

(19) 日本国特許庁 ( J P )

(12) 特 許 公 報 ( B 2 )

(11) 特許番号

特許第3266819号  
(P3266819)

(45) 発行日 平成14年3月18日 (2002. 3. 18)

(24) 登録日 平成14年1月11日 (2002. 1. 11)

(51) Int.Cl.<sup>7</sup>

識別記号

F I

G 1 0 L 21/04  
11/00

G 1 0 L 3/02  
9/00

A  
A

請求項の数13(全 36 頁)

(21) 出願番号 特願平8-344247

(22) 出願日 平成8年12月24日 (1996. 12. 24)

(65) 公開番号 特開平10-97287

(43) 公開日 平成10年4月14日 (1998. 4. 14)

審査請求日 平成8年12月24日 (1996. 12. 24)

(31) 優先権主張番号 特願平8-200845

(32) 優先日 平成8年7月30日 (1996. 7. 30)

(33) 優先権主張国 日本 ( J P )

(73) 特許権者 592179296

株式会社エイ・ティ・アール人間情報通信研究所

京都府相楽郡精華町光台二丁目2番地2

(72) 発明者 河原 英紀

京都府相楽郡精華町大字乾谷小字三平谷5番地 株式会社エイ・ティ・アール人間情報通信研究所内

(72) 発明者 増田 郁代

京都府相楽郡精華町大字乾谷小字三平谷5番地 株式会社エイ・ティ・アール人間情報通信研究所内

(74) 代理人 100064746

弁理士 深見 久郎 (外2名)

審査官 山下 剛史

最終頁に続く

(54) 【発明の名称】 周期信号変換方法、音変換方法および信号分析方法

1

(57) 【特許請求の範囲】

【請求項1】 周期信号のスペクトログラム上に表現される、基本周期の間隔と基本周波数の間隔とで決まる格子点の情報を用いて、区分的多項式で補間することで、平滑化されたスペクトログラムを得るステップと、前記平滑化されたスペクトログラムを用いて、前記周期信号を別の信号に変換するステップとを含み、周期信号のスペクトログラム上に表現される、基本周期の間隔と基本周波数の間隔とで決まる格子点の情報を用いて、区分的多項式で補間することで、平滑化されたスペクトログラムを得る前記ステップでは、周波数軸上での補間関数と前記周期信号のスペクトログラムを、周波数方向で畳み込み、さらに、時間軸上での補間関数と前記畳み込みで得られたスペクトログラム

2

を、時間方向で畳み込むことによって、前記平滑化されたスペクトログラムを得る、周期信号変換方法。

【請求項2】 位相調整成分と音のスペクトルとの積を用いてインパルス応答を求めるステップと、前記インパルス応答を時間軸上で変換後の音の基本周波数の逆数として決まる周期ずつ移動させながら加算していくことにより、前記音を別の音に変換するステップとを含み、前記位相調整成分を時間の関数とすることにより得られる音源信号は、インパルスと同じパワースペクトルを有し、時間的にエネルギーが分散している、音変換方法。

【請求項3】 前記位相調整成分 ( ) は、  
【数1】

10

$$\Phi(\omega) = \exp\left(j\rho(\omega) \sum_{k \in \Lambda} \alpha_k \cdot \sin(m_k \cdot \xi(\omega))\right)$$

であり、式中の  $\exp(\quad)$  は指数関数を示し、式中の  $\xi(\omega)$  は角周波数を示し、式中の  $\sin(\quad)$  は連続な奇関数を示し、式中の  $\Lambda$  は数字の集まりで、有限個の数字を集めたものを示し、式中の  $k$  は  $\Lambda$  の中から取出した1つの数字を示し、式中の  $\alpha_k$  は係数を示し、式中の  $m_k$  はパラメタを示し、式中の  $\rho(\omega)$  は重みを表わす関数を示す、請求項2に記載の音変換方法。

【請求項4】 前記位相調整成分は、周波数軸上で、乱数と帯域制限関数を畳み込み、帯域制限された乱数を求めるステップと、前記帯域制限された乱数と遅延時間の変動の目標値とを掛け合わせて、群遅延特性を求めるステップと、前記群遅延特性を周波数で積分することにより、位相特性を求めるステップと、前記位相特性と虚数単位とを掛け合わせて、指数関数の指数とすることにより、前記位相調整成分を得るステップとによって得られる、請求項2に記載の音変換方法。

【請求項5】 前記位相調整成分は、第1の成分と第2の成分との積であり、前記第1の成分  $\Phi_1(\omega)$  は、

$$\Phi_1(\omega) = \exp\left(j\rho_1(\omega) \sum_{k \in \Lambda_1} \alpha_k \cdot \sin(m_k \cdot \xi(\omega))\right)$$

であり、式中の  $\exp(\quad)$  は指数関数を示し、式中の  $\xi(\omega)$  は角周波数を示し、式中の  $\sin(\quad)$  は連続な奇関数を示し、式中の  $\Lambda_1$  は数字の集まりで、有限個の数字を集めたものを示し、式中の  $k$  は  $\Lambda_1$  の中から取出した1つの数字を示し、式中の  $\alpha_k$  は係数を示し、式中の  $m_k$  はパラメタを示し、式中の  $\rho_1(\omega)$  は重みを表わす関数を示し、

前記第2の成分は、周波数軸上で、乱数と帯域制限関数を畳み込み、帯域制限された乱数を求めるステップと、前記帯域制限された乱数と遅延時間の変動の目標値とを掛け合わせて、群遅延特性を求めるステップと、前記群遅延特性を周波数で積分することにより、位相特性を求めるステップと、前記位相特性と虚数単位とを掛け合わせて、指数関数の指数とすることにより、前記第2の成分を得るステップとによって得られる、請求項2に記載の音変換方法。

【請求項6】 時間とともに特性が変化するほぼ周期的な信号を生成する機構を表わす時間周波数曲面が、時間の区分的多項式と、周波数の区分的多項式との積で表わされると仮定するステップと、前記ほぼ周期的な信号から所定範囲を、窓関数を使って取出すステップと、

取出された前記所定範囲の前記ほぼ周期的な信号から第1のスペクトルを求めるステップと、前記窓関数の周波数領域での表現と、前記周波数の区分的多項式で表わされる空間の基底とから、周波数方向の最適な補間関数を求めるステップと、前記第1のスペクトルと、前記周波数方向の最適な補間関数を畳み込んで、第2のスペクトルを求めるステップとを含み、前記周波数方向の最適な補間関数は、前記第2のスペクトルと、前記時間周波数曲面の周波数軸に沿った断面との誤差を最小にする、信号分析方法。

【請求項7】  $\omega_1$  から  $\omega_2$  の領域を  $\omega_1$  から  $\omega_2$  の領域に写像する単調で滑らかな関数を用いて、前記第2のスペクトルを第3のスペクトルに変換するステップをさらに含む、請求項6に記載の信号分析方法。

【請求項8】 前記第1のスペクトルから、前記ほぼ周期的な信号の基本周波数の影響を除去して第4のスペクトルを求めるステップと、前記第1のスペクトルを、前記第4のスペクトルで割算して第5のスペクトルを求めるステップと、前記第3のスペクトルと、前記第4のスペクトルとを掛け合わせて、第6のスペクトルを求めるステップとをさらに含み、前記第2のスペクトルを求める前記ステップでは、前記第1のスペクトルの代わりに前記第5のスペクトルを用いて前記第2のスペクトルを求める、請求項7に記載の信号分析方法。

【請求項9】 前記窓関数の時間領域での表現と、前記時間の区分的多項式で表わされる空間の基底とから、時間方向の最適な補間関数を求めるステップと、任意の時間ごとに複数の前記第2のスペクトルを求めるステップと、前記複数の第2のスペクトルを時間方向に並べて第1のスペクトログラムを求めるステップと、前記第1のスペクトログラムと、前記時間方向の最適な補間関数を畳み込んで、第2のスペクトログラムを求めるステップとをさらに含み、前記時間方向の最適な補間関数は、前記第2のスペクトログラムと、前記時間周波数曲面との誤差を最小にする、請求項6に記載の信号分析方法。

【請求項10】 任意の時間ごとに複数の前記第2のスペクトルを求めるステップと、  
-  $\omega_1$  から  $\omega_2$  の領域を  $\omega_1$  から  $\omega_2$  の領域に写像する単調で滑らかな第1の関数を用いて、前記複数の第2のスペクトルを複数の第3のスペクトルに変換するステップと、  
50 前記複数の第3のスペクトルを時間方向に並べて第1の

スペクトログラムを求めるステップと、  
 前記窓関数の時間領域での表現と、前記時間の区分的多項式で表わされる空間の基底とから、時間方向の最適な補間関数を求めるステップと、  
 前記第 1 のスペクトログラムと、前記時間方向の最適な補間関数を畳み込んで、第 2 のスペクトログラムを求めるステップと、  
 - から + の領域を 0 から + の領域に写像する単調で滑らかな第 2 の関数を用いて、前記第 2 のスペクトログラムを第 3 のスペクトログラムに変換するステップとをさらに含み、  
 前記時間方向の最適な補間関数は、前記第 2 のスペクトログラムと、前記時間周波数曲面との誤差を最小にする、請求項 6 に記載の信号分析方法。  
 【請求項 1 1】 時間とともに特性が変化するほぼ周期的な信号を生成する機構を表わす時間周波数曲面が、時間の区分的多項式と、周波数の区分的多項式との積で表わされると仮定するステップと、  
 前記ほぼ周期的な信号から所定範囲を、窓関数を使って取出すステップと、  
 取出された前記所定範囲の前記ほぼ周期的な信号から第 1 のスペクトルを求めるステップと、  
 任意の時間ごとに複数の前記第 1 のスペクトルを求めるステップと、  
 前記複数の第 1 のスペクトルから、前記ほぼ周期的な信号の基本周波数の影響を除去して複数の第 2 のスペクトルを求めるステップと、  
 前記各第 1 のスペクトルを、対応する前記第 2 のスペクトルで割算して複数の第 3 のスペクトルを求めるステップと、  
 前記窓関数の周波数領域での表現と、前記周波数の区分的多項式で表わされる空間の基底とから、周波数方向の最適な補間関数を求めるステップと、  
 前記各第 3 のスペクトルと、前記周波数方向の最適な補間関数を畳み込んで、複数の第 4 のスペクトルを求めるステップと、  
 - から + の領域を 0 から + の領域に写像する単調で滑らかな第 1 の関数を用いて、前記複数の第 4 のスペクトルを複数の第 5 のスペクトルに変換するステップと、  
 前記各第 5 のスペクトルと、対応する前記第 2 のスペクトルとを掛け合わせて、複数の第 6 のスペクトルを求めるステップと、  
 前記複数の第 6 のスペクトルを時間方向に並べて第 1 のスペクトログラムを求めるステップと、  
 前記第 1 のスペクトログラムから、前記ほぼ周期的な信号の周期性に基づく時間的変動の影響を除去して第 2 のスペクトログラムを求めるステップと、  
 前記第 1 のスペクトログラムを、前記第 2 のスペクトログラムで割算して第 3 のスペクトログラムを求めるステ

ップと、  
 前記窓関数の時間領域での表現と、前記時間の区分的多項式で表わされる空間の基底とから、時間方向の最適な補間関数を求めるステップと、  
 前記第 3 のスペクトログラムと、前記時間方向の最適な補間関数を畳み込んで、第 4 のスペクトログラムを求めるステップと、  
 - から + の領域を 0 から + の領域に写像する単調で滑らかな第 2 の関数を用いて、前記第 4 のスペクトログラムを第 5 のスペクトログラムに変換するステップと、  
 前記第 5 のスペクトログラムと、前記第 2 のスペクトログラムとを掛け合わせて、第 6 のスペクトログラムを求めるステップとを含み、  
 前記周波数方向の最適な補間関数は、前記第 4 のスペクトルと、前記時間周波数曲面の周波数軸に沿った断面との誤差を最小にし、  
 前記時間方向の最適な補間関数は、前記第 4 のスペクトログラムと、前記時間周波数曲面との誤差を最小にする、信号分析方法。  
 【請求項 1 2】 第 1 の窓関数を用いて、時間とともに特性が変化するほぼ周期的な信号の第 1 のスペクトルを求めるステップと、  
 所定の窓関数を用いて、第 2 の窓関数を求めるステップと、  
 前記第 2 の窓関数を用いて、前記ほぼ周期的な信号の第 2 のスペクトルを求めるステップと、  
 前記第 1 のスペクトルと、前記第 2 のスペクトルとの平均値を、自乗あるいは単調で非負な関数による変換を介して求め、求めた自乗あるいは単調で非負な関数による変換を介した平均値を第 3 のスペクトルとするステップとを含み、  
 前記第 2 の窓関数を求める前記ステップは、  
 前記所定の窓関数を、原点の両側に、相互の間隔を基本周期分、離して配置するステップと、  
 前記配置された一方の所定の窓関数の符号を反転させるステップと、  
 前記符号を反転させた所定の窓関数と、前記配置された他方の所定の窓関数とを加えさせて前記第 2 の窓関数を求めるステップとを含む、信号分析方法。  
 【請求項 1 3】 任意の時間ごとに複数の前記第 3 のスペクトルを求めるステップと、  
 前記複数の第 3 のスペクトルを時間方向に並べて、スペクトログラムを求めるステップとをさらに含む、請求項 1 2 に記載の信号分析方法。  
 【発明の詳細な説明】  
 【0 0 0 1】  
 【発明の属する技術分野】本発明は、周期信号変換方法、音変換方法および信号分析方法に関し、特に、音を変換するための周期信号変換方法、音変換方法および音

を分析するための信号分析方法に関する。

【0002】

【従来の技術】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。

【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】このような目的を達成するための第1の従来技術は、たとえば、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，pp 527-534に開示されている。この先行技術文献では、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る方法が示されている。

【0005】また、第2の従来技術は、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)，pp 110-116に開示されている。この先行技術文献は、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だものである。

【0006】第3の従来技術として、PSOLA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する方法がある。

【0007】

【発明が解決しようとする課題】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。

【0008】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】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】本発明は、以上のような問題点を解決するためになされたもので、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変換方法を提供することを目的とする。

【0012】この発明の他の目的は、音の標本化周期より高い分解能で精密に音程を設定できる音変換方法を提供することである。

【0013】この発明のさらに他の目的は、過剰平滑化の影響を排除したスペクトルおよびスペクトログラムを求めることができる信号分析方法を提供することである。

【0014】この発明のさらに他の目的は、零となる点のないスペクトルおよびスペクトログラムを求めることができる信号分析方法を提供することである。

【0015】

【課題を解決するための手段】本発明の請求項1の周期信号変換方法は、離散的なスペクトルで与えられる周期信号のスペクトルを区分的多項式で表わされる連続的なスペクトルに変換するステップと、連続的なスペクトルを用いて、周期信号を別の信号に変換するステップとを含む。離散的なスペクトルで与えられる周期信号のスペクトルを区分的多項式で表わされる連続的なスペクトルに変換するステップでは、周波数軸上の補間関数と、離散的なスペクトルを畳み込むことにより、連続的なスペクトルを得る。

【0016】本発明の請求項2の周期信号変換方法は、周期信号のスペクトログラム上に表現される、基本周期の間隔と基本周波数の間隔とで決まる格子点の情報を用いて、区分的多項式で補間することで、平滑化されたスペクトログラムを得るステップと、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するステップとを含む。周期信号のスペクトログラム上に表現される、基本周期の間隔と基本周波数の間隔とで決まる格子点の情報を用いて、区分的多項式で補間すること

で、平滑化されたスペクトログラムを得るステップでは、周波数軸上での補間関数と周期信号のスペクトログラムを、周波数方向で畳み込み、さらに、時間軸上での補間関数と畳み込みで得られたスペクトログラムを、時間方向で畳み込むことによって、平滑化されたスペクトログラムを得る。

【0017】本発明の請求項3の音変換方法は、位相調整成分と、音のスペクトルとの積を用いてインパルス応答を求めるステップと、インパルス応答を時間軸上で、目的とする周期ずつ移動させながら加算していくことにより、音を別の音に変換するステップとを含む。位相調整成分から得られる音源信号は、インパルスと同じパワースペクトルを有し、時間的にエネルギーが分散している。

【0018】本発明の請求項4の音変換方法は、請求項3に記載のものであって、位相調整成分 ( ) は、

【0019】

【数3】

$$\Phi(\omega) = \exp\left(j\rho(\omega) \sum_{k \in \Lambda} \alpha_k \cdot \sin(m_k \cdot \xi(\omega))\right)$$

【0020】であり、式中の  $\exp( )$  は指数関数を示し、式中の  $\omega$  は角周波数を示し、式中の ( ) は連続関数を示し、式中の  $\Lambda$  は数字の集まりで、有限個の数字を集めたものを示し、式中の  $k$  は  $\Lambda$  の中から取出した1つの数字を示し、式中の  $\alpha_k$  は係数を示し、式中の  $m_k$  はパラメタを示し、( ) は重みを表わす関数を示す。

【0021】本発明の請求項5の音変換方法は、請求項3に記載のものであって、位相調整成分は、周波数軸上で、乱数と帯域制限関数を畳み込み、帯域制限された乱数を求めるステップと、帯域制限された乱数と遅延時間の変動の目標値とを掛け合わせて、群遅延特性を求めるステップと、群遅延特性を周波数で積分することにより、位相特性を求めるステップと、位相特性と虚数単位とを掛け合わせて、指数関数の指数とすることにより、位相調整成分を得るステップとによって得られる。

【0022】本発明の請求項6の音変換方法は、請求項3に記載のものであって、位相調整成分は、第1の成分と第2の成分との積である。第1の成分 ( ) は、

【0023】

【数4】

$$\Phi(\omega) = \exp\left(j\rho(\omega) \sum_{k \in \Lambda} \alpha_k \cdot \sin(m_k \cdot \xi(\omega))\right)$$

【0024】であり、式中の  $\exp( )$  は指数関数を示し、式中の  $\omega$  は角周波数を示し、式中の ( ) は連続関数を示し、式中の  $\Lambda$  は数字の集まりで、有限個の数字を集めたものを示し、式中の  $k$  は  $\Lambda$  の中から取出した1つの数字を示し、式中の  $\alpha_k$  は係数を示し、式中の  $m_k$

$k$  はパラメタを示し、( ) は重みを表わす関数を示す。

【0025】第2の成分は、周波数軸上で、乱数と帯域制限関数を畳み込み、帯域制限された乱数を求めるステップと、帯域制限された乱数と遅延時間の変動の目標値とを掛け合わせて、群遅延特性を求めるステップと、群遅延特性を周波数で積分することにより、位相特性を求めるステップと、位相特性と虚数単位とを掛け合わせて、指数関数の指数とすることにより、第2の成分を得るステップとによって得られる。

【0026】本発明の請求項7の信号分析方法は、時間とともに特性が変化するほぼ周期的な信号を生成する機構を表わす時間周波数曲面が、時間の区分的多項式と、周波数の区分的多項式との積で表わされると仮定するステップと、ほぼ周期的な信号から所定範囲を、窓関数を使って取出すステップと、取出された所定範囲のほぼ周期的な信号から第1のスペクトルを求めるステップと、窓関数の周波数領域での表現と、周波数の区分的多項式で表わされる空間の基底とから、周波数方向の最適な補間関数を求めるステップと、第1のスペクトルと、周波数方向の最適な補間関数を畳み込んで、第2のスペクトルを求めるステップとを含む。そして、周波数方向の最適な補間関数は、第2のスペクトルと、時間周波数曲面の周波数軸に沿った断面との誤差を最小にする。

【0027】本発明の請求項8の信号分析方法は、請求項7に記載のものであって、 $-$  から  $+$  の領域を  $0$  から  $+$  の領域に写像する単調で滑らかな関数を用いて、第2のスペクトルを第3のスペクトルに変換するステップをさらに含む。

【0028】本発明の請求項9の信号分析方法は、請求項8に記載のものであって、第1のスペクトルから、ほぼ周期的な信号の基本周波数の影響を除去して第4のスペクトルを求めるステップと、第1のスペクトルを、第4のスペクトルで割算して第5のスペクトルを求めるステップと、第3のスペクトルと、第4のスペクトルとを掛け合わせて、第6のスペクトルを求めるステップとをさらに含む。そして、第2のスペクトルを求めるステップでは、第1のスペクトルの代わりに第5のスペクトルを用いて第2のスペクトルを求める。

【0029】本発明の請求項10の信号分析方法は、請求項7に記載のものであって、窓関数の時間領域での表現と、時間の区分的多項式で表わされる空間の基底とから、時間方向の最適な補間関数を求めるステップと、任意の時間ごとに複数の第2のスペクトルを求めるステップと、複数の第2のスペクトルを時間方向に並べて第1のスペクトログラムを求めるステップと、第1のスペクトログラムと、時間方向の最適な補間関数を畳み込んで、第2のスペクトログラムを求めるステップとをさらに含む。そして、時間方向の最適な補間関数は、第2のスペクトログラムと、時間周波数曲面との誤差を最小に

10

20

30

40

50

する。

【0030】本発明の請求項11の信号分析方法は、請求項7に記載のものであって、任意の時間ごとに複数の第2のスペクトルを求めるステップと、 $-$  から $+$  の領域を0から $+$  の領域に写像する単調で滑らかな第1の関数を用いて、複数の第2のスペクトルを複数の第3のスペクトルに変換するステップと、複数の第3のスペクトルを時間方向に並べて第1のスペクトログラムを求めるステップと、窓関数の時間領域での表現と、時間の区分的多項式で表わされる空間の基底とから、時間方向の最適な補間関数を求めるステップと、第1のスペクトログラムと、時間方向の最適な補間関数を畳み込んで、第2のスペクトログラムを求めるステップと、 $-$  から $+$  の領域を0から $+$  の領域に写像する単調で滑らかな第2の関数を用いて、第2のスペクトログラムを第3のスペクトログラムに変換するステップとをさらに含む。そして、時間方向の最適な補間関数は、第2のスペクトログラムと、時間周波数曲面との誤差を最小にする。

【0031】本発明の請求項12の信号分析方法は、時間とともに特性が変化するほぼ周期的な信号を生成する機構を表わす時間周波数曲面が、時間の区分的多項式と、周波数の区分的多項式との積で表わされると仮定するステップと、ほぼ周期的な信号から所定範囲を、窓関数を使って取出すステップと、取出された所定範囲のほぼ周期的な信号から第1のスペクトルを求めるステップと、任意の時間ごとに複数の第1のスペクトルを求めるステップと、複数の第1のスペクトルから、ほぼ周期的な信号の基本周波数の影響を除去して複数の第2のスペクトルを求めるステップと、各第1のスペクトルを、対応する第2のスペクトルで割算して複数の第3のスペクトルを求めるステップと、窓関数の周波数領域での表現と、周波数の区分的多項式で表わされる空間の基底とから、周波数方向の最適な補間関数を求めるステップと、各第3のスペクトルと、周波数方向の最適な補間関数を畳み込んで、複数の第4のスペクトルを求めるステップと、 $-$  から $+$  の領域を0から $+$  の領域に写像する単調で滑らかな第1の関数を用いて、複数の第4のスペクトルを複数の第5のスペクトルに変換するステップと、各第5のスペクトルと、対応する第2のスペクトルとを掛け合わせて、複数の第6のスペクトルを求めるステップと、複数の第6のスペクトルを時間方向に並べて第1のスペクトログラムを求めるステップと、第1のスペクトログラムから、ほぼ周期的な信号の周期性に基づく時間的変動の影響を除去して第2のスペクトログラムを求めるステップと、第1のスペクトログラムを、第2のスペクトログラムで割算して第3のスペクトログラムを求めるステップと、窓関数の時間領域の表現と、時間の区分的多項式で表わされる空間の基底とから、時間方向の最適な補間関数を求めるステップと、第3のスペク

トログラムと、時間方向の最適な補間関数を畳み込んで、第4のスペクトログラムを求めるステップと、 $-$  から $+$  の領域を0から $+$  の領域に写像する単調で滑らかな第2の関数を用いて、第4のスペクトログラムを第5のスペクトログラムに変換するステップと、第5のスペクトログラムと、第2のスペクトログラムとを掛け合わせて、第6のスペクトログラムを求めるステップとを含む。そして、周波数方向の最適な補間関数は、第4のスペクトルと、時間周波数曲面の周波数軸に沿った断面との誤差を最小にし、時間方向の最適な補間関数は、第4のスペクトログラムと、時間周波数曲面との誤差を最小にする。

【0032】本発明の請求項13の信号分析方法は、第1の窓関数を用いて、時間とともに特性が変化するほぼ周期的な信号の第1のスペクトルを求めるステップと、所定の窓関数を用いて、第2の窓関数を求めるステップと、第2の窓関数を用いて、ほぼ周期的な信号の第2のスペクトルを求めるステップと、第1のスペクトルと、第2のスペクトルとの平均値を、自乗あるいは単調で非負な関数による変換を介して求め、求めた自乗あるいは単調で非負な関数による変換を介した平均値を第3のスペクトルとするステップとを含む。そして、第2の窓関数を求める前記ステップは、所定の窓関数を、原点の両側に、相互の間隔を基本周期分、離して配置するステップと、配置された一方の所定の窓関数の符号を反転させるステップと、符号を反転させた所定の窓関数と、配置された他方の所定の窓関数とを加えて第2の窓関数を求めるステップとを含む。

【0033】本発明の請求項14の信号分析方法は、請求項13に記載のものであって、任意の時間ごとに、複数の第3のスペクトルを求めるステップと、複数の第3のスペクトルを時間方向に並べて、スペクトログラムを求めるステップとをさらに含む。

【0034】

【発明の実施の形態】以下、本発明による周期信号変換方法および音変換方法としての音声変換方法について、原理、処理、具体的処理の順に説明する。

【0035】[実施の形態1]

(原理)本実施の形態では、音声信号の周期性を積極的に利用することにより、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、標準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。

【0036】次のような周期信号(音声信号) $f(t)$ を仮定する。すなわち、 $f(t) = f(t + n)$ である。ここで、 $t$ は時間であり、 $n$ は任意の整数であり、 $1$ は周期である。この信号のフーリエ変換を $F(\omega)$ とすると、 $F(\omega)$ は、 $2\pi / T$ を間隔とするパルス列と

なる。これを適当な補間関数  $h(\lambda)$  を用いて次のよう \* 【0037】  
 に平滑化する。 \* 【数5】

$$S(\omega) = \sqrt{g^{-1} \left( \int_{-\infty}^{\infty} h(\lambda) g(|F(\omega - \lambda)|^2) d\lambda \right)} \quad (1)$$

【0038】式(1)において、 $S(\lambda)$  は、平滑化されたスペクトルであり、 $g(\lambda)$  は適当な単調増加関数であり、 $g^{-1}(\lambda)$  は  $g(\lambda)$  の逆関数であり、 $\lambda$  は角周波数である。積分の範囲は  $-\lambda/2$  から  $\lambda/2$  としているが補間関数として例えば  $-\lambda/2$  から  $\lambda/2$  の範囲外では0となるようなものを用いることにより、 $-\lambda/2$  から  $\lambda/2$  とすることができる。ここで、補間関数が以下に示す直線復元条件を満たすことを要請する。この直線復元条件は、音色情報を表わすスペクトル包絡が「信号の周期性の影響を受けずしかも滑らかである」ということを合理的に定式化したものである。

【0039】直線復元条件について説明する。この条件

$$a\omega + b = \int_{-\infty}^{\infty} (a\omega + b) h_2(\lambda) \left( \sum_{k=0}^N \delta(\omega - \lambda - k\Delta\omega) \right) d\lambda \quad (2)$$

【0041】が区間  $(-\lambda/2, (N-2)\lambda/2)$  において 20  
 成立する。ここで  $a, b$  は任意の定数を表わし、 $\delta(x)$  はデルタ関数を表わす。また  $\Delta\omega$  は信号の周期に対応する周波数軸上での調波の間隔を角周波数で表わしたものである。なお、標本化関数として知られている  $\sin(x)/x$  も、パルス列が無限に一定値で続く場合や、一定の割合で変化し続ける場合には直線復元条件を満たす。しかし、実際の時間的に変化する信号ではそのように無限に同じ傾向が続くことはなく、直線復元条件は満たされない。

【0042】時間窓との相互作用について説明する。信号の短時間フーリエ変換を求める場合には、何らかの窓関数  $w(t)$  を用いて信号の一部を切出すことが必要となる。周期関数をこのような窓関数を用いて切出すとその短時間フーリエ変換は、周波数領域でのパルス列に窓関数のフーリエ変換である  $W(\omega)$  を畳み込んだものとなる。この場合でも、補間関数として直線復元条件を満たすパートレット窓関数 (Bartlett窓関数) を用いれば、最終的なスペクトル包絡は直線復元条件を満たす。

【0043】標本化周期より詳細な基本周期制御方式について説明する。以上のようにして、平滑化された実数スペクトルが求まれば、直接逆フーリエ変換することによって要素となる時間領域での直線位相のインパルス応答  $s(t)$  を求めることができる。具体的には、 $j$  を虚数単位 ( $j = \sqrt{-1}$ ) とすると、次式で表わされる。

【0044】

【数7】

$$s(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega$$

【0045】あるいは、次のようにして最小位相のパルス応答  $v(t)$  を作成することができる。

は、隣り合う複数のインパルスの高さが同一であるときに補間関数により平滑化された値が一定値になることを要請する。さらに、この条件は、インパルスの高さが一定の割合で変化していくときに補間関数により平滑化された値が直線になることを要請する。この条件を満たす補間関数  $h(\lambda)$  は、パートレット窓 (Bartlett窓) として知られる  $4/\lambda^2$  を幅とする三角形の補間関数  $h_2(\lambda)$  と、時間窓関数を周波数変換して得られるようなエネルギーが局在するような関数を畳み込むことにより作られる関数である。具体的には、 $S(\lambda)$  のうち、

【0040】

【数6】

【0046】

【数8】

$$c(q) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \log S(\omega) e^{-j\omega q} d\omega$$

$$g(q) = \begin{cases} 0 & (q < 0) \\ c(0) & (q = 0) \\ 2c(q) & (q > 0) \end{cases}$$

$$V(\omega) = \exp\left(\frac{1}{\sqrt{2\pi}} \int_0^{\infty} g(q) e^{j\omega q} dq\right)$$

$$v(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} V(\omega) e^{j\omega t} d\omega$$

30

40

50

【0047】直線位相のインパルス応答  $s(t)$  または最小位相のインパルス応答  $v(t)$  を、時間軸の上で目的とする周期ずつ移動させながら加算していくことにより、変換音声を作成することができる。しかし、信号を標本化により離散化している場合には、この方法では、標本化周波数により決まる標本化周期よりも細かく周期を制御することはできない。そこで、時間遅れが周波数領域では位相の直線的な変化として表わされることを利用して、波形の構成の際に標本化周期よりも細かな周期の補正分を求めて復元波形を変換することにより、この問題を解決する。具体的には、目的とする周期が、標本化周期  $T$  を用いると  $(m+r)T$  と表わされるものとする。ここで、 $m$  は整数であり、 $r$  は  $0 < r < 1$  の実数とする。このようにすると、具体的な位相調整分の値 (以下、「位相調整成分」という)  $\phi_r(\omega)$  は、次のようになる。

【0048】

【数 9】

$$\Phi_1(\omega) = e^{-j\alpha\Delta T}$$

(8)

【0049】直線位相のインパルスを用いる場合には、位相調整成分  $\phi_1(\omega)$  によって  $S(\omega)$  を位相調整して  $S_r(\omega)$  を作成する。具体的には、 $\phi_1(\omega)$  と  $S(\omega)$  とを掛け合わせて  $S_r(\omega)$  を作成する。そして、この  $S_r(\omega)$  を、式(3)の  $S(\omega)$  の代わりに用いることによって、直線位相のインパルス応答  $s_r(t)$  を求める。この直線位相のインパルス応答  $s_r(t)$  を、目的とする周期の整数分  $m \cdot T$  の位置に加算して波形を作成する。

【0050】最小位相のインパルス応答を用いる場合には、位相調整成分  $\phi_1(\omega)$  によって  $V(\omega)$  を位相調\*

$$\Phi_2(\omega) = \exp\left(j\rho(\omega) \sum_{k \in \Lambda} \alpha_k \cdot \sin(m_k \cdot \xi(\omega))\right)$$

(9)

【0053】ここで、 $\exp(\cdot)$  は指数関数を示し、 $\rho(\omega)$  は  $-\pi$  の範囲を  $\pi$  の範囲に移すような滑らかな連続な奇関数であり、範囲の両端である  $-\pi$  と  $\pi$  において  $\rho(\omega) = \pm\pi$  となるように拘束されている。また、 $\Lambda$  は、添字の集まりで、たとえば 1, 2, 3, 4 など、有限個の数字を集めたものである。このような式(9)は、 $\xi_2(\omega)$  が、 $\xi(\omega)$  によって非線形に伸縮された角周波数  $\omega$  の上の複数の異なった三角関数を、係数  $\alpha_k$  により重み付けしたものの和として

$$\xi(\omega) = \pi \cdot \text{sgn}(\omega) \left| \frac{\omega}{\pi} \right|^\beta$$

(10)

【0055】周波数軸上での位相回転の周波数微分が群遅延に相当することを利用すれば、平均値が0の乱数を積分したものを位相成分とすることによって、群遅延の分布を乱数により制御することができる。このような高い周波数の成分の位相の制御は、息づかいの混じった声を作り出すなど、合成音声の自然性の向上のために、非常に大きく貢献する。具体的には、位相調整成分  $\phi_3(\omega)$  により位相調整して音声合成を行なう。この位相調整成分  $\phi_3(\omega)$  は、次のようにして作成される。第1のステップとして乱数を発生する。第2のステップとして、周波数軸上で、第1のステップで発生した乱数と、帯域制限関数を畳み込む。そして、帯域制限された乱数を求める。第3のステップとして、どの周波数領域がどれだけの群遅延の変動を許すかを設計する。つまり、どの周波数領域がどれだけの遅延時間の変動を許すかを設計する。具体的には、遅延時間の変動の目標値を設計する。そして、帯域制限された乱数(第2のステップで求めたもの)と遅延時間の変動の目標値とを掛け合わせて、群遅延特性を作成する。第4のステップとして、群遅延特性を周波数で積分することにより、位相特性を作成する。第5のステップとして、位相特性と虚数

\* 整して  $V_r(\omega)$  を作成する。具体的には、 $\phi_1(\omega)$  と  $V(\omega)$  とを掛け合わせて  $V_r(\omega)$  を作成する。そして、式(7)の  $V(\omega)$  の代わりに  $V_r(\omega)$  を用いて、最小位相のインパルス応答  $v_r(t)$  を求める。この最小位相のインパルス応答  $v_r(t)$  を、目的とする周期の整数分  $m \cdot T$  の位置に加算して波形を作成する。

【0051】位相調整成分の他の例を示す。すなわち、位相調整成分の他の例  $\phi_2(\omega)$  は次式で表わされる。

【0052】

【数 10】

表わされることを示している。なお、式(9)中の  $k$  は  $\Lambda$  の中から取出した1つの数字を示し、式中の  $m_k$  はパラメタを示している。 $\rho(\omega)$  は、重みを表わす関数を示す。連続関数  $\rho(\omega)$  の具体例として、 $\text{sgn}(\omega)$  をパラメタとすると、次式で表わされるものがある。ここで  $\text{sgn}(\omega)$  は  $\omega > 0$  内が0または正の時に1、負の場合に-1となる符号を表わす関数である。

【0054】

【数 11】

単位 ( $j = -1$ ) とを掛け合わせて、指数関数の指数とすることにより、位相調整成分  $\phi_3(\omega)$  を得る。

【0056】三角関数を用いた位相の制御 ( $\phi_2(\omega)$  を用いる位相の制御) と、乱数を用いた位相の制御 ( $\phi_3(\omega)$  を用いた位相の制御) とは、周波数領域で表現されているので、 $\phi_2(\omega)$  と  $\phi_3(\omega)$  とを掛け算することにより、両方の性質を有する位相調整成分を作成することができる。すなわち、声門の開閉のイベントに相当する離散的なパルスの周辺で乱流や声帯振動の変動に起因する雑音的な変動がある音源を作成できる。また、 $\phi_1(\omega)$  と  $\phi_2(\omega)$  と  $\phi_3(\omega)$  とを掛け算することによっても位相調整成分を作成することができるし、 $\phi_1(\omega)$  と  $\phi_2(\omega)$  とを掛け算することによっても位相調整成分を作成することができるし、 $\phi_1(\omega)$  と  $\phi_3(\omega)$  とを掛け算することによっても位相調整成分を作成することができる。ここで、位相調整成分  $\phi_2(\omega)$ 、 $\phi_3(\omega)$ 、 $\phi_1(\omega) \cdot \phi_2(\omega) \cdot \phi_3(\omega)$ 、 $\phi_1(\omega) \cdot \phi_2(\omega)$ 、 $\phi_1(\omega) \cdot \phi_3(\omega)$  および  $\phi_2(\omega) \cdot \phi_3(\omega)$  による位相調整の仕方は、 $\phi_1(\omega)$  による位相調整の仕方と同様である。

【0057】図1は、位相調整成分 $\phi_2$ ( )によって得られた音源信号を示す図である。図1を参照して、横軸は時間を示し、縦軸は音圧を示している。ここで、位相調整成分 $\phi_2$ ( )を構成する連続関数 $\phi_1$ ( )として、式(10)を用いている。重み関数としては、 $\phi_1$ ( ) = 1という定数値を持つものを選んでいる。また、 $m_1$ は1個の数字からなり、 $k = 1$ 、 $m_1 = 30$ 、 $\phi_1 = 0.3$ 、 $\phi_2 = 1$ としている。図2は、位相調整成分 $\phi_3$ ( )によって得られた音源信号を示す図である。図3は、位相調整成分 $\phi_2$ ( )・ $\phi_3$ ( )によって得られる音源信号を示す図である。図2および図3を参照して、横軸は時間を示し、縦軸は音圧を示している。図1~図3を参照して、音源信号が、インパルスと異なりエネルギーが時間的に分散しているのが観測できる。ここで、音源信号は、位相調整成分を時間の関数にしたものである。具体的には、音源信号は、位相調整成分を逆フーリエ変換して、時間の関数にしたものである。

【0058】(処理)実施の形態1による音声変換方法は、以下の手順によって実現される。まず、分析の対象となる音声信号は、予め何らかの手段でデジタル化されているものとする。第1の処理として、音声の基本周波数(基本周期)の抽出について説明する。実施の形態1による音声変換方法では、分析対象とする音声信号の周期性を積極的に利用している。これらの周期性の情報は、式(1)、(2)の中の補間関数のサイズを決定するために用いられる。第1の処理では、音声信号から一部を次々と選び出しながら、その部分における基本周波数(基本周期)を抽出する。詳しくは、デジタル化した音声信号の標本化周期よりも精密な分解能で基本周波数(基本周期)を抽出する。また、周期的でない信号が含まれる部分では、その旨を何らかの形で抽出しておく。第1の処理で基本周波数を精密に抽出しておくことが、後述する第5の処理で重要になる。なお、このような基本周波数(基本周期)の抽出は、既存の一般的な方法を用いて行なう。必要があれば、音声波形を視認しながら手作業で基本周波数を決めてもよい。

【0059】基本周波数の情報を利用した補間関数の適応を行なう第2の処理について説明する。第2の処理では、式(2)の条件を満たす1次元の補間関数を用いて、式(1)により、周波数方向において音声信号のスペクトルと補間関数を畳み込むことにより、平滑化スペクトルを計算する。これにより、周波数方向の周期性の影響が小さくなる。

【0060】音声パラメタの変換を行なう第3の処理について説明する。第3の処理では、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波数情報)の周波数軸を圧縮したり、声の高さを変えるために、精密な基本周波数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波数情報)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。

【0061】変換された音声パラメタを用いて音声合成を行なう第4の処理について説明する。第4の処理では、平滑化スペクトルから、式(3)を用いて精密な基本周波数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り、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、式(8)を用いて計算した値 $\phi_1$ ( )を、式(1)の $S$ ( )に掛け算してから式(3)を用いて $s$ ( $t$ )で表わされる音源波形を作成することで、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波数の制御を行なうことが可能となる。

【0062】また、平滑化スペクトルから、式(4)、(5)、(6)、(7)を用いて精密な基本周波数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り、変換された音声を作成することもできる。その場合には、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、式(8)を用いて計算した値 $\phi_1$ ( )を、式(6)の $V$ ( )に掛け算してから式(7)を用いて $v$ ( $t$ )で表わされる音源波形を作成することで、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波数の制御を行なうことが可能となる。ここで、 $S$ ( )または $V$ ( )に掛け算する位相調整成分としては、 $\phi_1$ ( )を用いたが、位相調整成分としては、 $\phi_2$ ( )、 $\phi_3$ ( )、 $\phi_1$ ( )・ $\phi_2$ ( )・ $\phi_3$ ( )、 $\phi_1$ ( )・ $\phi_2$ ( )・ $\phi_3$ ( )または $\phi_2$ ( )・ $\phi_3$ ( )を用いることもできる。

【0063】第4の処理は、この部分だけを取り出しても利用することができる。すなわち、平滑化スペクトルは、2次元の濃淡画像であるに過ぎないし、精密な基本周波数は、その画像の横幅と同じ幅を有する1次元の曲線にすぎない。したがって、第4の処理を用いれば、そのような画像と曲線を情報を失うことなく音に変えることができる。つまり、音声信号の入力が不要で、画像と曲線があれば、音を作ることができる。

【0064】(具体的処理)図4は、本発明の実施の形態1による音声変換方法を実現するための音声変換装置を示す概略ブロック図である。図4を参照して、音声変換装置は、パワースペクトル計算部1、基本周波数計算部2、平滑化スペクトル計算部3、インタフェース部

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 および波形合成部 8 を備える。図 4 の音声変換装置を用いて、8 kHz 16 ビットで標準化された音声を変換する例を説明する。パワースペクトル計算部 1 では、30 ms の Hanning 窓を用いて、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が計算される。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。

【0065】図 5 は、図 4 のパワースペクトル計算部 1 によって求められたパワースペクトルの一例および平滑化スペクトル計算部 3 によって求められた平滑化スペクトルの一例を示す図である。横軸は、周波数を示し、縦軸は、強度を対数表示 (デシベル表示) を用いて示している。図 5 を参照して、矢印 a で示す曲線が、パワースペクトル計算部 1 で求めたパワースペクトルである。

【0066】再び、図 4 を参照して、図 5 に示されるようなパワースペクトルの調波構造の周期から、基本周波数計算部 2 において、音声の基本周波数  $f_0$  を求める。パワースペクトル計算部 1 および基本周波数計算部 2 は、上述した第 1 の処理 (音声の基本周波数の抽出) を行なう部分である。平滑化スペクトル計算部 3 では、基本周波数計算部 2 で求めた基本周波数  $f_0$  に基づいて、平滑化のための補間関数として幅が  $2 f_0$  であるような三角形の形状の関数を選ぶ。この補間関数を用いて、周波数軸上で円環畳み込みを実行することにより平滑化されたスペクトルを得る。

【0067】再び、図 5 を参照して、矢印 b に示す曲線が平滑化されたスペクトルである。ここでは、単調増加関数  $g(\cdot)$  として、平方根を求める関数を用いている。人間の知覚に近づけるために  $g(\cdot)$  としてパワーの 0.6 乗を計算する関数を用いることもできる。平滑化スペクトル計算部 3 は、上述した第 2 の処理 (基本周波数の情報を利用した補間関数の適応) を行なう部分である。平滑化スペクトル計算部 3 で求めた平滑化スペクトルは、平滑化スペクトル変換部 5 に渡され、基本周波数計算部 2 で得られた音源情報 (精密な基本周波数情報) は、音源情報変換部 6 に渡される。ここで、後からの利用のために、平滑化スペクトルおよび音源情報を格納しておくこともできる。インタフェース部 5 は、平滑化スペクトルと音源情報の計算段階と、変換・合成段階との

インタフェース部分である。  
【0068】平滑化スペクトル変換部 5 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  $S(\cdot)$  を  $V(\cdot)$  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  $S_m(\cdot)$  を得る。あるいは、変形した平滑化スペクトル  $S_m(\cdot)$  を、式 (4) ~ (6) を用いて、 $V(\cdot)$  に変換しておく。つまり、式 (4) の  $S(\cdot)$  の代わりに、 $S_m(\cdot)$  を用いて、 $V(\cdot)$  を求める。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  $S_m(\cdot)$  も、「 $S(\cdot)$ 」で表わす。音源情報変換部 6 では、平滑化スペクトル変換部 5 での変換と並行して、音源情報を目的に応じて変換する。平滑化スペクトル変換部 5 および音源情報変換部 6 での処理は、上述した第 3 の処理 (音声パラメタの変換) を行なう部分である。位相調整部 7 では、平滑化スペクトル変換部 5 および音源情報変換部 6 で変換されたスペクトル情報と音源情報を用いて、標準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準化周期  $T$  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  $\phi_r(\cdot)$  を求める。そして、 $S(\cdot)$  あるいは  $V(\cdot)$  の位相を調整する。波形合成部 8 では、位相調整部 7 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 6 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 7 および波形合成部 8 は、第 4 の処理 (変換された音声パラメタによる音声合成) を行なう部分である。

【0069】図 6 は、 $V(\cdot)$  を逆フーリエ変換した最小位相のインパルス応答  $v(t)$  の例を示す図である。図 6 を参照して、横軸は時間を示し、縦軸は音圧を示している。図 7 は、 $V(\cdot)$  を用いて、音源を変換して合成された信号波形を示す図である。図 7 を参照して、横軸は時間を示し、縦軸は音圧を示す。図 7 を参照して、標準化周期よりも細かに基本周波数が制御されているため、繰返される波形の形状やピークの高さが微妙に異なっている。

【0070】以上のように、実施の形態 1 による音声変換方法では、周期信号のスペクトルのピークが周波数軸上で等間隔に並ぶ性質を利用し、等間隔のスペクトルのピーク値が直線状に変化する場合に直線性を保存するような補間関数と、周期信号のスペクトルを畳み込むことにより平滑化されたスペクトルを得る。つまり、周期性の影響が小さいスペクトルを得ることができる。このため、実施の形態 1 による音声変換方法では、これまで不可能だった 500% にも及ぶ範囲での声の高さや速度、周波数帯域の変換を自然性を損なうことなく行なうことができる。

【0071】また、実施の形態 1 による音声変換方法では、信号の周期性のみを用いて直線が直線として復元されるという 1 つの合理的な基準の下で平滑化されたスペクトルを抽出しているため、スペクトルのモデルに基づくこれまでの方法とは異なり、どのような音源から発した音であっても高い品質を保ちながら変換することができる。

【0072】さらに、実施の形態 1 による音声変換方法では、音声の解析などを行なう場合、周期的成分によるスペクトル形状に対する干渉を大きく削減することができるため、平滑化されたスペクトルは、音声の診断に有

用である。

【0073】さらに、実施の形態1による音声変換方法では、音声の解析などを行なう場合、周期的成分によるスペクトル形状に対する干渉を大きく削減することができるため、平滑化されたスペクトルは、音声認識・話者認識における標準パターンの作成の精度を大きく向上させることができる。

【0074】さらに、実施の形態1による音声変換方法では、電子楽器などにおいても、標本化された信号そのものを格納するのではなく、平滑化されたスペクトル情報と音源情報（音源の周期や強度の情報）の形に分離して格納しておくことによって、精密な周期の制御や位相調整成分を用いた音色の制御によりこれまでになかった音楽表現を生み出すことができる。

【0075】さらに、実施の形態1による音声変換方法では、任意の濃淡画像を音として合成することが可能となるため、芸術表現、視覚障害者の情報提示、コンピュータのデータの音響提示による新しいユーザインタフェースなどへの応用が可能である。このような応用は、音声研究を根本的に変革するばかりではなく、コンピュータグラフィックスが映像の世界にもたらしたのと同様のインパクトを音の世界にもたらすことが予想される。

【0076】また、実施の形態1による音声変換方法を用いることによって、以下に示すようなことが実現される可能性もある。たとえば、猫の発声器官の寸法が人間の発声器官の寸法の1/4程度であることを利用して、猫の声を実施の形態1による音声変換方法により4倍の寸法の器官から発生されたもののように変換し、また、人間の声を実施の形態1による音声変換方法により1/4の寸法の器官から発生されたもののように変換することにより、これまで物理的な寸法の違いによって等身大のコミュニケーションが不可能であった異種生物間でのコミュニケーションが可能になるという可能性もある。

【0077】[実施の形態2]一般的なスペクトログラム（スペクトルの時間・周波数表現）の性質について言及する。まず、時間分解能が高い場合のスペクトログラムの性質を述べる。周波数を一定にして、スペクトログラムの時間方向の変化を観察する。この場合には、スペク

$$S_2(\omega, t) = \sqrt{g^{-1} \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_t(\lambda, u) g(|F_2(\omega - \lambda, t - u)|^2) d\lambda du \right)}$$

(11)

【0081】ここで  $\lambda$  は周波数に対応する積分変数、 $u$  は時間に対応する積分変数を表わす。 $S_2(\omega, t)$  は式(1)の $S(\omega, t)$ に対応する平滑化されたスペクトログラムであり、 $F_2(\omega, t)$  は式(1)の $F(\omega, t)$ に対応するスペクトログラムである。双1次曲面復元条件について説明する。実施の形態1の直線復元条件は、周波数軸上での話であった。信号の周期性は、時間方向に  $C_\omega \omega + C_t t + C_0 = 0$

\*クトログラムの時間表現には、音声の基本周期による影響が残っている。一方、時間を一定にして、スペクトログラムの周波数方向の変化を観察する。この場合には、スペクトログラムの周波数表現の変化が、本来のスペクトログラムの周波数表現の変化に比べ、潰れてしまっているのが観察できる。次に、周波数分解能が高い場合のスペクトログラムの性質について述べる。周波数を一定にしてスペクトログラムの時間変化を観察する。この場合には、スペクトログラムの時間表現の変化が、本来のスペクトログラムの時間表現の変化に比べ、潰れてしまっているのが観察できる。一方、時間を一定にして、スペクトログラムの周波数方向の変化を観察する。この場合には、スペクトログラムの周波数表現に、周期性の影響が残ってしまう。なお、周波数分解能を高くすれば、必然的に時間分解能は低くなり、時間分解能を高くすれば、必然的に周波数分解能は低くなる。

【0078】従来の音声変換方法では、分析するスペクトルに周期性の影響が大きく残っていたため、音声の加工の自由度は小さかった。そこで、実施の形態1による音声変換方法では、分析するスペクトルの周波数方向の周期性の影響を小さくするため、周波数方向に平滑化したスペクトルを得た。この場合、時間方向の周期性の影響を小さくするため、周波数分解能を高くして（時間分解能を低くして）、スペクトルを分析した。このように、周波数分解能を高くすると、スペクトルの時間方向の細かい変化が潰れてしまうという問題が生じる。実施の形態2による音声変換方法は、このような問題を解決するためになされたものである。

【0079】（原理）実施の形態2による音声変換方法の原理は、実施の形態1による音声変換方法の原理と同様である。ただし、実施の形態1による音声変換方法では、式(1)の補間関数 $h(\omega, t)$ は、直線復元条件を満たすことが要請されていたが、実施の形態2による音声変換方法では、式(11)の補間関数 $h_t(\omega, t, u)$ は、直線復元条件に加えてさらに双1次曲面復元条件を満たすことが要請される。

【0080】  
【数12】

も認められる。したがって、周期信号の場合には、周波数方向には基本周波数ごと、時間方向には基本周期ごとに格子点の情報が、信号の分析から得られることになる。ここで、実施の形態1で説明した1次元の条件を2次元に拡張すると、補間関数 $h_t(\omega, t, u)$ には、

【0082】  
【数13】

【0083】という双1次形式で表わされる面の保存を要請することが合理的である。ここで $C_1$ 、 $C_2$ 、 $C_3$ は双1次曲面を表わすパラメタであり、任意の定数値をとり得る。このような双1次曲面復元条件は、周波数方向では $4/\lambda$ の幅を有する三角形の補間関数と、時間方向では $2/\Delta t$ の幅を有する三角形の補間関数を、2次元で畳み込んだものを補間関数 $h_1(x, u)$ として用いることにより満たすことができる。

【0084】(処理)実施の形態2による音声変換方法の第1の処理、第3の処理および第4の処理は、それぞれ、実施の形態1による音声変換方法の第1の処理、第3の処理および第4の処理と同様である。また、実施の形態2による音声変換方法では、実施の形態1による音声変換方法の第1の処理と第2の処理との間に、特有の処理を行なう。この実施の形態2による音声変換方法の特有の処理を、「第1.5の処理」と呼ぶことにする。さらに、実施の形態2による音声変換方法の第2の処理は、実施の形態1による音声変換方法の第2の処理と異なっている。また、実施の形態2による音声変換方法の第3の処理では、実施の形態1による音声変換方法の第3の処理を行なうことができるとともに、他の処理も行なうことができる。

【0085】基本周期に適応した周波数分析を行なう第1.5の処理について説明する。第1.5の処理では、音声信号の基本周期の情報を用いて、時間窓の周波数分解能と基本周波数の比と、時間窓の時間分解能と基本周期の比とが同じになるような時間窓を設計して適応的なスペクトル分析を行なう。また、周期性の存在しない雑音などの部分では、聴覚的な時間分解能である数msを分析のための時間窓の長さとする。実施の形態2による音声変換方法の効果を最大限に生かすためには、第1.5の処理では、上述の条件を満たす時間窓を用いて、信号の基本周期よりも細かな周期(たとえば、基本周期の $1/4$ 以下)でスペクトル分析を行なうことが必要である。なお、固定した長さの時間窓で行なっても、その時間窓内に数個の基本周期が含まれるのであれば、後述する第2の処理でかなり回復することが可能である。

【0086】実施の形態2による音声変換方法の第2の処理について説明する。第2の処理で、第1.5の処理までで求められたスペクトルの時間周波数表現(たとえば、横軸を時間とし、縦軸を周波数とし、その平面上にスペクトルの強度を表わしたもの。声紋。)、すなわち、スペクトログラムを用いる。また、第2の処理では、式(2)および式(12)の条件を満たす補間関数を基本周波数の情報に基づいて作成する。そして、この補間関数とスペクトログラムを、時間・周波数の2次元方向において畳み込みを行なう。これによって、周期性の影響が除かれた平滑化スペクトログラムを得ることができる。さらに、周期信号を与えることのできる時間・周波数平面上の格子点の情報を自然な形で最も有効に抽

出した平滑化スペクトログラムを得ることができる。実施の形態2による音声変換方法の第3の処理は、実施の形態1による第3の処理を包含している。実施の形態2による音声変換方法の第3の処理では、さらに、たとえば、発声速度を速くするために、求められた音声パラメタ(平滑化スペクトログラムと精密な基本周波数情報)の時間軸を伸縮したりする。なお、処理は、第1の処理、第1.5の処理、第2の処理、第3の処理、第4の処理の順に行なう。

10 【0087】(具体的処理)図8は、実施の形態2による音声変換方法を実現するための音声変換装置である。図8を参照して、この音声変換装置は、パワースペクトル計算部1、基本周波数計算部2、適応的周波数分析部9、平滑化スペクトログラム計算部10、インタフェース部4、平滑化スペクトログラム変換部11、音源情報変換部6、位相調整部7および波形合成部8を備える。なお、図4と同様の部分については同一の参照符号を付しその説明は適宜省略する。

20 【0088】パワースペクトル計算部1では、音声信号をデジタル化する。そして、デジタル化された音声信号のうち、30msに相当する個数のデータをまとめたものに対して、時間窓を掛け算したものをFFT(高速フーリエ変換)などの手段により短時間スペクトルに変換し、絶対値スペクトルとして基本周波数計算部2に送る。基本周波数計算部2では、パワースペクトル計算部1から送られてきた絶対値スペクトルを用いて、600Hzの幅を有する周波数領域での平滑化窓を畳み込むことにより、平滑化したスペクトルを求める。このパワースペクトル計算部1から送られてきた絶対値スペクトルを、この平滑化スペクトルで、対応する周波数ごとに割算することにより、平坦化された絶対値スペクトルを求める。つまり、(パワースペクトル計算部1から与えられた絶対値スペクトル)/(基本周波数計算部2で求めた平滑化スペクトル)=(平坦化された絶対値スペクトル)、である。

30 【0089】次に、平坦化された絶対値スペクトルの1000Hz以下をガウス分布の形状を有する低域通過フィルタ特性とかけ合わせたものを2乗したものを逆フーリエ変換することにより、正規化され平滑化された自己相関関数を求める。この相関関数を、パワースペクトル計算部1で用いた時間窓の自己相関関数で正規化した正規化相関関数の最大値を探索することにより、音声の基本周期の初期推定値を求める。次いで、この正規化相関関数の最大値の前後の点を合わせた3点の値に放物線を当てはめることにより、音声信号のデジタル化のための標本化周期よりも詳細に基本周波数を推定する。また、パワースペクトル計算部1から与えられる絶対値スペクトルのパワーが少なかったり、正規化相関関数の最大値が小さいなどの理由で周期的な音声部分ではないと判定される場合には、基本周波数の値を0としておくことに

より、その旨を記録する。パワースペクトル計算部 1 および基本周波数計算部 2 は、第 1 の処理（音声の基本周波数の抽出）を行なう部分である。このような第 1 の処理を、1 ms ごとに繰返し連続的に行なう。

【0090】なおこの基本周波数計算部では、実施の形態 1 で説明したように、既存の一般的な手法を用いても、音声波形の視認による手作業によっても良い。

【0091】適応的周波数分析部 9 では、基本周波数計算部 2 で求めた基本周波数の値に基づいて、時間窓の周波数分解能と基本周波数の比と、時間窓の時間分解能と基本周期の比とが同じになるような時間窓を設計する。具体的には、時間窓の関数形を決めた後、時間分解能と周波数分解能の積が一定の値になることを利用する。時間窓の大きさは、スペクトルの分析を行なうごとに基本周波数計算部 2 で求められた基本周波数を用いて更新する。このようにして設計された時間窓を使って、スペクトルを求める。適応的周波数分析部 9 は、第 1 . 5 の処理（基本周期に適応した周波数分析）を行なう部分である。平滑化スペクトログラム計算部 10 では、信号の基本周波数についての情報に基づいて、信号の基本周波数の 2 倍の周波数幅を有する三角形の補間関数を求める。そして、この補間関数と、適応的周波数分析部 3 で求められたスペクトルを周波数方向で畳み込む。次いで、基本周期の 2 倍の時間長を有する三角形の補間関数を用いて、先に周波数方向で補間したスペクトルを時間方向で補間することにより、時間・周波数平面の格子点の間を双 1 次関数の曲面で埋めた平滑化スペクトログラムを求める。平滑化スペクトログラム計算部 10 は、第 2 の処理（基本周波数の情報を利用した補間関数の適応）を行なう部分である。平滑化スペクトログラム計算部 10 までの処理によって、音声信号は、平滑化スペクトログラムと、精密な基本周波数情報の 2 つに分解される。平滑化スペクトログラム変換部 11 および音源情報変換部 6 は、第 3 の処理（音声パラメタの変換）を行なう部分である。位相調整部 7 および波形合成部 8 は、第 4 の処理（変換された音声パラメタによる音声合成）を行なう部分である。

【0092】図 9 は、平滑化前のスペクトログラムを示す図である。図 10 は、平滑化スペクトログラムを示す図である。図 9 および図 10 を参照して、横軸は時間 (ms) を示し、縦軸は周波数を表わす指標を示す。図 11 は、図 9 の一部を立体的に示した図である。図 12 は、図 10 の一部を立体的に示す図である。図 11 および図 12 を参照して、A 軸は時間を示し、B 軸は周波数を示し、C 軸は強度を示している。

【0093】図 9 および図 11 を参照して、周波数成分の相互干渉による零点が観測できる。この零点は、図 9 では、「白点」になっており、図 11 では、「凹み」になっている。図 10 および図 12 を参照して、零点が消えているのが観測できる。すなわち、スペクトログラム

は平滑化されており、周期性の影響が除かれているのが観測できる。

【0094】以上のように、実施の形態 2 による音声変換方法では、分析するスペクトルの周波数方向だけでなく、時間方向に対しても平滑化を行なう。つまり、分析するスペクトログラムを平滑化する。したがって、分析するスペクトログラムの時間方向および周波数方向の周期性の影響を小さくできる。このため、周波数分解能をいわずらに高くする必要がなく、分析するスペクトログラムの時間方向の細かい変化が潰れることはない。つまり、バランスよく周波数分解能および時間分解能を決定できる。

【0095】また、実施の形態 2 による音声変換方法は、実施の形態 1 による音声変換方法の処理をすべて含んでいる。このため、実施の形態 2 による音声変換方法は、実施の形態 1 による音声変換方法と同様に効果を奏する。さらに、実施の形態 2 による音声変換方法では、スペクトルを平滑化するのではなく、スペクトログラムを平滑化している。このため、実施の形態 2 による音声変換方法では、実施の形態 1 による音声変換方法の効果と同様の内容の効果を奏するが、その効果は、実施の形態 1 による音声変換方法に比べて顕著である。

【0096】[実施の形態 3] 実施の形態 1 では、平滑化スペクトル計算部 3 における平滑化の対象とするスペクトルが、基本周波数計算部 2 での周波数分析のときに用いる時間窓により既に平滑化されているという問題を無視していた。このように既にある程度平滑化されているスペクトルを補間関数を用いた畳み込みによりさらに平滑化することで、平滑化が二重に行なわれることとなり、音声の時間周波数特性を表わす曲面（音声を生成する機構を表わす時間周波数曲面）の周波数軸に沿った断面（スペクトル）の微細な構造がならされてしまうという問題が生じる。微細構造がなされてしまうことの影響は、原音声との比較試聴により、音声の個人性の微妙なニュアンスの劣化、声の張りの劣化および音韻の明瞭性の劣化として認められる。

【0097】このような過剰平滑化の問題を回避するためには、「中島隆之・鈴木虎三，“パワースペクトル包絡 (PSE) 音声分析・合成系”，日本音響学会誌 44 巻 11 号 (1988)，pp. 824 - 832」(以下、「文献 1」と呼ぶ) に記載されているように、節点の値だけを用いて、スペクトルのモデルを適合させるという方法がある。しかし、実際の音声では信号が正確には周期的ではなくさまざまな揺らぎや雑音を含むことから必然的に、文献 1 の適用範囲が限られてくるという問題が生じる。実施の形態 3 による信号分析方法としての音声分析方法は、以上のような問題を解決するために、以下のような処理を行なう。

【0098】(処理) 処理 1 について説明する。元の時間周波数特性を表わす曲面（音声を生成する機構を表わ

す時間周波数曲面)が、スプライン信号空間として知られる区分的多項式により構成される空間の直積として表わされる空間の要素であると仮定する。そして、時間窓の影響を受けたスペクトログラムから元の時間周波数特性を表わす曲面を最適近似する曲面を計算する最適な補間関数を求める。この最適な補間関数を用いて時間周波数特性を計算する。以下、このような処理 1 について詳しく説明する。

【0099】音声の時間周波数特性を表わす曲面(音声を生成する機構を表わす時間周波数曲面)が、時間方向の区分的多項式により構成される空間と、周波数方向の区分的多項式により構成される空間との積として表わされる曲面であるとする。たとえば、実施の形態 1 では、音声の時間周波数特性を表わす曲面が、時間方向の区分的 1 次式と、周波数方向の区分的 1 次式との積で表わされるとした。このような多項式の平行移動により、「寅市和男・岩城護, 区分的多項式からなる信号空間における周期標本化双直交基底, 電子情報通信学会論文誌, 92/6, Vol. J75-A, No. 6, pp. 1003-1012」(以下、「文献 2」と呼ぶ)に記載されているように、ある有限な観測区間上の自乗可積分な関数が構成する空間 L 2 の部分空間に基底を構成することができる。以下では、説明を簡単にするために時間周波数表現の周波数軸に沿った断面である周波数スペクトルについて議論する。時間軸についても同様に議論を進めればよい。

【0100】周波数軸の最適な補間関数に要請される条件は、空間 L 2 の部分空間の要素である 1 つの基底に相当するスペクトルが時間窓操作に対応する周波数領域での平滑化操作によって平滑化されたスペクトルに変換されたものに対して最適な補間関数が適用されたときに、元の基底(空間 L 2 の部分空間の要素である 1 つの基底)に相当するスペクトルが回復されることである。文献 2 に記載されているように、空間 L 2 の部分空間の要素は、基底による展開係数からなるベクトルと等価である。したがって、最適な補間関数に要請される条件は、最適な補間関数を元の基底(空間 L 2 の部分空間の要素である 1 つの基底)に相当するスペクトルに時間窓操作に対応する周波数領域での平滑化操作を施して求められる平滑化されたスペクトルに適用した結果の節点上での値が 1 箇所だけ非零となるように最適な補間関数を決めることと等価になる。最適な補間関数は、同じ空間内にあるため、基底の組合せとして表現されている。すなわち、最適な補間関数は、時間窓操作を施して求められたスペクトルの節点上での値からなる係数ベクトルと畳み込んだときに最大値に対応する係数の部分だけが非負となり他は 0 となるようなベクトルの要素を係数として基底を組合せたものとして求められる。このようにして求められた周波数軸上の最適な補間関数を用いることで、過剰平滑化の影響を除去できる。

【0101】処理 2 について説明する。処理 2 は、処理 2 - 1 と処理 2 - 2 に分けられる。処理 1 で求めた周波数軸上の最適な補間関数は、負の係数を含んでいるため、元のスペクトルの形状によっては補間後のスペクトルにも負の部分が生ずることがある。スペクトルに負の部分が生ずると、直線位相の場合には問題はないが、最小位相のインパルスを求める際に位相の不連続による長時間の応答を生じ異音の原因となる。また、これを避けるために負の部分で置換すると正から負に移り変わる部分において導関数の不連続(特異点)が生じ、比較的長い時間の応答を生じ異音の原因となる。この問題を解決するため、処理 2 - 1 を行なう。処理 2 - 1 では、 $(-\infty, \infty)$  の領域を  $(0, \infty)$  の領域に写像する単調で滑らかな関数を用いて、周波数軸上の最適な補間関数で補間されたスペクトルを変換する。

【0102】しかし、処理 2 - 1 だけでは次のような問題が生じる。音声のスペクトルは周波数帯域によってその中に含まれるエネルギーが大きく異なり、その比は 10000 倍を超える場合もある。人間の知覚では、それぞれの帯域における変動は、その帯域の平均的なエネルギーとの相対的な比率に比例して知覚される。このため、エネルギーの小さな帯域では、近似の誤差に伴う雑音もはつきりと知覚されることになる。したがって、補間を行なう際にすべての帯域において同じ精度で近似を行なうと、エネルギーの小さな帯域での近似の誤差が目立つことになる。このような問題を解決するために、処理 2 - 2 を行なう。処理 2 - 2 では、元のスペクトルを平滑化したスペクトルで正規化する。

【0103】以上のことをまとめる。処理 2 - 2 で正規化したスペクトルを対象に、周波数軸上の最適な補間関数で補間を行なう。これによって、近似の誤差は各帯域で知覚的に一様となる。また、このような正規化によりスペクトルの平均値は 1 となるため、 $(-\infty, \infty)$  の領域を  $(0, \infty)$  の領域に写像する単調で滑らかな関数を用いて、周波数軸上の最適な補間関数で補間されたスペクトルを、非負かつスペクトル上に特異点を持たないスペクトルに変換することができる(処理 2 - 1)。

【0104】(具体的処理)図 13 は、本発明の実施の形態 3 による音声分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。図 13 を参照して、音声分析装置は、マイク 101、アナログ/デジタル変換器 103、基本周波数分析部 105、基本周波数適応周波数分析部 107、概形スペクトル計算部 109、正規化スペクトル計算部 111、平滑化変換正規化スペクトル計算部 113 および逆変換・概形スペクトル復元部 115 を備える。この音声分析装置は、図 4 のパワースペクトル計算部 1、基本周波数計算部 2 および平滑化スペクトル計算部 3 からなる周波数分析装置と置換することができる。この場合、図 4 の平滑化スペクトル変換部 5 では、平滑化スペクトルの代わりに最適補間

平滑化スペクトル 1 1 9 を用いることになる。

【 0 1 0 5 】 図 1 3 を参照して、音声は、マイク 1 0 1 により、音波に対応する電気信号に変換される。この電気信号は、そのまま用いても、あるいは、一旦何らかの録音装置に収録してから再生して用いてもよい。次に、マイク 1 0 1 からの電気信号は、アナログ/デジタル変換器 1 0 3 によって、標本化されデジタル化されて、一連の数値の列として表わされる音声波形となる。音声波形の標本化周波数としては、たとえば、高品質の拡声電話の場合には、1 6 kHz、音楽や放送への利用を考慮する場合には、3 2 kHz、4 4 . 1 kHz、4 8 kHz などを用いる。標本化に伴う量子化は、たとえば、1 6 ビットとする。

【 0 1 0 6 】 基本周波数分析部 1 0 5 において、アナログ/デジタル変換器 1 0 3 から与えられる音声波形の基本周波数あるいは基本周期が抽出される。基本周波数あるいは基本周期の抽出は、さまざまな方法を利用することができる。その一例を述べる。4 0 ms の  $\cos^2$  窓によって切り出された音声のパワースペクトルを、周波数方向の平滑化関数との畳み込みによって平滑化したスペクトルで割算する。このようにして計算された概形が平坦なパワースペクトルを、たとえば、1 kHz 以下に周波数方向のガウス窓によって帯域制限した後に、フーリ \*

$$w(t) = e^{-\pi(t/\tau_0)^2} \tag{13}$$

$$W(\omega) = \frac{\tau_0}{\sqrt{2\pi}} e^{-\pi(\omega/\omega_0)^2} \tag{14}$$

【 0 1 0 9 】 ここで、t は時間であり、 $\omega$  は角周波数であり、 $\omega_0$  は基本角周波数である。 $\omega_0 = 2 \pi f_0$  であり、 $\tau_0 = 1 / f_0$  である。 $f_0$  は、基本周波数あり、 $T_0$  は、基本周期である。

【 0 1 1 0 】 基本周波数適応周波数分析部 1 0 7 における周波数分析の結果得られたパワースペクトルは、概形スペクトル計算部 1 0 9 において、たとえば、基本周波数の 6 倍の幅を持つ三角形の周波数領域の窓関数との畳み込みにより高度の平滑化を受けて、基本周波数の影響を除去された概形スペクトルにされる。正規化スペクトル計算部 1 1 1 において、基本周波数適応周波数分析部 1 0 7 で求められたパワースペクトルを、概形スペクトル計算部 1 0 9 で求められた概形スペクトルで割算することにより、それぞれの帯域における近似誤差に対する知覚的感度が一樣となるような正規化スペクトルが求められる。こうして求められた正規化スペクトルは、大局的には平坦な周波数特性を有するが、音声の周期性に基づく細かな凸凹や音韻の特徴を表わすホルマントと呼ばれるスペクトル上の局所的な山の形が認められるものとなる。このように、正規化スペクトル計算部 1 1 1 では、上述した処理 2 - 2 を行なう。

【 0 1 1 1 】 正規化スペクトル計算部 1 1 1 で得られた正規化スペクトルは、平滑化変換正規化スペクトル計算

\* 工逆変換して得られる変形自己相関関数の最大値の位置を求める。この最大値の位置と前後の点からなる近傍の 3 点を用いた放物線補間によって詳細な最大値の位置を求めることで、精密に基本周期を求めることができる。この基本周期の逆数を求めれば、それが基本周波数となる。変形自己相関の値は、周期性が完全であれば 1 となるので、この値の大きさを周期性の确实さの指標として用いることができる。

【 0 1 0 7 】 このようにして抽出された基本周波数あるいは基本周期の情報 (音源情報 1 1 7) を利用して、アナログ/デジタル変換器 1 0 3 からの音声波形は、基本周波数適応周波数分析部 1 0 7 において、基本周波数に適応して窓の長さが決まる時間窓により周波数分析される。最適補間平滑化スペクトル 1 1 9 だけを求めるのであれば、窓の長さを基本周波数に適応させて変化させる必要はないが、後で最適補間平滑化スペクトログラムを求めることが必要な場合には、基本周波数に適応した長さを有するガウス窓を用いることが最適となる。具体的には、次のようにして計算される窓を用いる。この要請を満たす窓関数  $w(t)$  は次のようなガウス関数となり、そのフーリエ変換  $W(\omega)$  は、次式で与えられる。

【 0 1 0 8 】

【 数 1 4 】

部 1 1 3 において、各周波数の値に対する単調な非線形変換を受ける。そして、非線形変換を受けた正規化スペクトルは、時間窓と非線形変換によって決まる下表に示す最適な重み係数により結び付けられて構成される図 1 4 に示す周波数軸上での最適な平滑化関数 1 2 1 と畳み込まれて平滑化変換正規化スペクトルの初期値とされる。この周波数軸上での最適な平滑化関数は、上述した処理 1 によって求められる。つまり、周波数軸上での最適な補間関数は、窓関数の周波数領域での表現と、周波数方向の区分的多項式により構成される空間の基底とから求められ、平滑化変換正規化スペクトルの初期値と、音声の時間周波数特性を表わす曲面の周波数軸に沿った断面との誤差を最小にする。なお、下表は、窓関数がガウス窓である場合の最適値を示している。また、図 1 4 および下表の例は、音声のスペクトルが 2 次の周期スプライン信号空間の信号であることを仮定した場合の最適な平滑化関数である。同様な係数および係数によって決められる平滑化関数は音声のスペクトルが一般に m 次の周期スプライン信号空間の信号であることを仮定しても求めることができる。

【 0 1 1 2 】

【 表 1 】

31	
位置	係 数
-3	-0.0241
-2	0.0985
-1	-0.4031
0	1.6495
1	-0.4031
2	0.0985
3	-0.0241

【0113】以上のようにして求められた平滑化変換正規化スペクトルの初期値には負の値が含まれる場合があ \* 10

$$\eta(x) = \frac{x + \log(2 \cosh x)}{2}$$

【0115】この (x) を用いて、平滑化変換正規化スペクトルの初期値を適当な係数を掛けて正規化した後に常に正の値をとるように変換する。このような変換によって得られたスペクトルを、正規化に用いた係数で割ることにより、平滑化変換正規化スペクトルを得る。

【0116】この平滑化変換正規化スペクトルは、逆変換・概形スペクトル復元部 115 において、平滑化変換正規化スペクトル計算部 113 で用いた非線形変換の逆変換を受け、再度、概形スペクトルと掛け合せられることにより、最適補間平滑化スペクトル 119 にされる。また、音源情報 117 に付随する情報として、有声音の場合には、基本周波数あるいは基本周期の情報が記録され、無声音や声の存在しない区間においては、0 が記録される。最適化補間平滑化スペクトル 119 は、ほぼ完全に元の音声の細かな情報まで保有し、かつ、滑らかである。

【0117】以上のような一連の処理を行なうことは、音声分析・音声合成の品質改善にとって非常に効果的である。また、最適補間平滑化スペクトル 119 を音声合成・音声変換に利用することで、合成音声・変換音声の品質を自然の音声と区別ができない程度に非常に高くすることができる。さらに、最適補間平滑化スペクトル 119 には、発声者の個人性や細かなニュアンスまでを保存した正確な音韻情報が安定に滑らかな形で表現されているので、音声の機械認識での情報表現、話者認識のための情報表現として使用した場合に、大きく性能が向上するという効果が期待される。また、音源の時間的微細構造の影響をほぼ完全に分離しているため、最適補間平滑化スペクトル 119 を逆フィルタとして用いることにより、音源の時間的微細構造のみを高精度に抽出することができる。これは、声質の診断や状態の判定などの応用に非常に有効である。また、実施の形態 1 による音声分析方法は、駆動音源の影響をうけない高精度の音声分析方法である。

【0118】[実施の形態 4] 実施の形態 2 では、音声信号の周期性を積極的に利用して時間周波数領域でのスペクトログラムの適応的補間により信号の時間周波数特

\* る。ここで、人間の聴覚は主にスペクトルの山についての感度が鋭いという性質を利用して、平滑化変換正規化スペクトルの初期値を、( - , ) の区間を ( 0 , ) の区間に写像する単調で滑らかな関数を用いて変換する。つまり、上述した処理 2 - 1 を行なう。具体的には、変換前の値を x、変換後の値を (x) とすると、次の式が条件を満たす。

【0114】

【数 15】

(15)

性を表わす曲面を求める方法に基づく音声変換方法により非常に高品質の音声変換が可能となった。しかし、注意深くヘッドホンを用いて原音声と比較試聴すると、声の張りや音韻性の劣化が認められた。この問題の主要な原因は、スペクトログラムの計算で必要となる時間窓による必然的な平滑化と適応的補間による平滑化が重なることによる過剰平滑化にある。

【0119】このような過剰平滑化の問題を詳しく説明する。実施の形態 2 では、音声の時間周波数特性を表わす曲面が、周波数方向と時間方向それぞれにおいて基本周波数と基本周期を格子間隔とする区分的 1 次関数で表わされる双 1 次曲面であると仮定した。そして、格子点の情報が与えられたときに区分的 1 次関数を求める演算を時間周波数領域での補間関数を用いた平滑化として実現することにより、実際の音声で遭遇する不完全な周期や非周期的な信号の場合でも破綻せずに安定に曲面を求めることを可能とした。しかし、この演算では平滑化の対象とするスペクトログラムが分析のときに用いる時間窓により既に平滑化されているという問題を無視していた。なぜなら、実施の形態 2 でも、大局的には元の曲面を保存するという条件が満たされていたからである。

【0120】実施の形態 2 においては、このように既にある程度平滑化されているものを補間関数を用いた畳み込みによりさらに平滑化することで、平滑化が二重に行なわれることとなり、曲面の微細な構造がならされてしまうという問題が生じる。微細構造がならされてしまうことの影響は、原音声との比較試聴により、音声の個人性の微妙なニュアンスの劣化、声の張りの劣化、および音韻の明瞭性の劣化として認められる。

【0121】このような過剰平滑化の問題を回避するためには、文献 1 に記載されているように、節点の値だけを用いて、スペクトルのモデルを適応させるという方法がある。しかし、文献 1 の方法では、時間周波数特性を考慮せず、ある一時刻でのスペクトルのモデルを提案しているにすぎない。このような方法では、時間方向の分解能が低下し、時間的な早い変化を捉えることができない。また、実際の音声では信号が正確には周期的ではな

20

30

40

50

くさまざまな雑音を含むことから必然的にこのような方法の適応範囲が限られている。また、文献 1 に記載されている方法を拡大解釈して、時間周波数分解能が音声の基本周期とマッチするような最適なガウス窓を用いて、時間周波数領域で等方的な格子点での値を求めたとしても、その値には相互に隣接する格子点からの影響が含まれており、そのまま用いたのでは、本来の時間周波数特性を表わす曲面を正確に復元することはできない。

【0122】実施の形態 4 では、上述したような過剰平滑化の影響を除いて、正しい時間周波数特性を表わす曲面を計算する方法を提案し、実施の形態 2 による音声変換方法の分析部分を改良する。さらに、実施の形態 4 では、音声の分析を必要とするさまざまな応用に対して、駆動音源の影響を受けない高精度の分析方法を提供する。以下、実施の形態 4 による信号分析方法としての音声分析方法について詳しく説明する。

【0123】(処理) 処理 3 について説明する。処理 3 では、処理 1 と同様にして時間軸上の最適な補間関数を求める。つまり、窓関数の時間領域での表現と、時間方向の区分的多項式により構成される空間の基底とから、時間軸上の最適な補間関数を求める。処理 4 について説明する。処理 4 は、処理 4 - 1 と処理 4 - 2 に分けられる。処理 3 で求めた時間軸上の最適な補間関数は、負の係数を含んでいるため、元のスペクトログラムの形状によっては補間後のスペクトログラムにも負の部分が生ずることがある。スペクトログラムに負の部分が生ずると、直線位相の場合には問題がないが、最小位相のインパルスを求める際に位相の不連続による長時間の応答を生じる原因となる。また、これを避けるために負の部分を零で置換えると正から負に移り変わる部分において導関数の不連続(特異点)が生じ、比較的長い時間の応答を生じ異音の原因となる。この問題を解決するため、処理 4 - 1 を行なう。処理 4 - 1 では、 $(-\infty, \infty)$  領域を  $(0, \infty)$  の領域に写像する単調で滑らかな関数を用いて、時間軸上の最適な補間関数で補間されたスペクトログラムを変換する。しかし、処理 4 - 1 だけでは次のような問題が生じる。音声のスペクトルは周波数帯域によってその中に含まれるエネルギーが大きく異なり、その比は 1 万倍を超える場合もある。人間の知覚では、それぞれの帯域における変動は、その帯域の平均なエネルギーとの相対的な比率に比例して知覚される。このため、エネルギーの小さな帯域では、近似の誤差に伴う雑音もはっきりと知覚されることになる。したがって、補間を行なう際にすべての帯域において同じ精度で近似を行なうと、エネルギーの小さな帯域での近似の誤差が目立つことになる。このような問題を解決するために、処理 4 - 2 を行なう。処理 4 - 2 では、元のスペクトログラムを平滑化したスペクトログラムで正規化する。

【0124】以上のことをまとめる。処理 4 - 2 で正規化したスペクトログラムを対象に、時間軸上の最適な補

間関数で補間を行なう。これによって、近似の誤差は各帯域で知覚的に様となる。また、このような正規化によりスペクトログラムの平均値は 1 となるため、 $(-\infty, \infty)$  の領域を  $(0, \infty)$  の領域に写像する単調で滑らかな関数を用いて、時間軸上の最適な補間関数で補間されたスペクトログラムを、非負でかつスペクトログラム上に特異点を持たないスペクトログラムに変換することができる(処理 4 - 1)。

【0125】(具体的処理) 図 15 は、本発明の実施の形態 4 による音声分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。なお、図 13 と同様の部分については同一の参照符号を付しその説明を適宜省略する。図 15 を参照して、この音声分析装置は、マイク 101、アナログ/デジタル変換器 103、基本周波数分析部 105、基本周波数適応周波数分析部 107、概形スペクトル計算部 109、正規化スペクトル計算部 111、平滑化変換正規化スペクトル計算部 113、逆変換・概形スペクトル復元部 115、概形スペクトログラム計算部 123、正規化スペクトログラム計算部 125、平滑化変換正規化スペクトログラム計算部 127、逆変換・概形スペクトログラム復元部 129 を備える。この音声分析装置は、図 8 のパワースペクトル計算部 1、基本周波数計算部 2、適応的周波数分析部 9 および平滑化スペクトログラム計算部 10 からなる音声分析装置と置換えることができる。この場合には、平滑化スペクトログラム変換部 11 では、平滑化スペクトログラムの代わりに最適補間平滑化スペクトログラム 131 を用いる。

【0126】図 15 を参照して、最適補間平滑化スペクトル 119 の計算は、分析周期ごとに行なわれる。音声の基本周波数として 500 Hz までを対象とするものとするれば、1 ms ごとに分析を行えばよい。こうして、たとえば、1 ms ごとに計算される最適補間平滑化スペクトル 119 を時間の順にならべていくことにより、最適補間平滑化スペクトルに基づいたスペクトログラムを求めることができる。しかし、このスペクトログラムは、時間方向での最適な補間平滑化を行っていないので、最適補間平滑化スペクトログラム 131 ではない。概形スペクトログラム計算部 123、正規化スペクトログラム計算部 125、平滑化変換正規化スペクトログラム計算部 127 および逆変換・概形スペクトログラム復元部 129 は、最適補間平滑化スペクトル 119 に基づいたスペクトログラムから、最適補間平滑化スペクトログラム 131 を計算するための部分である。

【0127】概形スペクトログラム計算部 123 において、最適補間平滑化スペクトル 119 に基づいたスペクトログラムの中から、現在の分析時点の前後 3 基本周期(合計 6 基本周期分)の区間を選択し、現在の時点を中心とする三角形の加重関数を用いて加重加算を行なって現時点での概形スペクトルの値を計算する。こうして計

算されたスペクトルを時間方向に並べることによって概形スペクトログラムを求める。つまり、最適補間平滑化スペクトル 1 1 9 に基づくスペクトログラムから、音声信号の周期性に基づく時間的変動の影響を除去したものが概形スペクトログラムである。

【0 1 2 8】正規化スペクトログラム計算部 1 2 5 において、最適補間平滑化スペクトル 1 1 9 に基づいたスペクトログラムを、概形スペクトログラム計算部 1 2 3 によって得られた概形スペクトログラムで割算し、正規化スペクトログラムを得る。このようにすることで、局所的な変動は残るが時間方向において場所ごとのレベルに応じて正規化が行なわれ、近似誤差の知覚的な影響が一様になる。このように、正規化スペクトログラム計算部 1 2 5 は、処理 4 - 2 を行なっている。

【0 1 2 9】平滑化変換正規化スペクトログラム計算部 1 2 7 において、正規化スペクトログラム計算部 1 2 5 で得られた正規化スペクトログラムは適当な単調な非線形変換を受ける。この非線形変換によって得られたスペクトログラムは、時間窓と非線形変換によって決まる表（実施の形態 3 で示した表）に示す最適な重み係数により結び付けられて構成される図 1 6 に示す時間軸上の最適な平滑化関数 1 3 3 との加重計算により、平滑化変換正規化スペクトログラムのスペクトル断面の初期値の集合とされる。このような時間軸上の最適な平滑化関数 1 \*

$$\eta(x) = \frac{x + \log(2 \cosh x)}{2}$$

【0 1 3 3】この (x) を用いて、平滑化変換正規化スペクトログラムのスペクトル断面の初期値を適当な係数を掛けて正規化した後に、常に正の値をとるように変換し、この変換によって得られたスペクトルを正規化に用いた係数で割る。この処理を、平滑化変換正規化スペクトログラムのスペクトル断面の初期値のすべてに対して行ない、複数のスペクトルを得る。この複数のスペクトルを時間方向にならべたものを平滑化変換正規化スペクトログラムとする。

【0 1 3 4】逆変換・概形スペクトログラム復元部 1 2 9 において、正規化変換正規化スペクトログラムは、平滑化変換正規化スペクトログラム計算部 1 2 7 で用いた非線形変換の逆変換を受け、再度概形スペクトログラムと掛け合せられることにより、最適補間平滑化スペクトログラム 1 3 1 にされる。

【0 1 3 5】以上のように実施の形態 4 による音声分析方法では、実施の形態 3 による音声分析方法の処理をすべて含む。このため、実施の形態 4 による音声分析方法は、実施の形態 3 による音声分析方法と同様の効果を奏する。ただし、実施の形態 4 による音声分析方法では、周波数方向のみならず時間方向をも考慮した処理を行なっている。つまり、実施の形態 3 で説明した処理 1 および処理 2 に加えて、処理 3 および処理 4 を行なっている。このため、実施の形態 4 による効果は、実施の形態

\* 3 3 は処理 3 によって求められ、平滑化変換正規化スペクトログラムのスペクトル断面の初期値と、音声の時間周波数特性を表わす曲面のスペクトル断面との誤差を最小にする。

【0 1 3 0】図 1 6 および実施の形態 3 で示した表の例は、音声のスペクトログラムの時間変化が 2 次の周期スプライン信号空間の信号であることを仮定した場合の最適な平滑化関数である。同様な係数および係数によって決められる平滑化関数は音声のスペクトログラムの時間変化が一般に m 次の周期スプライン信号空間の信号であることを仮定しても求めることができる。

【0 1 3 1】以上のようにして求められた平滑化変換正規化スペクトログラムのスペクトル断面の初期値には負の値が含まれる場合がある。ここで、人間の聴覚は主に音の立上がりについての感度が鋭いという性質を利用して、平滑化変換正規化スペクトログラムのスペクトル断面の初期値を、( - , ) の区間を ( 0 , ) の区間に写像する単調で滑らかな関数を用いて変換する。つまり、上述した処理 4 - 1 を行なう。具体的には、変換前の値を x、変換後の値を (x) とすると、次の式が条件を満たす。

【0 1 3 2】  
【数 1 6】

(16)

3 による音声分析方法よりも顕著である。したがって、実施の形態 4 による音声分析方法を用いることで、実施の形態 3 による音声分析方法を用いる場合に比べ、音声分析・音声合成の品質はさらに改善され、特に、子音の開始部分や発声の開始部分の生々しさが向上する。

【0 1 3 6】[実施の形態 5] 時間分解能と周波数分解能が基本周期およびおよそ基本周波数に対して同じ比率となるような等分解能の時間窓を用いた場合、周期的信号の調波の間の干渉により、周期的に零となる点がスペクトログラム上に生ずる。この零となる点は、隣り合う調波の位相が 1 基本周期で一巡するために、平均的に逆相となる部分が周期的に生ずるためである。実施の形態 2 による図 1 2 の説明で、実施の形態 2 による音声変換方法を用いることで、スペクトログラムの零となる点が消えるということを示した。なお、零となる点は、振幅が 0 になる点である。

【0 1 3 7】以上のような問題を解決するには、ちょうど零となる点の部分で最大の値となるようなスペクトログラムを与える窓関数を設計すればよい。そのような窓関数は無数にあるが、次のようにすれば具体的に構成できる。対象とする窓関数を、原点の両側に、相互の間隔を音声信号の基本周期分、離して配置する。そして、配置された一方の窓関数の符号を反転させる。符号を反転させた窓関数と、配置された他方の窓関数とを加え合せ

10

20

30

40

50

て、新たな窓関数を作る。この新たな窓関数の振幅は元の窓関数の半分とする。このようにして得られた新たな窓関数を用いることにより計算されるスペクトログラムは、元の窓関数を用いて得られたスペクトログラムの零となる点の位置に最大値を有し、元の窓関数を用いて得られたスペクトログラムが最大値を有する位置に零となる点を有するものとなる。元の窓関数を用いて計算したパワー表示のスペクトログラムと、新しく作成した窓関数を用いて計算したパワー表示のスペクトログラムとを、単調で非負な関数を加えた後、加え合せ、逆変換することにより、それぞれの零となる点と最大値は打消し合い、平坦で滑らかなスペクトログラムが求められる。以下、図面を参照しながら詳しく説明する。

【0138】図17は、本発明の実施の形態5による音声信号分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。図17を参照して、この音声分析装置は、パワースペクトル計算部137、適応時間窓作成部139、相補パワースペクトル計算部141、適応相補時間窓作成部143および非零パワース\*

$$w(t) = e^{-\pi(t/\tau_0)^2} \tag{17}$$

$$W(\omega) = \frac{\tau_0}{\sqrt{2\pi}} e^{-\pi(\omega/\omega_0)^2} \tag{18}$$

【0141】ここで、 $t$ は時間、 $\omega$ は角周波数、 $\omega_0$ は基本角周波数、 $\tau_0$ は基本周期である。そして、 $\omega_0 = 2\pi f_0$ 、 $\tau_0 = 1/f_0$ であり、 $f_0$ は基本周波数である。適応相補時間窓作成部143において、適応時間窓作成部139における適応時間窓の作成と同時に、適応時間窓に対して相補的な時間窓（以下、「適応相補時間窓」と呼ぶ）を作成する。つまり、適応時間窓と同じ形の窓関数を、原点の両側に相互の間隔を基本周期分だけ

$$w_d(t) = \frac{1}{2} \left( e^{-\pi\left(\frac{t-\tau_0/2}{\tau_0}\right)^2} - e^{-\pi\left(\frac{t+\tau_0/2}{\tau_0}\right)^2} \right) \tag{19}$$

【0143】図18は、適応時間窓 $w(t)$ および適応相補時間窓 $w_d(t)$ を示す図である。図19は、適応時間窓 $w(t)$ および適応相補時間窓 $w_d(t)$ に対応する実際の音声波形を示す図である。図18および図19を参照して、縦軸は振幅を示し、横軸は時間(ms)を示す。図18の適応時間窓 $w(t)$ および適応相補時間窓 $w_d(t)$ は、図19の音声波形(女性の声「オ」の一部)135の基本周波数に対応する。

【0144】再び図17を参照して、パワースペクトル計算部137において、適応時間窓作成部139で作成した適応時間窓を用いて、音声波形135を周波数分析し、パワースペクトルを求める。同時に、相補パワースペクトル計算部141において、適応相補時間窓作成部

$$P_{nz}^2(\omega) = P^2(\omega) + P_c^2(\omega) \tag{20}$$

\*ペクトル計算部145を備える。図13および図15の基本周波数適応周波数分析部107は、図17の音声分析装置と置換えることができる。この場合には、図13の概形スペクトル計算部109および正規化スペクトル計算部111は、基本周波数適応周波数分析部107で得られたスペクトルの代わりに非零パワースペクトル147を用いることになる。なお、音源情報117は、図13の音源情報117と同じであり、音声波形135は、図13に示したアナログ/デジタル変換器103から与えられる。

【0139】音源情報117の基本周波数あるいは基本周期の情報に基づいて、適応時間窓作成部139において、基本周波数および基本周期に対する時間窓の時間分解能と周波数分解能が等しい関係になるような窓関数を作成する。この要請を満たす窓関数（以下、「適応時間窓」と呼ぶ） $w(t)$ は次のようなガウス関数となり、そのフーリエ変換 $W(\omega)$ は、次式で与えられる。

【0140】  
【数17】

け離して配置する。そして、配置した一方の窓関数の符号を反転させたものと、配置した他方の窓関数とを加え合せたものとして、適応相補時間窓 $w_d(t)$ を作成する。振幅は元の窓関数（適応時間窓）の半分とする。適応相補時間窓 $w_d(t)$ を、ガウス窓の場合について具体的に書けば、次のようになる。

【0142】  
【数18】

143によって作成した適応相補時間窓を用いて、音声波形135を周波数分析し、相補パワースペクトルを求める。

【0145】非零パワースペクトル計算部145において、パワースペクトル計算部137で求めたパワースペクトル $P^2(\omega)$ と、相補パワースペクトル計算部141で求めた相補パワースペクトル $P_c^2(\omega)$ とから次の計算により、非零パワースペクトル147を求める。ここで、非零パワースペクトル147を、 $P_{nz}^2(\omega)$ とする。

【0146】  
【数19】

【0147】こうして求めた複数の非零パワースペクトル147を時間的に並べることにより、非零パワースペクトログラムを求めることができる。

【0148】一定の周期のパルス列を分析した例を用いて、実施の形態5による音声分析方法の働きを示す。図20は、周期的パルス列に適応時間窓を用いて求められるパワースペクトル $P^2$  ( ) から構成される3次元スペクトログラム $P$  ( ) を示す図である。図21は、周期的パルス列に適応相補時間窓を用いて求められる相補パワースペクトル $P^2$  ( ) から構成される3次元相補スペクトログラム $P$  ( ) を示す図である。図22は、周期的パルス列の非零パワースペクトル $P_{nz}$  ( ) から構成される3次元非零スペクトログラム $P_{nz}$  ( ) を示す図である。図20～図22を参照して、AA軸は時間(尺度任意)を示し、BB軸は周波数(尺度任意)を示し、CC軸は、強度(振幅)を示している。図20を参照して、3次元スペクトログラム155は、零となる点の存在により、周期的に曲面の値が0に落ち込んでいる。図21を参照して、図20の3次元スペクトログラムにおいて零となる点の存在していた部分、3次元相補スペクトログラム157では、最大値となっている。図22を参照して、3次元スペクトログラム155および3次元相補スペクトログラム157の平均として得られた3次元非零スペクトログラム159は、零となる点がなく平坦に近い滑らかな形状となっている。

【0149】以上のように、実施の形態5による音声分析方法では、零となる点のないスペクトルおよび零となる点のないスペクトログラムを作成できる。このようにして作成された零となる点のないスペクトルを、図13の概形スペクトル計算部109および正規化スペクトル計算部111で用いることにより、実施の形態3による音声分析方法に比べて、音声の時間周波数特性を表わす曲面の周波数軸に沿った断面の近似精度をさらに改善することができる。また、零となる点のないスペクトログラムを、図15の概形スペクトル計算部109および正規化スペクトル計算部111で用いることにより、実施の形態4による音声分析方法に比べて、音声の時間周波数特性を表わす曲面の近似精度をさらに改善できる。なお、 $P^2$  ( ) の代わりに、 $P^2$  ( ) に  $(0 < C_f = 1)$  なる補正量を掛けたものを用いることにより、最終的に得られる最適補間平滑化スペクトログラムの近似を総合的に改善することができる。ここで、 $C_f$  は、位相の干渉を補正するための量である。

【0150】[実施の形態6] 実施の形態3～5では、適応的な窓の長さの調整を行なっている(図13および図15の基本周波数適応周波数分析部107ならびに図17の適応時間窓作成部139)。実施の形態6では、窓関数の長さの調整のための基本周波数が安定に求められない場合においても安定に動作することができるよう

に、分析位置の近傍における音声波形を駆動する事象の位置関係を用いて適応的に窓関数の長さを調整する方法を提案する。

【0151】本発明の実施の形態6による信号分析方法としての音声分析方法について簡単に説明する。実施の形態3および実施の形態4に示したような周波数軸上での最適な平滑化関数および時間軸上での最適な平滑化関数を用いて、過剰平滑化の影響を排除する場合において、その効果を最もよく発揮させるためには、音声波形を最初に分析する場合の窓の長さを音声の基本周波数に対して一定の関係に設定することが望ましい。この要請を満たす窓関数 $w(t)$ は、式(13)や式(17)のようなガウス関数となり、そのフーリエ変換 $W$  ( ) は、式(14)や式(18)のようになる。式(13)や式(17)の窓関数 $w(t)$ の中に入って実質的に分析結果に影響を及ぼすのは、最大で2基本周期分であり、大部分の場合は、1つの基本周期分の波形が入るだけである。したがって、実施の形態6による音声分析方法では、有声音のように主要な励振がはっきりとしている場合には、現在の分析中心を挟む2つの励振の時間間隔をとして用いる。以下、詳しく説明する。

【0152】図23は、本発明の実施の形態6による音声分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。図23を参照して、この音声分析装置は、駆動点抽出部161、駆動点依存適応時間窓作成部163および適応パワースペクトル計算部165を備える。図13および図15の基本周波数適応周波数分析部107ならびに図17の適応時間窓作成部139は、図23に示した音声分析装置で置換えることができる。この場合には、図13および図15の概形スペクトル計算部109および正規化スペクトル計算部111では、基本周波数適応周波数分析部107で得られたパワースペクトルの代わりに適応パワースペクトル167を用いることになる。なお、音源情報117は、図13の音源情報117と同様のものである。音声波形135は、図13および図15のアナログ/デジタル変換器103から与えられる音声波形と同様のものである。図24は、図23の音声波形135の一例を示す図である。図23を参照して、縦軸は振幅を示し、横軸は時間(ms)を示す。

【0153】図23の音声分析装置は、適応時間窓の作成において基本周波数情報ではなく、分析位置の近傍にある音声波形から波形の駆動時点の情報求めて、分析位置と駆動点の相対関係に基づいて適切な窓関数の長さを決める音声分析方法を実現する。駆動点抽出部161において、音源情報117から信頼できる値に基づいて、平均的な基本周波数を求め、その基本周波数の2倍、4倍、8倍、16倍に対応する適応相補窓関数(図18に示した適応相補窓関数 $w_a(t)$ と同じ方法によって作成された窓関数)を、振幅を2倍しながら組合

せて、声門閉止検出用の関数を作成する。そして、声門閉止検出用の関数と、音声波形（図 2 4 参照）を畳み込むことによって、声門閉止において極大値をとる信号を得る。この信号の極大値に基づいて駆動点を求める。駆動点は、周期的に声門が閉じる時刻である。図 2 5 は、声門閉止において極大値をとる信号を示す図である。縦軸は振幅を示し、横軸は時間（ms）を示している。曲線 1 6 9 は、声門閉止において極大値をとる信号を示す。

【0 1 5 4】再び図 2 3 を参照して、駆動点依存適応時間窓作成部 1 6 3 においては、駆動点抽出部 1 6 1 で得られた駆動点の情報に基づいて、現在の分析時点を挟む駆動点の間の時間間隔を基本周期  $T_0$  とみなして、窓の長さを適応的に決める。適応パワースペクトル計算部 1 6 5 においては、駆動点依存適応時間窓作成部 1 6 3 で得られた窓を用いて周波数分析を行ない、適応パワースペクトル 1 6 7 を求める。

【0 1 5 5】実施の形態 6 による音声分析方法を、実施の形態 3 ~ 実施の形態 5 による音声分析方法に適用することによって、適応的な窓関数の長さの調整のための基本周波数が安定に求められない場合においても、安定した効果を得ることができる。つまり、適応的な窓関数の長さの調整のための基本周波数が安定に求められない場合においても、実施の形態 3 ~ 実施の形態 5 による音声分析方法の効果が損なわれることはない。

【0 1 5 6】

【発明の効果】この発明の第 1 の発明に係る周期信号変換方法では、連続的なスペクトル、つまり、平滑化スペクトルを用いて周期信号を別の信号に変換している。このため、周波数方向の周期性の影響が小さくなる。

【0 1 5 7】この発明の第 2 の発明に係る周期信号変換方法では、平滑化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できる。

【0 1 5 8】この発明の第 3 の発明に係る音変換方法では、位相調整成分から得られる音源信号は、インパルスと同じパワースペクトルを有し、時間的にエネルギーが分散している。このため、自然な音色を与えることができる。しかも、このような位相調整成分を利用することで、音の標準化周期よりも高い分解能で、精密に音程を設定できる。

【0 1 5 9】この発明の第 4 の発明に係る信号分析方法では、最適な周波数方向の補間関数によって補間を行なうことで、過剰平滑化の影響が取り除かれ、スペクトルの微細な構造がならされてしまうという弊害を防止できる。

【0 1 6 0】この発明の第 4 の発明に係る信号分析方法では、好ましくは、最適な時間方向の補間関数を用いて

補間を行なうことで、過剰な平滑化の影響を排除ことができ、スペクトログラムの微細な構造がならされてしまうという弊害を防止できる。

【0 1 6 1】この発明の第 5 の発明に係る信号分析方法では、第 1 の窓関数を用いて得られた第 1 のスペクトルと、第 1 の窓関数に対し相補的な第 2 の窓関数を用いて得られた第 2 のスペクトルとの平均値を、自乗あるいは単調で非負な関数による変換を介して求め、求めた自乗あるいは単調で非負な関数による変換を介した平均値を第 3 のスペクトルとする。こうして求めた第 3 のスペクトルには、零となる点が存在しない。

【図面の簡単な説明】

【図 1】位相調整成分  $2$  ( ) を用いて作成した音源信号を示す図である。

【図 2】位相調整成分  $3$  ( ) を用いて作成した音源信号を示す図である。

【図 3】位相調整成分  $2$  ( ) と位相調整成分  $3$  ( ) とを掛け合わせることによって作り出した位相調整成分を用いて作成した音源信号を示す図である。

【図 4】本発明の実施の形態 1 による音声変換方法を実現するための音声変換装置を示す概略ブロック図である。

【図 5】図 4 のパワースペクトル計算部で求められたパワースペクトルおよび平滑化スペクトル計算部で求められた平滑化スペクトルを示す図である。

【図 6】最小位相のインパルス応答  $v(t)$  を示す図である。

【図 7】変換されて合成された信号を示す図である。

【図 8】本発明の実施の形態 2 による音声変換方法を実現するための音声変換装置を示す概略ブロック図である。

【図 9】平滑化前のスペクトログラムを示す図である。

【図 1 0】平滑化されたスペクトログラムを示す図である。

【図 1 1】図 9 のスペクトログラムの一部を、立体的に示す図である。

【図 1 2】図 1 0 のスペクトログラムの一部を、立体的に示す図である。

【図 1 3】本発明の実施の形態 3 による音声分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。

【図 1 4】図 1 3 の平滑化変換正規化スペクトル計算部で用いる周波数軸上での最適な補間平滑化関数を示す図である。

【図 1 5】本発明の実施の形態 4 による信号分析方法を実現するための信号分析装置の全体構成を示す概略ブロック図である。

【図 1 6】図 1 5 の平滑化変換正規化スペクトログラム計算部で用いる時間軸上での最適な補間平滑化関数を示す図である。

【図17】本発明の実施の形態5による音声分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。

【図18】図17の適応時間窓作成部で得られる適応時間窓 $w(t)$ および図17の適応相補時間窓作成部で得られる適応相補時間窓 $w_d(t)$ を示す図である

【図19】図17の音声波形の一例を示す図である。

【図20】周期的パルス列に、図18の適応時間窓 $w(t)$ を用いて求められるパワースペクトル $P^2(\ )$ から構成される3次元スペクトログラム $P(\ )$ を示す図である。

【図21】周期的パルス列に、図18の適応相補時間窓 $w_d(t)$ を用いて求められる相補パワースペクトル $P^2_c(\ )$ から構成される3次元相補スペクトログラム $P_c(\ )$ を示す図である。

【図22】図17の非零パワースペクトル計算部で得られた周期的パルス列の非零パワースペクトル $P^2_{nz}(\ )$ から構成される3次元非零スペクトログラム $P_{nz}(\ )$ を示す図である。

【図23】本発明の実施の形態6による音声分析方法を実現するための音声分析装置の全体構成を示す概略ブロック図である。

【図24】図23の音声波形の一例を示す図である。

【図25】図23の駆動点抽出部で得られた声門閉止において極大値をとる信号を示す図である。

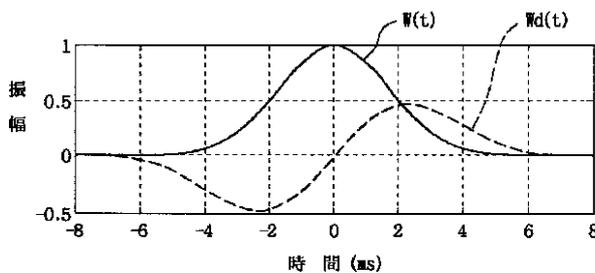
【符号の説明】

- 1 パワースペクトル計算部
- 2 基本周波数計算部
- 3 平滑化スペクトル計算部
- 4 インタフェース部
- 5 平滑化スペクトル変換部
- 6 音源情報変換部
- 7 位相調整部
- 8 波形合成部
- 9 適応的周波数分析部

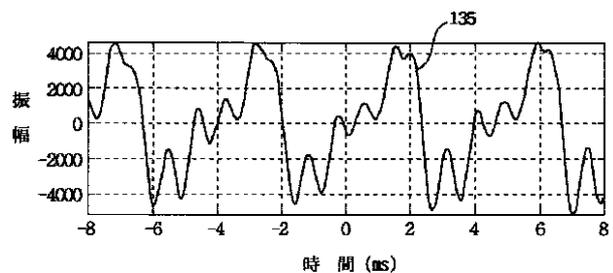
- \* 10 平滑化スペクトログラム計算部
- 11 平滑化スペクトログラム変換部
- 101 マイク
- 103 アナログ/デジタル変換器
- 105 基本周波数分析部
- 107 基本周波数適応周波数分析部
- 109 概形スペクトル計算部
- 111 正規化スペクトル計算部
- 113 平滑化変換正規化スペクトル計算部
- 115 逆変換・概形スペクトル復元部
- 117 音源情報
- 119 最適補間平滑化スペクトル
- 121 周波数軸上の最適な補間平滑化関数
- 123 概形スペクトログラム計算部
- 125 正規化スペクトログラム計算部
- 127 平滑化変換正規化スペクトログラム計算部
- 129 逆変換・概形スペクトログラム復元部
- 131 最適補間平滑化スペクトログラム
- 133 時間軸上の最適な補間平滑化関数
- 20 135 音声波形
- 137 パワースペクトル計算部
- 139 適応時間窓作成部
- 141 相補パワースペクトル計算部
- 143 適応相補時間窓作成部
- 145 非零パワースペクトル計算部
- 147 非零パワースペクトル
- 155 3次元パワースペクトログラム
- 157 3次元相補パワースペクトログラム
- 159 3次元非零パワースペクトログラム
- 30 161 駆動点抽出部
- 163 駆動点依存適応時間窓作成部
- 165 適応パワースペクトル計算部
- 167 適応パワースペクトル
- 169 声門閉止において極大値をとる信号

\*

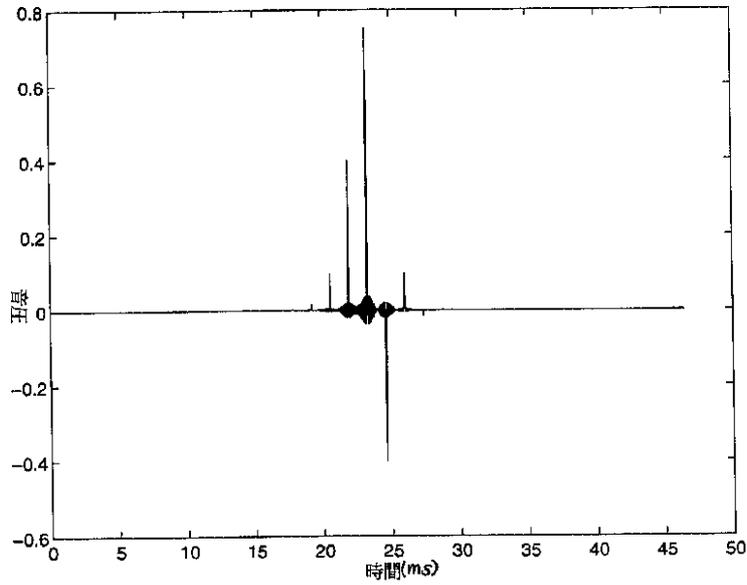
【図18】



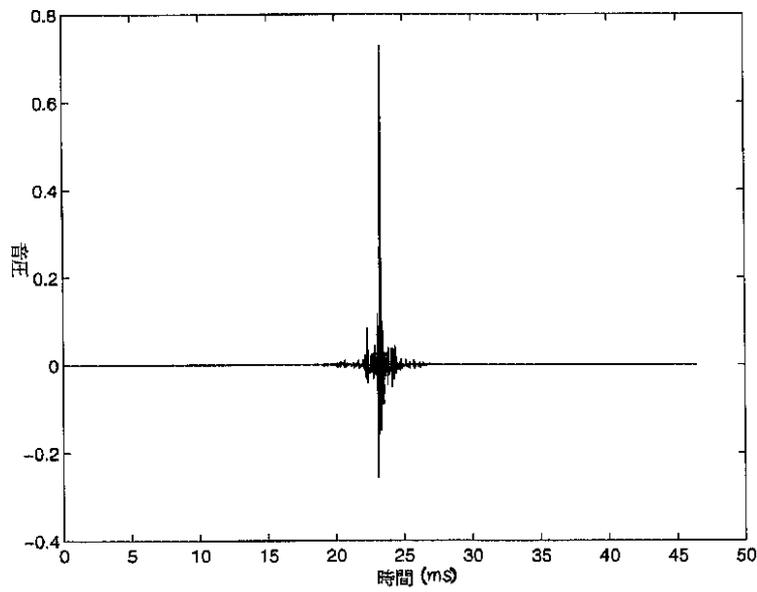
【図19】



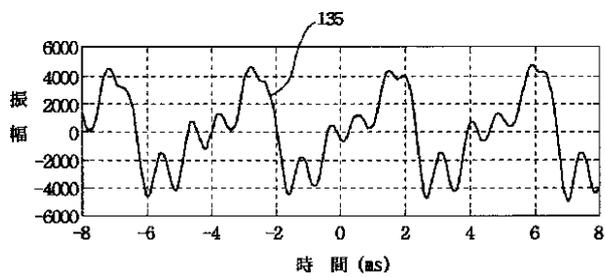
【図 1】



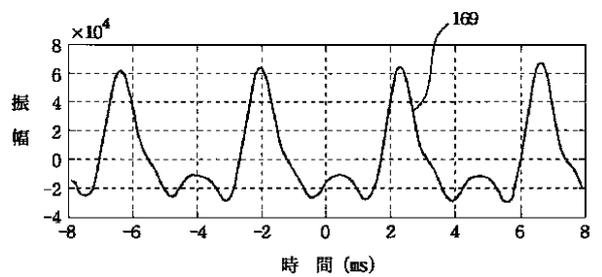
【図 2】



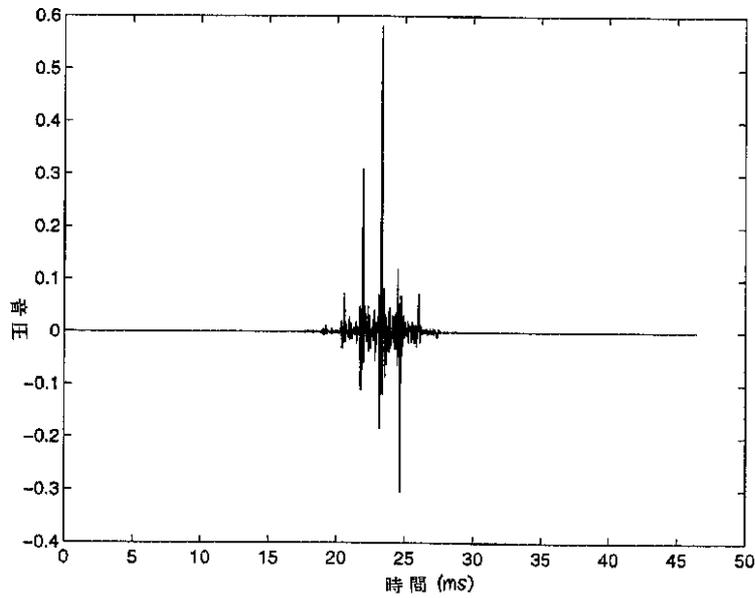
【図 2 4】



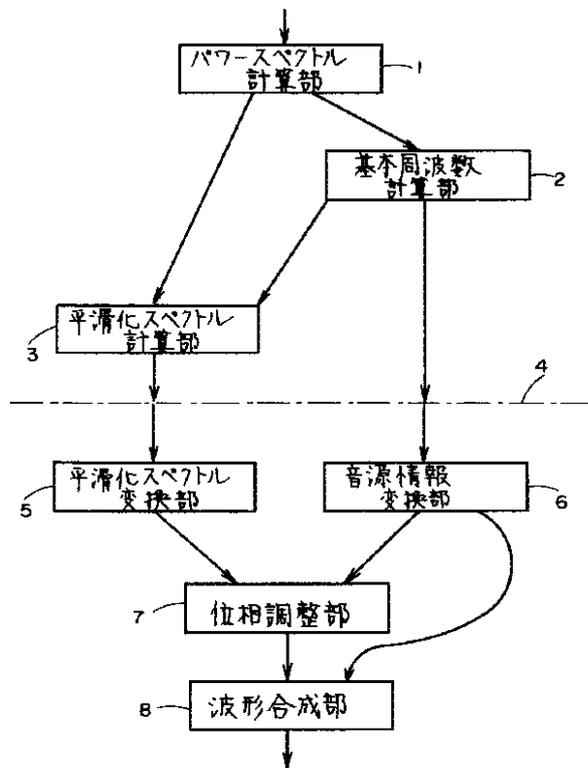
【図 2 5】



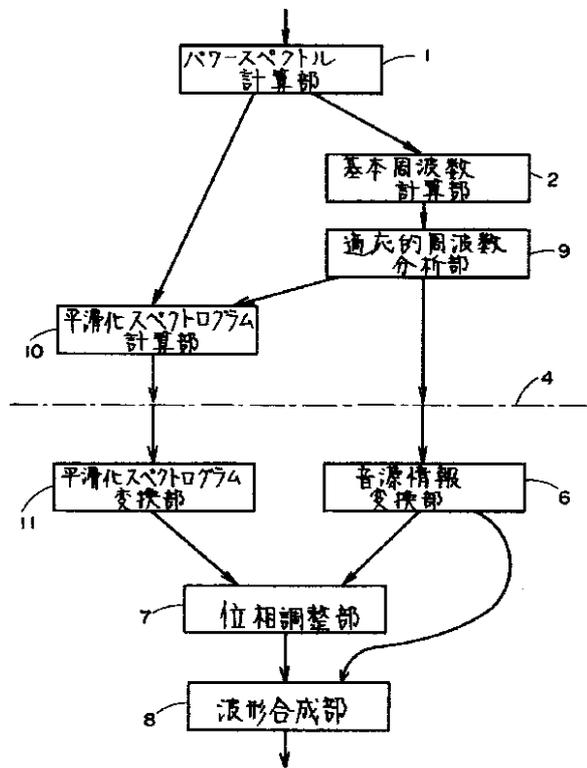
【図3】



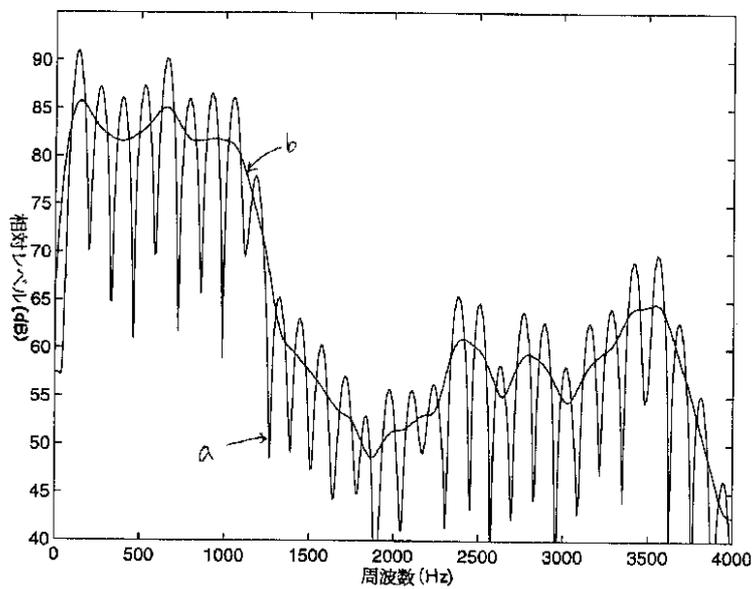
【図4】



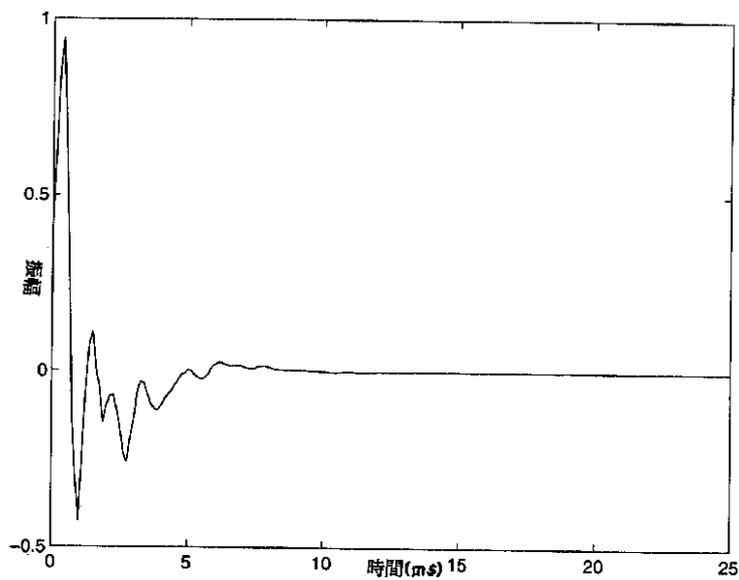
【図8】



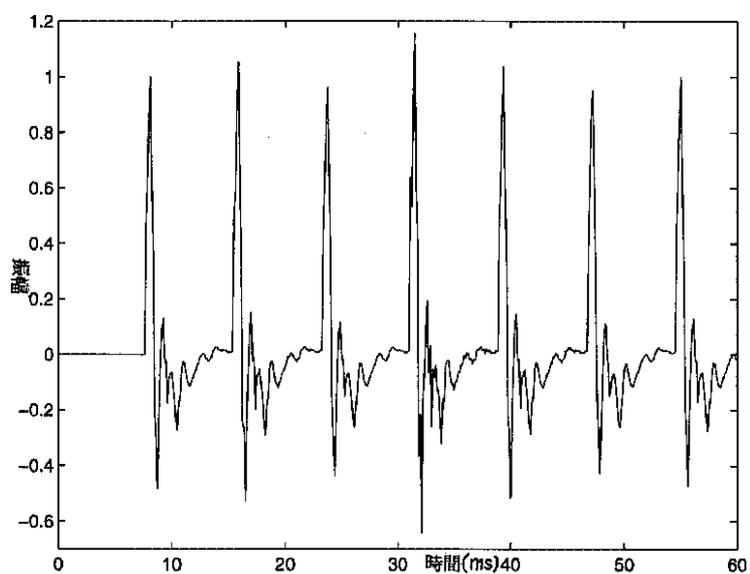
【図 5】



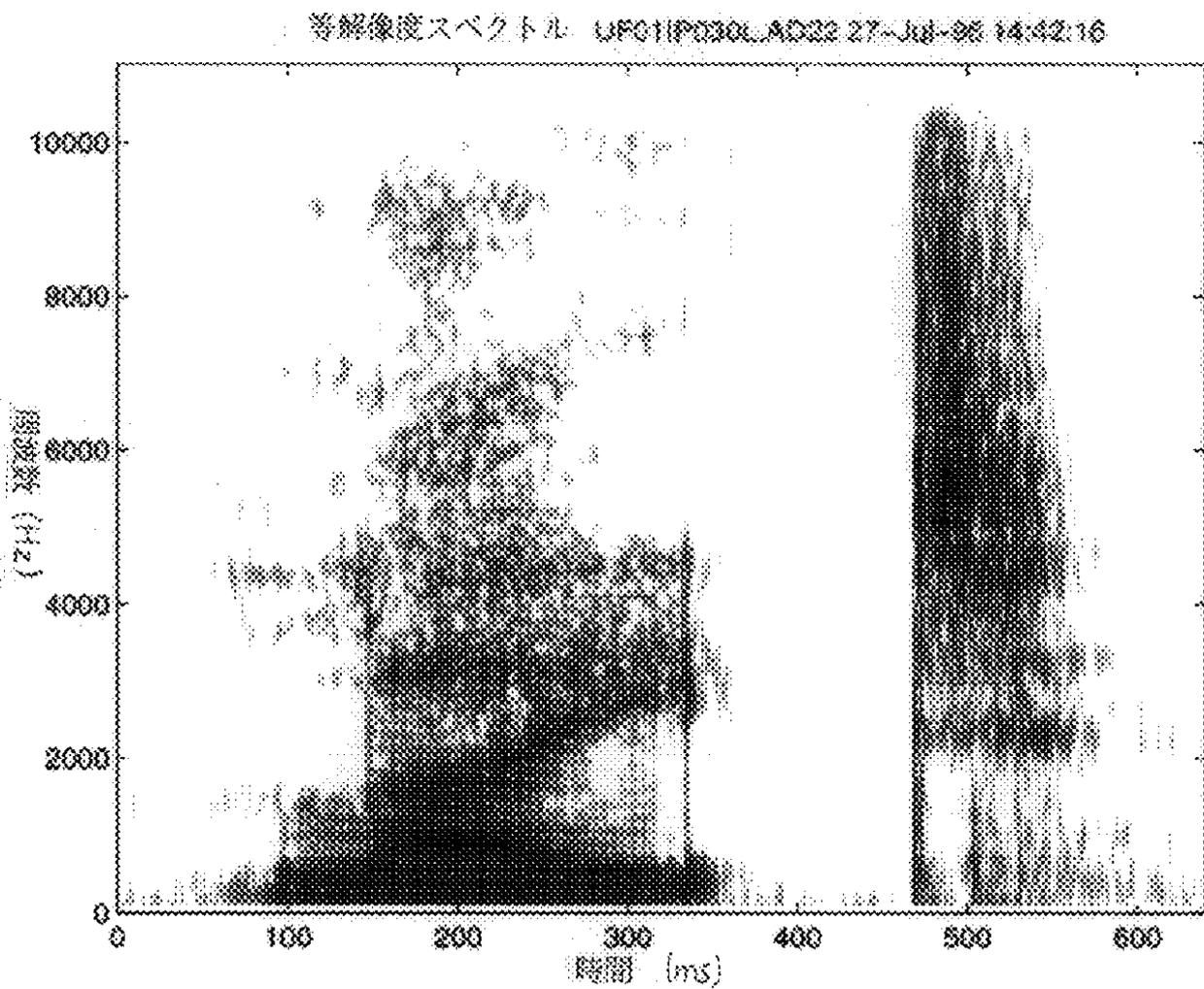
【図 6】



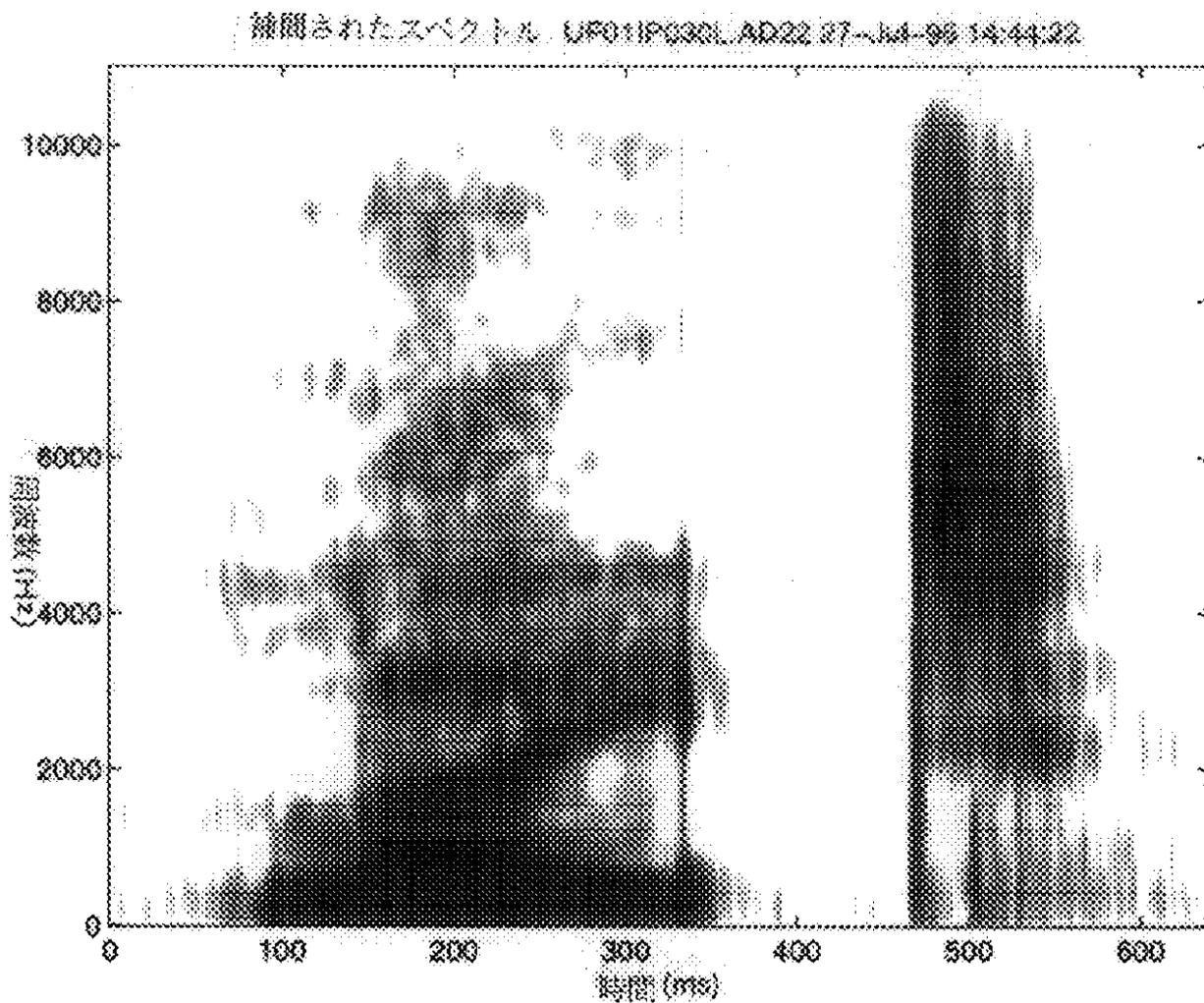
【図7】



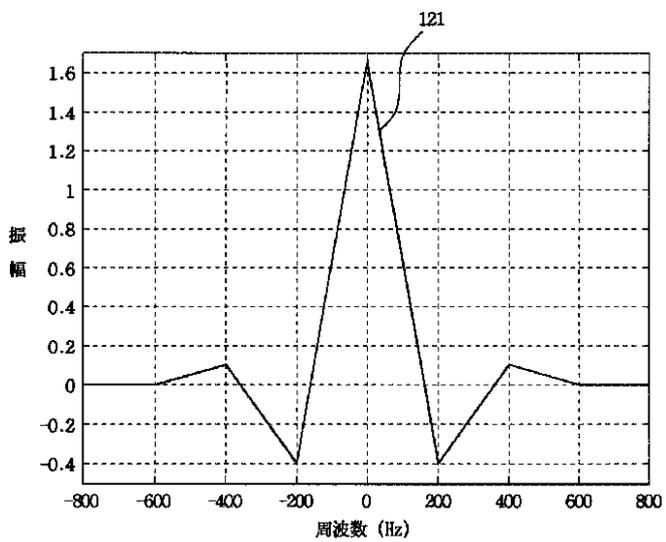
【図9】



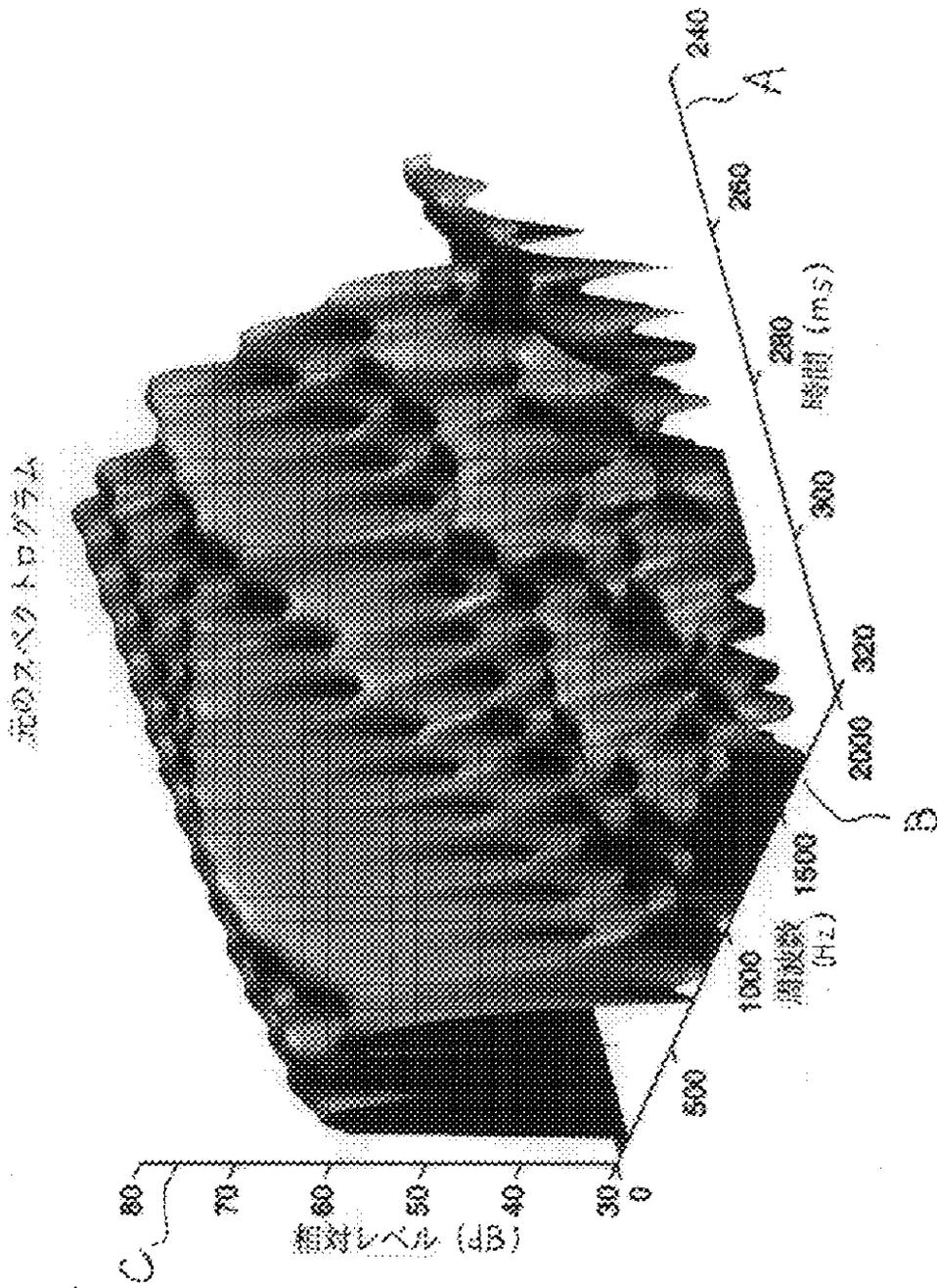
【図10】



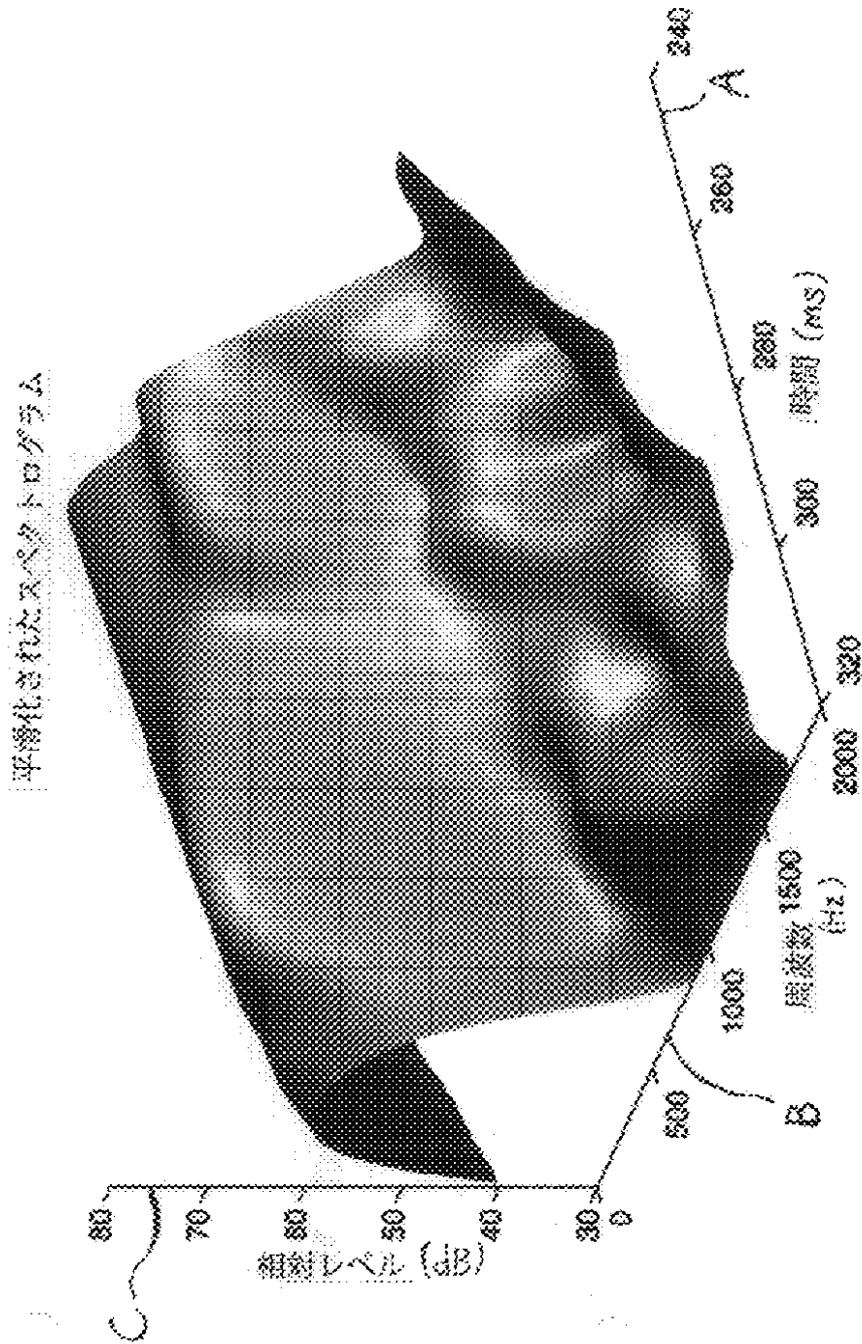
【図14】



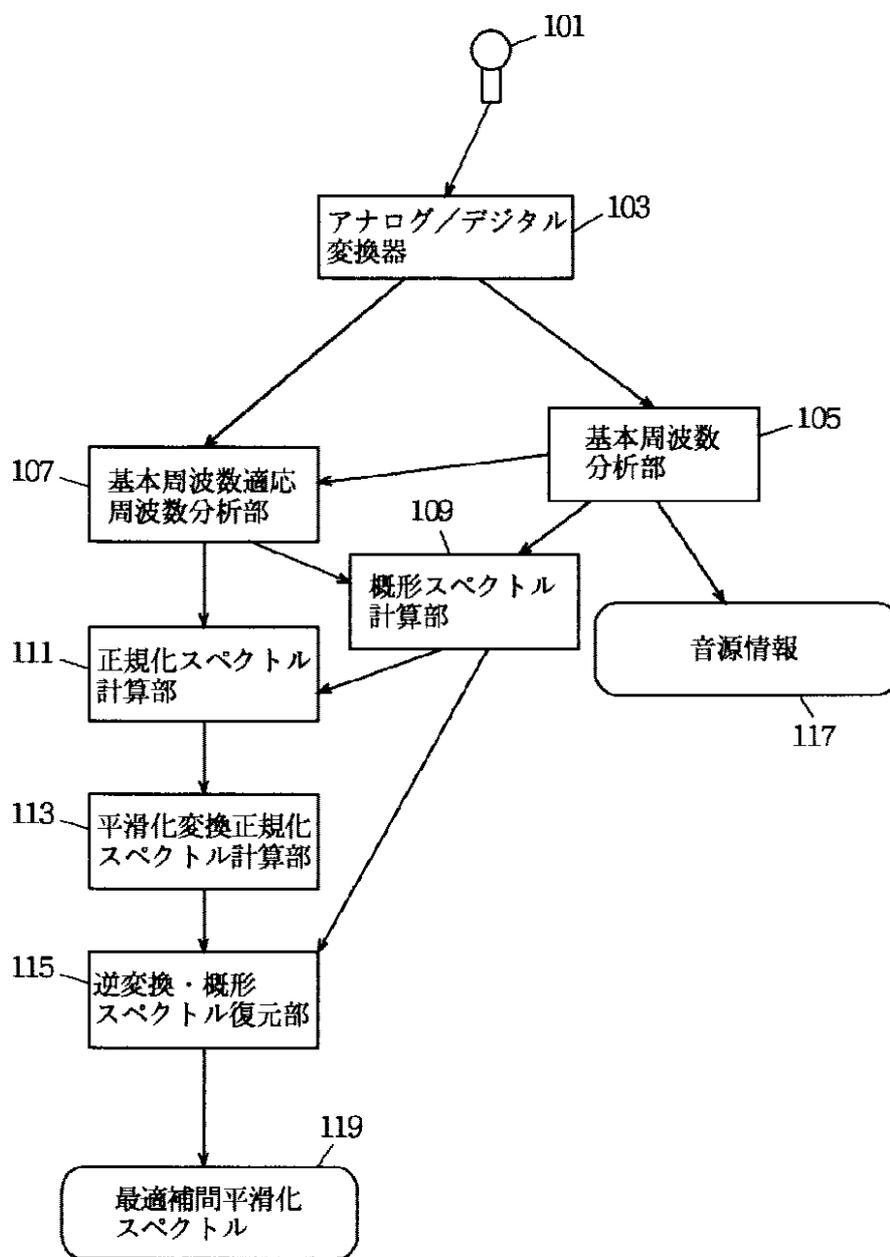
【図11】



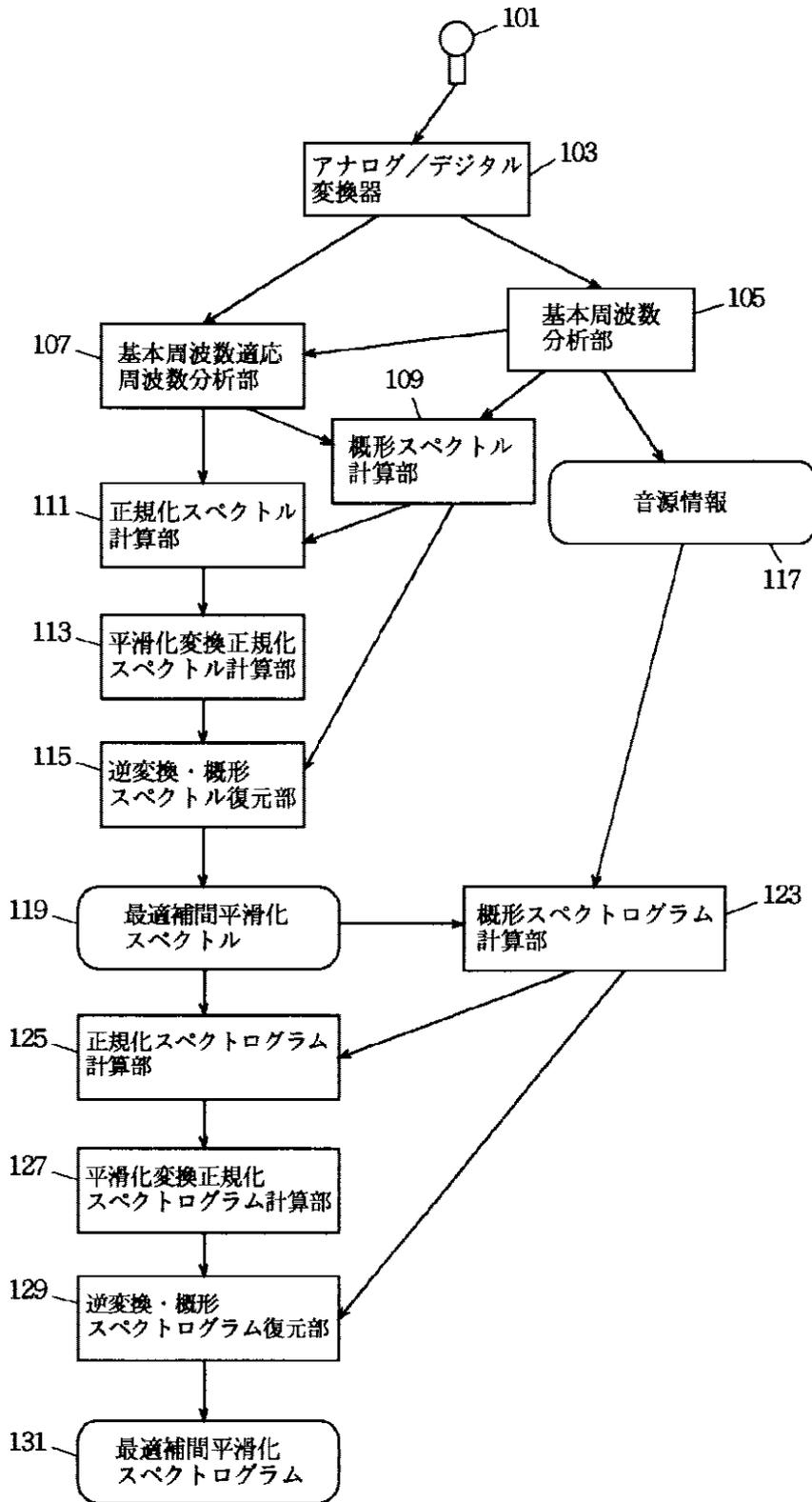
【図12】



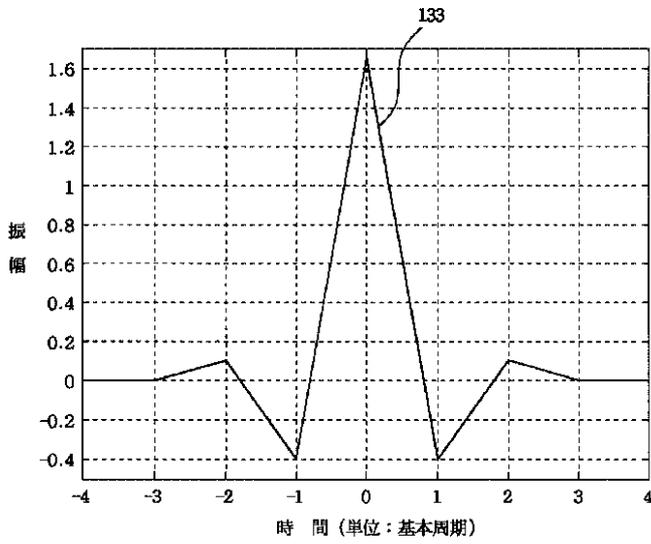
【図13】



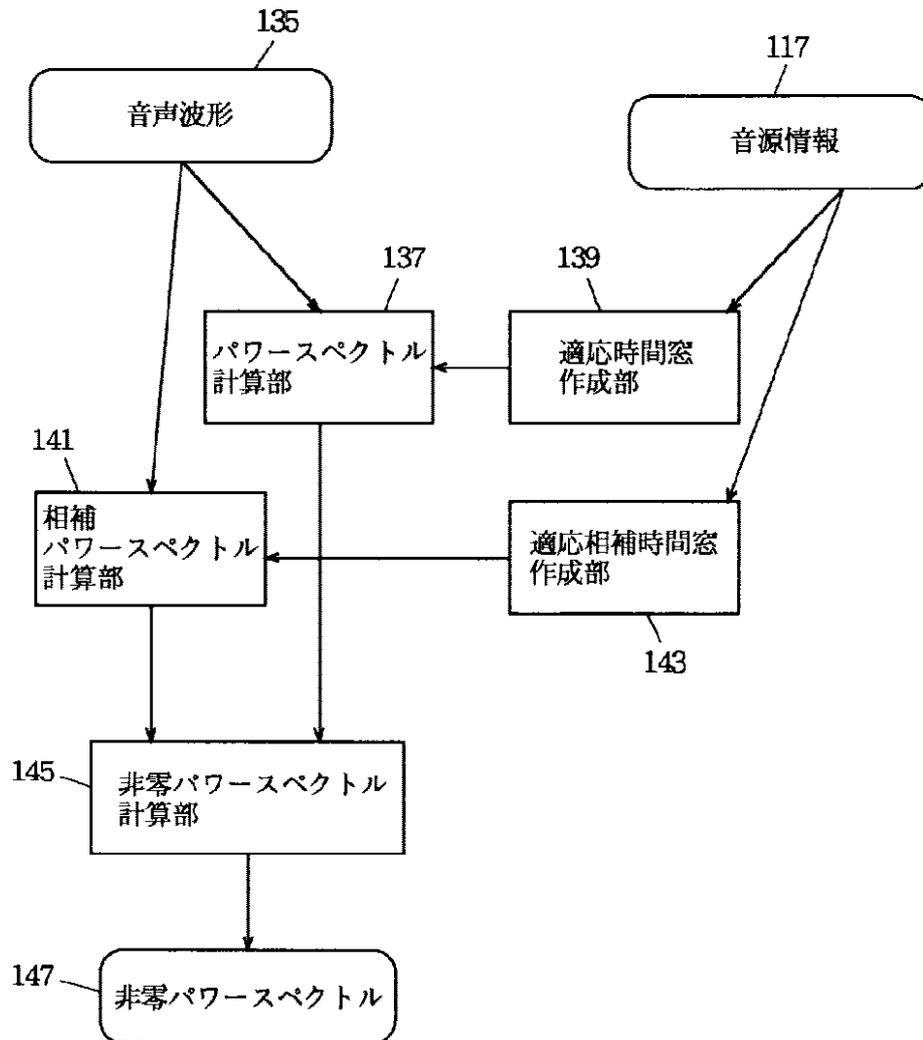
【図15】



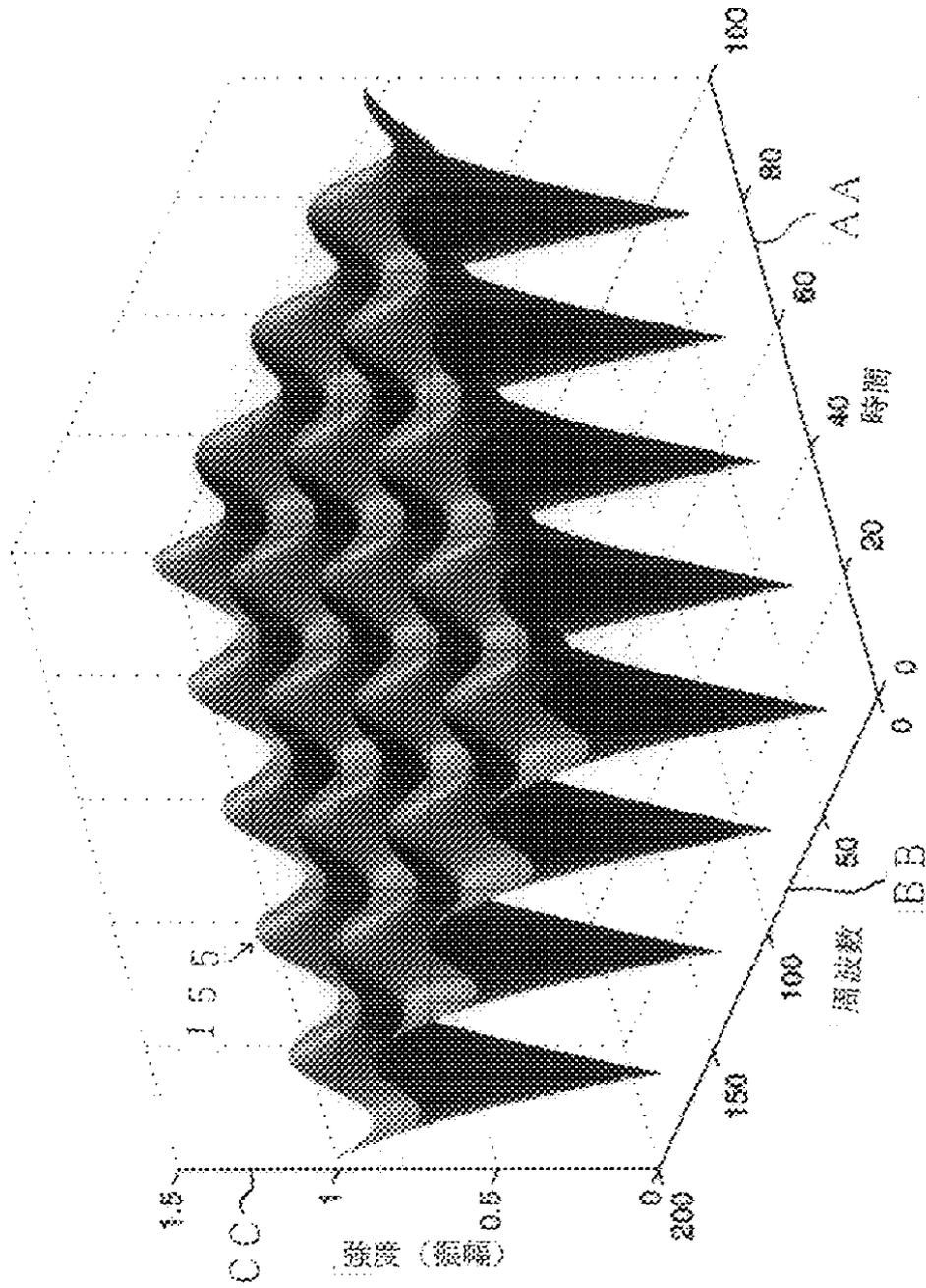
【図16】



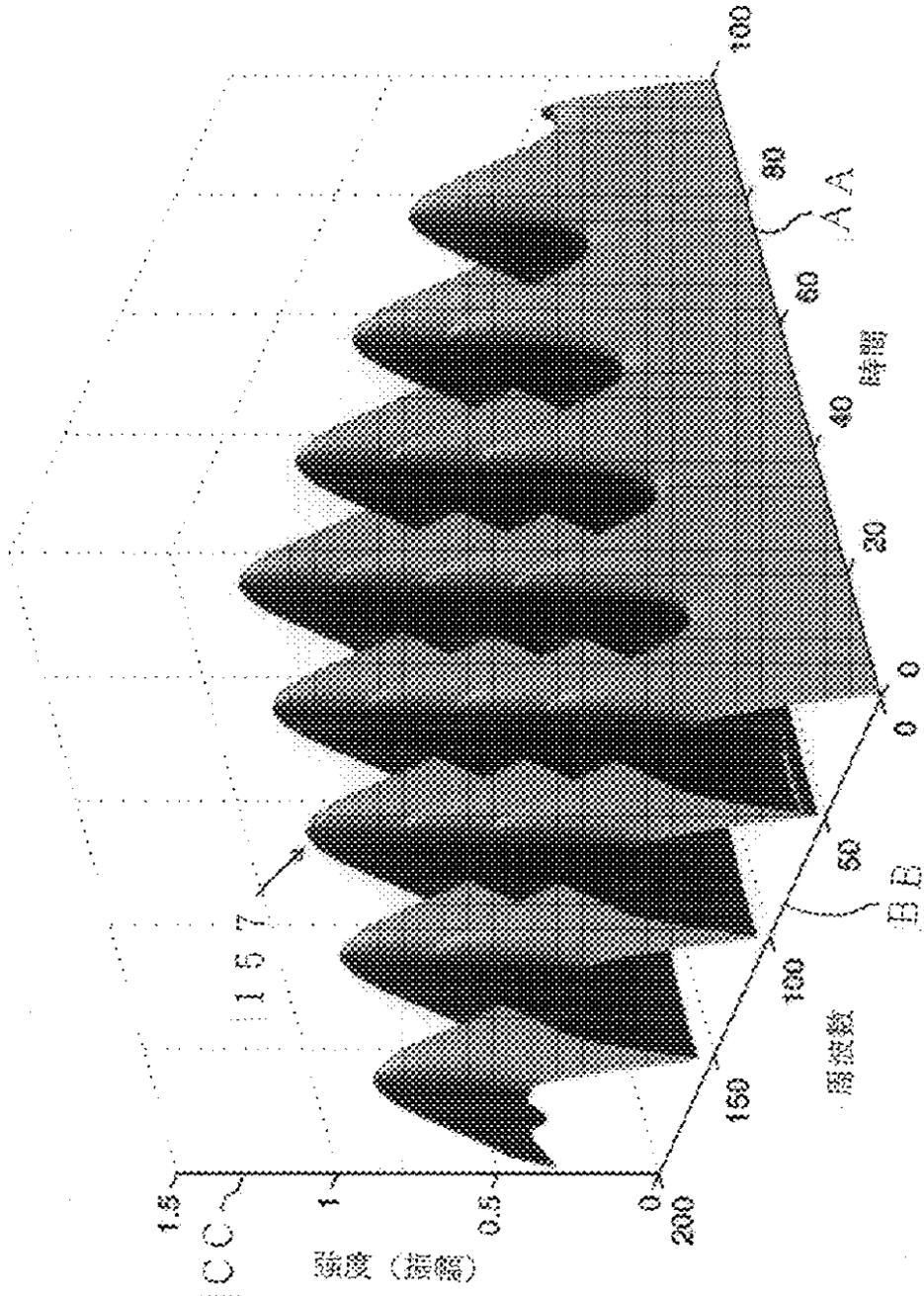
【図17】



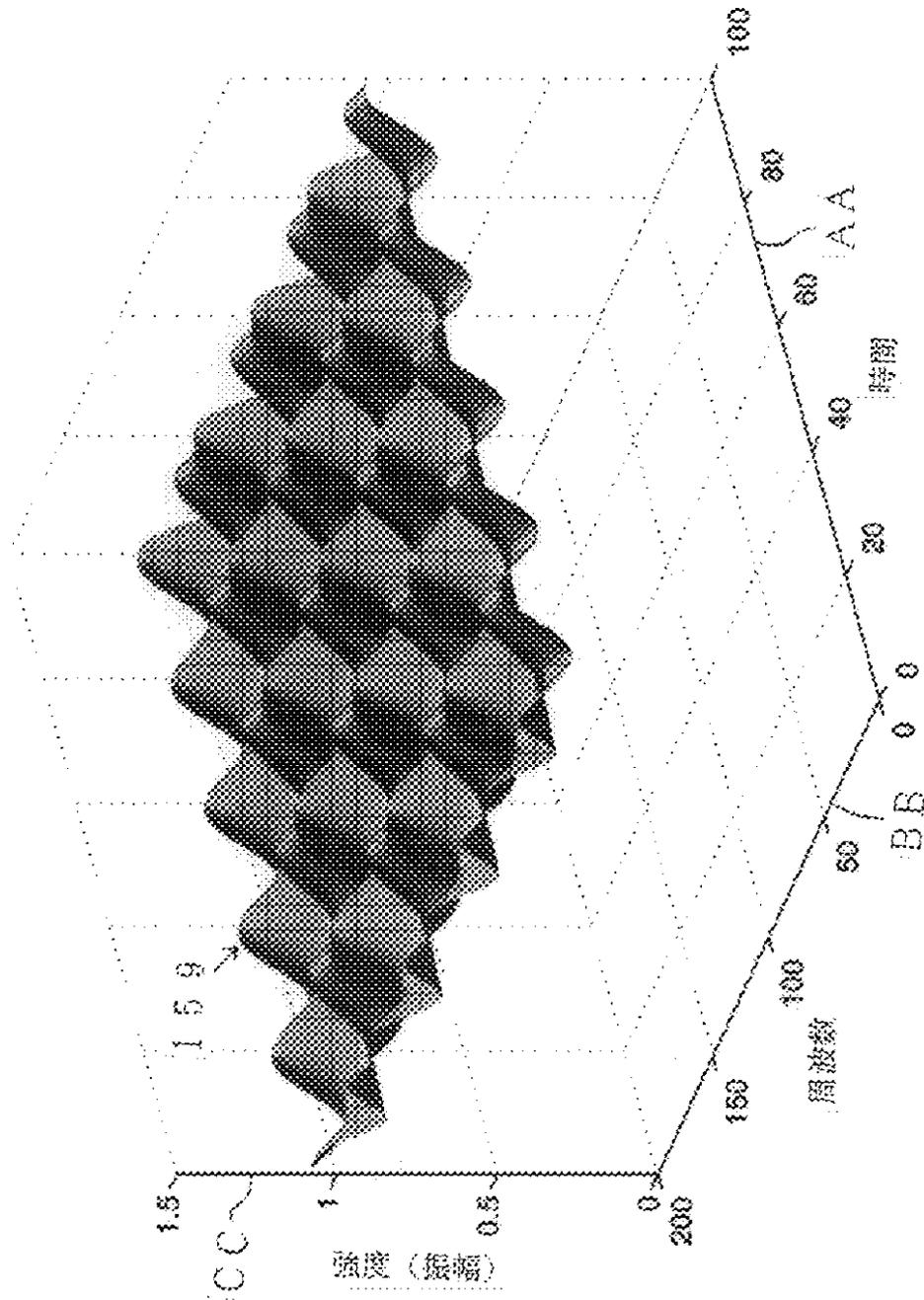
【図20】



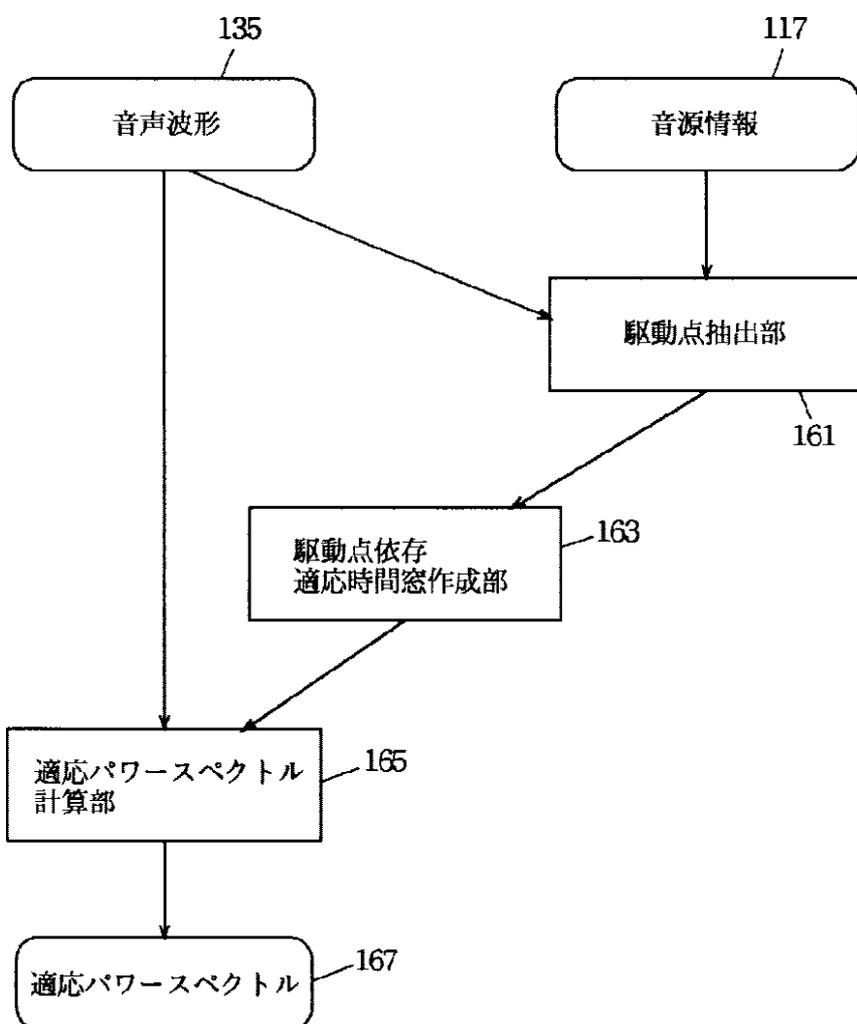
【図21】



【図22】



【図 2 3】



フロントページの続き

(56) 参考文献 特開 平 6 - 294830 ( J P , A )  
特開 昭 59 - 94795 ( J P , A )  
特開 平 3 - 259196 ( J P , A )  
特開 平 4 - 213500 ( J P , A )

(58) 調査した分野(Int.Cl.<sup>7</sup>, D B 名)  
G10L 11/00, 21/04  
G01R 23/16