係数を求めることで推定できることが明らかにされており, 反射音の特定に有効であることが知られている [11], [12]. 反射音に着目した残響抑圧法には, 田沢らによって提唱されたケプストラムを用いたエコーキャンセラ [13] がある. 1, 2 回程程度の反射音に対し, ケプストラム振幅のしきい値処理により反射音のケプストラムを特定し, 対象信号から反射音に起因する成分を除去することを試みている. しかし, 残響時間が長く多くの反射音がある場合においては, 限界が指摘されている.

そこで, 本論文では多くの反射音による影響を効果的に軽減することを検討した. 特に残響時間が 0.5 秒以上の長い室内環境下において, 単一マイクロホンを用いて音声発生直後の観測音声と, 反射音に伴い重畳されたそれ以降の観測音声 (残響音声) との類似性を逐次, フレーム間の相互相関関数によって定量的に抽出し, 相関値を重みとした差分処理によって残響の影響を軽減することを検討した. 信号を短い時間窓で処理 (フレーム処理) することや, フィルタバンクを用いることで, 長い残響環境下に対応させた. 相互相関関数を用いることで, 残響に起因するひずみの影響を定量的に抽出し, その定量値に基づき残響を抑圧する. 残響過程で過去に発生した信号に起因するひずみが重畳される観測信号と, 過去に観測した信号との相関関数を計算することで, その影響を推定し, その窓内の信号に対し逐次残響抑圧処理を施す手法を提案する [14]. 本論文では相関値に基づく重み付き差分を用いることで, 伝達特性の測定及び, 伝達関数の逆フィルタ処理を行わず, おおよその残響時間の把握のみで実現する残響抑圧法の可能性について実験的に調べる.

## 2. 相関関係とインパルス応答

原音声を  $x(t)$ , 室内インパルス応答を  $h(t)$  とすると残響音声  $y(t)$  は次式の畳込み演算で表されることが知られている.

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \\ &= x(t) * h(t) \end{aligned} \quad (1)$$

ただし,  $*$  は畳込み演算子である. このように任意の時間  $t$  における観測信号  $y(t)$  が音源信号  $x(t)$  の積分系で表される. この影響のため, 時間  $t$  以外に発生された多くの音源信号から影響を受けている. 一般に  $h(t)$  は  $t \geq 0$  で発生する反射信号のため, 音源信号  $x(t)$  は  $t$  以降の時間で観測信号に広く影響を与える. すなわち, 音源信号発生直後の観測信号には, 過去から積み重なった多くの反射音が含まれ, 音源信号の特徴を埋もれさせる.

室内インパルス応答  $h(t)$  は, 一つの直接音と複数の反射音で構成されるため, まず直接音と一つの反射音の和で表される単純なインパルス応答モデルで回復アルゴリズムを検討した. 図 1 に単純な室内インパルス応答モデルの残響過程を示す. 図 1 (a) は白色雑音を音源信号とした例である. 図 1 (b) に単純な室内インパルス応答モデル, 図 1 (c) に残響信号として音源信号と単純インパルス応答の畳込み演算信号をそれぞれ示す. 図 1 (d) は残響信号と音源信号の相互相関関数を次式により計算したものである.

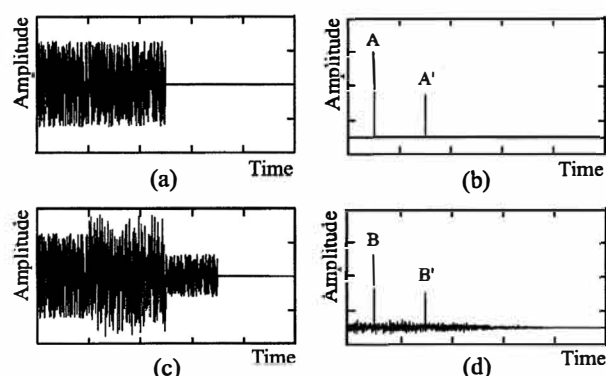


図 1 単純な音声波形モデル (a) 雑音信号, (b) インパルス応答, (c) 残響信号, (d) 相互相関

Fig. 1 Simple input-output model and simple impulse response: (a) source signal, (b) impulse response, (c) reverberant signal, (d) cross-correlation.

$$a(t) = \int_{-\infty}^{\infty} x(\tau)y(\tau+t)d\tau \quad (2)$$

図 1 (b), (d) より明らかなように, 図中 A 部を B 部に見立てると, A' 部が B' 部に対し, 相対的な振幅値と時間遅れがほぼ一致する. そこで, 式 (2) に示す相互相関関数によって, 室内インパルス応答の直接音と反射音の相対的な振幅比と時間遅れが抽出できると仮定した. 次式のように音源信号と観測信号の相互相関関数から, これら振幅情報と時間遅れの情報を抽出することで, 反射波の影響を取り除き音源信号  $x(t)$  を復元できる可能性がある.

$$\hat{x}(t) = y(t) - ax(t - \tau) \quad (3)$$

ただし,  $\hat{x}(t)$  は推定された音源信号,  $a$  は図 1 (d) の図中 B' 部に起因する相関係数,  $\tau$  は直接音と反射音の間における時間遅れである. ここで, 複数の反射波がある場合を考える.  $h(t)$  を次式のように単純に  $N$  個の反射パルスで構成されているとする.

$$h(t) = \delta(t) + \sum_{k=0}^N a_k \delta(t - \tau_k) \quad \delta(t) = \begin{cases} 1 & (t = 0) \\ 0 & (t \neq 0) \end{cases} \quad (4)$$

観測信号は  $x(t)$  と  $h(t)$  の畳込みで表され, 次式のように求めることができる.

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau \\ &= \int_{-\infty}^{\infty} \left\{ \delta(\tau) + \sum_{k=0}^N a_k \delta(\tau - \tau_k) \right\} x(t - \tau)d\tau \\ &= x(t) + \sum_{k=1}^N a_k x(t - \tau_k) \end{aligned} \quad (5)$$

したがって, 複数の反射波に関しても式 (3) を単純に拡張できると考えると, 次式により残響ひずみの回復

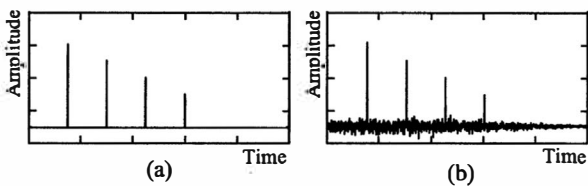


図 2 相対的な振幅値と時間遅れ (a) インパルス応答, (b) 相互相関  
Fig. 2 Relative amplitude value and time delay: (a) impulse response, (b) cross-correlation.

の可能性がある.

$$\hat{x}(t) = y(t) - \sum_{k=1}^N a_k \times x(t - \tau_k) \quad (6)$$

ただし,  $N$  は反射波の個数である. また, 図 2 (a) のインパルス応答のような複数のパルス列においても, その相互相関関数は図 2 (b) のようになるため, 式 (6) の係数  $a_k$  も残響信号の相互相関関数で近似できると考えた. 本論文では式 (6) に示される回復アルゴリズムを近似式により導出し, その回復効果を定量的に検討する.

### 3. 相関関係を用いた残響抑圧法

本論文では, ブラインド解法によって音源信号  $x(t)$  を回復することを目的としているため, 式 (1) の室内インパルス応答  $h(t)$  及び, 式 (6) の右辺第 2 項の音源信号  $x(t - \tau_k)$  を用いることなく音源信号を復元することを想定した. すなわち, 式 (6) で表される回復信号  $\hat{x}(t)$  は, 右辺第 2 項の音源信号  $x(t - \tau_k)$  で構成されているため, この  $x(t - \tau_k)$  を他の関数で近似することを考える. 残響効果は, 図 3 に示すように, 白色雑音等の音源信号を室内で発生すると, その観測信号の振幅は, 音源信号の発生開始, 継続, 停止に伴い, 順に成長過程, 平衡過程, 減衰過程に入る. そのため, 音源発生直後の観測信号は平衡状態の観測信号に比べ, 残響の影響が相対的に少ない. そこで観測信号  $y(t)$  を微小区間で区切ると, 音源発生直後の信号区間は, 残響の影響が少なくほぼ直接音のみで構成されるため, 音声信号  $x(t)$  として近似できると考えた. すなわち, 音源発生直後の観測信号  $y(t)$  を音源信号  $x(t)$  として, 残響の影響が取り除かれた回復信号  $\hat{x}(t)$  を順次式 (6) の第 2 項  $x(t - \tau_k)$  に代入し, 次式のように逐次的に回復信号を求める.

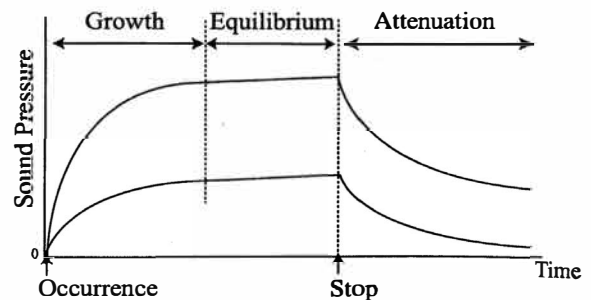


図 3 残響過程  
Fig. 3 Reverberation process.

$$\begin{aligned}
 \hat{x}_1(t) &= y_1(t) \\
 \hat{x}_2(t) &= y_2(t) - a_1 \hat{x}_1(t - \tau_1) \\
 \hat{x}_3(t) &= y_3(t) - a_2 \hat{x}_2(t - \tau_2) - a_1 \hat{x}_1(t - \tau_1) \\
 \hat{x}_k(t) &= y_k(t) - \sum_{m=1}^{k-1} a_m \hat{x}_m(t - \tau_m) \quad (7)
 \end{aligned}$$

なお、フレーム処理のため、微小区間に分別された音源信号  $x(t)$ 、観測信号  $y(t)$ 、回復信号  $\hat{x}(t)$  をそれぞれ  $x_k(t)$ 、 $y_k(t)$ 、 $\hat{x}_k(t)$  とした。また、 $y_k(t)$  は、観測信号  $y(t)$  のフレーム番号  $k$  における微小区間 ( $k$  番目のフレーム) であり、 $\hat{x}_k(t)$  は回復信号  $\hat{x}(t)$  の  $k$  番目の微小区間である。また、本論文では実時間処理を想定しており、観測信号  $y(t)$  と音源信号  $x(t)$  の相互相関関数を逐次的に近似する。そこで、暫定的に  $a_m$ 、 $\tau_m$  を  $\hat{x}_m$  と式 (7) の最下部の式の第 2 項との相互相関関数で求められた最大値、時間遅れ  $\tau$  で用いた。本手法の流れを図 4 に示す。図中  $g(n)$ 、 $f(n)$  は任意

の  $k$  に対しそれぞれ  $\hat{x}_{k-m}(t)$ 、 $y_k(t)$  に対応する。ただし、 $m$  は正の整数である。図中の矢印線に示すように回復対象フレームに対し、隣り合った過去の信号のフレームと回復対象フレームとの相互相関関数を求め、相関係数の大きさの最大値を重みとして回復対象フレームから反射波の影響を取り除く。このような処理を逐次的に行い、回復信号を計算する。図中  $M$  は回復対象フレームに対する最大比較フレーム数である。また、音声はよく知られているように広域信号のため、フィルタバンクを用いて狭域化を行い、処理効果を高めることが行われている。筆者らが既に室内伝達特性の測定を行わないブラインド処理の一例として提案した、パワーエンベロープ逆フィルタリング [5] でも狭域化の効果 [15] が明らかにされている。そこで、本論文でも図 5 に示すフィルタバンクを用いて、観測信号を分割しそれぞれの帯域で独立させて処理を試みることにした [16]。図中  $N$  は分割総数である。  $H_k(z)$ 、

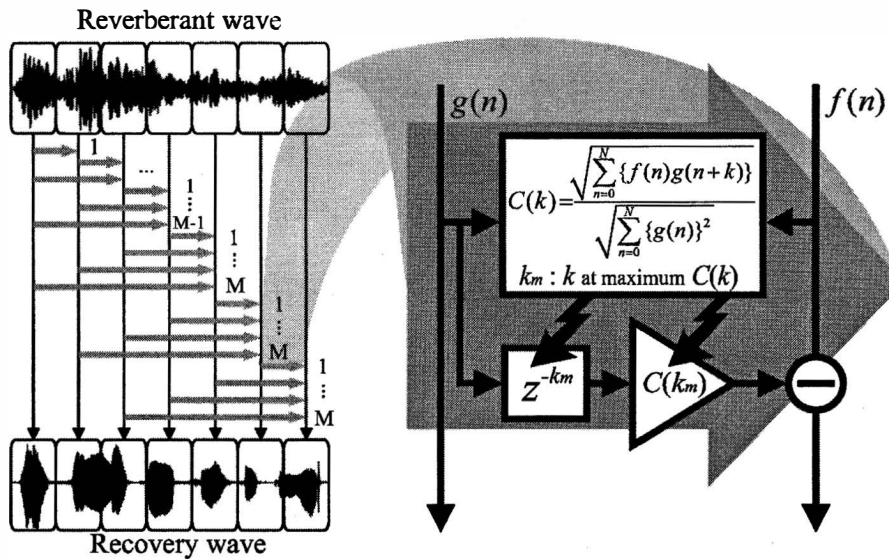


図 4 相関値を用いた残響抑圧の概略図

Fig. 4 Block diagram of reverberation suppression using correlation coefficients.

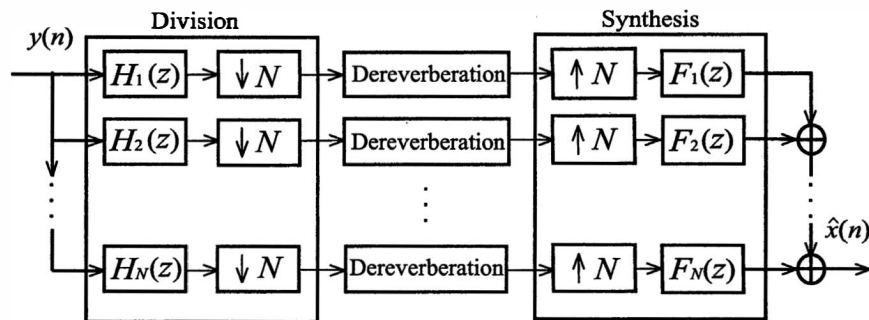


図 5 帯域分割処理を用いた処理の概略図

Fig. 5 Sub-band processing and dereverberation.

$F_k(z)$  ( $k = 1, 2, \dots, N$ ) はそれぞれ帯域抽出のための帯域フィルタである。

なお本回復手法は、信号計測開始直後の信号を音源信号と仮定しているため、特に過渡的な音の成長過程において必要以上に相関抑圧処理を行うと、残響の影響が小さいため、声帯振動に伴う周期信号同士の相関係数が大きく、原信号を逆にひずませる。この現象に対し抑圧処理自体を抑制させるため、過去の比較対象フレームのエネルギーに対し、回復対象フレームのエネルギーが大きい場合のみ処理を行うことにする。

#### 4. 相関処理による残響回復実験

図 6 はリスニングルームやコンサートホールに対応するインパルス応答である。残響時間はそれぞれ (a) 約 0.5 秒、(b) 約 1 秒である。本評価実験では、0.5 秒以上の比較的長い残響時間を有する室内環境を想定しており、図 6 に示されるインパルス応答を用いることにする。本章では、音声に代表される時間的な周波数変動を伴う音源信号回復例として、まず TSP (Time Stretched Pulse) 信号 [17] に対して回復効果を確認し、次に音声信号に対して回復実験を行った。

##### 4.1 TSP 信号を用いた帯域幅の検証

残響抑圧効果を検証するため、TSP 信号を用いて音声のような調波構造を形成する信号を音源信号として回復実験を試みた。TSP 信号は時間軸に対し線形に周波数変化する単純な信号である。実験には、TSP 信号に図 6 (b) に示す残響時間約 1 秒のインパルス応答を畳み込んだ残響信号を用いた。サンプリング周波数を

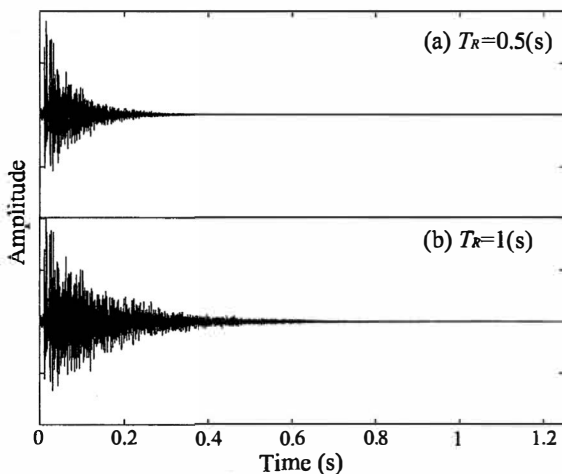


図 6 (a) 残響時間：約 0.5 秒、(b) 残響時間：約 1 秒の室内インパルス応答

Fig. 6 Room impulse responses for reverberation time (a)  $T_R = 0.5$  s, (b)  $T_R = 1$  s.

8 kHz とし、図 5 に示す等分割フィルタバンクにおいて  $N = 4$  として回復を行った。また、フレーム長を 64 ms とし、シフト長を 16 ms とした。分割後の帯域内に、連続的に変化するスペクトルをそれぞれ独立に回復処理を行った例を図 7 に示す。両図 (a), (b), (c) はそれぞれ原信号、残響信号、回復信号の各時間波形とスペクトログラムである。両図より明らかなように、帯域内にほぼ単独でスペクトルが存在している信号回復例 (図 7) の方が残響を軽減していることが分かる。しかし、図 8 のようにスペクトルの残響が重なっている場合においては、それぞれのスペクトルに相関が高くなるため、関係のないスペクトル成分と比較し、減算処理を行うことによって回復精度が悪くなる傾向がある。そのため実音声においても、帯域内のスペクトルが少ない方が回復精度が高いことが予想される。帯域内のスペクトル数を少なくするため、実音声の評価実験では、被験者のピッチ周波数を事前に統計的に調べ、フィルタバンクの帯域幅を定めることにした。

##### 4.2 評価指標

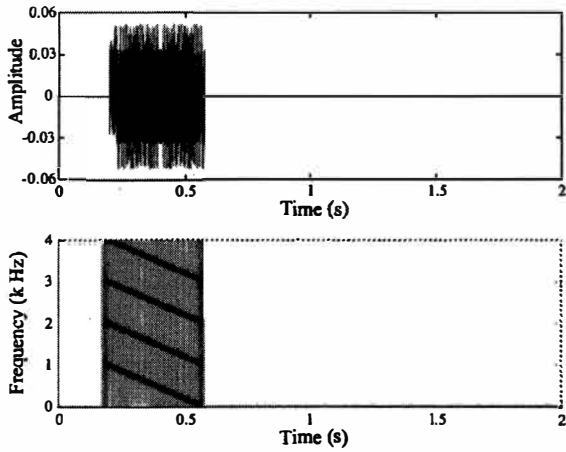
一般的に残響音声における残響感と音声の包絡線変化には高い相関があり、残響回復を定量的に検証するため信号包絡線に対する次式の二乗誤差を用いた。

$$I = 10 \log_{10} \frac{\int_0^T \{e_x(t) - e_y(t)\}^2 dt}{\int_0^T \{e_x(t) - \hat{e}_x(t)\}^2 dt} \quad (8)$$

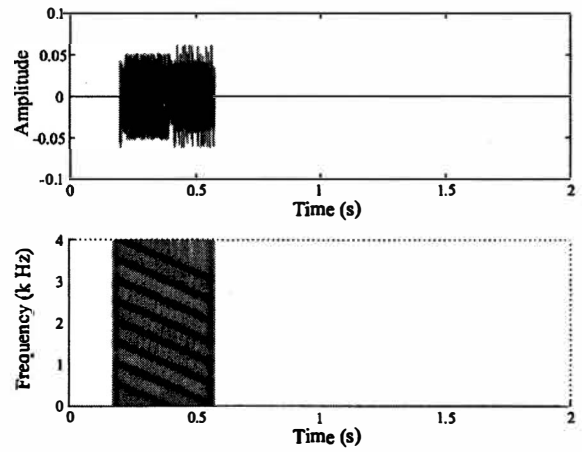
本論文では式 (8) をエンベロープ回復指標 (Improvement index of envelope distortion) と呼ぶことにする。 $e_x(t)$  は原音声、 $e_y(t)$  は残響音声、 $\hat{e}_x(t)$  は回復音声のエンベロープ、 $T$  は分析時間である。上式から明らかなように、回復効果が向上するにつれ、本回復指標値は大きくなる。本指標は、わずかな時間遅れを反映する厳密な指標ではないが、回復信号の大まかな回復効果を反映する指標である [5]。なお、エンベロープの抽出にはヒルベルト変換後、絶対値をとり、更に低域通過フィルタにより 10 Hz 程度の帯域制限を行った。

##### 4.3 回復アルゴリズムにおける比較フレーム数の最適化

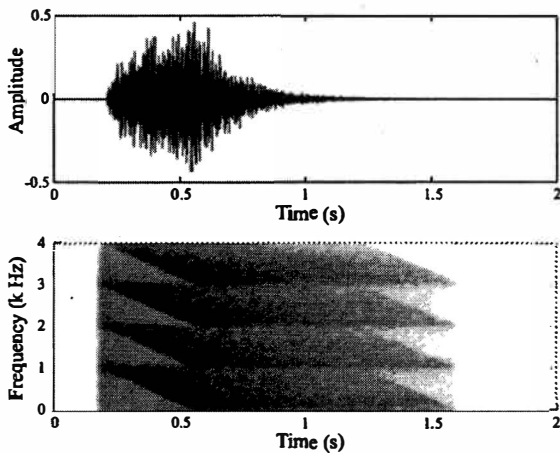
図 4 における回復対象フレームにおける過去のフレームとの比較フレーム数  $M$  について最適値を実験的に調べた。 $M$  が適正值以上に大きくなると回復対象フレームにおける減算処理が必要以上に多くなり、信号をひずませる要因となる。実験には、アメリカ人男女各 4 人の英語音声「A ship was torn apart



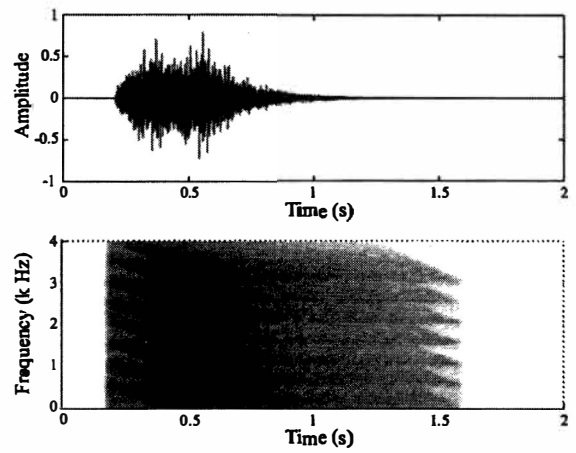
(a) TSP signal



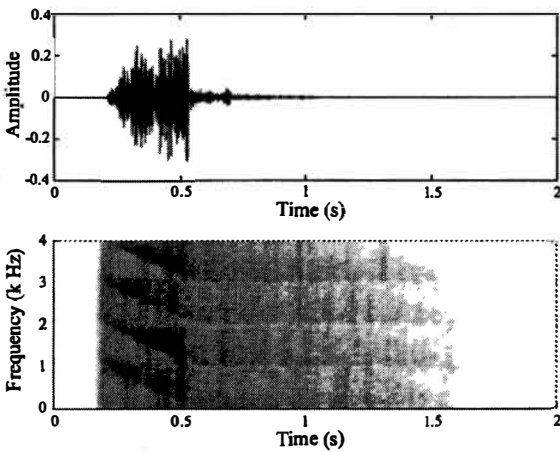
(a) TSP signal



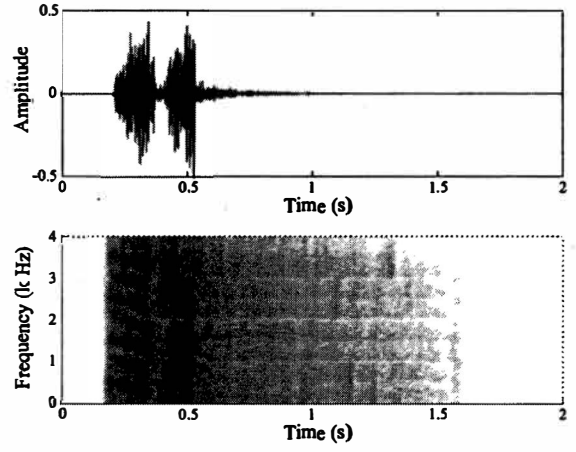
(b) Reverberant signal



(b) Reverberant signal



(c) Recovery signal



(c) Recovery signal

図7 波形とスペクトログラム (4 スペクトル)  
Fig.7 Waveforms and spectrograms (4 spectra).

図8 波形とスペクトログラム (8 スペクトル)  
Fig.8 Waveforms and spectrograms (8 spectra).

on a sharp reef] (約3秒)を用いた。サンプリング周波数を8kHzとし、帯域分割するフィルタバンクは前述したように、図5に示すものを用いた。図9に検証に用いた音声の基本周波数の平均値の分布を示

す。32分割以上の帯域幅 ( $4000/32 = 125\text{ Hz}$  以下)では、処理する帯域に無音帯が数多く存在するため、処理の効率化を図るため、暫定的に16分割 (帯域幅  $4000/16 = 250\text{ Hz}$ ) で帯域ごとに処理を行った。す

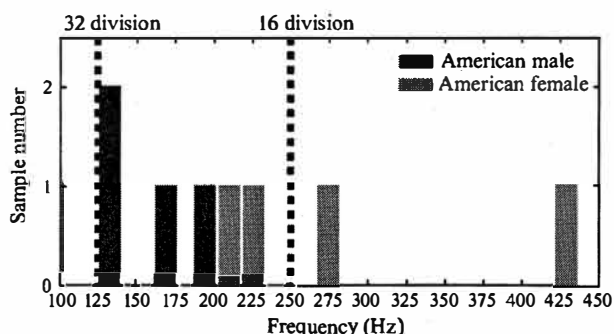


図 9 基本周波数の平均値の分布  
Fig. 9 Distribution of averaged pitch frequency.

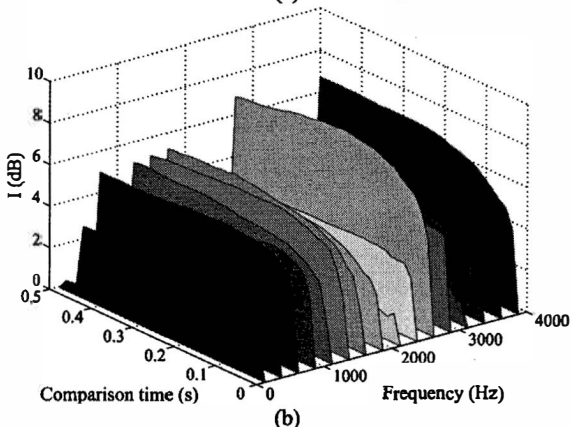
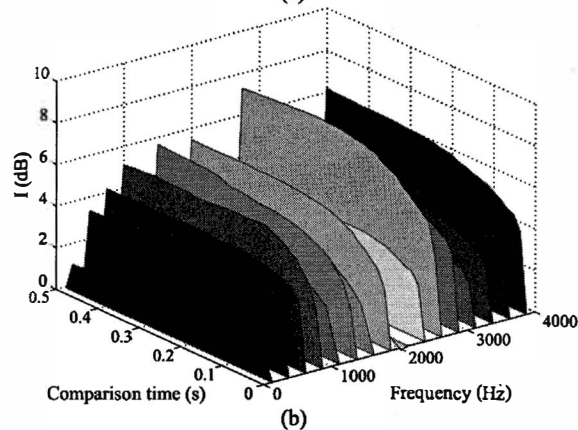
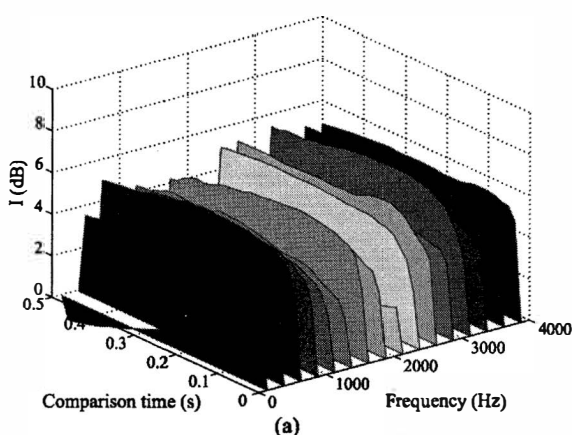
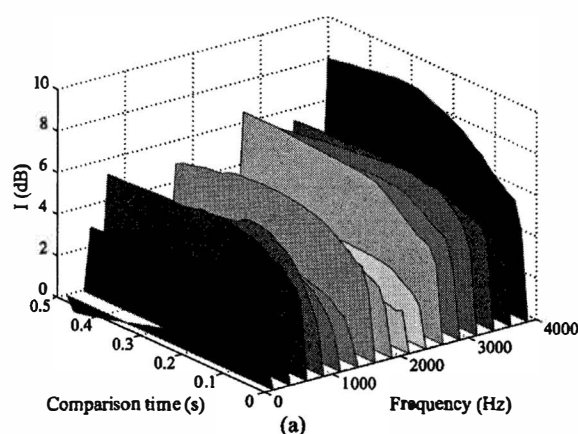


図 11 比較時間の長ささとエンベロープ回復指標 (残響時間約 1 秒) (a) アメリカ人男性, (b) アメリカ人女性  
Fig. 11 The comparison time and the recovery index ( $T_R = 1$ s): (a) American male, (b) American female.

図 10 比較時間の長ささとエンベロープ回復指標 (残響時間約 0.5 秒) (a) アメリカ人男性, (b) アメリカ人女性  
Fig. 10 The comparison time and the recovery index ( $T_R = 0.5$ s): (a) American male, (b) American female.

なわち、各帯域に 1 個ないし 2 個の倍音成分が含まれるようにした。回復処理するための微小区間のフレーム長を 64 ms、シフト長を 16 ms とした。図 10、図 11 はそれぞれ図 6 に示す残響時間約 0.5 秒、約 1 秒のインパルス応答を音源音声に畳み込んだ残響音声に対し回復実験を行い、そのエンベロープ回復指標を表示したものである。図 10(a)、図 11(a) は男声、

図 10(b)、図 11(b) は女声のエンベロープ平均回復度である。図 10 では回復対象フレームに対し、ばらつきはあるものの、図 12 に示すその平均値では残響時間の 1/5 に相当する過去との比較が最も良い。図 12 より、残響時間の 1/5 以上に長く比較すると、徐々に回復効果が減少する傾向が確認できる。また、図 10 に対しても、残響時間の 1/5 以上に比較フレームを用意すると、徐々に回復効果が減少する帯域が多い。このように、残響時間の 1/5 に相当する比較フレーム数を用意することで、より効果的な残響抑圧が確認できる。また、全体的に女声、男声とも同等な回復効果が確認できる。

残響時間約 1 秒のアメリカ人女性の残響音声の回復例を図 13 に示す。図 13(a)、(b)、(c) にそれぞれ音源音声、残響音声、回復音声の時間波形と、スペクトログラムを示す。図 13(c) は、図 12 において比較時間の最適と思われる 0.2 秒間として処理を行った。図より音源波形の概略が復元できることが分かる。表 1

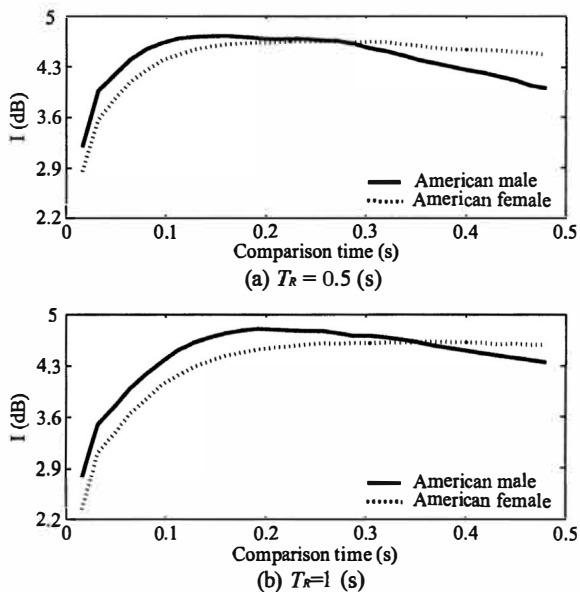


図 12 エンベロープ回復指標の平均値  
Fig. 12 Averaged recovery index of reverberant speech.

表 1 各帯域のエンベロープ回復値  
Table 1 Recovery index of sub-band processing.

Frequency bands (Hz)	Improvement index (dB)	Correlation value		
		recovery	reverberant	difference
		①	②	① - ②
0~250	3.5219	0.6926	0.5938	0.0988
250~500	4.1764	0.7554	0.6517	0.1036
500~750	5.1530	0.8654	0.5778	0.2876
750~1000	6.0755	0.8115	0.5057	0.3058
1000~1250	3.3686	0.7540	0.5472	0.2068
1250~1500	5.6924	0.8779	0.6219	0.2561
1500~1750	2.4188	0.8583	0.6802	0.1782
1750~2000	7.9419	0.8633	0.6049	0.2584
2000~2250	0.1109	0.7423	0.7481	-0.0058
2250~2500	4.6025	0.8875	0.7350	0.1525
2500~2750	8.0229	0.9299	0.7046	0.2254
2750~3000	2.7170	0.9082	0.7930	0.1152
3000~3250	4.9462	0.8531	0.5511	0.3020
3250~3500	3.4850	0.7416	0.4825	0.2590
3500~3750	3.7655	0.7738	0.5563	0.2175
3750~4000	7.8144	0.9077	0.6463	0.2613

の各項目に、それぞれ帯域ごとの式 (8) で定義された回復指標値、回復音声と音源音声のエンベロープの相関値①、残響音声と音源音声のエンベロープの相関値②、処理によって向上した相関値の差 ① - ② を示す。相関値とエンベロープ回復指標の増減がほぼ一致している。また、2000~2250 Hz の帯域で一部大きく改善できないものがあるものの、全体的にはほぼ同様な回復傾向にあり、平均で約 4.6 dB の回復が確認できた。

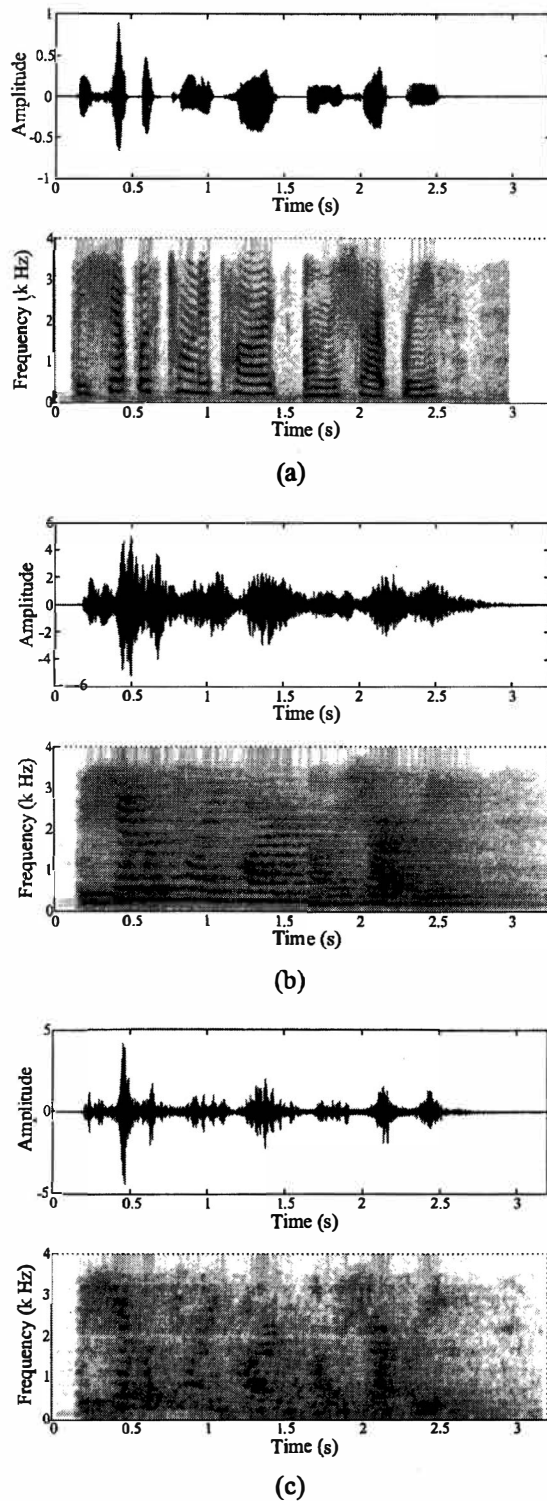


図 13 波形とスペクトログラム (a) 原信号, (b) 残響信号, (c) 回復信号  
Fig. 13 Waveforms and spectrograms: (a) source signal, (b) reverberant signal, (c) recovery signal.

## 5. むすび

本論文では室内インパルス応答の測定を必要としな



いブラインド逆問題について検討した。残響信号と音源信号との相関関数から、直接音と反射波の相対的な振幅情報と時間遅れを抽出して観測信号から逐次的な減算処理を施し、残響の影響を取り除き音源信号を推定するブラインド解法を提案した。実音声を用いた回復実験では、フィルタバンクを用いて広域信号を狭域化して、本提案処理を試みた。音源信号に、男女各4名の英語音声を用いた回復実験により、定量的にその回復効果を検証した。その結果、残響時間の1/5に相当する過去の時系列に対し、フレーム処理を行い、逐次減算することで、最も良い回復効果が得られることが分かった。全体的にほぼ同様な回復効果があり、最大8dB、平均4.6dBの改善が確認できた。TSP信号のように単純に周波数変動を伴う信号や、音声信号において回復効果が確認できた。

本提案手法は、室内インパルス応答を必要とせず、また、インパルス応答の概形にも余り影響されないため、例えば反射効果が異なる壁面を有する室内で観測される残響信号でも、回復できる可能性がある。今後は、本手法において帯域幅の変化やめいりょう度向上を評価する予定である。

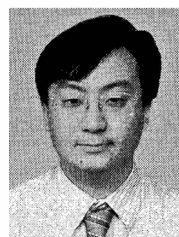
謝辞 本研究の一部は、テレコム先端技術研究支援センター研究助成金による。ここに御礼申し上げる。

## 文 献

- [1] S.T. Neely and J.B. Allen, "Invertibility of a room impulse response," J. Acoust. Soc. Am., vol.66, no.1, pp.165-169, 1979.
- [2] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-36, no.2, pp.145-152, 1988.
- [3] H. Nakajima, M. Miyoshi, and M. Tohyama, "Sound field control by indefinite MINT filter," IEICE Trans. Fundamentals, vol.E80-A, no.5, pp.821-824, May 1997.
- [4] 神沼充伸, 伊勢史郎, 鹿野清宏, "周波数領域における最小ノルム解を利用した多チャンネル音場再現システムにおける逆フィルタの設計," 音響誌, vol.57, pp.175-183, 2001.
- [5] 広林茂樹, 野村博昭, 小池恒彦, 東山三樹夫, "パワーエンベロープ伝達関数の逆フィルタ処理による残響音声の回復," 信学論 (A), vol.J81-A, no.10, pp.1323-1330, Oct. 1998.
- [6] 広林茂樹, 山淵龍夫, "帯域分割を用いたパワーエンベロープ逆フィルタ処理の残響抑圧効果," 信学論 (A), vol.J83-A, no.8, pp.1029-1033, Aug. 2000.
- [7] T. Houtgast, H.J.M. Steeneken, and R. Plomp, "Predicting speech intelligibility in room acoustics," ACUSTICA, vol.46, pp.60-72, 1980.
- [8] M. Unoki, M. Furukawa, K. Sakata, and M. Akagi, "A

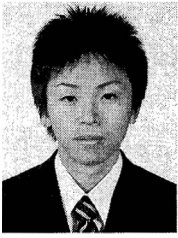
method based on the MTF concept for dereverberating the power envelope from the reverberant signal," Proc. ICASSP 2003, vol.I, pp.840-843, 2003.

- [9] 古川正和, 鶴木祐史, 赤木正人, "MTFに基づいた残響音声パワーエンベロープの回復方法," 信学技報, EA2002-15/SP2002-15, 2002.
- [10] 酒井恵吾, 鶴木祐史, 赤木正人, "MTFに基づいた残響音声の回復法の検討," 信学技報, SP2002-181, 2003.
- [11] 永田仁史, 劉 家祥, 安部正人, 城戸健一, "クロススペクトル法によるコヒーレンス関数及びインパルスレスポンス推定の際の窓関数の影響," 音響誌, vol.45, pp.950-956, 1989.
- [12] 福島 学, 城戸健一, "クロススペクトル法によるインパルスレスポンス推定における推定誤差の検討," 音響誌, vol.55, pp.265-274, 1999.
- [13] 田澤 徹, 大西 昇, 杉江 昇, "単一マイクロホンを用いた未知環境におけるエコーキャンセラー," 計測自動制御学会論文誌, vol.30, no.4, pp.460-466, 1994.
- [14] T. Hasegawa, S. Hirobayashi, T. Yamabuchi, and T. Yoshizawa, "Speech waveform recovery from a reverberant speech signal using a correlation coefficient," Proc. Third Asia-Pacific Conference on Industrial Engineering and Management Systems, pp.55-59, Dec. 2000.
- [15] H. Wang and F. Itakura, "Realization of acoustic inverse filtering through multi-microphone sub-band processing," IEICE Trans. Fundamentals, vol.E75-A, no.11, pp.1417-1483, Nov. 1992.
- [16] 真家仁志, マルチレート信号処理, pp.89-181, 昭晃堂, 東京, 1997.
- [17] 鈴木陽一, 浅野 太, 曾根敏夫, "音響系の伝達関数の模擬をめぐって (その2)," 音響誌, vol.45, pp.44-50, 1989.  
(平成15年9月26日受付, 16年5月10日再受付,  
7月20日最終原稿受付)



広林 茂樹 (正員)

平6工学院大工学研究科修士課程了。平7同大学院博士課程中退。同年4月金沢大学工学部助手。平11富山大学工学部講師。平12同大助教授。現在に至る。主として音響波動・振動論, 音響信号処理, 音場制御, 流体伝達系のモデル化に関する研究に従事。計測自動制御学会, 電気学会, 米国音響学会, 日本音響学会, IEEE各会員。



林 智也 (学生員)

平 15 富山大・工・知能情報卒。現在、同  
大大学院理工学研究科博士前期課程在学中。  
音響信号処理の研究に従事。



山淵 龍夫 (正員)

昭 40 東北大・工・通信卒，昭 42 同大  
大学院工学研究科修士課程了，昭 47 同大  
大学院工学研究科博士課程了。昭 47 富山大学  
工学部電気工学科助手，昭 54 同講師，昭  
58 同助教授，平元学科改組により電子情  
報工学科助教授，平 2 同教授，平 9 学科  
改組により知能情報工学科教授，現在に至る。工博。有限要素  
法を用いた圧電振動子や音場などの解析に関する研究に従事。  
IEEE member, 電気学会, 日本音響学会, 日本シミュレーショ  
ン学会, 情報処理学会各会員。



吉沢 寿夫 (正員)

昭 46 富山大・工・電気卒，同年タケダ  
理研工業(株)(現在，アドバンテスト)入  
社，昭 50 富山大学工学部教務補佐員，文  
部技官を経て現在，富山大学工学部知能情  
報工学科助手。電磁界，音場などの有限要  
素法やスパッタ薄膜生成のシミュレーショ  
ンに関する研究に従事。日本シミュレーション学会会員。