

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5275612号  
(P5275612)

(45) 発行日 平成25年8月28日 (2013. 8. 28)

(24) 登録日 平成25年5月24日 (2013. 5. 24)

(51) Int. Cl.		F I		
G 1 O L 25/18	(2013. 01)	G 1 O L 11/00	1 O 1 E	
G 1 O L 25/90	(2013. 01)	G 1 O L 11/04		
G 1 O L 25/45	(2013. 01)	G 1 O L 11/00	3 O 1	

請求項の数 8 (全 38 頁)

(21) 出願番号	特願2007-289006 (P2007-289006)	(73) 特許権者	504145283 国立大学法人 和歌山大学 和歌山県和歌山市栄谷930番地
(22) 出願日	平成19年11月6日 (2007. 11. 6)	(74) 代理人	100075557 弁理士 西教 圭一郎
(65) 公開番号	特開2009-42716 (P2009-42716A)	(74) 代理人	100072235 弁理士 杉山 毅至
(43) 公開日	平成21年2月26日 (2009. 2. 26)	(72) 発明者	河原 英紀 和歌山県和歌山市松江1487-5 レジ デンス21-102
審査請求日	平成22年11月5日 (2010. 11. 5)	(72) 発明者	森勢 将雅 和歌山県和歌山市中398-1 ル・マン A3番館311号
(31) 優先権主張番号	特願2007-187697 (P2007-187697)		
(32) 優先日	平成19年7月18日 (2007. 7. 18)		
(33) 優先権主張国	日本国 (JP)		

最終頁に続く

(54) 【発明の名称】 周期信号処理方法、周期信号変換方法および周期信号処理装置ならびに周期信号の分析方法

(57) 【特許請求の範囲】

【請求項 1】

周期性を有する信号のうち、時間方向の基本周期を  $n$  ( $n$  は、2 以上の整数) 分の 1 に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出し、

各時間窓によって取り出された複数の部分についてパワースペクトルを算出し、算出したパワースペクトルを同一の比率で加算することを特徴とする周期信号処理方法。

【請求項 2】

請求項 1 記載の周期信号処理方法によって得られたパワースペクトルに、周波数方向で基本周期の幅を有する矩形の平滑化関数を畳み込むことを特徴とする周期信号処理方法。

【請求項 3】

請求項 1 記載の周期信号処理方法によって、周波数方向の所定の範囲毎にパワースペクトルの累積和を求め、

前記周波数方向で定められた間隔を隔てた二点における前記所定の範囲の前記パワースペクトルの累積和の差を求めて線形補間を行うことを特徴とする請求項 1 に記載の周期信号処理方法。

【請求項 4】

前記線形補間によって得られる平滑化されたパワースペクトルを、対数変換し、予め定める補正を行い、指数変換することを特徴とする請求項 3 に記載の周期信号処理方法。

10

20

## 【請求項 5】

請求項 1 記載の周期信号処理方法によって求められたパワースペクトルを請求項 2 ~ 4 のいずれか 1 つに記載の周期信号処理方法によって求められたパワースペクトルで除算して得られたスペクトルから、1 を減算し、重み付きのフーリエ変換を計算することによって、基本周期の値を求めることを特徴とする周期信号の分析方法。

## 【請求項 6】

基本周期の周波数の瞬時周波数に反比例する割合で時間軸を伸縮することによって、見かけ上一定の基本周期の周波数を有する信号となるように変換した周期性を有する信号について、請求項 1 記載の周期信号処理方法によって求められたパワースペクトルを請求項 2 記載の周期信号処理方法によって求められたパワースペクトルで除算して得られたパワースペクトルから、1 を減算して得られる周期性に起因する成分だけを残したスペクトルに、予め定める基本周期の周波数を用いて設計した直交位相信号を畳込んで得られる信号の絶対値としてこの信号に含まれる周期成分の割合を求めることを通じて、この信号に含まれる非周期成分の割合を求めることを特徴とする周期信号の分析方法。

10

## 【請求項 7】

請求項 1 ~ 4 のいずれか 1 つに記載の周期信号処理方法によって得たスペクトルを用いて、前記周期信号を別の信号に変換することを特徴とする周期信号変換方法。

## 【請求項 8】

周期性を有する信号のうち、時間方向の基本周期を  $n$  ( $n$  は、2 以上の整数) 分の 1 に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出す取出手段と、

20

各時間窓によって取り出された複数の部分についてパワースペクトルを算出する算出手段と、

算出したパワースペクトルを同一の比率で加算する加算手段とを含むことを特徴とする周期信号処理装置。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、周期信号処理方法、周期信号変換方法および周期信号処理装置ならびに周期信号の分析方法に関し、特に音などの周期信号を処理するための周期信号処理方法および周期信号処理装置、音などの周期信号を変換するための周期信号変換方法、ならびに音などの周期信号の基本周期または非周期成分を分析する周期信号の分析方法とする。

30

## 【背景技術】

## 【0002】

音声の分析・合成などにおいて、音声のイントネーションを制御する場合や音声の編集合成において自然な音声の抑揚を与えるためには、元々格納されている音声の音色を保ちつつ音声の基本周波数を変えることが必要である。また、自然の音をサンプリングして電子楽器の音源として用いる場合も、音色を一定に保ちつつ基本周波数を変えることが必要である。また、基本周波数の変換においては、サンプリング周期で決定される分解能よりも詳細に基本周波数を設定する必要がある。他方、放送などにおいて情報提供者のプライバシーを守るために、個人性がわからないように音声を変換する場合には、音高を変えずに音色を変えたり、音色と音高の双方を変えたりする必要がある。

40

## 【0003】

また、異なった俳優の声を合成をすることによって、実際に声優を雇わなくても新しい声優の声を作り出すことなど、既存の音声資源の再利用が、ますます強く求められるようになってきている。高齢化社会を迎え、さまざまな聴覚障害や認知能力の障害などによりそのままでは音声や音楽の内容を聞き取ることが困難な人々の増加が予想されている。このような人たちの劣化した聴覚能力や認知能力に適合するように元の情報を失うことなく速度や、周波数帯域、声の高さを変換する方法は、強く要請されている。

## 【0004】

50

このような目的を達成するための第1の従来技術では、スペクトル包絡を表わすモデルを仮定して、モデルのパラメタを適当な評価関数の下でスペクトルのピークを重視して近似するようにパラメタを最適化することでスペクトル包絡を求めている（たとえば非特許文献1参照）。

【0005】

また、第2の従来技術では、周期信号であることを自己回帰モデルのパラメタ推定方式の中に組み込んでいる（たとえば非特許文献2参照）。

【0006】

また、第3の従来技術では、PSOLA（Pitch Synchronous OverLap Add）法のように時間領域での波形の伸縮と時間を移動させた重ね合わせにより音声を加工している。

【0007】

【非特許文献1】今井聖，北村正，「対数振幅特性近似フィルタを用いた音声の分析合成系」，電子通信学会論文誌，78/6，Vol. J61-A，No. 6，pp527-534

【非特許文献2】中田和男，「ピッチ周波数に影響されないホルマント抽出」，日本音響学会誌50巻2号（1994），pp110-116

【発明の開示】

【発明が解決しようとする課題】

【0008】

上述した第1および第2の従来技術のいずれも、特定のモデルを仮定しているためモデルを記述するパラメタの個数を適切に決定しなければ、正しいスペクトル包絡を推定することはできないという問題点がある。また、信号源の性質が想定したモデルと異なっている場合には、推定されたスペクトル包絡に周期性に基づく成分が混入してしまい逆に大きな誤差を生じてしまうという脆弱さを有するという問題点がある。さらに、第1および第2の従来技術では、最適化の過程で収束のための繰返し演算を必要としており、実時間処理のような時間的制約の大きい応用に不適切であるという問題点がある。

【0009】

さらに、第1および第2の従来技術において、周期性の制御について言及すると、音源をパルス列、スペクトル包絡をフィルタとして分離してしまっているため、標本化周波数で決定される時間分解能よりも高い精度で信号の周期を指定することができないという問題点がある。

【0010】

また第3の従来技術では、音源の周期を20%程度以上変化させると音声の自然さが失われてしまい、自由に音声を変換できないという問題点がある。

【0011】

また基本周波数の抽出において、従来技術では、音声合成を前提とした基本周波数の抽出に要求される条件を論理的に詰めずに設計されているため、合理的な設計が行われていない。時間分解能をどの程度にすべきかについても、原則は無く、時間窓のサイズなども試行錯誤などの方法で決められている。そのため、抽出された基本周波数を用いて合成した信号を再分析した場合、合成に用いたものとは異なった基本周波数が求められてしまうという問題がある。

【0012】

また従来技術では、非周期性に関連する物理属性を体系的に関連づけていなかったため、基本周波数の時間変化およびスペクトルの時間変化の影響を非周期成分として抽出してしまい、合成の際に用いるべき正確な値を抽出することができないという問題がある。

【0013】

したがって本発明の目的は、スペクトルのモデルに基づかず、かつ、周期性の影響を小さくできる周期信号処理方法、周期信号変換方法および周期信号処理装置、ならびに周期性を有する信号の基本周期および非周期成分を精度よく求めることができる周期信号の分析方法を提供することである。

10

20

30

40

50

**【課題を解決するための手段】****【0014】**

本発明は、周期性を有する信号のうち、時間方向の基本周期を $n$  ( $n$ は、2以上の整数)分の1に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出し、

各時間窓によって取り出された複数の部分についてパワースペクトルを算出し、

算出したパワースペクトルを同一の比率で加算することを特徴とする周期信号処理方法である。

**【0015】**

また本発明は、前記周期信号処理方法によって得られたパワースペクトルに、周波数方向で基本周期の幅を有する矩形の平滑化関数を畳み込むことを特徴とする。

10

**【0016】**

また本発明は、前記周期信号処理方法によって、周波数方向の所定の範囲毎にパワースペクトルの累積和を求め、

前記周波数方向で定められた間隔を隔てた二点における前記所定の範囲の前記パワースペクトルの累積和の差を求めて線形補間を行うことを特徴とする。

**【0017】**

また本発明は、前記線形補間によって得られる平滑化されたパワースペクトルを、対数変換し、予め定める補正を行い、指数変換することを特徴とする。

**【0018】**

20

また本発明は、周期性を有する信号のうち、時間方向の基本周期を $n$  ( $n$ は、2以上の整数)分の1に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出し、各時間窓によって取り出された複数の部分についてパワースペクトルを算出し、算出したパワースペクトルを同一の比率で加算すること周期信号処理方法によって求められた第1のパワースペクトルを、この第1のパワースペクトルに、周波数方向で基本周期の幅を有する矩形の平滑化関数を畳み込んで得られる第2のパワースペクトルで除算して得られたスペクトルから、1を減算し、重み付きのフーリエ変換を計算することによって、基本周期の値を求めることを特徴とする周期信号の分析方法である。

**【0019】**

30

また本発明は、基本周期の周波数の瞬時周波数に反比例する割合で時間軸を伸縮することによって、見かけ上一定の基本周期の周波数を有する信号となるように変換した周期性を有する信号について、前記第1パワースペクトルを前記第2パワースペクトルで除算して得られたパワースペクトルから、1を減算して得られる周期性に起因する成分だけを残したスペクトルに、予め定める基本周期の周波数を用いて設計した直交位相信号を畳み込んで得られる信号の絶対値としてこの信号に含まれる周期成分の割合を求めることを通じて、この信号に含まれる非周期成分の割合を求めることを特徴とする周期信号の分析方法である。

**【0020】**

また本発明は、前記周期信号処理方法によって得たスペクトルを用いて、前記周期信号を別の信号に変換することを特徴とする周期信号変換方法である。

40

**【0021】**

また本発明は、周期性を有する信号のうち、時間方向の基本周期を $n$  ( $n$ は、2以上の整数)分の1に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出す取出手段と、

各時間窓によって取り出された複数の部分についてパワースペクトルを算出する算出手段と、

算出したパワースペクトルを同一の比率で加算する加算手段とを含むことを特徴とする周期信号処理装置である。

**【発明の効果】**

50

## 【0022】

本発明によれば、周期性を有する信号に対して、分析位置に依存しないパワースペクトルを得ることができ、精度の高いパワースペクトルを求めることができる。周期性を有する信号のうち、時間方向の基本周期を $n$  ( $n$ は、2以上の整数)分の1に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出し、時間窓によって取り出された複数の部分についてパワースペクトルを算出し、算出したパワースペクトルを同一の比率で加算するという簡単な処理によって、分析位置に依存しないパワースペクトルを得ることができ、このような分析位置に依存しないパワースペクトルを得るために、複雑な計算、およびパラメタの調整をする必要がないか、あるいは極めて限定された少数のパラメタのみを設定するだけでよい。したがって、目的に応じた設計を容易に行うことができ、また簡単に計算できる関数のみを用いることができるので、短時間に簡単に、分析時刻に依存しないスペクトログラムを得ることができる。

10

## 【0023】

時間方向の基本周期を $n$  ( $n$ は、2以上の整数)分の1に分割する分割位置に中心を置くように時間窓をそれぞれ配置することによって、信号の時刻による変動をゼロ(0)とすることができる。

## 【0024】

また本発明によれば、分析位置に依存しないパワースペクトルを用いることができるため、分析位置に依存せずに、周波数方向の周期性を取り除いたスペクトルを求めることができる。このように時間方向と周波数方向の双方において周期性の影響が取り除かれたスペクトルを、音声合成、音声変換および音声認識などにおいて用いることによって、合成音または変換音の品質および音声認識の認識率を向上させるなどの効果を達成することができる。

20

## 【0025】

また本発明によれば、前記周波数方向の所定の範囲毎にパワースペクトルを求め、前記周波数方向で定められた間隔を隔てた二点における前記所定の範囲のパワースペクトルの差を求めて線形補間することによって、周波数方向に、さらに平滑化されたスペクトログラムを得ることができ、周波数方向の信号強度を平滑化して、ノイズを低減することができる。

## 【0026】

また本発明によれば、前記線形補間によって得られる平滑化されたパワースペクトルを、対数変換し、予め定める補正を行い、指数変換することによって、前述した各処理によって平滑化され過ぎた部分についてのパワースペクトルを元に戻すことができ、特に音声信号を処理するとき、音声に忠実なスペクトルを得ることができる。

30

## 【0027】

また本発明によれば、平滑化されたスペクトログラムを用いて、周期信号を別の信号に変換している。このため周波数方向および時間方向の周期性の影響が小さくなる。したがって、時間分解能および周波数分解能をバランスよく決定することができる。

## 【0028】

また本発明によれば、精度よく基本周期の値を求めることができる。基本周波数は、基本周期の値の逆数で表される。基本周波数に応じて適切なサイズの時間窓を選択すれば、音声合成に用いた場合に元の信号と同じ基本周波数が抽出されるような信号を合成することができる。また、複数の基本周波数を有する信号を適切に分析することができるようになるので、これまで分析合成を適切に行うことができなかつたみ声の分析合成が可能となる。

40

## 【0029】

また本発明によれば、非周期性を正確に推定することができる。正確に推定された非周期性を用いれば、音声合成および音声変換などにおいて、合成音声および加工音声の品質を改善することができる。また、非周期性の推定方法に根拠の曖昧な非線形処理を含まないため、声を利用した診断などに応用することができる。

50

## 【発明を実施するための最良の形態】

## 【0030】

図1は、本発明の実施の一形態の音声変換方法を実現するための周期信号変換装置1を示す概略ブロック図である。図2～4は、周期信号変換装置1が有するパワースペクトル取得部2を示す概略ブロック図である。前記音声変換方法は、周期信号処理方法を含んでいる。周期信号変換装置1は、音声信号の周期性を積極的に利用することによって、繰返しと収束の判定を含む計算を必要としない直接的な計算でスペクトル包絡を求めることを可能とする。また、そうして求めたスペクトル包絡から信号を再合成する際に位相を操作することにより、標本化周期よりも細かな分解能での周期の制御と音色の制御を実現する。周期信号変換装置1は、マイクロコンピュータによって実現され、CPU（中央演算処理装置）などの処理回路が、予め定めるプログラムを実行することによって実現される。

10

## 【0031】

周期信号変換装置1は、パワースペクトル取得部2、基本周期計算部3、平滑化スペクトル変換部4、音源情報変換部5、位相調整部6および波形合成部7を備える。これらの各部は、処理回路が、予め定めるプログラムを実行することによって機能する。周期信号変換装置1を用いて、22.05kHz、かつ16ビットで標本化された音声を変換する例を説明する。

## 【0032】

パワースペクトル取得部2は、周期性を有する信号のうち、時間方向に1周期の範囲であって、かつ時間方向に予め定める時間だけ異なる2つの範囲の部分、窓関数（時間窓）を用いて取り出し、窓関数によって取り出された2つの部分についてパワースペクトルを算出し、算出したパワースペクトルを同一の比率で加算し、加算した前記パワースペクトルの周波数方向の累積和に基づいてスペクトログラムを得る。パワースペクトル取得部2は、周期信号処理装置である。

20

## 【0033】

まず原理について、以下に説明する。図5は、入力信号である音声波形を示すグラフであり、図6は、窓関数を示すグラフである。図5および図6において横軸は、時間を表し、縦軸は振幅を表す。

## 【0034】

本発明の周期信号処理方法を用いることによって、パワースペクトル取得部2において、理論的に時間方向の変動を原理的には完全に取り除くことができることを保証することができる。この周期信号処理方法では、一種類の時間窓（窓関数）から求めたパワースペクトルと、それと同じ時間窓を時間方向に予め定める時間だけ移動させて求めたパワースペクトルとを同一の比率で加算することによって、目的とするパワースペクトルを求める。前記予め定める時間は、1周期（すなわち基本周期）の半分である。以後、一種類の時間窓（窓関数）から求めたパワースペクトルと、それと同じ時間窓を時間方向に予め定める時間だけ移動させた時間窓とを合わせて、TANDEM窓という場合がある。

30

## 【0035】

周期信号処理方法で用いる窓関数は、周期信号を分析した場合に、ある調波成分のパワースペクトルに対する隣接する調波成分およびそれ以上離れた調波成分からの影響が十分に小さい窓関数であれば、どのような窓関数が用いられても構わない。

40

## 【0036】

まず入力信号の一部を取り出すための時間窓を用意する。この時間窓の周波数特性は、低域通過型であり、かつ直流成分を通すものであるとする。帯域通過特性を有する場合には、中心周波数と同じ周波数の信号を用いて同期検波することによって中心周波数を直流に変換することができる。したがって、このように特性を指定することによって議論の一般性が失われることはない。この窓関数を $w(t)$ と表すことにする。また時間窓 $w(t)$ のフーリエ変換を $H(\omega)$ と表すことにする。ここで $\omega$ は、角周波数を表す。 $H(\omega)$ は、低域通過特性を有するので、ある角周波数 $\omega_0 = 2\pi f_0$ 以上の角周波数の成分は、通過しないものとみなす。なお、ここで $f_0$ は、 $\omega_0$ に対応する周波数を表す。実際には

50

、 $\omega_0$  以上の成分も、わずかではあるが通過する。その場合については、後に説明する。

【0037】

このような窓関数を用いて、基本周波数が  $f_0$  であるような周期関数  $x(t)$  を分析することとする。周期関数である  $x(t)$  は、次のようにフーリエ級数として表すことができる。

【0038】

【数1】

$$x(t) = \sum_{k \in Z} X_k e^{j \frac{\omega_k t}{T_0}} \quad \dots (1)$$

10

【0039】

ここで、 $Z$  は、整数全体の集合を表し、 $X_k$  は、一般に複素数となる。また、 $T_0 = 1/f_0$  は、基本周期を表す。

【0040】

窓関数を用いた短時間フーリエ変換は、この信号  $x(t)$  と、窓関数  $w(t - \tau)$  との積として表される信号  $s(t) = x(t) w(t - \tau)$  のフーリエ変換となる。窓関数が、時刻 0 を中心とする関数である場合には、 $\tau$  は、分析時の窓の中心時刻を表す。この時刻を明示的にパラメタとして用いて、時刻  $\tau$  を中心とする窓のフーリエ変換を  $H(\omega, \tau)$  と表すこととすると、 $H(\omega, \tau)$  は  $H(\omega)$  を用いて、次のように表される。

$$H(\omega, \tau) = H(\omega) e^{-j \omega \tau} \quad \dots (2)$$

20

【0041】

時間領域での積は、フーリエ変換によって周波数領域での畳込みに対応している。ここで、信号  $x(t)$  のフーリエ変換を求めておく。

【0042】

【数2】

$$X(\omega) = \sum_{k \in Z} X_k \delta(\omega - k\omega_0) \quad \dots (3)$$

【0043】

ここで  $\delta(\omega)$  は、ディラック (Dirac) のデルタ関数である。この周波数軸上で等間隔に配置されたデルタ関数の列として表される  $X(\omega)$  が、時刻  $\tau$  に置かれた窓関数のフーリエ変換である  $H(\omega, \tau)$  と畳込まれ、短時間フーリエ変換  $S(\omega, \tau)$  となる。

30

【0044】

ところで、 $H(\omega)$  は、 $\omega_0$  よりも高い角周波数成分を通さないように設定されている。したがって、ある角周波数  $\omega$  に注目したとき、 $S(\omega, \tau)$  に影響を与えるのは、 $\omega$  に最も近い角周波数の成分と、次に近い角周波数成分の2つの成分だけとなる。なお、2つの成分は隣接しているので、式における調波を表す番号は、片方の成分が偶数であれば、他方の成分は奇数となる。

【0045】

$S(\omega, \tau)$  のふるまいを調べるために、分析の対象となる信号のフーリエ変換  $X(\omega)$  を、以下のように片方の係数を1とした2つの複素指数関数からなる信号と考えると、一般性は失われない。

40

【0046】

【数3】

$$X(\omega) = \delta(\omega) + \alpha e^{j\beta} \delta(\omega - \omega_0) \quad \dots (4)$$

【0047】

この信号と、時刻  $\tau$  に置かれた窓関数のフーリエ変換  $H(\omega, \tau)$  とを畳込むことによって、分析時刻に依存したスペクトル  $S(\omega, \tau)$  が求められる。ここで、 $H(\omega, \tau)$  を  $H(\omega)$  と時間遅れを表す複素数を用いて表すこととする。

50

【 0 0 4 8 】

【数 4】

$$\begin{aligned} S(\omega, \tau) &= X(\omega) * H(\omega, \tau) \\ &= e^{-j\omega\tau} \left( H(\omega) + H(\omega - \omega_0) \alpha e^{j(\tau\omega_0 + \beta)} \right) \quad \dots (5) \end{aligned}$$

【 0 0 4 9 】

ここで、\*は畳込みを表す。この絶対値の自乗を求めて整理することによって、次のようにパワースペクトルが求められる。

【 0 0 5 0 】

10

【数 5】

$$\begin{aligned} |S(\omega, \tau)|^2 &= H^2(\omega) + \alpha^2 H^2(\omega - \omega_0) \\ &\quad + 2\alpha H(\omega) H(\omega - \omega_0) \cos(\omega_0\tau + \beta) \quad \dots (6) \end{aligned}$$

【 0 0 5 1 】

この式の右辺の第3項は、窓の時刻  $\tau$  の変化に応じて正弦波状に変化する成分を表している。

【 0 0 5 2 】

ここで、 $H(\omega, \tau)$  を、基本周期の半分だけ移動させて信号を切り出し、パワースペクトルを求めることを考える。すなわち、 $H(\omega, \tau - T_0/2)$  を用いて、パワースペクトルを求めることとする。整理すると、次式が得られる。

20

【 0 0 5 3 】

【数 6】

$$\begin{aligned} |S(\omega, \tau + T_0/2)|^2 &= H^2(\omega) + \alpha^2 H^2(\omega - \omega_0) \\ &\quad - 2\alpha H(\omega) H(\omega - \omega_0) \cos(\omega_0\tau + \beta) \quad \dots (7) \end{aligned}$$

【 0 0 5 4 】

ここで、 $|S(\omega, \tau)|^2$  と  $|S(\omega, \tau + T_0/2)|^2$  とを加えると、以下が得られる。

30

【 0 0 5 5 】

【数 7】

$$|S(\omega, \tau)|^2 + |S(\omega, \tau + T_0/2)|^2 = 2(H^2(\omega) + \alpha^2 H^2(\omega - \omega_0)) \quad \dots (8)$$

【 0 0 5 6 】

右辺には、窓の置かれた時刻  $\tau$  が含まれていない。すなわち、どの時刻で分析しても同じパワースペクトルを求めることができる。

【 0 0 5 7 】

次に、 $\omega_0$  よりも大きな角周波数からの影響について説明する。実質的には、それらの成分からの影響は、無視できる程度の大きさになる。たとえば、よく用いられるハニング(hanning)窓を例にとると、ハニング窓をここで説明した方法で用いる場合、窓の長さを分析対象とする信号の2倍とすることが合理的である。その場合、窓の振幅周波数特性の最小のサイドローブは、周波数の3乗に反比例して減衰する。ハニング窓のサイドローブは、正と負に極性を交互に変えながら減衰する。しかしながら、ここでは最悪条件を考慮するために同じ極性の場合について評価する。このように考えると、ハニング窓の場合、サイドローブ全体の寄与は、以下の級数の極限によって上限が抑えられる。

40

【 0 0 5 8 】



【数 8】

$$c_0 + c_0 \sum_{k=2}^n \frac{1}{k^3} \quad \dots (9)$$

【0059】

この値は  $2C_0$  を超えない。ここで  $C_0$  は、最初のサイドロープレベルを表す。結局、最悪の場合でも影響は  $-25 \text{ dB}$  を超えることはない。調波のレベルが等しい場合には、この影響は注目している調波のレベルを約  $0.5 \text{ dB}$  変化させる程度である。この程度の影響は、音声のスペクトルの時間的変動と比較すると、十分に小さいため、実質的に無視することができる。実際の信号の場合には、前述したようにサイドロープの極性が相殺し、また、成分間の位相も一般には一致しないため、影響はこの上限よりも遥かに少なくなる。なお、このように設計したハニング窓の場合、振幅周波数特性には、 $k f_0 / 2$  ( $k$  は、 $-1, 0, 1$  以外の整数) に零点があるので、 $n_1 f_0 / 2$  ( $n_1$  は、整数) のパワースペクトルには、誤差が全く含まれない。

10

【0060】

また、パワースペクトル取得部 2 では、スペクトルの正定値性を保証しかつ、新しい標本化定理の考え方に基づくことで一意性と最適性を保証することのできるスペクトル復元を行う。新しい標本化定理では、アナログ信号の標本化と、標本からのアナログ信号の復元とを一体のものとして考える。この標本化定理を説明する。

【0061】

ここでは、まず、対象とする系を定義しておく。標本化は、未知の入力信号 (関数) を  $f \in H$  を、ある関数  $\varphi_1(t)$  をインパルス応答とするような分析用関数により処理したものを、離散的に取り出す操作であると考え。また、標本からアナログ信号への復元は、積分が標本値となるようなデルタ関数である関数  $\varphi_2(t)$  をインパルス応答とするような合成用関数によって次々に処理していく操作であると考え。

20

【0062】

このように標本化と標本からの復元を定義した上で標本化定理を見直す。まず、分析合成の関数の相互関数  $a_{12}(k)$  を計算しておく。

【0063】

【数 9】

$$a_{12}(k) = \langle \varphi_1(t-k), \varphi_2(t) \rangle \quad \dots (10)$$

30

【0064】

なお、 $a(t), b(t)$  は、 $a(t)$  と  $b(t)$  の内積を表し以下のように定義される。

【0065】

【数 10】

$$\langle a, b \rangle = \int_{-\infty}^{\infty} b^*(t)a(-t)dt \quad \dots (11)$$

40

【0066】

これらの準備の下で、以下の標本化定理が成立する。

未知の入力信号 (関数)  $f \in H$  を考える。ここで、 $A_{12}(e^{j\omega}) > m$  を満たすような  $m > 0$  が存在するとすると、次式の意味で一貫性 (consistency) を満たす  $f$  の近似である  $V(\varphi_2)$  の要素が  $f$  が一意的に決まる。

【0067】

【数 11】

$$\forall f \in H, \quad c_1(k) = \langle f, \varphi_1(x-k) \rangle = \langle \tilde{f}, \varphi_1(x-k) \rangle \quad \dots (12)$$

なお、ここで、

50

【 0 0 6 8 】

【 数 1 2 】

$$A_{12}(e^{j\omega}) = \sum_{k \in Z} a_{12}(k)e^{-j\omega k}, \quad \dots (13)$$

である。また  $V(\quad_2)$  は、 $\quad_2$  により張られるベクトル空間を表す。

【 0 0 6 9 】

$c_1(k)$  は、標本化によって得られた標本値の系列である。短時間フーリエ変換は、窓関数を包絡線として持つ複素指数関数をインパルス応答とするフィルタ処理と等価であり、スペクトログラムは、窓関数の自乗を分析関数  $\quad_1$  とするフィルタ処理からの標本値を表していると解釈することができる。通常のスペクトログラムは、この  $c_1(k)$  をそのまま眺めていることに相当する。目的は、 $c_1(k)$  を用いて、近似関数  $f$  を再構成し、それを同様に分析関数を用いて分析したときに元の関数  $f$  を分析したときと同じ結果である  $c_1(k)$  が得られるようにすることである。これが consistent sampling である。

10

【 0 0 7 0 】

ここで、周期信号のパワースペクトルが式 8 として表されることに注意する。これは、TANDEM 窓によるパワースペクトルが、窓関数の振幅周波数特性の絶対値の自乗と、隣接する 2 つのデルタ関数の畳込みとして表されることを意味する。周期性の影響の除去のためには、底辺の大きさが基本周波数と等しい矩形の平滑化関数を用いればよい。矩形の平滑化関数を用いる計算は、実際に平滑化を行わずとも、累積和と線形補間から容易に計算することができる。これらにより、前述した標本化定理を満たす処理を、次の手順によって求めることができる。

20

【 0 0 7 1 】

1 . 分析関数と合成関数の相関を計算し、前述した標本化定理を満たす補正係数を求める。

2 . TANDEM 窓によって信号を分析し、パワースペクトルを求める。

3 . パワースペクトルの累積を求める。

4 . 累積和の線形補間によって求められる 2 つの周波数における累積和の値の差によって矩形の平滑化関数による平滑化の結果を計算する。

30

5 . 平滑化されたパワースペクトルを補正係数を用いて補正する。

【 0 0 7 2 】

求められたスペクトルを正弦波モデルによる音声合成に用いる場合には、基本周波数が一定であれば、合成用の関数はデルタ関数となる。スペクトルから FIR (Finite Impulse Response) フィルタを作成して合成に用いる場合には、FIR フィルタの計算に用いられる窓関数のパワースペクトルが合成用フィルタとなる。これらは、各フレームの分析に先立って、予め計算しておくことのできる値である。

【 0 0 7 3 】

補正されたパワースペクトルの正定値性を保証するために、次の性質を用いる。対数関数  $\ln(x)$  は、 $x = 1$  の付近でテイラー (Taylor) 展開することによって、 $(x - 1)$  の冪級数として表される。ここで、 $x = (x - 1)$  が十分に小さい場合には、1 次の項よりも高い次数の項を無視することができる。すなわち線形近似することができる。線形近似が成立している場合には、前述した補正係数をそのまま用いることができる。

40

【 0 0 7 4 】

補正係数は、厳密には複数個必要である。しかしながら、隣接する調波よりも遠い成分からの影響を考慮することは、実際の音声の処理では、様々な副作用があるので望ましくない。ここでは、隣接する調波だけを補正する場合に、節点における誤差が最小になるという条件で補正係数を求めることにより、副作用を避け、かつ、計算時間を短縮する方法を提案する。具体的には、補正係数  $q_k \{k \in \{0, 1\}\}$  から求める修正した補正係数をその文字の上に横棒を付した記号であらわすこととし、以下により求める。 $\quad_2$  を  $q_k$  の修

50

正した補正係数により重みづけて加えたものと  $q_1$  とを畳込んだ結果の、節点での値の自乗和が最小になるように、 $q_k$  の修正した補正係数に関する最小化問題を数値的に予め解いておく。

$q_k$  の修正した補正係数は、  
【 0 0 7 5 】  
【 数 1 3 】

$$\bar{q}_k \quad \dots (14)$$

で表され、

また、 $q_0$  の修正した補正係数は、  
【 0 0 7 6 】  
【 数 1 4 】

$$\bar{q}_0 = 1 - 2\bar{q}_1 \quad \dots (15)$$

として求められる。この修正した補正係数は、毎回計算する必要は無い。

【 0 0 7 7 】

式 1 6 は、前述した 1 ~ 5 の手続のうち、3, 4, 5 の手続きを具体的に数式を用いて表したものである。 $P_T(\omega)$  は、T A N D E M 窓により求められたパワースペクトルであり、 $C(\omega)$  は、パワースペクトルの累積である。累積する積分範囲の上限と下限は、0 からナイキスト周波数の範囲を上と下に、それぞれ  $2\omega_0$  だけ広げたものを用いる。式 1 6 は、基本角周波数  $\omega_0$  の幅を有する矩形の関数と T A N D E M 窓により求められたパワースペクトルを畳込んだ結果を対数変換したものを、このパワースペクトルの累積を用いて計算する方法を表している。パワースペクトルの累積から  $\omega_0$  だけ離れた 2 つの角周波数における値を線形補間を用いて精密に読み出して、高い角周波数における値から低い周波数における値を求めるだけで、畳込みを行ったものと同じ結果が得られる。それを、対数変換することで、対数の領域で表した平滑化スペクトル  $L_s(\omega)$  を得ている。式 1 6 の最後の式は、この平滑化スペクトルを、補正係数  $q_0$  の修正した補正係数と  $q_1$  の修正した補正係数を用いて組み合わせることで、補正した対数スペクトルを求め、指数変換することで、正值であることが保証された補正された平滑化パワースペクトルを求める具体的な方法を与えている。

【 0 0 7 8 】

【 数 1 5 】

$$C(\omega) = \int_{\omega_L}^{\omega_U} P_T(\omega) d\omega \quad \dots (16)$$

$$L_s(\omega) = \ln [C(\omega + \omega_0/2) - C(\omega - \omega_0/2)]$$

$$P_{TST}(\omega) = e^{[\bar{q}_1(L_s(\omega - \omega_0) + L_s(\omega + \omega_0)) + \bar{q}_0 L_s(\omega)]}$$

【 0 0 7 9 】

音声があるスペクトログラムから選択されたスペクトル断面から最小位相のインパルス応答を用いて合成されるものとする。この場合、それぞれの極に対応する減衰振動は、指数関数的に減衰する。一方、極の存在しない帯域での応答は、分析の窓関数の持続時間となり、また窓の自乗の応答となる。これが前述した標本化定理の合成用関数に対応する。

【 0 0 8 0 】

次に、図 2 ~ 4 を参照して、パワースペクトル取得部 2 の各構成について説明する。パワースペクトル取得部 2 は、処理の流れの順番に、第 1 ~ 第 3 部分 1 1 ~ 1 3 に分けられる。図 2 に第 1 部分 1 1 を示し、図 3 に第 2 部分 1 2 を示し、図 4 に第 3 部分 1 3 を示す。第 2 および第 3 部分 1 2, 1 3 は、スペクトログラム取得手段である。

【 0 0 8 1 】

第 1 部分 1 1 は、遅延部 2 1 と、第 1 および第 2 窓処理部 2 2, 2 3 と、第 1 および第

10

20

30

40

50

2 パワースペクトル計算部 2 4 , 2 5 と、パワースペクトル加算部 2 6 とを含んで構成される。遅延部 2 1 は、入力信号を、予め定める時間だけ遅延させて、第 2 窓処理部 2 3 に与える。入力信号は、遅延部 2 1 と第 1 窓処理部 2 2 に同時に与えられる。周期信号変換装置 1 に与えられる入力信号は、第 1 および第 2 窓処理部 2 2 , 2 3 のそれぞれに与えられるが、第 2 窓処理部 2 3 に与えられる入力信号は、遅延部 2 1 によって、第 1 窓処理部 2 2 に与えられる入力信号に対して、予め定める時間だけ遅延させることができる。遅延部 2 1 が、入力信号を遅延させる時間は、基本周期  $T_0$  の  $1/2$  である。基本周期に関する情報は、基本周期計算部 3 から与えられ、基本周期計算部 3 から与えられる基本周期に関する情報に応じて、遅延部 2 1 は遅延時間を決定する。遅延部 2 1、第 1 および第 2 窓処理部 2 2 , 2 3 は、取出手段である。

10

## 【 0 0 8 2 】

第 1 および第 2 窓処理部 2 2 , 2 3 は、与えられる入力信号の一部をハニング窓によって切り出す。第 1 窓処理部 2 2 によって切り出された信号は、第 1 パワースペクトル計算部 2 4 に与えられ、第 2 窓処理部 2 3 によって切り出された信号は、第 2 パワースペクトル計算部 2 5 に与えられる。ハニング窓の長さは、基本周期  $T_0$  の 2 倍に選ばれる。基本周期に関する情報は、基本周期計算部 3 から与えられ、基本周期計算部 3 から与えられる基本周期に関する情報に応じて、第 1 および第 2 窓処理部 2 2 , 2 3 はハニング窓の長さを決定する。

## 【 0 0 8 3 】

第 1 および第 2 パワースペクトル計算部 2 4 , 2 5 では、FFT (高速フーリエ変換) により、音声波形のパワースペクトルを計算する。このパワースペクトルには、音声の周期性による調波構造が観測される。第 1 および第 2 パワースペクトル計算部 2 4 , 2 5 は、算出手段である。

20

## 【 0 0 8 4 】

図 7 は、第 1 および第 2 パワースペクトル計算部 2 4 , 2 5 によって求められたパワースペクトルの一例を示すグラフである。図 7 のグラフにおいて、X 軸は時刻を示し、Y 軸は周波数を示し、Z 軸は強度を対数表示 (デシベル表示) を用いて示している。各軸の単位は、任意単位である。

## 【 0 0 8 5 】

第 1 および第 2 パワースペクトル計算部 2 4 , 2 5 によって計算されたパワースペクトルは、パワースペクトル加算部 2 6 に与えられる。パワースペクトル加算部 2 6 は、第 1 および第 2 パワースペクトル計算部 2 4 , 2 5 から与えられる各パワースペクトルを加算して、加算したパワースペクトル (出力パワースペクトル) を出力する。パワースペクトル加算部 2 6 は、加算手段である。

30

## 【 0 0 8 6 】

図 8 は、パワースペクトル加算部 2 6 から出力される出力パワースペクトルの一例を示すグラフである。図 8 のグラフにおいて、X 軸は周波数を示し、Y 軸は時刻を示し、Z 軸は強度を対数表示 (デシベル表示) を用いて示している。各軸の単位は、任意単位である。

## 【 0 0 8 7 】

出力パワースペクトルは、第 2 部分 1 2 に与えられる。第 2 部分 1 2 は、累積パワースペクトル計算部 3 1 と、第 1 および第 2 平滑化スペクトル計算部 3 2 , 3 3 と、対数変換部 3 4 , 3 5 と、最適周波数補償合成部 3 6 とを含んで構成される。出力パワースペクトルは、累積パワースペクトル計算部 3 1 に与えられる。累積パワースペクトル計算部 3 1 は、与えられる出力パワースペクトルの累積和を算出する。出力パワースペクトルの累積和は、第 1 および第 2 平滑化スペクトル計算部 3 2 , 3 3 に与えられる。

40

## 【 0 0 8 8 】

第 1 および第 2 平滑化スペクトル計算部 3 2 , 3 3 は、基本角周波数だけ異なった周波数の対について、それぞれの角周波数を中心とする基本角周波数の間隔を隔てた角周波数における累積パワースペクトルの値から、矩形の関数を畳込んだ結果に相当する平滑化ス

50

ペクトルを計算する。

【0089】

図9は、第1および第2平滑化スペクトル計算部32, 33の各部から出力される平滑化されたパワースペクトルの一例を示すグラフである。図9のグラフにおいて、X軸は周波数を示し、Y軸は時間を示し、Z軸は強度を対数表示(デシベル表示)を用いて示している。各軸の単位は、任意単位である。

【0090】

第1および第2対数変換部34, 35は、求められた平滑化スペクトルの値の対数変換を行う。

【0091】

最適周波数補償合成部36は、第1および第2対数変換部34, 35によって対数に変換された平滑化スペクトルの値を、最適な補正係数を用いて合成し、最適周波数平滑化対数パワースペクトルを出力する。

【0092】

図10は、最適周波数補償合成部36から出力される最適周波数平滑化対数パワースペクトルの一例を示すグラフである。図10のグラフにおいて、X軸は周波数を示し、Y軸は時間を示し、Z軸は強度を対数表示(デシベル表示)を用いて示している。各軸の単位は、任意単位である。

【0093】

最適周波数平滑化対数パワースペクトルは、第3部分13に与えられる。第3部分13は、3フレーム蓄積部41、最適時間補償合成部42、指数変換部43、第1および第2蓄積部44, 45を含んで構成される。

【0094】

3フレーム蓄積部41は、時間的に基本周期だけ離れた3つの時刻における最適周波数平滑化対数パワースペクトルの蓄積を行う。

【0095】

最適時間補償合成部42は、求められた最適時間周波数平滑化対数パワースペクトルを、指数変換部43および第1蓄積部44に与える。

【0096】

指数変換部43は、最適時間周波数平滑化対数パワースペクトルを、指数変換して、最適時間周波数平滑化パワースペクトルを出力する。

【0097】

第1蓄積部44は、最適時間周波数平滑化対数パワースペクトルを蓄積して、最適時間周波数平滑化対数パワースペクトログラムを出力する。

【0098】

第2蓄積部45は、最適時間周波数平滑化パワースペクトルを蓄積して、最適時間周波数平滑化対数パワースペクトログラムを出力する。

【0099】

前記パワースペクトル取得部2は、前述した信号の処理を、基本周期ごとに行なう。図7, 8, 9, 10の図は、方法の理解を助けるために1ms毎に計算した結果を示しているが、処理と処理の間の値は、処理により求められた値を線形補間したものを用いれば良い。

【0100】

再び図1を参照して、基本周期計算部3は、図5に示されるような音声波形の周期から、信号の基本周期 $T_0$ の抽出を行なう。基本周期計算部3は、たとえば、1msごとに信号の基本周期を抽出する。基本周期計算部3では、波形の自己相関関数を計算し、その最大値を与える時間間隔として基本周期 $T_0$ を抽出する。あるいは、基本波成分を分離するフィルタを用いて抽出した信号の瞬時周波数を求め、その逆数として基本周期 $T_0$ を抽出する。

【0101】

10

20

30

40

50

平滑化スペクトル変換部 4 には、パワースペクトル取得部 2 で得られた最適時間周波数平滑化パワースペクトルが与えられる。平滑化スペクトル変換部 4 では、最小位相のインパルス応答  $v(t)$  を作るために、平滑化スペクトル  $S(\quad)$  を  $V(\quad)$  に変換しておく。また、音色を操作したい場合には、平滑化スペクトルを目的に応じて操作して変形し、変形した平滑化スペクトル  $S_m(\quad)$  を得る。

【0102】

以下の説明では、平滑化されたスペクトルのみならず変形した平滑化スペクトル  $S_m(\quad)$  も、「 $S(\quad)$ 」で表わす。

【0103】

平滑化スペクトル変換部 4 および音源情報変換部 5 では、平滑化スペクトル変換部 4 での変換と並行して、音源情報を目的に応じて変換する。音源情報変換部 5 では、発声者の声の性質を変えるために（たとえば、女性の声を男性の声に変換するために）、求められた音声パラメタ（平滑化スペクトルと精密な基本周期情報）の周波数軸を圧縮したり、声の高さを変えるために、精密な基本周期に適当な係数を掛けたりすることを行なう。このように、音声パラメタを、目的に合わせて変えることが、音声パラメタの変換である。音声パラメタ（平滑化スペクトルと精密な基本周期情報）に対して操作を加えるだけであらゆるバリエーションの音声を作ることができる。

【0104】

位相調整部 6 では、平滑化スペクトル変換部 4 および音源情報変換部 5 で変換されたスペクトル情報と音源情報を用いて、標本化周期よりも高い分解能で周期を操作するための処理を行なう。つまり、目的とする波形を置く時間位置を標本化周期  $T$  を単位として計算し、整数部分と実数部分とに分け、実数部分を用いて位相調整成分  $1(\quad)$  を求める。そして、 $S(\quad)$  あるいは  $V(\quad)$  の位相を調整する。

【0105】

波形合成部 7 では、位相調整部 6 で位相調整された平滑化スペクトルおよび音源情報変換部 5 で変換された音源情報を用いて、波形を合成する。位相調整部 6 および波形合成部 7 は、平滑化スペクトルから、精密な基本周期から決まる周期ごとに音源波形を作成し、時間軸をずらしながら加え合わせていくことによって、変換された音声を作成する。つまり、音声合成をする。時間軸をずらすときには、信号がデジタル化される際の標本化周波数で決まる標本化周期よりも細かい精度でずらすことはできない。そこで、基本周期を積分して次々と得られる時間を標本化周期で割算したときの余りの部分（少数点以下の部分）については、計算した値  $1(\quad)$  に、余りの時間に応じた傾斜を有する、周波数に対して直線的に位相が変化する項を加えることで、標本化周期により決まる分解能よりも細かな精度で基本周期の制御を行なうことが可能となる。

【0106】

また、平滑化スペクトルから、精密な基本周期から決まる周期ごとに音源波形を作成し、時間軸をずらしながら加え合わせていくことによって、変換された音声を作成することもできる。

【0107】

以上のように周期信号変換装置 1 では、簡単な処理によって、スペクトログラムを得ることができ、複雑な計算、およびパラメタの調整をする必要がないか、あるいは極めて限定された小数のパラメタのみを設定するだけでよい。したがって、目的に応じた設計を容易に行うことができ、また簡単に計算できる関数のみを用いることができるので、短時間に簡単に、分析時刻に依存しないスペクトログラムを得ることができる。また周波数方向および時間方向に、さらに平滑化されたスペクトログラムを得ることができ、周波数方向の信号強度を平滑化して、ノイズを低減することができる。さらに平滑化されたスペクトログラムを用いて、周期信号を別の信号に変換している。このため周波数方向および時間方向の周期性の影響が小さくなる。したがって、時間分解能および周波数分解能をバランスよく決定することができる。

【0108】

10

20

30

40

50

本実施の形態では周期信号処理方法を、音声信号の合成に用いているが、本発明の周期信号処理方法が処理対象とする信号は、音声信号に限らず、たとえばエコー検査などで得られる種々の音響信号であってもよい。このような声に限らない信号の処理であっても、同様の効果を達成することができる。

【0109】

また本実施の形態では、パワースペクトル取得部2は、第1～第3部分11～13を備えているが、第1部分11のみによって構成されてもよく、また第1および第2部分11、12のみによって構成されてもよい。このような構成としても、初期の目的を達成することが可能である。

【0110】

また本実施の形態では、窓関数として、ハニング窓を用いているが、ハニング窓とパーレット(Bartlett)窓とを畳んだ窓を用いてもよい。この場合、パートレット窓の長さを基本周期の二倍とすることにより、ハニング窓の長さを基本周期と同じにしてもよい。パートレット窓の長さハニング窓の長さを同じく基本周期の二倍とすることで、より時間的変動を少なくすることが可能である。ただし、その場合には、時間方向の細かな変化に追従する性能が低下する。

【0111】

図11は、本発明の実施の他の形態の音声変換方法を実現するための周期信号変換装置50を示す概略ブロック図である。本実施の形態において前述の実施の形態の周期信号変換装置1の構成に対応する部分には、同様の参照符号を付してその説明を省略する場合がある。本実施の形態の音声変換方法は、周期信号処理方法および周期信号の分析方法を含んでいる。周期信号変換装置50は、処理回路が、予め定めるプログラムを実行することによって実現される。

【0112】

周期信号変換装置50は、基本的に周期信号変換装置1の構成に、非周期成分計算回路54を付加した構成であり、周期信号変換装置50は、パワースペクトル取得部2と、基本周期計算部3と、平滑化スペクトル変換部4と、音源情報変換部5と、位相調整部6と、波形合成部7と、非周期成分計算回路54とを備える。ただし、周期信号変換装置1とは、パワースペクトル取得部2、基本周期計算部3の構成が異なる。これらの各部は、処理回路が、予め定めるプログラムを実行することによって機能する。

【0113】

パワースペクトル取得部2は、周期性を有する信号のうち、時間方向の基本周期を $n$ ( $n$ は、2以上の整数)分の1に分割する分割位置に中心を置くように時間窓をそれぞれ配置して範囲の異なる複数の部分を取り出し、各時間窓によって取り出された複数の部分についてパワースペクトルを算出し、算出したパワースペクトルを同一の比率で加算する。またパワースペクトル取得部2は、加算した前記パワースペクトルの周波数方向の累積和に基づいてスペクトログラムを得る。すなわち時間方向に隣接する時間窓の中心位置は、時間方向の基本周期の $n$ ( $n$ は、2以上の整数)分の1の距離だけ離間している。前述した実施の形態のパワースペクトル取得部2では、 $n$ が、2に選ばれていることになるが、 $n$ は、2に限られない。

【0114】

パワースペクトル取得部2は、TANDEM回路55と、STRAIGHT回路56とを含んで構成される。

【0115】

図12は、TANDEM回路55の構成を示す概略ブロック図である。TANDEM回路55は、前述したパワースペクトル取得部2の第1部分11と同様であり、遅延部21、第2窓処理部23および第2パワースペクトル計算部25を、それぞれ $n-1$ 個備える。遅延部21、第2窓処理部23および第2パワースペクトル計算部25について、それぞれ添え字(1)～( $n-1$ )を付す。遅延部21(1)～( $n-1$ )が、入力信号を遅延させる時間は、基本周期 $T_0$ の $1/n$ である。

10

20

30

40

50

## 【0116】

Nが3以上に選ばれる場合、遅延部21(k1)に与えられる入力信号は、遅延部21(k1)によって基本周期 $T_0$ の $1/n$ だけ遅延させた後、遅延部21(k1+1)に与えられる。ここでk1は、自然数である。遅延部21(k1)に与えられた入力信号は、第2窓処理部23(k1)に与えられて切り出され、第2パワースペクトル計算部25(k1)によってパワースペクトルが計算される。

## 【0117】

第1および第2パワースペクトル計算部24, 25(1)~(n-1)によって計算されたパワースペクトルは、パワースペクトル加算部26に与えられ、パワースペクトル加算部26は、各パワースペクトルを加算して、加算したパワースペクトル(出力パワースペクトル)を出力する。出力パワースペクトルは、STRAIGHT回路56に与える。

10

## 【0118】

STRAIGHT回路56は、基本周期 $T_0$ に基づいて計算された分析位置に依存しないパワースペクトル(TANDEMスペクトル)について、周波数軸上で選択的な平滑化を行うことによって、周期性による干渉の影響の無いパワースペクトル(STRAIGHTスペクトル)を生成して出力する。STRAIGHT回路56は、前述した図3に示す第2部分12の累積スペクトル計算部31および平滑化スペクトル計算部分32とを含んで構成される。

## 【0119】

図13は、基本周期計算部3の構成を示す概略ブロック図である。基本周期計算部3は、複数の基本波成分周期性計算回路51と、周期性合成回路52と、基本波候補抽出回路53とを含んで構成され、入力信号の基本周期 $T_0$ の値を求める。基本周期 $T_0$ を求めることによって、基本周波数 $f_0$ が求められる。基本周期計算部3においては、基本周波数の候補を幾つか(具体的には、たとえば1オクターブに2個、4オクターブ分)を仮定しておき、それぞれの基本周波数の候補について、基本周期の関数として基本波の周期性の評価値を求め、それらを合成し、確率的な揺らぎによる偶然とは認められない確実な基本波成分の候補を分析し、抽出して、その周波数を基本周波数の候補として出力する。前述のような基本周波数の候補は、たとえば1オクターブに2個、4オクターブ分だけ仮定する場合には、基本波成分周期性計算回路51を8個用意されている。

20

## 【0120】

図14は、基本波成分周期性計算回路51の構成を示す概略ブロック図である。基本波成分周期性計算回路51は、TANDEM回路55aと、STRAIGHT回路56aと、変動スペクトル計算部61と、空間周波数加重部62と、逆フーリエ変換部64とを含んで構成される。TANDEM回路55aは、前述したTANDEM回路55と同様の構成であり、STRAIGHT回路56aは、前述したSTRAIGHT回路56と同様の構成である。基本波成分周期性計算回路51は、基本周波数の候補について、基本周期の関数として基本波の周期性の評価値(基本波成分周期性評価値)を求める。

30

## 【0121】

入力信号は、TANDEM回路55aに与えられて、TANDEM回路55aから出力されるTANDEMスペクトルが、STRAIGHT回路56aと、変動スペクトル計算部61とに与えられる。STRAIGHT回路56aは、与えられるTANDEMスペクトルについて、周波数軸上で選択的な平滑化を行うことによって、STRAIGHTスペクトルを生成して変動スペクトル計算部61に出力する。TANDEM回路55aと、STRAIGHT回路56aとは、予め仮定している基本周波数の候補が与えられる。前述のように基本周波数の候補を、たとえば1オクターブに2個、4オクターブ分だけ仮定する場合には、この4オクターブの範囲内で、隣接する基本周波数との対数周波数上での差分が等間隔となるような8個の基本周波数を選び、これらの基本周波数が、複数の基本波成分周期性計算回路51に1つずつ与えられる。

40

## 【0122】

変動スペクトル計算部61は、TANDEM回路55aによって与えられるTANDE

50



Mスペクトルを、S T R A I G H T回路56aによって与えられるS T R A I G H Tスペクトルで除算して、数値「1」を減算する。T A N D E MスペクトルをS T R A I G H Tスペクトルで各周波数において除算し、その結果から1を引くことによって、周期性に関連する変動のみを表す変動スペクトルを求めることができる。

【0123】

変動スペクトル計算部61からの出力(変動スペクトル)を $P_C(\omega)$ とすると、 $P_C(\omega)$ は、次式17で表される。

【0124】

【数16】

$$P_C(\omega) = \frac{P_T(\omega)}{P_{TST}(\omega)} - 1 \quad \dots (17)$$

10

【0125】

式17において、 $P_T(\omega)$ は、T A N D E Mスペクトルであり、 $P_{TST}(\omega)$ は、S T R A I G H Tスペクトルを表す。 $P_{TST}(\omega)$ は、式(16)で表されている。

【0126】

変動スペクトル $P_C(\omega)$ では、窓関数による周波数方向の帯域制限と、T A N D E M窓による相対的に大きな正のバイアス項によって、基本周波数に対応する空間周波数成分が支配的となる。また実際の音声などの入力信号では、パワースペクトルは平坦ではなく、基本周波数は一定ではない。前者の影響は、正規化に用いたS T R A I G H Tスペクトルに反映されているため、第一次近似としては無視することができる。後者の影響は、 $P_C(\omega)$ の周波数方向での振幅変調として顕われる。この振幅変調の変調空間周波数は、基本周期の半分の時間を隔てた時刻における基本周波数の差に比例する。この振幅変調では最大振幅の部分が周波数0に対応しているため、周波数0を中心として高域に向けて減衰するような周波数領域での窓 $w_{\omega 0, N}(\omega)$ を乗算してフーリエ(Fourier)変換を行うことによって、実質的に無視することができる。

20

【0127】

空間周波数加重部62には、重み係数 $w_{\omega 0, N}(\omega)$ が格納されており、 $P_C(\omega)$ の低周波成分を選択する。 $P_C(\omega)$ の低周波成分は、たとえば調波が4個程度となるように選ばれている。 $w_{\omega 0, N}(\omega)$ は、次式18に示される条件を満たすように設定されており、その一例を式19に示す。

30

【0128】

【数17】

$$w_{\omega 0, N}(\omega) = \begin{cases} 0 & |\omega| > N\omega_0 \\ w_{\omega 0, N}(-\omega) & |\omega| \leq N\omega_0 \end{cases} \quad \dots (18)$$

$$\int_{-\infty}^{\infty} w_{\omega 0, N}(\omega) d\omega = 1$$

【0129】

【数18】

$$w_{\omega 0, N}(\omega) = c_0 \left( 1 + \cos \left( \pi \frac{\omega}{N\omega_0} \right) \right) \quad \dots (19)$$

40

【0130】

逆フーリエ変換部64は、 $P_C(\omega)$ に重み係数 $w_{\omega 0, N}(\omega)$ を乗算して以下の式20に示すように、フーリエ変換して、周波数軸上での周期成分 $A(\omega)$ を求める。逆フーリエ変換することによって、基本波成分周期性評価値が、基本周期の関数として求められる。

50

【 0 1 3 1 】

【 数 1 9 】

$$A(\tau; T_0) = \int_{-\infty}^{\infty} w_{\omega 0, N}(\omega) P_C(\omega; T_0) e^{-j\omega\tau} d\omega \quad \dots (20)$$

【 0 1 3 2 】

式 2 0 では、 $P_C(\ )$  を  $P_C(\ ; T_0 )$  とし、 $A(\ )$  を  $A(\ ; T_0 )$  として、T A N D E M 窓の設計に必要な情報である基本周期  $T_0$  を明記している。以下では、必要に応じてこの表記方法を用いて記載する。逆フーリエ変換部 6 4 は、前記周期成分  $A(\ )$  を、基本波成分周期性評価値として出力する。基本波成分周期性評価値は、周期性合成回路 5 2 に与えられる。

10

【 0 1 3 3 】

再び図 1 3 を参照する。基本周波数は既知ではないので、複数の基本波成分周期性計算回路 5 1 によって、基本周波数を仮定して指標を計算し、周期性合成回路 5 2 によって、以下の式 2 2 を用いて、複数の基本波成分周期性計算回路 5 1 から与えられる周期成分  $A(\ )$  を統合して、周期成分を求める。

統合した周期成分は、

【 0 1 3 4 】

【 数 2 0 】

20

$$\bar{A}(\tau) \quad \dots (21)$$

で表され、計算式は、

【 0 1 3 5 】

【 数 2 1 】

$$\bar{A}(\tau) = \frac{1}{M} \sum_{k=1}^M w_{LAG}(\tau; T_L 2^{\frac{1-k}{L}}) A\left(\tau; T_L 2^{\frac{1-k}{L}}\right), \quad \dots (22)$$

30

【 0 1 3 6 】

で表される。ここで、 $T_L$  は、基本周期の探索の初期値にあたる最長の基本周波数を表し、 $L$  は、1 オクターブ内に仮定する基本周期の個数を表す。また、 $w_{LAG}(\ ; T_C )$  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。また式 2 2 のピークは、そのピーク付近の形状が放物線でよく近似できることを利用して、ピークを含む 3 点を用いた放物線補間によって求められることができる。

【 0 1 3 7 】

基本周期は、周期成分である式 2 1 が、 $\tau = T_C$  の場合に最大値をとることを利用することによって得られる。まずは、そのような性質を与えるためのパラメタを決定しておく。ある基本周期  $T_C$  を仮定した場合の  $A(\ ; T_0 )$  の振舞いを調べると、 $T_C$  を仮定して求めた  $A(\ ; T_0 )$  は、抽出したい成分以外のランダムな成分に起因するパワースペクトルの周波数軸上での変動も抽出してしまう。この抽出されるこの不要な成分と、本来の目的である周期成分との  $S/N$  比が最大になるように、T A N D E M 分析に用いる時間窓のサイズを設定する。具体的には、Blackman窓を用いた場合、長さが仮定する周期  $T_C$  の 4 倍において  $S/N$  比が最大となる。この条件の下で、重み関数  $w_{LAG}(\ ; T_C )$  を設計する。設計の目標は、元の窓のサイドローブや、長過ぎる時間窓を用いることでパワースペクトル上での空間周波数成分に生ずる非線形歪みに起因する不要なピークを、重み関数  $w_{LAG}(\ ; T_C )$  を用いて抑圧することである。重み関数の選定では、式 2 0 によって統合した結果が周波数方向で大きく変動しないことと、配置すべき帯域の数が多くなり過ぎないことの双方の条件を考慮する必要がある。ここでは、具体的な関数として以下

40

50

の式 2 3 を示す。帯域の配置密度は、オクターブ当たり 2 個としている。以下の式 2 3 の関数の定義域の幅は 2 オクターブであり、十分に重なり合っている。

【 0 1 3 8 】

【 数 2 2 】

$$w_{LAG}(\tau; T_0) = 0.5 + 0.5 \cos\left(\pi \log_2\left(\frac{\tau}{T_0}\right)\right), \quad \dots (23)$$

【 0 1 3 9 】

このようにして式 2 0 によって最終的に求められた式 2 1 のピークの分布は、注目する帯域では、ランダム入力に対するピーク値は周波数に依存していないものとなる。したがって、入力がランダムであると仮定した場合のピークの出現確率をピーク値の関数として表すことができる。図 1 5 に、ピークの出現確率をピーク値の関数として表すグラフの一例を示す。図 1 5 において、横軸は周期性の指標の値を示し、縦軸はランダムな揺らぎによって生じたピークを周期信号が存在する証拠であると誤判定する危険率を表わす。また図 1 5 には、2 次関数による近似曲線も示している。また窓関数には、Blackman を用いている。図 1 5 から見ても判るように、危険率として 1 % を許容する場合には、判定の閾値を 1 . 1 9 に設定すればよく、0 . 1 % の場合には、1 . 4 1、0 . 0 1 % の場合には、1 . 5 5 に設定すれば良いことが分かる。基本波候補抽出回路 5 3 では、判定の閾値が設定されており、この判定の閾値に基づいて精度の高い基本周波数を抽出する。

【 0 1 4 0 】

こうして求められる周期成分には、基本周期に相当するピークのみが存在し、半ピッチおよび倍ピッチの誤りは生じない。入力信号である音声であるとき、実際に声帯の振動にサブハーモニックが生じている場合には、繰返しの構造に応じて、複数の基本周期に対応するピークが生じる。

【 0 1 4 1 】

基本波候補抽出回路 5 3 は、周期性合成回路 5 2 によって求められた周期成分のピークのうち、どのピークに対応する基本周期に応じた基本周波数を抽出するのかが選択する。この選択は、利用者によって設定することができる。たとえば入力信号が音声の場合、最も大きな基本周波数のみを選択したり、最も大きな基本周波数と、この基本周波数の 2 分の 1 程度または 3 分の 1 程度となる基本周波数を選択したりする。最も大きな基本周波数と、この基本周波数の 2 分の 1 程度または 3 分の 1 程度となる基本周波数を選択する場合には、だみ声に含まれる複数の基本周波数を抽出することができる。このように基本周期計算部 3 では、単一の基本周波数を求められるだけでなく、基本周波数としての要件を備えている周波数が複数ある場合にも、その複数の周波数を抽出することができる。基本波候補抽出回路 5 3 は、選択した基本周波数を出力する。基本波候補抽出回路 5 3 から出力された基本周波数は、T A N D E M 回路 5 5、S T R A I G H T 回路 5 6 および非周期成分計算回路 5 4 に与えられ、与えられる基本周波数に応じて、これらの回路において用いられる基本周期  $T_0$  が設定される。

【 0 1 4 2 】

図 1 6 は、非周期成分計算回路 5 4 の構成を示す概略ブロック図である。非周期成分計算回路 5 4 は、入力信号の非周期成分を分析して求める。非周期成分は、基本周波数の軌跡と S T R A I G H T スペクトルの系列が既知であるとして、瞬時周波数としての基本周波数の逆数に比例して時間軸を伸縮させることによって見かけ上の基本周波数を一定とし、S T R A I G H T スペクトルの系列を用いて各周波数における分析区間内のスペクトルの変動を除いて、見かけ上一定である基本周波数から構成される直交位相信号を、この新しく時間軸を伸縮することで求められた周期信号から求められる変動スペクトルと畳込み、その結果求められる複素スペクトルの振幅として周期成分の相対的大きさを求め、それと T A N D E M スペクトルの計算に用いた窓関数固有の定数として求まる値に基づいて求められる。

## 【 0 1 4 3 】

非周期成分計算回路 5 4 は、時間軸変換部 7 1 と、T A N D E M回路 5 5 b と、S T R A I G H T回路 5 6 b と、変動スペクトル計算部 6 1 a と、直交位相畳込み部 7 3 と、非周期性計算部 7 4 とを含んで構成される。

## 【 0 1 4 4 】

時間軸変換部 7 1 は、入力信号について、基本周波数の瞬時周波数に反比例する割合で時間軸を伸縮することによって、見かけ上一定の基本周期の周波数を有する信号となるように変換する。時間軸変換部 7 1 は、目標となる設定周波数を分子として、現在の入力信号の周波数を分母において除算することによって、基本周波数の瞬時周波数に反比例する割合を求め、この割合で入力信号の周波数に乗算する。

10

## 【 0 1 4 5 】

さらに具体的には、時間的に変化する信号  $s(t)$  の基本周波数の瞬時周波数を  $f_0(t) = \omega_0(t) / 2\pi$  とすると、(振幅を無視した)基本波成分の波形  $s_0(t)$  は、以下の式 2 4 のように表される。なお、ここでは、基本波の位相  $\varphi(t)$  は、式 2 5 で表され、初期値を 0 と置いた。

## 【 0 1 4 6 】

## 【数 2 3】

$$s_0(t) = \sin \varphi(t) \quad \dots (24)$$

20

## 【 0 1 4 7 】

## 【数 2 4】

$$\varphi(t) = \int_0^t \omega_0(\tau) d\tau \quad \dots (25)$$

## 【 0 1 4 8 】

ここで、位相が一定の速度  $2\pi f_{TGT}$  で変化する場合の時間軸と解釈できる次の量  $\lambda(t)$  を式 2 6 によって求めておく。

## 【 0 1 4 9 】

## 【数 2 5】

$$\lambda(t) = \frac{\varphi(t)}{2\pi f_{TGT}} \quad \dots (26)$$

30

## 【 0 1 5 0 】

この時間軸を用いて  $s_0(t)$  を  $\lambda(t)$  の関数として表すと、その瞬時周波数は定数  $f_{TGT}$  となるのがわかる。したがって、基本周波数が分かっている信号があれば、それを式 2 6 の時間軸の上で表現することによって、固定した基本周波数定数  $f_{TGT}$  を有する信号に変換することができる。

40

## 【 0 1 5 1 】

T A N D E M回路 5 5 b は、前述した T A N D E M回路 5 5 と同様の構成であり、S T R A I G H T回路 5 6 b は、前述した S T R A I G H T回路 5 6 と同様の構成である。時間軸変換部 7 1 によって、時間軸が変換された入力信号は、T A N D E M回路 5 5 b に与えられて、T A N D E M回路 5 5 b から出力される T A N D E Mスペクトルが、S T R A I G H T回路 5 6 b と、変動スペクトル計算部 6 1 a とに与えられる。S T R A I G H T回路 5 6 b は、与えられる T A N D E Mスペクトルについて、S T R A I G H Tスペクトルを生成して変動スペクトル計算部 6 1 a に出力する。

## 【 0 1 5 2 】

変動スペクトル計算部 6 1 a は、変動スペクトル計算部 6 1 と同様の構成であり、T A

50

NDEM回路55bによって与えられるTANDEMスペクトルと、STRAIGHT回路56bによって与えられるSTRAIGHTスペクトルで除算して、数値「1」を減算して、得られた変動スペクトルを直交位相信号畳込み部73に与える。

【0153】

基本波が分かれば、前述したように時間軸を変換することによって基本周波数が任意の定数となるような信号に変換することができる。この任意に設定できる値を  $f_c = c / 2 = 1 / T_c$  と表すこととする。非周期成分計算回路54では、結局、この基本周波数成分についてだけ、非周期性を評価すればよいことになる。ただし、複数の基本周波数の候補がある場合や、サブハーモニックがある場合には、それらの周波数についても、併せて評価する必要がある。

10

【0154】

まず基本周波数成分による周波数軸上での周期構造の強さを調べるために、次式27のような直交位相信号を作成する。

【0155】

【数26】

$$h_N(\omega; T_C) = w_{\omega_C, N}(\omega) \exp(2\pi j\omega/\omega_C). \quad \dots (27)$$

【0156】

ここで、 $w_{\omega_C, N}(\omega)$  は、周期構造を調べる場合に用いる空間周波数方向での振幅包絡であり、たとえば、raised cosine 型の関数を用いて、式28に示すように表すことができる。

20

【0157】

【数27】

$$w_{\omega_C, N}(\omega) = c_0 (1 + \cos(\pi\omega/N\omega_C)) \quad \dots (28)$$

【0158】

この直交位相信号を用いて、変動スペクトル  $P_C(\omega; T_C)$  の中に含まれる  $\omega_C$  の速度で変動する成分の強さを表す

30

【0159】

【数28】

$$\tilde{\sigma}_{P.obs}^2(\omega; T_C)$$

【0160】

を計算する。まず、 $P_C(\omega; T_C)$  は、式17と同様であるが、以下の式29によって表される。

【0161】

【数29】

40

$$P_C(\omega; T_C) = \frac{P_T(\omega; T_C)}{P_{TST}(\omega; T_C)} - 1 \quad \dots (29)$$

【0162】

ここで、 $P_C(\omega; T_C)$  はTANDEMスペクトルを表し、 $P_{TST}(\omega; T_C)$  はSTRAIGHTスペクトルを表す。 $T_C$  は、用いた基本周期を明示するために付記されたものである。なお、非周期性の評価のために用いるTANDEMの計算にあたっては、 $f_0$  の推定の場合と同様に、周期性をもっとも良く評価できるように最初に用いる時間窓を設定する必要がある。たとえば、 $T_C$  の4倍の長さのBlackman窓を用いる。

【0163】

50

この変動スペクトル  $P_C(\omega; T_C)$  に、先ほどの直交位相信号  $h_N(\omega; T_C)$  を畳込むことにより、元の信号の周期性によって生ずる周波数軸上の周期性の強さを求めることができる。観測される信号であるので、

【 0 1 6 4 】  
【 数 3 0 】

$$\tilde{\sigma}_{P,obs}^2(\omega; T_C)$$

で表すことにする。

【 0 1 6 5 】

観測される信号には、本来の周期成分によるもの  $\tilde{\sigma}_{P,obs}^2(\omega)$  と、非周期成分の中で直交位相信号  $h_N(\omega; T_C)$  によって拾われるもの

【 0 1 6 6 】  
【 数 3 1 】

$$\varepsilon_{w_N} \tilde{\sigma}_N^2(\omega)$$

の両方が含まれている。ここで、

【 0 1 6 7 】  
【 数 3 2 】

$$\tilde{\sigma}_N^2$$

【 0 1 6 8 】

は、非周期成分の分散、 $\varepsilon_{w_N}$  は、直交位相信号によって非周期成分が拾われる割合を表す。 $\varepsilon_{w_N}$  は、包絡  $w_{C,N}(\omega)$  によって定まる。観測される信号は、式 3 0 によって表される。

【 0 1 6 9 】  
【 数 3 3 】

$$\begin{aligned} \tilde{\sigma}_{P,obs}^2(\omega; T_C) &= \left| \int_{-\infty}^{\infty} h_N(\lambda; T_C) P_C(\omega - \lambda; T_C) d\lambda \right|^2 \\ &= \sigma_{P,obs}^2(\omega) + \varepsilon_{w_N} \tilde{\sigma}_N^2, \end{aligned} \quad \dots (30)$$

【 0 1 7 0 】

これらは、直接観測することのできない量であるため、以下のように幾つかの近似を用いることにより、観測できる量から求める計算法を導く。直交位相信号による畳込みを、記号「 $\circ$ 」を用いて表すことにする。畳込みの結果の絶対値として得られる評価値（観測値）を  $Q_C$  と表すと、 $Q_C^2$  は、式 3 1 で与えられる。この  $Q_C^2$  の値は、式 3 0 と同じものを表す。

【 0 1 7 1 】  
【 数 3 4 】

$$\begin{aligned} Q_C^2 &= |h_N \circ P_C(\omega; T_C)|^2 \\ &= \left| h_N \circ \frac{P_T(\omega; T_C)}{P_{TST}(\omega; T_C)} - 1 \right|^2 \\ &= \left| h_N \circ \frac{P_T(\omega; T_C) - P_{TST}(\omega; T_C)}{P_{TST}(\omega; T_C)} \right|^2 \end{aligned} \quad \dots (31)$$

【 0 1 7 2 】

10

20

30

40

50

TANDEMスペクトルは、STRAIGHTスペクトルに $h_N$ で選択的に除去される周期的な変動分が加えられたものであること、また、その周期的な変動分には、信号の周期性に起因するものと、信号のランダムな変動に起因するものが含まれることに注意する。ここで、 $P_P$ を、信号の周期性に起因する変動分、 $P_R$ をランダムな変動に起因する変動分、 $P_P$ を、周期成分のSTRAIGHTスペクトル、 $P_R$ を、ランダムな成分のSTRAIGHTスペクトルと表記することとする。

【0173】

ここで、 $h_N$ の定義域の幅の中では $P_P(\ ; Tc)$ と $P_R(\ ; Tc)$ が定数と見なせると仮定する。すると、次式32が得られる。

【0174】

【数35】

$$Q_C^2 = \frac{V[h_N \circ \Delta P_P]}{P_P + P_R} + \frac{V[h_N \circ \Delta P_R]}{P_P + P_R} \quad \dots (32)$$

【0175】

周期信号の場合には、窓関数が決まれば、 $V[h_N \circ P_P]$ の値は、 $P_P$ の定数 $C_P$ 倍として一意に決まり、ランダムな成分の値 $V[h_N \circ P_R]$ についても、窓関数と $h_N$ が決まれば実効的なTB積から、 $P_R$ の定数 $C_R$ 倍として(期待値であるなら)一意に決まる。結局、以下の式33が得られる。

【0176】

【数36】

$$Q_C^2 = \frac{C_P P_P}{P_P + P_R} + \frac{C_R P_R}{P_P + P_R} \quad \dots (33)$$

【0177】

周期成分の自乗平均値の意味での平均振幅を $aPRD(\ )$ とし、非周期成分の平均振幅を $aRND(\ )$ とすると、以下の式34のように表される。

【0178】

【数37】

$$a_{RND}(\omega) = \sqrt{\frac{C_P - Q_C^2}{C_P - C_R}} \quad \dots (34)$$

$$a_{PRD}(\omega) = \sqrt{\frac{Q_C^2 - C_R}{C_P - C_R}}$$

【0179】

直交位相信号畳込み部73は、見かけ上一定である基本周波数から構成される直交位相信号と、変動スペクトル計算部61aから与えられる変動スペクトルとを畳込んで絶対値を求める。

【0180】

非周期性計算部74は、直交位相信号畳込み部73の演算結果から、周期成分の自乗平均値の意味での平均振幅を $aPRD(\ )$ 、および非周期成分の平均振幅を $aRND(\ )$ を求めて非周期成分評価値として出力する。この2つ、すなわち $aPRD(\ )$ と $aRND(\ )$ とが、音声の診断用の情報として利用されたり、音声合成の際に、パルス成分の帯域毎のパワーの決定とランダム成分の帯域毎のパワーの決定に用いられる。

【0181】

平滑化スペクトル変換部4、音源情報変換部5および位相調整部6を含むパラメータ変換部は、非周期成分計算回路54から与えられる非周期成分評価値をも考慮して、パラメータを調整する。非周期成分評価値は、音声合成で品質を向上するため用いられる。非周期成分評価値は、音声合成で品質を向上するため用いられる。非周期成分評価値は、平滑

10

20

30

40

50

化スペクトルの重みとして用いることにより、雑音により駆動されるフィルタ形状を決定するために用いられ、また、その残りとして周期信号により駆動されるフィルタ形状を決定したりして用いられる。

#### 【0182】

前述した a P R D ( ) と a R N D ( ) との計算のためには、測定で求められる値  $Q_c^2$  の他に、T A N D E M で用いる窓により決まる  $C_p$  と、分析条件により変化する  $C_R$  の統計的な性質が必要となる。たとえば、基本周期の 2 . 4 倍の Blackman 窓を用いた分析では、シミュレーションの設定により若干の差があるが、 $C_p = 0 . 56$  が得られた。ランダムな成分に対する係数  $C_R$  は、直交位相信号  $h_N ( ; T c )$  の周波数方向の広がりを表す  $N$  に依存する。図 17 に  $N = 2$  の場合と  $N = 16$  の場合の観測値  $Q_c$  の分布を示す。図 17 ( 1 ) が  $N = 2$  の場合であり、図 17 ( 2 ) が  $N = 16$  の場合である。図 17 では横軸に周期性を示し、縦軸に観測値を示す。図から明らかなように  $N = 2$  の場合には、分布が大きく広がっている。これは、実際の信号の分析においても推定値の分散が大きくなることを意味している。

10

#### 【0183】

この問題を避けるには、複数の分析フレームでの結果を平均することにより T B 積を大きくすることが必要となる。本実施の形態では、実際に利用される可能性のある範囲をカバーするように、分析フレーム周期、周波数方向の広がり  $N$ 、統合する分析フレーム数の全ての組み合わせについてシミュレーションを行い  $Q_c$  を求め、その平均値と分散を 3 次元のテーブルとして格納する。必要な  $C_R$  の値は、このテーブルから線形補間によって求める。実際の計算では、 $C_R$  の値は、該当する条件の  $Q_c$  の平均値に、 $Q_c$  の標準偏差の定数倍を加えたものとする。定数の具体的な値は、主観評価実験および評価値の consistency の条件を最適化する客観評価を用いたシミュレーション等によって決定する。

20

#### 【0184】

式 34 の  $Q_c$  にはランダムな成分が含まれているため確率的に揺らぐ。そのため、このまま使用した場合には、負のパワーや 100 % を超える非周期成分などの不合理な値となる場合がある。ここでは式 36 の根号内の値  $x$  を次式 35 によって変換する。

#### 【0185】

#### 【数38】

$$g(x) = \frac{1}{\alpha} \log \frac{1 + \exp(-\alpha x)}{1 + \exp(-\alpha(x-1))} + 1 \quad \dots (35)$$

30

#### 【0186】

ここで は、ソフトさを定める値であり、受聴試験などにより定める。

以上のように、周期信号変換装置 50 では、入力信号である音声信号の基本周波数が長くなったり短くなったりしても、そのときの基本周波数に応じた基本周波数を求めることができる。基本周波数が変化しても T A M D E M 窓の幅が基本周期に追従して短くなるので、基本周波数が変化しても基本周波数を正確に求めることができる。したがって、このような基本周波数を用いて、合成音または変換音が生成されるので、基本周波数に応じて適切なサイズの時間窓を選択すれば、音声合成に用いた場合に元の信号と同じ基本周波数が抽出されるような信号を合成することができ、合成音および変換音の品質を向上させることができる。また、抽出された基本周波数を用いて合成した信号を再分析した場合であっても、合成に用いたものと同じ基本周波数が得られるように設計することができる。また、複数の基本周波数を有する信号を適切に分析することができるようになるので、これまで分析合成を適切に行うことができなかつたみ声の分析合成が可能となる。

40

#### 【0187】

また基本周波数の時間変化およびスペクトルの時間変化の影響を非周期成分として抽出することが防止できるので、合成の際に用いるべき正確な基本周波数を抽出することができる。合成音声および加工音声の品質を改善することができる。また本発明では、非周期成分の推定方法に根拠の曖昧な非線形処理を含まないため、声を利用した医療診断などに

50



応用することができる。また、基本周波数の時間変化およびスペクトルの時間変化の影響を除いて非周期成分を求めることができ、合成の際に用いるべき正確な非周期性の値を抽出することができる。

【 0 1 8 8 】

また周期信号変換装置 5 0 では、基本波成分と、非周期成分とについて、確率として解釈できる評価指標を得ることができる。さらに周期信号変換装置 5 0 を実現するにあたって、実際の演算において、高速フーリエ変換を多用することができるため、高速な分析および合成を実現することができる。

【 0 1 8 9 】

前述した周期性合成回路 5 2 において求められる周期性のピークには、最初の T A N D E M 時間窓による時間遅れの関数となる窓がかかっているため、ピーク位置は短い時間遅れの方に偏る。周期性合成回路 5 2 は、この初期推定値を瞬時周波数を求めることによって改良してもよい。瞬時周波数の計算には Flanagan の式を用いる。直交位相信号を用いて、ある角周波数  $\omega_0$  における短時間 Fourier 変換の値  $X(\omega_0)$  を求めることができる。具体的には、式 ( 2 7 ) と同様の直交位相信号を作成する。 $X(\omega_0)$  を次のように虚数部と実数部を用いて表すこととする。

$$X(\omega_0) = a + j b \quad \dots (36)$$

この標記の下で、Flanagan の式は次式 3 7 のように表される。

【 0 1 9 0 】

【 数 3 9 】

$$\lambda(\omega) = \omega + \frac{a \frac{\partial b}{\partial t} - b \frac{\partial a}{\partial t}}{a^2 + b^2} \quad \dots (37)$$

ここで、Fourier 変換の次式 3 8 の性質を利用する。

【 0 1 9 1 】

【 数 4 0 】

$$\frac{\partial \mathcal{F}[x(t)]}{\partial t} = \mathcal{F}[tx(t)] \quad \dots (38)$$

【 0 1 9 2 】

具体的には、基本周波数の初期推定値  $\omega_0$  を用いて上記の直交位相信号を作成し、それを用いて  $\omega_0$  における瞬時周波数  $\lambda(\omega_0) = \lambda(\omega_0)$  を求める。こうして求めた瞬時周波数は、初期推定値よりも基本周波数の真値に近づいていると期待できる。しかし、初期推定値がバイアスを含むために、瞬時周波数にも、一般的にバイアスが残る。正しい周波数は、周波数から瞬時周波数への写像の不動点として求められる。そこで、初期推定値とは異なるもう一つの初期値  $\omega_1 = \omega_0$  に対応する瞬時周波数  $\lambda(\omega_1)$  を、同様に求めるとすると、以下の関係式 3 9 が成立する。

【 0 1 9 3 】

【 数 4 1 】

$$\begin{bmatrix} \lambda_0 \\ \lambda_1 \end{bmatrix} = \begin{bmatrix} \omega_0 & 1 \\ \omega_1 & 1 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} \quad \dots (39)$$

【 0 1 9 4 】

この式 3 9 から、係数行列の逆行列を求められた 2 つの瞬時周波数からなるベクトルに掛けることによって、周波数から瞬時周波数への写像の一次関数近似の係数  $u_0, u_1$  が求められる。ここで、不動点の条件  $\lambda(\omega) = \omega$  (もう一つの条件は、とりあえず外しておく) を用いることにより、改良された基本周波数の推定値  $\omega_{r1}$  を  $u_0, u_1$  を用いて以下の式 4 0 によって求めることができる。

【 0 1 9 5 】

10

20

30

40

50

## 【数 4 2】

$$\omega_{r1} = \frac{u_1}{1 - u_0} \quad \dots (40)$$

## 【0196】

このように求められた改良された基本周波数の推定値  $r_1$  を初期値として、更に、この初期値を挿むように上下の周波数における瞬時周波数を式 29 によって求めて、式 31 と式 32 とを計算することによって、更に改良された推定値  $r_2$  を求めることができる。基本周波数は誤差を含んでいるが、前述したように推定値を改良すれば、1 回の修正で、誤差を 1% 程度以下にすることができ、2 回の修正で、誤差を 0.2% 程度以下に

10

## 【0197】

また評価値と誤判定危険率との関係を定めておけば、基本波成分周期性評価値と、非周期成分評価値とを取得し、前記関係から、基本周波数がどれだけ信頼できるのかを求めることができる。たとえば、入力信号の基本周波数が「 $XX$ 」Hz であり、この基本周波数の誤判定危険率「 $XX$ 」% という情報を出力すれば、分析した基本周波数の信頼性を容易に判断することができる。評価値と誤判定危険率との関係は、基本周波数の抽出の仕組みができれば、実際にシミュレーションを行うことによって求めておけばよい。

## 【0198】

図 18, 図 19 および図 20 は、音声信号を基本周期計算部 3 によって分析した結果の一例を示す図である。ここでは、男性の発声した日本語連続母音「あいうえお」を試料として、周期成分(式 22)を各時刻について求めている。試料の標本化周波数は 22050 Hz である。ここでは、周期成分(式 22)のふるまいを詳しく調べるため、分析は 1ms 毎に行った。なお、仮定する基本周期の個数は 1 オクターブに 2 個とし、最長の基本周期を 32ms として全体で 9 個の基本周期を仮定した。図 18 は、直交位相信号の長さ  $N$  が、10 の場合の分析結果を示す。図 18 は、分析結果を濃淡画像で示したものであり、その横軸は時間を示し、縦軸は遅れ時間を示す。また図 18 では、周期性が強い部分ほど濃度が薄く(白く)なるように表している。基本周期に相当する時間遅れは、この図 18 から明瞭に分かる。図 19 は、それぞれの時刻について、周期性が極大値を示す位置を示したものである。図 19 では、横軸は時間を表し、縦軸は図 18 とは異なり、周波数(時間遅れの逆数)を表している。図 19 には、図中に印を用いて、周波数の最大値の軌跡を示している。図 19 を見ると、母音の開始部分と終了部分の一部を除き、正しく基本周波数が抽出されていることが分かる。図 20 は、それぞれの時刻における全ての極大値を示したものである。図 20 を見ると、基本波成分が卓越していることと、2 位の成分が目立っていることが分かる。

20

30

## 【0199】

また図 21 は、音声信号を非周期成分計算回路 54 によって分析した結果の一例を示す図である。音声信号の資料は、前述したものと同様である。図 21 は、分析結果を濃淡画像で示したものであり、その横軸は時間を示し、周波数を示す。また、非周期成分が強い部分ほど濃度が薄く(白く)なるように示している。

40

## 【0200】

以上では、周期信号変換装置 1, 50 について述べたが、本発明は、音声合成および音声変換以外にも、(a) 音声分析合成システムまたは音声符号化装置における基本周波数情報の抽出、(b) 音声分析合成システムまたは音声符号化装置における非周期性情報の抽出、音声認識システムにおける音声信号の検出、(c) 音声アーカイブへの付加情報(アノテーション)の付与における音声信号の検出ならびに基本周波数情報の抽出、(d) 鼻歌などによる楽曲検索システムにおける基本周波数情報の抽出、および(e) 声による発声障害の診断における音源情報(基本周波数および非周期性)の抽出、などに適用することができる。

## 【0201】

50

たとえば録音機に、前述した基本周期計算部 3 を備え、マイクロフォンで取得した音声信号から基本周波数を抽出して、人の声の周波数と一致するか否かを判断すれば、マイクロフォンのまわりで人の声がしているかどうかを判断して、人の声がしているときには、自動的に録音させる構成としてもよい。また本発明を用いれば、マイクロフォンで取得した音声信号から基本周波数を抽出して、人の声の周波数と一致するか否かを判断すれば、音声信号のうち、人がしゃべっているところを取り出すことができる。また本発明を用いれば、入力された信号が、完全にランダムな雑音であるか、周期性のあるものであるかどうかを検出することができる。また、本発明を用いれば、音声信号に含まれる基本周波数を正確に求めることができるので、声帯の異常の有無を判断することができる。

【 0 2 0 2 】

10

また本発明の他の実施の形態では、前述した実施の形態において組み合わせが可能な部分については組み合わせてもよく、たとえば、S T R A I G H T 回路 5 6 は、前述した図 3 に示す第 2 部分 1 2 および第 3 部分 1 3 を含んで構成され、最適時間周波数平滑化パワースペクトルを出力してもよい。

【図面の簡単な説明】

【 0 2 0 3 】

【図 1】本発明の実施の一形態の音声変換方法を実現するための周期信号変換装置 1 を示す概略ブロック図である。

【図 2】周期信号変換装置 1 が有するパワースペクトル取得部 2 を示す概略ブロック図である。

20

【図 3】周期信号変換装置 1 が有するパワースペクトル取得部 2 を示す概略ブロック図である。

【図 4】周期信号変換装置 1 が有するパワースペクトル取得部 2 を示す概略ブロック図である。

【図 5】入力信号である音声波形を示すグラフである。

【図 6】窓関数を示すグラフである。

【図 7】第 1 および第 2 パワースペクトル計算部 2 4 , 2 5 によって求められたパワースペクトルの一例を示すグラフである。

【図 8】パワースペクトル加算部 2 6 から出力される出力パワースペクトルの一例を示すグラフである。

30

【図 9】第 1 および第 2 平滑化スペクトル計算部 3 2 , 3 3 の各部から出力される平滑化されたパワースペクトルの一例を示すグラフである。

【図 10】最適周波数補償合成部 3 6 から出力される最適周波数平滑化対数パワースペクトルの一例を示すグラフである。

【図 11】本発明の実施の他の形態の音声変換方法を実現するための周期信号変換装置 5 0 を示す概略ブロック図である。

【図 12】T A N D E M 回路 5 5 の構成を示す概略ブロック図である。

【図 13】基本周期計算部 3 の構成を示す概略ブロック図である。

【図 14】基本波成分周期性計算回路 5 1 の構成を示す概略ブロック図である。

【図 15】ピークの出現確率をピーク値の関数として表すグラフの一例を示す。

40

【図 16】非周期成分計算回路 5 4 の構成を示す概略ブロック図である。

【図 17】 $N = 2$  の場合と  $N = 16$  の場合の観測値  $Q_c$  の分布を示す。

【図 18】音声信号を基本周期計算部 3 によって分析した結果の一例を示す図である。

【図 19】音声信号を基本周期計算部 3 によって分析した結果の一例を示す図である。

【図 20】音声信号を基本周期計算部 3 によって分析した結果の一例を示す図である。

【 0 2 0 4 】

【図 21】音声信号を非周期成分計算回路 5 4 によって分析した結果の一例を示す図である。

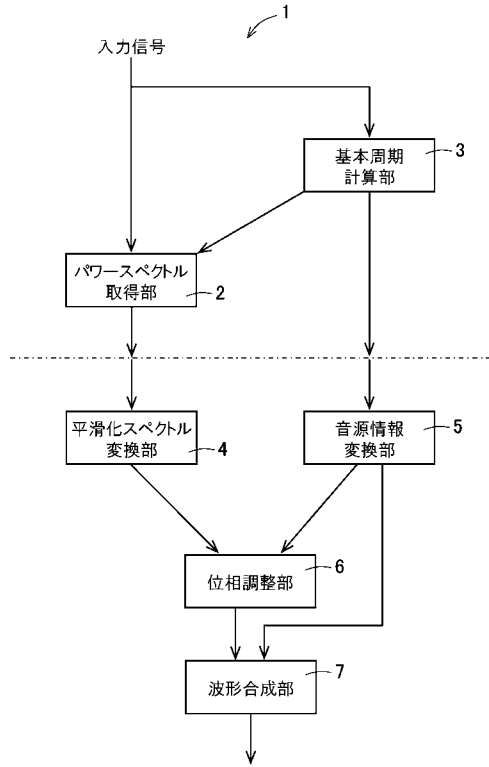
【符号の説明】

【 0 2 0 5 】

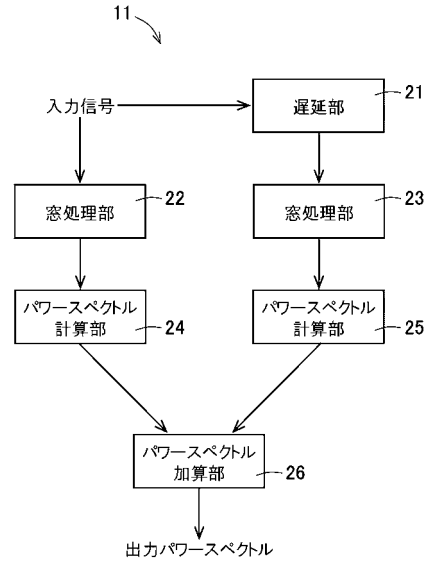
50

1 , 5 0	周期信号変換装置	
2	パワースペクトル取得部	
3	基本周期計算部	
4	平滑化スペクトル変換部	
5	音源情報変換部	
6	位相調整部	
7	波形合成部	
2 1	遅延部	
2 2	第 1 窓処理部	
2 3	第 2 窓処理部	10
2 4	第 1 パワースペクトル計算部	
2 5	第 2 パワースペクトル計算部	
2 6	パワースペクトル加算部	
3 1	累積パワースペクトル計算部	
3 2	第 1 平滑化スペクトル計算部	
3 3	第 2 平滑化スペクトル計算部	
3 4	対数変換部	
3 5	第 2 対数変換部	
3 6	最適周波数補償合成部	
4 1	3 フレーム蓄積部	20
4 2	最適時間補償合成部	
4 3	指数変換部	
4 4	第 1 蓄積部	
4 5	第 2 蓄積部	
5 1	基本波成分周期性計算回路	
5 2	周期性合成回路	
5 3	基本波候補抽出回路	
5 4	非周期成分計算回路	
5 5 , 5 5 a , 5 5 b	T A N D E M 回路	
5 6 , 5 6 a , 5 6 b	S T R A I G H T 回路	30
6 1 , 6 1 a	変動スペクトル計算部	
6 2	空間周波数加重部	
6 4	逆フーリエ変換部	
7 1	時間軸変換部	
7 3	直交位相畳込み計算部	
7 4	非周期性計算部	

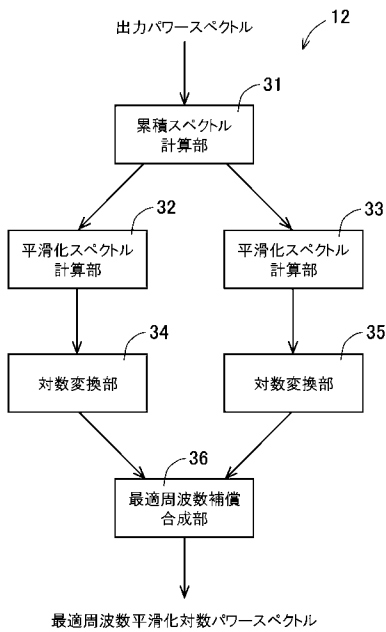
【図1】



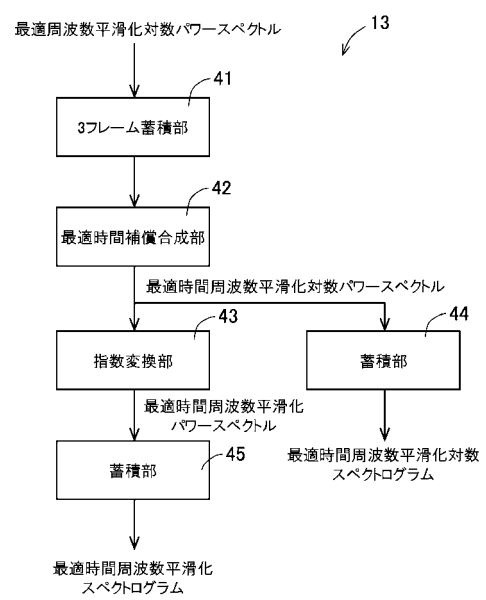
【図2】



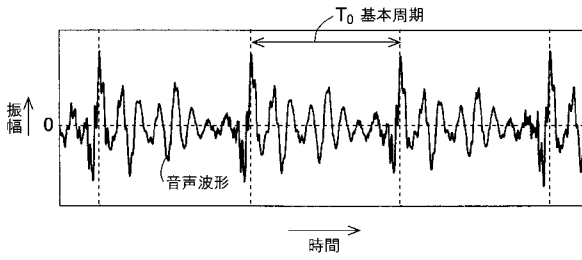
【図3】



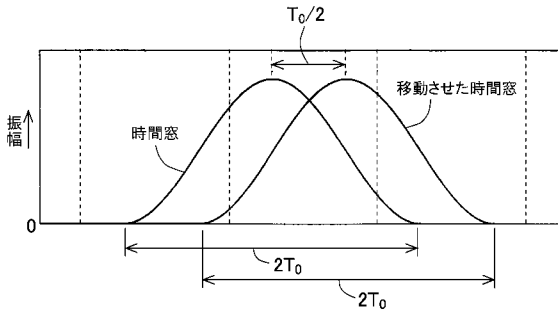
【図4】



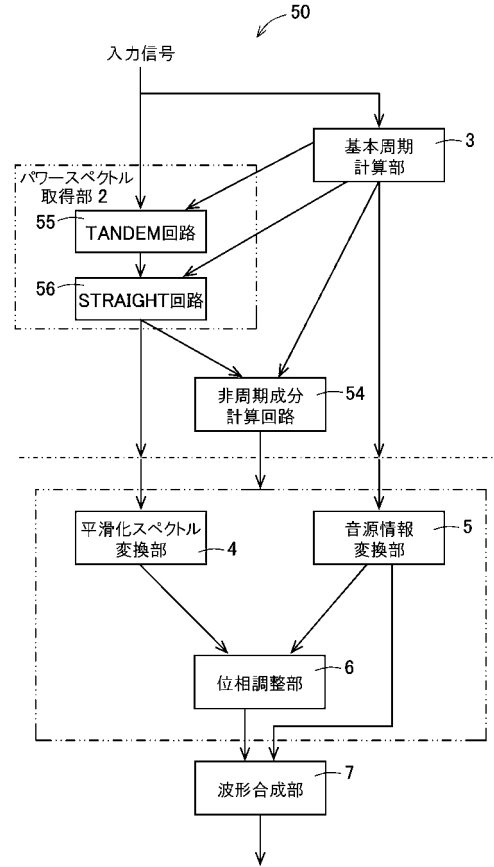
【図5】



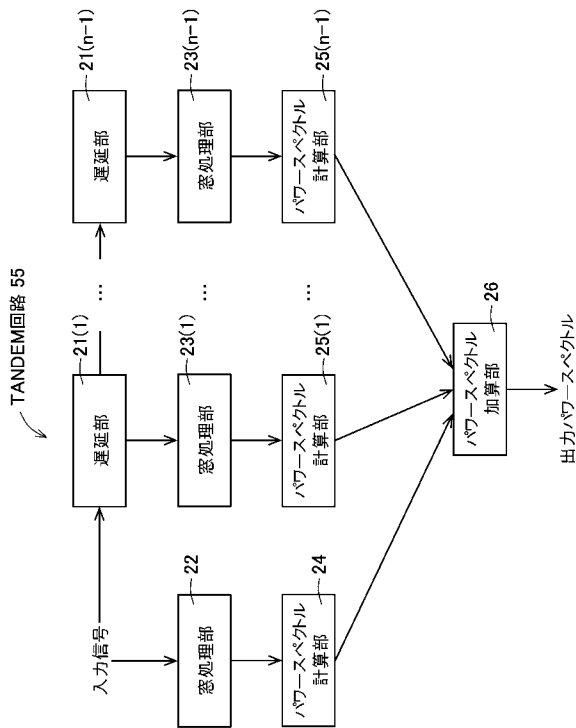
【図6】



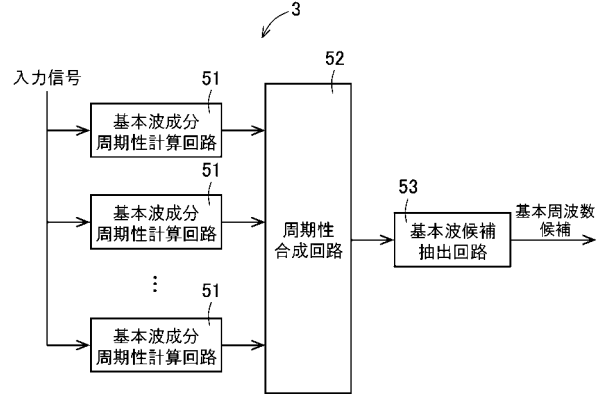
【図11】



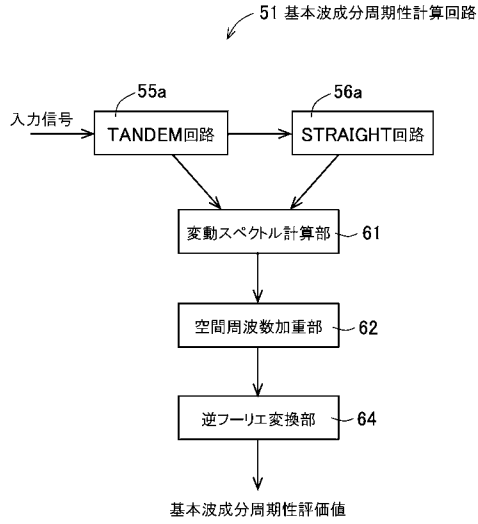
【図12】



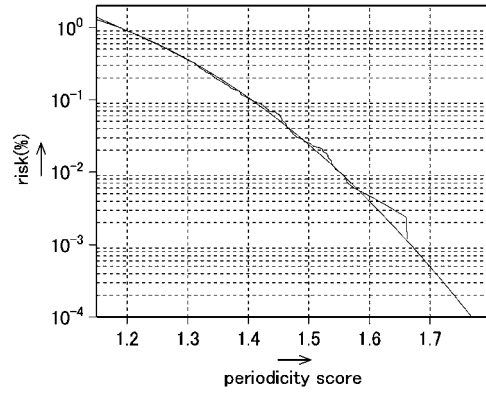
【図13】



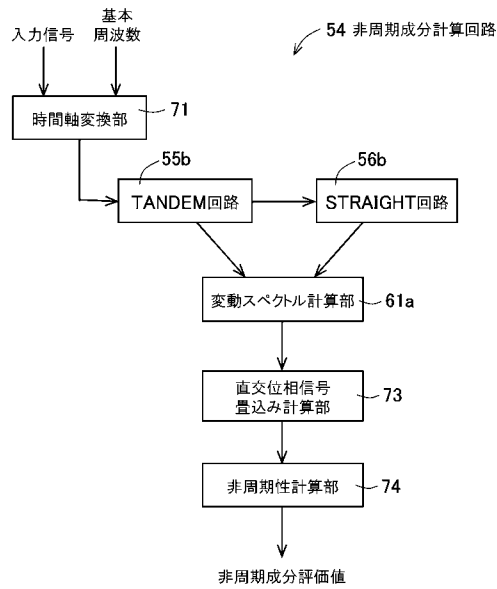
【図14】



【図15】

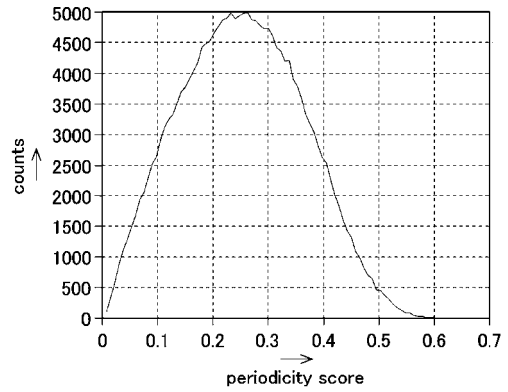


【図16】

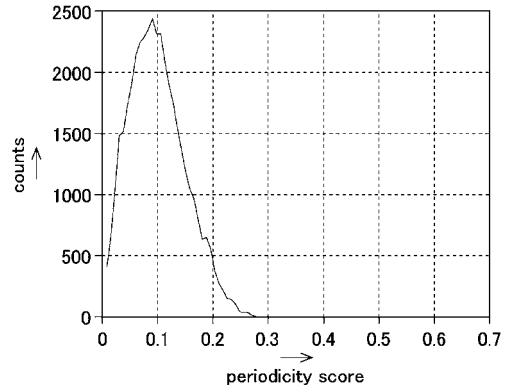


【図17】

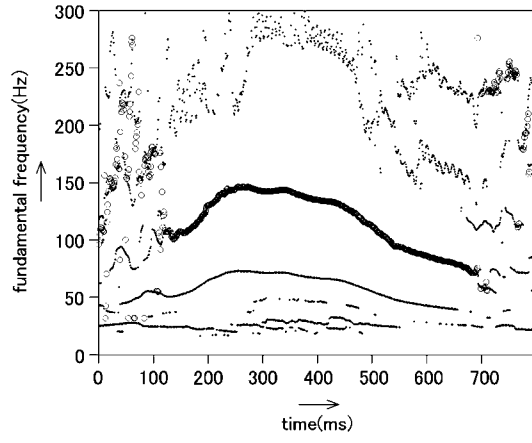
(1) noise nCycles=2 framePeriod=1(ms) nSmoothing=1 frames



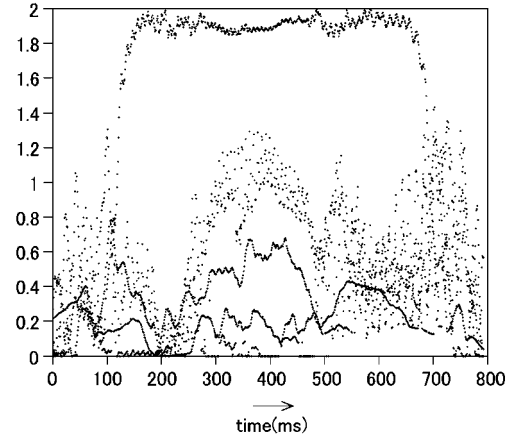
(2) noise nCycles=16 framePeriod=5(ms) nSmoothing=1 frames



【 19 】

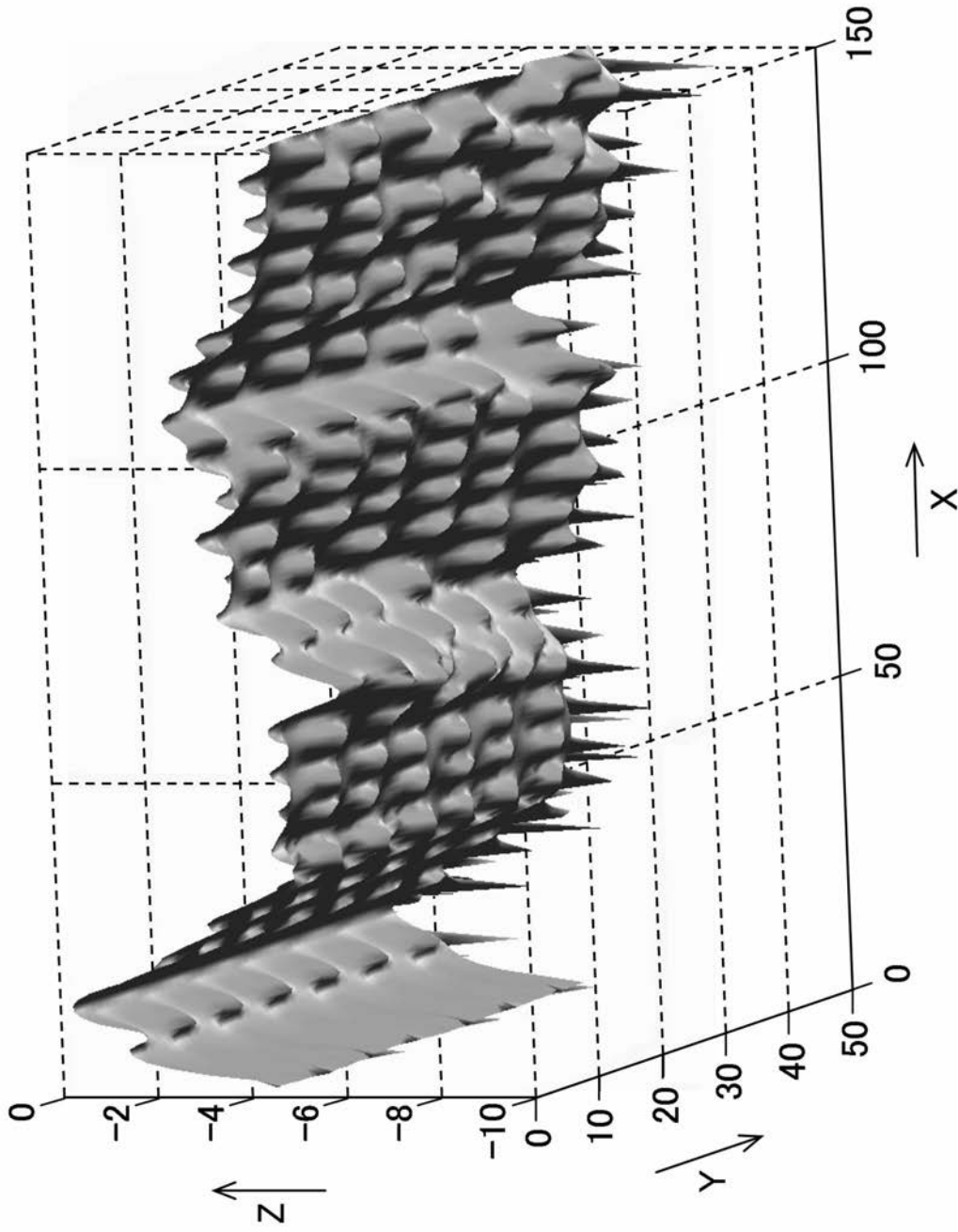


【 20 】

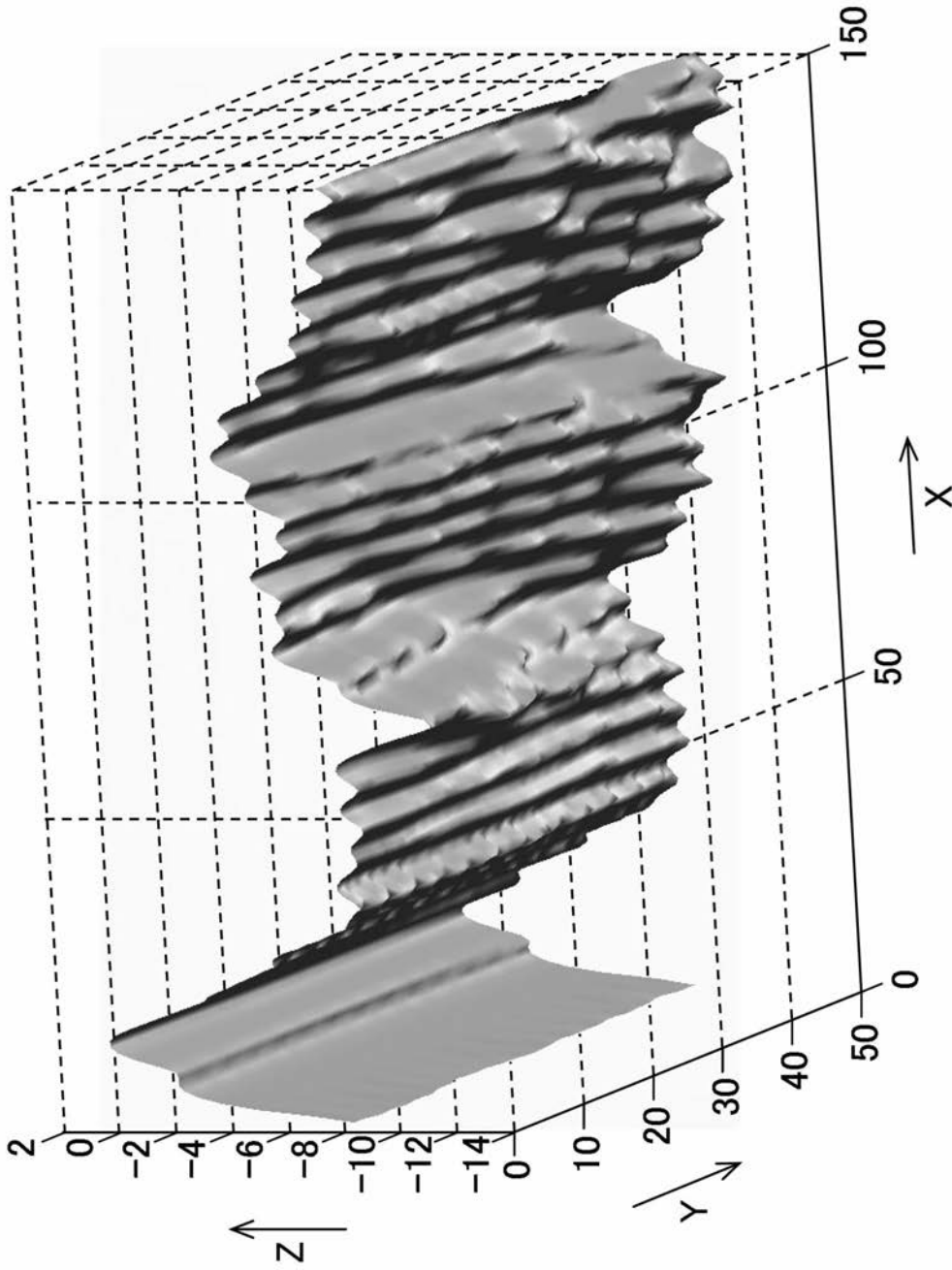




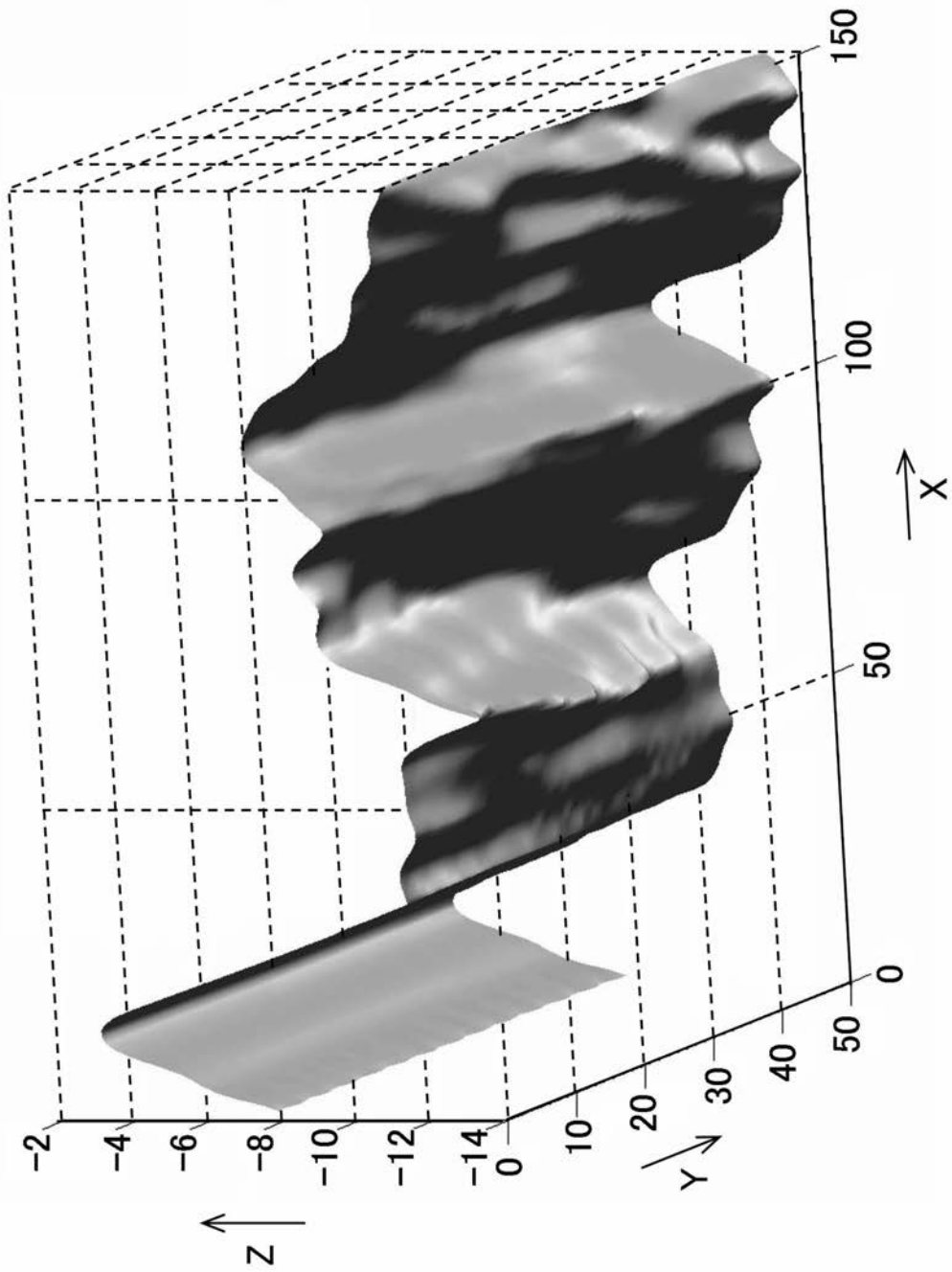
【図7】



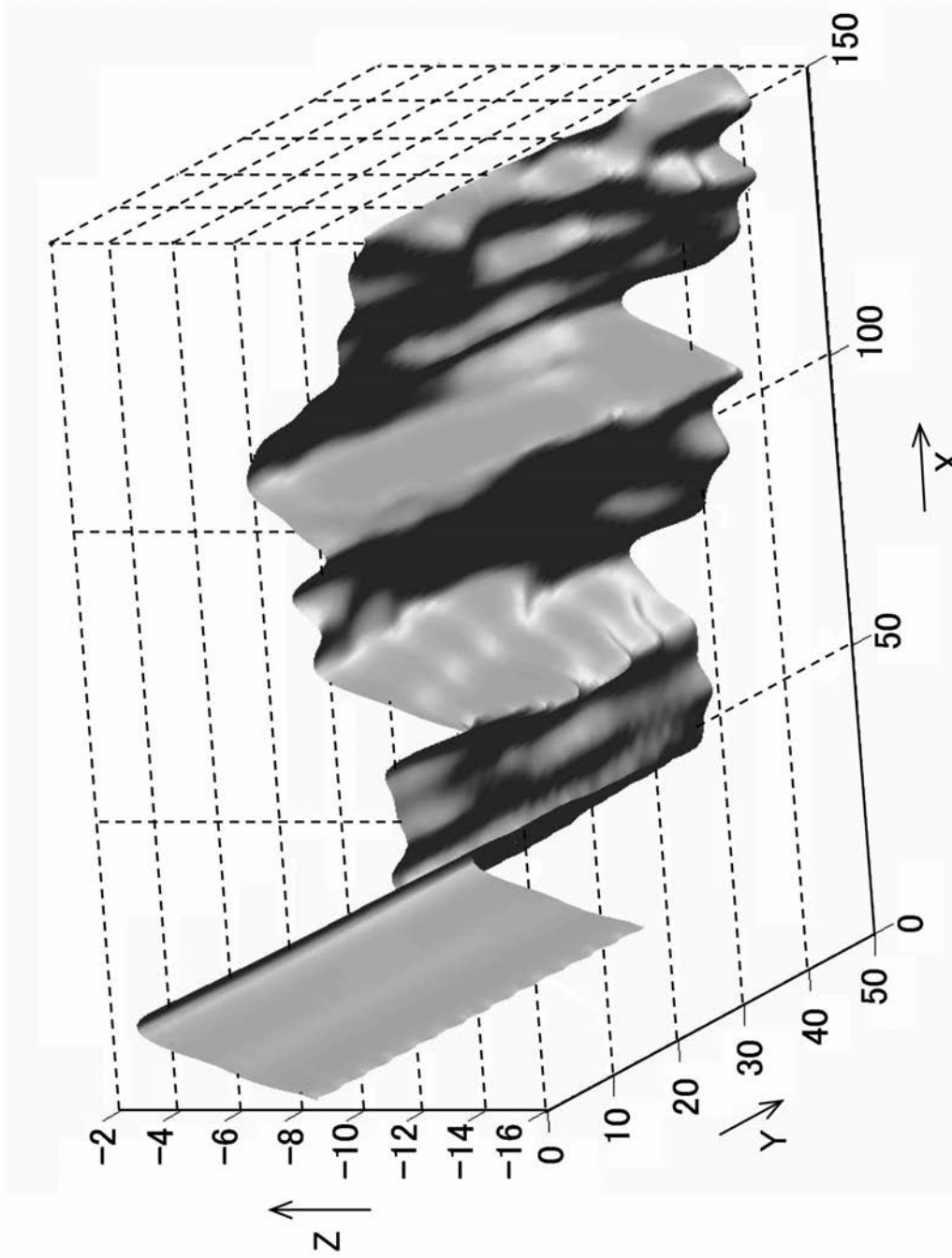
【図 8】



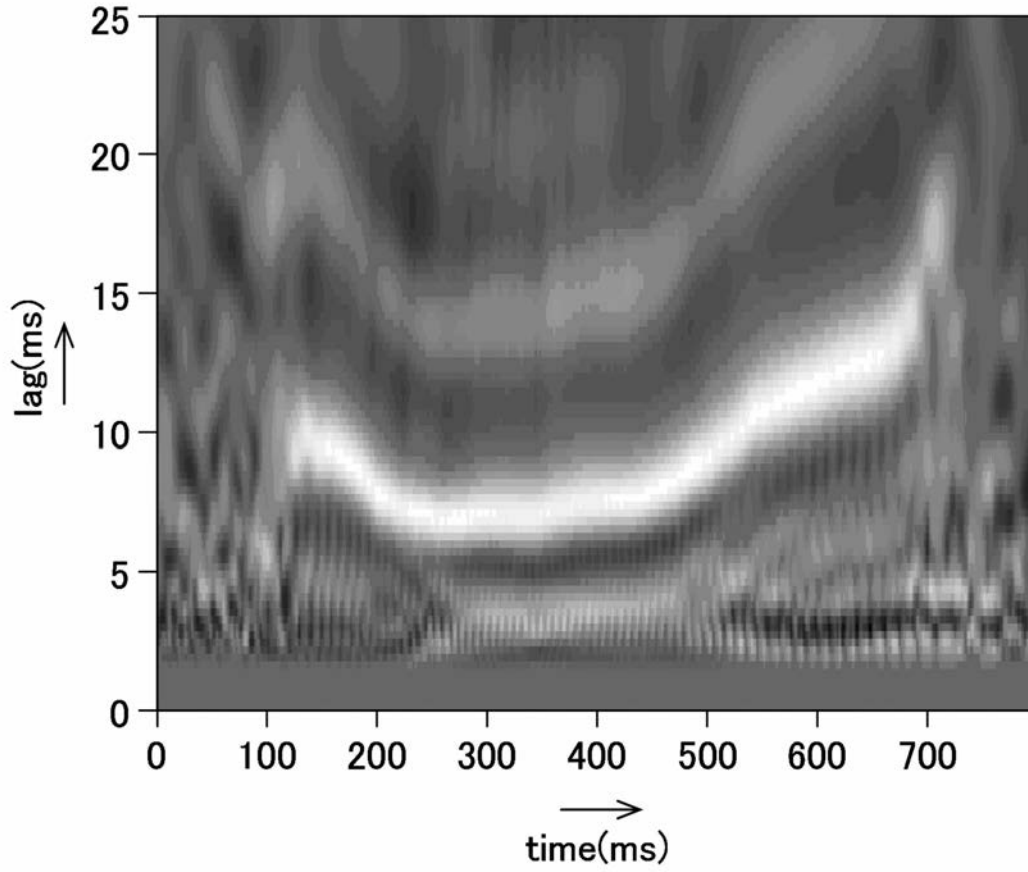
【図9】



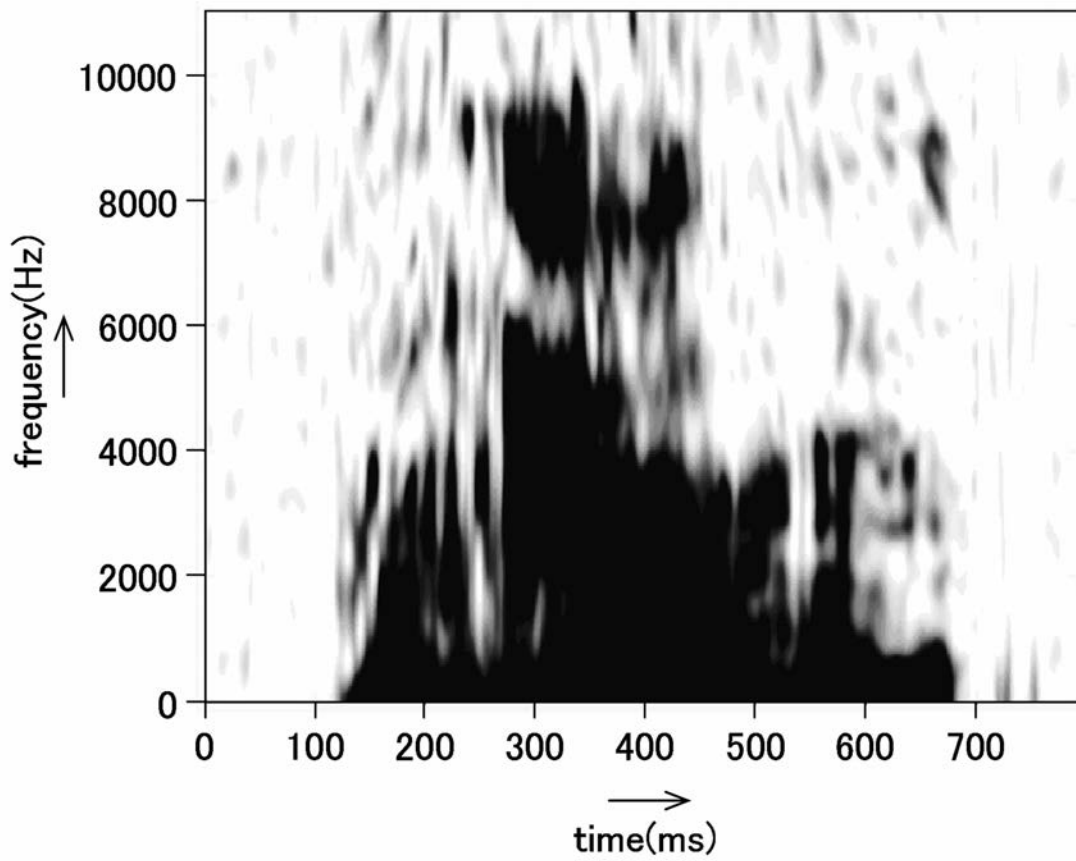
【図10】



【 図 1 8 】



【 図 2 1 】



---

フロントページの続き

- (72)発明者 高橋 徹  
和歌山県和歌山市宇治袋町10番地 メゾンホール彩603号
- (72)発明者 入野 俊夫  
大阪府泉南郡田尻町りんくうポート北5-17, 4-036

審査官 菊地 陽一

- (56)参考文献 特開2001-356787(JP,A)  
特開平10-097287(JP,A)  
特開2006-003517(JP,A)  
河原英紀 他, STRAIGHTにおける時間周波数分析の新しい定式化と実装について, 日本音響学会  
講演論文集CD-ROM, 2007年 9月12日, p.347-p.348

- (58)調査した分野(Int.Cl., DB名)
- |      |       |
|------|-------|
| G10L | 25/18 |
| G10L | 25/45 |
| G10L | 25/90 |