

(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2012-194653

(P2012-194653A)

(43) 公開日 平成24年10月11日(2012.10.11)

(51) Int.Cl.	F I	テーマコード (参考)
G 0 6 F 17/30 (2006.01)	G 0 6 F 17/30 3 2 0 Z	5 B 0 7 5
G 1 0 H 1/00 (2006.01)	G 0 6 F 17/30 1 7 0 E	5 D 3 7 8
	G 1 0 H 1/00 Z	

審査請求 未請求 請求項の数 8 O L (全 19 頁)

(21) 出願番号	特願2011-56655 (P2011-56655)	(71) 出願人	000004237
(22) 出願日	平成23年3月15日 (2011. 3. 15)		日本電気株式会社
		(74) 代理人	100102864
			弁理士 工藤 実
		(72) 発明者	芥川 愛子
			東京都港区芝五丁目7番1号 日本電気株式会社内
		Fターム(参考)	5B075 ND14 NK06 NK39 PP10 PP28 5D378 KK01 XX43

(54) 【発明の名称】 データ処理装置、データ処理システム、データ処理方法およびプログラム

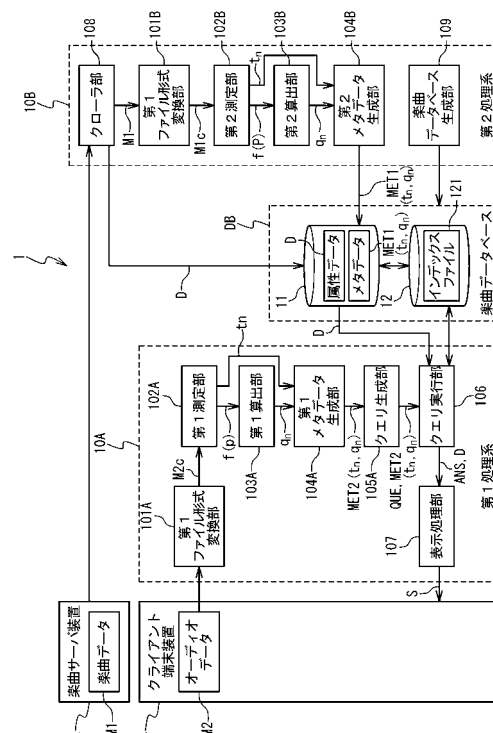
(57) 【要約】

【課題】探索に用いられるメタデータのデータ量を低減させ、楽曲を素早く探索すること。

【解決手段】第1および第2測定部102A、102Bは、楽曲のオーディオデータM1、M2の基本周波数 $f(p)$ を予め指定された時間間隔 t_n ごとに測定する。第1および第2算出部103A、103Bは、基本周波数 $f(p)$ を用いて、基準となる音の基準周波数に対する基本周波数の偏差 q_n を基本周波数に対応した音として各々算出する。第1および第2メタデータ生成部104A、104Bは、偏差を用いて、楽曲の特徴に関するメタデータMET1(t_n, q_n)、MET2(t_n, q_n)を生成する。

有する

【選択図】図3



【特許請求の範囲】

【請求項 1】

楽曲のオーディオデータの基本周波数を予め指定された時間間隔ごとに測定する測定部と、

前記測定部の測定ごとに得られた基本周波数を用いて、基準となる音の基準周波数に対する前記基本周波数の偏差を前記基本周波数に対応した音として各々算出する算出部と、

前記算出部によって各々算出された偏差を用いて、前記楽曲の特徴に関するメタデータを生成するメタデータ生成部と、

を有するデータ処理装置。

【請求項 2】

10

前記メタデータ生成部は、

各々の偏差を前記楽曲の開始からの経過時間に対応づけて、前記メタデータを生成する、

請求項 1 記載のデータ処理装置。

【請求項 3】

前記算出部は、

前記基準周波数を $f(A)$ と表記し、前記基本周波数を $f(p)$ と表記し、前記基準周波数 $f(A)$ に対する前記基本周波数 (p) の偏差を q と表記するとき、前記偏差 q を下記式を用いて算出する、

請求項 1 または 2 記載のデータ処理装置。

20

【数 1】

$$q = \log_{\sqrt[12]{2}} \frac{f(p)}{f(A)}$$

【請求項 4】

前記メタデータ生成部によって生成されたメタデータを、前記楽曲を含む複数の楽曲の属性が登録されたデータベースと照合する照合部を有し、

前記照合部は、

照合結果と共に前記楽曲の属性を出力する、

30

請求項 1 から 3 のいずれかに記載のデータ処理装置。

【請求項 5】

前記照合部は、

前記メタデータが前記データベース内のメタデータと一定の割合で一致した場合、前記メタデータが前記データベース内のメタデータと類似すると判断し、照合結果と共に前記楽曲の属性を出力する、

請求項 4 記載のデータ処理装置。

【請求項 6】

請求項 1 から 5 のいずれかに記載のデータ処理装置と、

前記データ処理装置にオーディオデータを送信する端末装置と、

40

を有するデータ処理システム。

【請求項 7】

楽曲のオーディオデータの基本周波数を予め指定された時間間隔ごとに測定する測定ステップと、

前記測定ステップでの測定ごとに得られた基本周波数を用いて、基準となる音の基準周波数に対する前記基本周波数の偏差を前記基本周波数に対応した音として各々算出する算出ステップと、

前記算出ステップで各々算出された偏差を用いて、前記楽曲の特徴に関するメタデータを生成するメタデータ生成ステップと、

を有するデータ処理方法。

50

【請求項 8】

楽曲のオーディオデータの基本周波数を予め指定された時間間隔ごとに測定する測定手順と、

前記測定手順での測定ごとに得られた基本周波数を用いて、基準となる音の基準周波数に対する前記基本周波数の偏差を前記基本周波数に対応した音として各々算出する算出手順と、

前記算出手順で各々算出された偏差を用いて、前記楽曲の特徴に関するメタデータを生成するメタデータ生成手順と、

をコンピュータに実行させるプログラム。

【発明の詳細な説明】

10

【技術分野】**【0001】**

本発明は、データ処理装置、データ処理システム、データ処理方法およびプログラムに関する。

【背景技術】**【0002】**

近年、パーソナルコンピュータ（以下「PC」という）や多機能携帯電話機などの普及により、楽曲の多くがデジタルデータ化されている。その際には、楽曲の探索や管理を容易にするため、楽曲名や作曲家名といったオーディオデータの属性が記述されたメタデータが生成されることが多い。

20

【0003】

また、ネットワークを利用した電子商取引の普及により、膨大な数の楽曲がコンテンツ提供者によって提供されている。例えば、個人ユーザがその中から所望する楽曲をPCを用いて探索する場合、楽曲名や作曲家名など、何らかのキーワードが必要である。こうしたキーワードが分からない場合、所望する楽曲を見つけ出すことが困難となる。

【0004】

この課題を解決すべく、特許文献1には、オーディオデータに高速フーリエ変換などの信号処理を所定時間ごとに施し、音程の特徴量をメタデータとして抽出することで、音（メロディ）をキーワードとして楽曲の探索を行うことができる技術が開示されている。この信号処理方法については、引用文献2に開示されている。

30

【先行技術文献】**【特許文献】****【0005】**

【特許文献1】特開2008-192102号公報

【特許文献2】特開2006-201614号公報

【発明の概要】**【発明が解決しようとする課題】****【0006】**

音をキーワードに、数ある同一異録音の楽曲の中から所望する一曲を探索する場合、探索に用いられるメタデータのデータ量が増大する。必然的に、探索対象および被探索対象のメタデータ同士を照合する際の演算量が増え、楽曲の探索に要する時間も増大する。

40

【0007】

本発明の目的は、探索に用いられるメタデータのデータ量を低減させ、楽曲を素早く探索することができる、データ処理装置、データ処理システム、データ処理方法およびプログラムを提供することである。

【課題を解決するための手段】**【0008】**

本発明のデータ処理装置は、楽曲のオーディオデータの基本周波数を予め指定された時間間隔ごとに測定する測定部と、前記測定部の測定ごとに得られた基本周波数を用いて、基準となる音の基準周波数に対する前記基本周波数の偏差を前記基本周波数に対応した音

50

として各々算出する算出部と、前記算出部によって各々算出された偏差を用いて、前記楽曲の特徴に関するメタデータを生成するメタデータ生成部と、を有する。

【0009】

本発明のデータ処理方法は、楽曲のオーディオデータの基本周波数を予め指定された時間間隔ごとに測定する測定ステップと、前記測定ステップでの測定ごとに得られた基本周波数を用いて、基準となる音の基準周波数に対する前記基本周波数の偏差を前記基本周波数に対応した音として各々算出する算出ステップと、前記算出ステップで各々算出された偏差を用いて、前記楽曲の特徴に関するメタデータを生成するメタデータ生成ステップと、を有する。

【0010】

本発明のプログラムは、楽曲のオーディオデータの基本周波数を予め指定された時間間隔ごとに測定する測定手順と、前記測定手順での測定ごとに得られた基本周波数を用いて、基準となる音の基準周波数に対する前記基本周波数の偏差を前記基本周波数に対応した音として各々算出する算出手順と、前記算出手順で各々算出された偏差を用いて、前記楽曲の特徴に関するメタデータを生成するメタデータ生成手順と、をコンピュータに実行させる。

【発明の効果】

【0011】

本発明によれば、探索に用いられるメタデータのデータ量を低減させ、楽曲を素早く探索することができる。

【図面の簡単な説明】

【0012】

【図1】図1は、本発明の実施の形態に係るデータ処理システムAの構成例を示す概略ブロック図である。

【図2】図2は、本発明の実施の形態に係るサーバ装置1のハードウェアの構成例を示すブロック図である。

【図3】図3は、図2に示すサーバ装置1の持つ機能を表したブロック図である。

【図4】図4は、本発明の実施の形態に係る算出処理を説明するための図である。

【図5】図5は、本発明の実施の形態に係るメタデータ生成処理によって得られたメタデータの構成を説明するための図である。

【図6】図6は、本発明の実施の形態に係るデータ処理システムAの動作例を示すシーケンス図である。

【発明を実施するための形態】

【0013】

以下、本発明の実施の形態を図面に関連づけて説明する。

【0014】

[データ処理システムAの構成例]

図1は、本発明の実施の形態に係るデータ処理システムAの構成例を示す概略ブロック図である。図1に示すように、データ処理システムAは、データ処理装置としてのサーバ装置1、楽曲サーバ装置2、クライアント端末装置3およびネットワーク4を有する。

【0015】

データ処理システムAは、サーバ装置1および楽曲サーバ装置2を所有するコンテンツ提供者が、ネットワーク4を介してクライアント端末装置3の所有者（以下「個人ユーザ」という）に対して個人ユーザが所望する楽曲を提供するシステムである。

【0016】

楽曲の提供に際して、データ処理システムAは、楽曲サーバ装置2に格納されている何千、何万という数の楽曲の中から個人ユーザが所望する楽曲（楽曲データM1）を探索する。その際に、楽曲名など、言葉のキーワードが分からなければ、探索が困難となる。そこで、データ処理システムAは、クライアント端末装置3に格納されているオーディオデータM2をサーバ装置1にアップロードすることで、言葉だけではではなく、音をキーワー

10

20

30

40

50

ドとして探索することができる。本実施の形態では、音をキーワードとして探索する場合を例に挙げる。

【0017】

楽曲の探索の便宜を図るため、データ処理システムAには、楽曲探索モードおよび類似探索モードがある。ここで、これら2つのモードについて説明する。

【0018】

(楽曲探索モード)

楽曲探索モードは、個人ユーザのオーディオデータM2を用いて、楽曲サーバ装置2に格納されている数ある楽曲の中から個人ユーザが所望する楽曲を探索するためのモードである。

10

【0019】

(類似探索モード)

類似探索モードは、オーディオデータM2のメロディに類似した楽曲を探索するためのモードである。

【0020】

例えば、クラシック音楽は、世界各国の演奏者により演奏され、録音されている。クラシック音楽のように、楽曲名は同一であるが、演奏者や録音の環境が異なるという、同一異録音の楽曲は無数に存在する。類似探索モードは、このような同一異録音の楽曲を探索する場合に好適である。

【0021】

20

サーバ装置1は、プログラムに従って演算処理を行うコンピュータを備えている。サーバ装置1は、ネットワーク4に接続され、楽曲サーバ装置2およびクライアント端末装置3と通信する。サーバ装置1は、クライアント端末装置3からの要求に従って、サーバ装置1の内部にて、以下の処理を行う。

【0022】

(楽曲探索処理)

サーバ装置1は、楽曲探索モードの実行を促す要求がクライアント端末装置3からあった場合、楽曲データM1をネットワーク4上の楽曲サーバ装置2の中から探索し、探索結果をクライアント端末装置3に送信する。

【0023】

30

(類似探索処理)

サーバ装置1は、類似探索モードの実行を促す要求がクライアント端末装置3からあった場合、オーディオデータM2に類似した楽曲データ(楽曲データM1とする)を探索し、探索結果をクライアント端末装置3に送信する。

【0024】

楽曲サーバ装置2は、探索対象の楽曲データM1を始め、種々の属性の楽曲データを複数格納している。楽曲サーバ装置2は、ネットワーク4に接続され、サーバ装置1およびクライアント端末装置3と通信する。

【0025】

クライアント端末装置3は、PC、デジタルオーディオ機器、PDA(Personal Data Assistance)、多機能/高機能携帯電話機などの電子機器である。クライアント端末装置3は、基本的に、以下の機能を持つ。

40

【0026】

第1に、クライアント端末装置3は、オーディオデータの取り扱いが可能である。第2に、クライアント端末装置3は、有線、無線を問わず、ネットワーク4に接続し、サーバ装置1および楽曲サーバ装置2と通信する。第3に、クライアント端末装置3は、表示画面に文字や画像を表示する。

【0027】

本実施の形態では、特に断りがない限り、PCをクライアント端末装置3の一例に挙げる。オーディオデータの取り扱い、ネットワーク4への接続、文字や画像の表示が可能な

50

電子機器であれば、クライアント端末装置 3 の構成は、特に限定されるものではない。

【0028】

以下、楽曲データ M 1 およびオーディオデータ M 2 は、デジタルオーディオデータであるとする。

【0029】

ネットワーク 4 は、例えば、通信プロトコルとしての TCP / IP (Transmission Control Protocol / Internet Protocol) を利用可能なインターネットである。

【0030】

[サーバ装置 1 のハードウェアの構成例]

10

図 2 は、本発明の実施の形態に係るサーバ装置 1 のハードウェアの構成例を示すブロック図である。図 2 に示すように、サーバ装置 1 は、CPU 10、第 1 HDD 11、第 2 HDD 12、ROM 13、RAM 14、I / O (Input / Output) ポート 15 および内部バス 16 を有する。サーバ装置 1 の各構成部は、内部バス 16 に各々接続されている。

【0031】

CPU 10 は、例えば、中央演算処理装置である。CPU 10 は、内部バス 16 を介して、ROM 13 に格納されているプログラム 131 に従って、第 1 および第 2 HDD 11、12、RAM 14 との間でデータの授受を行いながら、楽曲探索処理および類似探索処理を行う。

20

【0032】

第 1 HDD 11 は、記憶装置であって、例えば、ハードディスクである。第 2 HDD 12 も同様に、記憶装置であって、例えば、ハードディスクである。第 1 HDD 11 は、主に、楽曲探索処理および類似探索処理に用いられるメタデータを格納する。

【0033】

第 2 HDD 12 は、主に、楽曲探索処理および類似探索処理に用いられるインデックスファイルを格納する。

【0034】

ROM 13 は、不揮発性の記憶装置である。ROM 13 は、CPU 10 が楽曲探索処理および類似探索処理を実行するのに必要なプログラム 131 を格納している。

30

【0035】

RAM 14 は、例えば、DRAM (Dynamic Random Access Memory) ある。RAM 14 は、CPU 10 が演算処理のときに使用する一時的なデータを格納する。

【0036】

I / O ポート 15 は、外部機器を接続するためのポートであって、内部バス 16 を介して外部機器と CPU 10 との間のデータの授受を行う。I / O ポート 15 に接続可能な外部機器には、例えば、ネットワークカードやキーボードがある。本実施の形態では、サーバ装置 1 は、ネットワークカードが接続された I / O ポート 15 を介して、ネットワーク 4 上の楽曲サーバ装置 2 およびクライアント端末装置 3 と通信する。

40

【0037】

CPU 10 は、ROM 13 からプログラム 131 を読み出すと、第 1 HDD 11、第 2 HDD 12 および RAM 14 にアクセスしながら、プログラム 131 の処理手順に従って処理を行う。これにより、サーバ装置 1 が持つ各機能が実現される (図 3 参照) 。

【0038】

なお、図 2 に示すハードウェアの構成は一例であって、種々の改変が可能である。例えば、プログラム 131 を ROM 13、第 1 HDD 11 または第 2 HDD 12 に格納することができる。この他、例えば、1 個のハードディスクに、本実施の形態で用いられる各種データを格納することができる。

【0039】

50

[サーバ装置 1 の機能ブロック図]

図 3 は、図 2 に示すサーバ装置 1 の機能を表したブロック図である。図 3 に示すように、サーバ装置 1 は、第 1 処理系 10 A、第 2 処理系 10 B および楽曲データベース DB を有する。第 1 および第 2 処理系 10 A、10 B の各処理は、CPU 10 (図 2 参照) によってソフトウェア的に実行される。

【 0040 】

先ず、第 1 および第 2 処理系 10 A、10 B の概要について説明する。

【 0041 】

(第 1 処理系 10 A)

第 1 処理系 10 A は、第 1 ファイル形式変換部 101 A、第 1 測定部 102 A、第 1 算出部 103 A、第 1 メタデータ生成部 104 A、クエリ生成部 105 A、クエリ実行部 106 および表示処理部 107 を有する。なお、クエリ生成部 105 A およびクエリ実行部 106 は、照合部の一例である。 10

【 0042 】

第 1 処理系 10 A の目的は、楽曲探索処理および類似探索処理を行うため、個人ユーザのオーディオデータ M2 を、第 2 処理系 10 B によって作成された楽曲データベース DB と照合することである。第 1 処理系 10 A の基本的な処理は、以下の通りである。

【 0043 】

(測定処理)

第 1 に、第 1 処理系 10 A は、オーディオデータ M2 をクライアント端末装置 3 から入力し、オーディオデータ M2 の基本周波数 $f(p)$ を予め指定された時間間隔ごとに測定する。ここでいう、基本周波数とは、ある音 (例えば、音名 (以下、米名表記) でいう「A」) の音高 (音の高さ) を決定づける物理量であって、その音が持つ周波数成分の内 20
で最も低い周波数である。本実施の形態では、基本周波数 $f(p)$ を測定する時間間隔は、およそ 100 ms である。

【 0044 】

(算出処理)

第 2 に、第 1 処理系 10 A は、基準となる音 (基準音) の基準周波数 $f(A)$ に対する基本周波数 $f(p)$ の偏差を半音階単位で算出し、この偏差に対応した音を算出する。換言すれば、第 1 処理系 10 A は、基準周波数 $f(A)$ に対する基本周波数 $f(p)$ のズレ 30
に対応した音を算出する。ここでいう、基準音は、音名 (以下、米名) で表すと、「A」に対応する音である。その音の基準周波数 $f(A)$ は、望ましくは 440 Hz である。

【 0045 】

楽曲が基本周波数成分と種々の高調波成分で構成されていることは周知である。測定処理により基本周波数 $f(p)$ を測定し、算出処理により基本周波数 $f(p)$ が基準周波数 $f(A)$ からどの程度ズレているかを算出することで、基本周波数 $f(p)$ を測定した時点での音 (例えば、音名でいう「B」) が分かる。その詳細については後述する。

【 0046 】

(メタデータ生成処理)

第 3 に、第 1 処理系 10 A は、算出処理によって得た音を基に、オーディオデータ M2 40
の特徴、具体的には、メロディを表すメタデータを生成する。

【 0047 】

(楽曲探索処理 / 類似探索処理)

第 4 に、第 1 処理系 10 A は、メタデータを用いて、楽曲データベース DB から楽曲データ M1 の検索または類似探索をクライアント端末装置 3 からの処理要求に従って行い、その結果をクライアント端末装置 3 に送信する。

【 0048 】

(第 2 処理系 10 B)

第 2 処理系 10 B は、第 2 ファイル形式変換部 101 B、第 2 測定部 102 B、第 2 算出部 103 B、第 2 メタデータ生成部 104 B、クローラ部 108 および楽曲データベー 50

ス生成部 109 を有する。

【0049】

第2処理系10Bの目的は、楽曲データベースDBを作成することである。楽曲データベースDBは、楽曲探索処理および類似探索処理を行うため、楽曲データM1を始めとする、楽曲サーバ装置2内の楽曲データの情報を記録したものである。第2処理系10Bの基本的な処理は、以下の通りである。

【0050】

(収集処理)

第1に、第2処理系10Bは、楽曲サーバ装置2のディレクトリを巡回し、楽曲サーバ装置2に格納されている種々の楽曲データを収集する。

10

【0051】

(データベース作成処理)

第2に、第2処理系10Bは、測定処理、算出処理およびメタデータ生成処理を行って、楽曲データベースDBを作成する。

【0052】

以下、本実施の形態の特徴である、算出処理およびメタデータ生成処理について順を追って説明する。

【0053】

[算出処理における音の算出方法]

図4は、本発明の実施の形態に係る算出処理を説明するための図である。本実施の形態では、音階の表現方法に半音階を用いる。十二平均律における半音は、1オクターブを12等分したものである。1オクターブは、米名の表記を用いると、音高の低い方から順に、「C」、「D (フラット)」、「D」、「E (フラット)」、「E」、「F」、「F (シャープ)」、「G」、「A (フラット)」、「A」、「B (フラット)」、「B」の12音からなる。

20

【0054】

互いに隣接する2つの半音の周波数比は、(1)式で表される。

【0055】

【数1】

$$1:\sqrt[12]{2} \quad \dots(1)$$

30

【0056】

本実施の形態では、上述したように、「A」を基準音とし、その基準周波数は、 $f(A) = 440 \text{ Hz}$ である。例えば、音高が基準音「A」より半音高い「B $\flat$ 」音の基本周波数 $f(B\flat)$ は、(1)式の関係により、(2)式で表される。

【0057】

【数2】

$$f(A):f(B\flat)=1:\sqrt[12]{2} \quad \dots(2)$$

40

【0058】

(2)式より、周波数 $f(B\flat)$ は、(3)式で表される。

【数3】

$$f(B\flat)=440 \times \sqrt[12]{2} \div 466.16 \text{ Hz} \quad \dots(3)$$

【0059】

例えば、音高が「B $\flat$ 」音より半音高い「B」音の基本周波数 $f(B)$ は、基本周波数

50

$f(B)$ の算出と同様の議論により、(4) 式で表される。

【0060】

【数4】

$$f(Bb):f(B)=1:\sqrt[12]{2} \quad \dots(4)$$

【0061】

すなわち、周波数 $f(B)$ は、(3) 式を用いると、(5) 式で表される。

【0062】

【数5】

$$\begin{aligned} f(B) &= f(Bb) \times \sqrt[12]{2} \\ &= 440 \times \sqrt[12]{2} \times \sqrt[12]{2} \\ &= 440 \times \left(\sqrt[12]{2}\right)^2 \\ &\doteq 493.88\text{Hz} \quad \dots(5) \end{aligned}$$

10

【0063】

20

図4には、「C」音からそれよりも1オクターブ高い「C」音までの各音に対応した基本周波数 $f(p)$ が示されている。なお、変数 p は、「C」から「B」のいずれかの音を表す。音が高くなるほど、基本周波数 $f(p)$ も高くなる。

【0064】

基準音「A」の基準周波数 $f(A)$ と、基準音「A」より高く、かつ、これから2個の半音だけズレた「B」音の基本周波数 $f(B)$ との比(周波数比)は、(6) 式で表される。

【0065】

【数6】

$$f(A):f(B)=1:\left(\sqrt[12]{2}\right)^2 \quad \dots(6)$$

30

【0066】

図4には、基準音「A」を1としたとき、基準周波数 $f(A)$ と各音の基本周波数 $f(p)$ との比が示されている。

【0067】

(6) 式を参照すると、基準音「A」の基準周波数 $f(A)$ と、基準音「A」から q 個の半音だけズレた p 音の基本周波数 $f(p)$ との比を、一般化された(7) 式で表すことができる。

【0068】

【数7】

$$f(A):f(P)=1:\left(\sqrt[12]{2}\right)^q \quad \dots(7)$$

40

【0069】

(7) 式において、変数 q は、整数をとる。以下、変数 q を「ズレ個数」と呼ぶ。(7) 式を変形すると、(8) 式の関係式が得られる。

【0070】

【数 8】

$$\left(\sqrt[12]{2}\right)^q = \frac{f(p)}{f(A)} \quad \dots(8)$$

【0071】

(8)式において、底を $2^{1/12}$ とする対数をとると、ズレ個数 q は、(9)式で表される。

【0072】

【数 9】

$$q = \log_{\sqrt[12]{2}} \frac{f(P)}{f(A)} \quad \dots(9)$$

【0073】

ただし、ズレ個数 q は、整数となるように、四捨五入された値である。(9)式において、基準音「A」の基準周波数 $f(A)$ が既知であるので、基本周波数 $f(p)$ が分かれば、ズレ個数 q が求まる。したがって、半音で基準音「A」から何個ズレているかが分かる。

【0074】

図4には、ズレ個数 q と各音との対応関係が示されている。測定処理にて測定された基本周波数が $f(p) = 466.16 \text{ Hz}$ であった場合、ズレ個数は $q = 1$ となる。この演算結果は、基準音「A」から半音でズレ個数 $q = 1$ だけズレた「B」を表している。なお、音名は、オクターブに左右されずに、一定の表記(「C」、「D」、...「B」)がなされるが、本実施の形態では、音高そのものを表すことができる。

【0075】

このように、基本周波数 $f(p)$ が基準周波数 $f(A)$ からどの程度ズレているかを算出することで、測定された基本周波数 $f(p)$ がどの音に対応するかということが分かる。

【0076】

[メタデータの構成例]

図5は、本発明の実施の形態に係るメタデータ生成処理によって得られたメタデータの構成を説明するための図である。第1処理系10Aによる測定処理では、楽曲の開始から、およそ100msごとにオーディオデータM2の基本周波数 $f(p)$ が測定される。第2処理系10Bによる測定処理は、第1処理系10Aのものと同様であるので、ここでは、第1処理系10Aの処理を例に挙げて説明する。

【0077】

楽曲は和音で構成されることが多いため、一度に複数の基本周波数 $f(p)$ が測定されることがある。図5に示す例では、楽曲の開始から100ms経過した時間 t_1 において、2個の基本周波数 $f(p_1) = 293.7 \text{ Hz}$ 、 $f(p_2) = 349.2 \text{ Hz}$ が各々測定されている。

【0078】

算出処理の結果、時間 t_1 において、基本周波数 $f(p_1)$ におけるズレ個数が $q_1 = -7$ であったとすると、ズレ個数 q_1 に対応する音は「D」であることが分かる。同様に、基本周波数 $f(p_2)$ におけるズレ個数 q_2 に対応する音が「F」であることが分かる。

【0079】

メタデータ生成処理において、第1処理系10Aは、基本周波数 $f(p)$ の経過時間 t_n と共に、ズレ個数 q_n が時系列順に記述されたメタデータを生成する。メタデータは、経過時間 t_n およびズレ個数 q_n をパラメータとし、下記に示す構文1で記述される。

【 0 0 8 0 】

(構 文 1)

(t_n) , q_1 , q_2 , . . . , q_n , ,

【 0 0 8 1 】

変数 n は、正の整数 (1 , 2 , 3 , . . .) をとる。経過時間 t_n は、楽曲開始からの経過時間であって、括弧「 () 」内に記述される。経過時間 t_n の記述の後には、「 , (コンマ)」を於いてズレ個数 q_1 が記述される。ズレ個数 q_{n-1} の記述の後にズレ個数 q_n を記述する場合、前者の記述の後に「 , 」を於いて後者のズレ個数 q_n が記述される。最後のズレ個数 q_n の記述の後には、メタデータの終了を表す「 , (コンマ)」が記述される。

10

【 0 0 8 2 】

例えば、経過時間 $t_1 = 100 \text{ ms}$ における測定では、ズレ個数 $q_1 = -7$, $q_2 = -4$ であるので、「 (1 0 0) , - 7 , - 4 , , 」というメタデータが生成される。予めズレ個数 q_n を音名 (具体的には音高) に対応づけておけば、メタデータを参照するだけで、経過時間 $t_n = 100 \text{ ms}$ における楽曲データが「 D 」音および「 F 」音によって構成されていることが分かる。

【 0 0 8 3 】

15分程度の楽曲では、経過時間 t_n ごとに得られるメタデータのバイト数は、数バイトから20バイト程度であり、平均すると15バイト程度である。100msごとに15バイトのメタデータが得られると仮定すると、この楽曲から得られるメタデータの総バイト数は、およそ $15 \text{ バイト} \times (900000 \text{ ms} / 100 \text{ ms}) = 135 \text{ KB}$ となる。

20

【 0 0 8 4 】

上述したように、経過時間 t_n およびズレ個数 q_n を用いてオーディオデータ M2 のメロディを表すので、15分程度の楽曲であっても、メタデータのデータ量は、わずか135KB程度でしかない。サーバ装置1は、楽曲の特徴を的確に表しつつ、軽量のメタデータを生成することを可能にしている。

【 0 0 8 5 】

(第1処理系10Aの各構成部)

次に、図3を参照しながら、第1処理系10Aの各構成部について説明する。

【 0 0 8 6 】

第1ファイル形式変換部101Aは、オーディオデータM2をネットワーク4 (不図示) を介してクライアント端末装置3から時系列で入力し、これを所定のファイル形式に変換し、変換したオーディオデータM2cを第1測定部102Aに出力する。これは、処理対象のファイル形式を統一するために行われる。本実施の形態では、例えば、オーディオデータM2がMP3 (M P E G A u d i o L a y e r 3) のファイル形式で圧縮されている場合、第1ファイル形式変換部101Aは、このオーディオデータM2をMP4のファイル形式に変換する。

30

【 0 0 8 7 】

第1測定部102Aは、オーディオデータM2cを第1ファイル形式変換部101Aから時系列で取り込み、およそ100msごとに測定処理を行い、得られた基本周波数 $f(p)$ を第1算出部103Aに出力する。測定処理の際に、第1測定部102Aは、例えば、高速フーリエ変換 (F F T) を用いて、オーディオデータM2cを周波数スペクトルデータに変換し、およそ100msごとに基本周波数 $f(p)$ を算出する。更に、第1測定部102Aは、基本周波数 $f(p)$ を測定したときの測定時間を楽曲開始からの経過時間 t_n として第1メタデータ生成部104Aに出力する。なお、基本周波数 $f(p)$ の算出には、上述の他、ハーモニッククラスタリング法、くし形フィルタ法などを用いることができる。

40

【 0 0 8 8 】

第1算出部103Aは、基本周波数 $f(p)$ を第1測定部102Aから入力し、算出処理を行い、算出したズレ個数 q_n を第1メタデータ生成部104Aに出力する。算出処理

50

の際に、第1算出部103Aは、基準音「A」の基準周波数 $f(A)$ と、第1測定部102Aによって測定された基本周波数 $f(p)$ とを(9)式に代入し、整数でズレ個数 q_n を算出する。

【0089】

第1メタデータ生成部104Aは、経過時間 t_n を第1測定部102Aから入力すると共に、ズレ個数 q_n を第1算出部103Aから入力し、およそ100msごとにメタデータ生成処理を行い、生成したメタデータMET2(t_n, q_n)をクエリ生成部105Aに出力する。メタデータ生成処理の際に、第1メタデータ生成部104Aは、構文1に示すように、「(t_n), $q_1, q_2, \dots, q_n, ,$ 」という記述のメタデータを生成する。

10

【0090】

本実施の形態では、データの流れを明確にするため、第1メタデータ生成部104Aに第1測定部102Aから経過時間 t_n が入力されるものとしているが、第1メタデータ生成部104Aが経過時間 t_n を直接把握しても差し支えはない。このことは、第2処理系10Bについても同様である。

【0091】

クエリ生成部105Aは、メタデータMET2(t_n, q_n)を第1メタデータ生成部104Aから入力し、メタデータMET2(t_n, q_n)を用いてクエリを生成し、生成したクエリQUEをクエリ実行部106に出力する。その際に、クエリ生成部105Aは、メタデータMET2(t_n, q_n)もクエリ実行部106に出力する。

20

【0092】

クエリQUEは、楽曲探索処理または類似探索処理を実行するために必要なデータであって、楽曲データベースDBに対する処理要求が記述されている。処理要求には、サーバ装置1がクライアント端末装置3から受けた、楽曲探索モードまたは類似探索モードの実行がある。

【0093】

クエリ実行部106は、クエリQUEをクエリ生成部105Aから入力し、クエリQUEに従って楽曲データベースDBにアクセスしながら、以下の処理を行う。

【0094】

(楽曲探索モードの実行時)

30

楽曲探索モードの実行時には、クエリ実行部106は、メタデータMET2(t_n, q_n)を第2HDD12に格納されているインデックスファイル121と照合し、照合結果を楽曲探索結果ANSとして表示処理部107に出力する。

【0095】

インデックスファイル121は、例えば、図5に図示するように、第2処理系10Bによって生成されたメタデータMET1(t_n, q_n)を時系列順に記録したものである。なお、インデックスファイル121には、楽曲データM1以外のもののメタデータも記録されている。

【0096】

照合の際に、クエリ実行部106は、オーディオデータM2のメタデータMET2(t_n, q_n)をインデックスファイル121に記述されているメタデータMET1(t_n, q_n)と比較する。

40

【0097】

比較の結果、両者が一致する場合、クエリ実行部106は、楽曲データM1が楽曲サーバ装置2にあると判断し、楽曲データM1が見つかった旨を楽曲探索結果ANSとして、第1HDD11に格納されている属性データDと共に表示処理部107に出力する。一方、両者が不一致する場合、クエリ実行部106は、楽曲データM1が楽曲サーバ装置2にないと判断し、楽曲データM1が見つからなかった旨を楽曲探索結果ANSとして表示処理部107に出力する。

【0098】

50

(類似探索モードの実行時)

類似探索モードの実行時には、クエリ実行部106は、楽曲探索モードの場合と同様に、メタデータMET2(t_n , q_n)をインデックスファイル121と照合し、照合結果を探索結果ANSとして表示処理部107に出力する。ただし、以下の点が楽曲探索モードの場合と異なる。

【0099】

照合の際に、クエリ実行部106は、オーディオデータM2のメタデータMET2(t_n , q_n)がインデックスファイル121に記述されているメタデータMET1(t_n , q_n)と類似するか否かを判断する。本実施の形態では、前者のメタデータMET2(t_n , q_n)が後者のメタデータMET1(t_n , q_n)と一定の割合(例えば、80パーセント)一致した場合、クエリ実行部106は、両者が類似すると判断する。この類似の割合を表す割合は、サーバ装置1の所有者または個人ユーザが好適に設定すればよい。

10

【0100】

比較の結果、両者が類似する場合、クエリ実行部106は、オーディオデータM2が楽曲データM1に類似した楽曲であると推定し、その旨を類似探索結果ANSとして、属性データDと共に表示処理部107に出力する。一方、両者が類似しない場合、クエリ実行部106は、オーディオデータM2に類似する楽曲が楽曲サーバ装置2にはないと判断し、その旨を類似探索結果ANSとして表示処理部107に出力する。

【0101】

表示処理部107は、楽曲探索結果ANSまたは類似探索結果ANSと共に、属性データDをクエリ実行部106から入力し、両者をクライアント端末装置3に表示するための表示処理を行い、その処理結果Rをクライアント端末装置3にネットワーク4を介して送信する。表示処理は、例えば、表示画面のレイアウトなどを決める処理である。

20

【0102】

いずれのモードにおいても、探索対象が見つかった場合、クライアント端末装置3は、その旨と共に楽曲データM1の属性データDを表示画面に表示する。

【0103】

(第2処理系10Bの各構成部)

図3を参照しながら、第2処理系10Bの各構成部について説明する。

【0104】

クローラ部108は、収集処理として、ネットワーク4上にある楽曲サーバ装置2のディレクトリを巡回し、楽曲サーバ装置2に格納されているコンテンツデータのキャッシュデータを収集する。キャッシュデータには、楽曲データM1を始めとする種々の楽曲データやその属性データがある。属性データは、上述したように、楽曲名などのデータである。クローラ部108は、キャッシュデータの内、楽曲データM1を第2ファイル形式変換部101Bに出力し、属性データDを第1HDD11に出力する。

30

【0105】

第2ファイル形式変換部101Bは、楽曲データM1をクローラ部108から入力し、第1ファイル形式変換部101Aと同様の処理を行い、所定のファイル形式に変換した楽曲データM1cを第2測定部102Bに出力する。

40

【0106】

第2測定部102Bは、楽曲データM1cを第2ファイル形式変換部101Bから時系列で取り込み、第1測定部102Aと同様の測定処理を行い、この処理によって得られた基本周波数 $f(p)$ を第2算出部103Bに出力し、経過時間 t_n を第2メタデータ生成部104Bに出力する。

【0107】

第2算出部103Bは、基本周波数 $f(p)$ を第2測定部102Bから入力し、第1算出部103Aと同様の算出処理を行い、算出したズレ個数 q_n を第2メタデータ生成部104Bに出力する。

【0108】

50

第2メタデータ生成部104Bは、経過時間 t_n を第2測定部102Bから入力すると共に、ズレ個数 q_n を第2算出部103Bから入力し、第1メタデータ生成部104Aと同様のメタデータ生成処理を行い、生成したメタデータMET1(t_n, q_n)を第1HDD11に出力する。

【0109】

楽曲データベース生成部109は、楽曲データベースDBにアクセスし、第1HDD11に格納されたメタデータMET1(t_n, q_n)を用いて、インデックスファイル121を作成する。

【0110】

(楽曲データベースDB)

楽曲データベースDBについて説明する。楽曲データベースDBは、楽曲サーバ装置2にある種々の楽曲データと、これらの楽曲データのメタデータを基に作成されたインデックスファイル121とによって構成されている。

【0111】

第1HDD11は、第2メタデータ生成部104Bから入力されたメタデータMET1(t_n, q_n)と、クローラ部108から入力された属性データDとを格納している。第2HDD12は、インデックスファイル121を格納している。

【0112】

クローラ部108が収集処理に失敗した場合など、何らかの理由により、第1HDD11にクローラ部108から属性データDが入力されなかった場合、例えば、サーバ装置1の所有者による手動操作で、属性データDを第1HDD11に直接入力することができる。更に、手動操作で、属性データDを変更することができる。

【0113】

[データ処理システムAの動作例]

図6は、本発明の実施の形態に係るデータ処理システムAの動作例を示すシーケンス図である。

【0114】

ここでは、クライアント端末装置3の所有者である個人ユーザがネットワーク4上の楽曲サーバ装置2の中からオーディオデータM2と同じ楽曲を探索する場合を例に挙げる。なお、類似探索処理であっても、データ処理システムAの動作は同じである。

【0115】

サーバ装置1は、楽曲探索モードの実行を促す要求がクライアント端末装置3からあった場合、その要求を許可すると、以下の処理を行う。先ず、ステップS1からS5に示す、第1処理系10Aによる処理の流れについて説明する。

【0116】

第1ファイル形式変換部101Aは、オーディオデータM2をクライアント端末装置3から時系列で入力すると、これを所定のファイル形式に変換する(ステップS1)。

【0117】

ファイル形式の変換後、第1測定部102Aは、オーディオデータM2cにおける楽曲の開始部分を検知する。そして、第1測定部102Aは、およそ100msごとに測定処理を行い、基本周波数 $f(p)$ を得る(ステップS2)。

【0118】

測定処理後、第1算出部103Aは、基本周波数 $f(p)$ を用いて算出処理を行い、ズレ個数 q_n を得る(ステップS3)。

【0119】

算出処理後、第1メタデータ生成部104Aは、楽曲開始からの経過時間 t_n と、ズレ個数 q_n とを用いて、およそ100msごとにメタデータ生成処理を行い、メタデータMET2(t_n, q_n)を得る(ステップS4)。

【0120】

メタデータ生成処理、クエリ生成部105Aは、メタデータMET2(t_n, q_n)を

10

20

30

40

50

用いてクエリを生成し、クエリ Q U E を得る (ステップ S 5)。

【0121】

次に、ステップ S 6 から S 11 に示す、第2処理系 10B の処理の流れについて説明する。

【0122】

クローラ部 108 は、楽曲サーバ装置 2 に格納されているコンテンツデータのキャッシュデータを収集し、楽曲データ M 1 および属性データ D を得る (ステップ S 6)。

【0123】

収集処理後、第2ファイル形式変換部 101B は、クローラ部 108 から時系列で入力した楽曲データ M 1 を所定のファイル形式に変換する (ステップ S 7)。

10

【0124】

ファイル形式の変換後、第2測定部 102B は、測定処理を行い、基本周波数 $f(p)$ を得る (ステップ S 8)。

【0125】

測定処理後、第2算出部 103B は、基本周波数 $f(p)$ を用いて算出処理を行い、ズレ個数 q_n を得る (ステップ S 9)。

【0126】

算出処理後、第2メタデータ生成部 104B は、楽曲開始からの経過時間 t_n と、ズレ個数 q_n とを用いて、およそ 100ms ごとにメタデータ生成処理を行い、メタデータ M E T 1 (t_n, q_n) を得る (ステップ S 10)。

20

【0127】

メタデータ M E T 1 (t_n, q_n) が生成された後、楽曲データベース生成部 109 は、メタデータ M E T 1 (t_n, q_n) を用いて、インデックスファイル 121 を作成する (ステップ S 11)。

【0128】

上述したステップ S 6 から S 11 の処理によって、楽曲データベース D B が作成される。この一連の処理は、ステップ S 12 の処理が開始される前に完了していることが望ましい。また、この一連の処理は、定期的 (例えば、24時間ごと) に実行されればよい。

【0129】

次に、ステップ S 12 および S 13 に示す、第1処理系 10A による処理の流れについて説明する。

30

【0130】

クエリ実行部 106 は、オーディオデータ M 2 のメタデータ M E T 2 (t_n, q_n) をインデックスファイル 121 と照合する (ステップ S 12)。

【0131】

表示処理部 107 は、探索結果を得ると、これを基に表示処理を行い、その処理結果 R をクライアント端末装置 3 に送信する (ステップ S 13)。

【0132】

クライアント端末装置 3 は、処理結果 R をサーバ装置 1 から受けると、楽曲データ M 1 が見つかったか否かを表す探索結果を表示画面に表示する。

40

【0133】

なお、ステップ S 1 から S 13 の処理は、プログラム 131 (図2参照) に処理手順として記述されている。

【0134】

以上述べたように、本実施の形態によれば、サーバ装置 1 は、経過時間 t_n およびズレ個数 q_n を用いてメタデータを生成するので、以下の顕著な効果を得ることができる。

【0135】

メタデータは、上述したように、15分程度の楽曲であっても、135KB程度であり、非常に軽量である。したがって、クエリ実行部 106 がメタデータ同士を照合する際の演算量を著しく低減させることができる。このことは、データベースの使用容量の削減や

50

、楽曲探索処理および類似探索処理に要する時間の短縮につながる。

【0136】

更に、本実施の形態によれば、以下の顕著な効果を得ることができる。

【0137】

言語（日本語、英語など）に左右されず、的確に探索することができる。クラシック音楽のように、同一異録音の楽曲は、楽曲名などの属性が楽曲間で同じになることが多い上、多種多様な言語で記述される。そもそも、楽曲名などの属性の記述方式が統一されているわけではなく、第三者の手によって属性を変更することもできる。したがって、楽曲名などの言葉のキーワードを用いて、同一異録音の楽曲の中から一曲を探索することが難しい。しかしながら、本実施の形態では、メタデータに言葉のキーワードを記述しないので、メロディだけで探索を行うことができる。

10

【0138】

探索対象が見つかった場合、探索結果と共に探索対象の属性がクライアント端末装置3に送信される。したがって、個人ユーザは、探索対象の楽曲名などが分からなくても、探索対象が見つければ、その楽曲名などを直ちに知ることができる。

【0139】

算出処理において、基本周波数 $f(p)$ に着目し、これが基準周波数 $f(A)$ からのズレを算出するため、楽曲の演奏時における楽器の構成や、楽曲の音質に左右されにくく、楽曲の特徴を的確に表したメタデータを生成することができる。

【0140】

20

この他、データ処理システムAを用いれば、属性データが第三者の手によって故意に書き換えられ、他のサーバ等にアップロードされたような違法な楽曲の探索も可能となる。

【0141】

（実施の形態の変形例）

実施の形態の変形例について説明する。本変形例は、音程を「C」音から「E」音に一定にずらすなど、キー変更がある楽曲の探索に関する。以下、実施の形態と異なる点について説明する。

【0142】

類似探索モードにおいて、クエリ実行部106は、メタデータMET2(t_n, q_n)がメタデータMET1(t_n, q_n)に類似するか否かを判断する際に、以下のキー変更楽曲探索処理を行う。

30

【0143】

具体的には、クエリ実行部106は、すべての経過時間 t_n において、両者のズレ個数 q_n 同士が一定個数 q_w ずれている場合に、オーディオデータM2と楽曲データM1間でキー変更があると判断する。一定個数 q_w は、例えば、 $q_w = 3$ など、好適に設定することができる。クエリ実行部106は、キー変更がある旨を類似探索結果ANSとして、属性データDと共に表示処理部107に出力する。なお、両者のズレ個数 q_n 同士が一定個数 q_w ずれていない場合の処理は、実施の形態と同様である。

【0144】

以上述べたキー変更楽曲探索処理をキー変更楽曲探索モードとして、楽曲探索モードおよび類似探索モードに加えることができる。

40

【0145】

本変形例により、楽器を変えて演奏した楽曲や原曲の探索を容易に行うことができる。無論、実施の形態における効果を得ることができる。

【0146】

本発明は、その要旨を逸脱しない範囲内で種々の改変が可能である。例えば、サーバ装置1および楽曲サーバ装置2を家庭内LAN(Local Area Network)に接続し、クライアント端末装置3を用いて、家庭内で楽曲探索モード、類似探索モードおよびキー変更楽曲探索モードを実行することができる。

【0147】

50

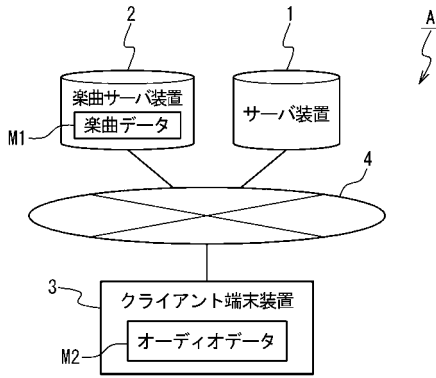
例えば、楽曲サーバ装置 2 が持つ機能をサーバ装置 1 に組み込み、両者を 1 つのサーバ装置とすることができる。

【符号の説明】

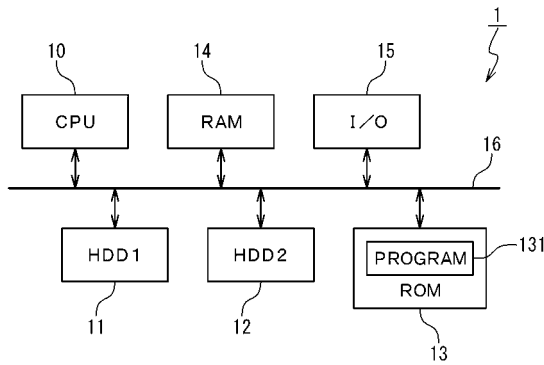
【 0 1 4 8 】

1 ...	サーバ装置	
2 ...	楽曲サーバ装置	
3 ...	クライアント端末装置	
4 ...	ネットワーク	
1 0 ...	C P U	
1 0 A ...	第 1 処理系	10
1 0 B ...	第 2 処理系	
1 1 ...	第 1 H D D	
1 2 ...	第 2 H D D	
1 3 ...	R O M	
1 4 ...	R A M	
1 5 ...	I / O ポート	
1 6 ...	内部バス	
1 0 1 A ...	第 1 ファイル形式変換部	
1 0 1 B ...	第 2 ファイル形式変換部	
1 0 2 A ...	第 1 測定部	20
1 0 2 B ...	第 2 測定部	
1 0 3 A ...	第 1 算出部	
1 0 3 B ...	第 2 算出部	
1 0 4 A ...	第 1 メタデータ生成部	
1 0 4 B ...	第 2 メタデータ生成部	
1 0 5 A ...	クエリ生成部	
1 0 6 ...	クエリ実行部	
1 0 7 ...	表示処理部	
1 0 8 ...	クローラ部	
1 0 9 ...	楽曲データベース生成部	30
1 2 1 ...	インデックスファイル	
1 3 1 ...	プログラム	

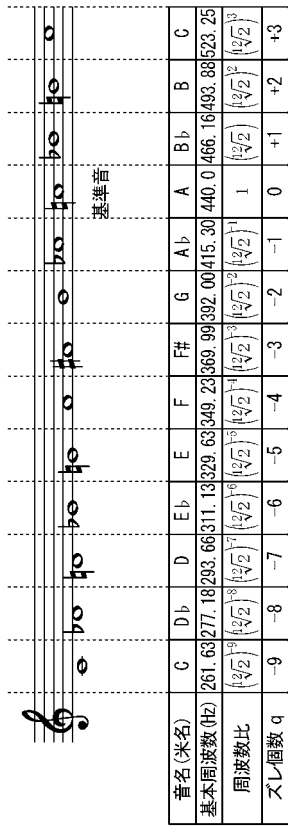
【図 1】



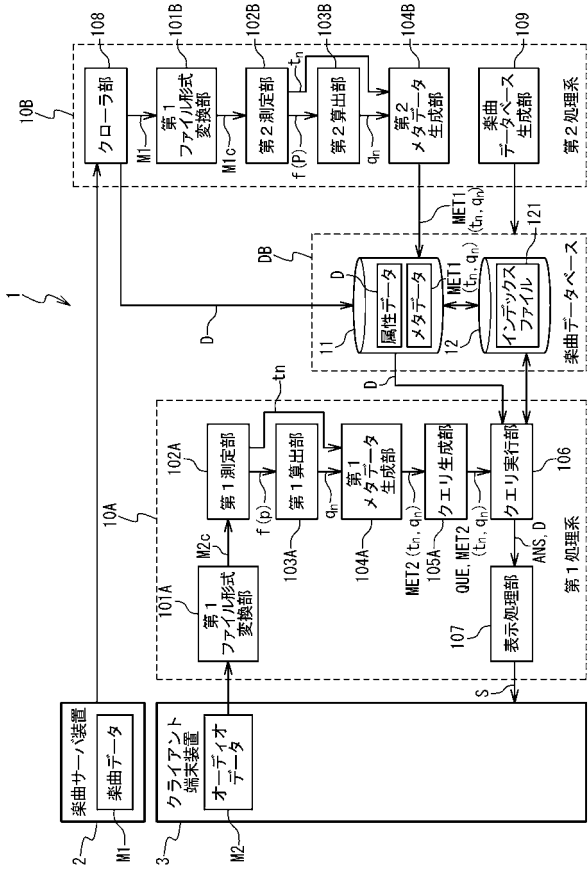
【図 2】



【図 4】



【図 3】



【図 5】

経過時間 t _n (ms)	基本周波数 f (p) (Hz)				音名 (米名)				ズレ個数 q _n				メタデータ
	P ₁	P ₂	P ₃	P ₄	D	F	G	A	q ₁	q ₂	q ₃	q ₄	
t ₁	293.7349.2				D	F			-7	-4			(100), -7, -4.,
t ₂	293.7349.2				D	F			-7	-4			(200), -7, -4.,
t ₃	293.7349.2440.0				D	F	A		-7	-4	0		(300), -7, -4.0.,
t ₄	293.7349.2440.0				D	F	A		-7	-4	0		(400), -7, -4.0.,
t ₅	293.7349.2440.0				D	F	A		-7	-4	0		(500), -7, -4.0.,
t ₆	349.2440.0				F	A			-4	-4	0		(600), -4, -0.,
t ₂₅	293.7392.0				D	G			-7	-2			(2500), -7, -2.,
t ₂₆	293.7392.0466.1				D	G	B ^b		-7	-2	1		(2600), -7, -2.1.,
t ₂₇	293.7392.0466.1659.3				D	G	B ^b	E	-7	-2	1	7	(2700), -7, -2.1.7.,
t ₆₀	293.7349.2				D	F			-7	-4			(6000), -7, -4.,
t ₆₁	293.7349.2440.0				D	F	A		-7	-4	0		(6100), -7, -4.0.,
t ₆₂	293.7349.2440.0				D	F	A		-7	-4	0		(6200), -7, -4.0.,
t ₆₃	349.2440.0698.5				F	A	F		-4	-4	0	8	(6300), -4, -0.8.,

【図 6】

