社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.
EA93-81 (1993-12)

# 最小位相成分とパワーエンベロープに着目した残響音場における音源波形回復

広林茂樹*　小池恒彦**　東山三樹夫***
Richard.H.Lyon ****

*,*** 工学院大学，　** NTTアドバンステクノロジ株式会社
**** マサチューセッツ工科大学

〒163-91 東京都新宿区西新宿 1-24-2 工学院大学 数理音響学研究室

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，音による機械故障診断や通信における音声明瞭度向上などのために重要である。本報告は，残響音場の伝達関数を測定することなく，ケプストラム処理によって音源波形を回復する可能性を検討したものである。音源信号を，パルス波，パルス列，さらに音声波形と分類し，それぞれに残響波形からの波形復元を試みた。その結果，パルス波や周期的なパルス列からなる音源信号は，最小位相成分のケプストラムだけを用いて波形の概略が復元できることがわかった。また，音声信号に対しては，複素ケプストラム処理による音源波形に加えて，信号のパワーエンベロープに着目した逆フィルタ処理が有効である見通しを得た。

**和文キーワード**　ケプストラム処理，音響信号処理，逆フィルタ処理，室内音響，機械診断

# Source Waveform Recovery from Reverberant Minimum Phase Components and Power Envelope

Shigeki HIROBAYASHI *　Tsunehiko KOIKE **　Mikio TOHYAMA ***
Richard H.Lyon ****

*,*** Kogakuin University,　** NTT Advanced Technology Corporation
**** Massachusetts Institute of Technology

Institute of Computer Acoustics & Hearing, Kogakuin University,
1-24-2, Nishi-Shinjuku, Shinjuku Tokyo, 163-91 Japan

**Abstract**　Source waveform recovery under reverberant conditions is quite important for machine diagnostics and speech quality enhancement. This article investigates cepstral dereverberation without using reverberant transfer functuins. Source waveforms are typically modeled here by a pulse-like signal, pulse-trains, and speech signals. The dereverberantion experiments demonstrate that a pulse-like waveform and a periodic pulse train are recovered from the reverberant minimum phase cepstrum without all-pass components. All-pass components do not include important signature elements for those source signals. However, aperiodic pulse trains and speech waveforms convey their signatures also in their non-minimum phase components. The recovery for those waveforms requires their all-pass components. The authors also describe how a power envelope inverse filtering method shows promise for speech waveform recovery.

**英文 key words**　Cepstrum, Acoustic Signal Processing, Inverse Filtering, Room Acoustics, Machine Diagnostics

# 1　はじめに

本研究は，残響音場伝達関数の正確な逆フィルタ処理を施すことなく，音源波形の特徴を回復することを目的とする。室内伝達関数は，一般には非最小位相特性を有するため，観測点の数を多点に拡張しない限り，正確な逆フィルタ処理は不可能である[1]。本報告では，室内の伝達関数を測定することなく，ケプストラム処理と残響時間から音源波形の概略を復元することを述べる。

ケプストラム処理の音響機械診断への応用は，本報告の共著者であるR.Lyonによって提唱されている[2]。一般に機械診断では，音源信号の変化を識別するため，音源信号から離れた観測点にて観測される音響振動伝達系の残響を伴う信号波形から，残響の影響を取り除いて，音源波形を復元することが望まれている。筆者らは既に，音源信号が最小位相の零点で構成されているならば，複素ケプストラムを用いることなく，最小位相ケプストラムから復元されること，さらに，最小位相ケプストラムには，残響波形観測点による変動が小さいことを，孤立パルス音源波形と室内伝達関数を例にとり実証した[3-5]。室内残響の伝達関数は多くの"All−Pass"成分を含むため，残響信号から最小位相成分を抽出するだけで，残響音の影響をある程度軽減することができる。

しかし，機械診断では，孤立パルス波形に加えて，対をなすパルス波形の組，あるいは，複数のパルス列などの音源波形復元も重要である。そこで本報告では，音源信号の例として，パルス波，パルス列，さらに連続信号の例として，機械振動ではなく音声波形をとりあげ，それぞれに最小位相成分に着目した残響波形からの波形復元を試みる。音声波形の復元は，通信における明瞭度向上のために重要であり，また，復元手法に機械振動の音源波形復元と同一の理論的枠組みが期待できる[6]。その結果，周期的なパルス列が孤立パルス波形と同様に最小位相ケプストラムから復元されること，さらに音声波形の復元では，非最小位相成分を含む複素ケプストラム処理に加えて，長い時間窓で観測した信号のパワーエンベロープ逆フィルタ処理の有効性などを明らかにする[7]。

# 2　ケプストラム処理と音源波形復元

## 2.1　ケプストラムの原理[8]

原信号を s(t)，室内インパルス応答を h(t)，残響信号を r(t) とすると，時間領域では畳込みの関係

$$r(t)=s(t)*h(t) \tag{1}$$

が成立し，周波数領域では，

$$R(\omega)=S(\omega)\ H(\omega) \tag{2}$$

のようにそれぞれの積で表される。従って，上式において両辺の複素ケプストラムをとれば，

$$\begin{aligned} F^{-1}\{\ln R(\omega)\} &= F^{-1}\{\ln S(\omega)+\ln H(\omega)\} \\ &= F^{-1}\{\ln S(\omega)\}+F^{-1}\{\ln H(\omega)\} \end{aligned} \tag{3}$$

となり原信号と室内インパルス応答のケプストラムの和となる。但し，ここで，右辺第1項は原信号のケプストラムであり，第2項はインパルス応答のケプストラムである。また，$F^{-1}$ は逆フーリエ変換を示す。実時間関数の複素ケプストラムは実関数となる。

一般に，室内インパルス応答のケプストラムは，ロータイムからハイタイムに広く分布する。一方，音源信号自身のケプストラムは，ロータイム部に集中する。従って，ケプストラムのロータイムウインドウイング処理により，室内残響の影響を軽減して，音源波形の特徴を抽出復元することが期待できる。

## 2.2　最小位相ケプストラム処理

室内伝達関数は，一般に非最小位相特性を有している。そのため，室内インパルス応答は，All−Pass成分を含んでいる。伝達関数 H(ω) は，

$$H(\omega)=H_{min}\ H_{ap} \tag{4}$$

のように，最小位相成分とAll−Pass成分の積で表すことができる。ここで，$H_{min}$は，最小位相成分を表し，$H_{ap}$は，All−Pass成分を示し，$|H_{ap}|=1$ である。ケプストラムは，最小位相とAll−Passのそれぞれの成分の和として表される。

音源信号が最小位相成分だけからなると仮定する。残響応答から観測されるケプストラムのAll−Pass成分は，音源信号ではなく室内伝達関数に起因するものである。従って，観測したケプストラムの最小位相成分のみに音源情報が含まれ，最小位相ケプストラムのロータイムウインドウイングから最小位相音源信号の特徴が抽

出される。最小位相ケプストラムと All−Pass ケプストラムの関係を図 1 に示す。

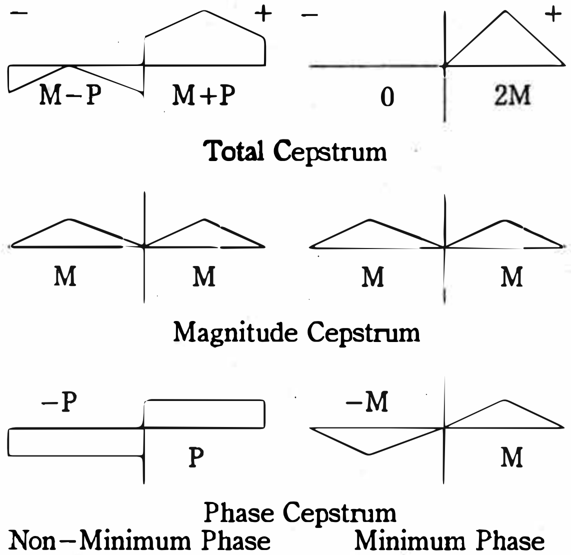



図 1　非最小位相と最小位相のケプストラム

## 3 パルス音源波形の復元 (3−5)

本報告で述べる復元実験は，室内で実測したインパルス応答（サンプリング周波数 :8kHz，データ点数 :4096 点）に基ずく計算機実験によるものである。図 2a は本実験に用いたインパルス応答である。このインパルス応答に音源信号を畳み込むことによって得た残響応答から，音源波形の特徴を復元することを試みる。

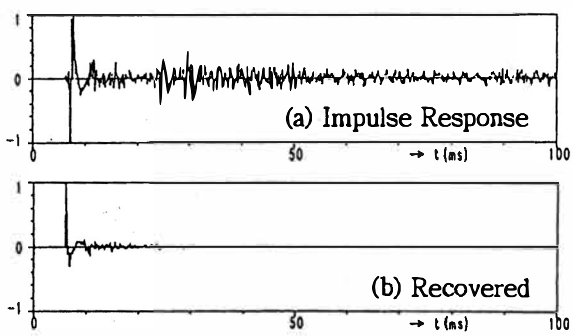



図 2　インパルス音源波形の回復

### 3．1 インパルス波形の復元

最小位相特性を持つ単純な音源波形の例として，理想的なインパルス波形がある。図 2a のインパルス応答は，そのような理想的なインパルスを音源とする残響応答波形と考えられる。従って，図 2b に示すように，最小位相ケプストラムのロータイムフィルタリング（カットオフタイム :32ms，窓関数 : ハニング窓）により，インパルス応答の波形が時間的に圧縮されて，インパルス音源波形の概略を復元できる。

### 3．2 周期的なパルス列の復元

音源波形が孤立波ではなく，あるタイミングで繰り返す場合が機械診断などでは考えられる。そこで，複数個のインパルス波形が等しい時間間隔で周期的に繰り返す場合を考える。図 3a にその例を示す。図 3b はこの音源信号に図 2a に示すインパルス応答を畳み込んだ残響波形である。

図 3c は，インパルス波形の復元と同様に最小位相ケプストラムから復元した音源波形である。周期的なパルス列の特徴が復元されていることがわかる。ケプストラムのカットオフタイムは図 3a のパルス列の継続時間を考慮して 32ms，窓関数はハニング窓である。

周期的なパルス列が最小位相ケプストラムから復元されることは，音源信号のもつ零点配置よって実証される。周期を T とした n+1 個のインパルス列の零点は，

$$S(z)=(1-z^{-nT})/(1-z^{-T}) \qquad (5)$$

の根である。この根は図 3d に示すようにすべて単位円周上にある。従って，最小位相ケプストラムからの音源波形抽出が可能である。

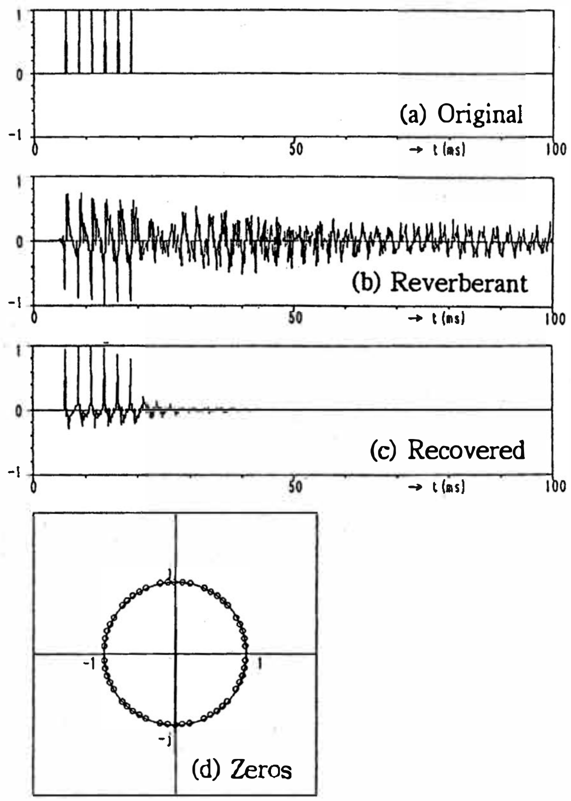



図 3　周期的なパルス列の復元と零点配置

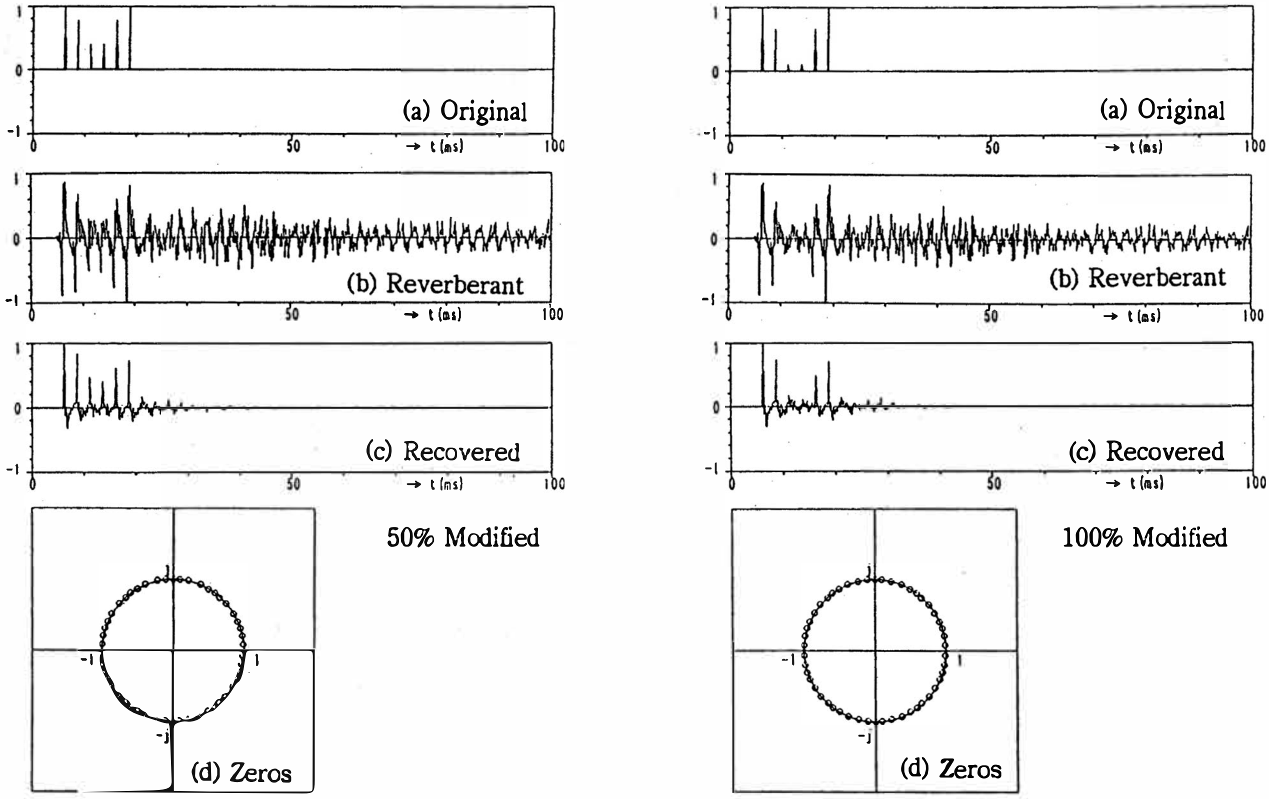



図 4　変調された周期的なパルス列の復元と零点配置

## 3．3 振幅変調された周期的パルス列の復元

正弦波で振幅変調された周期的パルス列も同様に最小位相ケプストラムから復元される。図 4 は，変調パルス列 (a)，残響波形 (b)，回復波形 (c)，音源信号の零点配置 (d) をそれぞれ変調度 50% と 100% の場合について示したものである。変調の効果は零点の分布を一様にさせることにあるが，いずれの零点も単位円周上に存在する。従って，最小位相ケプストラムから音源信号が復元される。

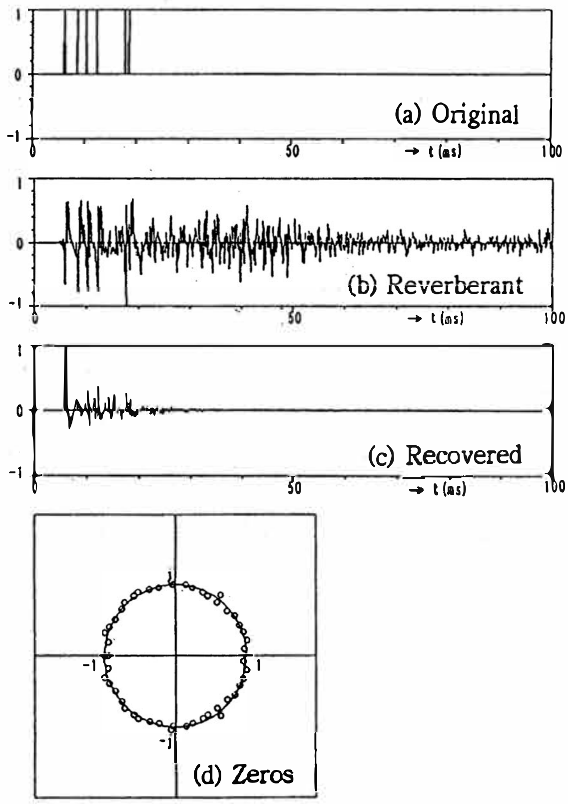



図 5　非周期的パルス列と零点配置

## 3．4 非周期的なパルス列の復元

図 5 は，不等間隔で繰り返すインパルス列の例 (a)，残響波形 (b)，回復波形 (c)，音源信号の零点配置 (d) である。零点は単位円の内外にも分布していることがわかる。従って，図 5(c) に示される通り，最小位相ケプストラムだけでは，この音源波形の特徴を復元できない。不等間隔で繰り返すインパルス列の復元には，非最小位相成分が必要である。

室内インパルス応答は一般に All−Pass 成分

を含んでいる。そのため，最小位相成分のみを取り出すだけでも残響の影響を観測信号から軽減することは可能である。しかし，音源自身の重要な特徴がその非最小位相成分によって表される場合には，最小位相ケプストラムのみによって音源波形を回復することは困難である。非周期的なパルス列はその一例である。

## 4 残響音声からの音声抽出

音源波形回復のもう一つの重要な応用に，残響音声の明瞭度向上がある。残響音の影響をケプストラムによって軽減する試みもいくつか既に行われているが[9]，未だ研究途上の段階にある。ここでは，最小位相ケプストラムからの回復の問題点を明らかにするとともに，音声のパワーエンベロープに着目した回復について述べる。

### 4．1 最小位相ケプストラムからの回復

残響を含む室内のインパルス応答は，(2)式の関係より音源信号のスペクトルを歪ませる。このスペクトル歪みは，(3)式に基づいてケプストラム処理によって軽減される。図6は，インパルス応答にケプストラムのロータイムウインドウイングを施すことによって伝達特性が平滑化される様子を示したものである。

音声の音韻の特徴はその短時間スペクトルの変化に現れる。音韻の継続時間は，一般に室内残響の時間，すなわち室内インパルス応答の継続時間に比べれば非常に短い。しかし，音声を定常的な信号であると考えると，音声の短時間スペクトルにも(2)式が成立し，短時間窓で分析したケプストラム処理によって残響の影響を軽減することが可能である。

音韻の特徴は波形よりもそのパワースペクトルにあると言われている。パワースペクトルは信号の最小位相成分のみによって表される。従って，音声信号の回復にも残響成分を多く含むAll−Pass成分を用いることなく，最小位相成分のみによる回復が期待できる。

図7は短時間区間の最小位相成分から回復を試みた一例である。室内インパルス応答は，図2aに示した通りであり，音源信号は「爆音が銀世界の高原に広がる」という女性の声の一部である。この結果を見ると最小位相成分からは，音声信号の特徴が回復されないのがわかる。これは，音声波形を変調されたパルス列と考えると，音声のピッチ変動によってパルス列が非周期的な繰り返しとなるためと考えられる。すなわち，音声の回復には，All−Pass成分が必要である。

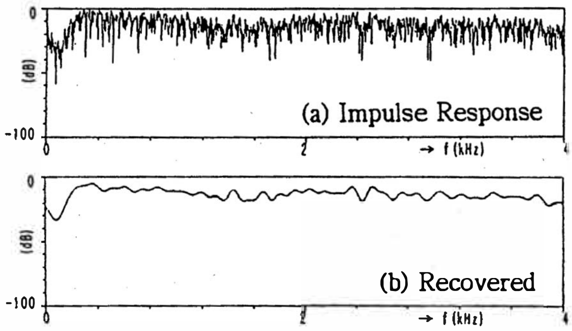



図6 伝達関数の平滑化

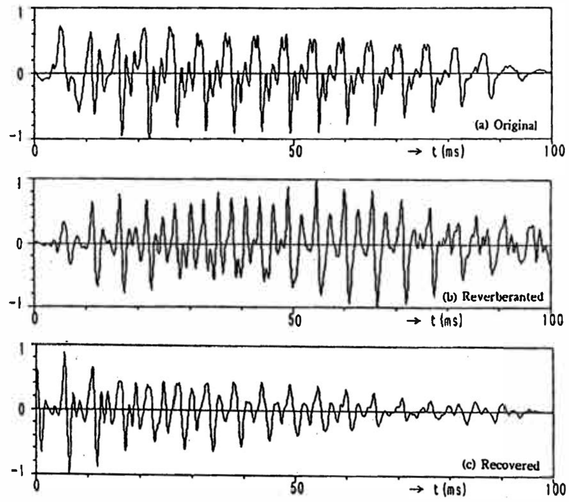



図7 最小位相ケプストラムからの回復結果

### 4．2 複素ケプストラム処理によるスペクトル歪みの回復

図8は図7と同様に短時間分析区間をとり，複素ケプストラム処理によって残響音によるスペクトル歪みを回復した例である。音声波形の全区間では，パワースペクトル(dB値)の原音声からの差の二乗平均値でスペクトル歪みの回復効果を評価するとおよそ1.0dB改善された。図7(c)の回復音声波形と比べると，All−Pass成分によって音声の特徴が保存されているのがわかる。しかし，音声の特徴はその短時間スペクトル構造だけでなく，音声のパワー変動に対

応したパワーエンベロープにも現れる。短時間分析によるケプストラム処理では，音声のパワーエンベロープから残響音の影響を軽減することは難しい。

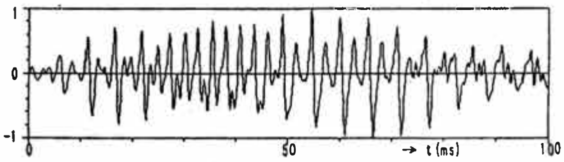



図 8　複素ケプストラムからの回復結果

## 5　パワーエンベロープの回復

近年，Houtgast [10] らによって提案されている MTF による音声明瞭度の予測理論は，音声のパワーエンベロープの歪みを評価するものである。パワーエンベロープの歪みが室内の残響音による影響である場合には，MTF は室内のインパルス応答から求めることができる [11]。そのため，王等 [7] は，室内伝達関数を用いて正確に原音声のパワーエンベロープを残響音声から回復する逆フィルタ理論を提唱実証している。ここでは，王等 [7] の結果を利用し，室内残響時間を用いて簡便にパワーエンベロープを回復すること試みる。

### 5．1　残響音声のパワーエンベロープと逆フィルタ処理

音声のパワーエンベロープの変動は音声信号自身の変動に比べて非常に遅い。そこで，MTF の理論に合わせて，音声信号を

$$s(t)=e_s(t)\ n_1(t) \quad (6)$$

のようにモデル化する。ここで $n_1(t)$ は白色雑音，$e_s(t)$ はゆっくり変動する包絡信号である。室内インパルス応答を上式 (6) のように

$$h(t)=e_h(t)\ n_2(t) \quad (7)$$

$$e_h(t)=a\,e^{-6.9/RT},\ RT は残響時間 (s) \quad (8)$$

と簡略化する [11]。(6) 式同様 $n_2(t)$ は，白色雑音，$e_h(t)$ はゆっくり変動する包絡信号である。但し，白色雑音は互いに無相関で，

$$<n_k(t)\ n_k(t+\tau)>=\delta(\tau),\ k=1,2 \quad (9)$$

が成立し，<*> は集合平均を表す。

ここで，上式 (6)(7) 式の畳み込みで表される室内残響音声からそのパワーエンベロープ（二乗エンベロープ）を求めると

$$e_r^2(t)=e_s^2(t)*e_h^2(t) \quad (10)$$

$$e_h^2(t)=a^2\ e^{-13.8/RT} \quad (11)$$

となり，原音声信号と室内インパルス応答のそれぞれのパワーエンベロープの畳み込みで記述することができる。従って，室内インパルス応答のパワーエンベロープの逆特性を用いることによって音声のパワーエンベロープを回復することができる。室内インパルス応答のパワーエンベロープを単純に指数関数で近似すれば，その逆フィルタは室内の残響時間から求めることができる。

### 5．2　パワーエンベロープの回復結果

図 9 は「爆音が銀世界の高原に広がる」という女性音声の，(a) 原音声，(b) 残響音声，(c) 回復音声を示したものである。パワーエンベロープ逆フィルタは，図 2a に示したインパルス応答から残響時間を 0.75 秒として実現し，分析時間窓長は 3.0 秒とした。波形回復は，フレームごとにまずパワーエンベロープを推定する。次に，推定されたパワーエンベロープと残響音声のパワーエンベロープとの比の平方根をとり，それを残響音声に掛け合わせる。全体の回復波形は，区間ごとに回復された波形を重ね合わせつないだものである。図 10 はソナグラム分析結果である。この結果から，パワーエンベロープの簡易逆フィルタ処理によって，音声のパワーエンベロープの特徴が概ね回復されることがわかる。しかし，包絡線処理によって音韻の一部に歪みが生じ，音質には改善すべき問題が多くある。

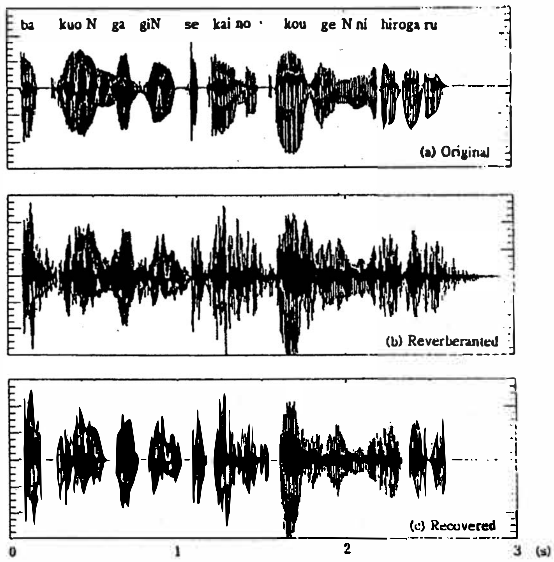
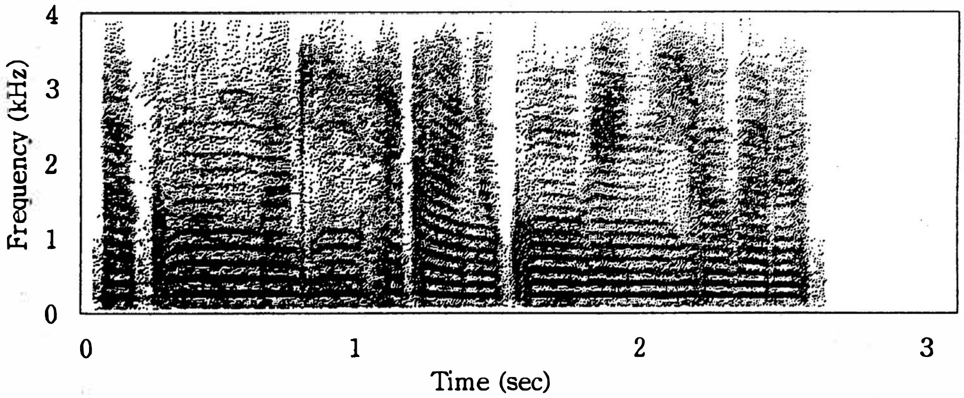
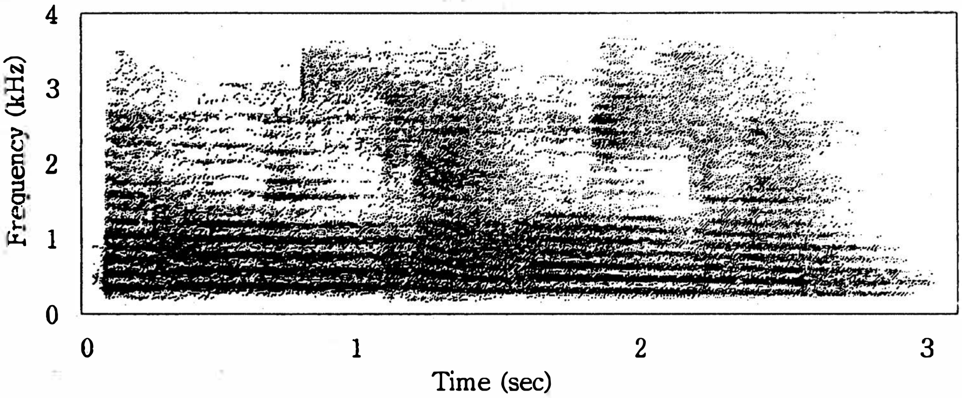


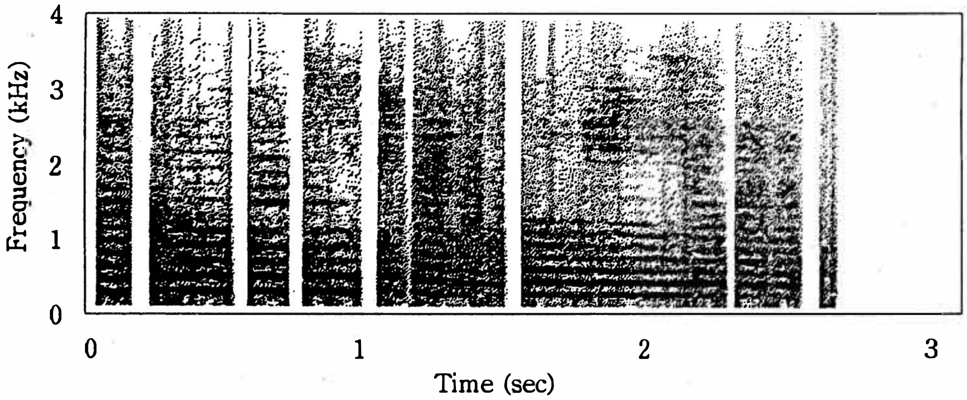

図 9　パワーエンベロープからの回復結果




(a) Original Speech




(b) Reverberant Speech (RT=0.75s)




(c) Recovered Speech

図 10 Sound Spectrogram

## 6 あとがき

本報告では，室内伝達関数の逆フィルタ処理を施すことなく音源波形を回復するため，ケプストラム処理について検討した。機械の故障診断などで着目されているパルス波形の回復では，複素ケプストラムではなく， All−Pass 成分を含まない最小位相ケプストラムから音源波形が回復できることを明らかにした。

しかし，非周期的に繰り返すパルス列の回復には，複素ケプストラムが必要であり，ピッチの揺らぐ音声信号もまた，最小位相ケプストラムでは回復できないことがわかった。また，音声信号の回復ではケプストラム処理に加えて音声のパワーエンベロープの逆フィルタ処理を検討し，回復の見通しを得た。本実験を進めるにあたり，室内インパルス応答のデータを御提供頂いた早大理工学総合センタの山崎教授に厚くお礼申し上げる。また，本研究の一部はサウンド技術振興財団からの研究助成による。重ねて厚くお礼申し上げる。

# 参考文献


(1) M.Miyoshi and Y.Kaneda, "Inverse filtering of room acoustics", IEEE Trans. ASSP. 36 145−152 (1988)

(2) R.Lyon, "Machinery Noise and Diagnostics", Butterworth Publishers (1987)

(3) M.Tohyama, R.Lyon and T.Koike, "Pulse Waveform Recovery in a Reverberant Condition", J. Acoust. Soc. Am. 91(5) 2805−2812 (1992)

(4) M.Tohyama, R.Lyon and T.Koike, "Inverse Filtering for Reverberant Transfer Functions", IUTAM Symposium '92 39−48 Springer−Verlag (1992)

(5) M.Tohyama, R.Lyon and T.Koike, "Reverberant Transfer Functions and Cepstrum Dereverberation", J. Acoust. Soc. Am. 92(4) part 2 1pPA64 (1992)

(6) M.Tohyama, R.Lyon and T.Koike, "Inverse Filtering for Reverberant Transfer Functions and Cepstrum Dereverberation" ICASSP'93 157−160 (1993)

(7) H.Wang and F.Itakura, "Realization of Acoustic Inverse Filtering through Multi−Microphone Sub−band Processing", IEICE Jpn, Trans. E75A 1474−1483 (1992)

(8) A.V. Oppenheim et al., "Nonlinear Filtering of Multiplied and Convolved Signals", Proc. IEEE 56 1264−1291 (1968)

(9) Y.Yamaski et al, "An Application of Homomorphic Filtering to Wide−range Audio Signals", Proc. Autumn Meeting of Acoust. Soc. Jpn. 315−316 (1982)

(10) T.Houtgast, H.J.M.Steeneken and R.Plomp, "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I.General Room Acoustics", ACUSTICA, 46, 60−72 (1980)

(11) M.R.Schroeder, "Modulation Transfer Functions: Definition and Measurement", ACUSTICA, 49, 179−182 (1981)