

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6286933号
(P6286933)

(45) 発行日 平成30年3月7日 (2018.3.7)

(24) 登録日 平成30年2月16日 (2018.2.16)

(51) Int.Cl.

F I

G 1 O G 3/04 (2006.01)

G 1 O L 25/51 (2013.01)

G 1 O G 3/04

G 1 O L 25/51 3 0 0

請求項の数 7 (全 18 頁)

(21) 出願番号	特願2013-171660 (P2013-171660)	(73) 特許権者	000001443
(22) 出願日	平成25年8月21日 (2013.8.21)		カシオ計算機株式会社
(65) 公開番号	特開2015-40970 (P2015-40970A)		東京都渋谷区本町1丁目6番2号
(43) 公開日	平成27年3月2日 (2015.3.2)	(74) 代理人	100074099
審査請求日	平成28年8月10日 (2016.8.10)		弁理士 大菅 義之
		(72) 発明者	山谷 崇史
			東京都羽村市栄町3丁目2番1号 カシオ
			計算機株式会社羽村技術センター内
		審査官	大野 弘
		(56) 参考文献	特開2012-168538 (JP, A)
)
			最終頁に続く

(54) 【発明の名称】 小節間隔推定およびその推定のための特徴量抽出を行う装置、方法、およびプログラム

(57) 【特許請求の範囲】

【請求項1】

入力する音響信号データに対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出する時間周波数解析部と、

前記単位時間毎に、前記複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出する最大振幅抽出部と、

前記最大振幅抽出部により抽出された前記単位時間毎の前記最大振幅信号から、前記最大振幅信号の時系列データの自己相関信号を導出し、当該自己相関信号を特徴量として抽出する自己相関導出部と、

を備えることを特徴とする特徴量抽出装置。

10

【請求項2】

請求項1に記載の特徴量抽出装置によって抽出された自己相関信号を特徴量として、前記音響信号データの小節間隔を推定する装置であって、

前記自己相関信号の所定位置から始まる時間範囲内で、第1の時間長を有する時間範囲である第1の検査範囲を移動させながら、当該第1の検査範囲で指定される各時間範囲内で最大値をとる前記自己相関信号の各ピークを抽出するピーク抽出部と、

前記音響信号データが表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第2の時間長を有する時間範囲である第2の検査範囲を移動させながら、当該第2の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを前記抽出された各ピークから選択し、該選択したピークに基づいて前記強弱パターン周期を推定する強

20

弱パターン周期推定部と、

前記抽出された各ピークに基づいて、拍間隔を推定する拍間隔推定部と、

前記強弱パターン周期と前記拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する小節間隔計算部と、

を備えることを特徴とする小節間隔推定装置。

【請求項 3】

前記小節間隔計算部は、前記拍間隔に基づいて前記強弱パターン周期の拍数を算出し、当該拍数が所定値に比較して小さければ前記強弱パターン周期を前記小節間隔として算出し、当該拍数が所定値に比較して大きければ前記強弱パターン周期の半分の周期を前記小節間隔として算出する、

10

ことを特徴とする請求項 2 に記載の小節間隔推定装置。

【請求項 4】

入力する音響信号データに対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出し、

前記単位時間毎に、前記複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出し、

前記単位時間毎の最大振幅信号の時系列データの自己相関信号を導出し、当該自己相関信号を特徴量として抽出する、

ことを特徴とする特徴量抽出方法。

【請求項 5】

20

請求項 4 に記載の特徴量抽出方法によって抽出された自己相関信号を特徴量として、前記音響信号データの小節間隔を推定する方法であって、

前記自己相関信号の所定位置から始まる時間範囲内で、第 1 の時間長を有する時間範囲である第 1 の検査範囲を移動させながら、当該第 1 の検査範囲で指定される各時間範囲内で最大値をとる前記自己相関信号の各ピークを抽出し、

前記抽出された各ピークに基づいて、拍間隔を推定し、

前記音響信号データが表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第 2 の時間長を有する時間範囲である第 2 の検査範囲を移動させながら、当該第 2 の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを前記抽出された各ピークから選択し、該選択したピークに基づいて前記強弱パターン周期を推定し、

30

前記強弱パターン周期と前記拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する、

ことを特徴とする小節間隔推定方法。

【請求項 6】

コンピュータに、

入力する音響信号データに対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出し、

前記単位時間毎に、前記複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出し、

前記抽出された前記単位時間毎の前記最大振幅信号から、前記最大振幅信号の時系列データの自己相関信号を導出し、当該自己相関信号を特徴量として抽出する、

40

処理を実行させることを特徴とする特徴量抽出プログラム。

【請求項 7】

請求項 6 に記載の特徴量抽出プログラムによって抽出された自己相関信号を特徴量として、前記音響信号データの小節間隔を推定するプログラムであって、

コンピュータに、

前記自己相関信号の所定位置から始まる時間範囲内で、第 1 の時間長を有する時間範囲である第 1 の検査範囲を移動させながら、当該第 1 の検査範囲で指定される各時間範囲内で最大値をとる前記自己相関信号の各ピークを抽出し、

前記抽出された各ピークに基づいて、拍間隔を推定し、

50

前記音響信号データが表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第2の時間長を有する時間範囲である第2の検査範囲を移動させながら、当該第2の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを前記抽出された各ピークから選択し、該選択したピークに基づいて前記強弱パターン周期を推定し、

前記強弱パターン周期と前記拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する、

処理を実行させることを特徴とする小節間隔推定プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、小節間隔を推定し、またその推定のための特徴量を抽出する装置、方法、およびプログラムに関する。

【背景技術】

【0002】

音楽を構成する要素にテンポ（拍間隔）や小節間隔がある。拍間隔は一般的にBPM（1分間に4分音符が何拍あるか）で表し、楽曲全体の速さに対応する。小節は一般的に音楽の構造の最小単位で、コード（和音）やリズムパターンが多くの場合1小節毎（2小節の場合もある）に割り当てられる。従って、小節間隔を推定することは、コードを抽出して電子楽器の自動演奏を行ったり、音楽構造の分析を行ったりするために必要な技術である。

【0003】

小節間隔・拍間隔を推定する第1の従来技術として、バスドラムやスネアドラムなど打楽器が一定時間以上一定のビートを刻むことを前提としたものが提案されている。より具体的には例えば、楽曲データから周波数帯別の音圧データを作成し、この音圧データからリズムを最も刻む周波数帯を特定し、この特定した周波数帯の音圧データにおける変化周期をもとにリズム成分を推定する手法が提案されている（例えば特許文献1）。

【0004】

小節間隔・拍間隔を推定する第2の従来技術として、次のような手法が提案されている（例えば特許文献2）。まず、フレーム毎のパワー（自乗平均）を計算して、その微分値（実施例によればフレーム間の差分ではなく、前後複数フレームのパワーを最小2乗法で直線近似したときの傾きを微分値としてロバスト性を高めている）の自己相関を計算する。この相関値の極大（小節間隔が期待される範囲内で最大）を与える位置を小節間隔とする。拍子数は既知として拍間隔を計算する。パワーの微分値が大きいフレームが発音時刻に相当し、自己相関と極大位置検出で周期の推定になっている。

【0005】

小節間隔・拍間隔を推定する第3の従来技術として、次のような手法が提案されている（例えば特許文献3）。まず、音楽信号から発音時刻を検出して発音時系列信号を生成し、この生成信号の自己相関を算出する。この自己相関のピーク位置とピーク値を検出する。このピーク位置とその値をグルーピング階層化することで音楽のビート構造を解析する。また自己相関のピーク値から算出したテンポ候補と解析したビート構造から、最も適切と思われるテンポを推定する。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2000-250534

【特許文献2】特開平05-027751

【特許文献3】特開2002-116754

【発明の概要】

【発明が解決しようとする課題】

【0007】

10

20

30

40

50

しかし、例えばピアノソロのように、楽曲中に打楽器が存在せず、また、発音時刻が揺らぐような楽曲では、小節間隔・拍間隔の正しい推定が難しい場合がある。

【0008】

具体的には、ピアノソロでは、特定の周波数帯にビートが集中することはないので、上述した第1の従来技術では、そのような楽曲に対して小節間隔・拍間隔を正しく推定することはできない。

【0009】

また、上述した第2または第3の従来技術のように発音時刻の周期性を利用する手法では、発音時刻が正確なビートを刻んでいるときには、自己相関がピーキーになるが、発音時刻に揺らぎがある場合は自己相関が低下するために小節間隔の検出精度が悪化する。

10

実際のピアノソロ曲（以下「楽曲1」と呼ぶ）を第2の従来技術の手法で自己相関信号を計算して得られる波形例を図17に示す。さらに、比較対象として、あるポップス曲（以下「楽曲2」と呼ぶ）を第2の従来技術の手法で自己相関信号を計算して得られる波形例を図18に示す。楽曲1は4分の3拍子、楽曲2は4分の4拍子である。

【0010】

図17または図18において、楽曲1または2の実際の拍間隔および小節間隔と、これらの楽曲から抽出された極大位置を図示している。ここで、極大位置とは、76BPM～140BPM程度の拍間隔において、小節間隔が存在する時間範囲、例えば、1.3秒～4秒程度の時間範囲内での、計算された自己相関値から抽出される最大値である。

【0011】

20

まず、図18の楽曲2についてみると、実際の小節位置と検出された極大位置が一致している。また、自己相関値のピークの時間間隔もほぼ等間隔で並んでおり、ビート構造もはっきりと現れている。

【0012】

一方、図17の楽曲1においては、小節間隔と極大位置の間にある小さなピークを加えて、検出された極大位置は5拍目に現れており、実際の小節間隔の3拍目とまったく関連がない。さらに、ビート構造らしきものも見当たらない。

【0013】

以上のように、従来技術では、小節間隔・拍間隔を正しく推定することができない場合があるという問題点を有していた。

30

【0014】

本発明は、ピアノソロなどのような楽曲でもビートの効いた楽曲でも、小節間隔や拍間隔の特徴がよく現れる特徴量抽出を行えるようにし、この特徴量を使って小節間隔の抽出精度を向上させるようにすることを目的とする。

【課題を解決するための手段】

【0015】

態様の一例の特徴量抽出装置は、入力する音響信号に対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出する時間周波数解析部と、単位時間毎に、複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出する最大振幅抽出部と、単位時間毎の最大振幅信号の時系列データの自己相関信号を導出し、その自己相関信号を特徴量として抽出する自己相関導出部とを備える。

40

【0016】

態様の他の一例の小節間隔推定装置は、特徴量抽出装置によって抽出された自己相関信号を特徴量として、その自己相関信号の所定位置から始まる時間範囲内で、第1の時間長を有する時間範囲である第1の検査範囲を移動させながら、第1の検査範囲で指定される各時間範囲内で最大値をとる自己相関信号の各ピークを抽出するピーク抽出部と、抽出された各ピークに基づいて、音響信号が表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第2の時間長を有する時間範囲である第2の検査範囲を移動させながら、第2の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを抽出

50

された各ピークから選択し、その選択したピークに基づいて強弱パターン周期を推定する強弱パターン周期推定部と、拍間隔を推定する拍間隔推定部と、強弱パターン周期と拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する小節間隔計算部とを備える。

【発明の効果】

【0017】

本発明によれば、ピアノソロなどのような楽曲でもビートの効いた楽曲でも、小節間隔や拍間隔の特徴がよく現れる特徴量抽出を行うことが可能となり、この特徴量を使って小節間隔の抽出精度を向上させることが可能となる。

【図面の簡単な説明】

10

【0018】

【図1】本発明による小節間隔推定装置の実施形態のハードウェア構成の一例を示す図である。

【図2】小節間隔推定装置の処理機能の構成例を示すブロック図である。

【図3】本実施形態に係る自己相関信号の波形例（楽曲1）を示す図である。

【図4】本実施形態に係る自己相関信号の波形例（楽曲2）を示す図である。

【図5】本実施形態に係る自己相関信号の波形例（楽曲3）を示す図である。

【図6】パワーの自己相関信号の波形例（楽曲1）を示す図である。

【図7】パワーの自己相関信号の波形例（楽曲2）を示す図である。

【図8】パワーの自己相関信号の波形例（楽曲3）を示す図である。

20

【図9】振幅の自己相関信号の波形例（楽曲1）を示す図である。

【図10】振幅の自己相関信号の波形例（楽曲2）を示す図である。

【図11】振幅の自己相関信号の波形例（楽曲3）を示す図である。

【図12】小節間隔推定装置の全体の処理例を示すフローチャートである。

【図13】特徴量抽出部の詳細な処理例を示すフローチャートである。

【図14】ピーク抽出部の詳細な処理例を示すフローチャートである。

【図15】強弱パターン周期推定部の詳細な処理例を示すフローチャートである。

【図16】小節間隔計算部の詳細な処理例を示すフローチャートである。

【図17】従来技術（特許文献2）による自己相関信号の波形例（楽曲1）を示す図である。

30

【図18】従来技術（特許文献2）による自己相関信号の波形例（楽曲2）を示す図である。

【発明を実施するための形態】

【0019】

以下、本発明を実施するための形態について図面を参照しながら詳細に説明する。

【0020】

図1は、本発明による小節間隔推定装置の実施形態のハードウェア構成の一例を示す図である。この小節間隔推定装置は、CPU（中央演算処理装置）101、ROM（リードオンリーメモリ）102、RAM（ランダムアクセスメモリ）103、入力部104、出力部105、音響入力部106、外部記憶装置107が、バス108によって相互に接続された構成を備える。

40

【0021】

音響入力部106は、音楽音響データを取り込み、必要に応じて、内部のRAM103やハードディスク記憶装置またはソリッドステートドライブ装置等の外部記憶装置107に記憶する。この音響入力部106は例えば、インターネット上の音楽配信サイトに接続して音楽音響データをダウンロードしてRAM103や外部記憶装置107に取り込むネットワーク接続装置を含む。また、音響入力部106は例えば、CD（コンパクトディスク）やDVD（デジタルバーサタイルディスク）等のディスク記録媒体に録音された音楽音響データを読み出し、必要に応じてRAM103や外部記憶装置107に記憶させるディスク記録媒体駆動装置を含む。あるいは、音響入力部106は例えば、SDメモリーカー

50

ド等の可搬記録媒体に録音された音楽音響データを読み出し、必要に応じてRAM 103や外部記憶装置107に記憶させる可変記録媒体駆動装置を含む。

【0022】

CPU 101は、ROM 102に記憶された制御プログラムに従って、当該小節間隔推定装置全体の制御を行う。RAM 103は、制御プログラムの実行に必要な各種制御データや、小節間隔が推定される対象となる音楽音響データを、一時的に記憶する。

【0023】

入力部104は、ユーザによるキーボード、マウス、各種スイッチ等による入力操作を検出し、その検出結果をCPU 101に通知する。

【0024】

出力部105は、CPU 101の制御によって送られてくるデータを液晶ディスプレイやプリンタに出力する。例えば、出力部105は、小節間隔の推定結果を液晶ディスプレイに表示する。

【0025】

外部記憶装置107は、音楽音響データや小節間隔の推定結果のデータ等を保存する。

【0026】

本実施形態による小節間隔推定装置の動作は、後述する図2の構成で示される処理機能を備え、それらの機能を実現する図12～図16に示される各フローチャートの処理を実行する制御プログラムを、CPU 101がROM 102から読み出して順次実行することで実現される。そのプログラムは、例えば外部記憶装置107に記録して配布してもよく、或いは特には図示しないネットワーク接続装置によりネットワークから取得できるようにしてもよい。

【0027】

図2は、図1のハードウェア構成例を有する小節間隔推定装置201の処理機能の構成例を示すブロック図である。本実施形態による小節間隔推定装置201は、特徴量抽出部202、ピーク抽出部206、強弱パターン周期推定部207、拍間隔推定部208、および小節間隔計算部209を備える。

【0028】

まず、特徴量抽出部202について、説明する。特徴量抽出部202は、小節間隔を推定するときの特徴量を、入力する音響信号のデジタルデータから抽出し、時間周波数解析部203、最大振幅抽出部204、および自己相関導出部205を備える。

【0029】

時間周波数解析部203は、例えば外部記憶装置107から入力する音響信号のデジタルデータに対して時間周波数解析を行って、離散時間（例えば分析フレーム）毎に複数の周波数成分毎の周波数成分信号を算出する。

【0030】

最大振幅抽出部204は、離散時間（例えば分析フレーム）毎に、複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出する。

【0031】

自己相関導出部205は、離散時間毎の最大振幅信号の時系列データの自己相関信号を導出し、その自己相関信号を特徴量として抽出する。

【0032】

以上のように、本実施形態では、特徴量抽出部202は、発音時刻のようなピーキー（時間変動が急峻）な信号の自己相関ではなく、もっと時間変動の緩やかな最大振幅信号の自己相関を特徴量とすることでロバスト性を高めている。

【0033】

具体的には、上述の最大振幅信号の時系列データを用いることにより、小節内の拍の強弱パターンを手掛かりに周期性が推定されるようになる。強弱パターンとしては4分の4拍子では、強・弱・中・弱やバックビートと呼ばれる弱・強・弱・中などがある。それ以外は1拍目に強がくることが多い。この強弱パターンの自己相関を求めれば、強弱パター

10

20

30

40

50

ン周期に対応した位置にピークが現れることが期待される。多くの場合、強弱パターン周期が小節間隔になるので、強弱パターン周期を示すピークに基づいて小節間隔を決定することが可能となる。強弱パターン周期が2小節間隔の楽曲の場合でも、ピーク位置が小節間隔の2倍の位置に検出されるが、このようなケースは後述する小節間隔計算部209の機能によって補正することが可能である。

【0034】

次に、強弱パターン周期の検出に用いるべき信号について説明する。音響信号の強度としてよく利用されるものに、パワーと振幅がある。音楽はダイナミックレンジが広いので、パワーの場合は強の影響が強くなりすぎ、弱の影響が弱くなりすぎるので適切とは言えない。よって、本実施形態では、振幅を強度信号とする。また、本実施形態では、振幅として、分析フレーム内のパワー合計を平方根したものでなく、例えば分析フレームを周波数解析して得られる複数の周波数成分信号のうち最大振幅を有するものを利用する。なぜなら、複数の楽器が演奏されるタイミングでは、それらのパワーが合算されるため強度信号としては適切でない。また、発音間隔が密な区間では前の音が減衰しきる前に新たな発音のパワーが次々加算されるので強弱パターンに規則性を乱す原因になるからでもある。本実施形態では、例えば分析フレーム毎に上述の最大振幅信号を抽出することにより、楽曲による強弱パターンが最も現れやすいような時系列データを生成することが可能となる。

【0035】

本実施形態においては、最大振幅の時系列データの自己相関を演算して自己相関信号を導出し、これを小節間隔・拍間隔情報に対応する強弱パターン周期を含む特徴量とする。

【0036】

図3、図4、および図5は、楽曲1、楽曲2、および楽曲3に対応して特徴量抽出部202が特徴量として抽出する各自己相関信号の波形例を示す図である。図3から図5において、楽曲1、2、または3の実際の強弱パターン周期、拍間隔、小節間隔と、これらの楽曲から後述するピーク抽出部206および強弱パターン周期推定部207により抽出される極大位置を図示している。

【0037】

図3～図5の例をみると、特徴量抽出部202が出力する最大振幅信号の時系列データの自己相関信号の特徴量を用いることにより、拍間隔は、時間ゼロ付近を除く時間ゼロから1秒程度の時間範囲内で最初に現れるピークとして容易に検出可能なことがわかる。図2の拍間隔推定部208は、このようなロジックに基づいて、拍間隔を推定する。

【0038】

また、図3または図4の例において、特徴量抽出部202が出力する最大振幅信号の時系列データの自己相関信号の特徴量を用いることにより、強弱パターン周期の存在が仮定される1.3秒～4秒程度の時間範囲内の極大位置が、強弱パターン周期に良く一致していることがわかる。

【0039】

本実施形態では、図2のピーク抽出部206が、自己相関信号の所定位置（例えば時刻ゼロ）から始まる時間範囲内で、比較的短い第1の時間長を有する時間範囲である第1の検査範囲を移動させながら、第1の検査範囲で指定される各時間範囲内で最大値をとる自己相関信号の各ピークを抽出する。これにより、本実施形態では、ノイズに起因するピークを取り除いたロバストなピーク検出を行うことが可能となる。

【0040】

さらに、図3～図5の例において、特徴量抽出部202が出力する最大振幅信号の時系列データの自己相関信号の波形は、時間が増加する方向に対して全体的に右肩下がりで、強弱パターン周期に対応するピーク毎にそのピークの周囲の波形を挟り取った形状になっていることがわかる。また、強弱パターン周期の自己相関値は、その2倍の時刻の自己相関値と比べて常に大きいことがわかる。そこで、本実施形態では、強弱パターン周期推定部207が、強弱パターン周期が存在すると仮定される例えば1.3～4秒程度の時間範

10

20

30

40

50

図内で、比較的長い第2の時間長を有する時間範囲である第2の検査範囲を移動させながら、第2の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを、ピーク抽出部206が抽出した各ピークから選択し、選択したピーク位置を強弱パターン周期として推定する。この結果、例えば図5の波形例において、「極大位置」と表示されているピークについては、そのピークの直前の第2の検査範囲内に大きなピークが存在し、かつその直前のピークは強弱パターン周期が仮定される例えば1.3～4秒程度の時間範囲よりも手前に存在する。従って、本実施形態では、この「極大位置」と表示されているピーク位置およびその直前のピーク位置ともに強弱パターン周期としては推定されないように動作させることができる。実際には「強弱パターン周期」と表示されているピーク位置が強弱パターン周期として推定される。このようにして、本実施形態では、特徴量抽出部202が出力する最大振幅信号の時系列データの自己相関信号の波形の特徴（右肩下がり）をうまく利用して、精度の高い強弱パターン周期の推定を行うことが可能となる。

10

【0041】

さらに、図5の波形例では、強弱パターン周期が、実際の小節間隔の2倍の位置に現れている。実際の楽曲でも、2小節で1つの強弱パターン周期が形成されるケースがある。そこで、本実施形態では、図2の小節間隔計算部209が、拍間隔推定部208が推定した拍間隔を用いて、強弱パターン周期推定部207が推定した強弱パターン周期の拍数（拍間隔の何倍か）を求める。そして、小節間隔計算部209は、この拍数が所定値に比較して小さければ、強弱パターン周期推定部207が推定した強弱パターン周期を小節間隔として算出する。一方、小節間隔計算部209は、上記拍数が所定値に比較して大きければ、強弱パターン周期推定部207が推定した強弱パターン周期の半分の周期を小節間隔として算出する。

20

【0042】

図6、図7、および図8はそれぞれ、図3、図4、および図5に対する比較として、楽曲1、楽曲2、および楽曲3に対応して、特徴量として各分析フレームのパワーの自己相関信号を算出したときの波形例を示す図である。また、図9、図10、および図11はそれぞれ、図3、図4、および図5または図6、図7、および図8に対する比較として、楽曲1、楽曲2、および楽曲3に対応して、特徴量として各分析フレームの振幅（パワー合計の平方根）の自己相関信号を算出したときの波形例を示す図である。楽曲1に対しては、図6および図9ともに、本実施形態に基づく図3の波形例と同様に、比較的正確に強弱パターン周期が抽出されている。しかし、楽曲2に対しては、図7のパワーの自己相関信号の波形例では強弱パターンの周期がなんとか分かる程度であり、図10の振幅の自己相関信号の波形例では強弱パターンは十分に分かるレベルである。但し、本実施形態に基づく図4の波形例ほど顕著ではない。楽曲3に対しては、図8および図11ともに、本実施形態に基づく図5の波形例に比較して、さらに強弱パターン周期が分かりにくい。また、全体的に、図6から図8のパワーの自己相関信号の波形例では、ディテールが損なわれている。図9から図11の振幅の自己相関信号の波形例では、図6から図8のパワーの自己相関信号の波形例に比べて少しよいが、図3から図5の本実施形態による波形例に比べれば、やはりディテールが損なわれている。

30

【0043】

以上のことより、本実施形態において特徴量抽出部202が抽出する最大振幅信号の自己相関信号は、拍間隔、強弱パターン周期、および小節間隔を推定するための特徴量として、優れているといえる。

40

【0044】

以下、図2の処理機能を実現する本実施形態の制御動作の詳細について、図12から図16の各フローチャートに沿って説明する。なお、各フローチャートの説明において、随時図1または図2の各部を参照するものとする。

【0045】

図12は、図1のハードウェア構成例および図2の処理機能構成例を有する小節間隔推定装置の全体の処理例を示すフローチャートである。このフローチャートは、CPU10

50

1 が、ROM 102 に記憶された小節間隔推定プログラムを実行する動作として実現される。

【0046】

図2の特徴量抽出部202、ピーク抽出部206、強弱パターン周期推定部207、拍間隔推定部208、および小節間隔計算部209の機能をそれぞれ実現する処理として、特徴量抽出処理(ステップS101)、ピーク抽出処理(ステップS102)、強弱パターン周期推定処理(ステップS103)、拍間隔推定処理(ステップS104)、および小節間隔計算処理(ステップS105)の順に、各制御処理が実行される。

【0047】

図13は、図12のステップS101の特徴量抽出処理の詳細な処理例を示すフローチャートである。

10

【0048】

まず、例えば外部記憶装置107から入力する音響信号のデジタルデータに対して時間周波数解析処理が実行されて、離散時間(例えば分析フレーム)毎に複数の周波数成分毎の周波数成分信号が算出される(ステップS201)。この処理は、図2の時間周波数解析部203の機能を実現する。時間周波数解析処理としては、短時間フーリエ変換、ウェーブレット変換、定Qフィルタバンクなどを利用することができる。これにより、離散時間毎の各周波数チャネルの複素信号が得られる。

【0049】

次に、最大振幅抽出処理が実行される(ステップS202)。この処理では、離散時間(例えば分析フレーム)毎に、ステップS201の時間周波数解析処理により得られた複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号が最大振幅信号として抽出される。この処理は、図2の最大振幅抽出部204の機能を実現する。このとき、各周波数成分の振幅を先に計算してから最大値を求めてもよいが、パワーの最大値を求め、これの平方根を計算した方が計算負荷的に有利である。

20

【0050】

最後に、自己相関演算処理が実行される(ステップS203)。この処理では、ステップS202で抽出された最大振幅信号の時系列データの自己相関が演算される。自己相関は例えば、最大振幅信号の時系列データのパワースペクトルの逆フーリエ変換にて求めることができる。

30

【0051】

図14は、図12のステップS102のピーク抽出処理の詳細な処理例を示すフローチャートである。ここでは、図12のステップS101の特徴量抽出処理で抽出された特徴量から、ピーク抽出が行われる。拍間隔、小節間隔、強弱パターン周期は、このピークに含まれる。

【0052】

まず、例えばRAM103上の変数である位置(図3から図5の横軸のうち、以下のフローのループで現在処理対象とする横軸位置を示すカウンタ)が、開始位置に設定され、ピークコンテナがクリア(空)される(ステップS301)。ピークコンテナは、抽出したピーク(位置)を格納するための、例えばRAM103上の配列変数である。

40

【0053】

次に、位置に対応する検査範囲(第1の検査範囲:図3から図5の横軸で、現在処理対象となっている位置に対して検査する横軸範囲)が決定される(ステップS302)。単純に隣接間の自己相関信号値の比較でピーク判定をするとノイズを拾うことがあるので、ノイズを拾いにくい時間範囲を検査するようにする。例えば、図5に例示した楽曲3の波形例では、拍間隔の前後にノイズによるピークが見られ、自己相関はゼロ付近で急激に大きくなるのでこのようなノイズを拾いやすくなるが、こういったノイズは小さな出っ張りであることが多いので、第1の検査範囲を適切に設定すれば簡単に除去できる。また、位置によって検査範囲は変わるので、このように現在処理対象となっている位置毎に第1の検査範囲を決定するようにしている。

50

【 0 0 5 4 】

そして、ステップ S 3 0 2 で決定された第 1 の検査範囲内における自己相関値の最大値が取得される（ステップ S 3 0 3）。

【 0 0 5 5 】

続いて、現在の処理対象位置における自己相関の値とステップ S 3 0 3 で取得された最大値が等しいか否かが判定される（ステップ S 3 0 4）。

【 0 0 5 6 】

現在の処理対象位置における自己相関の値とステップ S 3 0 3 で取得された自己相関値の最大値が等しくステップ S 3 0 4 の判定が Y e s ならば、現在の処理対象位置における自己相関の値はピークなので、位置をピークコンテナに追加する（ステップ S 3 0 5）。 10

【 0 0 5 7 】

一方、現在の処理対象位置における自己相関の値と、ステップ S 3 0 3 で取得された自己相関値の最大値が等しくなくステップ S 3 0 4 の判定が N o ならば、現在の処理対象位置における自己相関の値はピークでないので何もしない。

【 0 0 5 8 】

この位置での処理が終了したので、位置が最終位置であるか否かが判定される（ステップ S 3 0 6）。

【 0 0 5 9 】

最終位置でなくステップ S 3 0 6 の判定が N o ならば、まだ処理すべき位置があるので、位置が進められ（ステップ S 3 0 7）、ステップ S 3 0 2 の処理に戻って処理が継続される。 20

【 0 0 6 0 】

そして、全ての位置の処理が完了すると、ステップ S 3 0 6 の判定で位置が最終位置と等しくなるとステップ S 3 0 6 の判定が Y e s となり、図 1 4 のフローチャートの処理が終了して、図 1 2 のステップ S 1 0 2 のピーク抽出処理が終了する。

【 0 0 6 1 】

図 1 5 は、図 1 2 のステップ S 1 0 3 の強弱パターン周期推定処理の詳細な処理例を示すフローチャートである。

【 0 0 6 2 】

まず、図 1 2 のステップ S 1 0 2 のピーク抽出処理で抽出されたピーク群から、検査する範囲内のピークが取得される（ステップ S 4 0 1）。推定すべき強弱パターン周期は、ピーク抽出処理で抽出されたピーク位置のどれかに該当するので、強弱パターン周期が存在すると仮定される時間範囲内にあるピークを予め取得しておく。仮定される範囲としては、例えば 1 小節間隔が 1 . 3 ~ 4 秒程度の時間範囲内と仮定すると、強弱パターン周期は 1 または 2 小節間隔なので、1 . 3 ~ 8 秒程度の時間範囲となる。この時間範囲は、こういったテンポのこういった拍子の楽曲を想定するか等に応じて、自由に定める事が可能である。 30

【 0 0 6 3 】

次に、検査するピークが、ステップ S 4 0 1 で取得された最初のピークに設定される（ステップ S 4 0 2）。 40

【 0 0 6 4 】

そして、検査範囲（第 2 の検査範囲）が決定される（ステップ S 4 0 3）。前述したように自己相関信号の波形は全体的に右肩下がり、強弱パターン周期に対応するピーク毎にそのピークの周囲の波形を挟り取った形状になっている。また、強弱パターン周期の自己相関値はその 2 倍の時刻の自己相関値と比べて常に大きいので、適切な第 2 の検査範囲内で最大値になる。例えば、第 2 の検査範囲は、図 3 ~ 図 5 の図で、検査するピーク位置に相当する横軸の位置の半分より若干小さい横軸の値となる位置から、検査するピーク位置に相当する横軸の位置の 2 倍より若干大きい横軸の値となる位置に設定される。

【 0 0 6 5 】

次に、ステップ S 4 0 3 で決定された検査範囲内における自己相関の最大値が取得され 50

る (S 4 0 4)。

【 0 0 6 6 】

続いて、検査するピーク位置における自己相関の値と、ステップ S 4 0 4 で取得された自己相関の最大値が等しいかが判定される (ステップ S 4 0 5)。

【 0 0 6 7 】

検査するピーク位置における自己相関の値と、ステップ S 4 0 4 で取得された自己相関の最大値が等しくステップ S 4 0 5 の判定が Y e s ならば、検査するピーク位置は強弱パターン周期としての条件を満たすので、強弱パターン周期にこのピーク位置の横軸位置に基づく周期が設定される (ステップ S 4 0 6)。そして、図 1 5 のフローチャートの処理が終了して、図 1 2 のステップ S 1 0 3 の強弱パターン周期推定処理が終了する。

10

【 0 0 6 8 】

一方、検査するピーク位置における自己相関の値と、ステップ S 4 0 4 で取得された自己相関の最大値が等しくなくステップ S 4 0 5 の判定が N o ならば、このピーク位置は強弱パターン周期ではないので、処理が続行され、全ピークの検査が終了したか否かが判定される (ステップ S 4 0 7)。

【 0 0 6 9 】

全ピークの検査が終了しておらずステップ S 4 0 7 の判定が N o ならば、検査するピーク位置が次のピークに設定され (ステップ S 4 0 8)、ステップ S 4 0 3 の処理に戻って、検査処理が続行される。

【 0 0 7 0 】

20

そして、全ピークの検査が終了しステップ S 4 0 7 の判定が Y e s になると、エラー処理が実行される (ステップ S 4 0 9)。そして、図 1 5 のフローチャートの処理が終了して、図 1 2 のステップ S 1 0 3 の強弱パターン周期推定処理が終了する。強弱パターン周期が見つからなかったという状態は通常あり得ないので、エラーとなる。

【 0 0 7 1 】

以上のようにして、図 1 2 のステップ S 1 0 3 の強弱パターン周期推定処理が実現される。

【 0 0 7 2 】

次に、図 1 2 のステップ S 1 0 4 では、拍間隔が推定される。推定すべき拍間隔は、ステップ S 1 0 2 のピーク抽出処理で抽出されたピーク位置のどれかに該当するので、拍間隔が存在すると仮定する時間範囲内にあるピーク位置のうちで最大の自己相関信号値を持つものが、拍間隔と推定される。例えば、拍間隔が多くの楽曲が入るであろう 7 5 ~ 1 5 0 B P M の範囲に入るとすれば 0 . 4 ~ 0 . 8 秒程度の時間範囲、もう少し広めに取って拍間隔が 6 0 ~ 1 8 0 B P M の範囲に入るとすればならば 0 . 3 3 ~ 1 . 0 秒程度の時間範囲が設定される。予め楽曲のジャンルとそのジャンルでの一般的な B P M の範囲が判っているならば、それが適用されてもよい。

30

【 0 0 7 3 】

図 1 6 は、図 1 2 のステップ S 1 0 5 の小節間隔計算処理の詳細な処理例を示すフローチャートである。ここでは、図 1 2 のステップ S 1 0 3 の強弱パターン周期推定処理で推定された強弱パターン周期とステップ S 1 0 4 の拍間隔推定処理で推定された拍間隔を使って、小節間隔が計算される。

40

【 0 0 7 4 】

まず、強弱パターン周期の拍数が計算される (S 5 0 1)。具体的には、強弱パターン周期が拍間隔で除算されて、その除算結果が整数値に丸められる。

【 0 0 7 5 】

次に、ステップ S 5 0 1 で計算された拍数が所定値以上であるか否かが判定される (S 5 0 2)。多くの楽曲で 4 分の 6 拍子以上が使われることはまずないので、例えば所定値は 6 に設定される。もちろん、楽曲のジャンルが予め判ってれば、それに合わせた所定値が設定されてもよい。

【 0 0 7 6 】

50

計算された拍数が所定値未満でステップ S 5 0 2 の判定が N o ならば、強弱パターン周期は 1 小節間隔と推定できるので、小節間隔に強弱パターン周期が設定される（ステップ S 5 0 3）。そして、図 1 6 のフローチャートの処理が終了して、図 1 2 のステップ S 1 0 5 の小節間隔推定処理が終了する。

【 0 0 7 7 】

一方、計算された拍数が所定値以上でステップ S 5 0 2 の判定が Y e s ならば、強弱パターン周期は 2 小節間隔と推定できるので、強弱パターン周期を 0 . 5 倍したものが小節間隔に設定される（ステップ S 5 0 4）。なお、強弱パターン周期を 0 . 5 倍したものがさらに上記の所定値 6 以上である場合には、強弱パターン周期を 3 や 4 等で除した値を用いることとしても良い。特に、ある種の音楽の場合、4 小節でひとまとまりとなる場合も多いので、4 で除した値を用いることとしても良い。そして、図 1 6 のフローチャートの処理が終了して、図 1 2 のステップ S 1 0 5 の小節間隔推定処理が終了する。

10

【 0 0 7 8 】

以上のようにして、本実施形態では、時間周波数解析から得られる最大振幅信号の自己相関を特徴量として用いることにより、図 3 から図 5 より明らかなように、ピアノソロなどのような楽曲でもビートの効いた楽曲でも、小節間隔と関連の高い強弱パターン周期や拍間隔の特徴を特徴量中に良く保持することが可能となる。

【 0 0 7 9 】

これは、発音時刻のようなピーキー（時間変動が急峻）な信号の自己相関ではなく、もっと時間変動の緩やかな最大振幅信号の自己相関を特徴量とすることでロバスト性を高めたことに起因する。

20

【 0 0 8 0 】

また、本実施形態における強弱パターン周期の抽出方法は、従来からある極大値を見つけるものとは異なり、正確に強弱パターン周期を見つけることが可能となる。

【 0 0 8 1 】

これにより、本実施例では、小節間隔の推定精度を向上させることが可能となる。

【 0 0 8 2 】

以上の実施形態に関して、更に以下の付記を開示する。

（付記 1）

入力する音響信号に対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出する時間周波数解析部と、

30

前記単位時間毎に、前記複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出する最大振幅抽出部と、

前記単位時間毎の最大振幅信号の時系列データの自己相関信号を導出し、当該自己相関信号を特徴量として抽出する自己相関導出部と、

を備えることを特徴とする特徴量抽出装置。

（付記 2）

付記 1 に記載の特徴量抽出装置によって抽出された自己相関信号を特徴量として、前記音響信号の小節間隔を推定する装置であって、

前記自己相関信号の所定位置から始まる時間範囲内で、第 1 の時間長を有する時間範囲である第 1 の検査範囲を移動させながら、当該第 1 の検査範囲で指定される各時間範囲内で最大値をとる前記自己相関信号の各ピークを抽出するピーク抽出部と、

40

前記音響信号が表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第 2 の時間長を有する時間範囲である第 2 の検査範囲を移動させながら、当該第 2 の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを前記抽出された各ピークから選択し、該選択したピークに基づいて前記強弱パターン周期を推定する強弱パターン周期推定部と、

前記抽出された各ピークに基づいて、拍間隔を推定する拍間隔推定部と、

前記強弱パターン周期と前記拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する小節間隔計算部と、

50

を備えることを特徴とする小節間隔推定装置。

(付記 3)

前記小節間隔計算部は、前記拍間隔に基づいて前記強弱パターン周期の拍数を算出し、当該拍数が所定値に比較して小さければ前記強弱パターン周期を前記小節間隔として算出し、当該拍数が所定値に比較して大きければ前記強弱パターン周期の半分の周期を前記小節間隔として算出する、

ことを特徴とする付記 2 に記載の小節間隔推定装置。

(付記 4)

入力する音響信号に対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出し、

前記単位時間毎に、前記複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出し、

前記単位時間毎の最大振幅信号の時系列データの自己相関信号を導出し、当該自己相関信号を特徴量として抽出する、

ことを特徴とする特徴量抽出方法。

(付記 5)

付記 4 に記載の特徴量抽出方法によって抽出された自己相関信号を特徴量として、前記音響信号の小節間隔を推定する方法であって、

前記自己相関信号の所定位置から始まる時間範囲内で、第 1 の時間長を有する時間範囲である第 1 の検査範囲を移動させながら、当該第 1 の検査範囲で指定される各時間範囲内で最大値をとる前記自己相関信号の各ピークを抽出し、

前記抽出された各ピークに基づいて、拍間隔を推定し、

前記音響信号が表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第 2 の時間長を有する時間範囲である第 2 の検査範囲を移動させながら、当該第 2 の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを前記抽出された各ピークから選択し、該選択したピークに基づいて前記強弱パターン周期を推定し、

前記強弱パターン周期と前記拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する、

ことを特徴とする小節間隔推定方法。

(付記 6)

コンピュータに、

入力する音響信号に対して時間周波数解析を行って、単位時間毎に複数の周波数成分毎の周波数成分信号を算出し、

前記単位時間毎に、前記複数の周波数成分毎の周波数成分信号のうち、振幅が最大の周波数成分信号を最大振幅信号として抽出し、

前記単位時間毎の最大振幅信号の時系列データの自己相関信号を導出し、当該自己相関信号を特徴量として抽出する、

処理を実行させることを特徴とする特徴量抽出プログラム。

(付記 7)

付記 6 に記載の特徴量抽出プログラムによって抽出された自己相関信号を特徴量として、前記音響信号の小節間隔を推定するプログラムであって、

コンピュータに、

前記自己相関信号の所定位置から始まる時間範囲内で、第 1 の時間長を有する時間範囲である第 1 の検査範囲を移動させながら、当該第 1 の検査範囲で指定される各時間範囲内で最大値をとる前記自己相関信号の各ピークを抽出し、

前記抽出された各ピークに基づいて、拍間隔を推定し、

前記音響信号が表す楽曲に対応する強弱パターン周期が存在すると仮定される時間範囲内で、第 2 の時間長を有する時間範囲である第 2 の検査範囲を移動させながら、当該第 2 の検査範囲で指定される各時間範囲内で最初に最大値をとるピークを前記抽出された各ピークから選択し、該選択したピークに基づいて前記強弱パターン周期を推定し、

10

20

30

40

50

前記強弱パターン周期と前記拍間隔とに基づいて小節間隔を計算して推定された小節間隔として出力する、

処理を実行させることを特徴とする小節間隔推定プログラム。

【符号の説明】

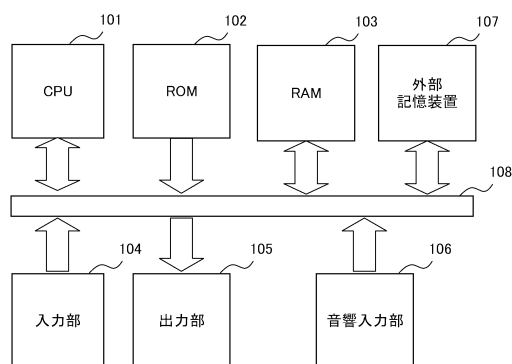
【 0 0 8 3 】

- 1 0 1 C P U
- 1 0 2 R O M
- 1 0 3 R A M
- 1 0 4 入力部
- 1 0 5 出力部
- 1 0 6 音響入力部
- 1 0 7 外部記憶装置
- 1 0 8 バス
- 2 0 1 小節間隔推定装置
- 2 0 2 特徴量抽出部
- 2 0 3 時間周波数解析部
- 2 0 4 最大振幅抽出部
- 2 0 5 自己相関演算部
- 2 0 6 ピーク抽出部
- 2 0 7 強弱パターン周期推定部
- 2 0 8 拍間隔推定部
- 2 0 9 小節間隔推定部

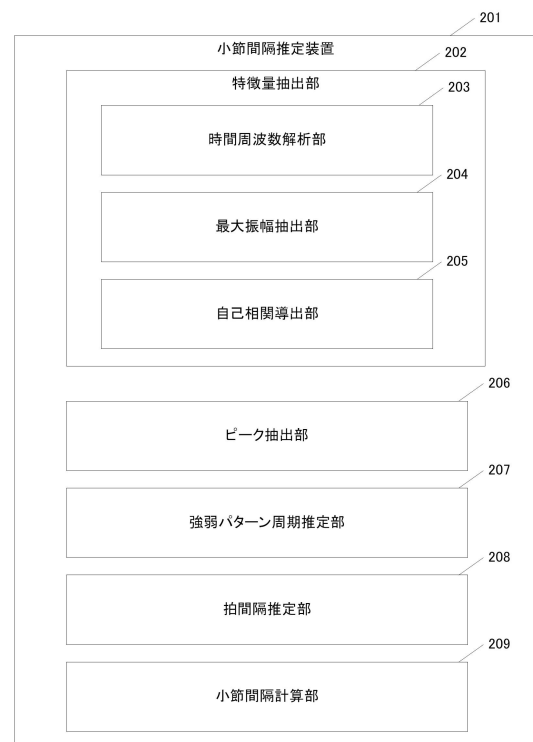
10

20

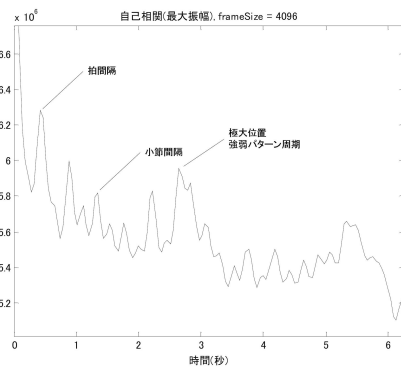
【図 1】



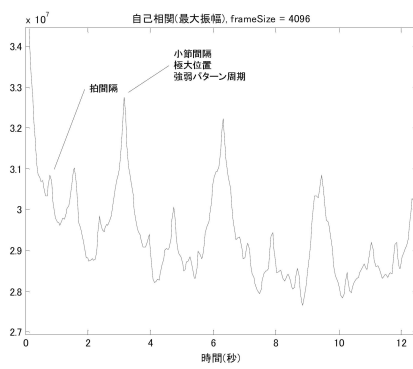
【図 2】



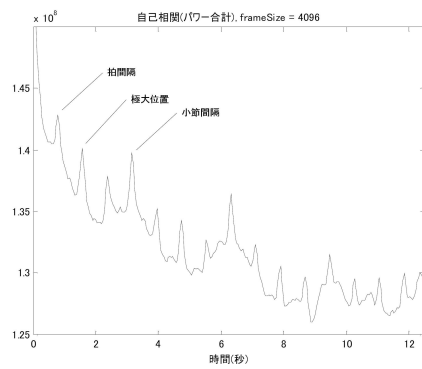
【図 3】



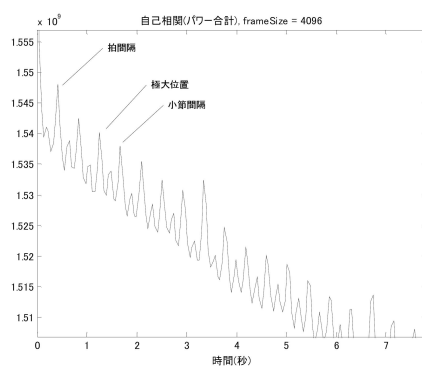
【図 4】



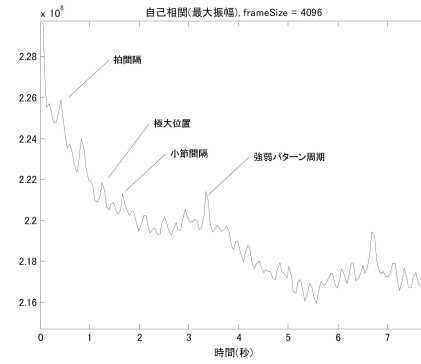
【図 7】



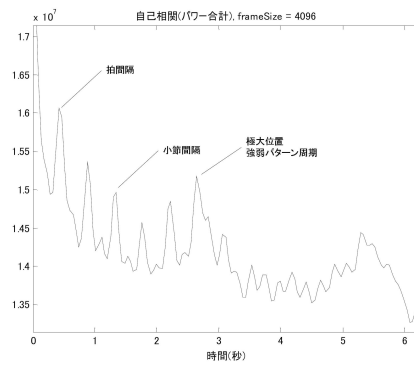
【図 8】



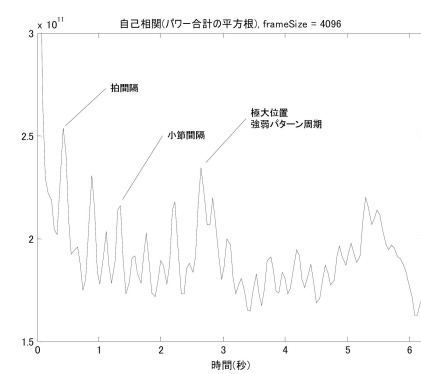
【図 5】



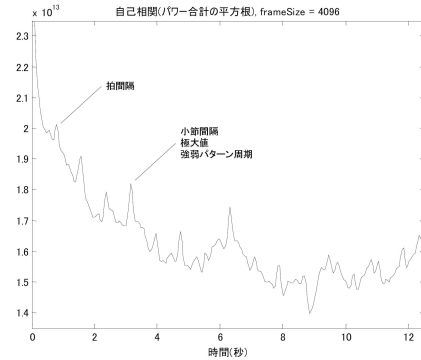
【図 6】



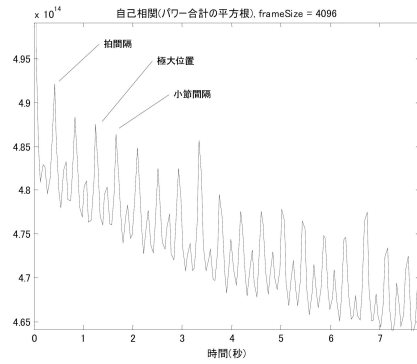
【図 9】



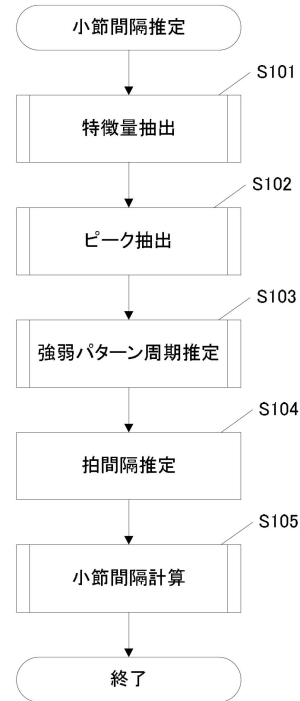
【図 10】



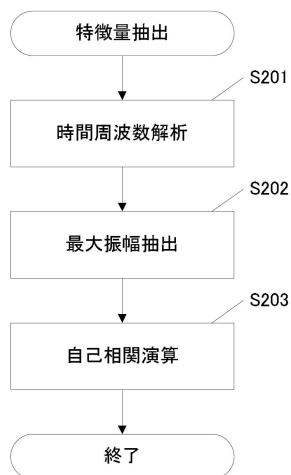
【図 1 1】



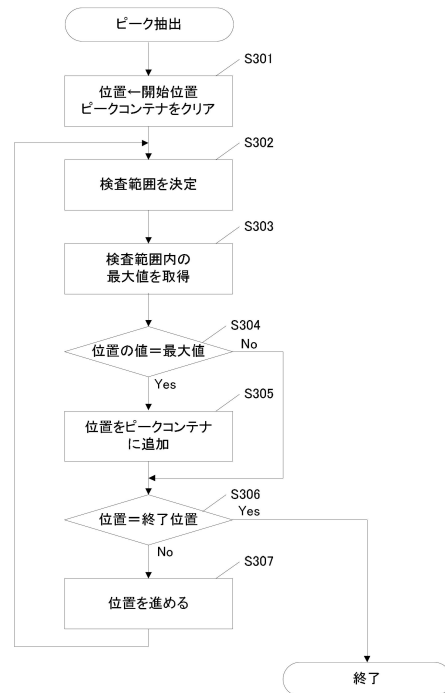
【図 1 2】



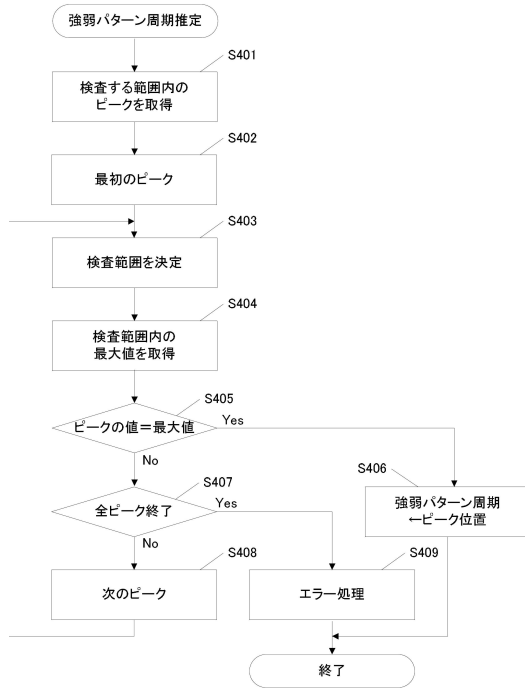
【図 1 3】



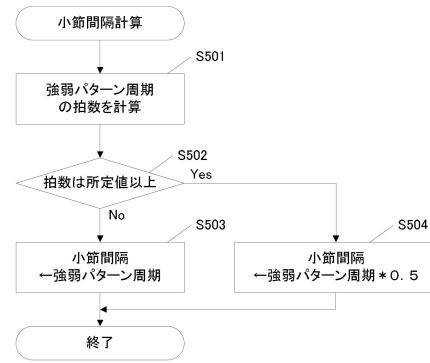
【図 1 4】



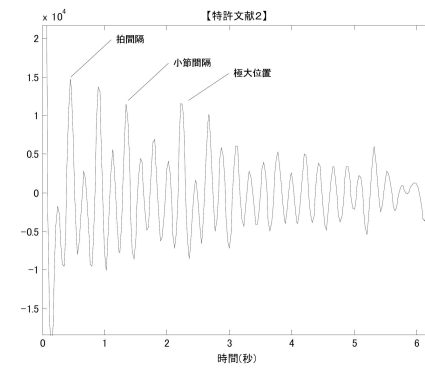
【図 15】



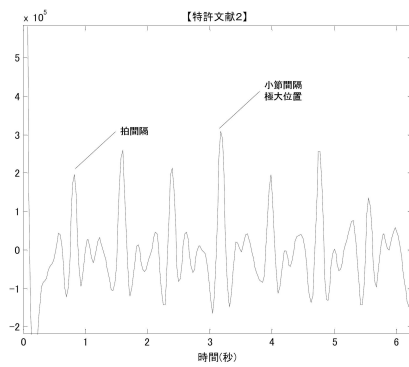
【図 16】



【図 17】



【図 18】



フロントページの続き

(58)調査した分野(Int.Cl. , D B 名)

G 1 0 G 3 / 0 4

G 1 0 L 2 5 / 5 1