(19)日本国特許庁(JP)　(12)公 開 特 許 公 報(A)　(11)特許出願公開番号
特開2010-60836
(P2010-60836A)
(43)公開日 平成22年3月18日(2010.3.18)

(51)Int.Cl.		F I			テーマコード（参考）
G10L 19/00	(2006.01)	G10L	19/00	330B	5D082
G10G 3/04	(2006.01)	G10G	3/04		

審査請求 有　請求項の数 13　ＯＬ　（全 23 頁）

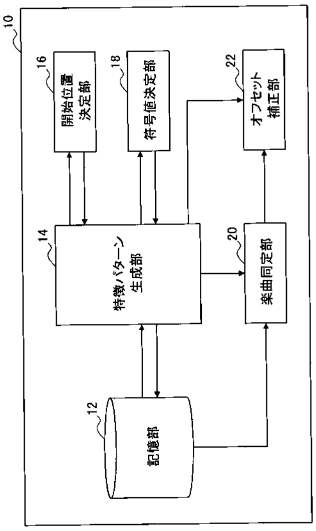
(21)出願番号　特願2008-226344 (P2008-226344)
(22)出願日　平成20年9月3日 (2008.9.3)

(71)出願人　000002185
ソニー株式会社
東京都港区港南１丁目７番１号
(74)代理人　100095957
弁理士　亀谷　美明
(74)代理人　100096389
弁理士　金本　哲男
(74)代理人　100101557
弁理士　萩原　康司
(74)代理人　100128587
弁理士　松本　一騎
(72)発明者　宮島　靖
東京都港区港南１丁目７番１号　ソニー株式会社内

Ｆターム(参考) 5D082 BB01 BB13

(54)【発明の名称】楽曲処理方法、楽曲処理装置、及びプログラム

(57)【要約】
【課題】オフセットのずれに影響されず楽曲データ又は楽曲データのメタデータを使用すること。
【解決手段】楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定するステップと、複数の区間にわたって決定された一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成するステップと、を含む楽曲処理方法を提供する。前記符号値は、例えば、ある区間における平均音量が前の区間における平均音量よりも増加しているか又は減少しているかを表していてもよい。
【選択図】図３



【特許請求の範囲】
【請求項１】
楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定するステップと；
複数の区間にわたって決定された一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成するステップと；
を含む楽曲処理方法。
【請求項２】
前記符号値は、ある区間における平均音量が前の区間における平均音量よりも増加しているか又は減少しているかを表す、請求項１に記載の楽曲処理方法。
【請求項３】
さらに、前記楽曲データのうち時間軸上の所定のウィンドウ幅における平均音量が当該楽曲データの全体又は一部分の平均音量を初めて上回るウィンドウ位置を求めることにより前記符号化開始位置を決定するステップを含む、請求項１に記載の楽曲処理方法。
【請求項４】
さらに、前記楽曲データについて生成された前記特徴パターンと任意の楽曲データの特徴パターンとを比較することにより当該楽曲データ間の特徴パターンの類似度を計算するステップを含む、請求項１に記載の楽曲処理方法。
【請求項５】
前記類似度は、２つの特徴パターンの間で符号値の一致するビット数の全ビット数に対する割合として計算される、請求項４に記載の楽曲処理方法。
【請求項６】
前記類似度は、２つの特徴パターンの間で符号値の一致するビットと符号値の一致しないビットの二項分布における発生確率を用いて計算される、請求項４に記載の楽曲処理方法。
【請求項７】
前記類似度は、２つの特徴パターンの間で符号値が連続して一致する一致部分の最大長を用いて計算される、請求項４に記載の楽曲処理方法。
【請求項８】
さらに、類似度の大きい２つの楽曲データの特徴パターンに対応する前記符号化開始位置を用いて、いずれかの楽曲データ又は当該楽曲データに関連するメタデータのオフセットを補正するステップを含む、請求項４に記載の楽曲処理方法。
【請求項９】
前記符号化開始位置を決定するステップにおいて、異なるウィンドウ幅を用いて複数の前記符号化開始位置を決定し、
前記特徴パターンを生成するステップにおいて、複数の前記符号化開始位置に基づいて、１つの楽曲データに対し複数の前記特徴パターンを生成する、
請求項３に記載の楽曲処理方法。
【請求項１０】
前記楽曲データについて生成された各特徴パターンと任意の楽曲データの特徴パターンとを比較することにより、特徴パターンごとに類似度を計算するステップと；
複数の前記特徴パターンのうち、計算された前記類似度が最も大きい特徴パターンに対応する前記符号化開始位置を用いて、いずれかの楽曲データ又は当該楽曲データに関連するメタデータのオフセットを補正するステップと；
を含む、請求項９に記載の楽曲処理方法。
【請求項１１】
楽曲データのうち時間軸上の所定のウィンドウ幅における平均音量が当該楽曲データの全体又は一部分の平均音量を初めて上回るウィンドウ位置を求めるステップと；
前記楽曲データの前記ウィンドウ位置に基づいて、当該楽曲データ又は当該楽曲データに関連するメタデータのオフセットを補正するステップと；

を含む楽曲処理方法。

【請求項１２】

楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定する符号値決定部と；

前記符号値決定部により決定された複数の区間にわたる一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成する特徴パターン生成部と；

を備える楽曲処理装置。

【請求項１３】

情報処理装置を制御するコンピュータを：

楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定する符号値決定部と；

前記符号値決定部により決定された複数の区間にわたる一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成する特徴パターン生成部と；

として機能させるための、プログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、楽曲処理方法、楽曲処理装置、及びプログラムに関する。

【背景技術】

【０００２】

近年、情報処理装置の性能向上に伴い、デジタル化された楽曲データの新たな利用形態が見られるようになってきた。新たな利用形態の１つは、リミックスである。リミックスとは、一般的には、複数の楽曲の全部又は一部を組み合わせて、全体として新たな楽曲を構成する手法を指す。リミックスそのものは旧来より行われてきた手法だが、専用の装置を持たない一般のユーザもＰＣ（Personal Computer）を用いて簡単にリミックスを行うことができるようになったことから、より多くのユーザがリミックスに参加し、リミックスされたデータの交換も行われ始めている。また、写真やスライドの上映と連動してシーンに応じて楽曲を再生させるスライドショーなども、楽曲データの新たな利用形態の一例である。スライドショーは、イベントでのプレゼンテーションや商品のプロモーションなどに広く利用されている。

【０００３】

楽曲データをリミックスやスライドショーなどに使用する場合、楽曲データに付加されたメタデータを活用するのが効果的である。メタデータには、曲名やアーティスト名などの静的なデータのみならず、楽曲データを時系列に解析して得られる、例えばリズム、ビート位置、小節の長さや位置、コード進行、又はメロディの種類などのデータも含まれる。そして、このような時系列のメタデータを用いることにより、楽曲データからの特定の小節の抽出、リズムやビート位置の調整などを簡単に行うことができる。

【０００４】

さらに、単純に楽曲を再生して楽しむ一般のユーザにとっても、メタデータを活用するメリットは存在する。例えば、ハードディスクやフラッシュメモリなどの大容量記録媒体に格納された大量の楽曲データの中から所望の楽曲データを素早く見つけ出すためには、メタデータにより表される楽曲の特徴を用いて検索する手法が有効である。

【０００５】

こうした状況から、楽曲データに関するメタデータの生成又は利用を支援するための技術開発が進められている。例えば、下記特許文献１には、楽曲データ内のビート位置や小節の頭の位置を自動的に抽出し、さらにユーザにより容易に補正可能としたメタデータ付与装置が開示されている。また、下記特許文献２には、楽曲に付与されたビート位置などのメタデータを用いて、楽曲データを再生するタイミングを適宜調整することのできる音楽編集装置が開示されている。

【０００６】

【特許文献１】特開２００７－２４８８９５号公報
【特許文献２】特開２００８－１６４９３２号公報
【発明の開示】
【発明が解決しようとする課題】
【０００７】
しかしながら、異なる環境において媒体に記録された楽曲データには、データ読取装置の違いやエンコードパラメータの違いにより、同一の楽曲であってもデータの開始点から演奏の開始点までのオフセットにずれが生じる場合がある。オフセットのずれは、例えば楽曲データとは別に時系列のメタデータを配布した場合には、メタデータが楽曲データと合致しない原因となる。そして、メタデータが楽曲データに合致しなければ、メタデータの使用を前提とした、リミックスやスライドショー又は楽曲データの検索などは、ユーザにより期待される結果を提供できない。
【０００８】
そこで、本発明は、上記問題に鑑みてなされたものであり、本発明の目的とするところは、オフセットのずれに影響されず楽曲データ又は楽曲データのメタデータを使用することのできる、新規かつ改良された楽曲処理方法、楽曲処理装置、及びプログラムを提供することにある。
【課題を解決するための手段】
【０００９】
上記課題を解決するために、本発明のある観点によれば、楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定するステップと、複数の区間にわたって決定された一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成するステップと、を含む楽曲処理方法が提供される。
【００１０】
前記符号値は、ある区間における平均音量が前の区間における平均音量よりも増加しているか又は減少しているかを表していてもよい。
【００１１】
前記楽曲処理方法は、さらに、前記楽曲データのうち時間軸上の所定のウィンドウ幅における平均音量が当該楽曲データの全体又は一部分の平均音量を初めて上回るウィンドウ位置を求めることにより前記符号化開始位置を決定するステップを含んでもよい。
【００１２】
前記楽曲処理方法は、さらに、前記楽曲データについて生成された前記特徴パターンと任意の楽曲データの特徴パターンとを比較することにより当該楽曲データ間の特徴パターンの類似度を計算するステップを含んでもよい。
【００１３】
また、前記類似度は、２つの特徴パターンの間で符号値の一致するビット数の全ビット数に対する割合として計算されてもよい。
【００１４】
また、前記類似度は、２つの特徴パターンの間で符号値の一致するビットと符号値の一致しないビットの二項分布における発生確率を用いて計算されてもよい。
【００１５】
また、前記類似度は、２つの特徴パターンの間で符号値が連続して一致する一致部分の最大長を用いて計算されてもよい。
【００１６】
前記楽曲処理方法は、さらに、類似度の大きい２つの楽曲データの特徴パターンに対応する前記符号化開始位置を用いて、いずれかの楽曲データ又は当該楽曲データに関連するメタデータのオフセットを補正するステップを含んでもよい。
【００１７】
また、前記符号化開始位置を決定するステップにおいて、異なるウィンドウ幅を用いて

複数の前記符号化開始位置が決定され、前記特徴パターンを生成するステップにおいて、複数の前記符号化開始位置に基づいて、１つの楽曲データに対し複数の前記特徴パターンが生成されてもよい。

【００１８】

前記楽曲処理方法は、さらに、前記楽曲データについて生成された各特徴パターンと任意の楽曲データの特徴パターンとを比較することにより、特徴パターンごとに類似度を計算するステップと、複数の前記特徴パターンのうち、計算された前記類似度が最も大きい特徴パターンに対応する前記符号化開始位置を用いて、いずれかの楽曲データ又は当該楽曲データに関連するメタデータのオフセットを補正するステップと、を含んでもよい。

【００１９】

上記課題を解決するために、本発明の別の観点によれば、楽曲データのうち時間軸上の所定のウィンドウ幅における平均音量が当該楽曲データの全体又は一部分の平均音量を初めて上回るウィンドウ位置を求めるステップと、前記楽曲データの前記ウィンドウ位置に基づいて、当該楽曲データ又は当該楽曲データに関連するメタデータのオフセットを補正するステップと、を含む楽曲処理方法が提供される。

【００２０】

上記課題を解決するために、本発明の別の観点によれば、楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定する符号値決定部と、前記符号値決定部により決定された複数の区間にわたる一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成する特徴パターン生成部と、を備える楽曲処理装置が提供される。

【００２１】

上記課題を解決するために、本発明の別の観点によれば、情報処理装置を制御するコンピュータを、楽曲データ内の時間軸上の所定の符号化開始位置から一定の区間ごとに、音量の変化に応じて符合値を決定する符号値決定部と、前記符号値決定部により決定された複数の区間にわたる一連の前記符号値を用いて、前記楽曲データの音量変化の特徴を表す特徴パターンを生成する特徴パターン生成部と、として機能させるためのプログラムが提供される。

【発明の効果】

【００２２】

以上説明したように、本発明に係る楽曲処理方法、楽曲処理装置、及びプログラムによれば、オフセットのずれに影響されず楽曲データ又は楽曲データのメタデータを使用することができる。

【発明を実施するための最良の形態】

【００２３】

以下に添付図面を参照しながら、本発明の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【００２４】

また、以下の順序にしたがって当該「発明を実施するための最良の形態」を説明する。

１．時系列のメタデータに関する課題

２．第１の実施形態

３．第２の実施形態

【００２５】

＜１．時系列のメタデータに関する課題＞

まず、図１を参照しながら、楽曲データに付与される時系列のメタデータについて説明する。図１は、楽曲データ、メタデータ、及びリミックスデータの関係を示す説明図である。

【００２６】

図１を参照すると、楽曲Ｇａの楽曲データＤａ、及び楽曲Ｇｂの楽曲データＤｂが、そ

れぞれ時間軸に沿った音声信号の波形として示されている。楽曲データＤａ及びＤｂは、例えば、ＷＡＶＥ又はＭＰ３（MPEG Audio Layer-3）など、任意の形式で表現された音声データであってよい。

【００２７】

また、楽曲データＤａにはメタデータＭａ、楽曲データＤｂにはメタデータＭｂが付与されている。図１の例において、メタデータＭａ及びＭｂには、それぞれ、“Ｃ＃”、“Ｃ６”、“Ｃ７”及び“Ｅ７”などのコード進行データ、時間軸に対する垂線で示されているビート位置データ、並びにオフセットＦａ及びＦｂにより示されているオフセットデータが含まれる。なお、本明細書において、オフセットとは、楽曲データの開始点から演奏の開始点（再生時に音が始めて認識され得る点）までの時間差を意味する。

【００２８】

図１には、さらにリミックスデータＲａｂも示されている。リミックスデータとは、複数の楽曲データの全部又は一部を組み合わせて全体として新たな楽曲を構成するための、いわゆる“レシピ”（構成）に関するデータである。図１を参照すると、リミックスデータＲａｂの内容の一例として、楽曲データＤａの再生中に楽曲データＤａをフェードアウトさせ、そのフェードアウトの途中から楽曲データＤｂをフェードインさせるという構成が示されている。なお、リミックスデータは、図１に示した例に限定されず、例えば、再生時間の指定、音量の指定、転調、又は繰返しなど、音声信号の編集に関する任意のデータであってよい。

【００２９】

ここで、リミックスデータＲａｂに基づいて楽曲データＧａ及びＧｂを正確に演奏するには、オフセットＦａ及びＦｂが、楽曲データＤａ及びＤｂのオフセットをそれぞれ正確に表していることが前提となる。これに対し、例えば、メタデータＭｂに含まれるオフセットＦｂが楽曲データＤｂのオフセットとずれていれば、リミックスデータＲａｂにおいて楽曲データＤｂをフェードインさせようとした時点で、正しく楽曲データＤｂがフェードインしない可能性がある。

【００３０】

図２は、楽曲Ｇａが異なる環境Ｅ１及びＥ２の下で録音される状況を示している。図２において、環境Ｅ１で録音された楽曲Ｇａの楽曲データＤａ’のオフセットと、環境Ｅ２で録音された楽曲Ｇａの楽曲データＤａ’’のオフセットには、例えばデータ読取装置の違いやエンコードパラメータの違いにより、差異が生じている。その結果、事前に楽曲データＤａに基づいて生成されたメタデータＭａのオフセットＦａは、楽曲データＤａ’のオフセットとは一致するが、楽曲データＤａ’’のオフセットとはズレている。即ち、環境Ｅ２で録音された楽曲データＤａ’’には、もはや別途入手したメタデータＭａを適用できない。

【００３１】

このような録音時の環境の相違等に起因するオフセットのずれは、楽曲データの利用形態が多様化し、メタデータが単独で流通する可能性の高い現状では、避けて通れない課題である。そこで、本明細書において説明するような、オフセットのずれに影響されることなく楽曲データ又は楽曲データのメタデータを使用可能とする装置の実現が望まれる。

【００３２】

＜２．第１の実施形態＞

図３は、本発明の第１の実施形態に係る楽曲処理装置１０の論理的な構成を示すブロック図である。図３を参照すると、楽曲処理装置１０は、記憶部１２、特徴パターン生成部１４、開始位置決定部１６、符号値決定部１８、楽曲同定部２０、及びオフセット補正部２２を備える。

【００３３】

［記憶部］

記憶部１２は、ハードディスク又はフラッシュメモリなどの記憶装置を用いて、楽曲処理装置１０の処理対象となる楽曲データ、及び後述する処理により生成される特徴パター

ンなどを記憶する。

【００３４】

［特徴パターン生成部］

特徴パターン生成部１４は、記憶部１２から楽曲データを取得し、取得した楽曲データの音量変化の特徴を表す特徴パターンを生成する。本実施形態において、特徴パターンは、楽曲データ内で時間軸上の一定の区間ごとに前の区間との間の相対的な音量変化を符号化した符号値の配列として与えられる。特徴パターン生成部１４は、特徴パターンの生成に際し、まず開始位置決定部１６を呼び出し、符号化を開始する時間軸上の位置、即ち符号化開始位置を決定させる。

【００３５】

［開始位置決定部］

開始位置決定部１６は、特徴パターン生成部１４から楽曲データを受け取り、当該楽曲データ内の時間軸上で音量変化の特徴の符号化を開始すべき符号化開始位置を決定する。例えば、符号化開始位置は、時間軸上の所定のウィンドウ幅における平均音量が楽曲データの全体又は一部分の平均音量を初めて上回るウィンドウ位置に応じて決定することができる。以下、図４及び図５を用いて、開始位置決定部１６による符号化開始位置の決定処理の一例について説明する。

【００３６】

図４に示すように、開始位置決定部１６は、まず楽曲データの所定の範囲内で、符号化開始位置の決定の基準となる基準平均音量Ｖｍを計算する。基準平均音量Ｖｍを計算する所定の範囲とは、図４において、ＴａからＴｂまでの期間として示されている。ＴａからＴｂまでの範囲には、楽曲データの全体が含まれていてもよく、又は楽曲データの一部分のみが含まれていてもよい。なお、量子化された楽曲データの音量には正負の符号が付されるが、基準平均音量Ｖｍは、音量の絶対値の平均値であってもよく、音量を二乗した後の平均値であってもよい。

【００３７】

次に、図５に示すように、開始位置決定部１６は、基準平均音量Ｖｍを決定した範囲の先頭位置Ｔａを始点とし、ウィンドウ幅Ｗｓを持つ時間軸上のウィンドウにおける平均音量Ｖｗが基準平均音量Ｖｍを初めて上回る位置を、時間軸に沿って探索する。

【００３８】

例えば、図５（Ａ）に示したウィンドウ位置では、ウィンドウ幅Ｗｓにおける平均音量Ｖｗは基準平均音量Ｖｍよりも小さい。そのため、開始位置決定部１６は、ウィンドウ位置を時間軸上の正の方向に移動させ、再びウィンドウ幅Ｗｓにおける平均音量Ｖｗと基準平均音量Ｖｍを比較する。ウィンドウ位置の移動は、符号化開始位置の決定の再現性を高めるためには、より小さい単位で行われるのが望ましい。例えば、ウィンドウ位置の移動は、楽曲データのサンプリングにおける１サンプルごととすることができる。

【００３９】

その後、ウィンドウ位置を時間軸上の正の方向に移動させていき、例えば、ウィンドウの始点が図５（Ｂ）に示した位置Ｔｓに到達したときに、ウィンドウ幅Ｗｓにおける平均音量Ｖｗが基準平均音量Ｖｍよりも初めて大きくなったとする。開始位置決定部１６は、このような位置Ｔｓを検出し、Ｔｓを符号化開始位置として特徴パターン生成部１４へ出力する。

【００４０】

特徴パターン生成部１４は、符号化開始位置Ｔｓが決定されると、符号値決定部１８を呼び出し、符号化開始位置Ｔｓから一定の区間ごとに特徴パターンを生成するための符号値を決定させる。本実施形態において、特徴パターンを生成するための符号値は、一定の区間ごとの前の区間との間の音量の変化に応じた値とする。

【００４１】

［符号値決定部］

図６は、符号値決定部１８による符号値決定処理の一例を説明するための説明図である

。図６を参照すると、符号化開始位置Ｔｓを起点とし、区間幅Ｗｅを持つ９つの符号化区間が示されている。また、各符号化区間において計算された区間別の平均音量が、符号化区間ごとの矩形の高さとして示されている。本実施形態において、符号値決定部１８は、図６に示したような区間別平均音量を順次計算し、計算した区間別平均音量が前の区間における平均音量よりも増加しているか又は減少しているかに応じた符号値を決定する。

【００４２】

図６の例では、符号値は、ある区間における平均音量が前の区間における平均音量よりも増加していれば“１”、同じか又は減少していれば“０”と決定されている。例えば、符号化開始位置Ｔｓを起点とする第１区間の平均音量よりも第２区間の平均音量は大きいため、第１の符号値（左端の符号値）は“１”となっている。同様に、第２区間の平均音量よりも第３区間の平均音量は大きいため、第２の符号値も“１”となっている。これに対し、第３区間の平均音量よりも第４区間の平均音量は小さいため、第３の符号値は“０”となっている。

【００４３】

特徴パターン生成部１４は、このように符号値決定部１８により決定された区間別の符号値を複数の区間にわたって結合し、一連の符号値としての特徴パターンを生成する。図６の例では、最終的に、各符号化区間について決定された符号値から、特徴パターンＣＰ＝“１１０１１０１０”が生成されている。なお、ここでは説明の便宜上、９つの符号化区間を例示しているが、符号化区間の数（即ち特徴パターンのビット数）は、特徴パターンを用いて楽曲データを相互に識別可能な程度に十分な数とするのが好適である。

【００４４】

また、特徴パターンは、時間軸上の相対的な音量変化を表す符号値であれば、二値ではなく多値の符号値によって生成されてもよい。また、符号値は、図６に示した一定の区間ごとの前の区間との間の音量の変化に応じた値ではなく、例えばある区間内での平均音量の基準平均音量に対する音量比など、任意の音量変化の特徴に応じた値であってよい。

【００４５】

以上の処理により生成された特徴パターンＣＰ、並びに特徴パターンＣＰの生成に使用されたウィンドウ幅Ｗｓ及び符号化区間幅Ｗｅは、後述する楽曲データの同定のために用いられる。また、符号化開始位置Ｔｓは、オフセットの補正のために用いられる。図７に示したように、特徴パターン生成部１４は、かかる特徴パターンＣＰ、ウィンドウ幅Ｗｓ、符号化区間幅Ｗｅ、及び符号化開始位置Ｔｓを一まとまりのデータセットとして、例えば楽曲データを識別する楽曲識別子により楽曲データと関連付けて、記憶部１２に格納する。

【００４６】

［処理フロー説明：特徴パターン生成］

図８～図１０は、本実施形態に係る楽曲処理装置１０による特徴パターンの生成の流れを説明するためのフローチャートである。

【００４７】

図８を参照すると、まず、特徴パターン生成部１４により、楽曲データが記憶部１２から取得される（Ｓ１１０）。次に、特徴パターン生成部１４から呼び出された開始位置決定部１６により、基準範囲Ｔａ～Ｔｂ内での基準平均音量Ｖｍが計算される（Ｓ１２０）。そして、開始位置決定部１６により、基準平均音量Ｖｍ及びウィンドウ幅Ｗｓを用いて符号化開始位置Ｔｓが決定される（Ｓ１３０）。

【００４８】

図９は、図８のＳ１３０における符号化開始位置の決定の流れをより詳細を示している。

【００４９】

図９を参照すると、まず、基準範囲の始点Ｔａを用いてウィンドウの初期位置が決定される（Ｓ３０２）。次に、当該ウィンドウ位置におけるウィンドウ内の平均音量Ｖｗが計算される（Ｓ３０４）。そして、計算されたウィンドウ内の平均音量Ｖｗと基準平均音量

Ｖｍが比較される（Ｓ３０６）。このとき、例えばウィンドウ内の平均音量Ｖｗの方が小さいか同等であれば、ウィンドウ位置を例えば１サンプル先へ進めた後、処理はＳ３０４へ戻る（Ｓ３０８）。一方、Ｓ３０６においてウィンドウ内の平均音量Ｖｗの方が大きければ、その時点のウィンドウ位置に応じて符号化開始位置Ｔｓが決定される（Ｓ３１０）。

【００５０】
図８に戻り、特徴パターンの生成の流れの説明を継続する。符号化開始位置Ｔｓが決定されると、特徴パターン生成部１４は、符号化開始位置Ｔｓ及び符号化区間幅Ｗｅを用いて、特徴パターンＣＰを生成する（Ｓ１４０）。

【００５１】
図１０は、図８のＳ１４０における特徴パターンの生成の流れをより詳細を示している。

【００５２】
図１０を参照すると、まず、前区間の平均音量Ｖｐの初期値として、第１区間の平均音量が設定される（Ｓ４０２）。次に、第ｎ区間の平均音量Ｖｎが計算される（Ｓ４０４）。初回の計算では、ｎ＝２である。そして、計算された第ｎ区間の平均音量Ｖｎが前区間の平均音量Ｖｐと比較される（Ｓ４０６）。このとき、例えば前区間の平均音量Ｖｐの方が大きいか同等であれば、第ｎ区間の符号値Ｃｎ＝“０”と決定される（Ｓ４０８）。また、第ｎ区間の平均音量Ｖｎの方が大きければ、第ｎ区間の符号値Ｃｎ＝“１”と決定される（Ｓ４１０）。その後、全区間の符号値の決定が終了したか否かが判定される（Ｓ４１２）。ここで、全区間の符号値の決定が終了していなければ、前区間の平均音量Ｖｐに第ｎ区間の平均音量が代入され、及びｎがｎ＋１に加算され、処理はＳ４０４へ戻る（Ｓ４１４）。一方、Ｓ４１２において全区間の符号値の決定が終了していれば、各区間の符号値Ｃｎ（ｎ＝２、３…）から特徴パターンＣＰが生成される（Ｓ４１６）。

【００５３】
図８に戻り、特徴パターンＣＰが生成されると、特徴パターン生成部１４は、特徴パターンＣＰ、ウィンドウ幅Ｗｓ、符号化区間幅Ｗｅ、及び符号化開始位置Ｔｓを一まとまりのデータセットとして、記憶部１２に格納する。（Ｓ１５０）。

【００５４】
ここまで、図４～図１０を用いて、本実施形態に係る特徴パターンの生成処理について主に説明した。本実施形態において生成される特徴パターンは、前述した通り、楽曲データの時間軸上の相対的な音量変化を符号化した符号値の配列である。楽曲データの音量は、データ読取装置の違いやエンコードパラメータの違いなど、環境の相違によって変動し得る。しかしながら、そうした環境の相違は、通常、楽曲データの全体に対して一様に影響を与えるものである。そのため、時間軸上の相対的な音量変化に基づいて符号値を決定することで、環境の相違による影響はキャンセルされる。このことから、本実施形態において生成される特徴パターンは、互いに異なる環境で録音された同一の楽曲を、環境の相違に影響されることなく同定するために適していることが理解される。

【００５５】
また、本実施形態において、特徴パターンの生成に用いた符号化開始位置は、同様に時間軸上の相対的な音量変化に基づいて決定される。即ち、同一の楽曲データであれば、符号化開始位置に差異を生じさせるのは、実質的には楽曲データ内の時間軸上の演奏の位置、即ちオフセットのみとなる。このことから、本実施形態において決定される符号化開始位置は、同一の楽曲データ又はそのメタデータを用いる際に、オフセットのずれを補正するために用いられ得ることが理解される。

【００５６】
［楽曲同定部］
次に、図３に戻り、ここまでに説明した特徴パターンを用いて楽曲を同定する楽曲同定部２０について説明する。

【００５７】

楽曲同定部２０は、ある楽曲データについて生成された特徴パターンと任意の楽曲データの特徴パターンを比較することにより、楽曲データ間の特徴パターンの類似度を計算する。楽曲同定部２０により比較の対象とされる特徴パターンは、同一の楽曲処理装置１０により生成された特徴パターンであってもよく、異なる装置により生成された特徴パターンであってもよい。但し、特徴パターンの生成に同一のウィンドウ幅Ｗｓ及び符号化区間幅Ｗｅが用いられていることを前提とする。

【００５８】

例えば、楽曲データＤの特徴パターンＣＰ、並びに特徴パターンＣＰの生成に用いられたウィンドウ幅Ｗｓ及び符号化区間幅Ｗｅが既知であったとする。このとき、別の楽曲データＤａについて、楽曲データＤとの同一性を判定するためには、楽曲データＤａの特徴パターンＣＰａをウィンドウ幅Ｗｓ及び符号化区間幅Ｗｅを用いて生成し、生成した特徴パターンＣＰａを特徴パターンＣＰと比較する。

【００５９】

図１１は、楽曲同定部２０による楽曲同定処理の一例を説明するための説明図である。

【００６０】

図１１において、楽曲データＤの特徴パターンＣＰ、楽曲データＤａの特徴パターンＣＰａ、及び楽曲データＤｂの特徴パターンＣＰｂが、同一のウィンドウ幅Ｗｓ及び符号化区間幅Ｗｅを用いて、既に特徴パターン生成部１４により生成されているものとする。かかる特徴パターンは、例えば記憶部１２から楽曲同定部２０へ読み込まれる。

【００６１】

そして、楽曲同定部２０は、楽曲データＤと楽曲データＤａとの間の同一性を判定するために、特徴パターンＣＰと特徴パターンＣＰａの間で、ビットごとの排他的論理和（ＥＸＯＲ）を計算している。図１１に示している通り、特徴パターンＣＰと特徴パターンＣＰａの間の排他的論理和は、“００００１１００”である。また、楽曲同定部２０は、楽曲データＤと楽曲データＤｂとの間の同一性を判定するために、特徴パターンＣＰと特徴パターンＣＰｂの間で、ビットごとの排他的論理和を計算している。特徴パターンＣＰと特徴パターンＣＰｂの間の排他的論理和は、“０１１０１１０１”である。

【００６２】

ここで、排他的論理和の計算結果のうち“０”のビットは、特徴パターンのそのビットの符号値が一致していたことを意味している。また、“１”のビットは、特徴パターンのそのビットの符号値が異なっていたことを意味している。即ち、排他的論理和の計算結果の中で“０”のビットの割合が大きいほど、特徴パターンは相互に類似している。

【００６３】

よって、楽曲同定部２０は、例えば、２つの特徴パターンの間で排他的論理和の計算結果が“０”であるビット数の、全ビット数に対する割合を、特徴パターンの類似度としてもよい。図１１の例では、特徴パターンＣＰと特徴パターンＣＰａの間の排他的論理和の計算結果“００００１１００”から、当該特徴パターンの類似度ＳＩＭａ＝６［ｂｉｔ］／８［ｂｉｔ］＝７５％と計算されている。同様に、特徴パターンＣＰと特徴パターンＣＰｂの間の排他的論理和の計算結果“０１１０１１０１”から、当該特徴パターンの類似度ＳＩＭｂ＝３［ｂｉｔ］／８［ｂｉｔ］＝３７．５％と計算されている。

【００６４】

また、図１１には示していないが、符号値の一致する“０”ビットの単純な割合ではなく、２つの特徴パターンの間で符号値の一致する“０”ビットと符号値の一致しない“１”ビットの二項分布における発生確率を考慮して類似度を計算してもよい。より具体的には、例えば、符号値の一致する“０”ビットの単純な割合を$Ｐ_0$とすれば、二項分布を考慮した類似度は、試行回数＝全ビット数、成功確率＝５０％の二項分布において、確率変数$Ｘ \geqq Ｐ_0$となる確率$Ｘ_0$に応じて与えられる（但し、この場合は確率$Ｘ_0$が小さいほど類似度が大きいことに注意すべきである）。

【００６５】

また、排他的論理和の計算結果において、符号値の一致する“０”ビットが連続して現

れる最大の長さを用いて特徴パターンの類似度を計算してもよい。例えば、異なる２つの環境で録音した同一の楽曲の特徴パターンを比較する場合を考える。このとき、一方の楽曲データの録音時にのみ一部に雑音が入ったとすると、雑音部分の特徴パターンの符号値は、５０％の確率で一致しない。雑音が入っていない部分では、理想的には全ての符号値が一致する。これに対し、異なる楽曲の特徴パターンを比較した場合には、符号値が一致しない箇所は、雑音部分など特定の範囲に偏らず、特徴パターン内の全体に散らばっているはずである。そのため、例えば、符号値の一致する“０”ビットが連続して現れる最大の長さを特徴パターンの類似度として扱えば、雑音に対しより堅牢性の高い楽曲の同一性の判定を行うことができる。

【００６６】

楽曲同定部２０は、このようないずれかの手法により計算した特徴パターンの類似度と例えば予め決定された閾値とを比較することにより、さらに２つの楽曲データの楽曲としての同一性を判定してもよい。また、楽曲同定部２０は、計算した特徴パターンの類似度を楽曲処理装置１０に接続された表示装置を介してユーザに呈示することにより、楽曲の同一性をユーザに判定させてもよい。

【００６７】

［処理フロー説明：楽曲の同定］

図１２は、本実施形態に係る楽曲処理装置１０による楽曲の同定の流れを説明するためのフローチャートである。

【００６８】

図１２を参照すると、まず、楽曲同定部２０は、２つの特徴パターンＣＰｘ及びＣＰｙを記憶部１２から取得する（Ｓ６１０）。次に、楽曲同定部２０は、特徴パターンＣＰｘ及びＣＰｙの間で排他的論理和を計算する（Ｓ６２０）。そして、楽曲同定部２０は、排他的論理和の計算結果から、特徴パターンＣＰｘ及びＣＰｙの類似度ＳＩＭを計算する（Ｓ６３０）。その後、楽曲同定部２０は、例えば、計算した類似度ＳＩＭを予め決定された閾値と比較して楽曲の同一性を判定し、又は類似度ＳＩＭを呈示して楽曲の同一性をユーザに判定させる（Ｓ６４０）。

【００６９】

ここまで、図１１及び図１２を用いて、本実施形態に係る楽曲処理装置１０による楽曲の同定処理について説明した。本実施形態に係る楽曲の同定処理は、前述した音量変化の特徴を表す特徴パターンを用いて行われる。そのため、例えば、環境の相違によるオフセットのずれなどに影響されることなく、楽曲の同定を実用的な精度で行うことが可能となる。

【００７０】

［オフセット補正部］

続いて、前述した楽曲の同定処理の結果、及び特徴パターンの生成に用いた符号化開始位置に基づいて、データの開始点から演奏の開始点までのオフセットのずれを補正する処理について説明する。

【００７１】

オフセット補正部２２は、類似度の大きい楽曲データの特徴パターンに対応する前述した符号化開始位置を用いて、楽曲データ又はメタデータのオフセットを補正する。図１３は、オフセット補正部２２によるオフセット補正処理の一例を説明するための説明図である。

【００７２】

図１３（Ａ）には、楽曲Ｇａの楽曲データＤａの波形と、楽曲データＤａに対して付与されたメタデータＭａが示されている。さらに、図１３（Ａ）には、メタデータＭａに含まれる楽曲データＤａのオフセットＴｏ、及び楽曲データＧａの符号化開始位置Ｔｓが示されている。

【００７３】

また、図１３（Ｂ）には、楽曲データＤａとは別の環境で録音された楽曲Ｇａの楽曲デ

ータＤａ’の波形が示されている。楽曲データＤａ’の波形を参照すると、楽曲データＤａ’の演奏開始のタイミングは、楽曲データＤａよりも遅い時点となっている。即ち、環境の相違の影響から楽曲データＤａと楽曲データＤａ’のオフセットにずれが生じている。そのため、楽曲データＤａに対して付与されたメタデータＭａをそのまま楽曲データＤａ’に適用すると、メタデータＭａに含まれるビート位置やコード進行などの時系列のデータは楽曲データＤａ’に合致しない。

【００７４】

そこで、オフセット補正部２２は、楽曲データＤａ’について開始位置決定部１６により決定された符号化開始位置Ｔｓ’を取得し、楽曲データＧａの符号化開始位置Ｔｓとの差に応じて、メタデータＭａを楽曲データＤａ’に適用する際のオフセットを補正する。補正後のオフセットＴｏ′は、次式により計算される。

【００７５】

【数１】

$$To' = To + (Ts' - Ts) \qquad (1)$$

【００７６】

図１３に示されているように、このようにメタデータＭａのオフセットを補正して楽曲データＤａ’に適用することで、メタデータＭａに含まれるビート位置やコード進行などの時系列のデータを楽曲データＤａ’に合致させることができる。

【００７７】

なお、図１３ではメタデータＭａのオフセットを補正値Ｔｓ’－Ｔｓを用いて補正する例について説明した。しかしながら、その代わりに、楽曲データＤａ’のオフセットを補正値Ｔｓ’－Ｔｓを用いて補正してもよい。即ち、楽曲データＤａ’の再生時に冒頭の再生時間を補正値Ｔｓ’－Ｔｓの分だけスキップして再生することで、メタデータＭａをそのまま再生された演奏に合致させることができる。

【００７８】

また、ここでは楽曲データＤａと楽曲データＤａ’が同一の楽曲Ｇａに由来することが分かっていることを前提として説明した。しかしながら、例えば、複数の楽曲データの候補の中から、補正対象の楽曲データとの間で特徴パターンの類似度が大きい楽曲データを自動的に又はユーザにより選択させ、選択された楽曲データの符号化開始位置に基づいてオフセットを補正してもよい。

【００７９】

［処理フロー説明：オフセット補正］

図１４は、本実施形態に係る楽曲処理装置１０によるオフセットの補正の流れを説明するためのフローチャートである。

【００８０】

図１４を参照すると、まず、オフセット補正部２２は、補正対象の楽曲の符号化開始位置Ｔｓ’を取得する（Ｓ７１０）。次に、オフセット補正部２２は、補正対象の楽曲との間で特徴パターンの類似度の大きい楽曲データの符号化開始位置Ｔｓを取得する（Ｓ７２０）。かかる符号化開始位置Ｔｓ及びＴｓ’は、例えば、開始位置決定部１６により事前に決定される。さらに、オフセット補正部２２は、オフセットの補正値Ｔｓ’－Ｔｓを計算する（Ｓ７３０）。そして、オフセット補正部２２は、補正対象の楽曲データ又は補正対象の楽曲データに適用するメタデータのオフセットを、補正値Ｔｓ’－Ｔｓを用いて補正する（Ｓ７４０）。

【００８１】

ここまで、図１３及び図１４を用いて、本実施形態に係る楽曲処理装置１０によるオフセットの補正処理について説明した。本実施形態に係るオフセットの補正処理は、楽曲データ内の時間軸上の演奏の位置を実質的に表す符号化開始位置を用いて行われる。それにより、楽曲データ又は楽曲データに付与されたメタデータを、環境の相違によるオフセットのずれを補正した上で使用することが可能となる。

【００８２】

以上、図３～図１４を参照しながら、本発明の第１の実施形態について詳細に説明した。本発明の第１の実施形態では、１つの楽曲データについて、時間軸上の相対的な音量変化の特徴を表す１つの特徴パターンを生成した。ここで、図６の説明からも理解されるように、特徴パターンに含まれる符号値は、開始位置決定部１６により決定される符号化開始位置に依存する。そのため、１つの楽曲データについて複数の符号化開始位置に対応する複数の特徴パターンを生成し、楽曲の同一性の判定の品質を高めることも考えられる。そこで、次節では、本発明の第２の実施形態として、１つの楽曲データについて異なるウィンドウ位置を用いて決定した複数の符号化開始位置に基づいて、複数の特徴パターンを生成する例について説明する。

【００８３】

＜３．第２の実施形態＞

図１５は、本発明の第２の実施形態に係る楽曲処理装置３０の論理的な構成を示すブロック図である。図１５を参照すると、楽曲処理装置３０は、記憶部１２、特徴パターン生成部３４、開始位置決定部３６、符号値決定部３８、楽曲同定部４０、及びオフセット補正部２２を備える。

【００８４】

［特徴パターン生成部］

本実施形態において、特徴パターン生成部３４は、記憶部１２から楽曲データを取得し、取得した楽曲データの音量変化の特徴を表す特徴パターンを複数生成する。特徴パターン生成部３４は、特徴パターンの生成に際し、まず開始位置決定部３６を呼び出し、異なるウィンドウ幅を用いて複数の符号化開始位置を決定させる。

【００８５】

［開始位置決定部］

図１６は、本実施形態における、開始位置決定部３６による符号化開始位置の決定処理を説明するための説明図である。

【００８６】

図１６を参照すると、開始位置決定部３６は、基準平均音量Ｖｍを計算した後、ウィンドウ幅$Ｗｓ_1$を持つ時間軸上のウィンドウにおける平均音量$Ｖｗ_1$が基準平均音量Ｖｍを初めて上回る位置$Ｔｓ_1$を決定する（図１６（Ａ））。次に、開始位置決定部３６は、ウィンドウ幅$Ｗｓ_1$とは異なるウィンドウ幅$Ｗｓ_2$を持つ時間軸上のウィンドウにおける平均音量$Ｖｗ_2$が基準平均音量Ｖｍを初めて上回る位置$Ｔｓ_2$を決定する（図１６（Ｂ））。このように、本実施形態において、開始位置決定部３６は、異なるウィンドウ幅$Ｗｓ_i$（ｉ＝１，２…）を用いて複数の符号化開始位置$Ｔｓ_i$を決定し、決定した複数の符号化開始位置$Ｔｓ_i$を特徴パターン生成部３４へ出力する。

【００８７】

［符号値決定部］

図１７は、本実施形態における、符号値決定部３８による符号値決定処理を説明するための説明図である。

【００８８】

図１７（Ａ）には、符号化開始位置$Ｔｓ_1$を起点とし、区間幅Ｗｅを持つ９つの符号化区間が示されている。また、図１７（Ｂ）には、符号化開始位置$Ｔｓ_2$を起点とし、同様に区間幅Ｗｅを持つ９つの符号化区間が示されている。符号値決定部３８は、このような各符号化区間について、区間別平均音量が前の区間における平均音量よりも増加しているか又は減少しているかに応じて符号値を決定する。

【００８９】

そして、特徴パターン生成部３４は、決定された区間別の符号値を一連の符号化区間にわたって結合し、符号化開始位置$Ｔｓ_i$ごとの特徴パターン$ＣＰ_i$を生成する。図１７の例では、符号化開始位置$Ｔｓ_1$について特徴パターン$ＣＰ_1$＝“１１０１１０１０”、符号化開始位置$Ｔｓ_2$について$ＣＰ_2$＝“００１１１０１０”が生成されている。

【００９０】
以上の処理により生成された特徴パターンＣＰ$_i$、並びに特徴パターンＣＰ$_i$の生成に使用されたウィンドウ幅Ｗｓ$_i$及び符号化区間幅Ｗｅは、後述する楽曲データの同定のために用いられる。また、符号化開始位置Ｔｓ$_i$は、オフセットの補正のために用いられる。図１８にｉ≦２の場合について示したように、特徴パターン生成部３４は、かかる特徴パターンＣＰ$_i$、ウィンドウ幅Ｗｓ$_i$、符号化区間幅Ｗｅ、及び符号化開始位置Ｔｓ$_i$を一まとまりのデータセットとして記憶部１２に格納する。

【００９１】
［処理フロー説明：複数の特徴パターンの生成］
図１９は、本実施形態に係る楽曲処理装置３０による特徴パターンの生成の流れを説明するためのフローチャートである。

【００９２】
図１９を参照すると、まず、特徴パターン生成部３４により、楽曲データが記憶部１２から取得される（Ｓ１１１０）。次に、特徴パターン生成部３４から呼び出された開始位置決定部３６により、基準範囲Ｔａ～Ｔｂ内での基準平均音量Ｖｍが計算される（Ｓ１１２０）。そして、開始位置決定部３６により、基準平均音量Ｖｍとウィンドウ幅Ｗｓ$_i$（ｉ＝１，２…）とを用いて、複数の符号化開始位置Ｔｓ$_i$が決定される（Ｓ１１３０）。そして、特徴パターン生成部３４により、複数の符号化開始位置Ｔｓ$_i$と符号化区間幅Ｗｅとを用いて、複数の特徴パターンＣＰ$_i$を生成する（Ｓ１１４０）。かかるＳ１１３０及びＳ１１４０は、全てのウィンドウ幅Ｗｓ$_i$について特徴パターンＣＰ$_i$が生成されるまで繰り返される。そして、全てのウィンドウ幅Ｗｓ$_i$について特徴パターンＣＰ$_i$が生成されると、特徴パターン生成部３４は、特徴パターンＣＰ$_i$、ウィンドウ幅Ｗｓ$_i$、符号化区間幅Ｗｅ、及び符号化開始位置Ｔｓ$_i$を一まとまりのデータセットとして、記憶部１２に格納する。（Ｓ１１５０）。

【００９３】
［楽曲同定部］
次に、このような複数の特徴パターンを用いて楽曲を同定する楽曲同定部４０について説明する。楽曲同定部４０は、ある楽曲データについて生成された各特徴パターンと任意の楽曲データについて生成された各特徴パターンとを比較することにより、特徴パターンごとに類似度を計算する。このとき、同一のウィンドウ幅Ｗｓ$_i$に基づいて生成された特徴パターン同士を比較するものとする。

【００９４】
図２０は、楽曲同定部４０による楽曲同定処理の一例を説明するための説明図である。

【００９５】
図２０を参照すると、楽曲データＤについて生成された２つの特徴パターンＣＰ$_1$及びＣＰ$_2$と、楽曲データＤｒｅｆについて生成された２つの特徴パターンＣＰｒｅｆ$_1$及びＣＰｒｅｆ$_2$が示されている。このうち、特徴パターンＣＰ$_1$とＣＰｒｅｆ$_1$、特徴パターンＣＰ$_2$とＣＰｒｅｆ$_2$はそれぞれ、同一のウィンドウ幅に基づいて生成された特徴パターンである。

【００９６】
楽曲同定部４０は、まず、かかる特徴パターンの間でそれぞれ、ビットごとの排他的論理和（ＥＸＯＲ）を計算する。図２０の例において、特徴パターンＣＰ$_1$とＣＰｒｅｆ$_1$の間の排他的論理和は、“００００００１”と計算されている。また、特徴パターンＣＰ$_2$とＣＰｒｅｆ$_2$の間の排他的論理和は、“１１１００００１”と計算されている。

【００９７】
そして、楽曲同定部４０は、排他的論理和の計算結果を用いて、ウィンドウ幅ごとの特徴パターンの類似度ＳＩＭ$_1$、ＳＩＭ$_2$を計算する。図２０の例では、“０”ビットの単純な割合を特徴パターンの類似度としており、特徴パターンＣＰ$_1$とＣＰｒｅｆ$_1$の類似度ＳＩＭ$_1$＝８７．５％、特徴パターンＣＰ$_2$とＣＰｒｅｆ$_2$の類似度ＳＩＭ$_2$＝５０％と計算されている。

【００９８】
図２０から理解されるように、同じ楽曲データの組合せである楽曲データＤと楽曲データＤｒｅｆの間であっても、符号化開始位置の決定に用いるウィンドウ幅によって、特徴パターンの類似度は変動し得る。このとき、例えば、複数の類似度のうち最も大きい類似度を、その楽曲データ間の類似度として扱うことができる。その場合、図２０の例では、類似度SIM_1の方が類似度SIM_2よりも大きいため、類似度SIM_1が楽曲データＤと楽曲データD_{ref}の間の類似度として採用される。
【００９９】
採用された類似度SIM_1の計算に用いられた特徴パターンCP_1と$CPref_1$の符号化開始位置の差は、楽曲データＤと楽曲データＤｒｅｆの間のオフセットのずれに相当する。そのため、特徴パターンCP_1と$CPref_1$の符号化開始位置の差を、後述するオフセットの補正処理において補正値として用いることができる。
【０１００】
［処理フロー説明：楽曲の同定及びオフセットの補正］
図２１は、本実施形態に係る楽曲処理装置３０による楽曲同定及びオフセット補正の流れを説明するためのフローチャートである。
【０１０１】
図２１を参照すると、まず、楽曲同定部４０により、補正対象の楽曲データの特徴パターンCP_i（ｉ＝１，２…）が取得される（Ｓ１２１０）。次に、楽曲同定部４０により、参照する楽曲データの特徴パターン$CPref_i$（ｉ＝１，２…）が取得される（Ｓ１２２０）。
【０１０２】
その後、楽曲同定部４０により、ウィンドウ幅ごとに特徴パターンCP_iと特徴パターン$CPref_i$との間の排他的論理和が計算される（Ｓ１２３０）。さらに、排他的論理和の計算結果から、ウィンドウ幅ごとの類似度SIM_iが計算される（Ｓ１２４０）。かかるＳ１２３０及びＳ１２４０は、全てのウィンドウ幅について類似度SIM_iが計算されるまで繰り返される。そして、全ての類似度SIM_iの計算が終了すると、楽曲同定部４０により、類似度SIM_iが最も大きくなるｉが決定される（Ｓ１２５０）。図２１においては、類似度SIM_iが最も大きくなるｉをｉ＝ｋとする。
【０１０３】
次に、オフセット補正部２２により、類似度SIM_kに対応する特徴パターンCP_kの符号化開始位置Ts_kと特徴パターン$CPref_k$の符号化開始位置$Tsref_k$とを用いて、オフセットの補正値が計算される（Ｓ１２６０）。そして、オフセット補正部２２により、計算された補正値を用いて、補正対象の楽曲データ又は楽曲データのメタデータのオフセットが補正される（Ｓ１２７０）。
【０１０４】
以上、図１５～図２１を参照しながら、本発明の第２の実施形態について説明した。本実施形態によれば、異なるウィンドウ幅を用いて決定された複数の符号化開始位置に基づいて、１つの楽曲データに対し複数の特徴パターンが生成される。そして、生成された複数の特徴パターンのうち、参照される特徴パターンとの間で最も類似度の大きい特徴パターンに対応する符号化開始位置を用いて、楽曲データ又は楽曲データのメタデータのオフセットが補正される。それにより、類似度に基づく楽曲の同一性の判定の品質が向上し、オフセットの補正の精度を高めることができる。
【０１０５】
なお、本明細書において説明した第１及び第２の実施形態に係る一連の処理は、典型的にはソフトウェアを用いて実現される。一連の処理又はその一部をソフトウェアで実行させる場合には、ソフトウェアを構成するプログラムが、例えば図２２に示した汎用コンピュータなどを用いて実行される。
【０１０６】
図２２において、ＣＰＵ（Central Processing Unit）９０２は、汎用コンピュータ

の動作全般を制御する。ＲＯＭ（Read Only Memory）９０４には、一連の処理の一部又は全部を記述したプログラム又はデータが格納される。ＲＡＭ（Random Access Memory）９０６には、処理の実行時にＣＰＵ９０２により用いられるプログラムやデータなどが一時的に記憶される。

【０１０７】

ＣＰＵ９０２、ＲＯＭ９０４、及びＲＡＭ９０６は、バス９１０を介して相互に接続される。バス９１０にはさらに、入出力インタフェース９１２が接続される。

【０１０８】

入出力インタフェース９１２は、ＣＰＵ９０２、ＲＯＭ９０４、及びＲＡＭ９０６と、入力装置９２０、出力装置９２２、記憶装置９２４、通信装置９２６、及びドライブ９３０とを接続するためのインタフェースである。

【０１０９】

入力装置９２０は、例えばマウス、キーボード、又はタッチパネルなどの入力装置を介して、ユーザからの指示や情報入力を受け付ける。出力装置９２２は、例えばＣＲＴ（Cathode Ray Tube）、液晶ディスプレイ、ＯＬＥＤ（Organic Light Emitting Diode）などの表示装置、又はスピーカなどの音声出力装置を介してユーザに情報を出力する。

【０１１０】

記憶装置９２４は、例えばハードディスクドライブ又はフラッシュメモリなどにより構成され、プログラムやプログラムデータなどを記憶する。通信装置９２６は、ＬＡＮ又はインターネットなどのネットワークを介する通信処理を行う。ドライブ９３０は、必要に応じて汎用コンピュータに設けられ、例えばドライブ９３０にはリムーバブルメディア９３２が装着される。

【０１１１】

第１～第２の実施形態に係る一連の処理をソフトウェアで実行する場合には、例えば図２２に示したＲＯＭ９０４、記憶装置９２４、又はリムーバブルメディア９３２に格納されたプログラムが、実行時にＲＡＭ９０６に読み込まれ、ＣＰＵ９０２によって実行される。

【０１１２】

以上、添付図面を参照しながら本発明の好適な実施形態について説明したが、本発明は係る例に限定されないことは言うまでもない。当業者であれば、特許請求の範囲に記載された範疇内において、各種の変更例又は修正例に想到し得ることは明らかであり、それらについても当然に本発明の技術的範囲に属するものと了解される。

【０１１３】

例えば、本明細書では楽曲データについてステレオデータとモノラルデータの違いを特に区別することなく説明したが、ステレオデータの一対の音声の片方ずつについて特徴パターンを生成してもよい。また、ステレオデータの一対の音声を合成した上で特徴パターンを生成してもよい。

【０１１４】

また、各フローチャートに記載された第１又は第２の実施形態に係る処理を、必ずしもフローチャートに記載された順序に沿って実行しなくてもよい。各処理ステップは、並列的あるいは個別に独立して実行される処理を含んでもよい。

【図面の簡単な説明】

【０１１５】

【図１】楽曲データ、メタデータ及びリミックスデータの関係を示す説明図である。

【図２】異なる環境で楽曲が録音される状況を説明するための説明図である。

【図３】第１の実施形態に係る楽曲処理装置の論理的構成を示すブロック図である。

【図４】第１の実施形態に係る基準平均音量の計算処理を説明するための説明図である。

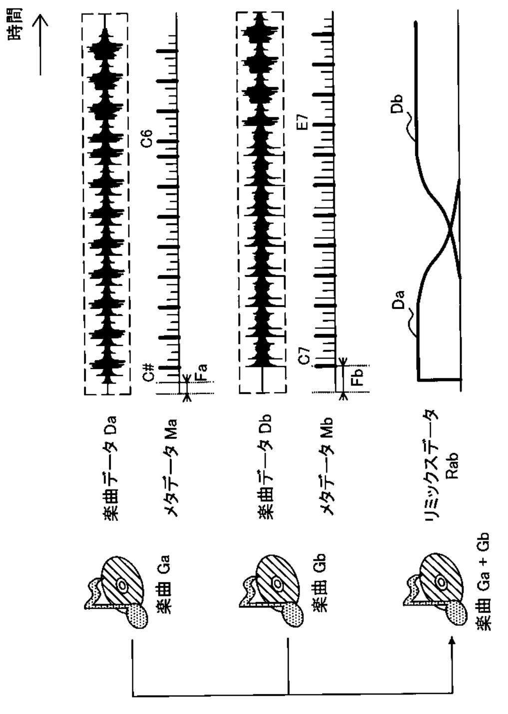
【図５】第１の実施形態に係る符号化開始位置の決定処理を説明するための説明図である。

【図６】第１の実施形態に係る符号値の決定処理を説明するための説明図である。

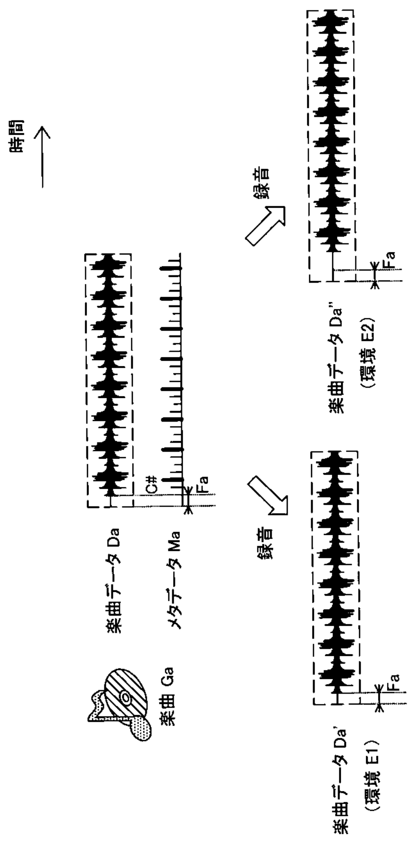
【図７】第１の実施形態に係る特徴パターン生成部により出力されるデータセットを説明するための説明図である。
【図８】第１の実施形態に係る特徴パターンの生成処理の一例を示すフローチャートである。
【図９】第１の実施形態に係る符号化開始位置の決定処理の一例を示すフローチャートである。
【図１０】第１の実施形態に係る符号値の決定処理の一例を示すフローチャートである。
【図１１】第１の実施形態に係る楽曲同定処理を説明するための説明図である。
【図１２】第１の実施形態に係る楽曲同定処理の一例を示すフローチャートである。
【図１３】第１の実施形態に係るオフセット補正処理を説明するための説明図である。
【図１４】第１の実施形態に係るオフセット補正処理の一例を示すフローチャートである。
【図１５】第２の実施形態に係る楽曲処理装置の論理的構成を示すブロック図である。
【図１６】第２の実施形態に係る符号化開始位置の決定処理を説明するための説明図である。
【図１７】第２の実施形態に係る符号値の決定処理を説明するための説明図である。
【図１８】第２の実施形態に係る特徴パターン生成部により出力されるデータセットを説明するための説明図である。
【図１９】第２の実施形態に係る特徴パターンの生成処理の一例を示すフローチャートである。
【図２０】第２の実施形態に係る楽曲同定処理を説明するための説明図である。
【図２１】第２の実施形態に係る楽曲同定処理及びオフセット補正処理の一例を示すフローチャートである。
【図２２】汎用コンピュータの構成例を示すブロック図である。
【符号の説明】
【０１１６】

１０、３０	楽曲処理装置
１２	記憶部
１４、３４	特徴パターン生成部
１６、３６	開始位置決定部
１８、３８	符号値決定部
２０、４０	楽曲同定部
２２	オフセット補正部

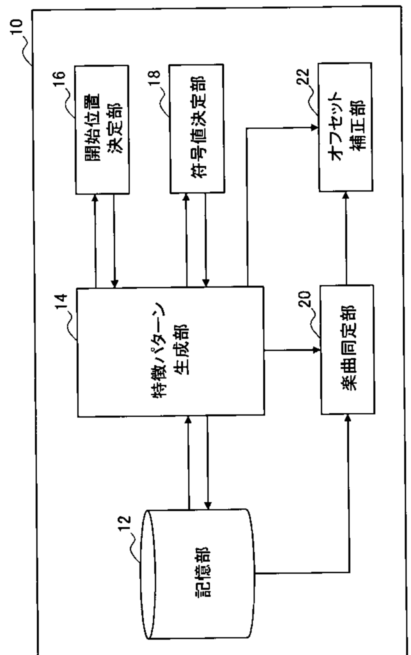
【図１】


時間
楽曲データ Da
メタデータ Ma
楽曲データ Db
メタデータ Mb
リミックスデータ Rab
楽曲 Ga
楽曲 Gb
楽曲 Ga + Gb
C#
C6
Fa
C7
E7
Fb
Da
Db

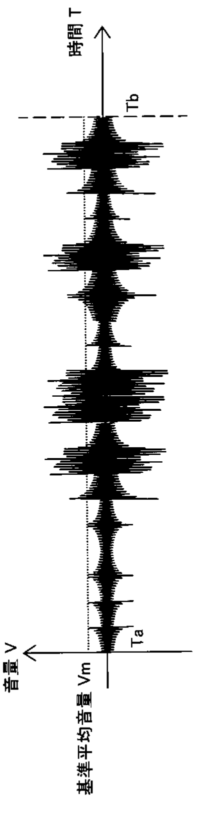
【図２】


時間
楽曲データ Da
メタデータ Ma
楽曲 Ga
C#
Fa
録音
録音
楽曲データ Da'
(環境 E1)
楽曲データ Da''
(環境 E2)

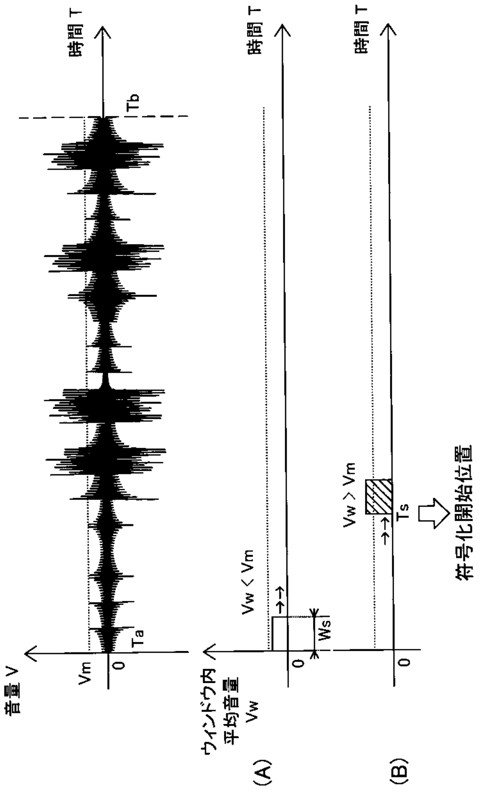
【図３】


10
16
開始位置決定部
18
符号値決定部
22
オフセット補正部
14
特徴パターン生成部
20
楽曲同定部
12
記憶部

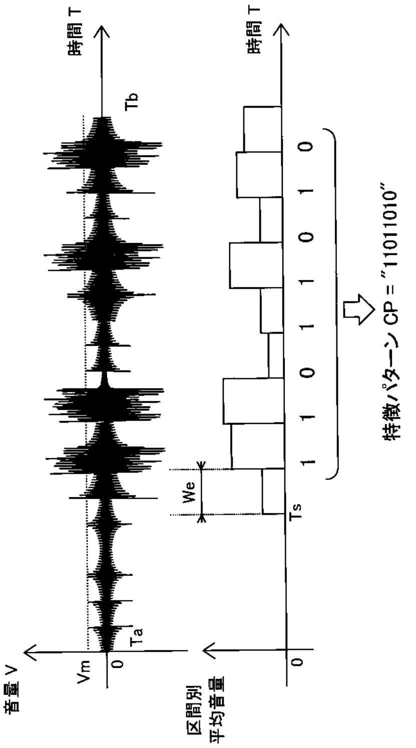
【図４】


時間 T
Tb
Ta
音量 V
基準平均音量 Vm

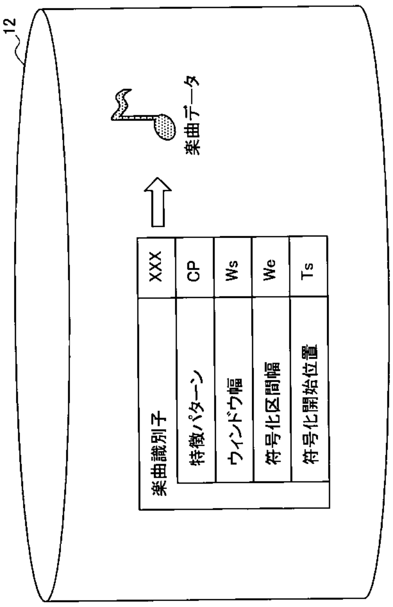
【図５】


時間 T
Tb
Ta
音量 V
Vm
0
ウィンドウ内平均音量 Vw
Vw < Vm
Ws
(A)
Vw > Vm
Ts
符号化開始位置
(B)

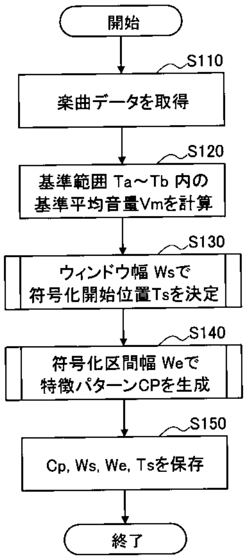
【図６】


時間 T
Tb
Ta
音量 V
Vm
0
区間別平均音量
We
Ts
1 1 0 1 1 0 1 0
特徴パターン CP = "11011010"

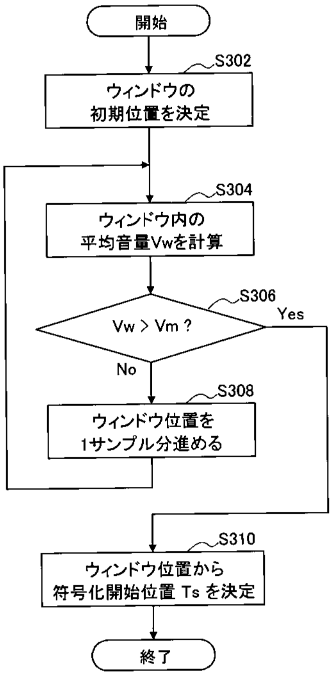
【図７】


12
楽曲データ
楽曲識別子 XXX
特徴パターン CP
ウィンドウ幅 Ws
符号化区間幅 We
符号化開始位置 Ts

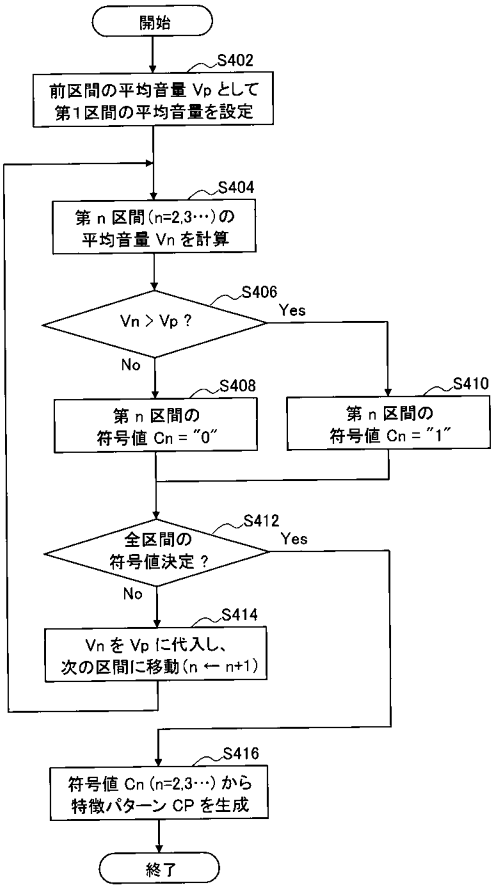
【図８】


開始
S110
楽曲データを取得
S120
基準範囲 Ta～Tb 内の基準平均音量Vmを計算
S130
ウィンドウ幅 Wsで符号化開始位置Tsを決定
S140
符号化区間幅 Weで特徴パターンCPを生成
S150
Cp, Ws, We, Tsを保存
終了

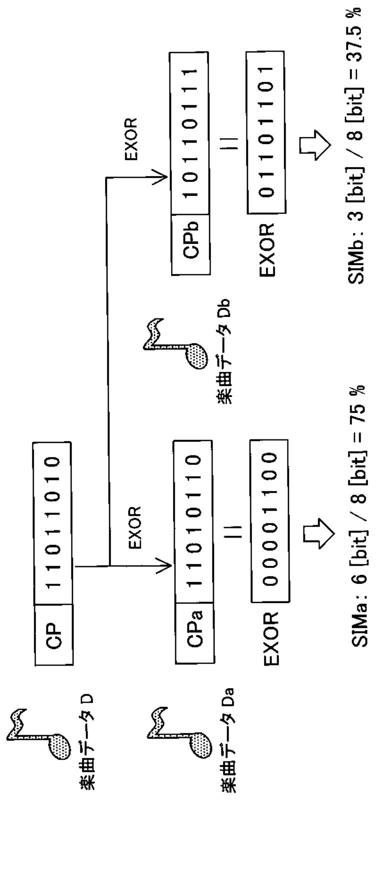
【図９】



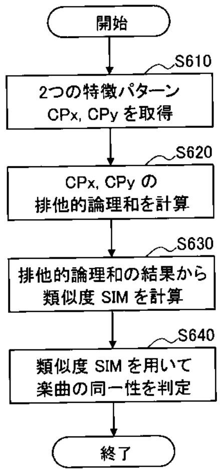
【図１０】



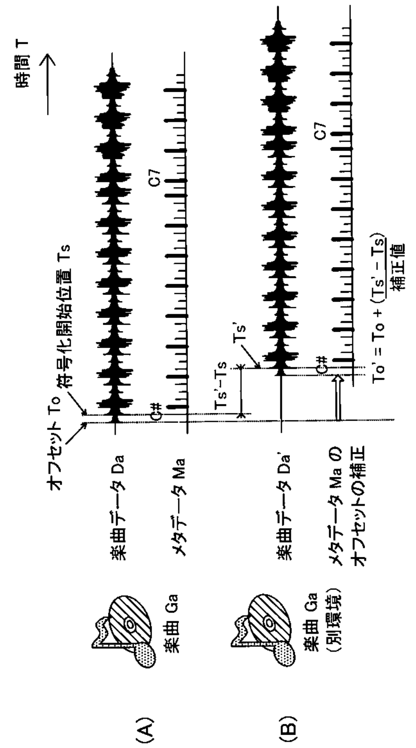
【図１１】



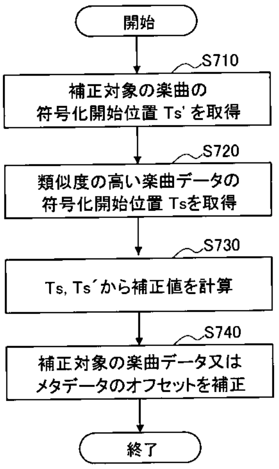
【図１２】



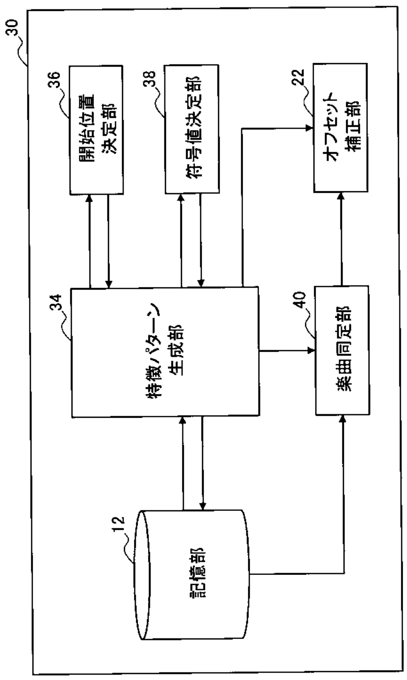
【図１３】



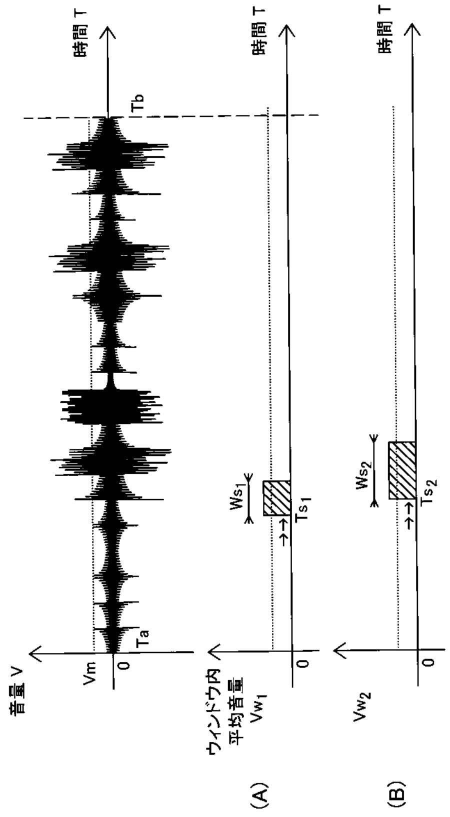
【図１４】



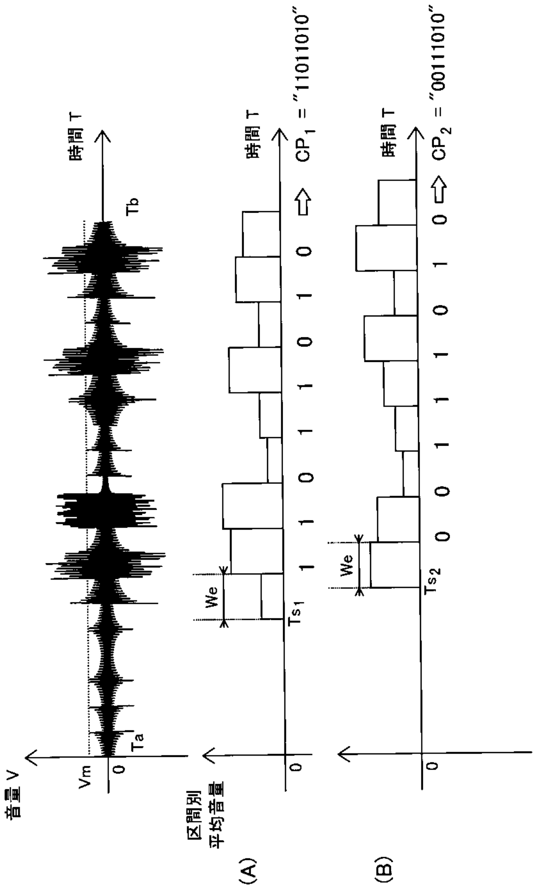
【図１５】



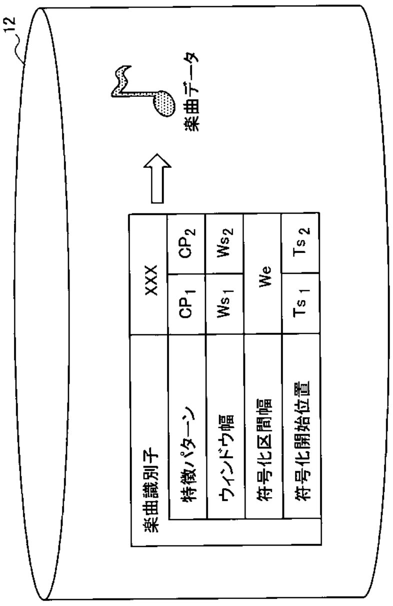
【図１６】



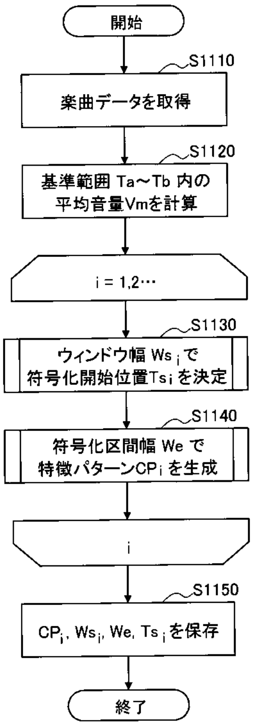
【図１７】



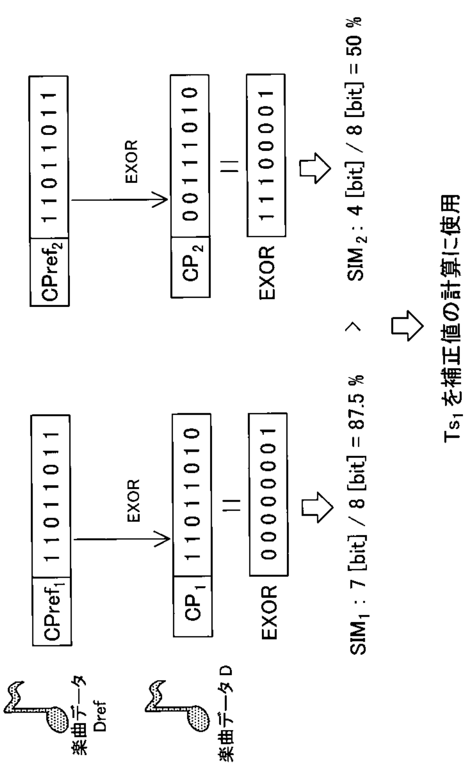
【図１８】



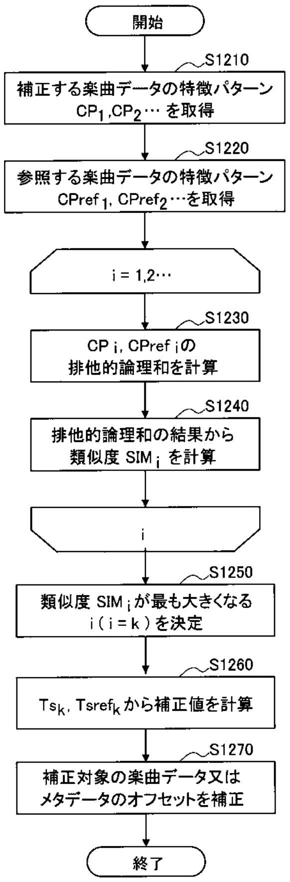
【図１９】



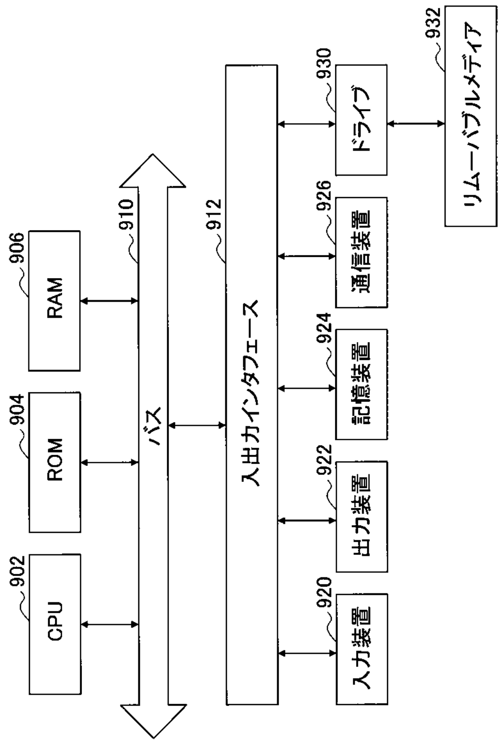
【図２０】



【図２１】


開始
S1210
補正する楽曲データの特徴パターン
CP1,CP2… を取得
S1220
参照する楽曲データの特徴パターン
CPref1, CPref2…を取得
i = 1,2…
S1230
CP i, CPref iの
排他的論理和を計算
S1240
排他的論理和の結果から
類似度 SIM i を計算
i
S1250
類似度 SIM i が最も大きくなる
i (i = k) を決定
S1260
Tsk, Tsrefk から補正値を計算
S1270
補正対象の楽曲データ又は
メタデータのオフセットを補正
終了

【図２２】


906
RAM
904
ROM
902
CPU
910
バス
912
入出力インタフェース
930
ドライブ
932
リムーバブルメディア
926
通信装置
924
記憶装置
922
出力装置
920
入力装置