



HOKKAIDO UNIVERSITY

Title	アルゴリズム選択型適応信号処理スペクトル分析システム
Author(s)	広重, 真人; Hiroshige, Makoto; 宮永, 喜一 他
Citation	北海道大學工学部研究報告, 146, 85-96
Issue Date	1989-05-31
Doc URL	https://hdl.handle.net/2115/42184
Type	departmental bulletin paper
File Information	146_85-96.pdf



アルゴリズム選択型 適応信号処理スペクトル分析システム

広重 真人 宮永 喜一 柄内 香次
(昭和 63 年 12 月 20 日受理)

A Selective System of Multiple Adaptive Algorithms for Speech Spectrum Estimation

Makoto HIROSHIGE, Yoshikazu MIYANAGA and Koji TOCHINAI
(Received December 20, 1988)

Abstract

This paper shows that two different kinds of adaptive algorithms must be applied to analyze voiced and unvoiced speech waveforms. The reason is based on the characteristics of the inputs of speech production model, i. e. a pseudo-periodical pulse train and white Gaussian process. As a suitable algorithm for voiced speech, we introduce a modified MRAS. As the other algorithm, we introduce the MIS for analyzing unvoiced speech. In this paper, a new method to separate voiced and unvoiced speech is first proposed. Then we apply the two adaptive algorithms to analyze them. Especially, it is shown that an accurate estimation parameters must be selected among the results given by using the modified MRAS since this algorithm shows an excellent accuracy only within a specific time interval. The selection can be performed by using the technique of knowledge engineering and a neural network. Using the above methods, we organized a system which automatically selects an appropriate analysis algorithm, and produces an output of a highly accurate spectrum. In some experiments, this paper shows partially that this proposed system is superior to only MIS, modified MRAS, and LPC at the point of spectra estimation.

1. ま え が き

音声のデジタル信号処理では、従来より FFT, LPC^{(1),(2)}, MEM⁽³⁾ 等の一括処理型アルゴリズムが提案されている。これらのアルゴリズムは音声信号を短い時間範囲で定常であると仮定し、音声生成モデルのパラメータを推定しているが、実際の音声はモデルが時間とともに変化する非常過程であると考えられる。従って一括処理型のアルゴリズムでは音声生成モデルの細かい変化に追従することは難しい。そこで近年音声の高精度スペクトル推定法として、各観測データについて 1 つのスペクトルパラメータを求める適応型の信号処理手法が提案されており、主なものに

SEARMA⁽⁴⁾, 拡張 MRAS⁽⁵⁾⁻⁽⁷⁾, MIS^{(8),(9)} 等がある。またこれらのアルゴリズムで用いられる音声生成モデルは、フォルマントのみを推定する AR モデルから、フォルマントとアンチフォルマントの両方を同時推定する ARMA モデルへと移行してきている。

上述のような従来の ARMA 型適応信号処理アルゴリズムでは、生成モデルの入力を白色ガウス過程および疑似周期パルス列としている。これは実際の音声が発声器官の構造に照らしてもこれらの入力混じりであったものと考えられるためである。この二つの入力は性質が異なっているため、入力をどちらか一方に想定したアルゴリズムでは、入力が実際と異なると考えられる部分で分析精度が悪化する。さらに二つの入力を組み合わせた場合の分析手法においても、十分な精度が得られないことが報告されている⁽⁸⁾。

そこで本論文では、音声の生成機構を根拠に、入力が白色ガウス過程で確率信号と考えられる /s/あるいは/t/等の無声音声と、入力が疑似パルス列で確定信号と考えられる有声音声とは、別々のアルゴリズムで分析すべきであることを述べ、経験則を用いた知識工学的手法により最適なアルゴリズムを選択して分析する、スペクトル分析システムを構成する。

2. では非定常な 1 次 AR モデルの係数を用いて経験的に無声音声と有声音声を分離するアルゴリズムを提案し、分離実験結果を示す。

次に 3. において無声音声向き適応信号処理アルゴリズムとして MIS 法^{(8),(9)} を用い、有声音声向き適応信号処理アルゴリズムとして拡張 MRAS 法⁽⁵⁾⁻⁽⁷⁾ を導入する。

4. では、これらのアルゴリズムのうち拡張 MRAS 法において、アルゴリズムの性質上スペクトルデータを取捨選択する必要があることを述べる。従来この様な選択作業は理論的、数式的に行なうことが難しいために、人間の目視による手作業で行なわれてきたが、本論文のシステムでは、ニューラルネットのノードを学習可能な if-then ルールとみなして応用した手法を用いて、学習による自動選択を行なう⁽¹⁰⁾。

5. では、以上の手法を 1 つの適応信号処理スペクトル分析システムとして構成し、そのシステムを用いた分析実験結果を示す。分析結果より、本論文で構成したシステムの有効性が示される。

2. 無声音声と有声音声の分離

2.1 音声の生成モデル

人間の音声には大きく分けて有声音声と無声音声がある。有声音声は声帯の振動を伴う音声である。声帯の振動とは即ち声帯の開閉であるから、この開閉によって呼気流はほぼ周期的に断続させられる。有声音声ではこの断続的な呼気流が音源となるため、音声の生成モデルにおいて図 1(a)のように入力を疑似周期パルス列とするのは合理的である。またこの様に音声生成フィルタをパルス列で励振した場合、フィルタの出力、即ち観測波形は、パルスが入力した時刻ではパルス成分が含まれ、他の時刻ではそのパルスによるフィルタの応答波形が観測される。よって有声音声では各観測信号の値に意味があり、近似的には確定信号と見なすことができる。

一方/s/, /t/等の無声音声では、声帯の振動は伴わず、調音点における乱気流が音源となる。乱気流は可聴周波数帯域全域にわたった連続なスペクトルを持つ。ここでこの雑音音源を、平坦な周波数特性をもつ白色過程で、かつ平均と分散の情報だけで性質の記述が可能なガウス過程であると仮定する。無声音声の生成モデルでは、この様に白色ガウス過程を音声生成フィルタの入力とする(図 1(b))。入力がガウス過程であるから出力信号もガウス過程となり、観測波形は 2 次のモーメントに特徴を持つ確率信号となる。

以上に述べたことから、有声音声と無声音声は、音声生成フィルタを励振する音源の性質が本質的に異なり、そのため観測される信号の性質も異なっている。従って音声を高精度に分析する場合、有声音声と無声音声では各々に適したアルゴリズムを使い分ける必要があると考えられる。実際、入力を白色ガウス過程もしくは疑似周期パルス列のどちらか一方に限定したアルゴリズムでは高い分析精度は得られず、また白色ガウス過程入力と疑似パルス列入力の両方に対応できるように工夫されたアルゴリズムを用いても、分析精度には限界があることが示されている⁽⁸⁾。

有声音声と無声音声で分析アルゴリズムを使い分けようとする場合、まず有声音声と無声音声を時間軸上で分離する必要がある。以下に有声音声と無声音声を分離する一手法を提案し、その手法による分離実験の結果を述べる。

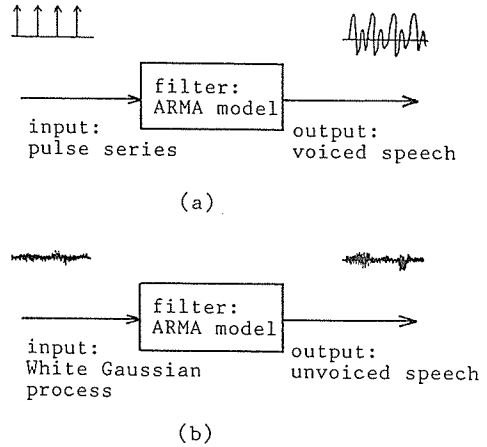


図1 音声の生成モデル
(a) 有声, (b) 無声

2.2 分離アルゴリズム

本節では有声音声と無声音声を分離するアルゴリズムについて述べる。最終的にはこのアルゴリズムで得られた分離点を用いて、有声音声・無声音声を異なったアルゴリズムで分析する。従ってこのアルゴリズムを実際の信号処理を行なう前段階において情報を得るという意味でパイロットアルゴリズム (PA) と名付ける。

2.2.1 分離に用いるパラメータ

有声音声と無声音声を分離するためには、有声音声と無声音声の違いを明確に表わすパラメータを見つけなければならない。本手法では、有声音声と無声音声のスペクトル包絡の大まかな傾きの違いに注目し、音声スペクトルの傾きを表わす1次の時変ARモデルを用いる。

このモデルのAR係数は、既に提案されている確率信号向き非定常ARMA分析アルゴリズムMIS法^{(8),(9)}を書き換えることにより容易に求められる。

MIS法の詳細は文献(9), (11)に譲り、ここでは生成フィルタの入力を白色ガウス過程とし、1次ARモデルを用いた場合のアルゴリズムを示す。

(定数)

$$\alpha = 50 \text{ サンプル}, C = 16, t_0 = 20 \text{ サンプル},$$

(初期値)

$$\hat{a}_1(0) = 0, f(0| -1) = 100, \\ q^2(0) = 0, \hat{\sigma}^2(0) = 0.1, t_i = 0$$

(回帰式)

$$v(k) = y(k) + y(k-1) \cdot \hat{a}_1(k-1) \tag{1}$$

$$f(k-1) = f(k-1|k-2) - \frac{y^2(k-2)f^2(k-1|k-2)}{1+y^2(k-2)f(k-1|k-2)} \tag{2}$$

$$\text{if } v^2(k) \geq C \cdot \hat{\sigma}^2(k-1) \cdot (1+y^2(k-1) \cdot f(k-1)) \tag{3}$$

$$\text{and } k \geq t_i + t_0 \text{ then} \tag{4}$$

$$\hat{q}^2(k-1) = \{v^2(k) - \hat{\sigma}^2(k-1) \cdot (1+y^2(k-1) \cdot f(k-1))\} / y^2(k-1) \tag{5}$$

$$v(k) = 0 \quad (6)$$

$$t_i = k \quad (7)$$

else

$$\hat{q}^2(k-1) = 0 \quad (8)$$

endif

$$f(k|k-1) = f(k-1) + \frac{\hat{q}^2(k-1)}{\hat{\sigma}^2(k-1)} \quad (9)$$

$$\hat{a}_1(k) = \hat{a}_1(k-1) - \frac{y(k-1)f(k|k-1)v(k)}{1+y^2(k-1)f(k|k-1)} \quad (10)$$

$$\alpha(k) = \min(k, \alpha) \quad (11)$$

$$\hat{\sigma}^2(k) = \hat{\sigma}^2(k-1) + \{v^2(k)/(1+y^2(k-1)f(k|k-1)) - \hat{\sigma}^2(k-1)\}/\alpha(k) \quad (12)$$

上式における \hat{a}_1 は次の評価関数を最小にするように求められる⁽⁹⁾。

$$V = \frac{\sigma_u^2}{2\pi j} \int |H(z^{-1})(1+\hat{a}_1 z^{-1})|^2 \frac{dz}{z} \quad (13)$$

ここで $H(z^{-1})$ は音声生成モデルであり、 σ_u^2 はモデルの入力分散である。上式より音声信号の場合、音声の声帯および放射モデルを考慮すれば、 \hat{a}_1 は 0 から -1 までの値をとり、無声音声では 0 に、有声音声では -1 に近いことが予想される。このアルゴリズムは適応型であるので、サンプルデータ 1 個につき 1 個の \hat{a}_1 が求められる。

2.2.2 \hat{a}_1 パラメータ-時間曲線

前項のアルゴリズムにテストデータをかけた例が図 2(b)である。縦軸が \hat{a}_1 、横軸が時間 t である。使用したデータは成人男子の実音声/ $t \int i$ /で、12 kHz サンプリング、12 bit A-D である。図 2(a)に示される原波形と比較すると、分離すべき無声音声/ $t \int /$ と有声音声/ i /の境目では、パラメータ \hat{a}_1 の値が急激に増加・減少しており、明らかに他の点とは異なった特徴を示している。この特徴ある増加・減少点を捕らえることにより無声音声と有声音声の分離が可能となる。

2.2.3 分離点の決定

前項で求められた \hat{a}_1-t 曲線はほぼ予想通り無声音部分では 0 に近く、有声音部分では -1 に近い値をとり、無声音から有声音への変化点で急激に増加・減少している。無声音と有声音を分離するにはこの急激な増加・減少点を捕らえればよい。このことは即ち無声音と有声音を分離する問題がグラフ上の特定のパターンを抽出する問題に変換されたことになる。この様な問題は純粋に理論的、数式的に解決することは難しく、経験的な法則をルールの形式で蓄えて解決する知識工学的な手法をとるのが適当と考えられる。

捕らえるべき分離点の \hat{a}_1-t 曲線の特徴は、まず① \hat{a}_1 の値が急激に増加(減少)していることと、② \hat{a}_1 は -1 (0) 付近から 0 (-1) 付近へ値が変化していることである。これらの特徴をまず大まかなルールの形で表現し、その後実験により経験的なルールを追加するという方法をとった。

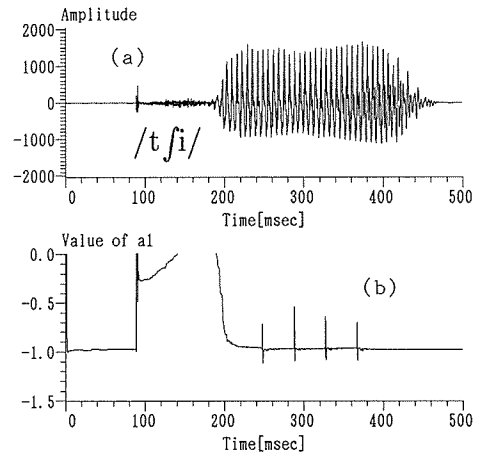


図 2 \hat{a}_1-t 曲線

[ルール1] 捕らえるべき分離点は \hat{a}_i の値の大きな増加（減少）がある程度続いている。

以下の手順によりこのルールを適用する。

- i) \hat{a}_i の差分 $\Delta\hat{a}_i$ をとる。
 ii) $\Delta\hat{a}_i$ を以下のように3段階に分類して記号化する。

$$a : 0.005 < \Delta\hat{a}_i$$

$$c : -0.005 < \Delta\hat{a}_i < 0.005$$

$$e : \Delta\hat{a}_i < -0.005$$

- iii) 「a」（「e」）が2個連続した点を候補点とする。

正しい分離点を落とさないためには連続点数は2点となった。

[ルール2] 候補点がいくつも密集している場合には1つにまとめる。

15 msec 以内に複数の候補点がある場合は1つにまとめた。

[ルール3] 捕らえるべき分離点は、増加（減少）開始点と増加（減少）終了点の \hat{a}_i の値の差が大きい。

増加（減少）開始点の前20点の平均値と、増加（減少）終了点の後20点の平均値の差が0.2(0.1)以上のものを候補点として残した。

[ルール4] 減少の場合、捕らえるべき分離点は、減少終了点の \hat{a}_i の値が-1の付近である。

上述の終了点付近の平均値が-0.8以下ならば「-1付近」であるとした。なおこのルールは増加点には適用しない。

[ルール5] \hat{a}_i の値が0（-1）にきわめて近い値を安定してとれば、明らかに無声音（有声音）である。

「変化」のみに注目しているルール1～4の明らかな誤りを除くルールであり、 \hat{a}_i の3 msecの平均が-0.1以上（-0.9以下）ならば無声音（有声音）とした。

上述の5つのルールにおける具体的な数値は、後述するテストデータから目視で経験的に決定した。

2.3 分離実験結果

前節で述べたアルゴリズム及びルールで、ルール抽出に用いたものと同様のデータを分析し、分離実験結果を得た。音声データは比較的分離が難しいとされる破裂性の無声子音を含む単音節音声であり、3名の成人男性話者による {/p/, /t/, /k/} + {/a/, /i/, /u/, /e/, /o/}, /t/ + {/a/, /e/, /o/}, /t/ + {/i/, /tsu/} (12 kHz サンプリング, 12 bit A-D) の計45個である。図3～5に結果の例を示す。横軸はすべて時間 t で、(a)は原波形、(b)は \hat{a}_i 、(c)はシステムが出力した分離実験結果である。45サンプル中19サンプル(42%)でほぼ妥当な点を捕らえた(図3)。25サンプル(56%)

については、破裂子音を対象としているため子音部がガウス性というよりはインパルス応答的な信号であり、そのため \hat{a}_1 の値にあまり変化がみられず、分離点は出力されなかった(図4)。しかしこのような子音は、MIS法ではなく拡張MRAS法で分析されるべきであり、この意味でシステムは正しい判断をしたと考える。残りの1サンプル(2%)については、増加点に対応する減少点が捕らえられず、ルール5によっても正しく修正されなかったため、誤った分離結果を得た(図5)。

以上の実験より、ルール抽出と同じデータではあるが、45サンプル中44サンプル(98%)でシステムはほぼ妥当な分離を行なうことが確かめられ、前節で述べたアルゴリズム及びルールを用いることによって、無声音声と有声音声を良好に分離できることが示された。本実験において無声音声と有声音声の分離問題における、経験則を用いた知識工学的手法の有効性を示した。

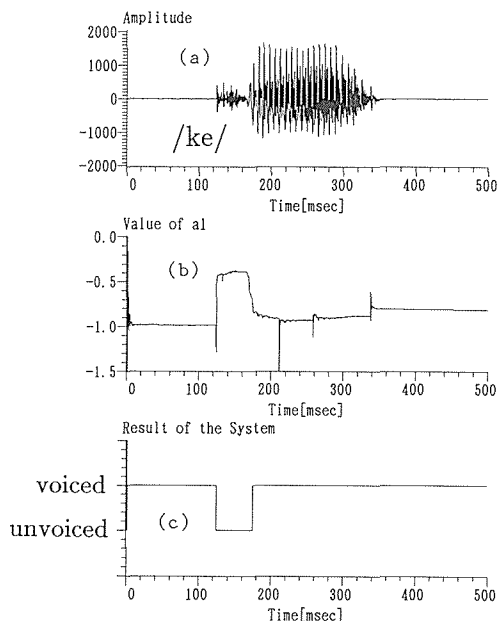


図3 分離実験結果

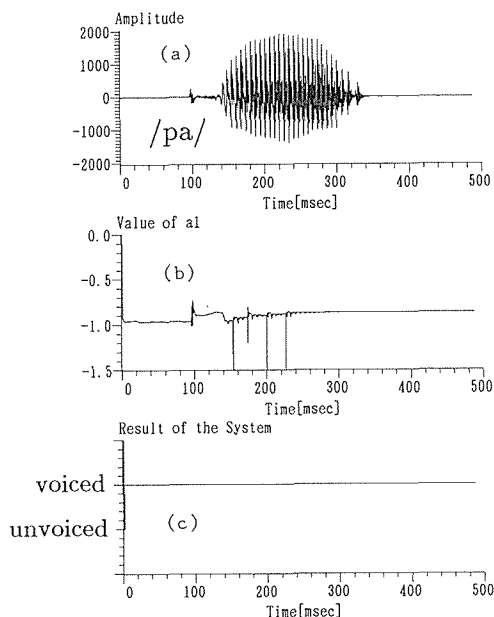


図4 分離実験結果

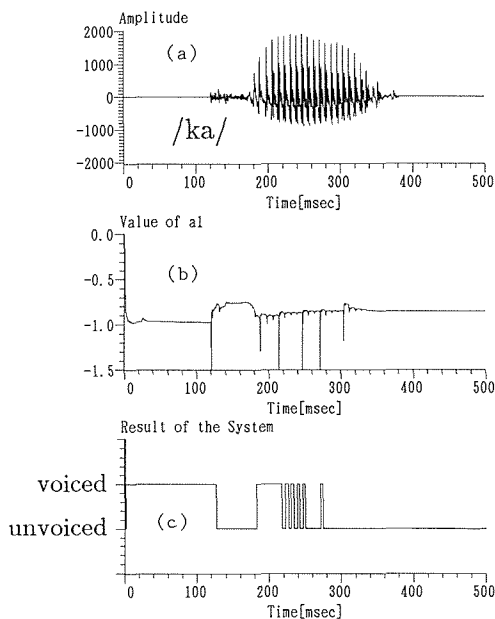


図5 分離実験結果

3. 二つのスペクトル推定手法

本章では、本論文で使用する二つの高精度非定常スペクトル推定手法について述べる。これらの二つのアルゴリズムは、共に音声の生成モデルを ARMA モデルとしているが、入力を白色ガウス過程とするか疑似周期パルス列とするかに違いがある。最初に規範モデルを、

$$y(k) = -\sum_{i=1}^n a_i y(k-i) + \sum_{j=1}^m b_j u(k-j) + u(k) \quad (14)$$

とする。但し、 $y(k)$ は出力であり、 $u(k)$ は入力で、MIS 法の場合白色ガウス過程、拡張 MRAS 法の場合疑似パルス列とする。

推定モデル（時刻 τ の推定値）を、

$$\hat{y}(\tau|k) = -\sum_{i=1}^n \hat{a}_i(k) y(\tau-i) + \sum_{j=1}^m \hat{b}_j(k) \hat{u}(\tau-j) \quad (15)$$

とする。

3.1 モデル同定システム (MIS 法)

MIS 法のアルゴリズムは既に文献(9), (11)に示されているため、その特徴のみを以下に述べる。

- (1) 係数推定時における入力は白色ガウス過程である。
- (2) 時変 ARMA パラメータの推定が可能である。
- (3) パラメータの時変性は推定誤差の変化に従っており、誤差があるしきい値を越えた時パラメータが変化すると検出し、その変化量は最尤推定法により求めている。
- (4) 信号の 2 次モーメントからモデルの係数を推定するため、確率信号に適している。

3.2 拡張 MRAS 法

拡張 MRAS 法のアルゴリズムについても既に文献(5), (11)において示されているため、以下にその特徴のみを述べる。

- (1) 入力はパルス列である。
- (2) 時変 ARMA パラメータの推定が可能である。
- (3) パルスの検出は予測誤差があるしきい値を越えた時に行なわれる。そのときのパルス値は同一時刻の予測誤差とする。
- (4) 現在の観測信号に重みをかけ、過去の観測信号の情報を徐々に忘却させる重み係数を導入している。
- (5) 拡張 MRAS 法は、重み係数の働きにより時間的に狭い範囲で観測された信号を用いた推定を行なうため、1 周期内のデータからスペクトル推定を行なわなければならない確定信号に適している。

4. 拡張 MRAS 法におけるデータ選択

4.1 データ選択の必要性

確定信号用アルゴリズム「拡張 MRAS 法」では、3.2 に述べた特徴(3)においてパルスを求めている。その後次のパルスがやってくるまでは、出力信号はそのパルスによって引き起こされた自由振動であるとする。音声生成モデルの周波数特性はこの自由振動の部分に含まれている。またパルスの直後にはパルスは入力されないという仮定の下で、ある時間（20 サンプル点）は予測誤差に関わらずモデルの入力を零としている。

このようなアルゴリズムにおいては本質的に、パルスが入力したと推定される時刻では推定され

たスペクトル包絡に大きな誤差があることになる。また実有声音を用いた実験によると、パルスが入力したと推定される時刻以降もしばらくの間予測誤差が大きくかつ不安定な状態が続き、その後誤差が小さくなり安定していく。

従って、拡張 MRAS 法を用いてスペクトル分析を行なう場合は、予測誤差が小さく、かつ安定している部分のスペクトル包絡を選び出さなければならない。このような部分のスペクトル包絡は、自由振動をしていると見なされる時刻で求められたことになるので、音声生成フィルタの特性を精度よく表わしていると考えられる。

4.2 データ選択手法

以上のようなデータ選択の問題は、2. の有声無声音の分離問題と同様に、純粋に理論的、数式的に解決することは難しい。従来このような選択はしきい値操作を行なう if-then ルールによって行なわれていたが、しきい値の客観的獲得に問題が残されていた⁽¹¹⁾。そこで本論文では、しきい値操作を行なう if-then ルールが、理論的にはニューラルネットワーク^{(12),(13)} で用いられている 1 個のノードに対応することに注目し、バックプロパゲーションアルゴリズム⁽¹³⁾ を用いてノードの学習を行なうことで、しきい値をより客観的に獲得し自動選択を行なう⁽¹⁰⁾。

4.3 データ選択実験

前節の手法を用いて、実音声におけるデータ選択実験を行なった。

使用した音声サンプルは成人男性、12 kHz サンプリグ、12 bit-AD の日本語母音である。

ノードへの入力パラメータは物理的に意味があると予想されるものを経験的に決定する。本論文では、拡張 MRAS 法で得られる誤差 $v(k)$ が大きく変動している箇所ので分析スペクトルも悪化していることから、ノードへの入力パラメータとして誤差 $v(k)$ の大きさを表わす $v^2(k)$ 、 $v(k)$ の分散の大きさを表わす $v^2(k)$ の短時間平均値 $\overline{v_s^2(k)}$ 、すなわち

$$\overline{v_s^2(k)} = \frac{1}{10} \sum_{i=k-9}^k v^2(i) \quad (16)$$

さらに信号のおおまかなレベルを表わす $v^2(k)$ の長時間平均値 $\overline{v_l^2(k)}$ 、すなわち

$$\overline{v_l^2(k)} = \overline{v_l^2(k-1)} + \frac{v^2(k) - \overline{v_l^2(k-1)}}{50} \quad (17)$$

を用いた。

学習データは音声サンプルの 1 つを拡張 MRAS 法で分析し、誤差 $v(k)$ の時間変化と求められたスペクトルの時間変化を表示して、それらを参考に目視でスペクトル選択を行なって作成した。ノードの学習回数は 2 万回である。

結果の一例を図 6 に示す。実験結果より、ほぼ良好な選択が行なわれることが示された。

5. アルゴリズム選択型適応信号処理スペクトル分析システム

本章では、前章までの有声・無声音分離システム、MIS 法および拡張 MRAS 法、拡張 MRAS 法におけるスペクトルデータ選択手法を合体させ、1 つのスペクトル分析システムを構成する。このシステムは、入力された信号の性質に適合する分析アルゴリズムを自動的に選択し、高精度な分析スペクトルを出力するものである。

5.1 システム構成

本システムのブロック図を図 7 に示す。入力された音声波形は、まず MIS 法および拡張 MRAS 法の両方でスペクトル分析される。拡張 MRAS 法の出力スペクトルは 4. で述べた選択手法によって選択された後、スペクトル平滑化のため選ばれたスペクトルのみの間で個数 15 個で平均操

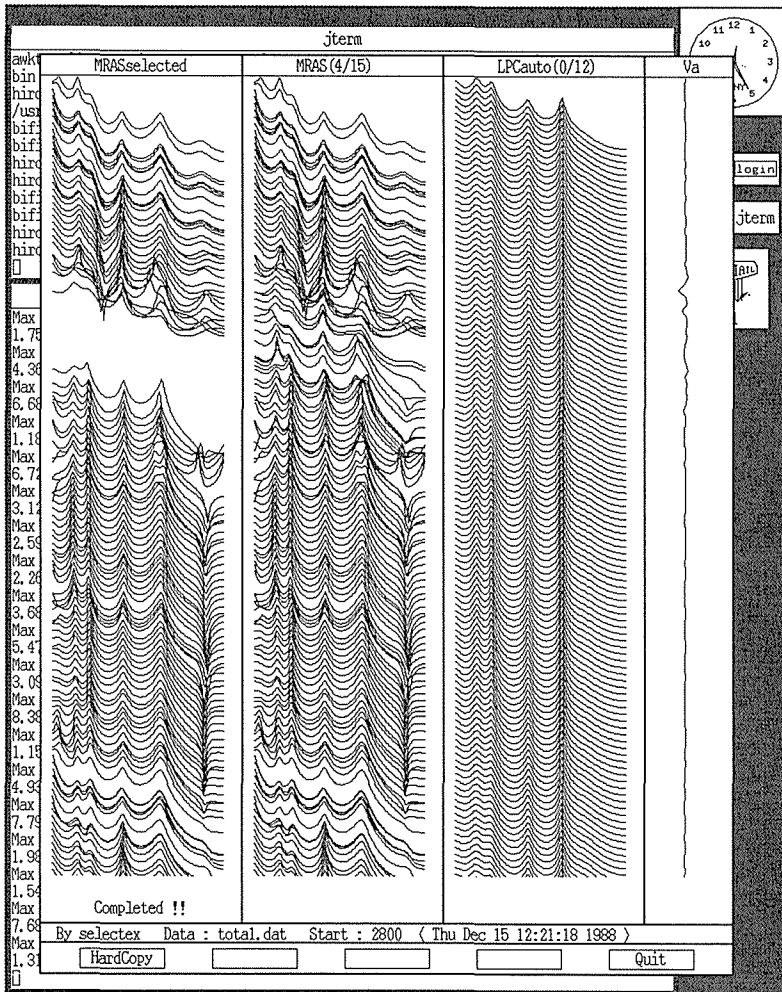


図6 拡張MRAS法スペクトルの選択実験 (/a/)

作が施される。一方入力波形は2. のPAにも同時に入力され、その出力 $\hat{a}_1(k)$ がマネジメントシステムに入力される。マネジメントシステムでは蓄えられた if-then ルールを参照し、 $\hat{a}_1(k)$ の値の変化から有声・無声の判断を行なう。判断結果はスペクトル分析アルゴリズムの下のスペクトル選択スイッチに伝えられ、高精度な分析スペクトルが選択されて本システムの最終出力となる。

スペクトル分析に適応信号処理アルゴリズムを用いているので、その利点を生かすためマネジメントシステムも適応的な動作をするよう考慮した。すなわち、基本的にはデータが1点入力される度に計算・判断をし、2.2.3の [ルール2] のように判断にある時間範囲を必要とする場合には、十分なデータがそろそろまで判断を保留する。また2つの分析アルゴリズム両方で分析した後、スペクトルを選択することによって、2つの分析アルゴリズムの次数を一致させる必要がなくなり、有声・無声音声それぞれに適した次数で分析が可能になる。

5.2 スペクトル分析実験

本システムに2.3で用いたものと同じ実音声データを入力し、分析結果を得た。結果の一例を

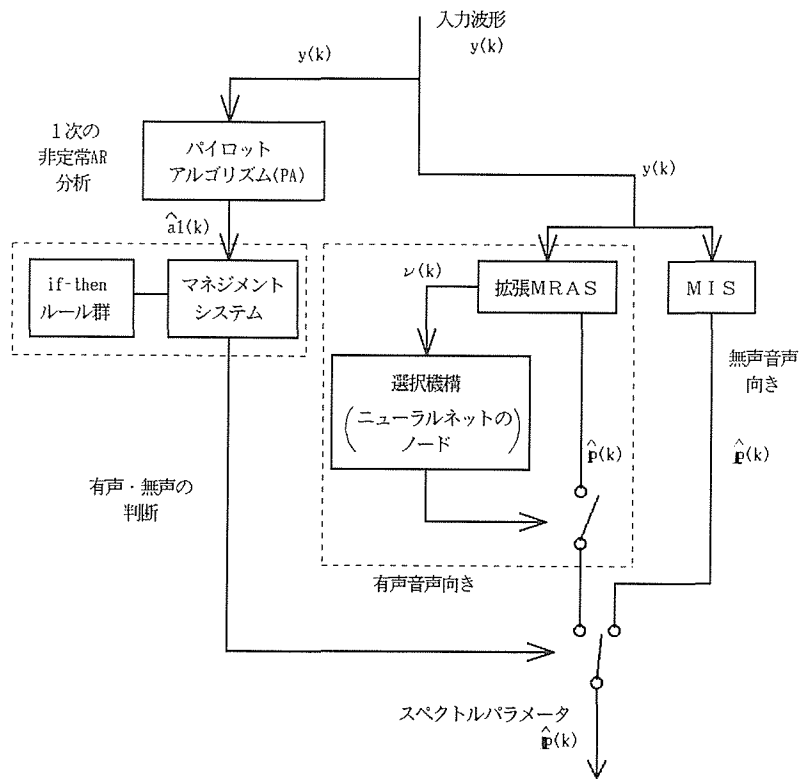


図7 本論文で提案するシステムのブロック図

図8に示す。

図8において、分析音声は/ta/であり、4枚のグラフは左からそれぞれ、MIS法(AR:15次, MA:6次)、比較のためのLPC(自己相関法, 12次)、拡張MRAS法(AR:15次, MA:4次, $\lambda=0.9$)、信号波形である。波形のグラフの横破線は、システムが出力した分析アルゴリズム切り替え点であり、MISとMRASのグラフの右側の四角はそのスペクトルがシステムによって選択されたことを示す。

図8より、本システムの最終出力スペクトルは、MISおよびMRAS単独で分析したスペクトルより優れており、LPCと比較しても時間分解能、高次のフォルマント、フォルマント帯域幅等で優れていることが示された。

本章の実験により、ガウス性の無声音声と周期パルス性の有声音声とを分離して、それぞれに適したアルゴリズムで分析することの有効性が確認された。

6. む す び

本論文ではまず、音声のスペクトル分析において、確率信号である無声音声と、確定信号とみなせる有声音声では、適する分析アルゴリズムが異なるという立場に立ち、無声音声と有声音声の分離を試みた。無声音声と有声音声の違いは1次の非定常ARモデリングにおける係数の値の変化にあらわれることを示し、経験則をルールの形で蓄積して問題を解決する知識工学的な手法が、無声音声と有声音声の分離に有効であることを示した。

次に無声音声に適したMIS法、有声音声に適した拡張MRAS法の二つの適応信号処理アルゴ

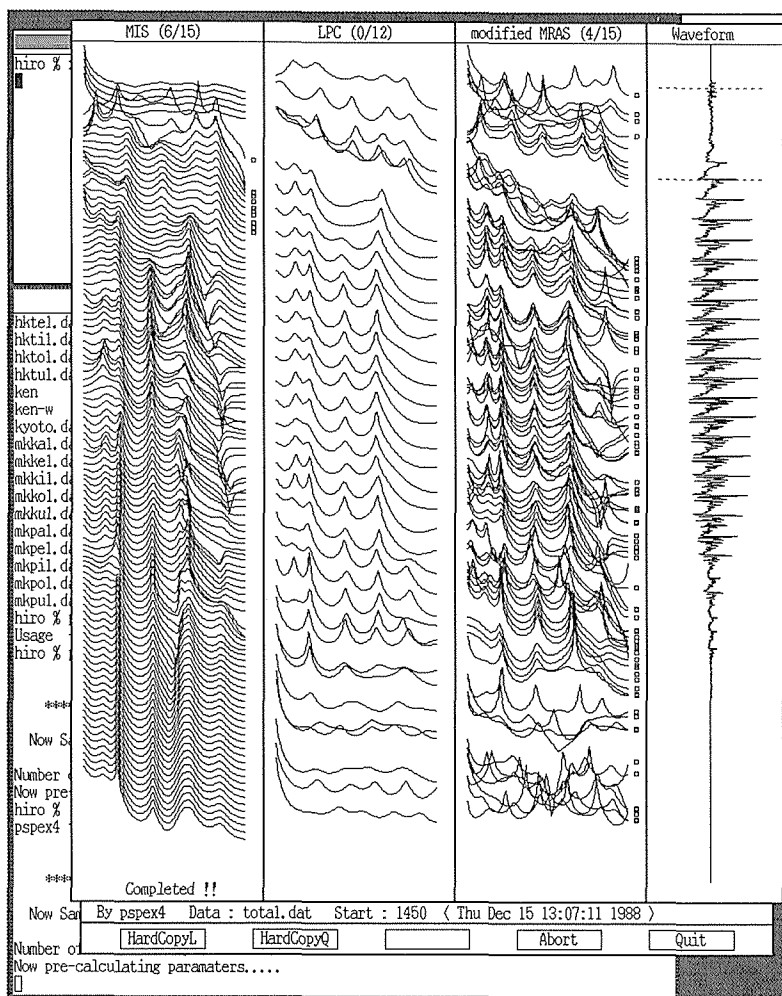


図 8 本システムでのスペクトル分析実験 (/ta/)

リズムを示した。

拡張 MRAS 法において、得られたスペクトルデータから近似精度のよい結果を選択する必要があることを述べ、ニューラルネットのノード学習可能な if-then ルールとみなして応用し、学習によって適切なデータ選択が可能であることを示した。

最後に以上の内容をふまえて、入力信号の性質に適合したアルゴリズムを自動的に選択しながら分析を行なう、適応信号処理スペクトル分析システムを構成した。本システムで得られたスペクトルは、それぞれのアルゴリズム単独の分析結果より優れており、無声音と有声音の分離の必要性を裏付けた。

今後の課題としては、各パラメータがより複雑な動きをする連続音声への対応、他の信号処理アルゴリズム中に数多くみられる経験則的部分に知識工学的手法を導入すること等が考えられる。

参考文献

- (1) 板倉, 齋藤: “統計的手法による音声スペクトル密度とホルマント周波数の推定”, 信学論 (A), 53-A, 1, pp. 35-42 (昭45-01).
- (2) J. D. Markel and A. H Gray Jr.: “Linear prediction of speech”, Springer-Verlag (1976).
- (3) S. Haykin (Ed): “Nonlinear Methods of Spectral Analysis”, Springer-Verlag (1979).
- (4) 森川, 藤崎: “SEARMA 法による音声分析における観測区間の適応的制御”, 音響学会誌, 39, 8, pp. 512-520 (1983).
- (5) 宮永, 三木, 永井: “ピッチ推定を含めた音声の ARMA パラメータの一推定法”, 信学論 (A), 63-A, 11, pp. 737-744 (昭55-11).
- (6) 宮永, 三木, 永井, 羽鳥: “時変 ARMA パラメータの適応的同時推定”, 信学論 (D), E 64-D, 4, pp. 308-315 (昭56-04).
- (7) Y. Miyanaga, N. Miki, N. Nagai and K. Hatori: “A speech analysis algorithm which eliminates the influence of pitch using model reference adaptive system”, IEEE Trans. ASSP, ASSP-30, 1, pp. 88-96 (Feb. 1982).
- (8) 宮永, 三木, 永井: “音声の ARMA 分析に関する一考察”, 音響学会誌, 40, 6, pp. 397-406 (1984).
- (9) Y. Miyanaga, N. Miki and N. Nagai: “Adaptive identification of time-varying ARMA speech model”, IEEE Trans. ASSP, ASSP-34, 3, pp. 423-433 (Jun. 1986).
- (10) 広重, 宮永, 柄内: “神経回路モデルを用いた高精度スペクトル選択法”, 信学技報, PRU 88-52, pp. 31-37 (1988-09).
- (11) 広重, 宮永, 柄内: “知識工学的手法を導入した適応信号処理”, 第2回デジタル信号処理シンポジウム, pp. 337-342 (1987-12).
- (12) R. P. Lippmann: “An Introduction to Computing with Neural Nets”, IEEE ASSP magazine, pp. 4-22 (Apr. 1987).
- (13) D. E. Rumelhart, J. L. McClelland: “Parallel Distributed Processing: Explorations in the Microstructure of Cognition Vol. 1”, MIT Press (1986).