窓長漸減型線形予測分析による過渡的音声のホルマント周波数抽出

メタデータ	言語: Japanese
	出版者:
	公開日: 2017-10-02
	キーワード (Ja):
	キーワード (En):
	作成者: 三好, 義昭, 大和, 一晴, 柳田, 益造, 角所, 収
	メールアドレス:
	所属:
URL	http://hdl.handle.net/2297/3350

論　文

窓長漸減型線形予測分析による過渡的音声のホルマント周波数抽出

正　員　三好　義昭[†]　　正　員　大和　一晴[†]
正　員　柳田　益造[††]　　正　員　角所　　収[†††]

Formant Frequency Extraction of Transient Parts of Speech by Linear Prediction with a Diminishing Window

Yoshiaki MIYOSHI[†], Kuzuharu YAMATO[†], Masuzo YANAGIDA[††]
and Osamu KAKUSHO[†††], *Members*

あらまし　近年，音声の分析手法として線形予測分析が広く用いられるているが，通常の線形予測法では分析窓内での定常性が仮定されている．このため音声の重要な音響パラメータであるホルマント周波数の急激な時間的変化を正確に追尾する場合には，分析窓長を数 ms 程度に短くする必要があった．しかしながら，有声音の場合，分析窓長を1ピッチ周期程度以下に短くすると，分析窓と励振点との相対位置の影響を大きく受けるといった問題が生じる．本論文では，通常の線形予測法を用いるが，分析窓長を短かくすることによる悪影響を受けることなく，音声の過渡部のホルマント周波数を精度よく推定できる窓長漸減型線形予測分析について述べる．本分析手法は通常の線形予測分析による極周波数推定値の分析窓長依存性の解析結果に基づき，分析窓の任意の点を固定して窓長を漸減させた一連の分析の結果から，窓長が零になる場合の値を外挿することにより，分析窓長を極端に短くすることなく，音声の過渡部の任意の時点のホルマント周波数を精度よく推定するものである．本手法を合成および自然有声破裂音のホルマント周波数追尾に適用することにより，その有効性が示されている．

1. まえがき

最近，ホルマント周波数の時間的変化に注目した知識工学的手法に基づく連続音声認識システム[(1)]が開発され，かなり高い認識率が得られている．これらのシステムでは，ホルマント周波数の追尾とその記述法が音韻識別のキーポイントとなっているように，ホルマント周波数の時間的変化を正確に追尾することは音声分析の重要な課題の一つである．このホルマント周波数の推定手法として，現在では線形予測分析[(2),(3)]が広く活用されているが，通常の線形予測法は分析窓内での定常性が仮定されているため，ホルマント周波数が急激に変化している音声の過渡部のホルマント周波数を正確に推定するには，分析窓長を数 ms 程度に短くする必要があると言える．しかしながら，有声音の場合，分析窓長を1ピッチ周期程度以下に短くすると，分析窓と励振点との相対位置の影響が生じるため[(4),(5)]，声門閉止区間分析[(6)~(9)]あるいは励振源を考慮した分析[(10),(11)]等が必要となる．しかしながら，声門閉止区間分析ではホルマント周波数の推定値がピッチ周期ごとにしか得られず，ホルマント周波数の急激な時間的変化を追尾するには時間分解能が不十分となり易い，また，励振源を考慮した分析では励振源パラメータの推定が不適切な場合には，ホルマント周波数の推定精度がかえって悪くなると言った問題がある．励振源の影響を軽減するために荷重を導入した線形予測モデル[(12)]も提案されているが，それを用いた過渡音に対する小区間分析の効果はまだ検討段階である．一方，分析窓内での非定常性を考慮した線形予測法も検討されてはいるが，予測係数の時間的変化の近似が必要であったり[(13),(14)]，時間的連続性が悪い[(15)]など，最適な近似空間の適否あるいは時間的追従性の問題がある．推定パラメータの時間的連続性に関しては改善が試みられているが[(16)]，一般に非定常性を考慮した分析手法は処理手順が複雑になるといった問題もあり，今後の研究課題であると言える．

† 姫路工業大学電子工学科，姫路市
Fuculty of Engineerring, Himeji Institute of Technology, Himeji-thi, 671-22 Japan
†† 郵政省通信総合研究所，小金井市
Communications Research Laboratory, Koganei-shi, 184 Japan
††† 大阪大学産業科学研究所，茨木市
Institute of Scientific and Industrial Research, Osaka University, Ibaraki-shi, 567 Japan

本論文では，通常の線形予測分析を用いて，分析窓の任意の点（始点，中心等）を固定し，窓長を徐々に短くしていった一連の分析結果に基づき，分析窓長が零になる場合の値を外挿すれば，分析窓長を極端に短くすることによる弊害を受けることなく，過渡部の任意の時点のホルマント周波数が安定に精度よく推定できることを示す．以下，2.において，窓長漸減型線形予測分析の概略を示し，3.において，その理論的基礎として線形予測分析による極周波数推定値の分析窓長依存性を解析的に考察する．4.において，合成音のシミュレーションにより本手法のホルマント周波数推定精度の改善度合を示し，5.では，本手法を実際に自然音声裂音のホルマント周波数軌跡推定に適用して，その有効性を示す．

2. 窓長漸減型線形予測分析

いま，分析窓長を T_a としたとき，分析窓の始端から γT_a（但し，$0 \leqq \gamma \leqq 1$）の分析窓中の時点（端点を含む）を音声の特徴パラメータ（ホルマント周波数等）の瞬時的な値を推定しようとする時点 t_0 に一致させ，γ を一定値に保ったまま，窓長 T_a を漸減した一連の分析窓を設定し，それらの窓に対応する各々の線形予測分析の結果から，分析窓長が零になる場合の特徴パラメータの値を外挿推定する手法を窓長減漸型線形予測分析と名付ける（図1参照）．勿論，γ は $0 \leqq \gamma \leqq 1$ の任意の値を取り得るが，窓長漸減型線形予測分析の典型としては，$\gamma=0$ とした始点固定型，$\gamma=0.5$ とした中心固定型，$\gamma=1$ とした終点固定型などが考えられる．

窓長減漸型線形予測分析は分析窓長を極端に短くすることなく，窓長が零になる場合の値を推定する手法であるため，分析窓長を実際を1ピッチ周期程度以下に短くした場合に生じる難点を避けることのできる分析手法と言える．したがって，本手法は音声の過渡部のホルマント周波数推定に，特に有用であると考えられるので，この点について詳細に検討した結果を以下に述べる．なお，始点固定型，中心固定型および終点固定型はそれぞれ語頭，語中および語尾における音声の特徴パラメータ推定に有用と考えられるが，終点固定型は基本的には始点固定型で時間軸を反転したものと言えるので，本論文では，始点固定型と中心固定型に関して検討した結果を述べる．

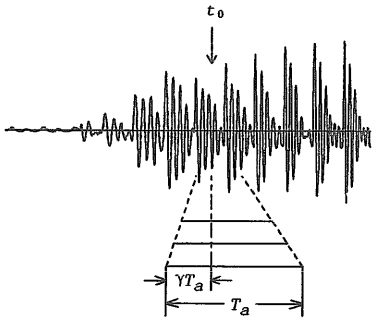

図1 窓長漸減型線形予測分析
Fig. 1 A linear prediction analysis with a diminishing window.

3. 線形予測分析による極周波数推定値の分析窓長依存性

本手法の基本的な特性を明らかにするため，振幅および周波数が時間と共に線形に変化している式(1)の過渡モデル音を用いて，通常の線形予測分析による極周波数推定値の分析窓長依存性の解析的検討を行う．

$$s(t)=\{1+(t-t_0)\Delta A\}\sin\{\omega_0 t+(t-t_0)^2\Delta\omega/2-\omega_0 t_0\} \tag{1}$$

但し，ΔA および $\Delta\omega$ はそれぞれ振幅および周波数の時間的変化率である．

通常の線形予測法では，周知のように，信号波の自己関係数に基づく正規方程式の解を係数とする高次方程式の根から極周波数が推定される．したがって，線形予測法による極周波数推定値の分析窓長依存性を解析するために，信号 $s(t)$ の任意の時刻 $t=t_s$ から $t=t_s+T_a$ までの区間における自己相関係数 $R(\tau, t_s, T_a)$ を式(2)のように定義する．

$$R(\tau, t_s, T_a)=\int_{t_s}^{t_s+T_a} s(t)s(t+\tau)dt \Big/ \int_{t_s}^{t_s+T_a} s^2(t)dt \tag{2}$$

いま，式(1)で表せる過渡モデル音を窓長 T_a，分析次数 $p=2$ で線形予測分析する場合を考える．

3.1 始点固定型

始点固定型は図1において $\gamma=0$ として分析窓の始点を特徴パラメータ推定時点 t_0 に固定し，分析窓長 T_a を変化させるので，この場合の自己相関係数は式(2)において $t_s=t_0$ とし，これに式(1)を代入すれば，

$$R(\tau, t_0, T_a)=\frac{\sin(\Delta\omega\tau T_a/2)}{\Delta\omega\tau T_a/2}\cos(\omega_0\tau+\Delta\omega\tau T_a/2+\phi_1) \tag{3}$$

但し，$\phi_1=\tan^{-1}\dfrac{2\Delta A}{(1+\Delta A T_a)\Delta\omega\tau}\{1-\Delta\omega\tau T_a/2\times\cot(\Delta\omega\tau T_a/2)\}$

となる（付録参照）．

一方，分析次数 $p=2$ の線形予測法による極周波数推定値 $\hat{F}$ は

$$\hat{F}=\frac{1}{2\pi T}\cos^{-1}\frac{-a_1}{2\sqrt{a_2}} \qquad (4)$$

但し，

$$a_1=\frac{r_1(r_2-r_0)}{r_0^2-r_1^2} \qquad a_2=\frac{r_1^2-r_0r_2}{r_0^2-r_1^2}$$

r_i：遅延 iT の自己相関係数

T：標本化周期

となる．そして，信号 $s(t)$ の振幅の時間的変化率 ΔA が比較的小さい場合には，Z 平面上の根はほぼ単位円上付近にあると言えるので，$a_2 \fallingdotseq 1$ となる．この場合，式(4)より，

$$\hat{F}=\frac{1}{2\pi T}\cos^{-1}(r_1/r_0) \qquad (5)$$

となる．したがって，式(1)の信号 $s(t)$ を分析次数 $p=2$，分析窓長 T_a の線形予測分析を行って得られる極周波数推定値 $\hat{F}$ は，式(3)，(5)より，

$$\hat{F}=\frac{1}{2\pi T}\cos^{-1}\left\{\frac{\sin(\Delta\omega TT_a/2)}{\Delta\omega TT_a/2}\times\cos(\omega_0 T+\Delta\omega TT_a/2+\phi_1)\right\} \qquad (6)$$

但し，$\phi_1=\tan^{-1}\dfrac{2\Delta A}{(1+\Delta A\,T_a)\Delta\omega T}\{1-\Delta\omega TT_a/2\cot(\Delta\omega TT_a/2)\}$

となる．ところで，$\Delta F=\Delta\omega/2\pi=10\text{Hz/ms}$（音声の過渡部では一般にこのオーダでホルマント周波数が変化していると言える），$T_a=20$ ms，$T=0.1$ ms の場合，$\Delta\omega TT_a/2=0.02\pi$ であるので $\sin(\Delta\omega TT_a/2)/(\Delta\omega TT_a/2)\fallingdotseq 1$ と近似できる．したがって，式(6)より，

$$\hat{F}=F_0+\Delta FT_a/2+\phi_1/2\pi T \qquad (7)$$

となる．すなわち，式(1)のように振幅および周波数が時間と共に線形に変化している過渡モデル音 $s(t)$ の場合，式(7)より，分析窓の始点を $t=t_0$ に固定した通常の線形予測分析による周波数推定値は「分析窓の中心位置での過渡モデル音の瞬時周波数と $\phi_1/2\pi T$ の和」となり，信号 $s(t)$ の振幅が変化せず（$\Delta A=0$），周波数のみが線形に変化する場合は，$\phi_1=0$ となるので，周波数推定値は分析窓の中心位置での瞬時周波数となることがわかる．

3.2 中心固定型

中心固定型は図1において，$\gamma=0.5$ として分析窓の中心を特徴パラメータ推定時点 t_0 に固定し，分析窓長 T_a を変化させるので，この場合の自己相関係数は式(2)において $t_s=t_0-T_a/2$ とし，これに式(1)を代入すれば，前節と同様の導出過程より

$$R(\tau,t_0,T_a)=\frac{\sin(\Delta\omega\tau T_a/2)}{\Delta\omega\tau T_a/2}\cos(\omega_0\tau+\phi_2) \qquad (8)$$

但し，$\phi_2=\tan^{-1}\dfrac{2\Delta A}{\Delta\omega\tau}\{1-\Delta\omega\tau T_a/2\cot(\Delta\omega\tau T_a/2)\}$

となる．

したがって，式(1)の信号 $s(t)$ を分析次数 $p=2$，分析窓長 T_a の線形予測分析を行って得られる極周波数推定値 $\hat{F}$ は中心固定型の場合，式(5)，(8)より，

$$\hat{F}=\frac{1}{2\pi T}\cos^{-1}\left\{\frac{\sin(\Delta\omega TT_a/2)}{\Delta\omega TT_a/2}\cos(\omega_0 T+\phi_2)\right\} \qquad (9)$$

但し，$\phi_2=\tan^{-1}\dfrac{2\Delta A}{\Delta\omega T}\{1-\Delta\omega TT_a/2\cot(\Delta\omega TT_a/2)\}$

となる．ここで，前節の式(6)から式(7)を導出したのと同様の近似を行えば，式(9)より，

$$\hat{F}=F_0+\phi_2/2\pi T \qquad (10)$$

となる．すなわち，式(1)の過渡モデル音 $s(t)$ の場合，式(10)より，分析窓の中心を $t=t_0$ に固定した通常の線形予測分析による周波数推定値は「$t=t_0$ での過渡モデル音の瞬時周波数と $\phi_2/2\pi T$ の和」となり，信号 $s(t)$ の振幅が変化せず（$\Delta A=0$），周波数のみが線形に変化する場合は，$\phi_2=0$ となるので，周波数推定値は周波数の時間的変化率にかかわらず $t=t_0$ での瞬時周波数 F_0 と一致することがわかる．

3.3 数値計算例

式(1)に示す過渡モデル音の時刻 $t=t_0$ における周波数推定値の分析窓長 T_a の依存性を図2に示す．但し，標本化周期 $T=0.1$ ms，$F_0=\omega_0/2\pi=1000$ Hz，ΔF

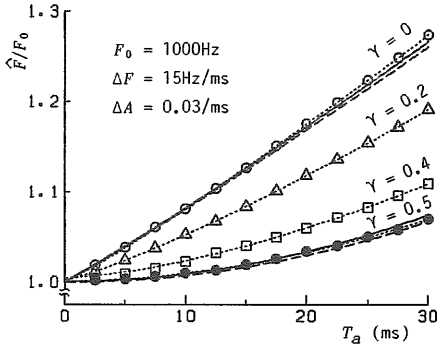

図2 周波数推定値 $\hat{F}$ の分析窓長 T_a 依存性—過渡モデル音—

Fig. 2 The T_a (window length) —dependency of frequency estimation for a time varying signal.

$=\Delta\omega/2\pi=15$ Hz/ms, $\Delta A=0.03$/ms とし，図中の○，△，□および●印はそれぞれ $\gamma=0, 0.2, 0.4$ および 0.5 とし分析窓中の γT_a の時点を $t=t_0$ に一致させ通常の線形予測分析（分析次数 $p=2$）を行って得られた値，また図中の○印近傍の実線は式(6)，破線は式(7)による計算値，そして●印近傍の実線は式(9)，破線は式(10)による計算値である（図中の点線は後述）．

図2より，式(1)に示す過渡モデル音を通常の線形予測分析して得られる周波数推定値は実線すなわち $\gamma=0$（始点固定型）の場合には式(6)による計算値，また $\gamma=0.5$（中心固定型）の場合には式(9)による計算値とほぼ一致し，本解析の妥当性が示されていると言える．また，破線（$\gamma=0$ の場合，式(7)，$\gamma=0.5$ の場合，式(10)による計算値）とも比較的よく一致していると言え，式(7)および式(10)はそれぞれ $\gamma=0$ および $\gamma=0.5$ に対する近似式として十分妥当であると言える，更に重要な特徴は，$\hat{F}/F_0$ は分析窓長 T_a と共に，$\gamma=0$ の場合にほぼ直線的に，また，$\gamma=0.5$ の場合には $T_a=0$ に対称軸を持つほぼ2次曲線的に変化し，$0<\gamma<0.5$ の場合にはその中間となっていることである．すなわち，分析窓中の γT_a の時点を固定にした窓長の異なる線形予測分析の結果を

$$f(T_a)=a(T_a)^{\zeta}+b \qquad (11)$$

但し，$\zeta_1 \leqq \zeta \leqq \zeta_2$

なる関数で最小自乗近似し，$T_a \to 0$ms の値を外挿すれば，非定常な場合でも分析窓長を極端に短くすることなく，通常の線形予測分析で正確な極周波数推定が可能になると期待できる（この関数において $\zeta_1=1, \zeta_2=2$ とし，ζ を 0.01 の精度で $T_a=10\sim30$ ms の分析結果を最小自乗近似した場合，図中の各点線となり，各点線は $T_a=0$ ms に於いていずれも $\hat{F}/F_0 \fallingdotseq 1$ となる）．なお，いまの場合，分析窓長を十分短く $T_a=2.5$ ms とすれば，ほぼ正確な推定値が得られているが，これは本過渡モデル音が励振源のないいわゆる AM・FM 音のため励振源の影響がないからである．

3.4 ζの範囲

式(11)による最小自乗近似は ζ に関して非線形なためでここでは 0.01 の精度で最適な ζ を求めたが，ζ_1 および ζ_2 を γ の関数とすることにより計算量を大幅に軽減することができる．すなわち，前節の結果より，基本的には γ のいかんにかかわらず $\zeta_1=1, \zeta_2=2$ とすれば十分であるが，図2から明らかなように，$\gamma=0$ および $\gamma=0.5$ 付近の最適な ζ はそれぞれ2および1程度になることはないと言える．図3に，$T_a=10\sim30$ ms の分析結果を最小自乗近似して得られる $T_a \to 0$ ms の外挿値が±1％以内の誤差となる ζ の範囲を γ の関数として示す．但し，$F_0=500\sim3000$ Hz, $\Delta F=-20\sim20$ Hz/ms, $\Delta A=0\sim0.3$/ms の範囲（但し，$F_0=500$ Hz のとき，ΔF の下限は −10Hz/ms）で変化させた計312個の過渡モデル音を用いて ζ のきざみ幅 0.01 の精度で求めた．

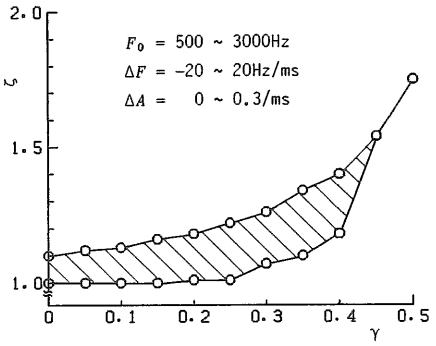

図3 過渡モデル音に対する ζ の適切は範囲（γ の関数として示す）
Fig. 3 Appropriate region of ζ as a function of γ for time varying signals.

図3より，最適な ζ を探索すべき範囲は非常に限られており，特に，$\gamma \geqq 0.45$ のときは最適な ζ を探索する必要がないことがわかる．なお，ΔF および ΔA の変化範囲は実際の音声の過渡部でのホルマント周波数ならびに振幅レベルの時間的変化を考慮したものであり（例えば，30 ms で振幅レベルが1～10倍，周波数が0～±600 Hz 変化），以後，式(11)の ζ_1 および ζ_2 として，それぞれ図3に示す領域の下限値および上限値を用いる．

以上の結果は正弦的振動波において振幅ならびに周波数が線形に変化している場合の解析結果であるが，実際の音声の過渡部では数個の極が一般には指数関数的に変化していると考えられる．しかし，このような場合にも同様のことが結論できるかどうかを解析的に導出するのは困難であるため，以下，合成音を用いたシミュレーションによりその検証を行った結果について述べる．

4. 合成音による検証

過渡的音声の代表例と言える有声破裂音を用いて前章の検証を行う．図4に合成有声破裂音/ga/における極周波数推定値の分析窓長 T_a 依存性を示す．但し，合成条件は標本化周波数 10 kHz，励振源：ピッチ周期8

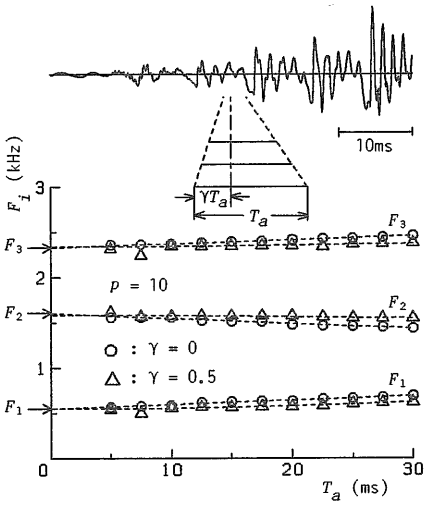

図4　極周波数推定値の分析窓長 T_a 依存性—合成有声破裂音/ga/—

Fig. 4　The T_a (window length) —depencency of pole frequency estimation for a synthetic syllable /ga/.

ms の Rosenberg 波[17](但し，破裂時点から 2 ms 長のノイズバースト付加)，ホルマント周波数：F_1～F_3 は時変（図 6 の実線参照），F_4=3437.5 Hz 一定（但し，図 6 で $t \geqq 20$ ms)，放射特性：6 dB/oct であり，分析は前処理として 1 階差分後，分析窓の始点から γT_a の時点を破裂時点から 17 ms 後の過渡部に固定し（図 4 上段参照)，分析次数 p=10 で通常の線形予備分析を行ったもので，○および△印はそれぞれ γ=0 (始点固定型）および γ=0.5（中心固定型）とした場合の結果である．なお，縦軸左側の矢印は破裂時点から 17 ms 後における合成音のホルマント周波数である．

図 4 より，複数の極が時間と共に指数関数的に変化している場合でも，前章の解析結果と同様，各窓長に対する個々の極周波数推定値は T_a と共に，γ=0 の場合には，ほぼ直線的に，また，γ=0.5 の場合には，ほぼ 2 次曲線的に変化していることがわかる．前章と同様，T_a=10～30 ms の分析結果を式(11)で最小自乗近似した場合，それぞれ図中の点線となり，これらの点線の T_a=0 ms における値はいずれも合成音のホルマント周波数とほぼ等しくなると言える．すなわち，分析窓の始点から γT_a の時点を固定にした分析窓長の異なる通常の線形予測分析結果を式(11)で最小自乗近似したときの $T_a \to 0$ ms における値を求めれば，分析窓長を極端に短くすることなく正確なホルマント周波数が推定できると言える．

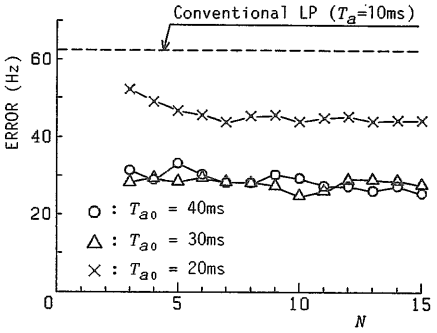

図5　ホルマント周波数推定誤差の外挿データ点数 N 依存性—合成有声破裂音/ga/—

Fig. 5　The N (number of extrapolation points) —dependency of the formant estimation error for synthetic syllable /ga/.

ところで，どのような窓長の分析結果に基づいて $T_a \to 0$ ms における値を外挿するのが適当であるかが問題になるが，一般に分析窓長の最短値が 1 ピッチ周期程度以下に短くなると，分析窓と励振点との相対位置の影響ならびに分析データ点数が少なくなることによる弊害が生じる．このことを考慮して，分析窓長の最短値は 10 ms とし，ホルマント周波数推定誤差の外挿データ点数 N 依存性を図 5 に示す．但し，図 5 は前処理として 1 階差分後，分析次数 p=10 とし，破裂時点から破裂時点後 10 ms までの区間について 2 ms 間隔ごとに外挿して求めた計 6 フレームの第 1 ～第 3 ホルマント周波数推定誤差の平均値で，○，△および×印は漸減窓長の初期値（最長値）T_{a0} をそれぞれ 40, 30 および 20 ms とし，$T_a \to 0$ ms の値を式(11)に基づき最小自乗近似外挿した場合の結果である．なお，γ は $T_a = T_{a0}$ のときに分析窓の始点が破裂時点以前とならないように式(12)により設定した．

$$\gamma = \frac{t_0 - t_b}{T_{a0}} \tag{12}$$

但し，t_0：分析時点，t_b：破裂時点であり，γ>0.5 となる場合には γ=0.5 とする．また，通常の線形予測分析（共分散法）において分析窓長を 10 ms とした場合の誤差を図中破線にて示す．

図 5 より，外挿データ点数 N が 7 点以上であればホルマント周波数推定誤差の外挿データ点数依存性はほとんどなく，また，漸減窓長の初期値が 20～40 ms のいずれでも，通常の線形予測分析において分析窓長を 10 ms (本手法での最短の分析窓長）とした場合よりホルマント周波数推定誤差が改善することがわかる．そして，いまの場合，漸減窓長の初期値 T_{a0}=30 ms で外

挿データ点数が10のとき，すなわち，窓長を30 msから10.2 msまで2.2 msずつ漸減して得られる10個の分析結果を最小自乗近似し，$T_a \to 0$ msにおける値を外挿すれば，ホルマント周波数推定誤差の平均値が62.5 Hzから24.9 Hzに大幅に改善することがわかる．なお，合成有声破裂音/ba/および/da/においても，ホルマント周波数推定誤差の平均値が通常の線形予測分析ではそれぞれ69.4 Hzおよび53.2 Hzであったのが上記と同じ外挿条件でそれぞれ24.7 Hzおよび14.5 Hzに改善した．

合成有声破裂音/ga/のホルマント周波数軌跡推定例を図6に示す．但し，前処理として1階差分後，分析次数p=10，フレーム間隔は2 msとし，○印は本手法による推定値(各分析フレームにおいて，γを式(12)により設定し，窓長を30 msから10.2 msまで2.2 msずつ減少させた線形予測分析を行って得られる極周波数を式(11)で最小自乗近似したときの$T_a \to 0$ msにおける値)，×印は通常の線形予測分析による推定値（分析窓長10 ms，分析窓の中心を分析時点とみなす）であり，実線は合成音のホルマント周波数を示す．

図6より，通常の線形予測分析において，分析窓長を10 msと1ピッチ周期以上にすると，分析位置と励振うとの相対位置関係が原因でホルマント周波数推定誤差が極端に大きくなるようなことは起こらないが(t=28 ms付近の第3ホルマント周波数推定値に若干の影響がみられる)，通常の方法では有声破裂音の相互識別に重要となるホルマントローカス（遷移開始時点）付近の推定誤差が大きいのに対して，本方法の最小の分析窓長は10.2 msといまの場合の通常の線形予測分析の分析窓長とほぼ同じに設定してあるにもかかわらず，本方法の方がより正確なホルマント軌跡が推定

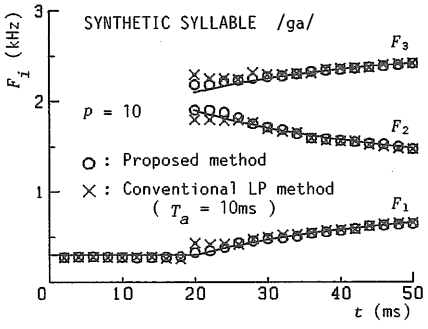

図6 ホルマント周波数軌跡推定の比較—合成有声破裂音/ga/—

Fig. 6 The comparison of formant trajectory extraction for synthetic syllable /ga/.

でき，特に，ホルマントローカス付近の推定誤差が大幅に改善されていると言える．

5. 自然音声への適用例

成人男性が発声した単音節/ga/における極周波数の分析窓長T_a依存性を図7に示す．但し，前処理として1階差分後，分析窓の始点からγT_aの時点を破裂時点から17 ms後の過渡部に固定し(図7上段参照)，通常の線形予測分析(分析次数p=12，窓長：30 msから10.2 msまで2.2 ms間隔で減少)を行ったもので，○および△印はγをそれぞれ0および0.5とした場合の結果である．図中の各点線は窓長を漸減して得られた極周波数を式(11)で最小自乗近似したものである．

図7より，いまの場合，本分析位置付近では第3ホルマントがほぼ定常状態となっているため（図8/ga/の例でt=27 ms付近参照)，第3ホルマント周波数推定値はγならびにT_aにかかわらずほぼ一定値となるが，第1および第2ホルマント周波数推定値のT_a依存性には前章の合成音の場合と同様の特性があると言える．但し，γ=0の第1ホルマント周波数換推定のJT_a依存性が合成音の場合と多少異なるため，$T_a \to 0$ msの外挿値にγ=0とγ=0.5では若干の差が生じている．これは本分析位置でγ=0とした場合，$T_a \geqq 15$ msにおいて第1ホルマントがほぼ定常状態となっている区間が主な分析対象区間となるからである．すなわち，始点固定型（γ=0）は分析位置以降のホルマン

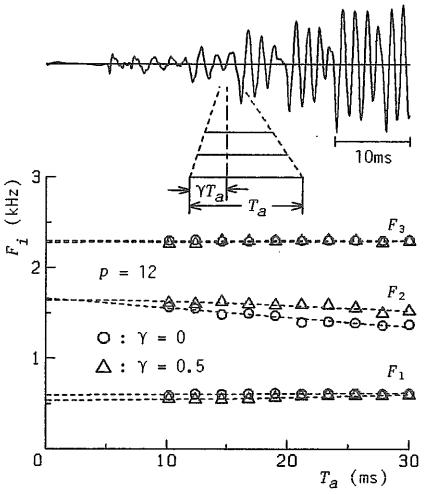

図7 極周波数推定値の分析窓長T_a依存性—自然有声破裂音/ga/—

Fig. 7 The T_a (window length) —dependency of pole frequency estimation for a natural syllable /ga/.

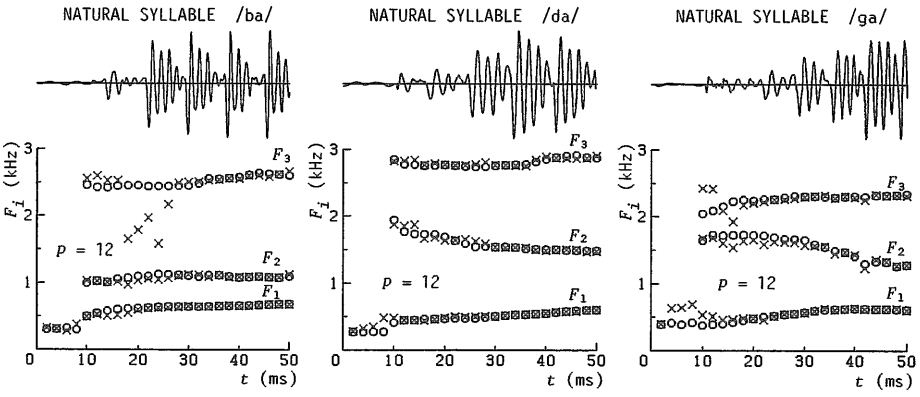

図8 ホルマント周波数軌跡推定の比較—自然有声破裂音/ba/, /da/, /ga/—○印：本方法による推定値 ×印：通常の線形予測分析（T_a=10 ms）による推定値

Fig. 8 A comparison of formant trajectory extraction for natural speech. ○: Proposed method. ×: Conventional LP method (T_a=10 ms).

ト変化のみに基づき分析位置での値を推定するのに対し，中心固定型（γ=0.5）の推定値は分析位置前後のホルマント変化に基づいているため，中心固定型による推定値の方がより信頼性があると考えられる．しかしながら，破裂音の破裂時点のように声道特性が急変する付近を中心固定型で分析すると，分析窓内に声道特性が急変する時点を含む窓と含まない窓が混在し推定値が不安定となるので，このような付近では始点固定型の方が良いと言える．以上のような特徴を考慮してγを設定すれば，自然音声においても，分析窓中の任意の点を固定にした分析窓長の異なる通常の線形予測分析の結果を式(11)により最小自乗近似したときの$T_a \to 0$ msにおける値を用いれば，より正確なホルマント周波数推定が可能であると推察される．

成人男性が発声した有声破裂音/ba/, /da/および/ga/のホルマント周波数軌跡推定例を図8に示す．但し，前処理として一階差分を行い，分析次数p=12，フレーム間隔2 msで分析した結果であり，図中の○印および×印の意味は図6と同じである．但し，実際の有声破裂音では，分析窓長を10 ms程度に短くすると，特に破裂時点付近において第3ホルマント周波数推定値のバラつきが大きくなる場合があるので，本手法では，窓長を漸減して得られる極が窓長を漸減する前の極と比較して，しきい値（いまの場合，±10 %）以内になければ最小自乗近似から除くと共に，γの設定を分析時点が視察により求めた破裂時点から10 msまでは（図8で10 ms≦t≦20 ms）零とし，それ以降は式(12)の分子を$t_0 - t_b - 10$ msとした．なお，破裂時点以前はγ=1.0とし，第1ホルマント周波数のみを推定した．

図8より，本手法によれば，/ba/のt=20 ms付近のように通常の線形予測分析では分析窓長を10 msとすると第3ホルマントが正しく推定できない場合でも妥当な第3ホルマントが推定でき，また/da/および/ga/の結果から明らかなように，有声破裂音の相互識別に重要となるホルマントローカス付近のホルマント軌跡が通常の線形予測分析よりも正確に推定できると言える．

6. むすび

音声の過渡部の一分析手法として，分析窓の任意の点を固定して窓長を漸減させた一連の線形予測分析の結果から，分析窓長が零になる場合の値を外挿する窓長漸減型線形予測分析を提案し，その基本型である分析窓の始点を固定した始点固定型ならびに分析窓の中心を固定した中心固定型について詳細な検討を行った．その結果，分析窓の始点あるいは中心を固定して窓長を漸減させた一連の線形予測分析を行って得られる極周波数推定値は音声の過渡部においては分析窓長と共にそれぞれほぼ直線的あるいは2次曲線的に変化することが過渡モデル音による解析ならびに合成音によるシミュレーションにより明らかとなった．この性質を利用すれば，分析窓長を極端に短くすることなく，分析窓長を零にした場合の値が推定できるため，特に語頭における音声の過渡部の任意の時点のホルマント周波数が分析窓と励振点の相対位置の影響等を受けずに安定かつより正確に推定できることを合成有声破裂音ならびに実際の有声破裂音のホルマント周波数

軌跡推定に適用することにより示した。

なお，本論文では窓長をピッチ周期とは独立に漸減させたが，漸減する各窓長をピッチ周期の整数倍とした場合の検討，また外挿のための特徴量として，今回は線形予測分析により得られる極周波数を用いたが，線形予測係数あるいはLPCケプストラム係数を用いた場合との比較等が今後の課題と言える。

謝辞 本研究に関し有益な御助言を頂いた阪大産研溝口理一郎助教授ならびに御討論頂いた阪大産研電子機器部門の各位に深く感謝する。

文　献

(1) 溝口，田中，福田，辻野，角所："連続音声認識エキスパートシステム—SPREX—"，信学論(D)，**J70-D**，6，pp. 1189-1198(昭62-06)．

(2) 板倉，斉藤："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)．

(3) B. S. Atal and S.L. Hanauer: "Speech analysis and synthesis by linear prediction of the speech wave", J. Acoust. Soc. Amer., 50, pp. 637-655(1971).

(4) 藤崎，佐藤："各種ホルマン周波数抽出方式における短区間分析の時間窓の影響"，音響学会講演論文集，2-2-2(昭49-06)．

(5) 三好，大和，角所："線形予測法による有声音の1ピッチ周期内分析"，信学技報，**EA76**-53(昭52-01)．

(6) 河原，栃内，永田："小区間の線形予測分析とその誤差評価"，日本音響学会誌，**33**，9，pp. 470-479(昭52-09)．

(7) 片桐，松井，牧野，城戸："高ピッチ音声に対する短区間線形予測分析の検討"，信学技報，**EA80**-31(昭55-08)．

(8) 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., **ASSP-22**, pp. 403-415(1974).

(9) K. Steiglitz and B. Dickinson: "The use of time-domain selection for improved linear prediction", IEEE Trans. Acoust., Speech & Signal Process., **ASSP-25**, pp. 34-39(1977).

(10) 深林太計志："線形予測法による音声分析の精度向上"，信学論(A)，**J61-A**，11，pp. 1168-1169(昭53-11)．

(11) M. Ljungqvist，藤崎："線形予測分析にもとづく声帯音源・声道パラメータの同時推定法"，音響学会音声研資，**S85**-21(昭60-06)．

(12) 柳田，角所："重み付き線形予測分析の検討"，音響学会音声研資，S 85-08(昭60-05)．

(13) 中島，鈴木："非定常態音声分析法"，音響学会講演論文集，2-7-2(昭55-05)．

(14) Y. Grenier: "Time-dependent ARMA modeling of non-stationary signals", IEEE Trans. Acoust., Speech & Signal Process., **ASSP-31**, pp. 899-911(1983).

(15) 宮永，三木，永井，羽島："時変ARMAパラメータの適応的同時推定"，信学論(D)，**J64-D**，4，pp. 308-315(昭56-04)．

(16) 芹沢，三木，宮永，永井："時変ARMAモデルに基づく適応的音声分析法"，信学論(A)，**J71-A**，2，pp. 434-442(昭63-02)．

(17) A. E. Rosenberg: "Effect of glottal pulse shape on the quality of natural vowels", J. Acoust. Soc. Amer., **49**, pp. 583-590(1971).

付　録

式(3)の導出

$t-t_0 \to t$ と変数変換を行い $(\Delta A)^2$ の項を省略すれば，

$$\begin{aligned}
&\int_{t_0}^{t_0+T_a} s(t)s(t+\tau)dt \\
&=\frac{1+\Delta A\tau}{2}\left\{\int_0^{T_a}\cos(\Delta\omega t+\Omega)\tau dt-\int_0^{T_a}\cos(\Delta\omega t^2+2\Omega t+\Omega\tau)dt\right\} \\
&\quad+\Delta A\left\{\int_0^{T_a} t\cos(\Delta\omega t+\Omega)\tau dt-\int_0^{T_a} t\cos(\Delta\omega t^2+2\Omega t+\Omega\tau)dt\right\} \\
&=\frac{1}{\Delta\omega\tau}[(1+\Delta A\tau)\sin(\Delta\omega\tau T_a/2)\cdot\cos(\Delta\omega T_a/2+\Omega)\tau+\Delta A\{T_a\sin(\Delta\omega T_a+\Omega)\tau \\
&\quad-\frac{2}{\Delta\omega\tau}\sin(\Delta\omega\tau T_a/2) \\
&\quad\times\sin(\Delta\omega T_a/2+\Omega)\tau-\tau\sin(\Delta\omega T_a^2/2+\Omega T_a) \\
&\quad\times\cos(\Delta\omega T_a^2/2+\Omega(T_a+\tau))\}] \\
&\quad-\frac{1}{2}\sqrt{\frac{\pi}{2\Delta\omega}}\left(1-\frac{2\Delta A}{\Delta\omega}\omega_0\right)[\{C(x_2)-C(x_1)\}\cos\theta \\
&\quad+\{S(x_2)-S(x_1)\}\sin\theta]
\end{aligned}$$

但し，$\Omega=\omega_0+\Delta\omega\tau/2$

$S(\cdot)$：正弦フレネル関数，$C(\cdot)$：余弦フレネル関数

$$x_1=\Omega\sqrt{\frac{2}{\Delta\omega\pi}},\ x_2=x_1+\sqrt{\frac{2\Delta\omega}{\pi}}T_a,$$

$$\theta=\frac{\omega_0^2-(\Delta\omega\tau/2)^2}{\Delta\omega}$$

となる．ところで，$F_0=\omega_0/2\pi=1000$ Hz，$\Delta F=\Delta\omega/2\pi=10$ Hz/ms，$T_a=20$ ms，$\tau=0.1$ ms のとき，$x_1=20.01$，$x_2=24.01$ となるので，フネレル関数の特徴より，$S(x_1)\fallingdotseq S(x_2)\fallingdotseq 0.5$，$C(x_1)\fallingdotseq C(x_2)\fallingdotseq 0.5$ となり，また，$T_a\gg\tau$ より，上式の下から1行目ならびに2行目を省略し，かつ $1+\Delta A\tau\fallingdotseq 1$，$\Omega\tau\fallingdotseq\omega_0\tau$ とすれば，

$$\begin{aligned}
&\int_{t_0}^{t_0+T_a} s(t)s(t+\tau)dt \\
&=\frac{1}{\Delta\omega\tau}\sqrt{(1+\Delta A T_a)^2\sin^2(\Delta\omega\tau T_a/2)}
\end{aligned}$$

$$+(\Delta A)^2\{\frac{2}{\Delta\omega\tau}\sin(\Delta\omega\tau T_a/2)$$
$$-T_a\cos(\Delta\omega\tau T_a/2)\}^2\cos(\omega_0\tau+\Delta\omega\tau T_a/2+\phi)$$

但し，$\phi=\tan^{-1}\frac{2\Delta A}{(1+\Delta A T_a)\Delta\omega\tau}\{1-\Delta\omega\tau T_a/2 \cot(\Delta\omega T_a/2)\}$

となる．そして，平方根中の第1項≫第2項(例，$\Delta F=10$ Hz/ms, $T_a=20$ ms, $\tau=0.1$ ms, $\Delta A=0.01$/ms のとき，第1項 $=5.68\times10^{-3}$, 第2項$=6.92\times10^{-8}$)より，

$$\int_{t_0}^{t_0+T_a}s(t)s(t+\tau)dt=\frac{1+\Delta A T_a}{\Delta\omega\tau}\sin(\Delta\omega\tau T_a/2)\cos(\omega_0\tau+\Delta\omega\tau T_a/2+\phi)$$

となる．同様に

$$\int_{t_0}^{t_0+T_a}s^2(t)dt=\frac{1}{2}(1+\Delta A T_a)T_a$$

となり，これより式(3)が導出される．

(昭和63年3月7日受付，5月27日再受付)

三好　義昭

昭42姫路工大・電気卒．同年同大電子工学科助手．音声の分析および認識，ディジタル信号処理などの研究に従事．日本音響学会会員．

大和　一晴

昭29姫路工大・電気卒．同年同大助手．講師，助教授を経て，昭46電子教授．工博．現在，多値論理，画像処理，音声認識および信頼性に関する研究に従事．電気学会，画像電子学会各会員．

柳田　益造

昭44阪大・工・電子卒．昭46同大学院修士課程了．同年NHK入局．昭53阪大大学院博士課程了．同年阪大産業科学研究所助手．同助教授を経て昭62郵政省電波研究所(現，通信総合研究所)通信技術部音声研究室長．昭53～54年オランダ国立グローニンゲン大学音声研究所客員研究員．聴覚，音声，ディジタル信号処理の研究に従事．工博．日本音響学会，情報処理学会，人工知能学会，IEEE各会員．

角所　収

昭25阪大・工・通信卒．昭32阪大産研勤務．現在，同研究所教授．工博．音声パターン認識，ネットワーク理論，信号処理および知的情報処理システムに関する研究に従事．1983年度Pattern Recognition Society論文賞受賞．日本音響学会，情報処理学会各会員．