

デジタルオーディオ信号のサンプル欠けの位置推定と修復

著者	高橋 弘太
雑誌名	電子情報通信学会論文誌. D, 情報・システム
巻	J94-D
号	7
ページ	1088-1100
発行年	2011-07-01
URL	http://id.nii.ac.jp/1438/00008895/

デジタルオーディオ信号のサンプル欠けの位置推定と修復

高橋 弘太^{†a)}

Detection and Restoration of Lost Samples in Digital Audio Data Streams

Kota TAKAHASHI^{†a)}

あらまし デジタルオーディオ信号データの転送の際、クロック同期の設定を誤ると、サンプルの欠けや重複を生じてしまう。本論文では、このうちサンプル欠けに焦点を絞り、修復を行う方法について述べる。まず、同期ミスによって起こるサンプル欠けの型を分類した後、一つの評価関数で、サンプル欠けの位置推定、サンプル欠けの型の判定、失われた信号値の復元を行う方法を提案する。次に、評価関数を得る方法として、線形予測誤差を規範とする方法、スパース性を規範とする方法、エントロピーを規範とする方法の三つを定式化する。最後に、それら三つを実験的に比較し、エントロピーを規範とする方法が最も優れていることを示す。

キーワード 消失サンプル, AR モデル, 線形予測法, スパース性, 最小エントロピー

1. ま え が き

デジタルデータを転送する際、送信側と受信側の同期クロックを別クロック源から取ってしまうと、クロック周波数の差異により、1 時間に数個から数十個程度のサンプル欠けやサンプル値重複を生じてしまう。

このような障害が発生したとき、クロックの設定を正常に直してデータを転送し直せばよいが、ライブ演奏の録音など、録り直しが不可能な場合もある。

同期ミスのような初歩的なミスは、プロのエンジニアはまず起こさないだろう。しかし、経験の未熟な技術者がデジタル機器を使って録音する機会が増えてきているので、同期ミスはいつでも起こり得る。本論文は、筆者の研究室内で実際に起こった同期ミスの修復を目的に開発したアルゴリズムを、他の方法との性能比較を含めてまとめたものである。

2. 本論文で取り扱う問題と損傷型の分類

ノイズに汚されたデータや、欠損したデータの修復技術は、信号処理分野の重要な研究テーマである。

最もよく研究されているのは、汚染前の原信号に対

して、それが AR (自己回帰) モデルによって生成されたという仮定をおく方法 [1], [2] であろう。汚染時区間の推定は、AR モデルの励起信号がしきい値を超えたことでも判定できるし、もっと正確に、MAP 推定によって決定する [3] こともできる。汚染時区間が特定できれば、信号値の復元は、AR モデルによる時系列生成で実現できる。更に、Gibbs サンプラによって失われた時区間のデータを発生させることで、より典型的な修復値系列を得ようとする方法 [2] も提案されている。

また、別な方法として、欠損前の信号がナイキスト周波数未満の周波数に帯域制限されているという先験的知識が得られている場合には、Papoulis-Gerchberg の反復法 (いわゆる超解像で使われるアルゴリズム) により信号を修復する方法 [4], [5] も知られている。

これらの信号修復法に対して、本論文で論じる修復法は、前提において以下の差異がある。

第一に、同期ミスによるデータ欠けは、1~2 サンプル程度の欠損に限られる。これは、上のような一般的な方法が数サンプル~数百サンプルにわたる欠損値をも復元しようとしているのとは根本的に異なる。

第二に、同期ミスによるデータ欠けは、サンプル値の欠損だけでなく、後続の値が前倒しにシフトされて詰められる現象を伴う。これは、上の修復法の多くが、加法的な雑音を前提としているのとは根本的に異なる。更に、後で述べるように、同期ミスの場合、欠けの直

[†] 電気通信大学大学院情報理工学研究所, 調布市
Graduate School of Informatics and Engineering, The
University of Electro-Communications, 1-5-1 Chofugaoka,
Chofu-shi, 182-8585 Japan

a) E-mail: kota@ice.uec.ac.jp

後に信号値重複が生じるなど独特の構造をもつ。

第三に、デジタルレコーディングを行うとき、多くの機器では、ほぼナイキスト周波数まで目一杯信号を収録するので、Papoulis-Gerchberg 法は使えない。

以上より、同期ミスに特化した信号修復法を論文としてまとめておくことには意味があると考えた。今回は、AR モデル法をアレンジした方法 (3.1) に加え、スパース性規範に基づく方法 (3.2)、エントロピー規範に基づく方法 (3.3) の三つの方法を定式化し、それらの性能を比較した。

ところで、送信側のクロックに対し、受信側のクロックがわずかに速い場合は、受信系列上に同じ数値が連続して出現することを手掛りに欠陥箇所を容易に絞り込むことができるし、修復もそのサンプル値を取り除くだけであるから、問題としては容易である。そこで、本論文で扱う問題を、送信側のクロックに対し、受信側のクロックがわずかに遅い場合に限定することにする。すなわち、サンプル欠けの場合である。簡単のため、以下、サンプル欠けのことを「損傷」と呼ぶ。

さて、サンプル番号 n に対し、送信側データを x_n とし、受信側データを y_n としよう。 $n = n_0$ で最初の損傷が起こった、すなわち、 x_{n_0} が失われたとしよう。 $n = n_0$ 直後は、送信側がデータを送り出す時刻と、受信側がデータを受ける時刻がほぼ同一であるため不安定な状態がある長さ続く。その長さを n_S としよう。すなわち、 $x_{n_0+n_S}$ 以降は次の損傷位置まで正常に転送が行えるが、それまでは不安定としよう。このとき、 $y_{n_0-1}, y_{n_0}, \dots, y_{n_0+n_S-1}$ として受信されるデータが送信データのどの部分に相当するかの対応関係の可能性を、 $n_S = 5$ の場合について、図 1 に列挙する。図中の Lost は、データが失われることを示し、D は同じデータが 2 回重複することを示している。

$y_{n_0+1}, y_{n_0+2}, y_{n_0+3}$ の三つのデータがそれぞれに近接する x_n の前後どちらの値を採用するかによって、信号の損傷は 8 通りのパターンに分けられる。本論文では、これを、M1 から M8 という損傷型名で呼ぶことにする。なお、 n_S を更に大きくとれば、このパターン数は更に増大する。 n_S は、送受信のクロック周波数の近接度とジッタの大きさによって決める必要がある。今回我々が対象とした機器では、M1 と M2 のみが生じていた。これを踏まえ、本論文では、これに多少余裕をもたせて $n_S=4$ での解析を行った。すなわち、M1 から M4 に絞った解析である。

ちなみに、本研究は AES3 や S/PDIF によるディジ

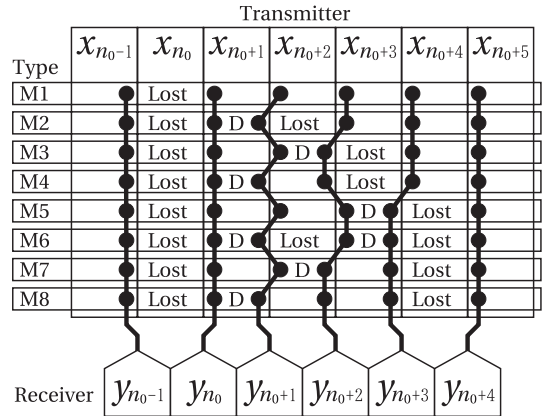


図 1 $n_S=5$ での損傷の八つの型 (M1~M8)
Fig. 1 The eight distinct types of lost samples.

タル伝送を想定しているが、これに限らず、バッファオーバーフローが 1 ワード単位で不確定な時刻で生じる伝送全般に適用可能である。なお、バッファのハードウェア構造によっては、1 ワード中の各ビットを構成するフリップフロップのクロック間の時間差により、前後のサンプルのビット値が錯綜してサンプル値自体が壊れてしまうこともあり得るが、本論文ではこのようなケースは扱わない。

3. 損傷の発見と値の復元の方法

解くべき問題は三つに分けられる。第一は、信号 y_n 上で損傷が発生している場所を発見することである。第二は、M1~M4 のどの型の損傷であるかを判定することである。第三は、失われた値 (図 1 において Lost と表示された x_n の値) の復元である。

本論文では、複数の規範間での比較を容易にするために、この三つの問題を規範ごとに決められた一つの評価関数 $L(n)$ で行うこととした。

まず、 y_{n-1} と y_n の間に最初の損傷があり、それを M1 型の損傷だとみなして Lost 値として z を挿入した場合の評価関数を、 $L_1(n, z)$ と書くことにする。M2~M4 についても、Lost 値を z_1, z_2 としたときの評価関数を、それぞれ、 $L_2(n, z_1, z_2)$, $L_3(n, z_1, z_2)$, $L_4(n, z_1, z_2)$ と書こう。これらを統合して、本手法で使う評価関数 $L(n)$ を次式のように定義する。

$$L(n) = \max \left\{ \max_z L_1(n, z), \max_{z_1, z_2} L_2(n, z_1, z_2), \max_{z_1, z_2} L_3(n, z_1, z_2), \max_{z_1, z_2} L_4(n, z_1, z_2) \right\} \quad (1)$$

発見・判定・復元は、以下のように行う。まず、 n を順にインクリメントしていきながら $L(n)$ を評価し、それが設定されたしきい値 C_L を超えたときに、地点 n を損傷位置として決定する。そして、損傷型を、 $L(n)$ の最大値を与える型 M_j と判定し、Lost 値を、 L_j を最大とする値 z (若しくは z_1, z_2) を用い復元する。

以上の枠組みを作ると、残すは評価関数を決めるのみである。以下、3 種類の評価関数を順に紹介する。

3.1 線形予測誤差を規範とする方法 (線形予測法)

本節では、AR モデル法による修復を、サンプル欠け問題用にアレンジする。なお、説明のしやすさを考え、AR モデルの代わりに線形予測モデルで記述する。

q 個の連続するサンプル値 $\xi = \{\xi_1, \xi_2, \dots, \xi_q\}$ を用意し、これに対する前向き k 次の線形予測を考える。 $\xi_{k+1}, \xi_{k+2}, \dots, \xi_q$ に対する予測誤差を、それぞれ $\delta_{k+1}, \delta_{k+2}, \dots, \delta_q$ と書き、予測係数を $\theta_1, \theta_2, \dots, \theta_k$ と書くことにすると、次式が成立する。

$$\begin{bmatrix} \delta_{k+1} \\ \delta_{k+2} \\ \vdots \\ \delta_q \end{bmatrix} = \begin{bmatrix} \xi_1 & \xi_2 & \dots & \xi_k \\ \xi_2 & \xi_3 & \dots & \xi_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{q-k} & \xi_{q-k+1} & \dots & \xi_{q-1} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_k \end{bmatrix} - \begin{bmatrix} \xi_{k+1} \\ \xi_{k+2} \\ \vdots \\ \xi_q \end{bmatrix} \quad (2)$$

上式を、 $\delta = G\theta - \zeta$ と書くことにすると、 δ のノルムを最小にする θ は、共役転置を「*」で表して、

$$G^*G\theta - G^*\zeta = 0 \quad (3)$$

の解となる。これを解くと、 $\theta = (G^*G)^{-1}G^*\zeta$ となるので、これを式 (2) に代入して、 δ について解くと、

$$\delta = (G(G^*G)^{-1}G^* - I)\zeta \quad (4)$$

が得られる。誤差のエネルギーの総和 $P = \delta^*\delta$ は、次式のように計算できる。

$$\begin{aligned} P &= \zeta^*(G(G^*G)^{-1}G^* - I)^*(G(G^*G)^{-1}G^* - I)\zeta \\ &= \zeta^*(I - G(G^*G)^{-1}G^*)\zeta \end{aligned} \quad (5)$$

ここで、 G^* を、 $G^* = U\Delta V^*$ と特異値分解すると、

$$P = \zeta^*(I - V\Delta U^*U\Delta^{-2}U^*U\Delta V^*)\zeta \quad (6)$$

$$= \zeta^*(I - VV^*)\zeta \quad (7)$$

となる。 P の大きさは、 ζ の要素数や要素の大きさに依存するので、これを規格化して、 ξ に対する k 次予

測の規格化残差 $Q(k, \xi)$ を次式のように定義する。

$$Q(k, \xi) = \frac{P}{\zeta^*\zeta} = \frac{\zeta^*(I - VV^*)\zeta}{\zeta^*\zeta} \quad (8)$$

今回、線形予測における修復法を決めるにあたり、 y_n を中心とする前後 D サンプルの時区間の範囲内で k 次の線形予測を行ったときの規格化残差 $Q(k, \xi)$ の、修正の前後での変化量を指標とすることにした。

修正の前後での ξ を定義しよう。M1 型の修復を行い、地点 n の直前に値 z を挿入したときの ξ を、

$$\begin{aligned} \xi_{M1,n,z,D} \\ = \{y_{n-D}, y_{n-D+1}, \dots, y_{n-1}, z, y_n, \dots, y_{n+D-1}\} \end{aligned} \quad (9)$$

と定義する。これに対するリファレンスとして、挿入を行わない場合の ξ を次式のように定義する。

$$\xi_{ref,n,D} = \{y_{n-D}, y_{n-D+1}, \dots, y_{n+D-1}\} \quad (10)$$

ここで、 k と D を固定したもとして考え、評価関数 $L_1(n, z)$ を、挿入前後での線形予測誤差の比をあらわす量として、次式のように定義した。

$$\begin{aligned} L_1(n, z) &= -\log Q(k, \xi_{M1,n,z,D}) \\ &\quad + \log Q(k, \xi_{ref,n,D}) \end{aligned} \quad (11)$$

$L_1(n, z)$ が大きな値をとることは、修復によって線形予測モデルへの一致度が高まったことを表すから、音声に対して有効な評価関数となることが期待できる。

なお、 ξ_{M2,n,z_1,z_2} , ξ_{M3,n,z_1,z_2} , ξ_{M4,n,z_1,z_2} についても同様に定義できるので、これらを使って、 $L_2(n, z_1, z_2)$, $L_3(n, z_1, z_2)$, $L_4(n, z_1, z_2)$ を、同様に定義する。以上で評価関数が導かれた。以下、この評価関数を用いた修復法を線形予測法と呼ぶことにする。

3.2 スパース性規範による方法 (ℓ^1 ノルム法)

n_0 を M1 損傷点としたとき、 $n \geq n_0$ で $y_n = x_{n+1}$ が成立するから、損傷を受けることは、雑音 e_n ,

$$e_n = y_n - x_n = \begin{cases} 0, & n < n_0 \\ x_{n+1} - x_n, & n \geq n_0 \end{cases} \quad (12)$$

の発生として解釈できる。 e_n のスペクトルについて考察しよう。 e_n に窓関数を乗算して FFT (高速フーリエ変換) したものは、「信号 $x_{n+1} - x_n$ を FFT したものの」と「窓関数の左半分を 0 で置き換え FFT したものの」の周波数領域での畳込みとなる。前者は、高域成

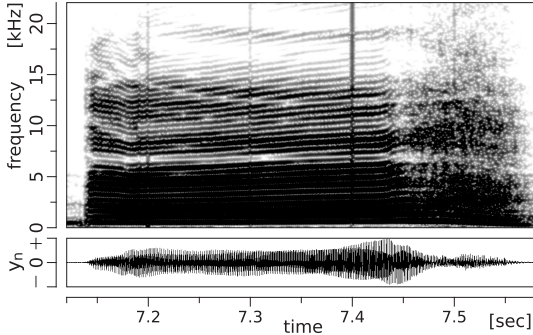


図 2 M1 損傷を受けた音声のスペクトログラムと波形
Fig. 2 A spectrogram and a waveform of signal y_n .

分が強調された原信号スペクトルである．後者は，窓中央の段差のせいで広帯域のスペクトルをもつ．後者を前者に畳み込むことで，周波数軸上でエネルギー散逸が起こる．そして，それは，原信号が高周波成分を強くもつほど顕著となる．例として，後に表 1 で定義する音声 VF の 7.2s, 7.3s, 7.4s, 7.5s に，M1 型損傷を与えた場合のスペクトログラムを図 2 に示す．損傷点におけるエネルギー散逸が黒い縦線として現れていることが分かる．

音声や音楽など，多くの音響信号は周波数領域でスパース性 [6] をもつから，損傷によりもともとエネルギーの小さかった領域へのエネルギー散逸が生じる．これは，スパース性の喪失で検出できるだろう．

なお，信号の雑音除去にスパース性の概念を持ち込んだ例としては，例えば中静 [7] が提案しているように，冗長な基底を用意し，音声と雑音異なる基底セットに帰着されることを利用する方法がある．これに対して，本手法は，単にスパース性の指標を用いるだけで信号の分解は行わない点や，基底が冗長でない点で本質的に異なっている．以下，提案手法を説明する．

まず，信号 ξ に窓関数を乗算する．今回は，サンプルの挿入の有無にかかわらず同じ特性の窓関数が適用され，かつ，挿入の有無で計算に組み入れるサンプルに差異が出ないように工夫した．これは， $|u| \geq D + \frac{1}{2}$ である u について $w(u)=0$ となる連続窓関数 $w(u)$ を用意し， $w(u)$ から互いに $\frac{1}{2}$ サンプルのずれを与えて値を採取することで得られる二つの離散窓関数を作り，挿入の有無に応じて両者を使いわけ窓関数 w_n ，

$$w_n = \begin{cases} w(n - D - 1), & \xi = \xi_{M1,n,z,D} \\ w(n - D - \frac{1}{2}), & \xi = \xi_{ref,n,D} \end{cases} \quad (13)$$

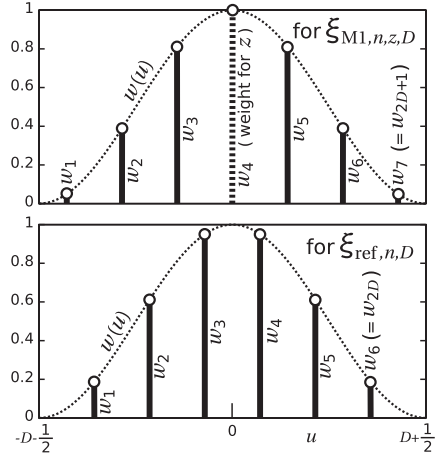


図 3 連続関数 $w(t)$ から，窓関数 w_n を導出する方法
Fig. 3 The relationship between continuous function $w(t)$ and induced discrete window w_n .

で実現できる． $D=3$ の場合を図 3 に図示する．

こうして定義した w_1, w_2, \dots を， ξ の第一要素から順に乘算して窓かけとする．窓かけした波形を N 点 ($N \geq 2D+1$) FFT したときの FFT ピン番号 k の成分値を Ξ_k とする．今回は，スパース性の指標として， ℓ^1 ノルムを用いることとし， ξ に対する ℓ^1 ノルムを，

$$\ell^1(\xi) = \frac{1}{N} \sum_{k=0}^{N-1} |\Xi_k| \quad (14)$$

と定義した．この値が小さいほどスパースである．

以上の準備のもと，評価関数 $L_1(n, z)$ を，挿入前後でのスパース性指標の差として，

$$L_1(n, z) = -\ell^1(\xi_{M1,n,z}) + \ell^1(\xi_{ref,n}) \quad (15)$$

と定義した．同様に， ξ_{M2,n,z_1,z_2} , ξ_{M3,n,z_1,z_2} , ξ_{M4,n,z_1,z_2} を使って，スパース性指標間の差として $L_2(n, z_1, z_2)$, $L_3(n, z_1, z_2)$, $L_4(n, z_1, z_2)$ が定義できる．以上で評価関数が導かれた．以下，この評価関数を用いた修復法を ℓ^1 ノルム法と呼ぶことにする．

なお，式 (14) の代わりに，規格化された ℓ^1 ノルム [8] を使う方法も考えられる．しかし，本手法に限っては，規格化しない方が良好な実験結果が得られた．

3.3 エントロピー規範の方法 (エントロピー法)

一方で「損傷によって信号の秩序が崩れる」と考えれば，エントロピー的な指標を用いるという発想も成立する．損傷によって，本来信号エネルギーが存在していなかった周波数帯域へのエネルギー出現が生じる

ことは、前節で説明した．これより、損傷は信号をエントロピー最大信号である白色雑音に一步近づける作用があるとみなせるからだ．

裏を返せば、損傷を受けたデータに対して、ある修復によってエントロピーが極端に減少すれば、その修復は正当であるということだ．すなわち、エントロピー最小化によって修復を行う方法である．

エントロピー最小化は、信号の修復や不連続点検出にしばしば応用されている．例えば、逆畳込み問題へ応用した例として（実際にはエントロピーの近似を用いているものの）Wiggins [9] の最小エントロピーデコンボリューション（MED）が古くから有名である．また、エントロピー指標を用いることで信号の不連続点を検出し、これを音声区間検出に応用した例 [10] も知られている．以上の成功例を考えると、サンプル欠け問題にもエントロピーが有効か否かをぜひ調べたいところだ．以下、提案方法を説明する．

ガウス性時系列信号の微分エントロピー H_e は、その信号のパワースペクトル $S(\omega)$ を使って、

$$H_e = \frac{1}{4\pi} \int \log S(\omega) d\omega + \frac{1}{2} \log(2\pi e) \quad (16)$$

と書ける [11]．すなわち、エントロピーの大小関係はパワーの周波数方向での積分で決まる．これを根拠に、本論文では、 $\xi = \{\xi_1, \xi_2, \dots, \xi_q\}$ に対する、エントロピー指標を、

$$H(\xi) = \frac{1}{N} \sum_{k=0}^{N-1} \log(|\Xi_k|^2) \quad (17)$$

と定義する．ここで、 Ξ_k は、窓がけをした信号の FFT 出力であり、前節で既に定義したものである．

ちなみに、分散が $|\Xi_k|^2$ である複素ガウス確率変数の微分エントロピーは、 $H_r = \log(|\Xi_k|^2) + \log \pi e$ であるので、式 (17) は、FFT の各周波数ピンの値が独立なガウス確率変数であると仮定したときのピン 1 個当たりの平均微分エントロピー（を線形変換した量）と見ることもできる．もちろん、正負の周波数成分間に複素数共役の関係が成立する他、今回は N よりも幅の狭い窓関数をかけることで隣り合う周波数成分間にも従属関係が生じているなど、式 (17) は、厳密なエントロピーではないが、エントロピー的指標としては妥当であると見てよいだろう．

ところで、修復アルゴリズムでは、 ξ の要素値 (z , または z_1 と z_2) を 1 LSB ステップで細かく調整す

る．これによって、特定の要素値で Ξ_k がほぼ 0 になることが起こり得る．このとき、 $H(\xi)$ は絶対値が非常に大きな負値となる．その結果、エントロピー最小の目的を外れて、「ある k に対して $\Xi_k = 0$ をめざす」という調整が行われてしまう可能性がある．今回は、これを防ぐために、微小量 ϵ ($\epsilon > 0$) を導入して、エントロピー指標を次式のように定義した．

$$H(\xi) = \frac{1}{N} \sum_{k=0}^{N-1} \log(|\Xi_k|^2 + \epsilon) \quad (18)$$

以上の準備のもと、評価関数 $L_1(n, z)$ を、挿入前後でのエントロピー指標の差として、

$$L_1(n, z) = -H(\xi_{M1, n, z}) + H(\xi_{ref, n}) \quad (19)$$

と定義した．同様に、 ξ_{M2, n, z_1, z_2} , ξ_{M3, n, z_1, z_2} , ξ_{M4, n, z_1, z_2} を使って、エントロピー指標間の差として $L_2(n, z_1, z_2)$, $L_3(n, z_1, z_2)$, $L_4(n, z_1, z_2)$ が定義できる．以上で評価関数が導かれた．以下、この評価関数を用いた修復法をエントロピー法と呼ぶことにする．

4. 実験による比較

実験では、六つの代表的波形を選び、三つの方法それぞれについて修復能力を調べ比較した．まず、実験対象とした信号と評価法について 4.1 で述べた後、発見率についての実験結果を 4.2 で述べ、損傷の型の判定率についての実験結果を 4.3 で述べ、復元誤差についての実験結果を 4.4 で述べる．なお、発見率と復元誤差については、四つの損傷型による差異はほとんどなかったため、4.2 と 4.4 では、M1 型の損傷に対する L_1 の評価結果のみを紹介する．

4.1 実験対象と実験方法

サンプリング周波数 $F_S = 44.1$ [kHz] で量子化ビット数 16 bit の長さ 9 秒の信号を 6 個用意し、詳細な比較を行った．用意した信号を表 1 に示す．このうち、音声と音楽については、スペクトログラムを図 4 に示す．音楽が 9 秒の時区間全てにわたって信号エネ

表 1 音源 x_n として選んだ六つの代表的な信号
Table 1 Six typical signals selected as signal x_n .

記号	内容
PM	男性ナレーターの新聞記事朗読（ピーク値 -3903）
PF	女性ナレーターの新聞記事朗読（ピーク値 -2991）
VF	女性声優のドラマ原稿発声（ピーク値 +15837）
MUSIC	音楽（白鳥の湖の「ヴェネチアの踊り」, 市販 CD）
SIN	1.403 kHz の正弦波（振幅 16768）
SIN+W	SIN に、標準偏差 32 の一様分布白色雑音を加算

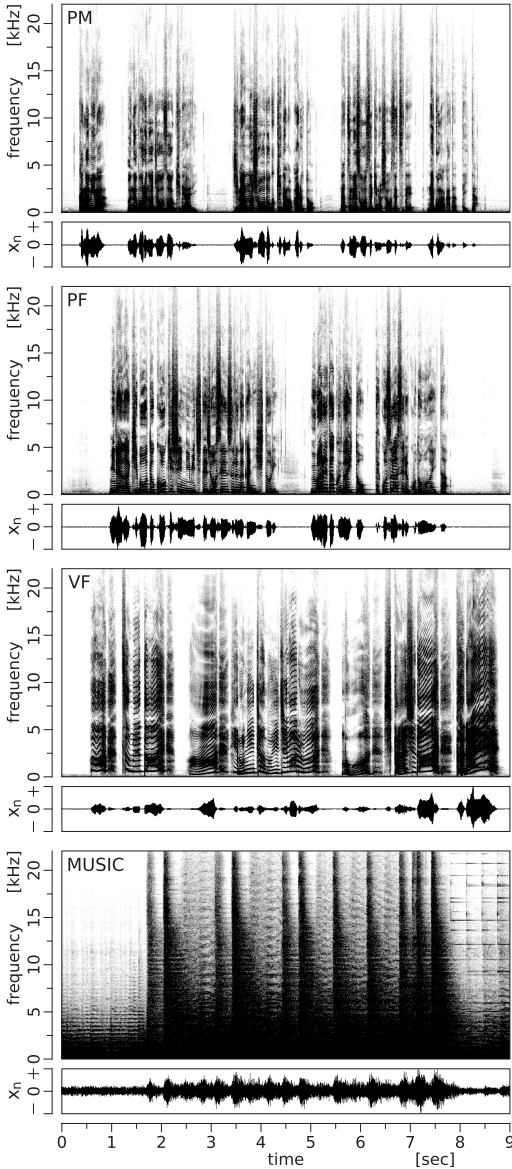


図 4 音源 x_n として選んだ信号のスペクトルと波形
Fig. 4 Spectrogram and waveforms of signal x_n .

ルギーが存在しているのに比較して、音声の場合は、10~35%程度の無音時間区間が存在している。

これら x_n に対して、0.1 秒間隔で 90 個の損傷を与えたものを損傷データ y_n とした。損傷位置による特異性が現れるのを防ぐため、損傷の開始位置を 0.0166 秒ずつずらして 6 通りの y_n を作り、この六つのデータに対して M1~M4 の四つの型の損傷を与えた。すなわち、ある音源 x_n に対し 24 通りの y_n を準備した。

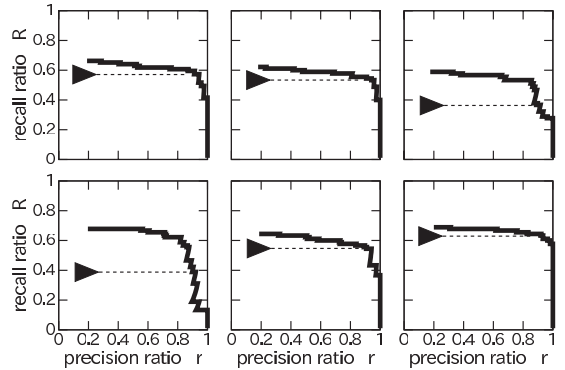


図 5 適合率 r (横軸) と再現率 R (縦軸) のトレードオフ関係の例。エントロピー法 ($D = 256, \epsilon = 1.0$) において、しきい値 C_L を変化させたときの軌跡。
Fig. 5 Trade-off curves between precision ratio r and recall ratio R . The trade-off parameter is C_L . (entropy method, $D = 256, \epsilon = 1.0$)

なお、修復にあたっては、式 (1) において L_1, L_2, L_3, L_4 の最大値を与える z, z_1, z_2 を求める必要がある。今回は、この計算は数値的探索で決定した。

4.2 発見率についての実験結果

しきい値 C_L を下げていけば、いずれ全ての損傷点を発見できるが、誤発見 (非損傷点を損傷と判断) も多く生じてしまう。すなわち、情報検索の用語を借用して表現すれば、 C_L の決定に際し、再現率 R (真の損傷点のうち何割を発見できるか) と適合率 r (検出された点のうち何割が真の損傷点か) のトレードオフの関係が存在する。例として、MUSIC から作った 6 通りの y_n の損傷点発見に関するトレードオフ関係を図 5 に示す。このトレードオフ関係の存在により、発見率を一元化した指標として定めるのは難しい。今回は、発見率の指標 R_d を「誤発見を三つまで許容したときの再現率の平均」として定義した。例えば、図 5 においては、各トレードオフ軌跡の下から四つ目の段差を示す六つの黒三角の縦座標の平均値が R_d である。

まず、線形予測法に対する結果から見てみよう。線形予測法では、設定が必要なパラメータとして、評価時間幅 D と、予測次数 k がある。これらを横軸にとり発見率 R_d を縦軸にとったグラフを描き解析した。

図 6 は、 $D = 64$ に固定して、 k と R_d の関係を見たものである。MUSIC では、 k は 10~40 程度の広い範囲でほぼ同一の発見率が得られるが、音声 (PM,PF,VF) の場合は、20~30 程度と、良好な範囲がやや狭くなる。

図 7 は、 $k = 23$ に固定して、 D と R_d の関係を見

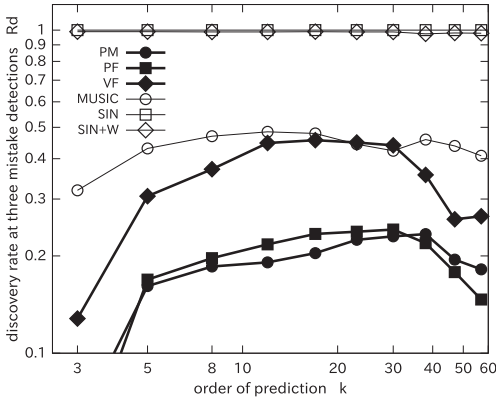


図6 線形予測法の発見率 R_d . 評価時間幅を $D=64$ に固定して, 予測次数 k との関係を示す.
Fig.6 The relationship between the order of prediction k and the discovery rate R_d . ($D=64$)

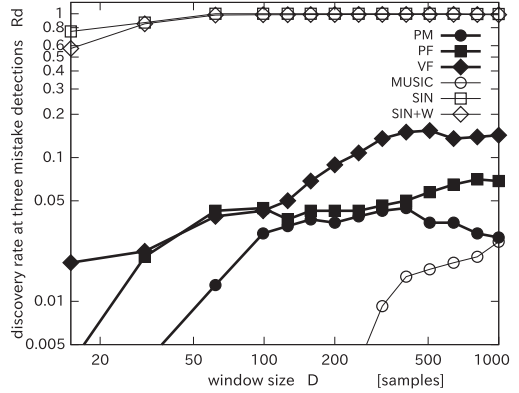


図8 ℓ^1 ノルム法の発見率 R_d . 窓幅 D との関係
Fig.8 The relationship between window size of the ℓ^1 -norm method D and the discovery rate R_d .

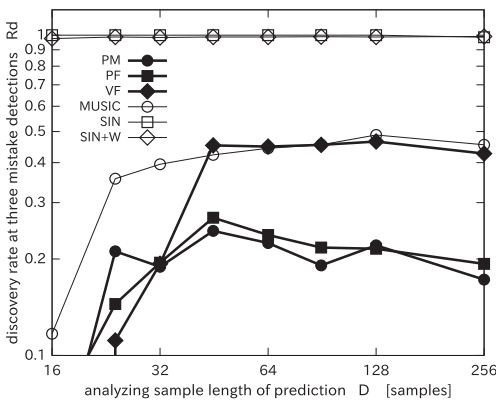


図7 線形予測法の発見率 R_d . 予測次数を $k=23$ に固定して, 評価時間幅 D との関係を示す.
Fig.7 The relationship between analyzing sample length of prediction D and the discovery rate R_d . ($k=23$)

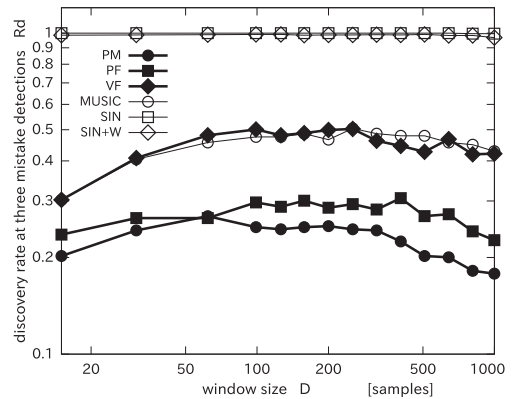


図9 エントロピー法の発見率 R_d . 窓幅 D との関係
Fig.9 The relationship between window size of the entropy method D and the discovery rate R_d .

たものである. D は 32 以上に設定すべきと分かる.

図6と図7を総合し, $D=64, k=23$ が最適であった. なお, SINとSIN+Wに対しては, どのパラメータであっても, ほぼ100%の発見が達成できた.

次に, ℓ^1 ノルム法について見てみよう. 窓幅 D と R_d の関係を, 図8に示す. 図6や図7とは, 縦軸の範囲が違うことに注意されたい. D を 500~1000 程度まで大きくしても, 線形予測法に比べて大きく劣っている. 特にMUSICの場合の性能低下が著しい.

最後は, エントロピー法である. パラメータとしては, 窓幅 D と微小量 ϵ がある. ここでは, 5.1 で述べる根拠に基づき, ϵ を係数 p を含む次式で決めた.

$$\epsilon = \frac{p}{12} \sum_n |w_n|^2 \tag{20}$$

$p=1.0$ とした場合の D と R_d の関係を図9に示す. ℓ^1 ノルム法より良好で, 線形予測法とほぼ同等である. 特に D について 50~500 程度の広いレンジにおいて, 良好な特性が安定して持続することは注目に値する.

4.3 損傷の型判定についての実験結果

損傷の型判定についての実験結果は, 紙面の都合で, エントロピー法についての結果のみ示す.

六つの音源別に見た判定の混同行列を表2に示す. これより, SINまたはSIN+Wを音源とした場合には, まず混同は起こらないことが分かる. これに対して, 音声(PM, PF, VF)やMUSICを音源とした場合には, 最大10%程度の割合で混同が起きている.

表 2 損傷型の判定の混同行列
Table 2 The confusion matrices of the damage types.

音源：PM	判定：M1	判定：M2	判定：M3	判定：M4
損傷型：M1	97.1%	2.2%	0.0%	0.7%
損傷型：M2	7.8%	75.7%	7.4%	9.1%
損傷型：M3	5.1%	10.8%	80.6%	3.5%
損傷型：M4	5.7%	8.5%	5.0%	80.8%

音源：PF	判定：M1	判定：M2	判定：M3	判定：M4
損傷型：M1	96.6%	2.7%	0.0%	0.8%
損傷型：M2	1.0%	81.5%	7.9%	9.6%
損傷型：M3	0.7%	9.2%	86.2%	3.9%
損傷型：M4	1.0%	10.7%	2.0%	86.3%

音源：VF	判定：M1	判定：M2	判定：M3	判定：M4
損傷型：M1	95.7%	2.8%	0.3%	1.3%
損傷型：M2	0.2%	87.1%	5.7%	6.9%
損傷型：M3	0.2%	6.3%	90.4%	3.0%
損傷型：M4	0.7%	12.2%	3.7%	83.4%

音源：MUSIC	判定：M1	判定：M2	判定：M3	判定：M4
損傷型：M1	98.9%	0.9%	0.0%	0.2%
損傷型：M2	0.2%	77.8%	11.5%	10.5%
損傷型：M3	1.4%	9.2%	85.0%	4.5%
損傷型：M4	0.6%	14.0%	3.5%	81.9%

音源：SIN	判定：M1	判定：M2	判定：M3	判定：M4
損傷型：M1	99.8%	0.0%	0.0%	0.2%
損傷型：M2	0.0%	100.0%	0.0%	0.0%
損傷型：M3	0.0%	0.0%	100.0%	0.0%
損傷型：M4	0.0%	0.9%	0.0%	99.1%

音源：SIN+W	判定：M1	判定：M2	判定：M3	判定：M4
損傷型：M1	99.6%	0.0%	0.0%	0.4%
損傷型：M2	0.0%	98.5%	0.7%	0.7%
損傷型：M3	0.0%	0.4%	99.6%	0.0%
損傷型：M4	0.0%	1.7%	0.0%	98.3%

M1 と M2 の間の判定に絞って、どのようなメカニズムで混同が起こっているのか、詳しく見ていこう。

まず、準備として、 $L_1(n) = \max_z L_1(n, z)$ と、 $L_2(n) = \max_{z_1, z_2} L_2(n, z_1, z_2)$ を定義しておく。図 10 は、音源 PF に対して M2 型の損傷を与えたとき、 $n = 1, 2, \dots$ に対して $L_2(n)$ と $L_1(n)$ の関係をプロットしたもので、いわゆる散布図である。プロットが点線より右側にあることは、その点が正しく損傷型判定されることを示している。図より、判定はほぼ完全であることが分かる。 L_2 は y_{n_0+1} を変化させた上で最大値を求めるが、 L_1 にはその変化の余地がない。これにより、M2 であれば常に $L_2 \geq L_1$ が成立するので、等号の場合を除いて混同は起こらないのである。なお、等号の場合、現在は L_1 と判定している。これによって誤判定とはなるが、 L_2 と正しく判定して y_{n_0+1} の新しい値を y_{n_0} とすると結果としては

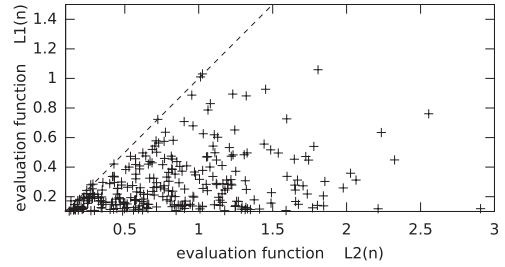


図 10 M2 型損傷時の評価関数 $L_2(n)$, $L_1(n)$ の関係
Fig. 10 The relationship between $L_2(n)$ and $L_1(n)$ in the case of M2-type damage.

同じ修復になるから、誤判定でも全く問題にはならない。すなわち、表 2 で混同と判定されていても、修復としては無害である部分があるということである。

一方、多くの損傷は M1 型であるので、真の損傷が M1 であるときに M2 と判定されてしまう誤判定の方が重要である。しかし、 L_2 が定義されるためには、 $y_{n_0+1} = y_{n_0}$ の条件が成立することが必要で、それは音源 x_n について隣り合う 2 サンプルが偶然同じ値でない限り成立しない。すなわち、多くの場合、 L_2 が定義されないという理由で混同が防がれる。データのビット長が長くなればなるほど、その偶然が起こる確率も小さくなるので、更に問題にならなくなる。ただし、音声収録での無音部では $y_{n_0+1} = y_{n_0}$ も起こり得る。しかし、このように信号値変化が小さな区間では、M2 の修正を行って y_{n_0+1} を調整してしまったとしても、結局は y_{n_0} に近い値のままとなり、多くの場合、問題にならないと考えられる。

M1 誤りと M2 誤りについての以上のような考察より、もともと誤判定の確率は小さく、しかも、起こっても結果には大きな影響を及ぼさないことが分かる。

4.4 値の復元誤差についての実験結果

本節では、損傷で失われてしまった信号値 x_{n_0} に対する復元値 z の復元誤差について調べる。本論文では、復元誤差 E を、誤差 $|z - x_{n_0}|$ の平均として定義する。

まず、線形予測法による復元誤差 E を図 11 (a), (b) に示す。図 11 (a) より、予測次数 k として、10~40 程度が適切であることが分かる。これは、発見率での結果とほぼ一致する。また、図 11 (b) より、評価時間幅 D についても、32 以上であればほぼ十分であることも分かった。

次に、 ℓ^1 ノルム法による復元誤差を図 11 (c) に示す。線形予測法に比較して、復元誤差が大きいことが分かる。特に MUSIC に対する復元値は極端に悪い。

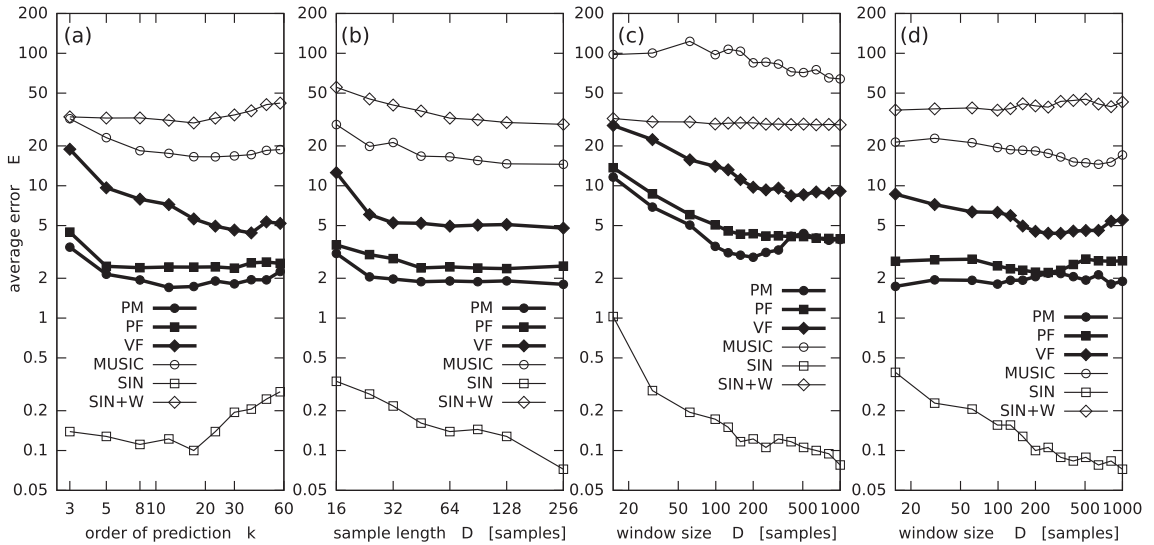


図 11 復元誤差 E での比較 . (a) 線形予測法 . $D=64$ に固定し , 予測次数 k を横軸にとった . (b) 線形予測法 . $k=23$ に固定し , 評価時間幅 D を横軸にとった . (c) ℓ^1 ノルム法 . 窓幅 D を横軸にとった . (d) エントロピー法 . 窓幅 D を横軸にとった .

Fig. 11 The relations between parameters and estimation errors. (a) LP method, $D=64$. (b) LP method, $k=23$. (c) ℓ^1 -norm method. (d) Entropy method.

最後に、エントロピー法による復元誤差を図 11 (d) に示す。線形予測法と同程度であることが分かる。

5. 考 察

5.1 エントロピー法の微量量 ϵ について

エントロピー法では、微量量 ϵ を決定する必要がある。式 (18) によれば、 ϵ が $|\Xi_k|^2$ に加算されてエントロピー評価が行われるので、 $|\Xi_k|^2$ のもつ誤差エネルギー程度の加算であればエントロピー評価に大きな影響は与えないであろうと考えた。

本論文では、 $|\Xi_k|^2$ のもつ誤差エネルギーを、信号 x_n の量子化誤差から見積もることにする。 x_n を整数とすれば、1 サンプル当りの量子化誤差のエネルギーの期待値は $\frac{1}{12}$ である。この誤差がサンプル間で独立であると仮定すると、これが窓関数 w_n で重みづけられ FFT された結果の誤差エネルギーは、式 (20) の右辺において $p = 1.0$ としたものと等しくなる。

p の最適値について論じよう。 p と発見率の関係を図 12 に示す。 $p > 1.0$ では性能低下が生じているが、PF のように $p = 1.0$ 付近で発見率が上がっている音源もある。次に、 p と復元誤差 E の関係を図 13 に示す。多くの音源で、 p が 1.0 より少し大きなところに最適値をもつ。両者を勘案すると、 $p = 1.0$ 付近が良好だ。1.0 という比較的大きな値が最適であったこと

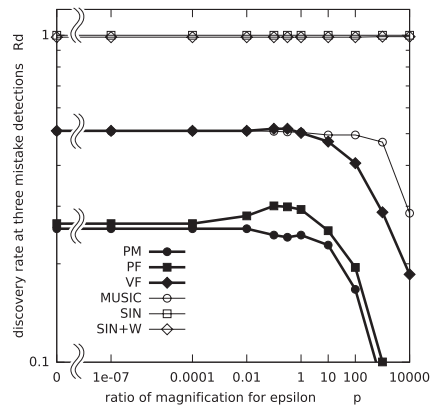


図 12 微量量 ϵ の修正倍率 p が発見率 R_d へ与える影響
Fig. 12 The relationship between p and R_d .

は、 ϵ の加算が、式 (18) における 0 の対数演算を回避する目的を超えて、積極的な機能を有することを示している。すなわち、量子化誤差などランダムな要因で生じたスペクトル上の凹凸を ϵ の加算で消すことができ、結果として、偶然に左右されない安定なエントロピー推定が実現できたものと考えている。以上の考察より、本論文では $p = 1.0$ を支持したい。なお、今回の実験に限っては、 $p = 0.0$ としても、計算不能や性能低下などの問題は生じなかった。

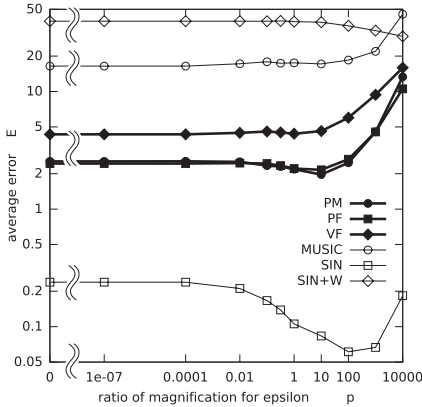


図 13 微小量 ϵ の修正倍率 p が復元誤差 E へ与える影響
Fig. 13 The relation between p and average error.

表 3 音源 VF (6 秒) の損傷の修復にかかる計算時間
Table 3 Computation times for restoration of VF.

方法	設定パラメータ	全点評価での計算時間
線形予測法	$D=64, k=23$	6723 秒
l^1 ノルム法	$D=253$	360 秒
エントロピー法	$D=253$	296 秒

2.66 GHz の Xeon, メモリ 3 GByte のマシン . C 言語で実装 .

5.2 性能比較

三つの方法を比較する . 図 6 ~ 図 9 と図 11 より判るように , l^1 ノルム法は明らかに性能が低く , 線形予測法とエントロピー法が , ほぼ同じ性能である .

エントロピー法は , 本質的には時間周波数分布をスパースにするよう機能していると考えられるので , エントロピー法と l^1 ノルム法で性能差が出たことは興味深い . エントロピーにおける対数演算が , 時間周波数分上のエネルギーがほぼ 0 の領域をより強く作る効果を生み出しており , これが良い効果を生んだと考えているが , 更なる解析が必要であろう .

次に , 三つの方法について , パラメータ D, k を最適に設定し , 計算時間を比較した結果を表 3 に示す . 線形予測法は , 計算コストが非常に高い .

更に , 特性が良好であった線形予測法とエントロピー法を詳しく比較しよう . パラメータを表 3 の値に固定したときの発見率 R_d での比較を図 14 に示し , 復元誤差 E での比較を図 15 に示す . 両手法が全く異なる原理に基づいているにもかかわらず , R_d と E の双方で , 両手法による値が高い相関をもっていることは興味深い . いずれにしても , 計算コストを考えれば , エントロピー法が最も優れている .

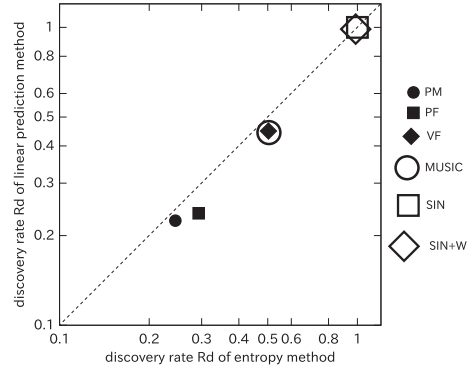


図 14 線形予測法 ($D=64, k=23$) とエントロピー法 ($D=253$) の発見率 R_d での性能差比較
Fig. 14 An R_d comparison between LP method ($D=64, k=23$) and entropy method. ($D=253$)

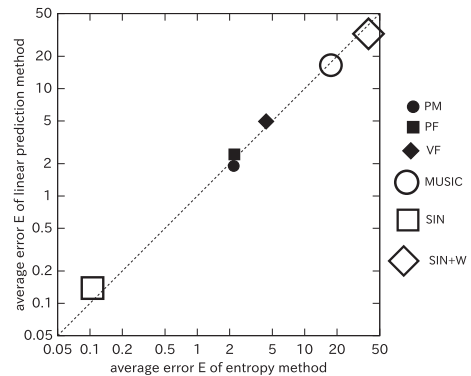


図 15 線形予測法 ($D=64, k=23$) とエントロピー法 ($D=253$) の復元誤差 E での性能差比較
Fig. 15 An error comparison between LP method ($D=64, k=23$) and entropy method. ($D=253$)

5.3 音楽のジャンルによる差異

4. の実験で用いた音楽 (MUSIC) は , オーケストラ演奏であった . このサンプル信号だけで音楽信号一般に対して本手法の有効性を示すのは不十分である .

具体的には , l^1 ノルム法とエントロピー法は , 損傷によって広帯域のエネルギー散逸が起こることを利用して検出と修復を行うので , もともとパルス的な音を多く含む音楽に対して有効に働くかという疑問が残る . また , 線形予測法では , 隣接サンプルの値が近似的に線形予測可能であるということを利用しているため , 調波成分のほとんど存在しない音波形に対しては良好な特性が得られないことが予想される .

そこで , 表 1 の MUSIC に加えて , 表 4 に示す四つの音楽信号を追加し , 三つの方法 (設定パラメータは表 3 に示す値に固定) を比較する追加実験を行った .

表 4 追加して実験した四つの音楽信号
Table 4 Additional four music signals.

記号	内容
MUSIC-2	日本オーディオ協会のテスト用 CD の Impact2 からドラム独奏 (Tr.3 の 231 秒 ~ 240 秒の強奏部, 右ch)
MUSIC-3	ロック・ビートルズのアルバム HELP から「HELP」(Tr.1 の 1 秒 ~ 10 秒の演奏冒頭部・モノラル音源)
MUSIC-4	ポピュラー・Mike のアルバム「懐かしのブルーライトヨコハマヨコスカから「恋のパカンス」(Tr.3 の 125 秒 ~ 134 秒のサビ部分, 右ch)
MUSIC-5	ジャズ・渡辺貞夫のオレンジ・エクスプレスから「ムバリ・アフリカ」(Tr.7 の 10 秒 ~ 19 秒の部分, 右ch)

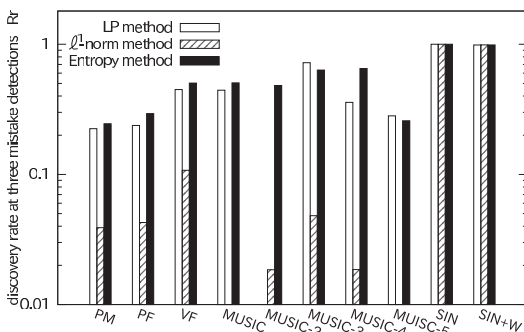


図 16 発見率 R_d の総括 (追加の音楽信号を含む比較)
Fig. 16 A summary of the discovery rates R_d . (including additional four music signals)

発見率での比較を図 16 に示す。ドラムの独奏 (MUSIC-2) は、調波成分のほとんど存在しない信号だが、予想どおり、MUSIC-2 に対して、線形予測法での発見率が極端に低下し、値としては 0 となった。これに対して、エントロピー法は、今回追加した四つの音楽信号に対しても良好な発見率を示した。

復元誤差での比較を図 17 に示す。四つの音楽信号に対しても、 l^1 ノルム法が最も劣っていることは変わらなかった。エントロピー法は、線形予測法と比較して、ほぼ同等か上回る性能を示した。

5.4 損傷位置の統計的性質の利用

損傷は、送信側と受信側のクロックの周波数差から生じることから、損傷位置は、ほぼ等間隔に現れる。この等間隔性を利用すれば、発見率を向上させられるだけでなく、誤発見を抹消することもできる。

損傷の時間間隔の変化を 10 分間実機で記録した結果を図 18 に示す。送信側は AURORA16 (LYNX 社)、受信側はパソコンに内蔵したボード (AES16-XLR) である。ソフトウェアは SoundForge10 (SONY) を

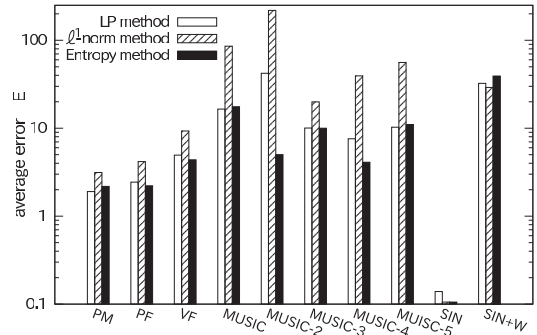


図 17 復元誤差 E の総括 (追加の音楽信号を含む比較)
Fig. 17 A summary of the estimation errors E . (including additional four music signals)

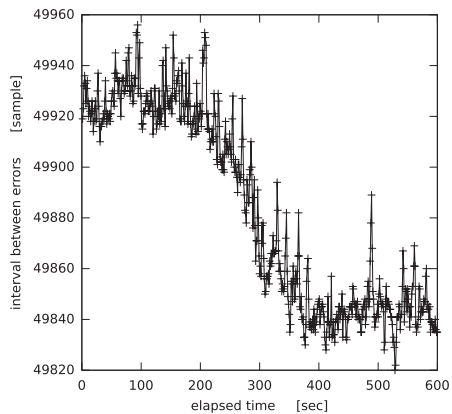


図 18 損傷の生じる間隔の時間変化
Fig. 18 Variations of time intervals of damages.

使っている。図 9 より正弦波の発見率は 100% であるので、正弦波を収録することで図 18 を描いている。

図 18 より、損傷間隔は、温度変化によると思われるドリフトと、ランダムなばらつきを加算値とみなせる。後者の大きさは、せいぜい ± 30 サンプル程度である。これより、ある時点の損傷を発見できたとき、次の損傷点は、候補地点の周り 60 サンプル程度を探せば見つかることになる。

損傷位置に対するこのような統計的情報を利用し、 C_L を適応的に変化させ、検出点を改良していくことが可能である。これを表 5 に示す。

本来、 C_L の最適値は信号依存だが、表 5 (1) における C_L を y_n だけから設定できることを図 19 で説明しよう。図 19 (f) が損傷を受けた音声波形 y_n であり、発話者は VF と同一である。エントロピー法 ($D = 256, \epsilon = 1.0$) を用いた場合の $L(n)$ を図 19 (a) に描く。ここで、しきい値を $C_L = 0.15$ とすると、発

表 5 損傷点位置の統計的性質を利用した損傷点の修正法
Table 5 A method for rejecting and adding points of lost samples using statistical knowledge.

- (1) しきい値 C_L を高めに設定して、損傷を発見してリストとして保存する。
- (2) そのリストを元に、損傷点の平均的時間間隔を推定する。
- (3) その平均的時間間隔から大きくずれる時間間隔を生じている損傷点候補を誤発見として抹消する。
- (4) 損傷点の平均的時間間隔を全体より細かな時区間で推定し、損傷点間隔をスプライン関数で近似する。
- (5) しきい値 C_L を低く設定し、スプライン関数で近似された間隔に一定の誤差以内で一致する間隔を生み出す損傷点候補であれば、新たに損傷点として登録する。
- (6) スプライン関数で近似された間隔に対し、一定の誤差以上の間隔を生み出す損傷点は、誤発見として抹消する。
- (7) (5), (6) での修正が一つでも起きていれば (4) に戻る。起きていなければ終了。

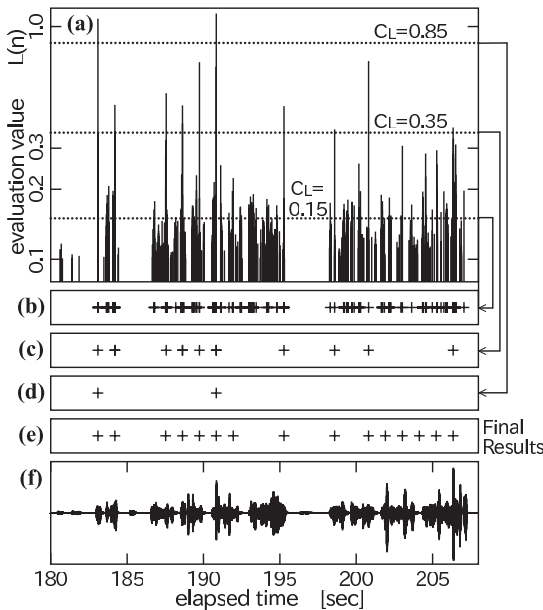


図 19 統計的情報を利用した修正点の追加と削除
Fig. 19 A method for addition and deletion based on statistical information.

見点は図 19 (b) となる。誤発見が多すぎて、損傷点の周期は推定不可能だ。そこでしきい値を上昇させる。 $C_L = 0.85$ まで上昇させると、図 19 (d) のように今度は発見数が少なすぎて損傷点周期が検出不可となるので、 $C_L = 0.35$ としたときの発見点が見点である。このとき、損傷点の周期 (図では 1 秒) が推定できるので、これを用いて、誤発見点を除去し、残った発見点を用いて更に精密に周期を推定した後、 C_L

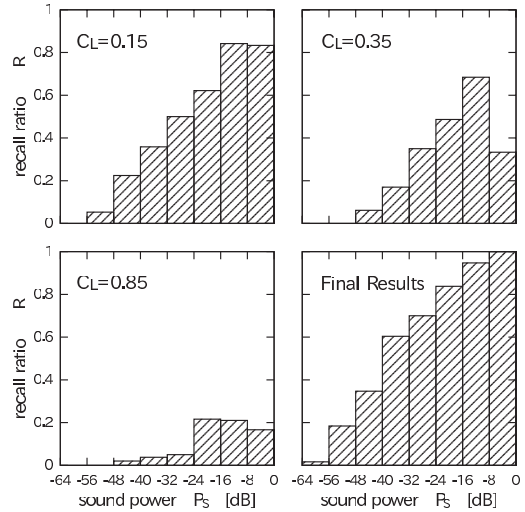


図 20 音声パワー別の再現率 R を、三つの C_L と統計的
情報を利用した修正後の結果について比較する。
Fig. 20 Relations between sound power P_S and recall
ratio R in the cases of three different thresh-
old values or the final detected points.

をより低い値に再設定し、推定周期と矛盾しない発見点を追加する処理を繰り返す。これが表 5 の手順である。最終的に得られた発見点を図 19 (e) に示す。音声が発声されている時区間に限れば、損傷が発見できていることが分かる。

なお、図 19 (e) では、193 秒と 194 秒付近の損傷点 2 個が発見できていないが、音素間の無音部に存在している損傷点であり、修復の重要度は低い。

そこで、音声強度別に解析してみよう。図 20 は、再現率 R を、図 19 (f) を含む 400 秒の信号について集計したものである。 C_L を 0.15 (左上)、0.35 (右上)、0.85 (左下) にとった場合と、表 5 の修正を行った場合 (右下) の四つのグラフを描いた。横軸は音声パワー P_S (損傷点を中心とした 0.01 秒の時区間で信号値の二乗平均) を表し、ピークレベルを 0 dB として 8 dB 間隔で集計してある。音声の最も強い (-8 dB ~ 0 dB) 時区間において、 $C_L = 0.35$ では全損傷点の 33% しか発見できないが、表 5 の手順を踏むことで、全損傷点が 100% 発見できている。なお、音声強度が -64 dB ~ -56 dB の時区間においては、表 5 の手順を踏んでも約 2% の発見にとどまっているが、ここは音声ほとんど発話されていない時区間であるので、修復の必要性は薄い。

次に、音声強度別の適合率 r を図 21 に集計する。図 19 では妥当に見えた $C_L = 0.35$ は、音声強度が大

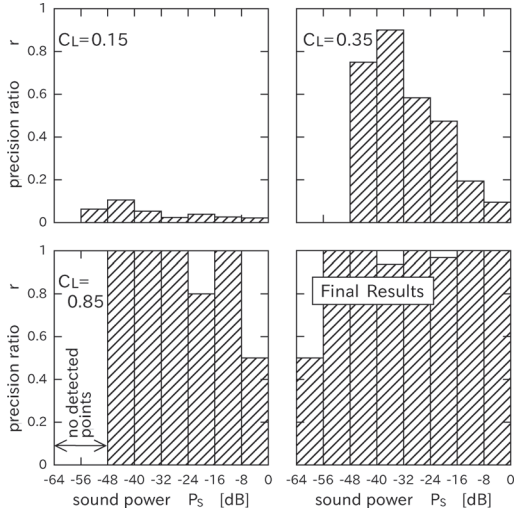


図 21 音声パワー別の適合率 r を、三つの C_L と統計的情報を利用した修正後の結果について比較する。

Fig. 21 Relations between sound power P_s and precision ratio r in the cases of three different threshold values or the final detected points.

きな時区間を中心に適合率の低下が見られた。しかし、表 5 の処理導入で、大幅な改善が得られている。

以上のように、損傷位置の統計的情報を利用すれば C_L を適応的に調整できるので、再現率と適合率の両方が向上する。更に、 $L(n)$ の計算を損傷の可能性のある点だけに限定して行うことも可能となり、表 3 の計算時間を 2~3 けた削減することもできる。

6. むすび

デジタルデータのサンプル欠けの修復法を論じた。送信側と受信側のクロックの非同期で生じる損傷の型を分類し、損傷位置の推定、型の判別、信号値復元を、単一の評価関数で行う方法を提案した。評価関数を得る方法として、線形予測法、 ℓ^1 ノルム法、エントロピー法の三つを定式化し、それらの性能を実験で比較した結果、エントロピー法が優れていることが分かった。更に、損傷位置の統計的性質を利用することで、修復性能を大幅に向上できることが分かった。

謝辞 本研究は科研費(20500111)の助成を受けて行われた音声収録を発端として遂行されたものである。

文 献

- [1] S.V. Vaseghi and P.J.W. Rayner, "A new application of adaptive filters for restoration of archived gramophone recordings," Proc. IEEE Conference on Acoustics, Speech, and Signal Processing, vol.5, pp.2548-2551, 1988.
- [2] J.J.K. O'Ruanaidh and W.J. Fitzgerald, "Restoration of missing samples in digital audio signals," in Numerical Bayesian Methods Applied to Signal Processing, pp.122-160, Springer, 1996.
- [3] S.J. Godsill and P.J.W. Rayner, "A Bayesian approach to the detection and correction of error bursts in audio signals," Proc. IEEE Conference on Acoustics, Speech, and Signal Processing, vol.2, pp.261-264, 1992.
- [4] R.J. Marks, II, "Restoring lost samples from an over-sampled band-limited signal," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-31, no.2, pp.752-755, 1983.
- [5] C.-Y. Hsu and T.-M. Lo, "Improved Papoulis-Gerchberg algorithm for restoring lost samples," Proc. IEEE International Symposium on Signal Processing and Information Technology, pp.717-721, 2005.
- [6] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," IEEE Trans. Inf. Theory, vol.47, no.7, pp.2845-2862, 2001
- [7] 中静 真, "反響雑音推定を伴うスパース信号分解による音声からのショット雑音抑圧," 信学技報, EA2006-13, 2006.
- [8] 小野順貴, 嵯峨山茂樹, "音声のスパース表現のためのフィルタバンクの検討と聴覚特性との比較," 音響講論集, 3-P-24, 2006.
- [9] R.A. Wiggins, "Minimum entropy deconvolution," Geoexploration, vol.16, pp.21-35, 1978.
- [10] F. Yingle, L. Yi, and W. Chuanyan, "Speech end-point detection based on speech time-frequency enhancement and spectral entropy," Proc. IEEE Conference on Engineering in Medicine and Biology Society, pp.4682-4684, 2005.
- [11] 甘利俊一, 長岡浩司, "情報幾何の方法," 岩波講座・応用数学, p.93, 1993.

(平成 22 年 11 月 6 日受付, 23 年 3 月 3 日再受付)



高橋 弘太 (正員)

1984 東大・工・計数卒。1986 同大大学院修士課程了。同年、東京大学工学部助手。1994 同大講師。同年、電気通信大学講師。1996 同大助教授。2007 同大准教授。博士(工学)。統計的信号処理、センサ信号のデジタル信号処理、マイクロホンアレーや独立成分分析などに特化した実時間センシングシステムの DSP の並列処理による実現、映像と音響情報によるマンマシンインタフェース、効率的視聴法と話速推定法の研究に従事。また、話速推定法の研究者用に、話速を厳密に制御して原稿を読み上げ収録した独自の音声データベース SRV-DB を製作し提供。日本音響学会、IEEE 各会員。