

一般化ケプストラムをパラメータとする音声のスペクトル推定

正員 徳田 恵一[†] 正員 小林 隆夫[†]
 非会員 山本竜太郎[†] 正員 今井 聖[†]

Spectral Estimation of Speech Based on Generalized Cepstral Representation

Keiichi TOKUDA[†], Takao KOBAYASHI[†], Members, Ryutaro YAMAMOTO[†],
 Nonmember and Satoshi IMAI[†], Member

あらまし 一般化ケプストラムで表現されたスペクトルモデルに対して、対数スペクトルの不偏推定におけるスペクトル評価を適用したスペクトル推定法を提案する。また、このときの最小化問題の解法を与えると共に、解の収束性、得られる伝達特性の安定性を示す。本方法では、べきパラメータ γ の値に従い、スペクトルのモデルを全極形からケプストラムによって表現された形まで、連続的に変化させることができる。従って、適当な γ の値によって、音声スペクトルの性質に適した形でスペクトルを表現することができ、音声の分析合成に適用した場合より良い品質の合成音を得られる。本論文では、提案するスペクトル推定法の性質についていくつかの観点から考察し、更に、合成信号、自然音声の分析例により本方法の有効性を示している。本方法は、線形予測法、対数スペクトルの不偏推定を含む手法であり、線形予測法とケプストラム法を統一的に扱う手法と考えることもできる。

1. ま え が き

線形予測法⁽¹⁾は、全極モデルによるスペクトル推定法として、音声分析に広く用いられている。しかし、一般に、音声信号には極のみでなく零も含まれており、全極モデルでは、特に鼻音などの零をもつ信号のスペクトルをうまく表現することができない。また、ピッチ周波数の高い信号では、スペクトルの共振点が高調波成分に適合するという問題が起こることもある。極零モデルにより音声を分析する方法も数多く提案されているが、伝達特性の安定性、解の収束性などが問題となることが多い。

一方、ケフレンシ領域では、極と零との違いはケプストラム係数の符号の違いとしてのみ現れるため、ケプストラムをパラメータとすることにより、極と零とを区別なく取り扱うことができる。ケプストラム法⁽²⁾によって得られるスペクトルは、ケプストラムをパラメータとしていることから、極と零とを同じ正確さで表現しており、音声スペクトルの表現に適していると考え

られる。しかし、反面、このようなケプストラムをパラメータとするスペクトルモデルは、低次数ではホルマントの鋭いピークを実現しにくくなるという問題をもつ。

ところで、一般化ケプストラム⁽³⁾は、ケプストラムを拡張したもので、べきパラメータ γ の値によって、 $\gamma=0$ でケプストラム、 $\gamma=-1$ でAR係数に対応し、 γ の値によって性質を連続的に変化させることができる。従って、一般化ケプストラムによりスペクトルを表現することにより、スペクトルのモデルを、全極モデルから、極零を同等に表現するケプストラムモデルまで、連続的に変化させることができ、 γ を適当な値に選ぶことにより全極形とケプストラムによって表現されたスペクトルとの長所を併せもったスペクトルモデルが得られると考えられる。

しかし、従来的一般化ケプストラムを用いたスペクトル分析法⁽⁴⁾により得られるスペクトルは、一般化対数目盛り上での線形平滑化のため、スペクトルの推定値としては、偏りをもったものとなる。この問題を避けるために、本論文では、対数スペクトルの不偏推定⁽⁵⁾におけるスペクトル評価関数を最小化する形でスペクトルの推定値を求めている。この評価関数は、最ゆ

[†] 東京工業大学精密工学研究所, 横浜市
 Research Laboratory of Precision Machinery and Electronics,
 Tokyo Institute of Technology, Yokohama-shi, 227 Japan

う推定法における評価関数と形式的に同じ形になっていることから、本方法は、 $\gamma = -1$ で線形予測法と等価となる。また、対数スペクトルの不偏推定では、スペクトルをケプストラムで表現していることから、本方法は $\gamma = 0$ で対数スペクトルの不偏推定と等価となる。

$\gamma = -1$ 以外では、本方法は非線形最小化問題となるが、効率的な繰返し計算により解くことができる。すなわち、各繰返しは、一般化ケプストラムの次数を M として $O(M^2)$ の演算量の高速アルゴリズムによって線形方程式を解くことに帰着される。また、解の収束性、得られる伝達関数の安定性が保証されることが示される。

いくつかの γ について合成信号および自然音声の分析例を示し、本方法の有効性を示す。

2. スペクトルのモデルと評価関数

2.1 一般化ケプストラム⁽³⁾によるスペクトルモデル

一般化対数関数 $s_\gamma(w)$ は、

$$s_\gamma(w) = \begin{cases} (w^\gamma - 1)/\gamma, & 0 < |\gamma| \leq 1 \\ \log w, & \gamma = 0 \end{cases} \quad (1)$$

と定義される。更に、一般化ケプストラム $c_\gamma(m)$ は、一般化対数スペクトル $s_\gamma(X(e^{j\omega}))$ のフーリエ係数で定義され、次の関係をもつ。

$$c_\gamma(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} s_\gamma(X(e^{j\omega})) e^{j\omega m} d\omega \quad (2)$$

$$s_\gamma(X(e^{j\omega})) = \sum_{m=-\infty}^{\infty} c_\gamma(m) e^{-j\omega m} \quad (3)$$

本論文では、音声のスペクトル $H(e^{j\omega})$ を、 M 次までの一般化ケプストラム $c_\gamma(m)$ によって、

$$s_\gamma(H(z)) = \sum_{m=0}^M c_\gamma(m) z^{-m} \quad (4)$$

とモデル化する。式(1)より、 $s_\gamma(w)$ の逆関数は、

$$s_\gamma^{-1}(w) = \begin{cases} (1 + \gamma w)^{1/\gamma}, & 0 < |\gamma| \leq 1 \\ \exp w, & \gamma = 0 \end{cases} \quad (5)$$

で与えられるので、 $H(z)$ は、

$$H(z) = s_\gamma^{-1} \left(\sum_{m=0}^M c_\gamma(m) z^{-m} \right) \quad (6)$$

と書くことができる。ここで、ゲイン項をくり出すと、

$$H(z) = K \cdot D(z) \\ = K \cdot s_\gamma^{-1} \left(\sum_{m=1}^M c'_\gamma(m) z^{-m} \right)$$

$$= \begin{cases} K \cdot \left(1 + \gamma \sum_{m=1}^M c'_\gamma(m) z^{-m} \right)^{1/\gamma}, & 0 < |\gamma| \leq 1 \\ K \cdot \exp \sum_{m=1}^M c'_\gamma(m) z^{-m}, & \gamma = 0 \end{cases} \quad (7)$$

但し、

$$K = s_\gamma^{-1}(c_\gamma(0)) \quad (8)$$

$$D(z) = s_\gamma^{-1} \left(\sum_{m=1}^M c'_\gamma(m) z^{-m} \right) \quad (9)$$

また、

$$c'_\gamma(m) = \begin{cases} 0, & m = 0 \\ c_\gamma(m)/(1 + \gamma c_\gamma(0)), & m \geq 1 \end{cases} \quad (10)$$

と書き換えることができる。 $c'_\gamma(m)$ は規格化一般化ケプストラム⁽³⁾と呼ばれるものであるが、以下ではこれを簡単のために一般化ケプストラムと呼び、 $c_\gamma(m)$ と表記することにする。

式(7)のスペクトルモデルは、 $\gamma = -1$ で全極形に、また $\gamma = 0$ でケプストラムで表現された形になる。従って、 γ を -1 から 0 の間の適当な値に選べば、ケプストラムで表されたスペクトルモデルと全極モデルの特長を併せもったスペクトルモデルを得ることができると思われる。

2.2 スペクトル評価関数

一般化対数領域でのスペクトル平滑化法⁽⁴⁾によれば、一般化ケプストラムによって表現されたスペクトルを得ることができる。しかし、得られたスペクトルは、スペクトル推定の観点からは偏りをもったものとなり、また、得られる伝達特性は γ が -1 に近いときには必ずしも安定なものとはならない。

対数スペクトルの不偏推定⁽⁵⁾では、相対値の意味で不偏なパワースペクトルの推定値 $|H(e^{j\omega})|^2$ が、評価関数

$$E = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\exp R(\omega) - R(\omega) - 1) d\omega \quad (11)$$

を最小化する形で得られることを示している。但し、

$$R(\omega) = \log |X(e^{j\omega})|^2 - \log |H(e^{j\omega})|^2 \quad (12)$$

であり、 $|X(e^{j\omega})|^2$ は弱定常過程 $x(n)$ の変形ピリオドグラムとする。

式(11)は特定のスペクトルモデルを仮定せずに導かれたものであるので、本論文では、前節で述べたモデルによりスペクトルを表し、式(11)の評価関数を適用することにする。

対数スペクトルの不偏推定では、スペクトルをケプストラムによって表現している。従って、 $\gamma = 0$ で一般化ケプストラムはケプストラムに等しいことから、このとき本方法は対数スペクトルの不偏推定⁽⁵⁾と等価と

なる。また、 $\gamma = -1$ において、本方法のスペクトルモデルは全極形となり、また式(12)が線形予測法における評価関数と形式的に同じ形になっていることから、本方法は線形予測法と等価となる。

ここで、 $H(z)$ を音声の合成フィルタと考えれば、 $H(z)$ は安定でなければならないことから、 $H(z)$ は最小位相を仮定する。このとき、

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log |H(e^{j\omega})|^2 d\omega = \log K^2 \quad (13)$$

の関係が成り立つので、評価関数、式(11)は、

$$E = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|D(e^{j\omega})|^2} d\omega / K^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |X(e^{j\omega})|^2 d\omega + \log K^2 - 1 \quad (14)$$

と書くことができる。従って、 E の $c_r(m)$ に関する最小化問題は、

$$\varepsilon = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|D(e^{j\omega})|^2} d\omega \quad (15)$$

の最小化問題と等価となる。

3. 解 法

3.1 決定方程式

$c_r(m)$ を決定するため、式(15)を $c_r(m)$ について微分して零と置くことにより、次の決定方程式を得る。

$$g(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|D^{1+\gamma}(e^{j\omega})|^2} D^{\gamma}(e^{j\omega}) e^{j\omega k} d\omega = 0, \quad k=1, 2, \dots, M \quad (16)$$

ε は、 $\mathbf{c} = [c_r(1) \ c_r(2) \ \dots \ c_r(M)]^T$ に関して、凸であることから(付録1.)、式(16)の解は一意に存在し、 ε を最小にする \mathbf{c} を与える。更に、式(16)の解 $c_r(m)$ によって定められる $H(z)$ は、必ず安定となる(付録2.)。

また、 ε の最小値を ε_{\min} とすれば、 E を最小にする K は、

$$K = \sqrt{\varepsilon_{\min}} \quad (17)$$

で与えられる。

3.2 決定方程式の解法

式(16)の決定方程式は、 $\gamma = -1$ では線形予測法での正規方程式に一致するが、 $\gamma = -1$ 以外では非線形方程式となるので、Newton-Raphson法により ε を最小化する \mathbf{c} を求める。つまり、 i 番目の近似値 $c_r^{(i)}(l)$ に対し、線形方程式

$$\sum_{l=1}^M \frac{\partial g(k)}{\partial c_r(l)} \Delta c_r(l) = -g(k), \quad k=1, 2, \dots, M \quad (18)$$

を $\Delta c_r(l)$ について解き、次回の近似値

$$c_r^{(i+1)}(l) = c_r^{(i)}(l) + \Delta c_r(l), \quad l=1, 2, \dots, M \quad (19)$$

を得る。

式(16)の微分を計算し、式(18)に代入すれば、次の方程式が得られる。

$$(\mathbf{R} + (1+\gamma)\mathbf{Q})\Delta\mathbf{c} = \mathbf{g} \quad (20)$$

但し、

$$\mathbf{R} = \begin{bmatrix} r(0) & \dots & r(M-1) \\ \vdots & \ddots & \vdots \\ r(M-1) & \dots & r(0) \end{bmatrix} \quad (21)$$

$$\mathbf{Q} = \begin{bmatrix} q(2) & \dots & q(M+1) \\ \vdots & \ddots & \vdots \\ q(M+1) & \dots & q(2M) \end{bmatrix} \quad (22)$$

$$\Delta\mathbf{c} = [\Delta c_r(1) \ \Delta c_r(2) \ \dots \ \Delta c_r(M)]^T \quad (23)$$

$$\mathbf{g} = [g(1) \ g(2) \ \dots \ g(M)]^T \quad (24)$$

また、

$$r(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|D^{1+\gamma}(e^{j\omega})|^2} e^{j\omega k} d\omega \quad (25)$$

$$q(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|D^{1+2\gamma}(e^{j\omega})|^2} D^{2\gamma}(e^{j\omega}) e^{j\omega k} d\omega \quad (26)$$

式(20)の線形方程式は、係数行列が対称テプリーブツ行列とハンケル行列の和になっているため、 $O(M^2)$ の高速算法⁽⁶⁾によって解くことができる。従って、演算量は、方程式の係数を求めるためのフーリエ変換によるものが支配的となる。

式(15)の ε は、 $\mathbf{c} = [c_r(1) \ c_r(2) \ \dots \ c_r(M)]^T$ に関して凸(付録1.)、つまり ε のヘッセ行列が正定であることから、式(20)および式(19)の繰返しにより2次収束解を得ることができる。

3.3 初期値

$\gamma = -1$ のときには、 ε は2次関数になるので、どのような初期値を与えても、1回の繰返しで解が得られる。特に、初期値として $\mathbf{c} = [0 \ \dots \ 0]^T$ を与えたときには、式(20)は線形予測法での正規方程式に一致する。

$\gamma = -1$ 以外のときの初期値は、できるだけ収束値に近いことが望ましいが、このような初期値として、一般化対数スペクトルの線形平滑化⁽⁴⁾によって得られる一般化ケプストラムを用いることができる。この一般化ケプストラムは、再帰式⁽⁷⁾

$$c_r(m) = c(m) + \sum_{k=1}^{m-1} \frac{k}{m} \gamma c(k) c_r(m-k), \quad m \geq 1 \quad (27)$$

によって、ケプストラム法により得られるケプストラム $c(m)$ から簡単に計算することができる。また、同様に、線形予測法により得られる予測係数 $a(m)$ から変換された一般化ケプストラムを用いることもでき、その再帰式は、

$$c_r(m) = -a(m) - \sum_{k=1}^{m-1} \frac{k}{m} (\gamma a(k) c_r(m-k))$$

$$+c_r(k)a(m-k), \quad m \geq 1 \quad (28)$$

で与えられる⁽⁷⁾.

4. 本スペクトル推定法におけるいくつかの性質

以上で述べたスペクトル推定法は、対数スペクトルの不偏性の条件に基づいた方法であるが、評価関数の形は、信号をガウス過程とした場合の最尤推定法におけるスペクトル評価関数と同じ形となっている。その他にもいくつかの興味深い性質をもっているの、これらについて、図1にそって簡単な検討を加える。

但し、図1において、 $e(m)$ は、信号 $x(m)$ を逆フィルタ $1/D(z)$ に通したときの出力であり、 $h(m)$ は、 $H(z)=K \cdot D(z)$ のインパルス応答とする。また、 $D^{1+\gamma}(z)$ と $D^{-\gamma}(z)$ は、 $D(z)=D^{1+\gamma}(z) \cdot D^{-\gamma}(z)$ の関係を持ち、信号 $x(m)$ を逆フィルタ $1/D^{1+\gamma}(z)$ に通した

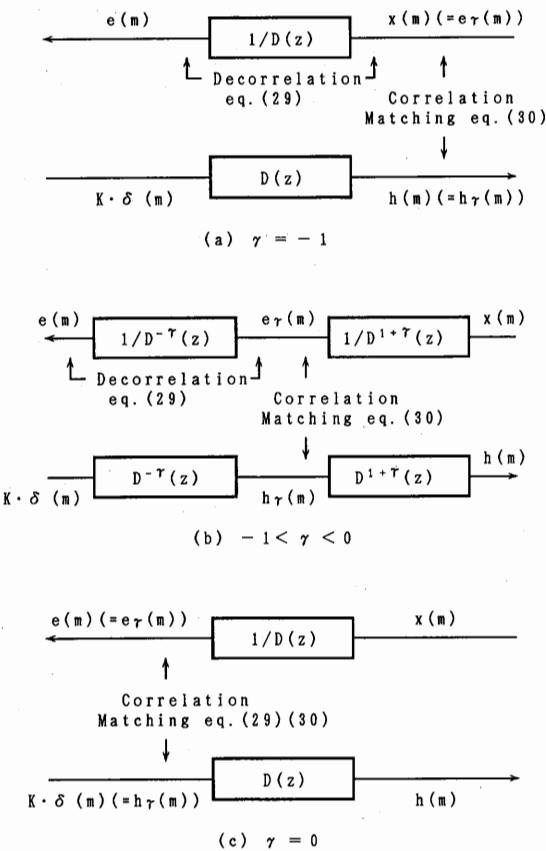


図1 一般化ケプストラムをパラメータとするスペクトル推定におけるいくつかの性質

Fig. 1 Properties of the estimated spectrum based on the generalized cepstral representation.

ときの出力を $e_r(m)$, $K \cdot D^{-\gamma}(z)$ のインパルス応答を $h_r(m)$ とする。

[逆フィルタ分析] 式(15)から明らかなように、 ϵ の最小化は、逆フィルタ $1/D(z)$ の出力 $e(m)$ のエネルギー最小化問題と等価となる。

[無相関化] 式(16)を

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{X(e^{j\omega})}{D(e^{j\omega})} \right] \left[\frac{X(e^{j\omega})}{D^{1+\gamma}(e^{j\omega})} \right]^* e^{j\omega k} d\omega = K^2 \cdot \delta(k), \quad k=0, 1, \dots, M \quad (29)$$

と書き換えれば、図1(b)に示すように、逆フィルタ $1/D(z)$ の出力 $e(m)$ とフィルタ $1/D^{1+\gamma}(z)$ の出力 $e_r(m)$ の相互相関は、1次から M 次までが零、0次が K^2 となることがわかる。特に、 $\gamma=-1$ のときには、図1(a)に示すように、信号 $x(m)$ と逆フィルタ出力 $e(m)$ の相互相関が1次から M 次まで零となる。

[相関整合] 線形予測法、つまり本方法において $\gamma=-1$ のときには、図1(a)に示すように、信号 $x(m)$ と合成フィルタ $H(z)=K \cdot D(z)$ の自己相関は M 次まで一致することが知られている⁽¹⁾。 $\gamma=-1$ 以外のときにも、式(16)から、 $1/C(z)=D^{-\gamma}(z)$ が因果的であることを用いて同様の関係式

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|D^{1+\gamma}(e^{j\omega})|^2} e^{j\omega k} d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |K \cdot D^{-\gamma}(e^{j\omega})|^2 e^{j\omega k} d\omega, \quad k=0, 1, \dots, M \quad (30)$$

を導出することができる。従って、図1(b)に示すように、フィルタ $1/D^{1+\gamma}(z)$ の出力 $e_r(m)$ とフィルタ $K \cdot D^{-\gamma}(z)$ の自己相関が M 次まで一致することになる。特に、 $\gamma=0$ のときには、図1(c)に示すように、逆フィルタ出力 $e(m)$ と $K \cdot \delta(m)$ の自己相関が一致するので、逆フィルタ出力 $e(m)$ の自己相関は、 M 次まで零となる。このような性質は、逆フィルタ出力 $e(m)$ の符号化などで有利な点をもたらすと予想される。

5. スペクトルの分析例

5.1 合成信号の分析

本方法による合成信号の分析例を図2に示す。標準化周波数は10 kHz、時間窓は25.6 msのハミング窓を用いた。分析次数は $M=12$ である。信号はデジタルフィルタを正規乱数列で励振した出力で、フィルタの中心周波数(Hz)/帯域幅(Hz)は、共振点1000/200, 3500/250, 反共振点2500/200である。初期値は、従来のケプストラム法によって得られるケプストラムから

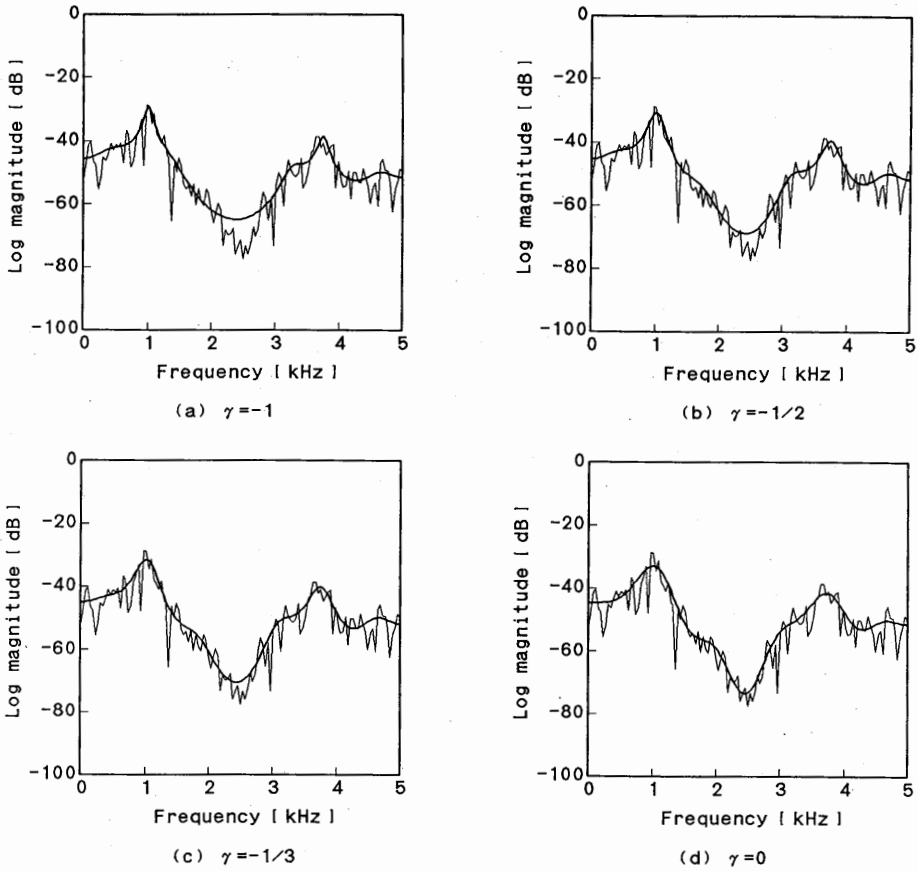


図 2 合成信号のスペクトル推定例 ($M=12$)
 Fig. 2 Spectral estimates of a synthetic signal for several values of γ ($M=12$).

再帰式により求めた一般化ケプストラムを用いている。

図 2 (a) の $\gamma = -1$ のスペクトルは、線形予測法によって得られるスペクトルと同一であり、(d) の $\gamma = 0$ のスペクトルは、対数スペクトルの不偏推定によって得られるスペクトルと同一である。図より、 γ が -1 に近いほど極によるスペクトルのピークが、逆に γ が 0 に近いほど零による谷が鋭くなっている。図 2 (c)、(d) から、 γ を -1 から 0 の間の適当な値に選ぶことにより、零によるスペクトルの谷をとらえながら、極付近である程度鋭いピークを実現できることがわかる。

図 2 の例について、 E の収束の様子を図 3 に示す。 $\gamma = -1$ のときには、繰返しなしで解を得ることができる。その他の γ の値に対しては、図 3 からわかるように収束は非常に速く、一般に繰返しは数回で十分である。

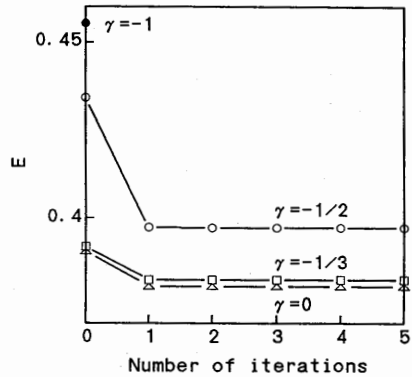


図 3 繰返し回数に対する評価関数 E の変化
 Fig. 3 Error E versus the number of iterations.

5.2 自然音声の分析

式 (11) の評価式は、信号を弱定常過程としているが、有声音などの周期信号に適用することもできる。これは、式 (11) の評価式と、周期信号のスペクトル包絡抽

出のための評価式⁽⁶⁾が、類似した形をしていることから予想できることである。

図4に女性話者による/ne/の分析例を示す。分析条件は、分析次数を $M=15$ としている点を除いて、図2の場合と同じである。

$\gamma=-1$ では、信号の高調波成分にスペクトルのピークがマッチしたため、ホルマントの帯域幅が狭くなり過ぎており、また2 kHz付近の零をとらえていない。 $\gamma=0$ のときには、零をよくとらえているが、逆にホルマントの帯域幅が大きくなっている。図4(c)より、 $\gamma=-1/3$ では零による谷を見落とさず、また極の部分でも適当な帯域幅となっていることがわかる。

上記すべての実験、および他の連続音声の分析においても、すべての分析フレームで、得られる伝達特性が安定となることが確認されている。

5.3 音声の分析合成と最適な γ の値に関する検討

γ の値もスペクトルを表すパラメータの一つと考え

れば、評価式 E を γ に関しても最小化しなければならないが、評価式 E を最小化する γ の値は、信号の性質、例えば音声の場合、音韻の種類などによって異なる。例として、図5に、男性話者による/a/と/n/について、 γ の値に対する E の値を示す。伝達特性の共振点はその音韻を特徴づける/a/では、 $\gamma=-1$ で E の値が小さくなり、逆に反共振点はその音韻を特徴づける/n/では、 $\gamma=0$ に近い値で E の値が小さくなっている。

しかし、 γ の値を可変とすることは容易ではないので、実際には、 γ として一定値を用いることになるが、その際、音声合成あるいは認識において最も良い結果を与える値を用いることが望ましい。この値の選択は、 $\gamma=-1$ の全極モデルの利点と、 $\gamma=0$ のケプストラムによるスペクトルモデルの利点とのトレードオフと言える。ここでは、このような γ の値を見つけるため、 $\gamma=-1, -1/2, -1/3, 0$ とした音声の分析合成音の非公式な聴取試験を行った。

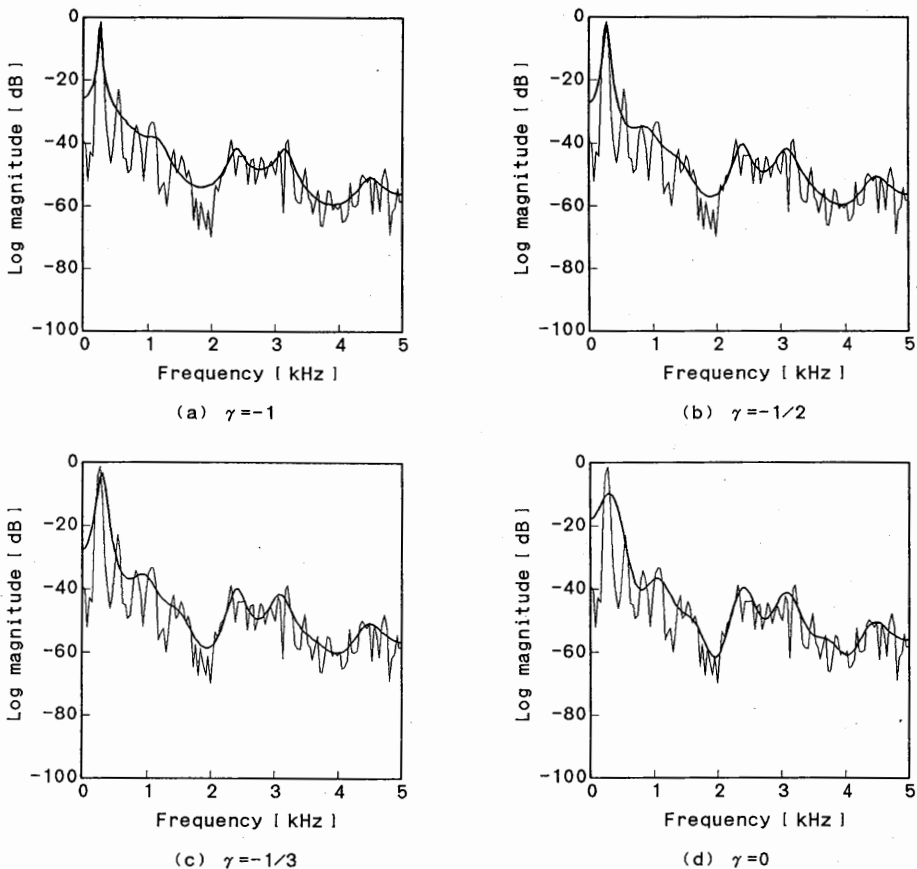
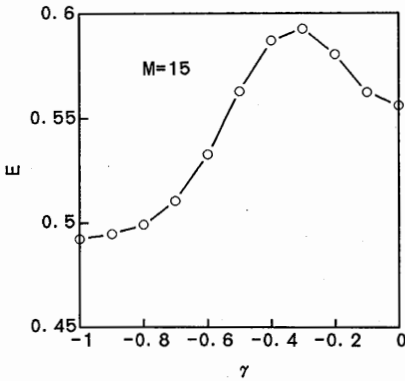
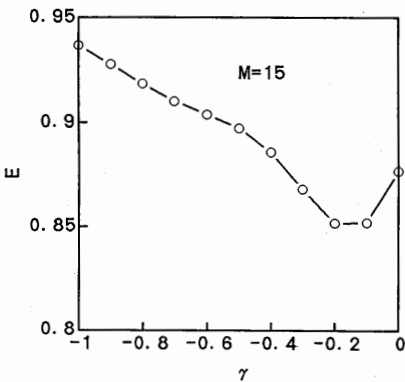


図4 自然音声のスペクトル推定例 ($M=15$)
Fig. 4 Spectral estimates of a natural speech /ne/ uttered by a female for several values of γ ($M=15$).



(a) /a/



(b) /n/

図5 γ に対する最小化された E の値Fig. 5 Minimized error E versus the value of γ

式(7)の合成フィルタは、 $\gamma=0$ ではLMAフィルタ⁽⁹⁾により、任意の γ に対してはGLSAフィルタ⁽¹⁰⁾により高い精度で実現できるが、特に n を自然数として、 $\gamma=-1/n$ のときには、全極フィルタの n 段縦続という簡単な構造で実現することができる⁽¹¹⁾。

合成音は、 γ が-1に近いほど歯切れが良い音に聞こえ、逆に0に近いほど滑らかな音に聞こえた。分析合成音の簡単なプレファレンステストでは、 $\gamma=-1/2$ または $-1/3$ で高いスコアを示した。但し、規則合成などでは音の滑らかさが重要となるので、最適な γ の値はより0に近い値となることが予想される。

6. むすび

本論文では、一般化ケプストラムで表現されたスペクトルモデルに対して、対数スペクトルの不偏推定におけるスペクトル評価を適用したスペクトル推定法を提案した。また、このときの最小化問題の解法を与え、

解の収束性、得られる伝達特性の安定性を示した。本方法は、スペクトルのモデルを全極形から、ケプストラムで表現された形まで連続的に変化させることができるので、信号に適したモデルを選ぶことができる。これらの特長から、本方法は、新しいスペクトル推定法として、音声以外の信号のスペクトル分析にも有用であると考えられる。

評価式 E あるいは ε の値と、合成音の好ましきさ、あるいは音声認識における認識率との関係の考察、次数 M の選択、より詳しい γ の最適値についての検討などは、今後の課題である。

本方法は、線形予測法、対数スペクトルの不偏推定法を含む手法であり、線形予測法とケプストラム法を統一的に扱う手法と考えることができる。従って、線形予測法の応用として定式化されている適応分析法⁽¹³⁾などをケプストラム法にまで拡張できると予想される⁽¹⁴⁾が、このような点については稿を改めて報告したい。

文 献

- (1) J. D. Markel and A. H. Gray, Jr.: "Linear Prediction of Speech", Springer-Verlag Berlin Heidelberg (1976).
- (2) A. V. Oppenheim and R. W. Schaffer: "Digital Signal Processing", Prentice-Hall, Englewood-Cliffs, NJ (1975).
- (3) T. Kobayashi and S. Imai: "Spectral analysis using generalized cepstrum", IEEE Trans. Acoust., Speech & Signal Process., **ASSP-32**, 5, pp. 1087-1089 (Oct. 1984).
- (4) 小林隆夫, 今井 聖: "一般化対数目盛における音声スペクトルの平滑化法", 信学論(A), **J64-A**, 6, pp. 473-474 (昭56-06).
- (5) 今井 聖, 古市千枝子: "対数スペクトルの不偏推定", 信学論(A), **J70-A**, 3, pp. 471-480 (昭62-03).
- (6) G. A. Merchant and T. W. Parks: "Efficient solution of a Toeplitz-plus-Hankel coefficient matrix system of equations", IEEE Trans. Acoust., Speech & Signal Process., **ASSP-30**, 1, pp. 40-44 (Feb. 1982).
- (7) 徳田恵一, 小林隆夫, 今井 聖: "メル一般化ケプストラムの再帰的計算法", 信学論(A), **J71-A**, 1, pp. 128-131 (昭62-01).
- (8) 徳田恵一, 小林隆夫, 今井 聖: "スペクトル包絡の抽出のための非一様スペクトル荷重によるケプストラム分析", 信学論(A), **J70-A**, 6, pp. 952-959 (昭62-06).
- (9) 今井 聖: "対数振幅近似(LMA)フィルタ", 信学論(A), **J63-A**, 12, pp. 886-893 (昭55-12).
- (10) 小林隆夫, 今井 聖: "一般化ケプストラムによる音声の合成", 信学論(A), **J65-A**, 3, pp. 245-252 (昭57-03).
- (11) 徳田恵一, 小林隆夫, 今井 聖: "一般化ケプストラムをパラメータとする音声合成", 昭63信学春季全大, A-34.
- (12) 美多 勉: "ディジタル制御理論", 昭晃堂 (昭59).
- (13) S. T. Alexander: "Adaptive signal processing",

Springer-Verlag, New York (1986).

(14) 徳田恵一, 小林隆夫, 塩本祥司, 今井 聖: “適応ケブストラム分析の検討”, 信学技報, SP88-53 (1988-09).

付 録

1. ϵ が凸であることの証明

ϵ が c に関して凸, つまり

$$\epsilon(sc_1 + (1-s)c_2) \leq s\epsilon(c_1) + (1-s)\epsilon(c_2), 0 \leq s \leq 1 \tag{A\cdot 1}$$

であることを示す. c_1, c_2 は任意に選んだ二つの c である.

$\gamma \neq -1$ のとき, 式(A\cdot 1)の左辺は,

$$\begin{aligned} &\epsilon(sc_1 + (1-s)c_2) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|s\gamma^{-1}(sF_1(e^{j\omega}) + (1-s)F_2(e^{j\omega}))|^2} d\omega \end{aligned} \tag{A\cdot 2}$$

但し,

$$F_1(z) = \sum_{m=1}^M c_1(m)z^{-m} \Big|_{c=c_1} \tag{A\cdot 3}$$

$$F_2(z) = \sum_{m=1}^M c_2(m)z^{-m} \Big|_{c=c_2} \tag{A\cdot 4}$$

と書ける.

ここで, $w = x + jy$ の関数

$$f(w) = 1/|s\gamma^{-1}(w)|^2 \tag{A\cdot 5}$$

を考え, $f(w)$ のヘッセ行列 J

$$J = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \tag{A\cdot 6}$$

をつくると,

$$|J| = -4\gamma(2+\gamma)/|s\gamma^{-1}(w)|^{4(1+\gamma)} \tag{A\cdot 7}$$

であり, また,

$$\frac{\partial^2 f}{\partial x^2} = 2 \frac{(2+\gamma)(1+\gamma x)^2 - \gamma(\gamma y)^2}{|s\gamma^{-1}(w)|^{2(1+2\gamma)}} \tag{A\cdot 8}$$

となるので, J は $-1 \leq \gamma \leq 0$ で半正定となることがわかる. 従って, この範囲で $f(w)$ は, 凸関数であり, 任意の w_1, w_2 に対して, 不等式

$$f(sw_1 + (1-s)w_2) \leq sf(w_1) + (1-s)f(w_2) \tag{A\cdot 9}$$

が成り立つ.

従って, この不等式を式(A\cdot 2)に適用することにより,

$$\begin{aligned} &\epsilon(sc_1 + (1-s)c_2) \\ &\leq s \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|s\gamma^{-1}(F_1(e^{j\omega}))|^2} d\omega \\ &\quad + (1-s) \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{|s\gamma^{-1}(F_2(e^{j\omega}))|^2} d\omega \end{aligned}$$

$$= s\epsilon(c_1) + (1-s)\epsilon(c_2) \tag{A\cdot 10}$$

となり, ϵ は, $-1 \leq \gamma \leq 0$ の範囲で凸であることが示される.

2. 解の安定性の証明

$\gamma = 0$ のとき, $H(z)$ は最小位相ケブストラムで表現されており, 因果性を仮定するならば, $c_\gamma(m)$ にどのような値を与えても安定となる.

$-1 \leq \gamma < 0$ のときには, $C(z)$ を

$$C(z) = 1 + \gamma \sum_{m=1}^M c_\gamma(m)z^{-m} \tag{A\cdot 11}$$

とおいて, $1/C(z)$ が安定ならば,

$$H(z) = K \cdot \left[\frac{1}{C(z)} \right]^{-1/\gamma} \tag{A\cdot 12}$$

で与えられる $H(z)$ も安定となる. $C(z)$ のコンパニオン行列 A は,

$$A = \begin{bmatrix} -\gamma c_\gamma(1) \cdots -\gamma c_\gamma(M-1) & -\gamma c_\gamma(M) \\ 1 & 0 \\ & \ddots \\ 0 & 1 & 0 \end{bmatrix} \tag{A\cdot 13}$$

で与えられるので, 式(16)の決定方程式より

$$A^T R A - R = -P \tag{A\cdot 14}$$

のリアプノフ方程式⁽¹²⁾が得られる. 但し

$$P = \begin{bmatrix} K^2 & 0 \cdots 0 \\ 0 & \\ \vdots & 0 \\ 0 & \end{bmatrix} \tag{A\cdot 15}$$

であり, R は式(21), K^2 は式(17)で与えられたものとする. 定義から, P が半正定値行列, R は正定値であるので, $C(z)$ の零は単位円内にあり, $1/C(z)$ つまり $H(z)$ は最小位相となる. 従って, 因果性を仮定すれば, $H(z)$ は安定である.

(昭和63年6月21日受付, 10月17日再受付)



徳田 恵一

昭59名工大・工・電子卒, 昭61東工大大学院修士課程了. 現在, 同大学院博士後期課程在学中. デジタル信号処理, 音声情報処理の研究に従事.



小林 隆夫

昭52 東工大・工・電気卒。昭57 同大大学院博士課程了。同年東工大精密工学研究所助手。工博。デジタル信号処理、音声情報処理の研究に従事。日本音響学会、IEEE 各会員。



山本竜太郎

昭61 東工大・工・電気卒。昭63 同大大学院修士課程了。在学中、音声の分析、合成の研究に従事。現在、東京電力株式会社勤務。



今井 聖

昭34 東工大・工・電気卒。昭39 同大大学院博士課程了。同年東工大・精密工学研究所助手。昭43 同大助教授。昭54 同大教授。工博。デジタル信号処理、音声の合成および認識の研究に従事。45年度精機学会論文賞受賞。著書「デジタル信号処理」など、計測自動制御学会、日本音響学会、IEEE、ASA 各会員。