原 著

基本周期ゆらぎのスペクトル分析とそのモデル化

遠藤 康男 粕谷 英樹

要 約 : 喉頭に疾患がある人や疾患がない人が発声した持続母音から測定した基本周期系列のゆらぎの特徴を表すために、自己回帰移動平均 (ARMA) モデルを提案する。モデルにおいて、自己回帰の項が系列のスペクトルを特徴づけるために用いられる。一方、移動平均部分は系列のトレンド成分が除去されているため、係数が固定されている。自己回帰パラメータの推定には線形予測分析法を用いる。実験により、求められたパラメータは知覚される声質だけでなく、系列のスペクトルの特徴を良く反映していることを示す。

索引用語 : 嗄声, ゆらぎ, 声質

Spectral Analysis of Fundamental Period Perturbation and its Modeling

Yasuo Endo, Hideki Kasuya

Abstract : An autoregressive moving average (ARMA) model was proposed to represent the perturbation properties of the fundamental period sequence which has been measured from a vowel sustained by subjects with or without laryngeal diseases. In the model, the AR term was introduced to characterize the spectrum of the sequence, whereas the MA part had a fixed coefficient as a result of removing a trend component in the sequence. The linear prediction analysis method was used to estimate the AR parameters. Experiments showed that the estimated parameters reflected well the spectral characteristics of the sequences as well as the perceived quality of the vowel.

Key words : hoarseness, perturbation, perceived quality of the vowel

I. はじめに

これまで基本周期系列のゆらぎの大きさを定量化するためにさまざまなゆらぎパラメータが提案されてきた¹⁾。また、われわれは新しいゆらぎパラメータを提案した²⁾。これらはいわば基本周期系列のゆらぎの性質をマクロな面からみている。一方、基本周期系列のゆらぎの性質をより詳しく、すなわち、よりミクロな面からみるためには、基本周期系列のゆらぎの性質を複数のパラメータで表す必要がある。

そこで、本報告では基本周期系列のゆらぎをモデル化することを考える。まず初めに基本周期のゆらぎのスペクトル分析を行う。本論文では、スペクトル分析の精度を上げるため、トレンド成分を除去した基本周期系列を扱う²⁾。まず実際の基本周期系列の周波数スペクトルの例をあげて、多くの場合ゆらぎのスペクトルが1つの極によって特徴づけられることを示す。また、トレンド除去の効果はスペクトルの直流成分の欠如として現れるので、その効果を表すために、モデルに零点も導入する。以上を総合して、ゆらぎの自己回

帰自己回帰移動平均モデルを提案する。さらに、モデルのパラメータと聴覚的印象との関係について調べる。

II. ゆらぎのスペクトル分析

ゆらぎのスペクトル分析は、平間ら³⁾が典型的な rough の音声に対して行っており、スペクトル上にピークがみられることを示している。本報告では典型的な rough の音声に限らず、正常音声も含めたいろいろな病的音声のゆらぎのスペクトル分析を行った。音声信号から抽出した基本周期系列からトレンド成分を除去し、平均基本周期で規格化する²⁾。そのようにして得られた基本周期系列とその周波数特性の例を図1に示す。図1において(a)が正常音声の例を、(b)が病的音声の例を示している。各図の左側がトレンド成分を除去し、基本周期系列の平均値で規格化した基本周期系列で、横軸が基本周期番号、縦軸が基本周期のトレンド成分からの規格化された偏差を意味する。縦軸の値は規格化された値である。右側がその周波数特性で、横軸が規格化周波数²⁾、縦軸が相対振幅である。正常音声

では基本周期系列のゆらぎがかなり小さく、その変化も比較的ゆっくりしている。このような系列の周波数特性をみると、まず規格化周波数 0.0 すなわち直流に大きな零点を持つのが分かる。これはトレンド成分が除去されたためである。また、規格化周波数 0.04 付近にピークがみられる。これは基本周期系列の 25 サンプルごとの周期性に対応している。一方病的音声では、基本周期系列のゆらぎがかなり大きく、その変化も 2, 3 サンプルごとという非常に速いものである。これに対応して周波数特性も規格化周波数 0.5 に最大ピークをもち、0.33 付近にそれよりやや小さいピークをもつ。このように 2 つのピークをもつのは基本周期系列の非定常性によるものと考えられる。すなわち、周波数特性の 0.5 のピークが基本周期系列の前半部の 2 サンプルごとの周期性に対応し、0.33 のピークが後半部の 3 サンプルごとの周期性に対応している。このような非定常の影響を取り除けば、周波数特性は、多くの場合規格化周波数 0.0 の一つの零点と、ある周波数の一つの極とで特徴づけられる。

図1 トrend除去した基本周期系列とその周波数特性。a：正常音声の例，b：病的音声の例

図2 ゆらぎのモデル

III. ゆらぎの自己回帰移動平均モデル

ゆらぎのスペクトルのモデル化については平間⁴⁾、今泉⁵⁾、小室⁶⁾、小林⁷⁾らの研究があげられる。平間のモデルはゆらぎが正弦波であるという理想的な場合のモデルである。また、今泉は病的音声の特定の聴覚的尺度とゆらぎとの関係を調べるためにモデルを用いている。小室と小林らは正常音声のゆらぎのモデルを提案している、これらに対し、われわれはあらゆる正常音声、病的音声のゆらぎを表せるモデルを考える。

前節のスペクトル分析の結果より、周波数特性は固定した周波数の一つの零点と、ある周波数の極により特徴づけられるので、図2のようなゆらぎの自己回帰移動平均モデルを考える。すなわち、平坦な周波数特性を持つ白色雑音 $w(n)$ をあるシステム関数 $H(z)$ に通すことにより、固定した一つの零点と、可変の極によりゆらぎのスペクトルを特徴づけ、これによってさまざまな周波数特性をもつゆらぎが得られる。ここでシステム関数 $H(z)$ は次式で与えられる。

$$H(z) = \frac{G(1 - z^{-1})}{(1 - a z^{-1})(1 - a^* z^{-1})} \quad (1)$$

ここで $z = e^{j\omega}$ とおくと $H(e^{j\omega})$ はゆらぎの周波数特性を表す。また、 $\frac{\omega}{2\pi}$ が規格化周波数となる⁸⁾。この式は $z = 1$ に固定した零点をもち、 $z = a$ 、 a^* に周波数の等しい一対の極をもつことを意味している。ここで a は複素数で、 a^* がその共役である。また、 G がゆらぎの平均パワーを調整するパラメータである。(1)式はさらに次のように書き換えられる。

$$H(z) = \frac{G(1 - z^{-1})}{1 - a_1 z^{-1} - a_2 z^{-2}} \quad (2)$$

すなわち $H(z)$ は2次の自己回帰、1次の移動平均モデルということになる⁹⁾。 $H(z)$ はさらに、図2のように移動平均モデル(反共振器)と自己回帰モデル(共振器)とに分けられる。ここで移動平均モデルはパラメータが固定されているので、 $x(n)$ を積分した系列を $v(n)$ とすると

$$v(n) = v(n-1) + x(n) \quad (3)$$

図3 ゆらぎのモデルを適用した例。a：正常音声の例、b：病的音声の例A、c：病的音声の例B

これを(2)式に代入すると

$$v(n) = a_1 v(n-1) + a_2 v(n-2) + Gw(n) \quad (4)$$

となり、自己回帰モデルとなる。結局、モデルのパラ

メータを求めるには $v(n)$ に対し線形予測分析を用いれば良いことになる。このようにモデルは自己回帰移動平均であるが分析は自己回帰であることに注意する。

このモデルを実際の基本周期系列に適用した例を図3に示す。図の点線がトレンド除去した基本周期系列の周波数特性で、実線がモデルの周波数特性 ($H(e^{j\omega})$) である。図の横軸が規格化周波数で、縦軸が相対振幅である。(a)が正常音声の例である。モデルが基本周期系列の周波数特性を比較的良く近似しているのが分かる。スペクトル上で、周囲より振幅が相対的に大きくなっている部分が極で、その位置が極周波数である。図では実線のピークの位置が極周波数に対応している。正常音声の場合、規格化周波数 0.04 付近に極が現れている。一方(b)の病的音声の場合規格化周波数 0.5 に極が現れている。このように病的音声の方が正常音声より極周波数が高い。また、(a)と(b)を比較するとモデルのゲイン(周波数特性の平均的なレベル)が病的音声の例において著しく大きい。また、(c)の病的音声の例の場合、極周波数が低い、バンド幅が非常に大きい。

IV. モデルの妥当性の評価

前節で述べたモデルが妥当であるか検討する。提案したモデルは入力として平坦な周波数特性を持つ白色雑音を仮定している。したがって分析により求められる $w(n)$ (以後残差系列と呼ぶ) は、モデルが妥当であれば平坦な周波数特性を持つと考えられる。そこで、実際の音声から得られる残差系列 $w(n)$ のスペクトル平坦性を以下のようにして調べてみる。

音声資料として声門癌、反回神経麻痺、声帯ポリープ患者が手術前、手術後に発声した持続母音/e/(手術前 24 例、手術後 28 例) の 52 例²⁾を用いた。これらの音声から基本周期系列を抽出し、トレンド成分を除去する。これらの系列に対しモデルのパラメータ a_1, a_2 を求める。そして(4)式より $w(n)$ を得る。得られた $w(n)$ より次式のようなスペクトル平坦尺度を求める¹⁰⁾。

$$\sigma = \frac{\left(\prod_{k=1}^N S_w(k) \right)^{\frac{1}{N}}}{\frac{1}{N} \sum_{k=1}^N S_w(k)} \quad (5)$$

ここで $S_w(k)$ が $w(n)$ のパワースペクトルである。この尺度はスペクトルが完全に平坦な場合 1 となり、平坦でなくなるにつれ小さな値をとる。

結果の 1 例を図 4 に示す。上がトレンド除去した基

図4 トrend除去した基本周期系列とその残差系列の周波数特性。a:トレンド除去した基本周期系列。b:残差系列。

本周期系列のスペクトルで、下が対応する残差系列のスペクトルである。残差系列のスペクトルがかなり平坦であるのが分かる。また、平坦尺度の 52 例の平均値は、トレンド除去した基本周期系列が 0.02 であるのに対し、残差系列では 0.53 である。ガウス性白色雑音の平坦性尺度値は、理論的にオイラー定数 γ ($= 0.5772\cdots$) であるので¹¹⁾、0.53 はそれに十分近い値であり、スペクトルが平坦化されているとみなせる。以上より、今回提案したモデルが妥当であるといえる。

V. モデルのパラメータと聴覚的印象との関係

次にモデルのパラメータが聴覚的印象とどのように関係しているのかを調べる。モデルのパラメータとしては前の考察よりゲイン、極周波数、極バンド幅を考える。今回は聴覚的印象として GRBAS 尺度¹²⁾の R に関する評点を用いた。音声資料は前節と同じものを用いた。結果を図 5 に示す。図 5 の (a) はパラメータがゲインと極周波数の場合で、(b) はゲインと極バンド幅の場合である。ここで極周波数と極バンド幅が規格化周波数で表されていることに注意されたい。各図

中の○, △, ×がRに関する評点で, それぞれ0, 0.5, 1~3である。GRBAS尺度は0から3の4段階で評定されるが, 今回用いた音声資料は特別に0, 0.5, 1, 1.5, 2, 2.5, 3の7段階で評定されている。図より, ×印は, ゲインが大きいことが目につくが, roughな音声の音響的特徴は様ではなく, (1)極周波数が0.5と高い, つまり基本周期が一周期ごとに規則的に変動する場合や, (2)バンド幅が大きい, つまり基本周期の変動の規則性が低い場合がある。こうした特徴はARMAのような分析方法を用いることにより明らかになる。また, 図5でゲインが比較的小さいにもかかわらず, Rの評点が1~3である音声資料について, その原因を調べた。結果を図6に示す。(a)が基本周期系列に対するゲイン極バンド幅とRに関する評点との関係で, (b)が振幅系列に対するゲイン極バンド幅とRに関する評点との関係である。図中の■が, Rの評点が1~3であるにもかかわらず基本周期系列に対するゲインが比較的小さいサンプルを示す。これらのサンプルの振幅系列に対するゲインをみると, 比較的大きいことが分かる。よってこれらのサンプルは振幅系列のゆらぎがより大きく聴覚的印象に寄与していると

考えられる。このように, 基本周期系列のゆらぎだけでなく, 振幅系列のゆらぎも聴覚的印象に関係していると考えられるので, 今後は, 振幅系列のゆらぎも考慮した解析を行う必要がある。

VI. 結 論

ゆらぎの性質をより詳しく調べるためにゆらぎの生成モデルを考えた。まず初めにゆらぎのスペクトル分析を行い, ゆらぎのスペクトルが1つの零点, 1つの極によって特徴づけられることを示した。そして, この結果をもとにゆらぎの自己回帰移動平均モデルを考えた。さらに, このモデルを実際の音声に適用し, モデルのパラメータと聴覚的印象との統計的関係について調べた。その結果 rough がゲインにも関係していることが分かった。

今後は本モデルによる合成音声を用いて聴覚実験を行い, 自己回帰パラメータが聴覚的印象にどのように寄与しているかについて詳細に検討したい。

図5 モデルのパラメータとRに関する評点との関係。a:ゲインと極周波数, b:ゲインと極バンド幅

図6 ゲイン, 極バンド幅とRに関する評点との関係。a:基本周期系列, b:振幅系列

音声資料を提供して下さった久留米大学医学部耳鼻咽喉科学教室の平野実教授、日比正史講師に感謝する。また、有益なご意見を頂いた査読者(匿名)に感謝する。なお研究の一部は文部省科学研究費によった。

文 献

- 1) Pinto, N. B., Titze, I. R. : Unification of perturbation measures in speech signals. *J. Acoust. Soc. Am.*, 87 : 1278-1289, 1990.
- 2) 遠藤康男, 粕谷英樹 : 嗄声のゆらぎパラメータについての考察. *音声言語医学*, 34 : 1-6, 1993.
- 3) 平間淳司, 垣田有紀 : ゆらぎのパワースペクトルに着目した典型的な Rough の病的音声 1 例の特徴. *音声言語医学*, 30 : 225-230, 1989.
- 4) 平間淳司, 垣田有紀 : F_0 のゆらぎが病的音声のカテゴリ (rough, voice tremor) に与える影響— F_0 のゆらぎに FM モデルを適用した場合—. *音声言語医学*, 33 : 2-10, 1992.
- 5) Imaizumi, S., Gauffin, J. : Acoustical and perceptual characteristics of pathological voices : rough, creak, fry and diplophonia. *ANN. BULL. RILP*. 25 : 109-119, 1991.
- 6) 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. *音響学会誌*, 47 : 928-934, 1991.
- 7) 小林哲則, 関根英敏 : 合成音声の自然性に対する基本周期の揺らぎの役割. *音響学会誌*, 47 : 539-544, 1991.
- 8) Oppenheim, A. V., Schafer, R., W. : *Discretetime signal processing*. Prentice Hall, New Jersey : 202-289.
- 9) 中溝高好 : 信号解析とシステム同定. コロナ社, 東京, pp. 50-54.
- 10) Markel, J., D., Gray, A., H., Jr. : 音声の線形予測. コロナ社, 東京, pp. 166-168, 1980.
- 11) Gray, A., J., Jr. : Log spectra of Gaussian signals. *J. Acoust. Soc. Am.*, Vol. 55, No. 5 : 1028-1033, 1974.
- 12) 日本音声言語医学会編 : 声の検査法, 医歯薬出版, 東京, pp. 182-195, 1979.

別刷請求先 : 〒 321 栃木県宇都宮市石井町 2753
宇都宮大学工学部
粕谷英樹