

Title	スペクトルピーク追跡モデルを用いたスペクトル予測 追跡に関する研究
Author(s)	坂口, 博文
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1153">http://hdl.handle.net/10119/1153</a>
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

# スペクトルピーク追跡モデルを用いたスペクトル 予測・追跡に関する研究

坂口 伯文

北陸先端科学技術大学院大学 情報科学研究科

1998年2月13日

キーワード： 1次聴覚野スペクトル表現、音韻修復、音脈形成、信号分離.

## 1 はじめに

近年、一連の音刺激を1つのまとまった音脈として捉える Stream Segregation の効果や、その形成機構の一要因とされる周波数変化音追跡機能についての研究 [1,2] が行なわれている。

相川ら [1] はピッチの追跡のメカニズムが2次系の特性で記述されると報告した。しかし、これらのモデルは、ピーク周波数のみの予測・追跡にとどまっている。一方、増田ら [2] は母音の遷移部に雑音が後続する音声の音韻知覚を調べることによって、スペクトル変化の外挿の存在を示唆し、IFIS を用いてスペクトルを外挿する手法で音韻修復を工学的に実現している。

本稿では、Shamma らによって提案された1次聴覚野 (Auditory Cortex 1:A1) モデル [3] を用いてスペクトルを表現し、スペクトルピークを表現する4つのパラメータ (周波数、振幅、帯域幅、対称性) の時間変化を予測・追跡することにより、雑音の重畳した信号から、信号スペクトルを予測・追跡する手法を提案する。

## 2 スペクトルピーク追跡モデル

### 2.1 1次聴覚野 (A1) におけるスペクトル表現

図1のAに示すように、Shamma らによって提案された、1次聴覚野 (A1) におけるスペクトル表現を実現するために、A1における神経細胞群の応答を Gabor 関数 $\psi(\omega)$  で近

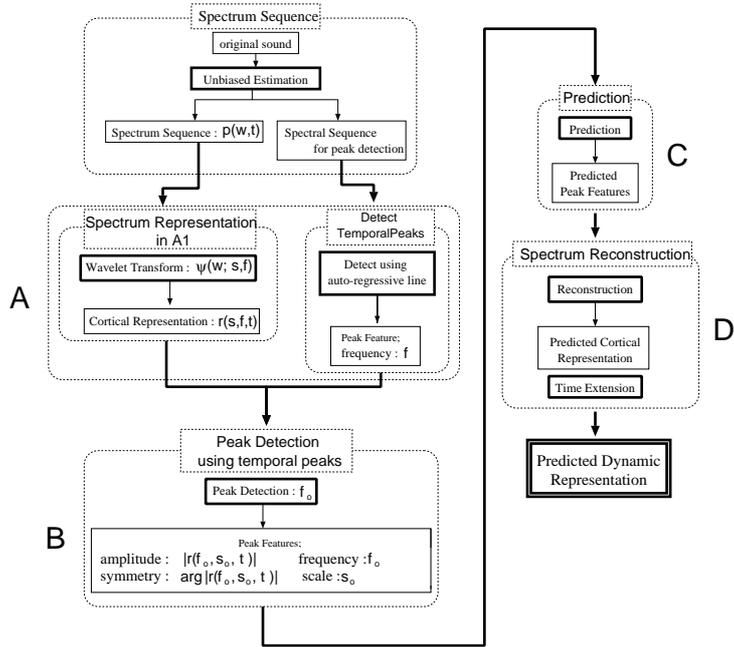


図 1: スペクトルピーク追跡モデル

似し、これを基本 wavelet として入力対数振幅スペクトル  $p(\omega)$  に対する周波数方向への wavelet 変換を行なう。そして、時刻  $t$  における応答  $r(s, f, t)$  を式 (1) で記述する。

$$r(s, f, t) = \frac{1}{|s|^{\frac{1}{2}}} \int_{-\infty}^{\infty} \psi^*\left(\frac{\omega - f}{s}\right) p(\omega, t) d\omega \quad (1)$$

ここで、 $s$ 、 $f$  はそれぞれスケール、A1 での周波数位置を、 $r(s, f, t)$  の絶対値は振幅、位相項は対称性を表す。また、 $*$  は複素共役を表す。なお、入力対数振幅スペクトルには、不偏推定法によるケプストラム平滑化対数スペクトルを用いた。

## 2.2 スペクトルピーク抽出アルゴリズム

図 1 の B で、入力となるスペクトル包絡から、自己回帰直線を用いて、ピーク周波数を求める。そして、これを振幅  $|r(s, f, t)|$  の極大値を与える  $f$  の候補とする。候補  $f$  の近傍で  $|r(s, f, t)|$  の極大値を探索し、極大値となる  $|r(s_0, f_0, t)|$  から帯域幅を与える  $s_0(t)$ 、ピーク周波数  $f_0(t)$ 、ピーク振幅  $|r(s_0, f_0, t)|$ 、対称性  $\arg\{r(s_0, f_0, t)\}$  を決定する。

## 2.3 パラメータの予測法

図 1 の C では、式 (2) のような 2 階微分方程式で表されるシステムによって、4 パラメータ各々についての予測を行なう。

$$a(1 - w)y''(t) + \{b(1 - w)$$

$$+cw\}y'(t) + y(t) = (1 - w)x(t) \quad (2)$$

ここで、 $a$ 、 $b$ 、 $c$  は定数であり  $x(t)$ 、 $y(t)$  は 4 パラメータに対応する入力、出力、 $w$  は重み関数を表す。

### 2.3.1 通常区間での予測機構

雑音のない区間では、 $w = 0$  として中心差分演算により式 (2) を式 (3) のように変形する。

$$y[n] = Gx[n - 1] - \alpha_1 y[n - 1] - \alpha_2 y[n - 2] \quad (3)$$

但し、 $G = \frac{2}{2a+b}$ 、 $\alpha_1 = \frac{2(1-2a)}{2a+b}$ 、 $\alpha_2 = \frac{2a-b}{2a+b}$  であり  $G = 1 + \alpha_1 + \alpha_2$  の関係がある。このように、任意の時刻  $n$  における出力  $y[n]$  はその時刻の 1 時刻前の入力  $x[n - 1]$  と、1 時刻および 2 時刻前の出力  $y[n - 1]$  と  $y[n - 2]$  の線形結合によって求めることができる。

### 2.3.2 雑音区間での予測機構

倉片ら [4] の報告によると、上昇 ( 下降 ) スイープ音の終端に雑音を付加した場合の知覚結果は、雑音がない場合の時よりもより高く ( 低く ) 知覚される。このことから、雑音内では、雑音前の音情報を保持しつつ、追跡を行なっていると思われる。そこで周波数、位相追跡には増 ( 減 ) が一定である変化を記述するために式 (4) を用いて追跡を行なう。

$$y[n] = y[n - 1] + (y[n - 1] - y[n - 2]) \quad (4)$$

また、雑音がある程度の時間続くとピークがぼけることから振幅は減少、帯域幅は大きくなっていくと思われる。そこで、振幅の追跡には  $w = 1$  として後退差分演算をすることによって式 (2) を式 (5) のように変形した式を用いて追跡を行なう。また、帯域幅は振幅  $\times$  帯域幅 = 一定の関係から求めた。

$$y[n] = \frac{c}{c + 1} y[n - 1] \quad (5)$$

## 2.4 複数ピークの追跡

複数ピークが存在する場合、入力をどのピークから選ぶかが問題となる。そこで、入力  $x[n - 1]$  を決定する際、一時刻前の出力  $y[n - 1]$  の最も近傍にある入力  $x[n - 1]$  を選択する。これは Bregman の発見的規則の音脈の形成に関わる規則である、「よい連続の原理や周波数近傍の変理の概念」の考えに従っている。これによって、任意の時刻における複数ピークの追跡を行なう。これによりバウンド効果をシミュレートできる。

## 2.5 スペクトルの復元

図 1 の D では、各ピークごとに予測・追跡された 4 つのパラメータ  $\hat{\hat{s}}_0, \hat{\hat{f}}_0, t$  を統合して逆 wavelet 変換を行なうことによって、スペクトルを再構成する。

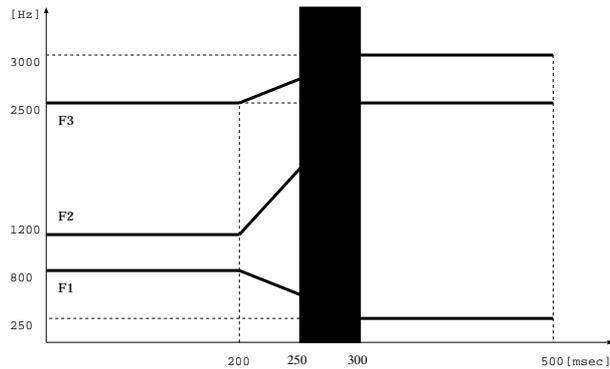


図 2: 入力の概図

### 3 周波数変化予測追跡シミュレーション

- 実験信号...ホルマントシンセサイザを使用して合成した、連続母音 /a/-/i/を用いた。合成音は F1 ~ F3 を持ち、それぞれの帯域幅は 80, 120, 150Hz である。サンプリング周波数は 20kHz、基本周波数は 140Hz とした。/a/-/i/の 100msec の遷移部の中心から終端までを長さ 50msec の白色雑音で置換した。
- 入力スペクトル...フレーム長を 25.6msec、シフト幅を 6.4msec、ケプストラム次数は 60 次までを使用して、不偏推定法より求めた。
- スペクトル予測・追跡...通常区間では 2 次系のシステムとして固有周波数が 20Hz、減衰定数が 1 となるように  $a, b, c$  を設定した。雑音区間では  $c = 0.0064$  とした。

図 2 は合成音声/a/-/i/の遷移部に雑音を付加した場合の略図。図 3 は/a/-/i/の遷移部の中心からその終端まで、長さ 50msec の白色雑音を付加した合成音のスペクトル例。図 4 は図 3 の/a/-/i/を入力とした時のモデルの出力である。スペクトルピークの 4 パラメータの情報を用いて予測追跡を行なうことによって、雑音置換部から信号が予測・追跡されている。

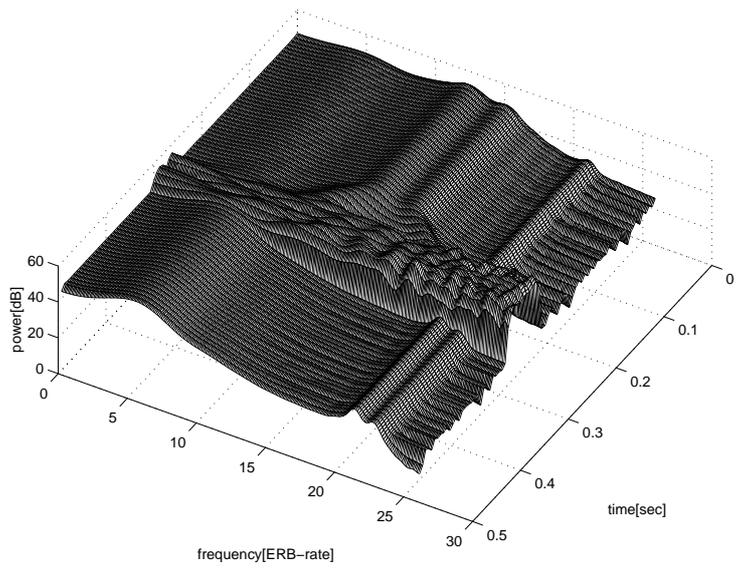


図 3: 入力スペクトル

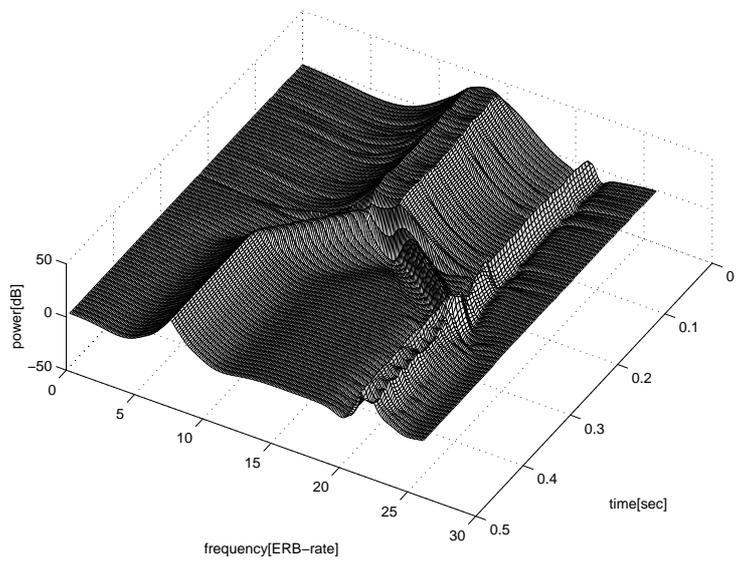


図 4: 出力スペクトル

## 4 まとめ

Shamma らによって提案された 1 次聴覚野モデルを用いて、スペクトルピークを 4 パラメータで表し、それを Bregman によって提唱された発見則と 2 次系線形システムを用いて予測追跡することによって、雑音が重畳した信号から、信号のスペクトルを予測・追跡することができた。

## 参考文献

- [1] 相川, 津崎, 河原, "スイープ音追跡系の動特性, 音響学会聴覚研資, H-95-31, 1-8, 1995.
- [2] 増田, 相川, 津崎, "周波数変化音追跡モデルに基づくスペクトル予測法, 信学技報, vol.96, SP96-2, pp.9-16, 1996.
- [3] K. Wang, S. A. Shamma, "Spectral Shape Analysis in the Central Auditory System, IEEE Trans. Speech Audio Processing, vol.3, no.5, pp.382-395, 1995.
- [4] 倉片, 松井, 西村, "周波数変化音に対する連続聴効果の知覚的軌跡" 聴覚研究会資料, H-94-71, 音楽音響研究会資料, MA94-30, 1994.