

Title	サロゲート法に基づく持続発声母音の基本周期揺らぎの高次相関解析
Author(s)	徳田, 功; 宮野, 尚哉; 合原, 一幸
Citation	電子情報通信学会論文誌 A, J87-A(3): 355-363
Issue Date	2004-03-01
Type	Journal Article
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/4722">http://hdl.handle.net/10119/4722</a>
Rights	Copyright (C)2004 IEICE. 徳田功, 宮野尚哉, 合原一幸, 電子情報通信学会論文誌 A, J87-A(3), 2004, 355-363. <a href="http://www.ieice.org/jpn/trans_online/">http://www.ieice.org/jpn/trans_online/</a>
Description	

サロゲート法に基づく持続発声母音の基本周期揺らぎの高次相関解析

徳田 功<sup>†a)</sup> 宮野 尚哉<sup>††</sup> 合原 一幸<sup>†††</sup>

Higher-Order Correlation Analysis of Pitch Fluctuations in Sustained Normal Vowels by the Method of Surrogate Data

Isao TOKUDA<sup>†a)</sup>, Takaya MIYANO<sup>††</sup>, and Kazuyuki AIHARA<sup>†††</sup>

あらまし 有声音の持続発声で観測される、ピッチの基本周期の揺らぎは、音声が入間的で自然な発音として知覚されるために必要不可欠であることが知られている。音声合成を行う際に、原音声の品質を保持するためには、ピッチの基本周期揺らぎのどのような統計的性質を再現することが重要なのか、慎重に検討する必要がある。これまでの研究から、2次の相関構造を保存することが、自然な基本周期揺らぎを実現するために重要であることが分かっているが、3次以上の高次相関の重要性については分かっていない。本研究では、サロゲート法を用いて、基本周期揺らぎの高次相関解析を行い、3次以上の相関構造が、母音の自然性に与える影響について調べる。得られた結果から、ピッチの基本周期揺らぎを合成する際には、2次の相関に加えて、3次以上の高次の相関構造を再現することが重要であることを指摘し、非線形モデル等によるアプローチの可能性について議論を行う。キーワード 持続発声母音、基本周期揺らぎ、サロゲート法、自然性の知覚、一対比較法

1. ま え が き

人の発する有声音は、ピッチと呼ばれるほぼ同じ形状をした単位波が繰り返される構造をもつ。このピッチは、一見すると整然として周期的に見えるが、実際には、ピッチの基本周期及び波形は時間発展による繰返しとともに微妙に揺らぐことが古くから知られている[1]。このピッチ揺らぎは、音声が入間的で自然な発音として知覚されるために必要不可欠であることが、音声研究者の様々な実験結果から分かっている。小室らは、ピッチの基本周期の揺らぎが、合成音の品質に大きな影響をもち、原音声の品質を保つには、基本周期揺らぎの低スペクトル成分を保存することが重要であることを指摘している[3]。小林らは、基本周期揺

らぎのスペクトル構造を再現し、合成に用いるためのARフィルタを提案している[2]。一方、伊福部らは、ピッチの波形揺らぎも重要であることを指摘し[4]、青木らは、波形揺らぎを実現するためのフラクタル合成法などを提案している[5]。また、非線形力学系理論に基づく音声信号の解析も近年活発に行われており[6]~[10]、佐藤ら[11]、徳田ら[12]~[14]、池口ら[15]は、母音の非線形力学系構造を解析し、その揺らぎのモデル化を検討している。

規則合成によって自然な音声を合成するためには、ピッチ揺らぎの構造を更に掘り下げて解析することが重要であり、その研究をもとに、ピッチ揺らぎを効果的に生成するモデルを開発してゆく必要がある。ピッチの基本周期揺らぎのスペクトル構造に着目した小室ら[3]及び小林ら[2]の研究は、この意味で非常に重要な結果といえるが、スペクトルよりも高次の相関構造の重要性についてはいまだ明確にされていないのが現状である。そこで、本研究では、サロゲート法[21],[22]を用いて、基本周期揺らぎの高次相関解析を行い、3次以上の相関構造が、母音の自然性に与える影響について調べることにする。サロゲート法は、非線形力学系理論に基づく時系列解析[18],[19]の結果を検証する方法として発展してきたが、Schreiber[23]

<sup>†</sup> 室蘭工業大学工学部情報工学科, 室蘭市  
Department of Computer Science and Systems Engineering, Muroran Institute of Technology, Muroran-shi, 050-0071 Japan

<sup>††</sup> 立命館大学 COE 推進機構, 草津市  
Center for Promotion of the COE Program, Ritsumeikan University, Kusatsu-shi, 525-8577 Japan

<sup>†††</sup> 東京大学大学院情報理工学系研究科数理情報学専攻, 東京都  
Department of Mathematical Informatics, The University of Tokyo, Bunkyo-ku, Tokyo, 113-8656 Japan

a) E-mail: tokuda@agnld.uni-potsdam.de

によって、高次相関を保存するデータを生成するアルゴリズムが提案され、近年ますます、応用の幅に広がりを見せている。実験では、Schreiber のアルゴリズムに基づいて、高次相関を保存した基本周期揺らぎを合成し、高次相関を保存することが、合成音の品質を保つために重要であるかを検証する。

以下、2. では、本研究で解析に用いる音声データについて紹介し、その基本周期揺らぎの解析を行う。3. では、サロゲート法に基づいて、基本周期揺らぎの高次相関を保存したデータの作成を行い、4. では、その合成音の音質を評価する聴覚心理実験を行う。5. で、実験結果に対する考察を行い、今後の展望について述べる。

## 2. 基本周期揺らぎの解析

### 2.1 音声データ

音声データには、男性話者 2 名 (oo, ao), 女性話者 2 名 (ka, yo) がそれぞれ発声した母音/a/の持続発声音を用いた。4 名の話者は、青森放送株式会社に勤務するアナウンサーで、声帯等に病的異常のない健全な男女とし、録音は青森放送スタジオ内で行われた。音声データはカットオフ周波数 10 kHz, 16 ビット解像度, 22.05 kHz サンプル周波数でデジタル計測を行った。母音の定常的な発声部分を解析の対象とするため、発声初期の不安定な立ち上がり部分と終端の減衰部分は除去したデータを用いることにする。定常部分の発声の長さは、表 1 にまとめたとおりである。

### 2.2 ピッチ抽出

まず、音声データのピッチ情報を取り出すため、零交差計数法 [20] に基づいて、音声信号からピッチの基本周期列

$$\{T_i : i = 1, 2, \dots, N\} \quad (1)$$

を抽出する (図 1)。零交差解析を行った後、視覚による検証も行い、ピッチの欠損がないことを確認したところ、4 名の話者 (oo, ao, ka, yo) の音声に対して、 $N = 137, 71, 223, 254$  の基本周期がそれぞれ抽出された。一例として、図 2 (a) に話者 oo から抽出された基本周期列を示す。

次に、基本周期列の揺らぎの大きさを評価するため、以下に定義される正規化分散 (coefficient of variation) [16]

$$\alpha = \frac{100}{\bar{T}} \cdot \sqrt{\frac{1}{N} \sum_{i=1}^N (T_i - \bar{T})^2} \quad [\%] \quad (2)$$

表 1 4 人の話者 (oo, ao, ka, yo) から計測された持続発声音 /a/ のデータ長, 抽出されたピッチの数, 平均のピッチ間隔, 及び, ピッチ揺らぎの正規化分散

Table 1 Data length, number of extracted pitches, mean pitch period, and coefficient of variation of a sustained vowel /a/ recorded from 4 subjects (oo, ao, ka, yo).

話者	データ長	ピッチ数	平均ピッチ間隔	ピッチ揺らぎ
oo	800 msec	137	5.86 msec	0.873 %
ao	450 msec	71	6.26 msec	0.778 %
ka	900 msec	223	4.06 msec	0.773 %
yo	900 msec	254	3.56 msec	1.10 %

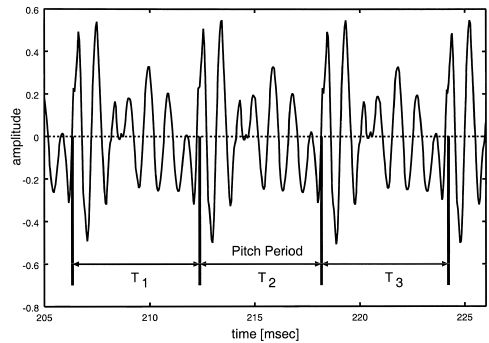


図 1 零交差点法によるピッチとその基本周期列  $\{T_1, T_2, \dots\}$  の抽出

Fig. 1 Pitch extraction via the zero-crossing method and generation of the pitch period sequence  $\{T_1, T_2, \dots\}$ .

を計算した。ただし、 $\bar{T} = \frac{1}{N} \sum_{i=1}^N T_i$  とする。4 名の話者 (oo, ao, ka, yo) の音声に対する基本周期の平均及び正規化分散は、表 1 にまとめたとおりである。文献 [17] によると、健全な男女の基本周期揺らぎの正規化分散は、 $1.05 \pm 0.40\%$  の範囲に分布することが報告されている。4 名の話者の正規化分散は、正常な値の範囲にあり、このことから、4 名の話者の音声は、標準的な基本周期揺らぎを有したデータとして適切な解析対象であることが確認された。

## 3. サロゲート法

非線形力学系理論に基づく時系列解析とその実データへの応用 [18], [19] が近年脚光をあびているが、この非線形解析の結果の信頼性を検証するための方法として導入されたのがサロゲート法 (method of surrogate) [21], [22] である。サロゲート法は統計的仮説検定 (statistical hypothesis testing) の一種であるが、仮説検定を行う際に、オリジナルの時系列データのある種の統計量を厳密に、あるいは近似的に、保存する

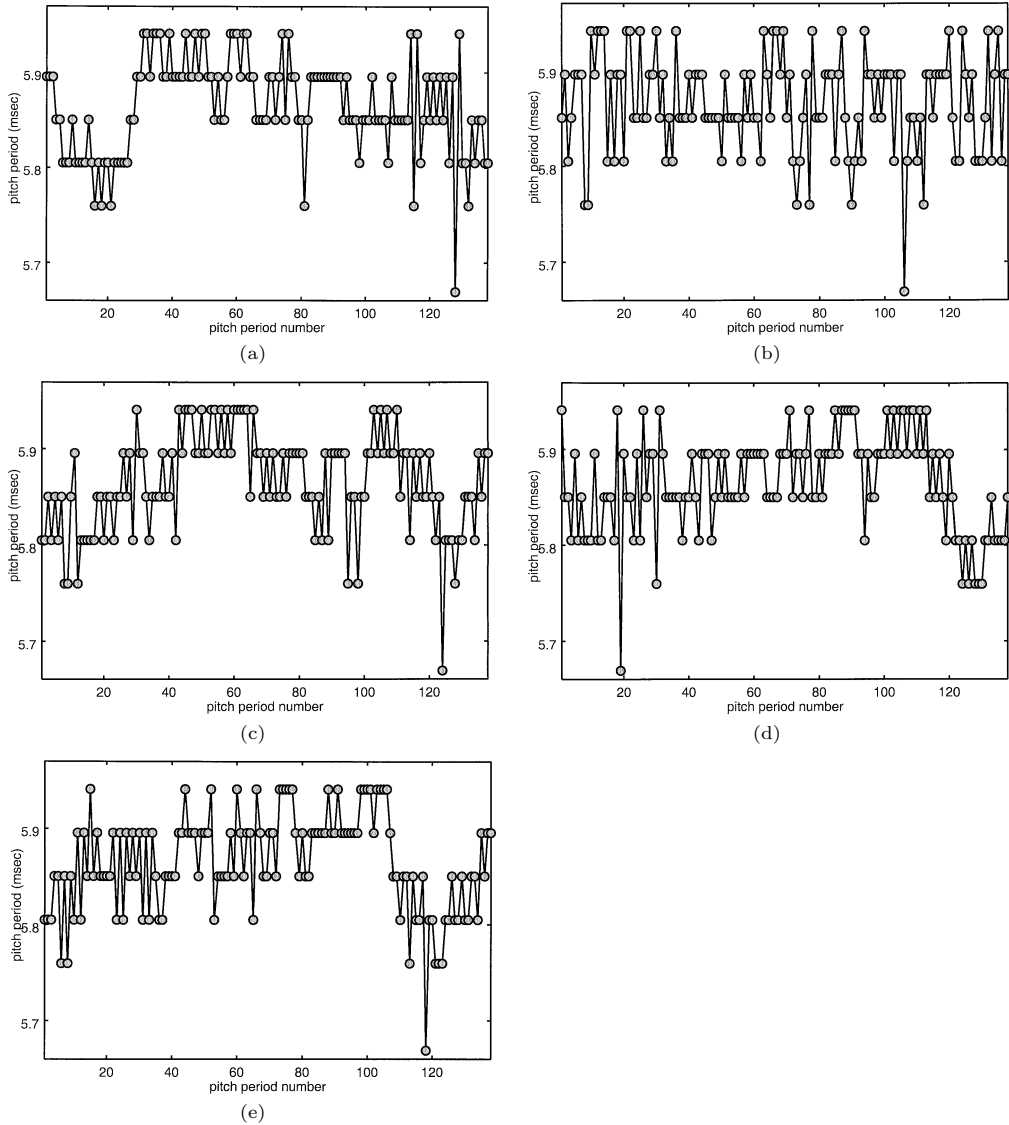


図 2 (a) 話者 oo から抽出されたオリジナルの基本周期列  $\{T_1, T_2, \dots, T_{137}\}$ . (b) (a) の基本周期列をランダムに並べ換えたデータ. (c) 2 次相関関数保存型サロゲートデータ. (d) 3 次相関関数保存型サロゲートデータ. (e) 4 次相関関数保存型サロゲートデータ.

Fig. 2 (a) Original pitch period sequence  $\{T_1, T_2, \dots, T_{137}\}$  extracted from the subject oo. (b) Surrogate data generated by random shuffling of the original data (a). (c) Surrogate data that preserves the 2nd order correlation function. (d) Surrogate data that preserves the 3rd order correlation function. (e) Surrogate data that preserves the 4th order correlation function.

ようなサロゲートデータを作る点に特徴がある．このような性質を利用すれば，基本周期揺らぎのどのような統計的性質が，自然な音声の知覚には重要であるかを検定することができる．ランダムシャッフル型サロ

ゲート，FT 変換型サロゲート [21] など，様々なサロゲートデータ作成法がこれまでに提案されているが，本実験では，Schreiber [23] の方法に基づいて，オリジナルの基本周期列  $\{T_i\}$  の高次相関を保存するサロ

ゲートデータを作成する．サロゲートデータに対して一対比較法による音質評価を行い，高次相関を保存することが，自然な基本周期揺らぎを実現するのに重要であるかについて，検証を行う．

3.1 高次相関関数

サロゲートデータを作成する前に，高次相関関数の定義を行う．まず，基本周期列  $\{T_i\}$  に対して，その平均を  $\bar{T}$ ，分散を  $\sigma$ ，遅れ時間を  $\tau$  としたときの中心モーメントは

$$\begin{aligned} & \underbrace{\mu_{rst\dots}(\tau)}_{m \text{ times}} \\ &= \frac{1}{N - (m-1)\tau} \cdot \frac{1}{\sigma^{r+s+t+\dots}} \\ & \cdot \sum_{t=1}^{N-(m-1)\tau} (T_i - \bar{T})^r \cdot (T_{i+\tau} - \bar{T})^s \\ & \cdot (T_{i+2\tau} - \bar{T})^t \dots \end{aligned} \quad (3)$$

と書ける．このとき，2次のキュムラント  $\kappa_2(\tau)$  は，

$$\begin{aligned} \kappa_2(\tau) &= \mu_{11}(\tau) \\ &= \frac{1}{\sigma^2} \cdot \frac{1}{N - \tau} \cdot \sum_{i=1}^{N-\tau} (T_i - \bar{T}) \cdot (T_{i+\tau} - \bar{T}) \end{aligned} \quad (4)$$

となり，自己相関関数に等しくなる．これを2次の自己相関関数：

$$\kappa_2(\tau) = \mu_{11}(\tau) \quad (5)$$

と定義する．同様に，3次及び4次の自己相関関数は以下のように定義できる [18], [24]．

3次自己相関関数：

$$\kappa_3(\tau) = \mu_{111}(\tau) \quad (6)$$

4次自己相関関数：

$$\begin{aligned} \kappa_4(\tau) &= \mu_{1111}(\tau) - \mu_{0011}(\tau) \cdot \mu_{1100}(\tau) - \mu_{1010}(\tau) \\ & \cdot \mu_{0101}(\tau) - \mu_{1001}(\tau) \cdot \mu_{0110}(\tau) \end{aligned} \quad (7)$$

本研究では，4次の相関関数までを扱うことにする．更に高次の相関関数の定義については，文献 [18], [24] を参照されたい．

3.2 高次相関保存型サロゲートデータ

音声信号から抽出された，オリジナルの基本周期列  $\{T_1, T_2, \dots\}$  が与えられたとき，その高次の自己相関関数を保存するサロゲートデータ  $\{T'_1, T'_2, \dots\}$  は，

以下のようにして作成される．いま， $d$  次までの相関関数  $\kappa_d$  を保存する場合について考える．

[  $d$  次相関関数保存型サロゲートデータの作成法 ]

(C1) ランダムシャッフリング法 [21] を用いて，オリジナル系列  $\{T_i\}$  をランダムに並べ換えた系列  $\{T'_i\}$  を生成する．この変換によって，系列の順序は変化するが，基本周期の分布関数は保存されることに注意する．

(C2) オリジナル系列  $\{T_i\}$  とランダム系列  $\{T'_i\}$  の  $d$  次までの自己相関関数を  $\kappa_k, \kappa'_k$  ( $k = 2, \dots, d$ ) とそれぞれ書くとき，その誤差関数を以下のように定義する．

$$E(\{T'\}) = \sum_{k=2}^d \sum_{\tau} \{\kappa_k(\tau) - \kappa'_k(\tau)\}^2 \quad (8)$$

シミュレーテッドアニーリング (SA) 法に基づいて，誤差関数 (8) を極小化するようなランダム系列  $\{T'\}$  を求め，これをサロゲートデータとする．ただし，SA 法では，各アニーリングステップごとに，二つの異なるデータ点  $(T_i, T_j) (i \neq j)$  をランダムに選択し，その入れ換え，すなわち， $\{\dots, T'_i, \dots, T'_j, \dots\} \rightarrow \{\dots, T'_j, \dots, T'_i, \dots\}$  を確率的に行う．SA 法の温度係数は，アニーリングステップ  $l$  に対して， $t = 1/\log(l)$  と冷却し，各ステップごとの誤差変動  $\Delta E(T_i \leftrightarrow T_j)$  に対して， $1/\{1 + \exp(\Delta E/t)\}$  の確率でデータ点の入れ換えを行う．

話者 oo から抽出された基本周期列  $\{T_i\}$  に対して，4次までの自己相関関数を保存するサロゲートデータ  $\{T'_i\}$  を生成したときの誤差関数の変化を図 3 に示す．アニーリングステップ  $l \approx 2.5 \cdot 10^4$  で SA 法はほぼ収束し，誤差の極小化が行われている．実際，生成さ

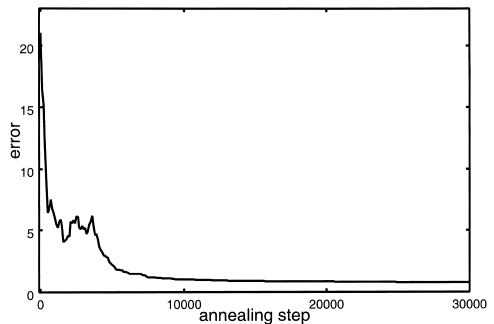


図 3 シミュレーテッドアニーリング法による誤差関数 (8) の極小化のプロセス

Fig. 3 Process of minimizing the error function (8) by the simulated annealing method.

れたサロゲートデータ ( 図 2 (e) ) はオリジナル系列 ( 図 2 (a) ) とは異なる時系列構造を有しているにもか

かわらず, 図 4 に示すとおり, 2 次, 3 次, 及び, 4 次の自己相関関数はオリジナルをよく近似していることが分かる .

#### 4. 実 験

前節で作成したサロゲートデータをもとにして, 本章では, 合成音の自然性に関する聴覚心理実験を行う . 高次相関保存型のサロゲートデータを基本周期列に用いて音声合成を行い, 合成音の自然性をサーストンの一対比較法に基づいて評価する .

##### 4.1 音声合成法

現在主流となっている音声合成法には, 線形予測法 ( LPC ) に代表される分析合成型と音声波形をそのまま符号圧縮する波形符号化型の二つのタイプが存在する [20] . 一般に, 符号量を減少させるのには, 分析合成型が有効であるとされ, 通信等に幅広く応用されているが, 音質面では符号量の多い波形符号化型の方が優れていると考えられている . 本研究では, 基本周期揺らぎの高次相関構造の微妙な差異を知覚できるように, 音質の優れた波形符号化のアプローチを採用することにする . 基本周期列  $\{T_i\}$  が与えられたとき, その合成音は以下のようにして作成する ( 図 5 参照 ) .

[ 音声合成法 ]

(S1) オリジナルの音声信号から単一のピッチ波を一つ抽出し, 基本波形とする .

(S2) 基本波形のコピーを  $N$  個用意して, その基本

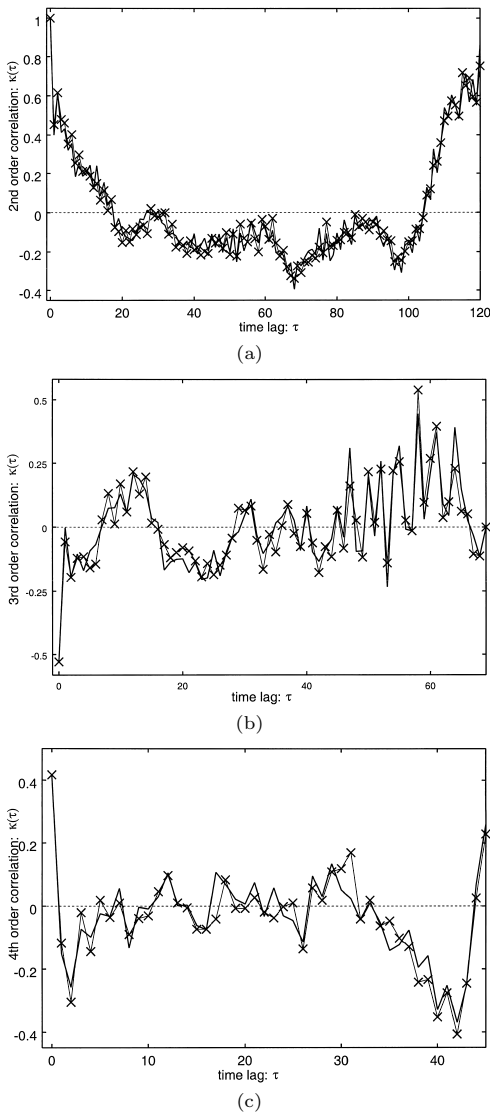


図 4 話者 oo のオリジナルの基本周期列の自己相関関数 ( 太線 ) と 4 次相関関数保存型サロゲートデータの自己相関関数 ( × 印 ) . (a) 2 次の自己相関関数  $\kappa_2(\tau)$  , (b) 3 次の自己相関関数  $\kappa_3(\tau)$  , (c) 4 次の自己相関関数  $\kappa_4(\tau)$  .

Fig. 4 Higher order correlation functions of the subject oo's original pitch sequence ( bold line ) and the surrogate data that preserves the 4th order correlation function ( crosses ) . (a) The 2nd order correlation functions  $\kappa_2(\tau)$  , (b) The 3rd order correlation functions  $\kappa_3(\tau)$  , (c) The 4th order correlation functions  $\kappa_4(\tau)$  .

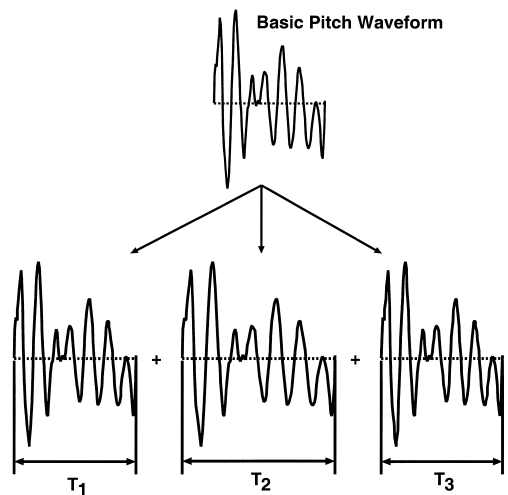


図 5 波形編集による音声合成法

Fig. 5 Speech synthesis based on waveform editing.

周期間隔を系列  $\{T_1, T_2, \dots, T_N\}$  に従って伸縮する。ただし、音声の振幅方向には、伸縮を行わない。  
(S3) 伸縮したコピー波を系列の順序に従って再結合し、音声信号とする。

本方式は、ピッチ波形を編集することを基本としており、伊福部らの実験 [4] に準拠したものととらえることができる。ステップ (S3) でコピー波を再結合する際に、信号の不連続性が生じ、ノイズの原因となる危険性があるが、この問題を回避するために、基本となるピッチ波形の始点と終点はほぼ連続になるように抽出を行い、再結合点の周辺には平均値フィルタを施し、平滑化を行った。ステップ (S2) で基本となるピッチ波形を伸縮する係数としては、最大振幅と基本周期の二つの要素が考えられる。ただし、これまでの予備実験から、自然性の知覚には基本周期の方が最大振幅よりもはるかに重要であることが分かっている [25]。具体的には、上記の方法に基づいて原音声から合成音を作成し、比較聴覚実験を行ったところ、音声の最大振幅を一定に固定しても原音声の自然性はあまり損なわれないのに対して、基本周期を一定に固定した場合には、原音の自然性が著しく損なわれることが確認された。このことから、最大振幅よりも基本周期の揺らぎの方が聴覚的な影響が大きいと考え、本実験では、ピッチ波形の最大振幅は一定に固定し、基本周期のみに揺らぎを加えることにした。

#### 4.2 実験

各話者の音声データに対して、以下の 5 種類の基本周期列：

- (a) 音声データから抽出されたオリジナルの周期列
- (b) オリジナルの周期列をランダムに並べ換えた周期列
- (c) オリジナルの周期列の 2 次の自己相関関数を保存した周期列
- (d) オリジナルの周期列の 3 次までの自己相関関数を保存した周期列
- (e) オリジナルの周期列の 4 次までの自己相関関数を保存した周期列

を用意して、前小節の方法に従って合成音を作成し、聴覚心理実験を行った。各合成音の自然性は一対比較法に基づいて評価した。一対比較法は 5 種類の合成音に対して、二つずつのすべての組合せを評定者に順次提示し、「自然性」を判断基準として、一方の音声を選択してもらおう方法である。複数の評定者の判定結果をサーストンの比較判断の法則に基づいて評定すること

により、合成音の自然性の相対尺度値を求めることができる [26]。実験では、聴覚実験に関する訓練を受けていない 20 名の一般の男女に評定を依頼した。

4 名の話者 (oo, ao, yo, ka) のそれぞれの音声から作成された 5 種類の合成音の自然性を評価した実験結果を図 6 に示す。4 名の話者のすべての結果において、基本周期をランダムに並び換えた合成音は自然さを著しく損なっていることが分かる。これは、音声の基本周期揺らぎが何らかの力学規則に従っており、全くランダムに揺らいでいるものではないことを示している。自然性は 2 次の相関を保存することで、大きく改善され、更に高次の相関を保存することで徐々にオリジナルの周期列の自然性に近づいてゆく。2 次の相関を保存することで自然性が大きく修復されることは、基本周期のスペクトル構造が重要であることを指摘した小室ら [3] 及び小林ら [2] の結果の妥当性を立証している。ただし、3 次以上の高次相関を保存することで音質が更に改善した本実験結果は、音声の基本周期揺らぎが小林らの提案した AR 過程としてもとらえきれないことを指摘している。線形ガウス過程よりも複雑な非線形の力学プロセスや非ガウスの確率過程が背後に存在することを示唆している。

話者 ao の実験で、4 次の相関関数保存型データが低い自然性を示しているが、これは ao から抽出された基本周期の数が他の話者に比べて少なく (表 1 参照)、ピッチ揺らぎの相関構造の相違が知覚しにくかったためと考えられる。実際に、基本周期列  $\{T_1, \dots, T_{71}\}$  及びそのサロゲート系列  $\{T'_1, \dots, T'_{71}\}$  を反復して、2 倍の長さの系列  $\{T_1, \dots, T_{71}, T_1, \dots, T_{71}\}$ ,  $\{T'_1, \dots, T'_{71}, T'_1, \dots, T'_{71}\}$  にして、同様の比較実験を行ったところ、図 7 に示すとおり、他の話者と同様の結果が得られた。このことから、高次相関構造の重要性は、男女の話者に依存しない、共通の性質とすることができる。

#### 5. むすび

音声信号の基本周期揺らぎの高次相関構造に着目し、自然な音声を作成するために重要な統計的性質を、サロゲート法に基づいて解析した。聴覚実験から得られた結果と考察をまとめると以下のとおりとなる。

(1) 基本周期をランダムに並べ換えた場合、合成音は自然さを著しく損なう。これは、音声の基本周期が全く独立でランダムに揺らぐものではないことを意味する。

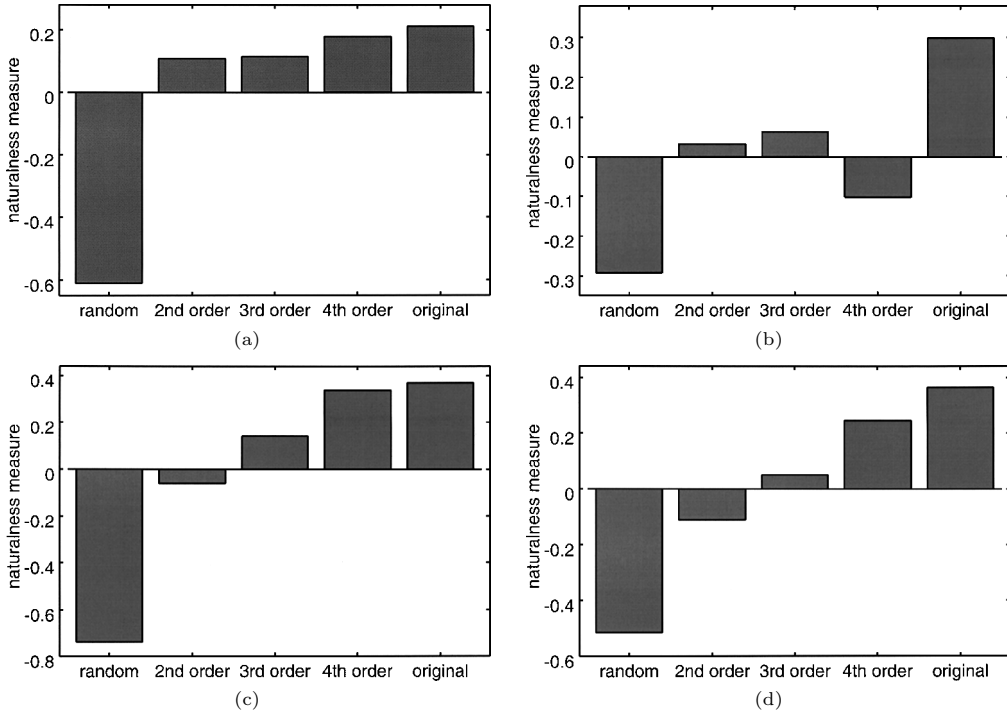


図 6 一対比較法による自然性の評価．左から，ランダム型サロゲート，2次相関保存型サロゲート，3次相関保存型サロゲート，4次相関保存型サロゲート，オリジナルの基本周期列．(a) 話者 oo，(b) 話者 ao，(c) 話者 yo，(d) 話者 ka．

Fig. 6 Measurement of the naturalness of the synthesized sounds. From left side, random shuffled surrogate data, 2nd order surrogate data, 3rd order surrogate data, 4th order surrogate data, and the original data. (a) subject oo, (b) subject ao, (c) subject yo, (d) subject ka.

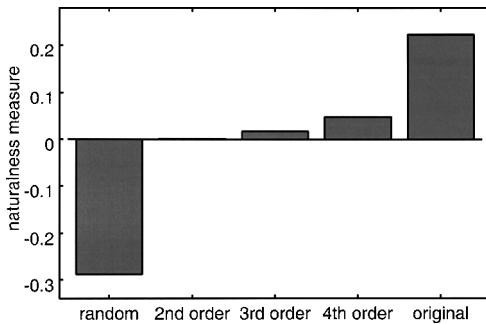


図 7 基本周期列を 2 倍にした場合の話者 ao に対する合成音の自然性の評価

Fig. 7 Measurement of the naturalness of the synthesized sounds from the subject ao in the case when the pitch sequence was repeated twice.

(2) 基本周期の 2 次の相関を保存することで，合成音の自然性は大きく修復される．これは，基本周期

のスペクトル構造が音声の自然性に寄与しているとす  
る，小室ら [3] 及び小林ら [2] の結果を裏づけるもので  
ある．

(3) 基本周期の 3 次以上の高次相関を保存するこ  
とで，合成音の音質は更に改善され，オリジナルの基本  
周期の自然性に近づいてゆく．これは，基本周期揺ら  
ぎが，線形ガウス過程や AR モデルよりも複雑な力学  
的性質をもっており，非ガウス性や非線形な力学構造  
が背後に存在する可能性を示唆している．

これまでに，基本周期のスペクトル構造，すなわち，  
2 次の相関構造に着目した研究はなされてきたが，高  
次相関に関する研究例はない．高次相関解析を行い，  
その重要性を指摘したのが，本研究の新規な結果であ  
る．音声信号の非線形解析が近年行われ [6] ~ [10]，哺  
乳類 [27] にも応用が進んでいるが，線形力学過程の限  
界を指摘した本結果は，基本周期揺らぎに対しても非  
線形力学のアイデアを適用することが可能であること



を示唆している．基本周期の非線形性については今後の更なる検討が必要であるが，研究を進展させることで，非線形の技術が自然な基本周期の合成に貢献できると期待できる．

今後の課題としては，以下の3項目を予定している．

( I ) 今回の実験では，発声の訓練を受け，ピッチの比較的安定したアナウンサーを話者とした．発声の訓練を受けない，一般の話者についても同様の結果が得られるか実験を行う．

( II ) 音声合成法には，波形符号化に基づいた方法を用いたが，携帯電話やインターネット通信で広く利用されているLPC等の分析合成型の方法でも，同様の結果が得られるか実験を行う．

( III ) 基本周期の揺らぎに非線形の力学構造が存在するならば，非線形予測モデルや非線形ARモデルが基本周期の合成には有効であると考えられる．これらの非線形モデルの，基本周期合成装置としての有効性を今後検討してゆく．

謝辞 本研究を進めるにあたり，適切なアドバイスと有意義な議論を頂いた池口徹先生（埼玉大学）と青木直史先生（北海道大学）に感謝致します．また，音声データを御提供頂いた青森放送株式会社，録音に御協力頂いた同社放送部の米澤章子氏及び辻拓哉氏に感謝致します．最後に，聴覚心理実験をお手伝い頂いた山本裕一氏と宮川竜二氏に感謝致します．

## 文 献

- [1] L. Dolansky and P. Tjernlund, "On certain irregularities of voiced-speech waveforms," *IEEE Trans. Audio Electroacoust.*, vol.16, pp.51-56, 1968.
- [2] 小林哲則, 関根英俊, "合成音の自然性に対する基本周期のゆらぎの役割," *音響誌*, vol.47, no.8, pp.539-544, 1991.
- [3] 小室 修, 粕谷英樹, "基本周期のゆらぎの性質とそのモデル化に関する検討," *音響誌*, vol.47, no.12, pp.928-934, 1991.
- [4] 伊福部達, 橋場参生, 松島純一, "母音の自然性における「波形ゆらぎ」の役割," *音響誌*, vol.47, no.12, pp.903-910, 1991.
- [5] N. Aoki and T. Ifukube, "Analysis and perception of spectral  $1/f$  characteristics of amplitude and period fluctuations in normal sustained vowels," *J. Acoust. Soc. Am.*, vol.106, no.1, pp.423-433, 1999.
- [6] H. Herzel, D. Berry, I.R. Titze, and M. Saleh, "Analysis of vocal disorders with method from nonlinear dynamics," *J. Speech Hear. Res.*, vol.37, pp.1008-1019, 1994.
- [7] W. Mende, H. Herzel, and I.R. Titze, "Bifurcations and chaos in newborn cries," *Phys. Lett. A*, vol.145, pp.418-424, 1990.
- [8] S.S. Narayanan and A.A. Alwan, "A nonlinear dynamical systems analysis of fricative consonants," *J. Acoust. Soc. Am.*, vol.97, no.4, pp.2511-2524, 1995.
- [9] A. Behrman, "Global and local dimensions of vocal dynamics," *J. Acoust. Soc. Am.*, vol.105, no.1, pp.432-443, 1999.
- [10] M. Banbrook, S. McLaughlin, and I. Mann, "Speech characterization and synthesis by nonlinear methods," *IEEE Trans. Speech Audio. Process.* vol.7, no.1, pp.1-17, 1999.
- [11] M. Sato, K. Joe, and T. Hirahara, "APOLONN brings us to the real world," *Proc. Int. Joint Conf. Neural Networks*, vol.1, pp.581-587, 1990.
- [12] I. Tokuda, R. Tokunaga, and K. Aihara, "A simple geometrical structure underlying speech signals of the Japanese vowel /a/," *Int. J. Bif. Chaos*, vol.6, no.1, pp.149-160, 1996.
- [13] I. Tokuda, T. Miyano, and K. Aihara, "Surrogate analysis for detecting nonlinear dynamics in normal vowels," *J. Acoust. Soc. Am.*, vol.110, no.6, pp.3207-3217, 2001.
- [14] 徳田 功, 池口 徹, 宮野尚哉, 合原一幸, "サロゲート法に基づく音声知覚心理実験," *信学技報*, NLP99-152, 2000.
- [15] T. Ikeguchi and K. Aihara, "Estimating correlation dimensions of biological time series with a reliable method," *J. Int. Fuzzy Sys.*, vol.5, no.1, pp.33-52, 1997.
- [16] N.B. Pinto and I.R. Titze, "Unification of perturbation measures in speech signals," *J. Acoust. Soc. Am.*, vol.87, pp.1278-1289, 1990.
- [17] R.C. Scherer, V.J. Vail, and C.G. Guo, "Required number of tokens to determine representative voice perturbation values," *J. Speech Hear. Res.*, vol.38, pp.1260-1269, 1995.
- [18] 合原一幸(編), 池口 徹, 山田泰司, 小室元政, *カオス時系列解析の基礎と応用*, 産業図書, 2000.
- [19] 松本 隆, 徳永隆治, 宮野尚哉, 徳田 功, *カオスと時系列*, 培風館, 2002.
- [20] 古井 貞, *デジタル音声処理*, 東海大学出版会, 1985.
- [21] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J.D. Farmer, "Testing for nonlinearity in time series: The method of surrogate data," *Physica D*, vol.58, pp.77-94, 1992.
- [22] J. Theiler and D. Prichard, "Constrained-realization Monte-Carlo method for hypothesis testing," *Physica D*, vol.94, pp.221-235, 1996.
- [23] T. Schreiber, "Constrained randomization of time series data," *Phys. Rev. Lett.*, vol.80, no.10, pp.2105-2108, 1998.
- [24] A. Stuart and J.K. Ord, *Kendall's Advanced Theory of Statistics*, Fifth Ed., Charles Griffin, 1987.
- [25] 山本裕一, 山本耕一郎, 徳田 功, 長島知正, "母音ゆらぎに関する一考察," *信学技報*, CAS2000-41, 2000.
- [26] 大串健吾, 中山 剛, 福田忠彦, *画質と音質の評価技術*,

昭晃堂, 1991.

- [27] I. Tokuda, T. Riede, J. Neubauer, M.J. Owren, and H. Herzel, "Nonlinear analysis of irregular animal vocalizations," J. Acoust. Soc. Am., vol.111, no.6, pp.2908-2919, 2002.

(平成 15 年 6 月 16 日受付, 9 月 21 日再受付,  
11 月 4 日最終原稿受付)



徳田 功 (正員)

平 4 筑波大・自然・物理卒・平 6 同大大学院修士課程了・同年室蘭工大・工・情報・教務職員・平 8 同助手・平 15 同助教授・工博・非線形力学の応用, 時系列解析, 生体信号処理に関する研究に従事・非線形問題研究会・研究専門委員・平 14 アレクサンダー・フォン・フンボルト財団リサーチフェロウシップ授賞・



宮野 尚哉 (正員)

昭 60 京大大学院理学研究科博士課程了・理博・同年住友電気工業(株)勤務・平元住友金属工業(株)勤務・平 11 弘前大学理工学部知能機械システム工学科助教授・平 15 立命館大学 COE 推進機構教授・カオス・複雑系科学, 非線形予測の研究に従事・日本物理学会, 日本神経回路学会, 日本公衆衛生学会, IEEE 各会員・



合原 一幸 (正員)

昭 52 東大・工・電気卒・昭 57 同大大学院工学系研究科電子工学専門課程博士課程了・工博・現在, 東京大学生産技術研究所教授・ニューロコンピューティングの応用研究及びカオスの研究に従事・日本エム・イー ME 学会, 生物物理学会, 電気学会, 電設学会, 神経回路学会, INNS, SMB 各会員・