

雑音補正による音声の LPC 分析の改善

趙 奇方^{†*} 島村 徹也[†] 鈴木 誠史[†]

Improvement of LPC Analysis of Speech by Noise Compensation

Qifang ZHAO^{†*}, Tetsuya SHIMAMURA[†], and Jouji SUZUKI[†]

あらまし 白色雑音の混入した音声信号から自己相関法を用いて予測係数を求める場合、自己相関関数 $R(k)$ の $k=0$ の付近にその雑音の影響は集中する。従って、原理的には、自己相関関数 $R(0)$ から雑音パワーを引くことにより雑音補正が行われ、LPC 分析の耐雑音性が向上される。しかし、従来の雑音補正法では雑音パワーの引きすぎが原因で LPC フィルタが不安定になることがある。従って、その実際の応用は困難と思われる。一方で、スペクトル推定における同様な問題を解決するため、島村らは反復アルゴリズムを利用した改良雑音補正 AR 係数推定法を提案した。本論文ではこの改良雑音補正 AR 係数推定法を音声の LPC 分析の雑音低減に応用する。評価実験からこの方法はプリエンファシス (pre-emphasis) されていない音声に対しては有効であるが、プリエンファシスされた音声に対して改善が見られないことが明らかとなった。その理由として、プリエンファシスによって雑音の影響が自己相関関数 $R(k)$ の $k=1$ の部分にも及んだことを理論解析で示した。そして、プリエンファシスに影響されない、 $R(0)$ と $R(1)$ の双方から雑音パワーを引き去る反復アルゴリズムを導出した。その有効性を計算機シミュレーションで確認している。更にこの方法を拡張し、白色雑音以外の雑音にも対応できる改善法を提案している。

キーワード 線形予測分析, 自己相関法, 雑音補正, プリエンファシス

1. ま え が き

線形予測 (LPC) [1] は音声分析の基本的手法として幅広く利用されている。その耐雑音特性は音声の低ビットレート伝送, 音声認識など多くの分野において重要な課題である [2], [3]。

LPC 分析の主な役割は予測係数を求めることである。音声に雑音を含まないとき、自己相関法によって LPC 係数を正確に求めることができる。しかも LPC フィルタの安定性を保証することもできる。しかし、音声に雑音を加わった場合は、分析精度が大幅に低下する [2], [4]。その解決策として、最小絶対値規範 [5] や weighting された最小 2 乗規範 [6] などを用いた LPC 分析法が提案されている。これらの方法により LPC 分析の雑音耐性はある程度改善されるが、安定性やアルゴリズムの複雑さなどの問題が残る。一方で、実際に最も多く利用されている最小 2 乗規範の LPC 分析法である自己相関法は、その雑音耐性の直

接的な改善法を見出していない。Tierney [4] は LPC 分析の次数を増やす方法を提案したが、実質的な改善は見られなかった。

そこで、我々は雑音補正を利用した LPC 分析方式に着目する。自己相関法で予測係数を求める場合は、音声の自己相関関数 $R(k)$ の $k=0 \sim p$ (p は LPC の次数) の部分だけを利用する。このとき、もし $R(k)$, $k=0 \sim p$ から雑音成分のみを除去できれば予測係数の推定精度の改善が期待できる。特に白色雑音の場合は、雑音成分は $R(0)$ 付近に集中する性質をもつため、 $R(0)$ だけから雑音成分を引くことにより、LPC 分析の雑音耐性を改善することが考えられる。しかし、この雑音補正 LPC 分析法には一つの大きな問題点がある。それは、雑音パワーの引きすぎが原因で LPC フィルタが不安定になることである。このとき、LPC フィルタのシステム関数 $H(z)$

$$H(z) = \frac{G}{1 - \sum_{j=1}^p a_j z^{-j}} \quad (1)$$

の極点は単位円の外側に出てしまう。ここで、 a_j と G はそれぞれ予測係数と利得である。

従来、雑音補正については主にスペクトル推定の分

[†] 埼玉大学工学部情報システム工学科, 浦和市
Faculty of Engineering, Saitama University, Urawa-shi, 338-8570
Japan

* 現在, (株)NTT データ

野において研究されてきた。さまざまな方法が提案されてきたが[7],[8], 安定性の保証については依然困難をきたしている。しかし, 近年, 島村らは改良雑音補正 AR 係数推定法を提案し[9], 反復アルゴリズムによって安定性の問題を解決した。そして, 正弦波の周波数推定実験で白色雑音に対して雑音低減効果を示した。

本論文ではまずこの改良雑音補正 AR 係数推定法を音声信号に適用し, シミュレーション実験によりその有効性を評価する。そして, その結果からプリアンファシスとの関係を調べ, 改良雑音補正 AR 係数推定法はプリアンファシスされていない音声に対しては有効であるが, プリアンファシスされた音声に対しては有効でないことを明らかにする。この結果をかんがみ, 理論解析を行い, そこから新しい雑音補正 LPC 分析法を導出する。更には本提案法を拡張し, 白色雑音以外の雑音にも対応できる改善法へと発展させる。これら二つの提案法の有効性はケプストラム距離を評価することにより確かめられる。

2. 改良雑音補正 AR 係数推定法とその評価

雑音が混入した音声信号を $x(n)$ とすると

$$x(n) = s(n) + w(n) \quad (2)$$

となる。ここで, $s(n)$ は音声信号であり, $w(n)$ は雑音信号である。LPC 分析の自己相関法は $x(n)$ の自己相関関数 $R(k)$, $k=0 \sim p$ を利用して予測係数を求める。 $w(n)$ が白色雑音である場合, 信号 $s(n)$ と雑音 $w(n)$ は無相関になる。それで

$$R(k) = \begin{cases} R_s(k) + \sigma_w^2 & k=0 \\ R_s(k) & k \neq 0 \end{cases} \quad (3)$$

となる。ここで $R_s(k)$ は音声信号 $s(n)$ の自己相関関数であり, σ_w^2 は雑音のパワーである。式(3)から雑音の影響が $R(k)$ の $k=0$ のところに集中することは明らかである。そこで, 雑音を低減するために, 次のように $\bar{R}(0)$ から雑音のパワーを引くことが考えられる[10]。

$$\bar{R}(k) = \begin{cases} R(k) - \sigma_w^2 & k=0 \\ R(k) & k \neq 0 \end{cases} \quad (4)$$

ここで, σ_w^2 は雑音パワーの推定値を意味する。予測係数を求めるとき, $R(k)$, $k=0 \sim p$ の代わりに, $\bar{R}(k)$, $k=0 \sim p$ を利用すれば, 精度が向上される。しかし, 式(4)のような雑音低減法を実際に適用する場合, 最も大きな問題点が雑音パワーの引きすぎによ

って音声の LPC フィルタが不安定になることである。この問題の解決にあたって島村らは改良雑音補正 AR 係数推定法を提案した[9]。本論文では, まずそれを音声に適用することを考える。

2.1 改良雑音補正 AR 係数推定法[9]

改良雑音補正 AR 係数推定法は反射係数を判断基準とし, 反復アルゴリズムを用いて AR スペクトル推定の安定性の問題を解決した。この方法を利用して白色雑音の混入した音声から予測係数を求める場合, その手順は次のようになる。

(i) 雑音のパワーを推定し, その結果を σ_w^2 とする。

(ii) 雑音が混入した音声信号 $x(n)$ の自己相関関数を計算する。

$$R(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n)x(n+k), \quad k=0 \sim p \quad (5)$$

ここで, N は分析窓長である。

(iii) アルゴリズムの反復回数 i の初期設定を行う。

$$i=0 \quad (6)$$

(iv) 遅延 $k=0$ における雑音成分の補正を次のように行う。

$$\bar{R}(k) = \begin{cases} R(k) - (\beta - \alpha i) \sigma_w^2 & k=0 \\ R(k) & k \neq 0 \end{cases} \quad (7)$$

ここで, α , β はあらかじめ設定すべきパラメータである。 β は σ_w^2 の推定誤差を考慮し, $1.0 < \beta < 1.5$ を設定する。 α は σ_w^2 の量を増大させる刻み幅に対応し, 本アルゴリズムの反復回数を左右する。ここでの α の選択に関しては, 計算量と推定精度のトレードオフが存在することになるが, 多くの場合, $0.1 \leq \alpha \leq 0.3$ 程度を設定するとよい。結局, 式(7)により, アルゴリズムの反復ごとに $R(0)$ から雑音パワーを徐々に引くことになる。

(v) 式(7)の結果 $\bar{R}(k)$, $k=0 \sim p$ を用いて Levinson-Durbin 法で予測係数と反射係数を求める。

(vi) 反射係数を K_j とするとき,

$$|K_j| < 1, \quad j=1 \sim p \quad (8)$$

を満たすかどうかを調べ, 満足していれば, (v) で得られた予測係数を解とし, そうでなければ,

$$i=i+1 \quad (9)$$

として, (iv)へ戻る。

上記の(i)~(vi)が改良雑音補正 AR 係数推定法の手順である。本法において, (iv)から(vi)までの反復操作は

$$\beta = c\alpha \quad (10)$$

を満足すれば必ず収束する。但し、 $\alpha > 0$, $\beta > 0$, c はある正の整数である。なぜなら、アルゴリズムの反復ごとに

$$\beta - \alpha i \rightarrow 0 \quad (11)$$

となるならば、

$$\tilde{R}(0) \rightarrow R(0) \quad (12)$$

となるからである。少なくとも $\tilde{R}(0) = R(0)$ のとき、式(8)は必ず満足される。

2.2 実験評価

音声信号に対する改良雑音補正 AR 係数推定法の有効性を実音声を用いたシミュレーション実験で評価する。改良雑音補正 AR 係数推定法は反復アルゴリズムによって LPC フィルタの安定性を保証することができるため、我々の関心は LPC 分析の雑音耐性の改善に向けられる。

音声の LPC 分析にはプリエンファシスが必要とされる場合が多い。そのため、我々はプリエンファシスをする場合としない場合の双方についてシミュレーションを行うことにする。

2.2.1 実験条件

2種類の音声試料を使用する。

A. 男性の発声した、長さが1秒ずつの日本語母音 /i/, /e/, /a/, /o/, /u/ である (Data A)。

B. 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」に収録されている男女各4名の発

声した単文である (Data B)。

A によって各改善法を細かく評価することができる。一方、B によって音声信号全体に対して各方法の効果がわかる。A, B はともに 3.4 kHz に帯域制限され、標本化周波数 10 kHz で標本化されている。また、付加雑音は白色ガウス雑音であり、その振幅値を変えて音声に加えることにより SN 比が調整される。ここでは雑音を帯域制限しないことにする。

無声区間は評価実験から除外される。無声区間を判別するために信号を 5 ms のフレームに区切った。そして、それぞれのフレームにおける信号の最大振幅が、信号全体の最大振幅の -30 dB 以下であるとき、無声区間とする。

反復アルゴリズムのパラメータ α と β の値は予備実験の結果からそれぞれ 0.1 と 1.2 に決定した。このとき、最悪の場合でも 12 回の反復を経てアルゴリズムは収束することになる。

表1は、本シミュレーション実験における諸定数をまとめている。なお、プリエンファシスをする場合は、プリエンファシス係数を 0.95 とする。

2.2.2 評価方法

雑音の影響がどの程度低減されるかをケプストラム距離で評価する。具体的には

(1) Levinson-Durbin 法で音声の予測係数を求める。

(2) 音声に雑音を加え、Levinson-Durbin 法で予測係数を求め、(1)とのケプストラム距離 cdn を計算する。

(3) (2)と同様にし、今度は改良雑音補正 AR 係数推定法を用いて予測係数を求め、(1)とのケプストラム距離 $cden$ を計算する。

(4) ケプストラム距離の改善度

表1 実験の諸定数

Table1 Experimental parameter specification.

標本化周波数	10kHz
分析窓	Hamming, 25.6ms
分析周期	5ms
LPC 次数	12

表2 従来法(改良雑音補正 AR 係数推定法)によるケプストラム距離の改善度 (dB) (白色雑音、左段プリエンファシスなし、右段プリエンファシスあり)

Table2 Improvement degree of cepstrum distance by conventional method (dB). (white noise. The two value for each SNR are for the cases of not using and using pre-emphasis)

SNR (dB)	-5		0		5		10		15		20	
i	-0.09	-0.05	-0.32	-1.90	-0.30	-1.01	0.09	-1.17	0.09	-1.55	0.28	-2.26
e	0.41	-0.18	0.95	-1.70	1.80	-2.34	2.42	-3.16	3.00	-3.28	2.84	0.70
a	0.85	-0.06	1.29	-1.59	1.50	-2.51	2.05	-3.30	2.50	-3.20	3.03	1.02
o	0.70	-0.69	0.96	-1.92	1.19	-2.96	1.06	-2.05	0.77	-1.62	0.57	-1.18
u	0.33	-0.11	0.35	-2.28	0.56	-2.53	1.16	-3.11	1.88	-3.39	1.64	-3.19
Average	0.44	-0.22	0.65	-1.92	0.95	-2.27	1.36	-2.56	1.55	-2.61	1.67	-1.18
Data B	0.18	-0.43	0.56	-1.21	0.97	-1.88	1.42	-2.03	1.83	-1.66	2.08	-1.16

$$Cep = cdn - cden \quad (13)$$

を計算する。式(13)の改善度で各方法を評価する。改良雑音補正 AR 係数推定法は、 $Cep > 0$ のとき有効であるが、 $Cep \leq 0$ のとき有効でない。

2.2.3 実験結果

表2は、改良雑音補正 AR 係数推定法から得られる推定結果のケプストラム距離改善度をまとめたものである。

プリアンファシスをしない場合、表2からは、全体的にケプストラム距離が改善され、改良雑音補正 AR 係数推定法が有効であることが見てとれる。しかし、母音/i/については SN 比が低いときに改善が見られていない。この理由としては次の二つが考えられよう。一つは、i/のスペクトルは振幅の変化が大きいので、ほかの母音より雑音に影響されやすいことである。もう一つは、実際に使用した白色雑音は理論的なものと多少差があり、また自己相関関数を有限区間で算出していることである。0以外の遅延においても自己相関関数の値は完全に0になっていない。それらの振幅の大きさは $R(0)$ と比べて極めて小さいが、SN 比が低下するにつれて音声自身に対して悪い影響が次第に強まってしまう。このため0遅延の自己相関関数から雑音を引いたとしても実際に予測係数の推定が改善できるとは限らない。

一方、プリアンファシスを行った場合、表2からはケプストラム距離の改善はほとんど見られていない。改良雑音補正 AR 係数推定法は有効でないと言える。

3. 提案法

プリアンファシスをする場合は、改良雑音補正 AR 係数推定法が有効でないことが2.での実験結果で示された。本章はその理由を理論的に解析した上、プリアンファシスをする場合としない場合の双方に有効な新しい雑音補正 LPC 分析法を提案する。

3.1 理論的解析

式(2)のように音声信号 $s(n)$ に雑音信号 $w(n)$ を加える。ここで再び記述すれば

$$x(n) = s(n) + w(n) \quad (2)$$

である。但し、ここでは $w(n)$ を白色雑音とする。

プリアンファシスされた音声信号を $\hat{s}(n)$ とすると

$$\hat{s}(n) = s(n) - \mu s(n-1) \quad (14)$$

となる。ここで、 μ はプリアンファシス係数である。

$\hat{s}(n)$ の自己相関関数を $\hat{R}_s(k)$ とすると

$$\hat{R}_s(k) = \sum_{n=-\infty}^{\infty} \hat{s}(n) \hat{s}(n+k) \quad (15)$$

となる。

同様に、プリアンファシスされた白色雑音 $\hat{w}(n)$ とその自己相関関数 $\hat{R}_w(k)$ を次の式で表すことができる。

$$\hat{w}(n) = w(n) - \mu w(n-1) \quad (16)$$

$$\hat{R}_w(k) = \sum_{n=-\infty}^{\infty} \hat{w}(n) \hat{w}(n+k) \quad (17)$$

プリアンファシスする前の白色雑音 $w(n)$ の自己相関関数を $R_w(k)$ とすると

$$R_w(k) = \sum_{n=-\infty}^{\infty} w(n) w(n+k) \quad (18)$$

となる。但し

$$R_w(k) = \begin{cases} \sigma_w^2 & k=0 \\ 0 & k \neq 0 \end{cases} \quad (19)$$

である。

式(16)を式(17)に代入し、式(18)を利用すれば、式(17)は

$$\hat{R}_w(k) = (1 + \mu^2) R_w(k) - \mu R_w(k-1) - \mu R_w(k+1) \quad (20)$$

と書き換えられる。更に式(19)を式(20)に代入すれば

$$\hat{R}_w(k) = \begin{cases} (1 + \mu^2) \sigma_w^2 & k=0 \\ -\mu \sigma_w^2 & k=1 \\ 0 & k > 1 \end{cases} \quad (21)$$

となる。

雑音の混入した音声信号 $x(n)$ をプリアンファシスすると

$$\begin{aligned} \hat{x}(n) &= x(n) - \mu x(n-1) \\ &= s(n) - \mu s(n-1) + w(n) - \mu w(n-1) \\ &= \hat{s}(n) + \hat{w}(n) \end{aligned} \quad (22)$$

となる。従って、 $\hat{x}(n)$ の自己相関関数は式(21)を用いて

$$\begin{aligned} \hat{R}(k) &= \sum_{n=-\infty}^{\infty} \hat{x}(n) \hat{x}(n+k) \\ &= \hat{R}_s(k) + \hat{R}_w(k) \\ &= \begin{cases} \hat{R}_s(k) + (1 + \mu^2) \sigma_w^2 & k=0 \\ \hat{R}_s(k) - \mu \sigma_w^2 & k=1 \\ \hat{R}_s(k) & k > 1 \end{cases} \end{aligned} \quad (23)$$

となる。式(23)に示されるように、プリアンファシスをする場合は、自己相関関数 $\hat{R}(0)$ だけではなく、 $\hat{R}(1)$ も雑音に影響される。このため、 $\hat{R}(0)$ のみから雑音パワーを引くのは雑音の低減に不十分とわかる。

プリアンファシスされた白色雑音 $\hat{w}(n)$ のパワーを P_w とすれば、式(21)より

$$P_w = \bar{R}_w(0) = (1 + \mu^2) \sigma_w^2 \quad (24)$$

となる。式(24)を用いると、式(23)は

$$\bar{R}(k) = \begin{cases} \bar{R}_s(k) + P_w & k=0 \\ \bar{R}_s(k) - \gamma P_w & k=1 \\ \bar{R}_s(k) & k>1 \end{cases} \quad (25)$$

と書き換えられる。但し、ここで

$$\gamma = \frac{\mu}{1 + \mu^2} \quad (26)$$

である。

3.2 提案法 1

白色雑音が混入した音声に対するプリエンファシスの悪影響は、これまでの実験結果と理論解析で裏づけられよう。3.1での結果は、明らかに正確なLPC係数を求めるためには、音声の自己相関関数 $\bar{R}(0)$ だけでなく、 $\bar{R}(1)$ からも雑音成分を引き去ることが不可欠であることを示唆する。そこで、白色雑音が混入した音声信号のLPC分析を改善するため、我々は次の改善法を提案する。

(i) 雑音信号 $\hat{w}(n)$ のパワーを推定し、その結果を \bar{P}_w とする。

(ii) 雑音が混入した音声信号 $\hat{x}(n)$ の自己相関関数 $\bar{R}(k)$, $k=0 \sim p$ を計算する。

$$\bar{R}(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} \hat{x}(n) \hat{x}(n+k) \quad (27)$$

(iii) アルゴリズムの反復回数 i の初期設定を行う。

$$i=0 \quad (28)$$

(iv) 式(25)に基づき、雑音補正を行う。

$$\bar{R}(k) = \begin{cases} \bar{R}(k) - (\beta - \alpha i) \bar{P}_w & k=0 \\ \bar{R}(k) + \gamma(\beta - \alpha i) \bar{P}_w & k=1 \\ \bar{R}(k) & k>1 \end{cases} \quad (29)$$

β と α は2.1と同じ方法で設定すればよい。 γ は式(26)により決められる。

(v) (iv)の結果 $\bar{R}(k)$ $k=0 \sim p$ を用いてLevinson-Durbin法で予測係数と反射係数を求める。

(vi) 反射係数 K_j が

$$|K_j| < 1 \quad j=1 \sim p \quad (30)$$

を満たすかどうかを調べ、満足していれば、(v)で得られた予測係数を解とし、そうでなければ、

$$i=i+1 \quad (31)$$

として、(iv)へ戻る。

音声信号をプリエンファシスしない場合、式(29)は式(7)と同一となり、本提案法は改良雑音補正AR係数推定法と一致する。本提案法によって、プリエン

ファシスされた音声とされない音声の双方に対処することが可能になる。

3.3 提案法 2

改良雑音補正AR係数推定法と提案法1は白色雑音を対象とする雑音補正LPC分析法である。しかし、もし雑音が白色雑音でない場合には、その雑音の影響は自己相関関数 $R(k)$, $k=0 \sim p$ 全体に及んでしまうことが多い。従って、さまざまな雑音環境に対処するため、音声の自己相関関数 $R(k)$, $k=0 \sim p$ から全体的に雑音を引き去る必要がある。

そこで、我々は提案法1を拡張し、新たに提案法2を提案する。本提案法は提案法1と同じ反復アルゴリズムを用いる。しかし、自己相関関数 $R(k)$, $k=0 \sim p$ から全体的に雑音を引き去ることによって、白色雑音以外の多くの雑音にも対応することが可能になる。提案法2の具体的な手順は次のようである。

(i) 雑音信号 $\hat{w}(n)$ の自己相関関数を推定し、その結果を $\bar{R}_n(k)$, $k=0 \sim p$ とする。

(ii) 雑音が混入した音声信号 $\hat{x}(n)$ の自己相関関数 $\bar{R}(k)$, $k=0 \sim p$ を式(27)を用いて計算する。

(iii) アルゴリズムの反復回数 i の初期設定を行う。

$$i=0 \quad (32)$$

(iv) 雑音補正を行う。

$$\bar{R}(k) = \bar{R}(k) - (\beta - \alpha i) \bar{R}_n(k), \quad k=0 \sim p \quad (33)$$

以下は提案法1と同じである。

本提案法は自己相関関数 $R(k)$, $k=0 \sim p$ から全体的に雑音を引くことによって雑音の低減を図っているため、プリエンファシスに影響されにくい。また、白色雑音の場合においても、本提案法は提案法1より多くの改善をもたらすと予想できる。なぜなら、2.2.3でも述べたように、実際の白色雑音は0遅延以外でも自己相関関数の値が完全に0になっていないために、提案法1での雑音低減は不十分と考えられるからである。

4. 実 験

提案法1と提案法2をシミュレーション実験を通して評価する。実験条件、評価方法は2.2と同様である。但し、提案法1は白色雑音のみで評価し、提案法2は白色雑音のほか、帯域制限された白色雑音と周期雑音も用いて評価する。

4.1 提案法 1

表3は白色雑音を用いたときの提案法1の実験結果

表 3 提案法 1 によるケプストラム距離の改善度 (dB) (白色雑音, 左段プリエンファシスなし, 右段プリエンファシスあり)

Table 3 Improvement degree of cepstrum distance by proposal 1 (dB). (white noise. The two value for each SNR are for the cases of not using and using pre-emphasis)

SNR(dB)	-5		0		5		10		15		20	
i	-0.09	-0.09	-0.32	-0.06	-0.30	-0.32	0.09	-0.45	0.09	-0.23	0.28	0.04
e	0.41	0.32	0.95	0.76	1.80	1.57	2.42	2.13	3.00	2.77	2.84	2.77
a	0.85	0.74	1.29	1.07	1.50	1.48	2.05	1.90	2.50	2.35	3.03	2.94
o	0.70	0.60	0.96	0.65	1.19	0.92	1.06	1.10	0.77	1.09	0.57	0.80
u	0.33	0.42	0.35	0.53	0.56	0.58	1.16	0.80	1.38	1.09	1.64	1.57
Average	0.44	0.40	0.65	0.59	0.95	0.85	1.36	1.10	1.55	1.43	1.67	1.62
Data B	0.18	0.33	0.56	0.65	0.97	1.00	1.42	1.41	1.83	1.81	2.08	2.09

表 4 提案法 2 によるケプストラム距離の改善度 (dB) (白色雑音, 左段プリエンファシスなし, 右段プリエンファシスあり)

Table 4 Improvement degree of cepstrum distance by proposal 2 (dB). (white noise. The two value for each SNR are for the cases of not using and using pre-emphasis)

SNR(dB)	-5		0		5		10		15		20	
i	1.31	1.66	1.11	1.58	1.36	1.22	1.74	0.87	1.89	1.07	2.13	1.95
e	1.18	0.93	2.05	1.72	3.12	2.78	3.97	3.74	4.35	4.18	4.10	4.23
a	2.23	2.07	3.01	2.66	3.24	2.88	3.55	3.27	4.03	3.88	4.31	4.21
o	2.24	1.19	2.85	2.43	3.07	2.80	2.98	2.92	2.66	2.60	2.41	2.21
u	1.30	1.51	1.43	1.56	2.02	1.65	2.79	2.17	3.27	2.84	3.49	3.40
average	1.65	1.47	2.09	1.99	2.56	2.27	3.01	2.71	3.24	2.91	3.29	3.20
Data B	1.17	1.05	1.85	1.72	2.55	2.35	3.20	2.97	3.73	3.54	4.00	3.89

である。プリエンファシスをしない場合は、提案法 1 が 2.1 の改良雑音補正 AR 係数推定法と等しくなり、実験結果は同一となる。表 3 の結果を見ると、提案法 1 は改良雑音補正 AR 係数推定法と異なり、プリエンファシスをして、しない場合と同程度の雑音低減をもたらすことができることがわかる。従って、本提案法によって、改良雑音補正 AR 係数推定法のプリエンファシスされた音声に対処できないという問題点が解決され、雑音の混入した音声からの LPC 係数の推定精度を改善することができると言える。

4.2 提案法 2

表 4 は白色雑音を用いたときの実験結果を表している。表 4 から、提案法 2 は提案法 1 よりケプストラム上でかなりの改善をもたらすことがわかる。特に、母音/i/でも大幅に改善される点は注目される。

図 1 と図 2 は提案法 2 によって得られるスペクトルの改善を示している。図 1 は母音/a/ (SNR = 5 dB)、図 2 は母音/i/ (SNR = 20 dB) の 10 フレームを平均した LPC スペクトルである。プリエンファシス係数は 0.95 であり、用いた雑音はともに白色雑音である。

表 5 は帯域制限された白色雑音を用いたときの実験結果である。白色雑音を音声信号と同様に 3.4 kHz に

帯域制限した後、音声に加えてシミュレーションを行った。実験結果は提案法 2 の有効性を示している。この結果は、提案法 2 が多くの音声信号処理システムに実際に適用可能であることを示唆している。

周期雑音に対しての提案法 2 の有効性を確かめるために、次の周期雑音

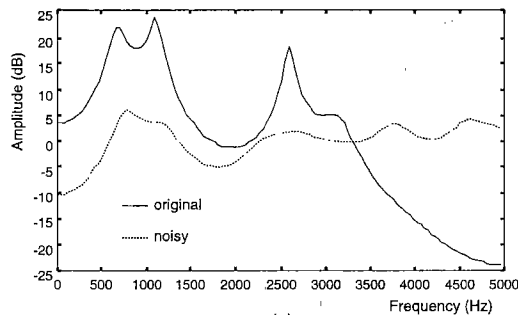
$$u(n) = A \cos(2\pi f_0 n / f_s) \quad (34)$$

ここで f_0 , f_s はそれぞれ

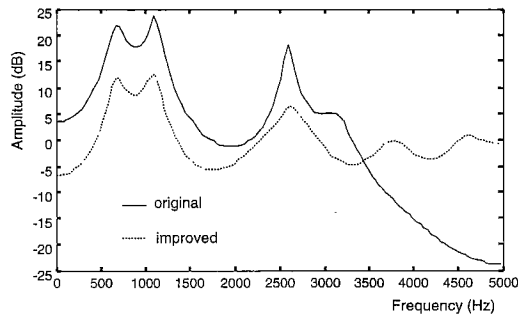
$$f_0 = 3000 \text{ Hz}, \quad f_s = 10000 \text{ Hz} \quad (35)$$

を音声信号に加え、シミュレーションを行った。音声信号の高帯域部分が低帯域部分より雑音に影響されやすい傾向があるので、ここでは周期雑音の周波数を 3000 Hz にした。シミュレーション実験のとき、SNR の値は振幅 A を変えることによって調整された。また、プリエンファシス係数は 0.95 とした。表 6 はケプストラム距離の改善度をまとめている。図 3 は SNR = 10 dB のときに得られたスペクトルの比較を表している。表 6 と図 3 の結果から提案法 2 が周期雑音にも十分に対応できることがわかる。

以上の実験結果からわかるように、プリエンファシスをする場合としない場合の双方において提案法 2 はほぼ同程度の改善を与える。また、提案法 2 により LPC 分析はさまざまな雑音に対処することが可能に

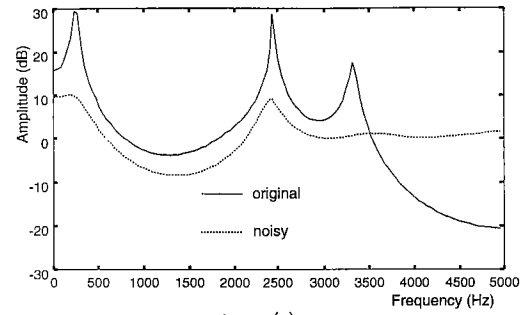


(a)

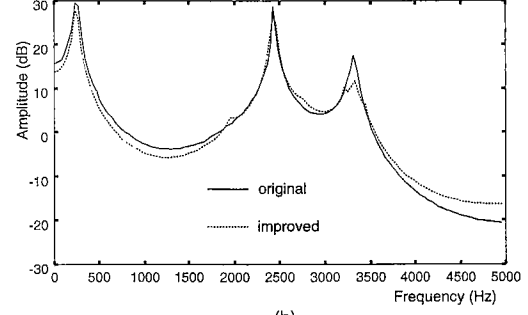


(b)

図 1 母音/a/のLPCスペクトルの比較(a)改善前(b)提案法2による改善後 (SNR=5 dB, 白色雑音)
Fig.1 Comparison of LPC spectra for Japanese vowel /a/. (a) before improvement (b) after improvement by proposal 2 (SNR=5 dB, white noise)



(a)



(b)

図 2 母音/i/のLPCスペクトルの比較(a)改善前(b)提案法2による改善後 (SNR=20dB, 白色雑音)
Fig.2 Comparison of LPC spectra for Japanese vowel /i/. (a) before improvement (b) after improvement by proposal 2 (SNR=20 dB, white noise)

表 5 提案法2によるケプストラム距離の改善度 (dB) (帯域制限された白色雑音, 左段プリアンファシスなし, 右段プリアンファシスあり)

Table 5 Improvement degree of cepstrum distance by proposal 2 (dB). (low-pass filtered white noise. The two value for each SNR are for the cases of not using and using pre-emphasis)

SNR(dB)	-5		0		5		10		15		20	
i	3.17	3.05	2.65	2.64	2.68	2.80	3.04	2.98	2.85	2.75	2.31	2.10
e	0.78	0.88	1.53	1.57	2.08	2.02	2.21	2.24	1.86	1.94	1.11	1.20
a	1.07	1.20	2.20	2.06	2.41	2.28	1.75	1.74	0.69	0.57	0.18	0.08
o	1.09	1.11	2.01	1.71	2.98	2.65	3.76	3.59	3.76	3.69	3.32	3.09
u	1.93	2.14	2.58	2.62	3.11	2.83	3.35	3.09	3.20	3.23	3.18	3.19
Average	1.61	1.68	2.19	2.12	2.65	2.52	2.82	2.73	2.47	2.44	2.02	1.93
Data B	0.76	0.87	1.24	1.26	1.60	1.50	1.79	1.64	1.68	1.59	1.36	1.32

なる。

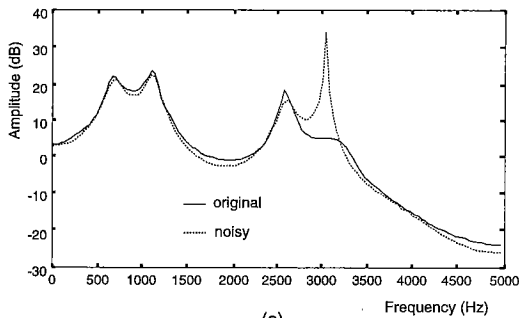
一方, 各方法のケプストラム距離を比較するために, 白色雑音を音声データ B に加えてシミュレーション実験を行った。図 4 はその結果を示している。但し, ここではプリアンファシス係数を 0.95 とし, プリアンファシスを行った。図 4 において, standard method は従来の LPC 分析の自己相関法を指している (雑音補正はない)。また, conventional method

は改良雑音補正 AR 係数推定法を指している。図 4 からは, 提案法 2 が最も優れていることがわかる。より具体的には, 図 4 によると, 提案法 2 が SN 比 10 dB の音声で得られたケプストラム距離は, 従来の LPC 法が SN 比 20 dB の音声で得られたケプストラム距離とほぼ同じ値をもつ。すなわち, 提案法 2 によって音声信号は 10 dB ほど改善されている。一方, 提案法 1 においても SN 比 10 dB の音声信号で 5 dB

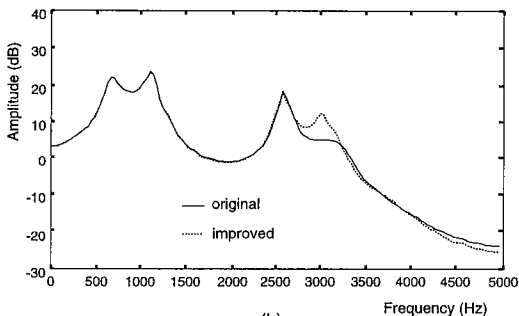
表 6 提案法 2 によるケプストラム距離の改善度 (dB) (周期雑音, 左段プリエンファシスなし, 右段プリエンファシスあり)

Table 6 Improvement degree of cepstrum distance by proposal 2 (dB). (periodic noise. The two value for each SNR are for the cases of not using and using pre-emphasis)

SNR(dB)	-5		0		5		10		15		20	
i	1.85	1.69	1.82	2.38	1.73	1.90	1.60	1.55	2.09	1.86	2.64	2.52
e	1.61	1.64	0.98	1.14	1.53	1.42	1.98	1.83	2.10	2.17	1.19	1.16
a	1.87	1.42	0.77	0.09	0.74	0.39	1.05	0.98	1.29	1.26	0.66	0.67
o	1.36	1.16	1.28	0.97	1.17	0.69	1.45	1.34	1.06	1.09	1.18	0.95
u	3.51	2.65	3.04	2.75	2.76	3.14	2.43	2.56	2.69	2.29	3.41	2.74
Average	2.04	1.71	1.58	1.47	1.59	1.51	1.70	1.65	1.85	1.73	1.82	1.61
Data B	2.07	1.80	1.41	1.23	1.11	0.91	1.05	0.90	0.97	0.91	0.79	0.76



(a)



(b)

図 3 母音/a/のLPCスペクトルの比較(a)改善前(b)提案法 2 による改善後 (SNR=10 dB, 周期雑音)

Fig. 3 Comparison of LPC spectra for Japanese vowel /a/. (a) before improvement (b) after improvement by proposal 2 (SNR=10 dB, periodic noise)

ほどの改善が得られている。

4.3 討 論

(a) 雑音の推定

本論文は雑音の推定問題を扱っていない。なぜなら、実際の場合には、無音区間を利用した推定法などを容易に用いることができるためである。本論文での実験は、雑音を既知と仮定したため、雑音パワーおよび自己相関関数の推定も音声に付加する白色ガウス雑音から直接計算することによって行った。

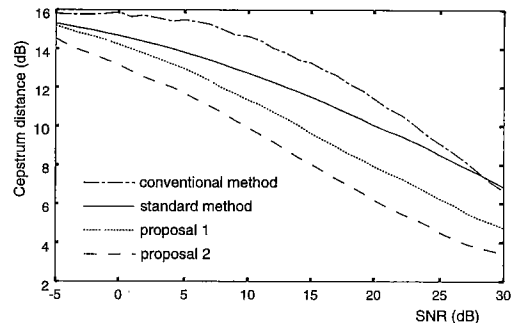


図 4 各方法のケプストラム距離の比較 (白色雑音)

Fig. 4 Comparison of cepstrum distance. (white noise)

(b) 計算量の問題

反復アルゴリズムを用いることにより、提案法には従来のLPC分析法よりも多くの計算量が必要となる。例えば、反復アルゴリズムのパラメータ α と β をそれぞれ0.1と1.2とすると、最悪の場合は12回の反復が必要である。しかし、実際には、ほとんどの場合、12回より少ない回数で反復アルゴリズムは収束する。また、反復されるのはLevinson-Durbin法の再帰計算のみであり、自己相関関数の計算はそこに含まれていない。しかも、自己相関関数の計算量は再帰計算のそれよりはるかに大きい。従って、現在のDSP (Digital Signal Processor) の処理能力を考慮すれば、提案法による計算量の増大はそれほど大きな問題にはならないと予想される。

提案法 2 は雑音低減効果と適用範囲において提案法 1 よりも優れている。しかし、雑音の自己相関関数を多くの遅延量まで計算する必要があるために、計算量が提案法 1 よりも多くなってしまふ。従って、実際に適用する場合に提案法 1 と提案法 2 のどちらを選択するかは、雑音環境と計算量の双方を考慮した上で決定する必要がある。

5. む す び

本論文は雑音の混入した音声信号からのLPC係数の推定問題について検討した。まず従来の改良雑音補正AR係数推定法を音声に適用し、シミュレーション実験を通して改良雑音補正AR係数推定法の問題点、すなわちプリエンファシスされた音声に対処できないという性質を明らかにした。そして、プリエンファシスの影響を詳しく分析し、新しい雑音補正LPC分析方式を提案した。これにより、LPCフィルタの安定性を保ちながら、雑音を低減することができた。更に、この提案法を拡張し、より効果的に雑音を低減できる雑音補正LPC分析法を提案した。本法により、白色雑音だけでなく、周期雑音などにも対処できるようになった。

今後は、本論文での提案法を音声のLPC分析合成システムに応用する予定である。従来雑音に弱いとされるLPC分析合成システムの雑音耐性を提案法によってどれほど改善できるか、視聴試験を通して評価する。

文 献

- [1] J. Markel, "Linear prediction: A tutorial review," Proc. IEEE, vol. 63, pp. 561-580, April 1975.
- [2] M. R. Sambur and N. S. Jayant, "LPC analysis/synthesis from speech inputs containing quantizing noise or additive white noise," IEEE Trans. Acoust. Speech and Signal Process., vol. ASSP-24, no. 6, pp. 488-494, Dec. 1976.
- [3] R. P. Ramachandran, M. S. Zilovic, and R. J. Mammone, "A comparative study of robust linear predictive analysis methods with applications to speaker identification," IEEE Trans. Speech and Audio Process., vol. 3, no. 2, pp. 117-125, March 1995.
- [4] J. Tierney, "A study of LPC analysis of speech in additive noise," IEEE Trans. Acoust. Speech and Signal Process., vol. ASSP-28, no. 4, pp. 389-397, Aug. 1980.
- [5] E. Denoel and J-P Solvay, "Linear prediction of speech with a least absolute error criterion," IEEE Trans. Acoust. Speech and Signal Process., vol. ASSP-33, no. 6, pp. 1397-1403, Dec. 1985.
- [6] C. H. Lee, "On robust linear prediction of speech," IEEE Trans. Acoust. Speech and Signal Process., vol. 36, no. 5, pp. 642-650, May 1988.
- [7] L. Marple, "High resolution autoregressive spectrum analysis using noise power cancellation," Proc. IEEE Int. Conf. Acoust. Speech and Signal Process., pp. 345-348, 1978.
- [8] S. M. Kay, "Noise compensation for autoregressive spectral estimates," IEEE Trans. Acoust. Speech and

Signal Process., vol. ASSP-28, no. 3, pp. 292-303, June 1980.

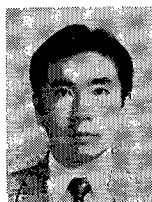
- [9] 島村徹也, 繆 衛国, 鈴木誠史, "データ拡張を利用する2次元スペクトル推定法とその改良," 信学論(A), vol. J78-A, no. 8, pp. 965-976, Aug. 1995.
- [10] S. M. Kay, "Autoregressive spectral estimation: methods," in Modern Spectral Estimation, pp. 217-270, Prentice Hall, New Jersey, USA, 1987.

(平成9年9月29日受付, 10年3月11日再受付)



趙 奇方 (正員)

昭60中国北京郵電大・通信工学・通信工学卒。昭63同大大学院修士課程了。同年中国郵電省経済技術発展研究センター入社。平10埼玉大学大学院理工学研究科博士課程単位取得退学。現在、(株)NTTデータ勤務。音声強調、音声符号化、音声合成の研究に従事。日本音響学会会員。



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。現在に至る。この間、平7ラフバラ大学、平8ペ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。スペクトル解析、適応信号処理に関する研究に従事。計測自動制御学会、日本音響学会、IEEE各会員。



鈴木 誠史 (正員)

昭30東工大・工・電気卒。同年郵政省電波研究所(現、通信総合研究所)入所。以来音声情報処理、音声通信、通信システム等の研究に従事。昭63通信総合研究所長。平1埼玉大・工・教授。平10日本工大教授。現在に至る。工博。日本音響学会、映像情報メディア学会、IEEE、アメリカ音響学会各会員。