

## ポップス系歌唱音声における基本周波数軌跡の楽譜からの変位について

吉田 有里<sup>†</sup>, 森勢 将雅<sup>†</sup>, 高橋 徹<sup>††</sup>, 西村 竜一<sup>††</sup> 入野 俊夫<sup>††</sup> 河原 英紀<sup>††</sup>

<sup>†</sup> 和歌山大学大学院システム工学研究科 <sup>††</sup> 和歌山大学システム工学部

STRAIGHT の基本周波数抽出法を、周期信号用の新しいパワースペクトル推定法である TANDEM を用いて再構成することにより、長時間の歌唱音声を容易に扱うことができるようになった。ここでは、実際のプロ歌手による様々な演奏スタイルの歌唱音声の基本周波数軌跡を求め、演奏スタイルによる物理特性の変化を、規範となる楽譜からの変位として調べた。また、歌唱音声に顕著に認められる大きな周波数遷移での基本周波数推定の問題について検討した。

### F0 trajectory deviations from nominal musical transcription in Pop singing

Yuri YOSHIDA<sup>†</sup> Masanori MORISE<sup>†</sup> Toru TAKAHASHI<sup>††</sup> Ryuichi NISIMURA<sup>††</sup> Toshio IRINO<sup>††</sup>  
Hideki KAWAHARA<sup>††</sup>

<sup>†</sup> Graduate School of Systems Engineering, Wakayama University

<sup>††</sup> Faculty of Systems Engineering, Wakayama University

A reformulation of STRAIGHT F0 extractor based on a new power spectrum estimation method for periodic signals called TANDEM made it practical to extract whole F0 trajectory of a singing voice of an actual performance. This article reports a first attempt for representing effects of singing style in terms of deviations from a nominal musical transcription, using a singing database that consists of various types of singing performance played by professional pop singers. F0 extraction issues caused by fast F0 transitions commonly found in singing voices are also discussed.

## 1 まえがき

デジタル処理技術の進化により、音楽の楽しみ方が変化している。ユーザが、自分の好みに合わせてコンテンツを作成したり加工するという、能動的な音楽鑑賞<sup>12)</sup>を実現するための様々な技術が提案されている。著者らが検討を進めている歌唱デザイン<sup>11)</sup>の転写は、そのような技術の重要な一つである。

ポップスなどのポピュラー音楽には西洋のクラシック音楽に見られるような規範としての楽譜が存在しないことが多い。古典的音楽理論に従って作曲されたものは多くなく、演奏記号も作曲者による明確な指定があることは少ない。ポピュラー音楽における歌手の個性は、先天的な特徴を反映する声質だけではなく、テンポの取り方、アクセントの置き方や音程の揺らし方などといった表現方法にも多くを負うことになる。歌唱デザイン<sup>11)</sup>の転写では、これらをユーザの意図に応じて選択的に操作することが必要となる。

ここでは、歌唱デザイン<sup>11)</sup>の転写のための基礎データ収集の第一歩として、後者の大きな部分を占める歌唱音声の基本周波数軌跡の演奏スタイルによる変化を調べることにした。

## 2 背景

歌唱デザインの転写において、声質と歌い回しは、それぞれスペクトル形状と基本周波数軌跡に密接に関連していることを報告してきた<sup>11)</sup>。この研究手段として用いた音声の部分モーフィング<sup>8)</sup>は、音声分析変換合成システム STRAIGHT<sup>7), 3)</sup>に基づいている。

基本周波数の推定は STRAIGHT の重要な構成要素である。そのために、「基本周波数らしさ」を指標とした方法<sup>7)</sup>、周波数から瞬時周波数への写像の不動点を用いた方法<sup>9)</sup>、複数の手掛かりを用いた方法<sup>6)</sup>が開発され改良が続けられてきた。しかし、これらのいずれの方法も一括処理を前提としており、また大量の計算を必要とする。また、スペクトル包絡の抽出も大量の計算と記憶容量を必要とするものであった。これらの問題により、従来の STRAIGHT では、歌唱デザインの転写研究に必要となる一つの曲の演奏全体の分析が困難であった。

これらの問題は、時間変動の無いスペクトルを計算する新しい方法 TANDEM<sup>10)</sup> の STRAIGHT への導入<sup>4), 5)</sup>により解消されることとなった。この方法に基づく基本周波数抽出<sup>2)</sup>は、逐次計算に適した構成を持つ。本資料は、この新しい基本周波数抽出法に基づく実音声資料の分析の最初の報告

となる。基本周波数抽出法の詳細は、付録に譲る。

### 3 分析対象

#### 3.1 音声素材

分析用の楽曲には、「さきがけ21」の研究の際に片寄により用意された、オリジナルのJ-POPの女声用楽曲「Love Affair 2007」と男声用楽曲「Ride 2007」が用いられた。歌唱デザイン転写<sup>1)</sup>のための研究用に、これらの曲が男性2名女性2名の合計4名のプロ歌手により新たに演奏され、収録された。収録では、演奏方法についての様々な指示が与えられ、複数のバリエーションが制作された。バリエーションとファイル名と歌手の対応を表1、表2に示す。

演奏方法への指示は、3種類に大別される。一つは、人間による物まねが、どのような物理的特徴を制御して行われているかを調べることと目的とするものである。ここでは、歌手自身のオリジナルなスタイルに加え、できるだけ対象歌手に近づけることを狙ったいわゆる「物真似」、対象歌手の歌い回しだけに注目して似せることを狙った「歌真似」、対象歌手の声質だけに注目して似せることを狙った「声真似」の四種類のバリエーションが収録された。

もう一つは、ジャンルによる歌唱スタイルの違いを調べることを目的とするものであり、ここでは、「Pop」「Rock」「唱歌」「演歌」の四種類のバリエーションが収録された。最後は、演奏の表情による違いを調べることを目的とするものであり、ここでは、「喜」「怒」「哀」「楽」「重」「軽」の六種類のバリエーションが収録された。(表では省略)ただし、ジャンルと表情についての収録は一名の男性歌手だけについて行われた。

コンテンツの制作には、ProTools (Digidesign) が用いられた。以下の分析では、歌唱音声を取録した ProTools のデータから歌唱音声のトラックを選択し、wave 形式に変換したものを用いた。ファイルの標準化周波数は 44,100 Hz、量子化ビット数を 16bit とした。「Love Affair 2007」の演奏時間は、4分7秒、「Ride 2007」の演奏時間は、3分13秒であった。同一の伴奏を全てのバリエーションで共通に用いているため、全てのファイルの時間軸も共通となっている。

#### 3.2 楽譜情報の作成

分析結果の比較対象とするために、歌唱音声の旋律から以下の手順により作成された MIDI データを楽譜情報として用いた。まず Love Affair と RIDE のそれぞれに対して、一つの基準となる演奏を選択し、楽譜を起こした。(Love Affair では) この楽譜からシーケンサーを用いて MIDI ファイルを作成し、音声ファイルと MIDI ファイルとを同時に読み込むことのできるシーケンサーを用いて、実際の歌唱と MIDI ファイルとの同期化を行った。具体的には、全体のテンポをできるだけ調整した後、音符の発音時刻を直接操作した。ただし、音

Table 1 File identifier and corresponding singing style for two professional female singers

ファイル名	歌唱スタイル	歌手
001	オリジナル	歌手 A
002	物真似	
003	歌真似	
004	声真似	
005	オリジナル	歌手 B
006	物真似	
007	歌真似	
008	声真似	

Table 2 File identifier and corresponding singing style for two professional male singers

ファイル名	歌唱スタイル	歌手
001	オリジナル	歌手 C
002	物真似	
003	歌真似	
004	声真似	
005	オリジナル	歌手 D
006	物真似	
007	歌真似	
008	声真似	
015	Pop	
016	Rock	
017	唱歌	
018	演歌	

符の終了時刻の判定が困難だったため、持続時間は、最初に作成した MIDI ファイルのものをそのまま用いることとした。こうして作成した同期化された MIDI ファイルを、Matlab により分析するため、miditoolbox<sup>1)</sup> の readmidi 関数を用いて読込んだ。

### 4 基本周波数軌跡の抽出

付録に説明するように、TANDEM-STRAIGHT における基本周波数抽出では、基本周波数の存在する範囲の想定に基づいて複数の検出器を配置する。本研究では、男性の歌唱では 60 Hz から 500 Hz、女性の歌唱では 100 Hz から 700 Hz を想定し、一オクターブに2個の検出器を配置した。基本周波数の抽出は、5 ms 毎に行った。検出器は、各フレーム毎に5個の候補を、それぞれの候補の周期性の指標とともに出力する。この指標を用いて基本周波数を検出するための閾値は、ランダムな信号を周期信号と誤認識する危険率が0.1%以下となるように設定した。なお、大部分のフレームにおいて、基本周波数に相当する候補が卓越した値を示した。各フレームにおいて周期性の指標が最大値となる候補を連続して接続したものを、基本周波数軌跡

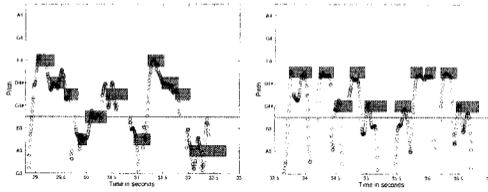


Fig. 1 Nominal note information and extracted F0 trajectory. (Left) Love Affair, (Right) Ride. Both represent initial part of the songs.

とした。

#### 4.1 旋律の特徴の抽出

図1に、基本周波数軌跡と、演奏と同期化された楽譜情報との対応関係の例を示す。このように、演奏から聴き取られる楽譜情報と物理的なパラメータは、一対一ではなく、さまざまなズレ(変位)が含まれている。ここでは、まず、以下のような指標を用いて演奏と物理パラメータとの変位の関係を調べることにした。

##### 4.1.1 音高の変位

Hzを単位として求められている基本周波数 $f_0$ を、次式を用いてMIDI note number  $M$ に変換し、以後の分析で用いた。

$$M = 12 \log_2(f_0/440) + 69 \quad (1)$$

前述の閾値により抽出された基本周波数軌跡の各時刻の $M$ の値のヒストグラムと、楽譜に表れるMIDI note numberのヒストグラムとを求めた。

##### 4.1.2 音程遷移時刻の変位

各音符の開始時刻を基準として、基本周波数軌跡における音程遷移時刻を表すことにした。なお、同じ高さの音符が連続する場合には、基本周波数軌跡から音程遷移時刻を求めることが困難であるため、対象から除外した。こうして選択されたそれぞれの音程遷移について、前の音符に対応する周波数と遷移につづく音符に対応する周波数の幾何平均の周波数を、同じ方向に基本周波数軌跡が横切る時刻を、基本周波数から求めた音程遷移時刻とした。なお、音程遷移の存在を調べる範囲を、遷移の前の音符の中央から、後の音符の中央の時刻までとした。複数の遷移がこの区間に生じた場合には、対象から除外した。

##### 4.1.3 音程遷移速度の変位

前節の手続きで求められた音程遷移時刻の前後の分析フレームにおける基本周波数 $f_0$ の比を音程遷移速度 $v$ とした。単位は、octave/sである。今回のフレームは5ms毎であるので、次式で表される。

$$v = 200 \log_2(f_0(n+1)/f_0(n)) \quad (2)$$

ここで $n$ は、遷移時刻の直前の分析フレーム番号を表す。

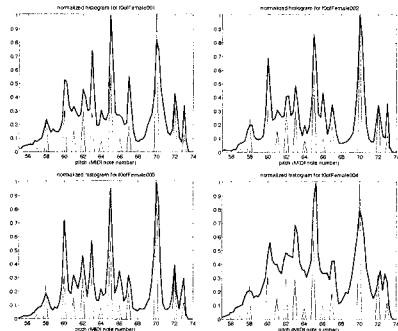


Fig. 2 Normalized histogram of F0 represented in terms of MIDI note number. Singer A. From left to right and top to bottom: original, imitation of a famous singer, imitation only in singing style, imitation only in voice character.

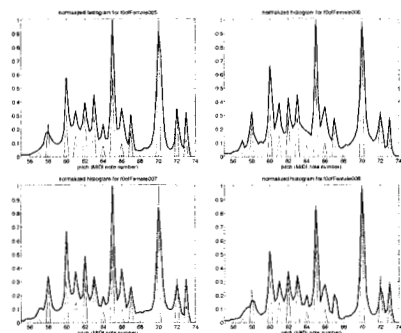


Fig. 3 Normalized histogram of F0 represented in terms of MIDI note number Singer B. Conditions are same as Fig. 2

## 5 分析結果

### 5.1 音高の分布

#### 5.1.1 女性歌手

図2、図3に女性歌手の音高のヒストグラムを示す。左上からZ字の筆順に沿って「オリジナル」「物真似」「歌真似」「声真似」の結果を載せた。図の横軸はMIDI note numberで表した音高である。ヒストグラムのbinの幅は1/4半音とした。

歌手Aのオリジナルの歌い方はクセが強い。この傾向を反映して、規範とする周波数から外れた高さの出現頻度が相対的に高くなっている。声質だけを真似るようにした「声真似」では、このオリジナルと同様の歌い方が行われており、同様に、規範とする周波数から外れた高さの出現頻度が相対的に高くなっている。ここで「真似」の対象とされた有名な歌手は、歌い回しに特徴はあるもの

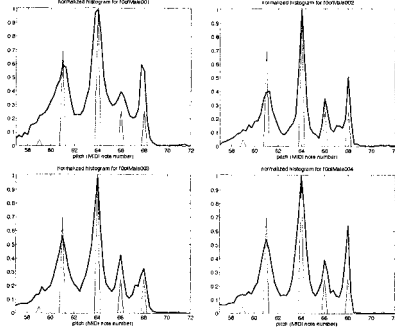


Fig. 4 Normalized histogram of F0 represented in terms of MIDI note number. Singer C. Conditions are same as Fig. 2

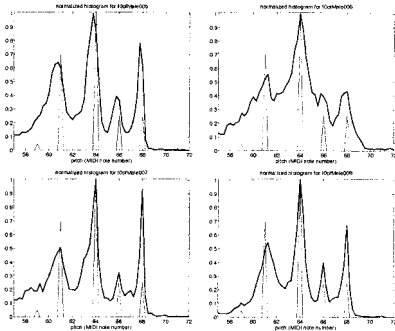


Fig. 5 Normalized histogram of F0 represented in terms of MIDI note number. Singer D. Conditions are same as Fig. 2

の、歌手 A と比較すると歌い回しのクセは少ない。図に示した結果でも、「物真似」と「歌真似」では規範とする周波数から外れた高さの出現頻度が相対的に低くなっている。

歌手 B の場合には、逆の傾向が認められる。歌手 B の歌い方は、「真似」の対象とされた歌手と比較するとクセが少ない。歌手 B 自身のオリジナルの歌い方も、歌手 A と比較すると素直である。ヒストグラムも、この傾向を反映し、「オリジナル」と「声真似」において、規範とする周波数から外れた高さの出現頻度が相対的に低くなっている。

なお、主観的な印象では歌手 B が真似た有名な歌手の方が歌手 A よりもクセが強いように感じられる。分布がこの傾向を直接反映していないように見えるのは興味深い。

### 5.1.2 男性歌手

次に示す図 4、図 5、図 6 は、楽譜と基本周波数の音高の変位について、真似の方法 (図 4、図 5) や音楽ジャンルによる歌唱スタイル (図 6) によ

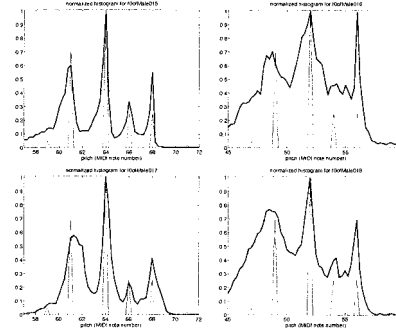


Fig. 6 Normalized histogram of F0 represented in terms of MIDI note number. Singer D. Musical genre for each performance is listed from left to right top to bottom; Pop, Rock, plain, and Enka (an emotionally rich Japanese singing style)

てまとめたものである。

図 4 と図 5 の表示の並べ方は、女性歌手の場合と同じである。歌手 C においては、オリジナルで歌った場合と比較すると、何らかの真似をした方が、ピッチの揺らぎが小さくなる傾向にある。歌手 D においては、「物真似」を行うとピッチの揺らぎが大きくなる傾向にある。実際に「物真似」の歌唱音声を聴いてみると、全体的にファルセットがかかっており、トリルのようなビブラートが多用されている印象を受ける。模倣の対象である有名な歌手の特徴はファルセットとビブラートである。歌手 D は模倣対象である歌手の声質をファルセットととらえたようだが、ビブラートを歌い回しとしてとらえなかったようで、「歌真似」、「声真似」とオリジナルの間には大きな差異は見られなかった。

図 6 では、左上から Z 字の筆順に Pop、Rock、唱歌、演歌のスタイルに対応する分布を示す。Rock と演歌についてはキーを 1 オクターブ下げて演奏された。両方ともに全体的に基本周波数の揺らぎが大きい。Pop と唱歌に関して実際に音声を聞き比べた場合は、唱歌の方が楽譜に忠実に歌っているような印象を受けたが、実際には差異は見られなかった。これらの結果は非常に興味深い。

### 5.2 遷移時刻と遷移速度

遷移時刻の変位と遷移速度の変位については、外れ値が多いため、詳しい統計処理を行っていない。しかし、幾つかの傾向を認めることができる。例として、知覚的には大きな違いがある歌手 B による「オリジナル」と「物真似」の散布図を示す。図 7 は、遷移の前後の音高の差と遷移時刻、図 8 は、同様に音程差と遷移速度の散布図である。

遷移時刻は、図に示されるように広い範囲に値が分布している。演奏での四分音符に対応する時間は、0.1 秒よりもやや少ない。この図からは、基

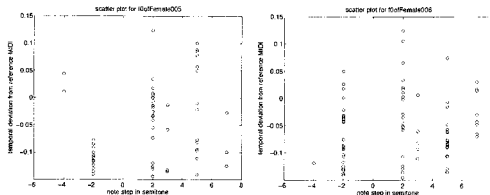


Fig. 7 Transition step vs. temporal deviation.

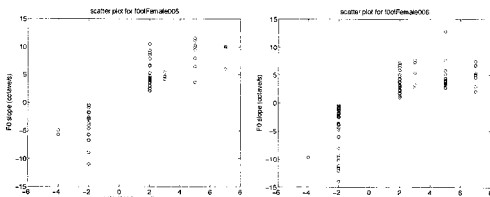


Fig. 8 Transition step vs. transition speed.

本周波数が上昇する遷移の場合には、知覚的に設定された MIDI note の発音開始時刻が、基本周波数の軌跡から求められた遷移時刻と、平均的に一致している傾向が認められる。一方、基本周波数が下降する遷移の場合には、基本周波数の軌跡から求められる遷移時刻は、知覚的に設定された発音開始時刻に、平均的には先行する傾向が認められる。なお、これらの傾向は「オリジナル」と「物真似」に共通して認められており、それらの間の差異は明らかではない。これらの傾向が実際に存在するのか、それらが存在するとしたらどのような知覚的特性と信号の物理的性質によるものかを明らかにする必要がある。

遷移速度の分布では、遷移時刻と比較するとよりはっきりとした傾向が認められる。特に、全音の幅の遷移では、「オリジナル」と比較すると「物真似」ではより遷移速度の絶対値が小さいことが分かる。特に、基本周波数が下降する遷移において、よりはっきりと認められる。これは、聴いた印象とも整合している。

## 6 まとめ

新しい STRAIGHT を用いることで、数分に及ぶ演奏時間を有する実演奏の歌唱音声进行分析することが可能になった。様々な歌唱スタイルの基本周波数軌跡を、MIDI note により表された規範的な演奏情報と比較し、歌唱デザイン転写のための特徴の抽出について検討した。より精密な結果の分析を通じて、有効な特徴とそれらの抽出法を明らかにして行きたい。

## 謝辞

本研究は、科学技術振興機構による戦略的創造研究推進事業のデジタルメディア領域 CrestMuse プロジェクトの支援を得て行われた。本研究で用いた MIDI データを作成し提供して下さった、関西学院大学理工学部情報科学科 4 年の吉村直起氏と北原鉄朗博士に深く感謝いたします。なお、TANDEM-STRAIGHT の開発は、科学技術研究費補助金 (A)19200017 による支援を得て行われた。

## 参考文献

- 1) Tuomas Eerola and Petri Toivola. *MIDI Toolbox: MATLAB Tools for Music Research*. University of Jyväskylä, Jyväskylä, Finland, 2004.
- 2) 河原英紀, 森勢将雅, 高橋徹, 西村竜一, 入野俊夫, 坂野秀樹. Tandem-straight に基づく基本周波数の抽出と評価について. 信学技報 SP2007-26, Vol. 37, No. 6, pp. 395-400, 2007.
- 3) 河原英紀. Vocoder のもう一つの可能性を探る—音声分析変換システム STRAIGHT の背景と展開—. 日本音響学会誌, Vol. 63, No. 8, pp. 442-449, 2007.
- 4) 河原英紀, 森勢将雅, 高橋徹, 西村竜一, 入野俊夫, 坂野秀樹. 時間方向および周波数方向の周期性の影響を除去した周期信号のパワースペクトルの表現について. 信学技報 SP2007-26, Vol. 37, No. 6, pp. 395-400, 2007.
- 5) 河原英紀, 森勢将雅, 高橋徹, 西村竜一, 坂野秀樹, 入野俊夫. 分析位置に依存しない周期信号のパワースペクトル推定法に基づく音声分析変換合成法 STRAIGHT の再構成について. 信号処理シンポジウム, 2007.11. [CD-ROM].
- 6) H. Kawahara, A. de Cheveigné, H. Banno, T. Takahashi, and T. Irino. Nearly defect-free f0 trajectory extraction for expressive speech modifications based on STRAIGHT. In *Inter-speech'05*, pp. 537-540, Lisboa, 2005.
- 7) H. Kawahara, I. Masuda-Katsuse, and A. de Cheveigné. Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction. *Speech Communication*, Vol. 27, No. 3-4, pp. 187-207, 1999.
- 8) H. Kawahara and H. Matsui. Auditory morphing based on an elastic perceptual distance metric in an interference-free time-frequency representation. In *Proc. ICASSP 2003*, Vol. 1, pp. 256-259, Hong Kong, 2003.
- 9) Hideki Kawahara, Haruhiro Katayose, Alain de Cheveigné, and Roy D. Patterson. Fixed point analysis of frequency to instantaneous frequency mapping for accurate estimation of F0 and periodicity. In *Proc. Eurospeech'99*, Vol. 6, pp. 2781-2784, 1999.

- 10) 森勢将雅, 高橋徹, 河原英紀, 入野俊夫. 窓関数による分析時刻の影響を受けにくい周期信号のパワースペクトル推定法. 電子情報通信学会誌, Vol. J90-D, No. 12, pp. 3265–3267, 2007.
- 11) 河原英紀, 生駒太一, 森勢将雅, 高橋徹, 豊田健一, 片寄晴弘. モーフィングに基づく歌唱デザインインタフェースの提案と初期的検討. 情報処理学会論文誌, Vol. 48, No. 12, pp. 3637–3648, 2007.
- 12) 片寄晴弘, 後藤真孝. 音楽のデザイン転写技術の開発にむけて - crestmuse プロジェクトの「価値」創出視点からの紹介 -. 人工知能学会 近未来チャレンジ, pp. 1D1-4, 2006.

## A TANDEM-STRAIGHT による基本周波数の推定

ここでは、実験に用いた基本周波数抽出方法について説明する。

### A.1 スペクトルと包絡計算

基本周期の半分の時間差のある一組の時間窓を用いてパワースペクトルを計算し、それらの平均を求めることにより、時間方向に変動しないパワースペクトル (TANDEM スペクトル) を得ることができる<sup>10)</sup>。さらに、consistency sampling に基づくことにより、周波数軸上における信号の周期性の影響が除去されたパワースペクトル (STRAIGHT スペクトル) を求めることができる<sup>4, 5)</sup>。これらの手続きは、以下のようにまとめられる。

$P(\omega, t)$  を、時刻  $t$  を中心とする窓を用いて計算したパワースペクトルとする。TANDEM スペクトル  $P_T(\omega, t)$  は、以下のように表される。

$$P_T(\omega, t) = (P(\omega, t - T_0/4) + P(\omega, t + T_0/4)) / 2 \quad (3)$$

基本周波数推定に用いる STRAIGHT スペクトル  $P_{TST}(\omega)$  は、実質的に過剰平滑化の影響を無視することで、以下のように簡単化される。

$$C(\omega) = \int_{\omega_L}^{\omega} P_T(\lambda) d\lambda$$

$$P_{TST}(\omega) = [C(\omega + \omega_0/2) - C(\omega - \omega_0/2)] / \omega_0(4)$$

なお、ここでは、分析時刻の記載を省略した。

### A.2 変動スペクトルの計算

TANDEM スペクトルは、STRAIGHT スペクトルに周期性に起因する周波数方向の周期的変動が重畳されたものである。したがって、TANDEM スペクトルを STRAIGHT スペクトルによって正規化することにより、信号の周期性に依存する成分  $P_P(\omega)$  だけを求めることができる。TANDEM-STRAIGHT での基本周波数の抽出<sup>2)</sup> では、この  $P_P(\omega)$  に注目する。

$$P_P(\omega) = \frac{P_T(\omega)}{P_{TST}(\omega)} - 1 \quad (5)$$

時間窓を用いた処理は、周波数軸上での低空間周波数成分通過フィルタに相当し、TANDEM スペクトルを STRAIGHT スペクトルにより正規化する処理は、高空間周波数成分通過フィルタに相当する。したがって、TANDEM-STRAIGHT の単一の設計パラメタである時間窓長を調整することで、目的とする空間周波数成分を選択的に抽出する周期性検出器を作ることができる。ここでは、Blackman 窓を用い、窓長を基本周期の 2.5 倍とし、次のようにして具体的な検出器を作成した。

$$A(\tau; T_C) = \int_{-\infty}^{\infty} w_{\omega_0, N}(\omega) P_C(\omega; T_C) e^{-j\omega\tau} d\omega(6)$$

$$w_{\omega_0, N}(\omega) = c_0 (1 + \cos(\pi\omega/N\omega_0)), \quad (7)$$

ここで、 $c_0$  は、正規化のための定数である。考慮する調波の個数を定める  $N$  は、信号が定常である場合には大きな値とすることで基本周波数の推定精度を向上させる。しかし、基本周波数の変化速度が大きな場合には、誤差が増加する原因となる。ここでは、歌唱での大きな音程遷移やビブラートにおける基本周波数軌跡の正確な追跡を目的として、予備検討により、 $N = 4$  とした。基本周波数の変化速度が 40 oct/s 以下であれば、性能劣化の生じない程度である。

### A.3 複数の検出器の統合

単一の周期性検出器では、仮定した基本周期から大きく外れた周期信号を検出することができない。ここでは、異なった基本周波数を仮定して設計した複数の検出器を統合することにより、目的とする範囲の基本周波数について、基本周波数依存性の無い検出器を作成した。

それぞれの帯域での計算結果を、時間遅れ領域での荷重  $w_{LAG}(\tau; T_C)$  を掛けて統合することとする。

$$\bar{A}(\tau) = C_0 \sum_{k=1}^M w_{LAG}(\tau; T_L 2^{\frac{1-k}{L}}) A(\tau; T_L 2^{\frac{1-k}{L}}), \quad (8)$$

ここで、 $w_{LAG}(\tau; T_C)$  は、周期  $T_C$  で値が 1 となるような単峰性の重み関数である。 $L$  は、オクターブあたりに設ける帯域の数を表す。 $C_0$  は、基本周波数成分の周期性を表す指標である  $\bar{A}(\tau)$  を正規化するための係数であり、周期的信号のみを分析した場合に  $\bar{A}(\tau)$  が 1 となるように設定した。

こうして求められた指標  $\bar{A}(\tau)$  は、目的とする基本周波数の範囲では、基本周波数の値に依存せず、周期性の程度を反映した値をとる。乱数の系列を入力した場合にも、信号の統計的揺らぎにより、 $\bar{A}(\tau)$  にはピークが生ずる。乱数により生ずるピークの値は、前述の上限と比較して十分に小さい。ここでは、統計的な揺らぎに起因するピークを周期信号によるものと誤認する危険率に基づいて、信頼できる基本周波数軌跡の範囲を選択することとした。(例えば危険率を 1% と設定した場合、閾値は 0.61 となる。)