

音声信号のための順序統計を用いた適応非線形予測器と反復法によるその特性改善

田中 啓文^{†a)} 早川 晴子[†] 島村 徹也[†]

Adaptive Nonlinear Predictor Based on Order Statistics for Speech Signals and Its Performance Improvement by Iterative Approach

Hirobumi TANAKA^{†a)}, Haruko HAYAKAWA[†], and Tetsuya SHIMAMURA[†]

あらまし 線形予測の抱えるインパルス性残留誤差の問題を解決することを目的とし、本論文では、音声信号に非線形予測を施すための新しい適応フィルタ、適応非線形予測器を提案する。順序統計処理が用いられつつ、Pitasらによる LMS-L フィルタに、入力信号の時間情報を保持する改良が加えられる。合成音声を用いた実験では、提案する非線形予測器が、線形予測器だけではなく、ポルテラ級数予測器よりも優れた予測性能をもつことが示される。しかしながら、本非線形予測器をもってしても、実音声を対象とする実験では予測精度が劣化してしまう。そこでこれを解決すべく、新たに反復法を提案する。反復法は音声の周期性に着目して音声データを再利用し、多くのデータサンプルを用いたときと同様の予測精度が得られる方法である。連続実音声を用いた実験の結果、反復法とともに提案する予測器の予測精度が改善され、その有効性が確認された。

キーワード LMS アルゴリズム, 順序統計, 非線形予測, 反復法

1. ま え が き

線形予測技術 [1] は、多くの音声処理システムに適用されている。それは、音声信号をモデル化する際、線形予測が少ないパラメータで正確な表現を与え、またそのパラメータを求める計算が効率的であるためである。しかし、線形予測にも解決されるべき問題点がある。それは、有声音の分析において、線形予測の結果が音声信号のピッチ周期の影響を受けてしまう点にある [2], [3]。具体的には、短時間予測に線形予測法が用いられる場合、インパルス状の残留誤差が生じてしまう。音声のモデル化において、残留誤差の大きさはその達成度を表す一評価量とみなせるため、インパルス状の誤差は不必要なものと考えられる。

Thyssen ら [9] は、線形予測によってもたらされるインパルス状の残留誤差は音声信号が有する非線形性

の現れであり、それは非線形予測によって取り除かれることを提唱している。Thyssen らは、非線形予測には多層パーセプトロンが有効であるとする立場をとっている。一方 Diaz-de-Maria ら [10] は、同目的においてラジアル基底関数の使用を提案している。また最近、Varoglu ら [4] は、リカレントニューラルネットワークによる新しい予測分析モデルを報告している。しかし、ニューラルネットワークは一般に、そのフィルタ構成、すなわち階層構造を決定するのにあたり、事前に十分な経験的検索を必要とする。また、収束に多くの時間がかかることに加え、局所的な最小値への収束が起こり、所望とする解が必ずしも得られる保証がないという性質を有する。

上記の線形及び非線形予測の方法はすべてバッチ処理である。バッチ処理では、音声分析結果として、分析区間での代表的パラメータ値が 1 セット得られる（付加的な処理を加えることにより、分析区間長あるいはそれより短い区間長の予測値（予測誤差）が算出され得る）。しかし、一方で適応処理による音声分析もまた可能であり、その有効性も知られている [5]。適応処理では、直接的に入力信号のサンプル数と同数の

[†] 埼玉大学大学院理工学研究科, さいたま市
Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama-shi, 338-8570 Japan

a) E-mail: tanahiro@sie.ics.saitama-u.ac.jp

予測値（予測誤差）及びパラメータセットが得られる．各 n 時刻で、予測パラメータセットをそれぞれ与える適応処理は、分析区間内で予測パラメータが 1 セット与えられるバッチ処理より、算出されるパラメータ数という意味においてより多くの予測分析情報を与えると考えられる．

Gibson ら [6] は、DPCM に適応線形予測器を導入することにより ADPCM を実現し、音声品質を保持しつつ、伝送システムにおけるビットレートを大幅に低減できることを報告している．また、その特性改善を目的として、Mumolo ら [8] はボルテラ級数フィルタ構成に基づく ADPCM のための適応非線形予測器を提案している．ADPCM では、予測分析時に生じる予測誤差の低減が可能となれば、それが同時に送信ビット数の低減となり、伝達効率の向上につながる．また、適応予測のために最小平均 2 乗 (LMS) アルゴリズムを用いれば、計算量は少なくすすみ、大変効率的となる．

予測誤差を低減するために、適応予測器にピッチ予測器を併用する方式が知られている [12]．しかし、音声信号のピッチ予測は元来難しい問題であり、適応非線形予測器を使用するのみで予測誤差に含まれるインパルス性成分を除去できれば、ピッチ予測器の必要性はなくなり、より低いビットレートでの音声伝送が実現できる可能性が生まれる．また、音声信号を合成する場合においても、生成フィルタへの入力信号をより少ないビット数で表現でき、高品質な音声信号を生成するシステムの構築が期待できる．更には、適応非線形予測器のパラメータがより高精度な音声信号表現を与えるなら、音声分析や音声認識などへのより効果的な利用も考えられる．

本論文では、適応処理における予測誤差を低減させることを目的とし、新しい適応非線形予測器を提案する．本予測器は、Pitas ら [7] によって提案された順序統計処理を用いた適応フィルタ、LMS-L フィルタを入力ベクトルの時間情報が保持される形態に改良するものである．LMS-L フィルタは、インパルス雑音を除去するという潜在的性質をもつことが知られている．その除去能力は順序統計処理特有の非線形性によってもたらされる．そこで本予測器では、線形予測が生成するインパルス状の残留誤差を不必要成分ととらえ、順序統計処理を施すことにより残留誤差を低減し、音声信号の予測精度の向上を図る．本予測器の適応アルゴリズムでは、従来の係数ベクトルを更新する適応処

理とは異なり、係数行列を更新する．しかし、各時刻において更新される予測係数の数は LMS アルゴリズムと同じである．したがって、LMS アルゴリズムの特徴である計算効率性は本予測器においても保持される．以降では、この提案する適応非線形予測器を順序統計 LMS (OSLMS) 予測器と呼ぶことにする．

まず提案する OSLMS 予測器を合成音声を用いて、Gibson ら [6] の LMS を適応アルゴリズムとする予測器 (文献 [6] では確率近似法と記されている)、Mumolo ら [8] のボルテラ級数フィルタ構成に基づく適応非線形予測器、Pitas ら [7] によって提案された LMS-L フィルタ構成に基づく適応非線形予測器と実行結果を比較検討する．そしてその比較結果をもとに、OSLMS 予測器の特性を明らかにする．OSLMS 予測器は、従来の予測器より卓越した実行結果を与える．しかし、実音声を対象にした予測分析では、誤差特性を収束させるために多くのデータサンプルを必要とするだけでなく、予測精度も劣化してしまうという問題を指摘する．そこでこの問題を解決する目的で、少ないデータサンプルしか与えられない場合でも、予測精度を低下させない工夫を講じる．具体的には、反復処理による予測分析法を提案し、連続実音声を用いた実験を介してその有効性を検証する．

以下、本論文の構成は次のようになる．まず、2. で線形予測及び非線形予測の概念に触れ、適応予測のための従来法を記し、3. で OSLMS 予測器を具体的に説明する．そして合成音声を用いた実験の結果をもとに、順序統計処理により残留誤差がどのように低減されるかを明らかにする．ここでは、更に実音声を用いた実験を行い、OSLMS 予測器による予測精度が劣化してしまう性質を指摘する．これを解決すべく、4. では新たに反復法を提案する．そして連続音声を用いた実験を行う．ここでは、反復法と分析区間の関係を実験を介して検証し、反復法の有効性を明らかにする．そして最後に 5. で結ぶことにする．

2. 線形予測と非線形予測

本章では、線形予測と非線形予測の違いを示しつつ、従来の予測器の適応アルゴリズムを記述しておく．

まず、図 1 で示すような適応予測器を考えてみる．ここで、 $s(n)$ はサンプリングされた音声信号、 $\hat{s}(n)$ はその予測値を表している．一般に、音声は過去のサンプル値から予測されると仮定されるため、ここでは音声予測を

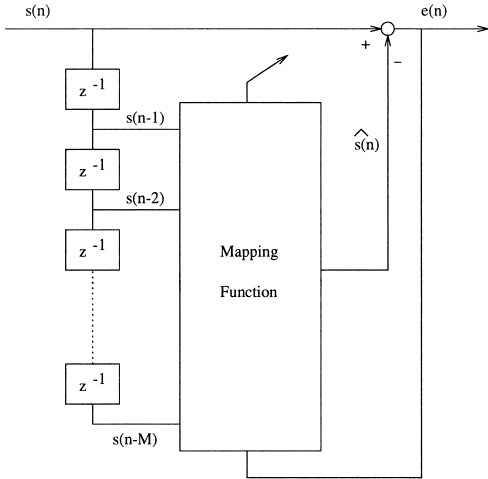


図 1 一般化適応予測器の構成

Fig. 1 Configuration of general adaptive predictor.

$$\hat{s}(n) = \phi(s(n-1), s(n-2), \dots, s(n-M)) \quad (1)$$

と数式化することにする。ただし、 $\phi(\cdot)$ は調整可能なパラメータを含むマッピング関数を表している。図 1 では、予測誤差 $e(n)$ を用いてこのマッピング関数のパラメータを時間とともに逐次調整する、いわゆる適応処理の形態を表している。もし、マッピング関数 $\phi(\cdot)$ が線形のみならず非線形に信号を処理する機能を備えるなら、この図 1 の予測器は一般化適応予測器とみなすことができる。

2.1 LMS 予測器

適応線形予測器は、上記の一般化予測器の特別な場合と考えられる。線形予測による音声信号の推定値は、

$$\hat{s}(n) = \sum_{i=1}^M a_i(n)s(n-i) \quad (2)$$

と表される。ここで $a_i(n), i = 1, 2, \dots, M$ は、予測係数に対応する。

LMS アルゴリズム (正規化) が適応処理に用いられる場合、LMS 予測器の適応手順は次のようになる [6]。

$$e(n) = s(n) - \mathbf{s}(n)^T \mathbf{a}(n) \quad (3)$$

$$\mathbf{a}(n+1) = \mathbf{a}(n) + \frac{\mu}{\mathbf{s}(n)^T \mathbf{s}(n) + \beta} \mathbf{s}(n)e(n) \quad (4)$$

ここで、 T は転置、 μ はステップサイズ、 β は正規化 LMS アルゴリズムのための安定化パラメータを表している。また、 $\mathbf{s}(n)$ は入力ベクトル

$$\mathbf{s}(n) = [s(n-1), s(n-2), \dots, s(n-M)]^T \quad (5)$$

を示し、 $\mathbf{a}(n)$ は反復回数 n での係数ベクトル

$$\mathbf{a}(n) = [a_1(n), a_2(n), \dots, a_M(n)]^T \quad (6)$$

である。 $\mathbf{a}(n)$ の収束は、ステップサイズ μ によって制御される。

2.2 VSLMS 予測器

式 (1) のマッピング関数 $\phi(\cdot)$ が非線形であるとき、図 1 の形態で適応非線形予測器は実現される。Mumolo ら [8] は、関数 $\phi(\cdot)$ としてボルテラ級数展開を用いている。Mumolo らのボルテラ級数展開においては、音声信号の予測値が

$$\begin{aligned} \hat{s}(n) = & \sum_{i=1}^{K_1} a_i(n)s(n-i) \\ & + \sum_{i=1}^{K_2} \sum_{j=i}^{K_2} b_{ij}(n)s(n-i)s(n-j) \end{aligned} \quad (7)$$

と表される。ここで、 $a_i(n), i = 1, 2, \dots, K_1$ は 1 次予測係数、 $b_{ij}(n), i, j = 1, 2, \dots, K_2$ は 2 次予測係数である。また、 K_1, K_2 はそれぞれ 1 次及び 2 次予測次数に対応している。式 (7) では 2 次予測係数の対称性 $b_{ij}(n) = b_{ji}(n)$ が考慮され、2 次項における冗長性は取り除かれている。

式 (7) の 1 次及び 2 次予測係数を LMS アルゴリズムで適応するボルテラ級数予測器を、以降ではボルテラ級数 LMS (VSLMS) 予測器と呼ぶことにする。VSLMS 予測器の適応手順は次のようになる。

$$e(n) = s(n) - \mathbf{q}(n)^T \mathbf{h}(n) \quad (8)$$

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{\mu}{\mathbf{q}(n)^T \mathbf{q}(n) + \beta} \mathbf{q}(n)e(n) \quad (9)$$

ここで、 $\mathbf{q}(n)$ は入力ベクトル

$$\begin{aligned} \mathbf{q}(n) = & [s(n-1), s(n-2), \dots, s(n-K_1), \\ & s(n-1)^2, s(n-1)s(n-2), \dots, \\ & s(n-K_2)^2]^T \end{aligned} \quad (10)$$

を示し、 $\mathbf{h}(n)$ は反復回数 n での係数ベクトル

$$\begin{aligned} \mathbf{h}(n) = & [a_1(n), a_2(n), \dots, a_{K_1}(n), \\ & b_{11}(n), \dots, b_{K_2 K_2}(n)]^T \end{aligned} \quad (11)$$

である。

2.3 LMS-L 予測器

LMS-L フィルタは, Pitas ら [7] により, インパルス性の付加雑音を低減する一種の適応 FIR フィルタとして考案された. したがって, 直接的に適応予測器としては考えられていない. しかしながら, 本論文で提案する適応非線形予測器は, この LMS-L フィルタを基本フィルタとしている. そこで, LMS-L フィルタを予測器として用いる LMS-L 予測器をここで導出しておくことにする. そして次章での提案予測器の記述へとつなげる.

LMS-L 予測器では, 式 (1) におけるマッピング関数 $\phi(\cdot)$ として, 順序統計を用いることになる. LMS-L フィルタ構成に基づく音声信号の予測値は,

$$\hat{s}(n) = \sum_{i=1}^M a_i(n)x_i(n-i) \quad (12)$$

となり, LMS-L 予測器の適応手順は次のようになる.

$$e(n) = s(n) - \mathbf{x}(n)^T \mathbf{a}(n) \quad (13)$$

$$\mathbf{a}(n+1) = \mathbf{a}(n) + \frac{\mu}{\mathbf{x}(n)^T \mathbf{x}(n) + \beta} \mathbf{x}(n)e(n) \quad (14)$$

ここで, $\mathbf{x}(n)$ は入力ベクトル $s(n)$ を, 要素の大きさを小さい順に並べ換えた順序統計ベクトル

$$\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_M(n)]^T \quad (15)$$

である. すなわち, このベクトルには

$$x_1(n) \leq x_2(n) \leq \dots \leq x_M(n) \quad (16)$$

の関係が存在することになる. $\mathbf{a}(n)$ は反復回数 n での係数ベクトルであり, 式 (6) と同様に表せる.

3. OSLMS 予測器

本章では, OSLMS アルゴリズムとそれに基づく予測器を説明し, 合成音声と実音声データに対して, OSLMS 予測器を用いた実験を行う. そして実験結果から提案法について考察する.

3.1 OSLMS 予測器の適応手順

Pitas ら [7] により考案された LMS-L フィルタは効率的かつ効果的に付加されたインパルス雑音を除去することを可能にしている. しかし, このフィルタでは, 式 (15) で与えられる入力ベクトルの順序統計が LMS 適応のための入力ベクトルとして直接用いられ, 入力ベクトルの要素の順序にかかわる時間情報を保持す

ることができない. 音声予測の前提が, ある時刻での音声信号の値をその過去の値から予測する式 (1) に基づいていることから, この LMS-L フィルタの時間情報を失う性質は, 音声処理において不適切である. そこで, 順序統計処理を行いつつ, 時間情報を保持する OSLMS 予測器を提案する.

OSLMS 予測器の適応アルゴリズム, OSLMS アルゴリズムでは, 係数ベクトルの代わりに係数行列

$$\mathbf{C}(n) = \begin{bmatrix} c_{11}(n) & c_{12}(n) & \dots & c_{1M}(n) \\ c_{21}(n) & c_{22}(n) & \dots & c_{2M}(n) \\ \vdots & \vdots & \ddots & \vdots \\ c_{M1}(n) & c_{M2}(n) & \dots & c_{MM}(n) \end{bmatrix} \quad (17)$$

を用意する. これらの要素 $c_{ij}(n)$, $i, j = 1, 2, \dots, M$ は, $n = 0$ のとき, すべて 0 値に初期化される.

適応においては, $\mathbf{C}(n)$ の $M \times M$ 個の要素のうち, M 個の要素のみが各反復ごとに選択され, 更新される. 具体的には, 反復回数 n のとき, $c_{m(j)j}(n)$, $j = 1, 2, \dots, M$ が選択される. そして, 係数ベクトル

$$\mathbf{c}(n) = [c_{m(1)1}(n), c_{m(2)2}(n), \dots, c_{m(M)M}(n)]^T \quad (18)$$

が形成され, 更新される. ただし, ここでの $m(j)$ は, 入力ベクトル

$$\mathbf{s}(n) = [s(n-1), s(n-2), \dots, s(n-M)]^T \quad (19)$$

が, 要素の大きさを小さい順に並べ換えた順序統計ベクトル $\mathbf{x}(n)$ (式 (15)) へ変換されるときに順番に対応する. すなわち, $m(j)$ は $i, j = 1, 2, \dots, M$ において

$$m(j) = i \quad \text{if } x_i(n) = s(n-j) \quad (20)$$

によって決定される整数を表す. 係数更新は,

$$e(n) = s(n) - \mathbf{s}(n)^T \mathbf{c}(n) \quad (21)$$

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \frac{\mu}{\mathbf{s}(n)^T \mathbf{s}(n) + \beta} \mathbf{s}(n)e(n) \quad (22)$$

によってなされる. 更新された係数ベクトル $\mathbf{c}(n+1)$ の要素は, 係数行列 $\mathbf{C}(n+1)$ に入れ直される. これにより, 1 回の係数更新が終了する. 次の係数更新は, また, 順序統計からの係数選択から始まり, 新たに選

択された係数からなる係数ベクトルが、再び更新される。これらの係数更新が繰り返されるたびに、最終的には係数行列の要素すべてが更新されることになる。

OSLMS 予測器は、基本的には式 (21), (22) により係数更新されるため、LMS 予測器と同じフィルタ構成で実現可能である。これは、図 1 で考えれば、マッピング関数において係数行列、式 (17) を伴うものの、実際のフィルタリングでは LMS 予測器と同様タップ遅延線フィルタの構成で、用いられる係数の数は LMS 予測器と同じであることを意味する。したがって、LMS 予測器の特徴である計算効率性は、本予測器でも保たれる。

3.2 実験

実験について述べる。まず合成音声/o/を使用し、OSLMS 予測器の比較対象として LMS 予測器、LMS-L 予測器、VSLMS 予測器の三つの予測器を取り上げることにする。次に実音声/a/を使用し、OSLMS 予測器の有効性について考える。

3.2.1 合成音による実験

ここで使用する合成音声は母音/o/である。以下の式に基づいて生成された。

$$u(n) = \delta(n) + \delta(n - P) + \delta(n - 2P) + \dots$$

$$n = 0, 1, 2, \dots \quad (23)$$

$$s(n) = \sum_{i=1}^M a_k(n)s(n - m) + Gu(n) \quad (24)$$

ここで、 $\delta(\cdot)$ はデルタ関数を表す。また P はピッチ周期のサンプル数に相当し、 a_k と G は一定のパラメータである。その具体的な値は、以下に記される。

$$G = 0.1345, a_1 = 1.53527, a_2 = -0.97789,$$

$$a_3 = 1.48396, a_4 = -1.78023, a_5 = 0.71704,$$

$$a_6 = 0.73514, a_7 = 0.76348, a_8 = 0.12135,$$

$$a_9 = -0.15552, a_{10} = -0.178143,$$

サンプリング周波数 = 10 kHz,

ピッチ周波数 = 2 ms ($P = 20$)

この合成音は、文献 [11] の中で、ピッチ周期 8 ms ($P = 80$) で使用されている。しかしここでは、生成されるインパルス雑音誤差をより明確に調べるために、より短いピッチ周期 2 ms ($P = 20$) を設定してある。ここでの評価には、予測誤差の収束特性と、収束時点での予測利得を考えることにする。誤差がより小さい値を

示せば、高精度の予測分析が行われ、予測利得が上昇すると考えられる。予測利得は、 n_1, n_2 を収束状態にあるように調整し、以下のように算出した。

$$SNR(\text{dB}) = 10 \log_{10} \frac{\sum_{n=n_1}^{n_2} s(n)^2}{\sum_{n=n_1}^{n_2} e(n)^2} \quad (25)$$

3.2.2 実験結果

図 2 は、LMS 予測器の収束特性を示している。ここで予測次数は、合成音声の生成過程に基づき $M = 10$ としてある。また、ステップサイズは $\mu = 1.0$ 、安定化パラメータは $\beta = 0.05$ としてある。図 2 から、LMS 予測器は音声がほぼ定常とみなせる 30 ms ($n = 300$) の間に収束しており、ここでのステップサイズ及び安定化パラメータの設定は適当と考えられる。しかし、収束しても線形予測共通の問題点であるインパルス状の予測残差は明らかに存在している。

図 3 には、VSLMS 予測器の収束特性を示している。ここでは、Thyssen ら [9] の線形予測による逆フィルタリングを線形処理、残留されるインパルスの抑制処理を非線形処理ととらえる立場を基本姿勢とし、予測次数を設定している。すなわち、式 (7) で表される線形部 (1 次) の予測次数は LMS 予測器と同様に $K_1 = 10$ とし、また非線形部 (2 次) の予測次数は文献 [8] に準じて $K_2 = 3$ としてある。また、ここでの適応は VSLMS 予測器の線形部 (1 次項) が LMS 予測器と同じ構成であることから、対等に比較するために、ステップサイズ、安定化パラメータはそれぞれ LMS 予測器と同じく $\mu = 1.0, \beta = 0.05$ にした。図 3 を図 2 と比較すると、VSLMS 予測器は、LMS 予測器より

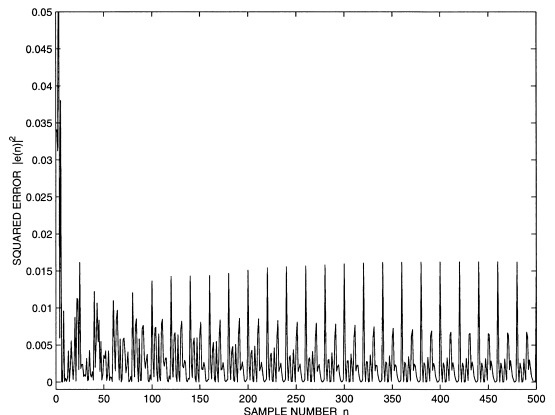


図 2 合成母音/o/における LMS 予測器の収束特性
Fig. 2 Convergence of the LMS predictor on the synthetic vowel /o/.

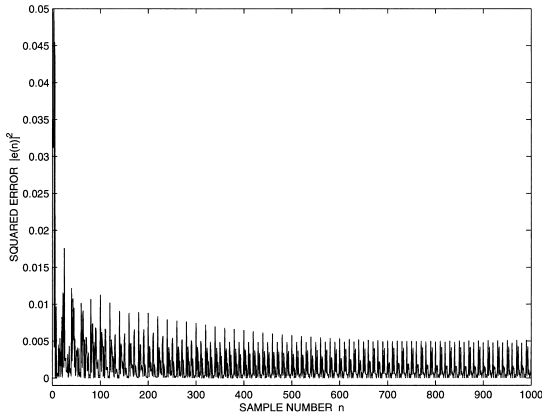


図3 合成母音/o/におけるVSLMS予測器の収束特性
Fig.3 Convergence of the VSLMS predictor on the synthetic vowel /o/.

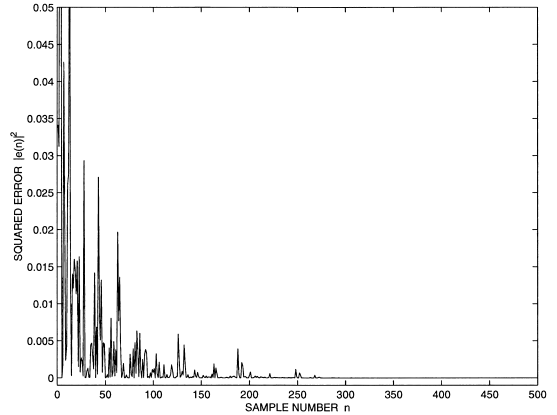


図5 合成母音/o/におけるOSLMS予測器の収束特性
Fig.5 Convergence of the OSLMS predictor on the synthetic vowel /o/.

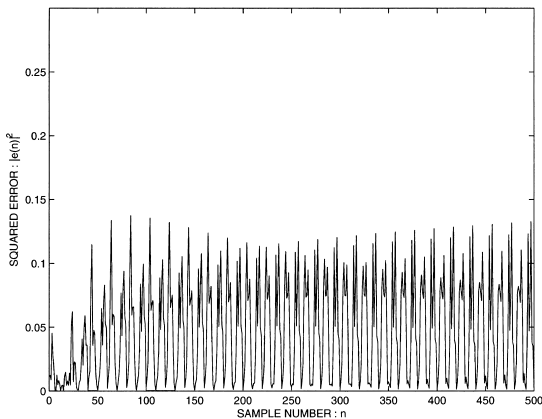


図4 合成母音/o/におけるLMS-L予測器の収束特性
Fig.4 Convergence of the LMS-L predictor on the synthetic vowel /o/.

も収束に達するまでにより多くの時間がかかることが分かる(図3の反復回数に対応する横軸のスケールは図2の2倍になっていることに注意されたい)。しかし、予測誤差の大きさが比較的小さくなっているのは、VSLMS予測器の非線形処理の結果と考えられる。

図4は、LMS-L予測器の収束特性を示している。予測次数は他の予測器同様、 $M = 10$ としてある。また、ステップサイズは $\mu = 1.0$ 、安定化パラメータは $\beta = 0.05$ と設定してある。文献[7]においては、インパルス雑音が低減できることが述べられているが、この図からは判断できない。そればかりか、LMS予測器に比べ予測残差の量も増加している。この原因として考えられるのは、入力ベクトルから順序統計ベクト

表1 合成音声における各予測器の予測利得

Table 1 SNR with each predictor on synthetic vowel.

predictor	SNR (dB)
LMS	9.0
VSLMS	12.40
OSLMS	101.80

ルへの変換の際の、入力信号が有する時間情報の損失である。

図5は、OSLMS予測器の収束特性である。OSLMS予測器のフィルタ構成がLMS予測器と基本的に同じであることから、ここでもLMS予測器と同様に予測次数 $M = 10$ 、ステップサイズ $\mu = 1.0$ 、安定化パラメータ $\beta = 0.05$ としてある。図2と比較することにより、OSLMS予測器はLMS予測器と同程度の収束速度を有することが見てとれる。しかし、音声の定常範囲内で予測誤差が十分に減衰しており、LMS予測器より収束状態での予測誤差の大きさが微小であることが分かる。また、図3のVSLMSの予測器と比較しても、明らかに収束後のOSLMS予測器の予測誤差は小さくなっている。このことは、予測利得を介しても確認することができる。表1には、式(25)によって算出した予測利得を示す。なお、ここではLMS-L予測器の結果は他の予測器より明らかに劣化することがうかがえるので明記していない。以上の結果より、OSLMS予測器は、線形予測器が抱える問題であるインパルス状の残留誤差を除去することを可能にし、線形予測器のみならず、VSLMS予測器をも上回る予測性能をもたらすことが分かる。

ここで、合成音声のような完全に周期的な音声波形

の予測において、OSLMS 予測器が上記のような卓越した実行結果を与える要因を考えてみる．線形予測器では、時刻 n において、時刻 $n-1$ での係数ベクトルに基づいて係数更新が行われるため、波形がピッチの駆動のような急激な変化を見せるときには、予測が不正確になってしまう．その結果がインパルス状の予測誤差として現れると考えられる．これに対して、VSLMS 予測器は付加される 2 次項により、インパルス状の予測誤差をある程度抑制することが可能である．しかし、除去するまでには至っていない．2 次予測次数を増大し、更なる実験を行ったが、予測誤差の改善は得られなかった．これにより、インパルス状の予測誤差を除去する目的においては、式 (7) に基づく 2 次項までのボルテラ級数展開ではフィルタ構成的に無理があると判断できる．

一方、OSLMS 予測器では、式 (18) での要素選択において時間情報を保持しながら順序統計処理を行うため、基本的には予測に必要な入力ベクトルが同一となる時刻、つまり波形のピッチ周期のサンプル数を P としたときの $n-P$ 時刻での係数ベクトルに基づいて、時刻 n での係数更新がなされると考えられる．すると、 n と $n-1$ 時刻に入力ベクトルの要素の大きさの順番が全く同等にならない限り、OSLMS 予測器では n 時刻と $n-1$ 時刻で異なる係数要素の更新がなされることになる．そのため、OSLMS 予測器では、線形予測器のような n 時刻での $n-1$ 時刻依存性は回避され、波形の急激な変化に影響されずに信号予測が可能となるわけである．また OSLMS 予測器では、適応が進むにつれて係数更新の規則性が明確となり、収束後の係数行列はほぼ一定になることが観察されている．これは、波形から選択される係数要素の値が固定されていることを意味する．このとき、波形の急激な変化を抑える係数決定がなされ、その係数を用いてフィルタリングが行われると考えれば、インパルス状の誤差は発生しなくなると容易に類推することができる．

上記のような OSLMS 予測器の係数選択とその更新を考えると、複数の線形予測器をあらかじめ用意しておき、入力信号ベクトルに応じてそれらのうちのひとつを選択しつつ、LMS アルゴリズムで係数更新することによって、OSLMS 予測器と同様の効果が得られるようにも思われる（結果は未確認である）．しかし、その場合の入力ベクトルに応じた予測器の選択は、明らかに容易ではない．もし、OSLMS 予測器をこのような複数の LMS 予測器との関連性としてとらえるなら、

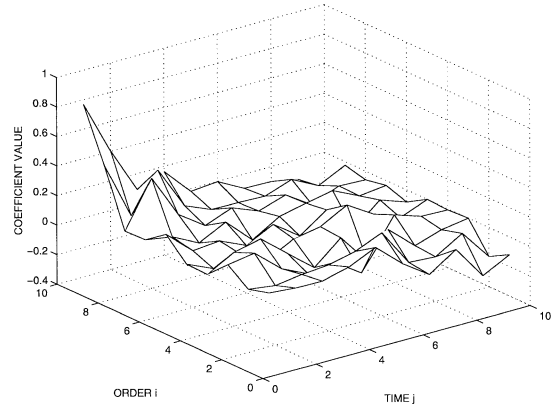


図 6 定常状態での OSLMS 予測器における係数行列
Fig. 6 Coefficient matrix for the OSLMS predictor at steady state.

いずれにしても係数行列を扱うことになり、OSLMS 予測器は係数行列からの係数選択を自動的に行うことから、効率性に優れた一方式と考えられる．

図 6 は、OSLMS 予測器における収束後の係数行列の一例を示している．時間軸 (Time j) 方向から順序軸 (Order i) 方向を見たとき、各係数の値がそれぞれの j において等しければ、そのときの OSLMS 予測器は線形予測器と同様線形処理のみを施すと考えられる．しかし、明らかに図 6 は非線形処理を施した結果を示している．図 5 の結果を加味すれば、図 6 の係数行列は、ここで取り上げた合成音声を高精度に表現する係数の組合せと解釈できる．また、もとの波形と OSLMS 予測器によって図 6 の係数行列から得られた予測波形を比較したところ、ほぼ一致した結果が得られたことをここに付記しておく．

3.2.3 実音声による実験 (単音声)

次に、実音声を用いて実験を行った．使用した音声はある男性によって発声された実母音 /a/ である．サンプリング周波数は 10 kHz である．LMS 予測器、VSLMS 予測器及び OSLMS 予測器を用いて予測分析を行い、予測誤差の収束特性と予測利得を評価する．使用した予測次数、ステップサイズ、安定化パラメータは LMS 予測器では $M = 10$, $\mu = 1.0$, $\beta = 0.05$, 文献 [8] を参考に VSLMS 予測器では $K_1 = 10$, $K_2 = 3$, $\mu = 1.0$, $\beta = 0.05$, OSLMS 予測器では $M = 10$, $\mu = 1.0$, $\beta = 0.05$, であり、優良の結果が得られるものである．ここで実音声を用いた予測分析では、合成音声のときのように完全な誤差の収束状態が得られない．そのため予測利得を算出する際には、式 (25) に

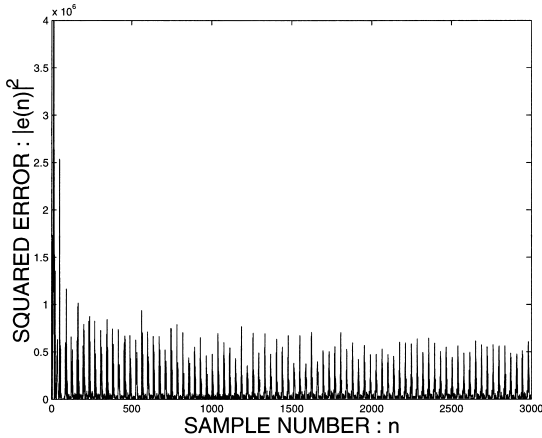


図 7 LMS 予測器における男性音/a/での予測誤差収束
Fig. 7 Residual errors for the LMS predictor on the real vowel /a/.

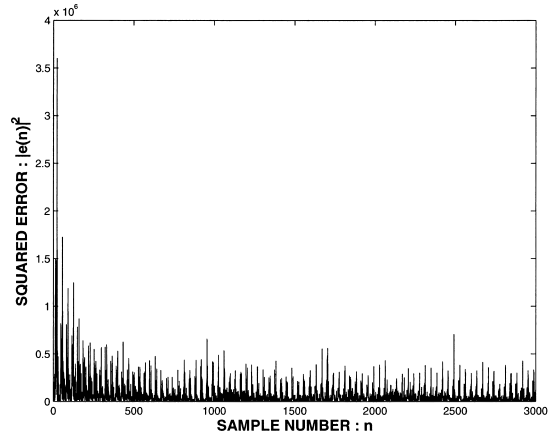


図 8 VSLMS 予測器における男性音/a/での予測誤差収束
Fig. 8 Residual errors for the VSLMS predictor on the real vowel /a/.

おける利用データ範囲を $n_1 = 2001, n_2 = 3000$ とし、この間を各予測器において十分な予測分析がなされた範囲と考えた。

3.2.4 実験結果

図 7, 図 8, 図 9 はそれぞれ, LMS 予測器, VSLMS 予測器, OSLMS 予測器の収束特性を示している. 図 7, 図 9 において双方の収束状態にある時点 (ここでは *sample number* $n = 2000 \sim 3000$) を比較すると, 合成音を用いた実験同様, OSLMS 予測器は LMS 予測器よりも予測誤差を低減させていることが分かる. また, VSLMS 予測器との比較においても, OSLMS 予測器はインパルス状の予測誤差を低減させている. しかしその反面, 収束状態を得るためには多くの音声サンプルを必要とすることも見てとれる. この原因としては, 合成音は機械的に作られ, 完全に周期的な音声波形を形成し得るが, 実音声は, 定常状態といえども時間の経過とともに周期波形の振幅が微妙に変化してしまい, OSLMS アルゴリズムの係数更新の規則性をかき乱すためと考えられる.

表 2 には, 予測利得を示している. 合成音を用いた実験では, OSLMS 予測器による大幅な予測利得の改善が見られたにもかかわらず, 実音声を用いた実験では改善度がそれほど顕著ではない. またここでは, LMS-L 予測器による実験も行っているが, 他の予測器に比べ明らかに劣化した結果となったため明記していない.

3.2.5 実音声による実験 (連続音声)

次に, より一般的な評価を得るために連続音声を用

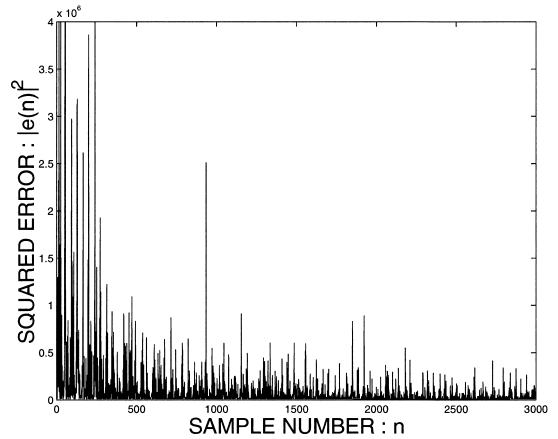


図 9 OSLMS 予測器における男性音/a/での予測誤差収束
Fig. 9 Residual errors for the OSLMS predictor on the real vowel /a/.

表 2 実音声における各予測器の予測利得
Table 2 SNR with each predictor on the real vowel /a/.

predictor	SNR (dB)
LMS	11.06
VSLMS	11.55
OSLMS	11.58

いた実験を行った. 使用した音声は, NTT アドバンステクノロジー (株) の「20ヶ国語音声データベース」に収録されている日本人話者の男女それぞれ 1 人によって発声された 10 秒程度の 1 短文, 「彼は以前から科学技術の進歩と人間の勇気が, 遙か

なる宇宙への旅を可能にしたのだと考えていました」(以降この文を“sentence1”とする)

であり、サンプリング周波数は 10 kHz である。LMS 予測器、VSLMS 予測器及び OSLMS 予測器を用いて予測分析を行い、予測利得を評価する。使用した予測次数、ステップサイズ、安定化パラメータは LMS 予測器では $M = 10$, $\mu = 1.0$, $\beta = 0.05$, VSLMS 予測器では $K_1 = 10$, $K_2 = 3$, $\mu = 0.5$, $\beta = 0.05$, OSLMS 予測器では $M = 10$, $\mu = 0.3$, $\beta = 0.05$ であり、最良の結果が得られるものである。ここでは、式(25)における利用データ範囲を $n_1 = 1$, $n_2 = 95000$ として予測利得を算出する。

3.2.6 実験結果

表 3 には、連続音声を用いた実験の予測利得を示している。単母音を用いた実験では、LMS 予測器に比べ、OSLMS 予測器が優れた予測精度を示していた。しかしながら連続音声を用いた実験では、LMS 予測器に比べ、OSLMS 予測器が 0.3~0.9 dB 予測精度が劣る結果となった。また、男性音声を対象にした実験では、VSLMS 予測器が三つの予測器の中で最も優れた予測精度を示したが、対象音が女性音声の場合は数値的に発散してしまった。またここでも、LMS-L 予測器による実験を行ったが、明らかに劣化する結果となったので明記していない。このことにより、以降の章では、LMS-L 予測器による結果を示さないことにする。

ここまでの合成音声と実音声の単母音と連続音声を用いた実験から、OSLMS 予測器は、合成音データのような完全周期的なデータサンプルを予測分析する場合は、その予測誤差が短時間で収束し、誤差量をも減少させ得るが、実音声のように完全な周期性をもたないデータサンプルを予測する場合は、短時間での予測誤差の収束は困難となるのが分かる。しかしながら OSLMS 予測器は、インパルス性残留誤差を低減する能力を明らかに有し、高い予測分析能力をもつこと

を示唆している。

4. 反復法

実音声、特に連続音声では、ほぼ定常とみなせる時間が、20~30 ms といわれており、この時間に対応するサンプルデータ数で予測分析が完了することが望まれる。しかしながら、前章で示したように、OSLMS 予測器は、実音声を高精度に予測分析するために、定常である多くのデータサンプルを必要とする。データサンプル数の低減は予測誤差の増大につながるため、結局 OSLMS 予測器でのデータサンプル数と予測精度の間にはトレードオフの関係が存在することになる。そこで、データサンプル数を低減し、かつ予測精度を向上させ得る方法として、反復処理を用いた予測分析法を本章では発展する。以下これを反復法と呼ぶことにする。反復法は、音声周期性を有することに着目して音声データを再利用し、少ないデータサンプルのみが与えられるときでも、多くのデータサンプル数を用いたときと同様の予測分析精度が得られる方法である。次節では、この反復法の具体的な手順を示す。

4.1 反復法の手順

反復法は 5 ステップからなる。

- (1) 予測係数行列の要素をすべて 0 に初期化する ($C(1) = 0$).
- (2) $n = 1$ から N まで適応処理を行う ($n = 1$ から N までが反復処理のための分析区間).
- (3) 最終的に求められた係数行列と入力ベクトルを格納する ($C(N)$ と $s(N)$).
- (4) (3) で得られた係数行列と入力ベクトルを初期値とする。
- (5) (2) に戻る。

定常とみなせる少ないデータサンプルからなる分析区間での予測誤差は、まだ収束に至っておらず、処理が不十分な状態にあると考えられる。しかし、その時点での予測係数行列、入力ベクトルを格納し、それらを初期値として使い、同一区間で再び予測分析をすることにより、予測分析区間を 2 倍に拡張した適応処理をすることが可能となる。同時にこれは、完全周期性により近いデータサンプルを適応処理していることも解釈できる。このような処理を繰り返し、その反復回数を増大すれば、定常なデータサンプルを多く用いた適応処理を行うことと同等の意味となる。すなわち、少ないデータサンプルのみを用いて、多くのデータサンプルを用いた予測分析と同等の予測分析精度を

表 3 連続音声における各予測器の予測利得
Table 3 SNR with each predictor on continuous speech.

predictor		SNR (dB)
LMS	male	9.66
	female	12.04
VSLMS	male	10.27
	female	-
OSLMS	male	8.81
	female	11.77

得るための解決策になると考えられる。一方、反復法では分析区間を用いているが、1. で述べたようにデータ数に対応する複数の予測パラメータを算出する意味において、適応処理の形態を保持していると考えられる。

次節以降では、反復法の有効性を示すため、男性と女性が発声した連続音声を用いて行った実験結果が示される。まず、反復法に基づく OSLMS 予測器の各パラメータの関係を調べる実験が行われる。その後、優良な結果を与えるパラメータを用いて、LMS 予測器と VSLMS 予測器を比較対象とし、反復法に基づく OSLMS 予測器の予測分析を行う。ここでは同時に、1 分析区間長の変化が予測精度に与える影響を調べる実験も行う。更には、反復法における予測利得算出方法を変化させる実験も行う。すべての実験の評価には、収束時点でのサンプル値を用いて算出された予測利得を用いる。

4.2 反復法と各パラメータの関係

反復法を適用した OSLMS 予測器の有効性と反復回数、ステップサイズ、予測次数依存性を調べるために、連続音声を用いた実験を行った。使用した音声は、3.2.5 で用いたものであり、男女 1 人ずつによって発声され、サンプリング周波数は 10 kHz である。具体的に使用した部分は、“科学技術の進歩と”の 3 秒程度の部分である。まず、反復回数を 5 回、10 回として、予測次数、ステップサイズを変化させた実験を行った。安定化パラメータは $\beta = 0.05$ としている。反復法のための 1 分析区間長は、連続音声における定常区間と考えられる 30 ms (300 サンプル) とし、連続した 100 区間分を分析する。ただし、分析区間のオーバーラップはしていない。そして合成音での実験同様、予測誤差がほぼ収束した時点での予測利得を以下の式（ここでの添字 w は、1 分析区間の切出しを意味する）に基づいて算出し、その平均を示すことにする。

$$SNR(\text{dB}) = 10 \log_{10} \frac{\sum_{n=n_1}^{n_2} s_w(n)^2}{\sum_{n=n_1}^{n_2} e_w(n)^2} \quad (26)$$

反復法を用いた予測分析を従来の適応予測分析と比較させるため、本実験では一例として、反復法を用いた予測分析を図 10 に示すようにして行った。まず、複数（予測利得算出方法によって異なるが、図 10 では三つ）の予測器を並列に配置する。ただし、連続音声において定常とみなせる区間（ここでは 30 ms）を 1 分析区間長とする。そして音声データが第 1 番目の予

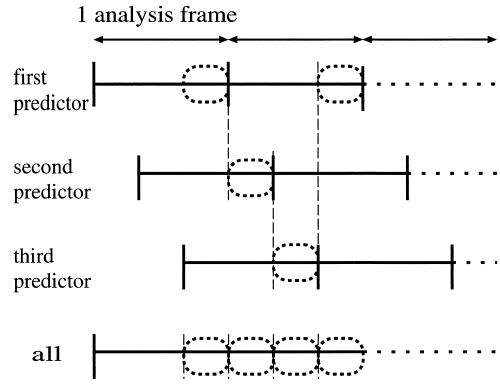


図 10 反復法の構成
Fig. 10 Configuration of iterative approach.

測器に入力され、反復処理が施されたのち、予測誤差が収束した時点での予測データを出力する（図 10 では収束時点を 20 ~ 30 ms とし、点線の で囲まれた範囲がそれを示している）。第 2 番目の予測器には 10 ms 後の音声データから入力されるように設定し、1 番目の予測器同様、反復処理の後、予測誤差が収束した時点での予測データが出力される。3 番目の予測器には 20 ms 後の音声から入力されるように設定して、1, 2 番目の予測器同様の処理が行われる。そして、また 1 番目の予測器が次の音声データ（はじめから考えると 30 ms 後のデータ）を処理するように設定する。これら一連の処理を繰り返し出力したデータをつなぎ合わせたものを、反復法を用いた実験結果として示すことにする。

表 4 には反復回数を 5 回、表 5 には反復回数 10 回としてステップサイズと予測次数を変化させたときに、式 (26) に基づき、 $n_1 = 201$, $n_2 = 300$ として算出した全 100 区間の予測利得の平均が示されている。これらの表から、反復回数の増加は、予測利得の上昇に関係することが分かる。また、ステップサイズに着目すると、次数、分析対象音にかかわらず、比較的小さな値のときに予測利得が最大値を示している。ステップサイズには、その値が比較的小さいときには、予測誤差の収束速度を減衰させるが、予測誤差の量を減少させるという特性が存在する。このようなステップサイズの特性を考慮すれば、ステップサイズが小さな値のときに予測利得の最大値を示したことは、反復法が、短時間予測を長時間予測と同様の予測に変換させ得たことを示唆するものである。そして、反復法を適用するときは、ステップサイズを小さい値にすることが有

表 4 反復回数 5 回での OSLMS 予測器における予測利得とステップサイズ, 予測次数の関係

Table 4 Effects of step size and predictive order on SNR in the case of 5 iterations for the OSLMS predictor.

	order	step size μ			
	M	0.1	0.3	0.5	1.0
male speech	10	11.80	12.49	12.24	10.52
	15	10.88	12.00	11.93	10.40
	20	10.02	11.49	11.55	10.12
female speech	10	10.11	10.98	10.81	8.94
	15	9.90	11.29	11.27	9.51
	20	9.75	11.43	11.46	9.79

表 5 反復回数 10 回での OSLMS 予測器における予測利得とステップサイズ, 予測次数の関係

Table 5 Effects of step size and predictive order on SNR in the case of 10 iterations for the OSLMS predictor.

	order	step size μ			
	M	0.1	0.3	0.5	1.0
male speech	10	12.99	13.44	13.06	11.11
	15	12.20	13.03	12.82	11.04
	20	11.38	12.56	12.48	10.77
female speech	10	11.05	11.67	11.38	9.34
	15	10.82	11.90	11.73	9.69
	20	10.68	11.91	11.76	9.75

効であることが分かる。ここでは、反復回数と予測次数を、示した値よりも更に増加させた実験も行っている。しかし、予測利得の大幅な改善が見られなかったため、ここでは明記していない。また、反復回数が 1 回のときでも予測利得が上昇したことをここに付記しておく。

4.3 LMS 予測器, VSLMS 予測器と反復法に基づく OSLMS 予測器との比較

次に、LMS 予測器と VSLMS 予測器による予測分析を従来法として、反復法を適用した OSLMS 予測器による予測分析との比較実験を行う。分析対象音声は、上記の実験で用いた連続音声 1 短文に加え、以下に示す 10 秒程度の 1 短文

「人々のびょう風絵と如来像に対する興味は、八百年の年月によって生じた表面の微妙な色彩変化にある」(以降この文を“sentence2”とする)

の合わせて 2 短文を男女それぞれ 3 人ずつが発声した短文、計 12 文とする。サンプリング周波数はすべて 10 kHz である。OSLMS 予測器で使用するパラメータは上記の実験結果を参考にして、ステップサイズ $\mu = 0.3$, 反復回数 10 回, 予測次数は、対象音声は男性音の場合は $M = 10$, 女性音の場合は $M = 20$ とする。そして、優良の結果が期待できる LMS 予測器で

表 6 LMS 予測器における予測利得とステップサイズ, 予測次数の関係

Table 6 Relation of step size and predictive order on SNR for the LMS predictor.

	order	step size μ			
	M	0.1	0.3	0.5	1.0
male speech	10	5.76	7.39	8.30	9.15
	15	5.56	7.39	8.20	9.15
	20	5.76	7.39	8.07	9.15
female speech	10	6.22	7.60	8.16	8.78
	15	6.23	7.54	8.16	8.90
	20	6.42	7.74	8.39	9.16

用いるパラメータを求めるために、LMS 予測器における予測利得とステップサイズ, 予測次数の関係を調べる実験を行った。分析対象音, 予測利得の求め方は先の OSLMS 予測器の実験と同様であり、反復回数を 0 回としたものである。表 6 がその結果を示している。

表 6 は、LMS 予測器における予測利得とステップサイズ, 予測次数の関係が示している。ここでは、ステップサイズ $\mu = 1.0$ 以上の結果を示していないが、これは $\mu = 1.0$ 以上において数値変化が誤差程度であったためである。表 6 の結果より、LMS 予測器で用いるパラメータは、ステップサイズ $\mu = 1.0$, 安定化パラメータ $\beta = 0.05$, 予測次数は分析対象音声は男性音の場合は $M = 10$, 女性音の場合は $M = 20$ として実験を行った。また VSLMS 予測器においても予測利得とステップサイズ, 予測次数の関係を調べる実験を行ったが、しばしば数値的に発散してしまった。そこで、極力、発散が起こらず、最も予測利得の改善が見込める、ステップサイズ $\mu = 1.0$, 線形部の予測次数 $K_1 = 10$, 非線形部の予測次数 $K_2 = 3$, 安定化パラメータ $\beta = 0.05$ を用いて実験を行うことにした。

連続音声を用いた実験における 1 分析区間長は 30 ms として、評価は式 (26) に基づき、 $n_1 = 201$, $n_2 = 300$ として算出された予測利得を用いた。更にここでは分析区間長と音声波形の周期との関係を考えて、分析区間長と音声波形の周期がちょうど一致した場合には、反復法は非常に良い結果を与えることが期待できる。しかし分析区間長と音声周期が一致しない場合には、反復法は予測精度を低下させる可能性もある。そこで 1 分析区間長を 24 ms と 36 ms としたときの実験を行い、そのときの予測利得も評価する。ただし、予測分析長の変更に伴い、予測利得算出時において、1 分析区間長が 24 ms のときは $n_1 = 161$, $n_2 = 240$, 36 ms のときは $n_1 = 241$, $n_2 = 360$ として予測利得を算出した。

表 7 男性が発声した連続音声の予測分析

Table 7 SNR on continuous speech uttered by male speakers.

		frame length (ms)	24	30	36	average
sentence1	LMS	adaptive (conventional)	8.26	8.90	9.26	8.93
	VSLMS	adaptive (conventional)	9.30	9.52	10.28	9.70
	OSLMS	adaptive	7.64	8.81	9.22	8.56
		no iterations	9.27	8.87	9.38	9.17
		10 iterations	11.54	11.26	11.76	11.52
sentence2	LMS	adaptive (conventional)	7.36	7.68	8.12	7.72
	VSLMS	adaptive (conventional)	9.10	8.83	8.27	8.73
	OSLMS	adaptive	8.89	9.34	9.99	9.41
		no iterations	10.51	10.32	10.02	10.28
		10 iterations	13.16	12.93	12.51	12.87

表 8 女性が発声した連続音声の予測分析

Table 8 SNR on continuous speech uttered by female speakers.

		frame length (ms)	24	30	36	average
sentence1	LMS	adaptive (conventional)	9.16	9.01	9.25	9.14
	VSLMS	adaptive (conventional)	9.27	10.50	11.01	10.26
	OSLMS	adaptive	9.62	10.25	10.89	10.25
		no iterations	9.48	9.90	10.27	9.88
		10 iterations	10.91	10.82	11.09	10.94
sentence2	LMS	adaptive (conventional)	8.29	8.07	8.48	8.28
	VSLMS	adaptive (conventional)	9.67	10.34	10.22	10.08
	OSLMS	adaptive	8.90	9.72	9.50	9.37
		no iterations	8.72	9.01	8.86	8.86
		10 iterations	12.48	12.17	12.54	12.40

表 7 は分析対象音声を 3 人の男性音声, 表 8 は分析対象音声を 3 人の女性音声として実験を行い, 1 分析区間長が 24 ms のときは連続する 417 区間, 30 ms のときは 333 区間, 36 ms のときは 278 区間で予測利得を求め, その平均を計算し, それらを用いて算出した 3 人の平均が示されている. ただし VSLMS 予測器における結果については, 短文によって数値的発散により正確なデータが得られない場合が生じた. そのため, 表 7 における “sentence2” の結果及び, 表 8 における “sentence1” と “sentence2” の結果は 2 人のみの平均を示すことにした. 表中にある, “adaptive” とは 10 秒程度の全音声データに適応処理を行ったものである. また, “no iterations” とは, 各分析区間で分析を始めるたびに, 反復は行わずに, 予測係数行列の初期値を 0 として分析を行ったものである. これらの表から, 反復法は, 分析対象音声, 分析区間の長さ, 音声の周期性に関係なく, 予測精度を向上させていることが分かる. 特に, 男性音声を予測分析した結果では, OSLMS 予測器は LMS 予測器よりも予測精度が劣っていたにもかかわらず, 反復法によってそれを逆転している. また, VSLMS 予測器をも上回る結果となっている. ここから, OSLMS 予測器のインパルス状の誤差を取り除く高い性能と, 反復法による少サンプル

数での予測精度の改善が確認できる.

一方, 分析区間で初期値を 0 としただけの “no iterations” の予測利得が連続処理に比べ, 増大するものと減少するものがあり, 分析対象音によって異なった結果が示されたことは興味深い事実である. 増大した原因として, 適応予測器における, 現在処理中のデータ以前のデータの影響を, 初期化により取り除いたからであると考えられる. 逆に減少した原因としては, 以前のデータと現在処理中のデータとの間に強い相関があったにもかかわらず, 初期化によってそれを取り除いてしまったからであると考えられる.

4.4 反復法と予測利得算出方法の関係

これまでの実音声における実験では, 反復法のための 1 分析区間を 30 ms とし, 式 (26) に基づいて, $n_1 = 201$, $n_2 = 300$ としたとき, 100 データサンプルをベースとして予測利得を算出していた. しかしながら, 予測利得の算出に使用するデータサンプル数の変化によって予測利得が異なった結果となることも考えられる. また使用データサンプル数の変化に伴う, 図 10 における予測器数の変化が予測利得に影響を与えることも考えられる. そこで 1 分析区間を 30 ms としたとき, 予測利得算出に用いるデータサンプル数を 50 サンプルと 150 サンプルに変化させた実験を行っ

表 9 50 サンプルのときの予測利得
Table 9 SNR with 50 samples.

		speaker	male	female	average
OSLMS	sentence1	no iterations	9.12	10.02	9.57
		10 iterations	11.44	12.61	12.03
	sentence2	no iterations	10.12	9.00	9.56
		10 iterations	12.83	12.32	12.58

表 10 150 サンプルのときの予測利得
Table 10 SNR with 150 samples.

		speaker	male	female	average
OSLMS	sentence1	no iterations	9.03	10.34	9.69
		10 iterations	11.62	12.97	12.30
	sentence2	no iterations	9.80	9.01	9.41
		10 iterations	12.54	12.20	12.37

た．使用したパラメータは先の実験と同様とした．ただし，データサンプル数が 50 サンプルのときは図 10 と同様に，六つの予測器の結果，150 サンプルのときは二つの予測器の結果をつなぎ合わせて，全体の結果として，予測利得を算出した．

表 9 は，予測利得算出に用いるデータサンプル数を，式 (26) に基づき， $n_1 = 251$ ， $n_2 = 300$ として 50 データサンプル，表 10 は $n_1 = 151$ ， $n_2 = 300$ として 150 データサンプルを用いて，連続する 333 区間で算出した予測利得の平均を示している．これらを表 7，表 8 と比較してみると，データサンプル数や予測器数を変化させても，予測利得に大きな変化は見られないことが分かる．しかしながら，データサンプル数を増加させるほどに，ごく微小ではあるが予測利得が減少している場合が見られる．これは，データサンプル数を増やすことによって，まだ予測誤差が収束しきれていない部分のデータサンプルが予測利得算出のために用いられたためと考えられる．

5. む す び

本論文では，音声信号のための適応非線形予測器を提案した．Pitas らによって導出された LMS-L フィルタに，入力信号の時間情報を保持するといった改良を加え，OSLMS 予測器を導いた．その適応アルゴリズムである OSLMS アルゴリズムは，係数行列からの更新する係数の選択を必要とするものの，基本的には LMS アルゴリズムと同一の係数更新式からなり，予測器としてのフィルタ構成の規模及び計算量は従来の線形予測器と同様少なくすむ．そして合成音を用いた実験において，予測誤差の収束特性と予測利得を示すことにより，OSLMS 予測器が LMS 予測器及び

VSLMS 予測器よりも優れていることを示した．

しかしながら，実音声を用いた実験では OSLMS 予測器は，その優れた能力を十分に発揮できなかった．これは，実音声で完全周期をもたない信号であるため，OSLMS アルゴリズムの係数行列の更新の規則性が保たれにくくなるのが原因であった．そこで，連続音声に対処することを考え，少ないデータサンプルを用いたときでも，予測精度を改善することを目的として，新たに反復法を提案し，これを OSLMS 予測器に組み入れることを試みた．そして，反復法を適用した OSLMS 予測器の有効性を確かめるべく連続実音声を用い，比較対象を LMS 予測器，VSLMS 予測器とした実験を行った．更には分析区間長と音声波形の周期，それに伴う予測精度の算出方法の変化を考慮した実験も行った．結果としては，反復法により予測精度が大幅に改善され，単独の適応予測では，LMS 予測器や VSLMS 予測器に劣っていた予測精度をも覆すことが可能とした．

本論文における連続音声での実験は，無音区間，有音区間を区別せず，全区間を予測分析対象として行った．3. における合成音声，実単音声をを用いた実験から，有音に対する OSLMS 予測器の予測分析は優れていることが確認できる．しかし，表 7，表 8 からは，有音区間，無音区間にかかわらず，反復法を用いた OSLMS 予測器は，予測精度を向上させる潜在能力があると期待できる．今後は，音声の種類に対する実行依存性を詳細に検討する予定である．

文 献

- [1] B.S. Atal and S. Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave," J. Acoust. Soc. Am., vol.50, no.2, pp.637-655, 1971.
- [2] S. Chandra and W.C. Lin, "Experimental comparison between stationary and nonstationary formulations of linear prediction applied to voiced speech analysis," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-22, no.6, pp.403-415, 1974.
- [3] L.R. Rabiner, B.S. Atal, and M.R. Sambur, "LPC prediction error — Analysis of its variation with the position of the analysis frame," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-25, no.5, pp.434-442, 1977.
- [4] E. Varoglu and K. Hacioglu, "Recurrent neural network speech predictor based on dynamical systems approach," IEE Proc. Vision Image Signal Process, vol.147, no.2, pp.149-156, 2000.
- [5] J.D. Gibson, J.L. Melsa, and S.K. Jones, "Digital speech analysis using sequential estimation techniques," IEEE Trans. Acoust. Speech Signal Process.,

vol.ASSP-23, no.4, pp.362-369, 1975.

- [6] J.D. Gibson, S.K. Jones, and J.L. Melsa, "Sequentially adaptive prediction and coding of speech signals," IEEE Trans. Commun., vol.COM-22, no.11, pp.1789-1797, 1974.
- [7] I. Pitas and A.N. Venetsanopoulos, "Adaptive filters based on order statistics," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-39, no.2, pp.518-522, 1991.
- [8] E. Mumolo, A. Carini, and D. Francescato, "AD-PCM with non linear predictors," Proc. EUSIPCO94, pp.387-390, 1994.
- [9] J. Thyssen, H. Nielsen, and S.D. Hansen, "Non-linear short-term prediction in speech coding," Proc. ICASSP94, pp.185-188, 1994.
- [10] F. Diaz-de-Maria and A.R. Figuiras-Vidal, "Nonlinear prediction for speech coding using radial basis functions," Proc. ICASSP95, pp.788-791, 1995.
- [11] K.K. Paliwal and P.V.S. Rao, "A modified autocorrelation method of linear prediction for pitch-synchronous analysis of voiced speech," Signal Process., vol.3, no.2, pp.181-185, 1981.
- [12] 守谷健弘, 音声符号化, 電子情報通信学会, 1998.
(平成 15 年 6 月 23 日受付, 11 月 4 日再受付,
16 年 3 月 1 日最終原稿受付)



島村 徹也 (正員)

昭 61 慶大・理工・電気卒。平 3 同大大学院博士課程了。工博。同年埼玉大・工・助手。平 10 同助教授, 現在に至る。この間, 平 7 ラフバラ大学, 平 8 ベルファーストクイーンズ大学(ともに連合王国)客員研究員。デジタル信号処理とその音声, 通信システムへの応用に関する研究に従事。日本音響学会, IEEE, EURASIP 各会員。



田中 啓文 (学生員)

平 14 埼玉大・工・情報システム卒。現在, 同大学院博士前期課程に在学中。デジタル音声処理に関する研究に従事。



早川 晴子

平 11 埼玉大・工・情報システム卒。同年, (株)インテックに入社。現在(株)インテックコミュニケーションズ。在学中はデジタル音声処理に関する研究に従事。