

Title	周波数変化音に対する予測追跡の知覚モデルに関する研究
Author(s)	河本, 晃一
Citation	
Issue Date	1997-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1017
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

周波数変化音に対する予測追跡の知覚モデルに関する研究

河本 晃一

北陸先端科学技術大学院大学 情報科学研究科

1997年2月14日

キーワード： スイープ音知覚、周波数変化音追跡、1次聴覚野スペクトル表現.

1 序論

音情報が雑音などによって欠落するような場合、計算機による音声認識は難しい。しかしこのような環境下においても、人間は音韻修復やカクテルパーティ効果などのように欠落した情報を補完し音声によるコミュニケーションを正確に行なうことができる。このような補完現象は、知覚対象とする音を連続した1つの音脈として知覚し、その音脈において欠落している情報を補うことにより生起するものと考えられる。また、音脈の形成において時間的周波数変化の連続性がその要因の1つあり、このことから聴覚には周波数変化を予測し追跡する機能の存在が示唆される。

1.1 スイープ音知覚と音声知覚

周波数変化の知覚に焦点を当てた研究として、相川らにより時間に対して周波数が対数的に変化する対数線形スイープ音の知覚が、湾曲した形状に知覚されることが報告されている。また、途中に周波数停留部を含む段つきスイープ音に対する知覚において、周波数停留部におけるピッチ知覚に揺らぎが生じることがわかっている。

倉片らは途中に雑音を置換した様々な変化パターンのスイープ音の軌跡知覚に関する心理物理実験の結果、雑音呈示中に知覚される音の軌跡は雑音の開始点と終了点の周波数を単純に補間するのではなく、複雑な連続知覚が行なわれており、その1要因として周波数変化を追跡、予測する聴覚機能が存在する可能性が報告されている。

1.2 1次聴覚野におけるスペクトルの表現

1次聴覚野に対する生理学的な知見をもとに、Wang と Shamma[1] は1次聴覚野におけるスペクトルを周波数的な位置、帯域幅、対称性の3つの特徴的な軸によって表す知覚モデルを提案している。このスペクトル表現モデルにおける処理は入力スペクトルに対する多重度解析と見なすことができ、スペクトル包絡の局所的な情報をそれぞれ独立して扱うことが可能であるという利点を持つ。

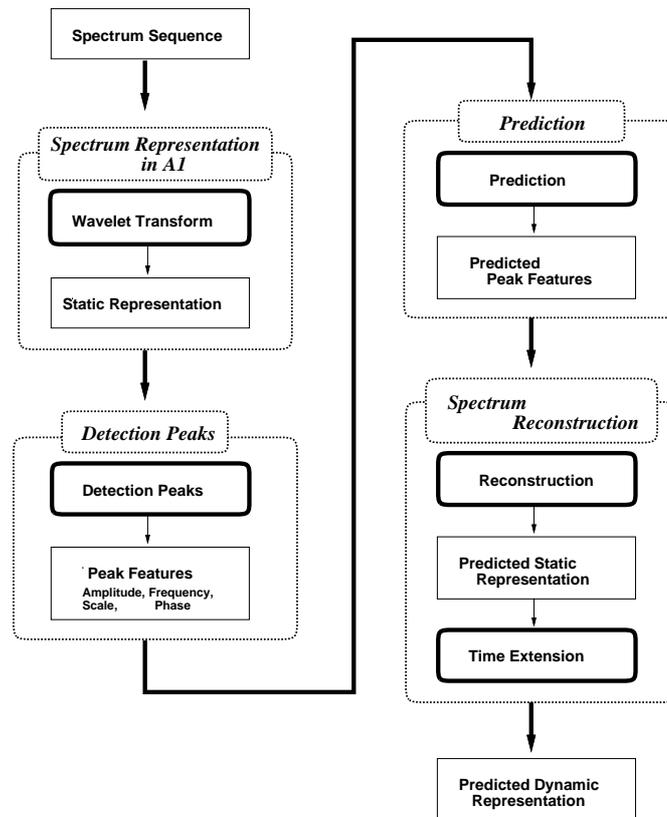


Figure 1: 周波数変化予測追跡モデルのフローチャート

2 周波数変化予測追跡モデル

周波数変化予測追跡モデルの流れを Figure 1 に示す。

本モデルでは、1 次聴覚野スペクトル表現モデルの利点を生かして、予測追跡の対象となるスペクトルピークの 4 パラメータによって表現される情報を、時間的に静的なスペクトル表現から独立して抽出する。それら情報の時間的変化に対して 2 次系のシステムにより実現した機能を作動させることにより、予測追跡を行なう。また、予測追跡したピーク情報を用いてスペクトルを再構成して時間方向へ拡張し、予測追跡を考慮したスペクトルの時系列を求める。

2.1 入力スペクトル

入力スペクトルは入力音響信号より改良ケプストラム法を用いて求めた。なお周波数は、基底膜振動など聴覚末梢系における生理学的及び心理学的知見と対応関係の良い ERBrate に変換し表す。このスペクトル包絡を式 (1) に示す入力 $p(x)$ とする。

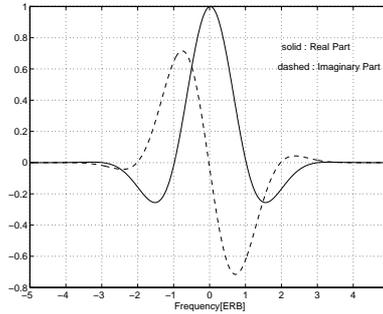


Figure 2: 周波数応答関数の概形

2.2 1次聴覚野におけるスペクトル表現モデル

Shamma らによると1次聴覚野における神経細胞群は特徴周波数付近に局在した周波数応答特性を有し、その周波数応答特性の対称性と帯域幅によって異なった機能を持つ。本モデルでは、この周波数応答特性を側抑制などを考慮して Figure 2に示すような Gabor 関数によって表現する。

入力されたスペクトル包絡と細胞群の周波数応答特性とが一致する場合、その細胞群は強い反応を示すことが生理学的実験からの知見として報告されている。よって、任意の時刻におけるスケール a 、周波数 b における周波数応答 $r(a, b)$ は入力スペクトル $p(x)$ と周波数応答関数 $\psi(x)$ の内積として表現できる。本モデルにおいては、式 (1) のウェーブレット変換として内積を定義する。

$$\begin{aligned} r(a, b) &= \langle p(x), \psi(x; a, b) \rangle_x \\ &= \frac{1}{|a|^{\frac{1}{2}}} \int_{-\infty}^{\infty} \psi^*\left(\frac{x-b}{a}\right) p(x) dx \end{aligned} \quad (1)$$

ここで、 a はスケール変換のためのパラメータ、 b はシフト変換のためのパラメータであり、それぞれスケール方向、周波数方向の位置を表す。また*は復素共役を表す。

式 (1) を用いて入力スペクトルを周波数方向へウェーブレット分析することにより、時間的に静的なスペクトル表現を得る。

また、特徴周波数に近い刺激を与えた場合、神経細胞群は刺激の立ち上がりでのみ強く反応し、その後は時間とともに反応の振幅と帯域幅がともに減衰していく。本モデルでは、この時間応答特性を指数減衰として式 (2) のような関数によって表す。

$$\lambda(t) = \exp\left(\frac{-t}{\tau}\right) \quad (2)$$

ここで、 τ は時定数である。

各細胞群の周波数応答に対して式 (2) に示す時間応答関数を畳み込むことにより、時間的な特性を含んだ時間-周波数応答が得られる。また、任意の時刻における神経細胞群全体の応答は、各細胞群の時間-周波数応答をそれぞれの時刻において加算することにより求められる。

2.3 ピーク情報によるスペクトル表現

本モデルでは、予測追跡に関してスペクトルピークの4パラメータで表現される情報を扱っている。そのピーク情報のみを用いた場合においてもピークおよびその付近の周波数帯域におけるスペクトルの概形を表現可能であることを確認した。

2.4 周波数変化に対する予測追跡機能

本モデルにおける予測追跡機能は、式(3)で表されるような伝達関数を持つ2次系のシステムによって実現した。

$$\begin{aligned} H(z) &= \frac{Gz^{-1}}{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2}} & (3) \\ \alpha_1 &= -2 \exp^{-\zeta\lambda} \cos(\sqrt{1 - \zeta^2} \cdot \lambda) \\ \alpha_2 &= \exp^{-2\zeta\lambda} \\ \lambda &= 2\pi \frac{f_n}{f_s} \\ G &= 1 + \alpha_1 + \alpha_2 \end{aligned}$$

ここで、 α_1, α_2 は線形予測係数、 G はゲイン定数、 f_s はサンプリング周波数を表し、 ζ, f_n はそれぞれ2次系の減衰定数、固有角周波数を表す。

また、追跡対象信号に後続して雑音が付加された場合、対象信号のピーク情報が抽出できないため、雑音入力中は1時刻前の入力の代わりに1時刻前の出力を用いて予測追跡を行なう。

3 周波数変化予測追跡シミュレーション

様々な対数線形スイープ音を対象として、周波数変化予測追跡モデルによるシミュレーションを行なった。段つき対数線形スイープ音と雑音中断のある対数線形スイープ音を対象としたシミュレーション結果を Figure3に示す。なお、それぞれのシミュレーションにおけるパラメータは図中に示す。

Figure 3の(a)、(b)より、段つき線形スイープ音の周波数停留部においてピッチの揺らぎが発生していることを確認できる。これは相川らによる心理物理実験の結果と同様な傾向である。また、固有振動数と時定数を変化させることにより、ピッチの揺らぎの程度や回数を制御することができる。

また、雑音中断のあるスイープ音における雑音中断時間は(c)100msec、(d)200msecである。雑音による中断が(c)のように短い場合は雑音中を内挿するように予測追跡を行なうことが可能である。しかし、(d)のように長い場合は追跡を継続できずに雑音中で中断する。倉片らによる心理物理実験の結果、外挿による追跡は雑音入力開始後100msec程度である。本モデルにおいては、雑音入力以前の情報のみを用いた予測追跡を行なっており、上記の結果は倉片らの実験結果と同等な傾向があるものと考えられる。

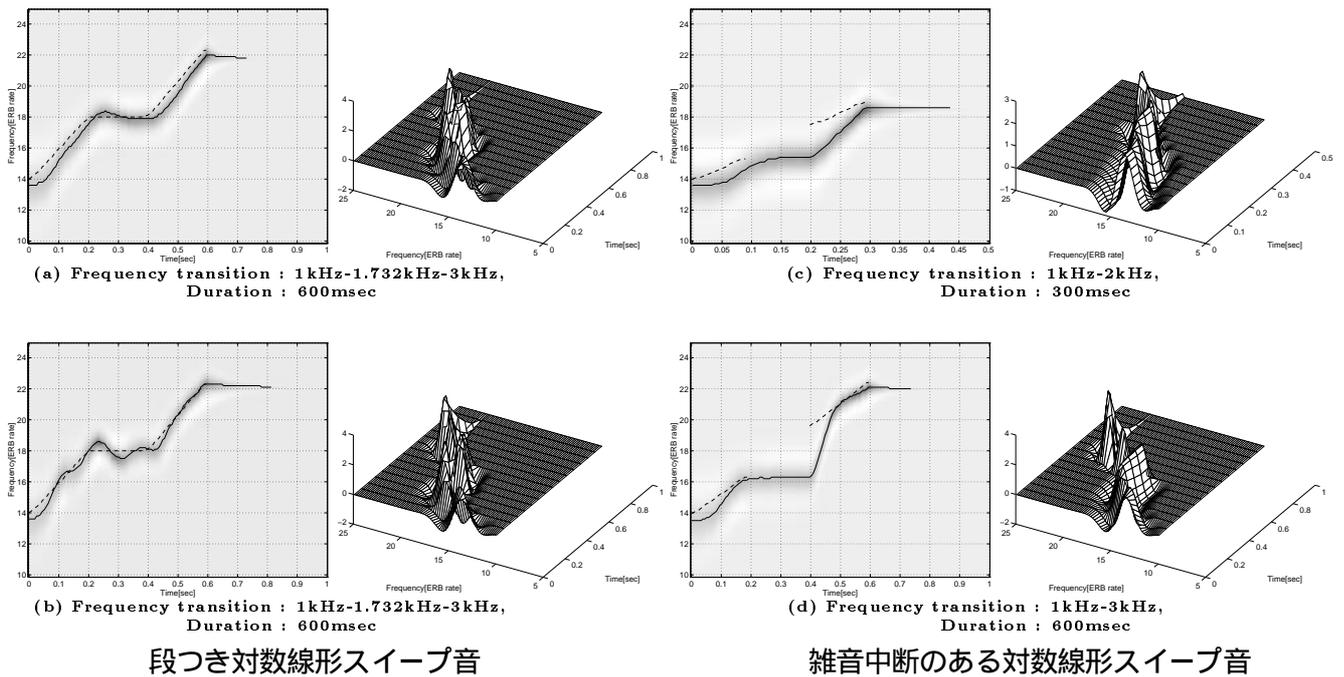


Figure 3: 予測追跡シミュレーション結果

4 結論

本研究では、音韻修復など人間の聴覚における音情報の補完現象の実現を目的として、1次聴覚野におけるスペクトル表現の知覚モデルに基づき周波数変化音に対する予測追跡機能をモデル化した。

本モデルにおいては1次聴覚野におけるスペクトル表現を用いたため、スペクトルピークを振幅、周波数、帯域幅、位相の4パラメータにより表現でき、さらに各ピークの情報よりスペクトルの概形を再現可能である。このような点から、スペクトルにおける情報を圧縮して表現していると言える。また、過去の周波数変化の履歴より以後の軌跡を予想し追跡する機能は、2次系のシステムによって実現した。

モデルを検証するため、様々な対数線形スイープ音を対象として予測追跡シミュレーションを行なった。その結果、モデルのパラメータを変化させることにより、対数スイープ音の湾曲知覚現象や雑音が付加されたスイープ音の補完追跡などの心理物理学実験の結果と類似した結果を得ることができ、人間の知覚の信号処理レベルにおける近似表現が可能であることを確認した。

参考文献

- [1] 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.