

(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2007-233078

(P2007-233078A)

(43) 公開日 平成19年9月13日(2007.9.13)

(51) Int. Cl.		F I		テーマコード (参考)
G10K 15/04 (2006.01)		G10K	15/04 302D	5D108
G10L 11/00 (2006.01)		G10L	11/00 402A	
G09B 15/00 (2006.01)		G09B	15/00 D	

審査請求 未請求 請求項の数 13 O L (全 15 頁)

(21) 出願番号 特願2006-55329 (P2006-55329)
 (22) 出願日 平成18年3月1日(2006.3.1)

(71) 出願人 000004075
 ヤマハ株式会社
 静岡県浜松市中区中沢町10番1号
 (74) 代理人 100098084
 弁理士 川▲崎▼ 研二
 (72) 発明者 入山 達也
 静岡県浜松市中区中沢町10番1号 ヤマハ株式会社内
 Fターム(参考) 5D108 BF04

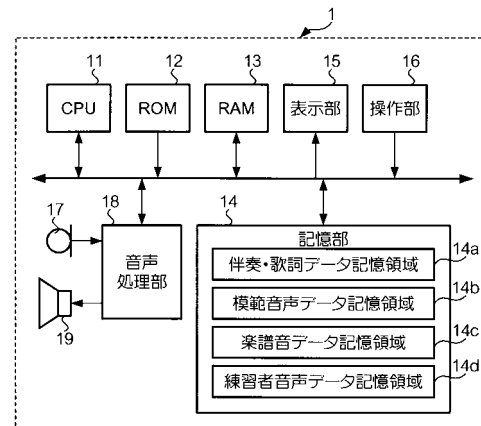
(54) 【発明の名称】 評価装置、制御方法及びプログラム

(57) 【要約】

【課題】 一定の評価基準に偏ることなく、練習者の歌唱や演奏を総合的に評価する。

【解決手段】 カラオケ装置1の記憶部14には、伴奏・歌詞データ記憶領域14aと、模範音声データ記憶領域14bと、楽譜音データ記憶領域14cと、練習者音声データ記憶領域14dとが設けられている。カラオケ装置1のCPU11は、楽譜に規定された内容と練習者の歌唱を比較する機械比較と、模範となる歌唱者の歌唱と練習者の歌唱とを比較する情感比較とを行い、これら双方の比較結果の関係から、練習者の歌唱を評価する。よって、一定の評価基準に偏ることなく、練習者の歌唱を総合的に評価することが可能となる。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

楽譜によって規定された音を表す楽譜音データを記憶する第 1 の記憶手段と、
模範となる音を表す模範音データを記憶する第 2 の記憶手段と、
收音手段によって收音された音を表す音データを練習者音データとして記憶する第 3 の
記憶手段と、

前記第 1 の記憶手段によって記憶されている楽譜音データと前記第 3 の記憶手段によっ
て記憶されている練習者音データとを比較して、両者の相違内容を特定する第 1 の比較手
段と、

前記第 2 の記憶手段によって記憶されている模範音データと前記第 3 の記憶手段によっ
て記憶されている練習者音データとを比較して、両者の相違内容を特定する第 2 の比較手
段と、

前記第 1 の比較手段によって特定された相違内容と前記第 2 の比較手段によって特定さ
れた相違内容との関係から、前記練習者音データに対する評価結果を求める評価手段と、
前記評価手段によって求められた評価結果を出力する出力手段と
を備えることを特徴とする評価装置。

10

【請求項 2】

前記出力手段は、前記練習者に前記評価結果を報知することを特徴とする請求項 1 記載
の評価装置。

【請求項 3】

前記第 1 の比較手段は、前記楽譜音データが表す音のピッチと前記練習者音データが表
す音のピッチとの差分を両者の相違内容として特定し、

前記第 2 の比較手段は、前記模範音データが表す音のピッチと前記練習者音データが表
す音のピッチとの差分を両者の相違内容として特定し、

前記評価手段は、前記第 1 の比較手段によって特定された差分と前記第 2 の比較手段に
よって特定された差分との大小関係から、前記練習者音データに対する評価結果を求め
ることを特徴とする請求項 1 記載の評価装置。

20

【請求項 4】

前記第 1 の比較手段は、前記楽譜音データが表す音のピッチと前記練習者音データが表
す音のピッチとの高低関係を両者の相違内容として特定し、

前記第 2 の比較手段は、前記模範音データが表す音のピッチと前記練習者音データが表
す音のピッチとの高低関係を両者の相違内容として特定し、

前記評価手段は、前記第 1 の比較手段によって特定された高低関係及び前記第 2 の比較
手段によって特定された高低関係に基づき、前記楽譜音データが表す音のピッチ、前記練
習者音データが表す音のピッチ及び前記模範音データが表す音のピッチのそれぞれの高低
関係を特定し、特定した高低関係から前記練習者音データに対する評価結果を求めるこ
とを特徴とする請求項 1 記載の評価装置。

30

【請求項 5】

前記評価手段は、前記練習者音データが表す音のピッチが、前記楽譜音データが表す音
のピッチと前記模範音データが表す音のピッチとの間にある場合には、高い評価結果を求
めることを特徴とする請求項 4 記載の評価装置。

40

【請求項 6】

前記第 1 の比較手段は、前記楽譜音データが表す音の発音タイミングと前記練習者音デ
ータが表す音の発音タイミングとの差分を両者の相違内容として特定し、

前記第 2 の比較手段は、前記模範音データが表す音の発音タイミングと前記練習者音デ
ータが表す音の発音タイミングとの差分を両者の相違内容として特定し、

前記評価手段は、前記第 1 の比較手段によって算出された差分と前記第 2 の比較手段に
よって算出された差分との大小関係から、前記練習者音データに対する評価結果を求め
ることを特徴とする請求項 1 記載の評価装置。

【請求項 7】

50

前記第 1 の比較手段は、前記楽譜音データが表す音の発音タイミングと前記練習者音データが表す音の発音タイミングとの時間的な前後関係を両者の相違内容として特定し、

前記第 2 の比較手段は、前記模範音データが表す音の発音タイミングと前記練習者音データが表す音の発音タイミングとの時間的な前後関係を両者の相違内容として特定し、

前記評価手段は、前記第 1 の比較手段によって特定された前後関係及び前記第 2 の比較手段によって特定された前後に基づき、前記楽譜音データが表す音の発音タイミング、前記練習者音データが表す音の発音タイミング及び前記模範音データが表す音の発音タイミングの時間的な前後関係を特定し、特定した前後関係から前記練習者音データに対する評価結果を求めることを特徴とする請求項 1 記載の評価装置。

【請求項 8】

前記評価手段は、前記練習者音データが表す音の発音タイミングが、前記楽譜音データが表す音の発音タイミングと前記模範音データが表す音の発音タイミングとの間にある場合には、高い評価結果を求めることを特徴とする請求項 7 記載の評価装置。

【請求項 9】

前記第 2 の比較手段は、前記模範音データが表す音のパワーと前記練習者音データが表す音のパワーとの差分又は大小関係を両者の相違内容として特定し、

前記評価手段は、前記第 2 の比較手段によって特定された差分又は大小関係に応じた前記評価結果を求めることを特徴とする請求項 1 記載の評価装置。

【請求項 10】

前記第 2 の比較手段は、前記模範音データが表す音のスペクトルと前記練習者音データが表す音のスペクトルとの差分を両者の相違内容として特定し、

前記評価手段は、前記第 2 の比較手段によって特定された差分に応じた前記評価結果を求めることを特徴とする請求項 1 記載の評価装置。

【請求項 11】

前記評価手段は、前記第 1 の比較手段によって特定された相違内容と前記第 2 の比較手段によって特定された相違内容とのうち、いずれか一方を他方よりも過大に参酌して前記評価結果を求めることを特徴とする請求項 1 記載の評価装置。

【請求項 12】

楽譜によって規定された音を表す楽譜音データを記憶する第 1 の記憶手段と、模範となる音を表す模範音データを記憶する第 2 の記憶手段と、收音手段によって收音された音を表す音データを練習者音データとして記憶する第 3 の記憶手段と、制御手段とを備えた評価装置の制御方法であって、

前記制御手段が、前記第 1 の記憶手段によって記憶されている楽譜音データと前記第 3 の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第 1 のステップと、

前記制御手段が、前記第 2 の記憶手段によって記憶されている模範音データと前記第 3 の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第 2 のステップと、

前記制御手段が、前記第 1 のステップにおいて特定された相違内容と前記第 2 のステップにおいて特定された相違内容との関係から、前記練習者音データに対する評価結果を求める第 3 のステップと、

前記制御手段が、前記第 3 のステップにおいて求められた評価結果を出力する第 4 のステップと

を備えることを特徴とする制御方法。

【請求項 13】

楽譜によって規定された音を表す楽譜音データを記憶する第 1 の記憶手段と、模範となる音を表す模範音データを記憶する第 2 の記憶手段と、收音手段によって收音された音を表す音データを練習者音データとして記憶する第 3 の記憶手段とを備えたコンピュータに、

前記第 1 の記憶手段によって記憶されている楽譜音データと前記第 3 の記憶手段によっ

10

20

30

40

50

て記憶されている練習者音データとを比較して、両者の相違内容を特定する第1の比較機能と、

前記第2の記憶手段によって記憶されている模範音データと前記第3の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第2の比較機能と、

前記第1の比較機能によって特定された相違内容と前記第2の比較機能によって特定された相違内容との関係から、前記練習者音データに対する評価結果を求める評価機能と、

前記評価機能によって求められた評価結果を出力する出力機能と
を実現させるプログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は歌唱や演奏を評価する技術に関する。

【背景技術】

【0002】

カラオケ装置においては、歌唱者の歌唱の巧拙を評価するための方法が種々提案されている。その評価方法の1つに、楽譜の音符と歌唱者の歌唱音とを比較し、両者の音程やタイミングがどの程度一致しているかということに基づいて歌唱を評価するものがある（例えば特許文献1, 2参照）。また、模範となるような歌唱者の歌唱音と、練習者の歌唱音とを比較し、両者の一致度を基準として練習者の歌唱を評価する方法も提案されている（例えば特許文献3参照）。

20

【0003】

ところで、歌手のように熟練した歌唱者は、楽譜の内容に沿って機械的に歌を歌うことはほとんどなく、その多くが、歌い始めや歌い終わりを意図的にずらしたり、ピブラートやこぶし等の歌唱技法を用いたりして、歌のなかに感情の盛り上がり（情感）を表現する。カラオケ装置で歌唱を練習する者のなかには、このような意図的なタイミングのずれや歌唱技法を真似て歌いたいという要望を持つ者が少なくない。前述した特許文献1, 2に記載された評価方法ではこのような要望に応えることができないが、特許文献3に記載された評価方法によれば、歌唱に現れる情感を評価の基準とすることができる。

30

【特許文献1】特開昭62-040488号公報

【特許文献2】特許第2890659号公報

【特許文献3】特公平06-007293号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、特許文献3に記載された評価方法では、模範となる歌唱をそっくり真似て、その歌唱者に特有の癖や特徴までも忠実に再現しなければ、高い評価を得ることができない。模範となる歌唱がどのような練習者にとっても常に良いというわけでもないから、これをいたずらに真似するばかりでは練習者の歌唱技術が向上しない場合も考えられる。

40

このような問題は、歌唱を練習する場合に限らず、模範演奏を真似て楽器の演奏を練習する場合であっても同様に発生する。

【0005】

本発明は上述した背景に鑑みてなされたものであり、その目的は、一定の評価基準に偏ることなく、練習者の歌唱や演奏を総合的に評価することにある。

【課題を解決するための手段】

50

【0006】

上記課題を解決するため、本発明は、楽譜によって規定された音を表す楽譜音データを記憶する第1の記憶手段と、模範となる音を表す模範音データを記憶する第2の記憶手段と、收音手段によって收音された音を表す音データを練習者音データとして記憶する第3の記憶手段と、前記第1の記憶手段によって記憶されている楽譜音データと前記第3の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第1の比較手段と、前記第2の記憶手段によって記憶されている模範音データと前記第3の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第2の比較手段と、前記第1の比較手段によって特定された相違内容と前記第2の比較手段によって特定された相違内容との関係から、前記練習者音データに対する評価結果を求める評価手段と、前記評価手段によって求められた評価結果を出力する出力手段とを備えることを特徴とする評価装置を提供する。

10

【0007】

この評価装置によれば、楽譜音データと練習者音データとの相違内容と、模範音データと練習者音データとの相違内容との関係から練習者音データに対する評価結果を求める。よって、例えば楽譜音データと練習者音データとの相違内容だけを評価基準にするとか、模範音データと練習者音データとの相違内容だけを評価基準にする、といったように一定の評価基準に偏ることなく、練習者の歌唱や演奏を総合的に評価することができる。

【0008】

また、本発明は、楽譜によって規定された音を表す楽譜音データを記憶する第1の記憶手段と、模範となる音を表す模範音データを記憶する第2の記憶手段と、收音手段によって收音された音を表す音データを練習者音データとして記憶する第3の記憶手段と、制御手段とを備えた評価装置の制御方法であって、前記制御手段が、前記第1の記憶手段によって記憶されている楽譜音データと前記第3の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第1のステップと、前記制御手段が、前記第2の記憶手段によって記憶されている模範音データと前記第3の記憶手段によって記憶されている練習者音データとを比較して、両者の相違内容を特定する第2のステップと、前記制御手段が、前記第1のステップにおいて特定された相違内容と前記第2のステップにおいて特定された相違内容との関係から、前記練習者音データに対する評価結果を求める第3のステップと、前記制御手段が、前記第3のステップにおいて求められた評価結果を出力する第4のステップとを備えることを特徴とする制御方法を提供する。

20

30

【0009】

この制御方法によれば、楽譜音データと練習者音データとの相違内容と、模範音データと練習者音データとの相違内容との関係から練習者音データに対する評価結果を求める。よって、一定の評価基準に偏ることなく、練習者の歌唱や演奏を総合的に評価することができる。

【0010】

さらに、本発明は、コンピュータに対して機能を実現させるプログラムとしての形態も採り得る。なお、本発明において「発音」という用語には、人が歌唱するときに出せられる音声のほか、楽器を演奏することで出せられる演奏音も含むものとする。

40

【発明の効果】

【0011】

本発明によれば、一定の評価基準に偏ることなく、練習者の歌唱や演奏を総合的に評価することができる。

【発明を実施するための最良の形態】

【0012】

次に、本発明を実施するための最良の形態を説明する。

1. 構成

図1は、この発明の一実施形態に係る評価装置としてのカラオケ装置1のハードウェア構成を示したブロック図である。CPU (Central Processing Unit) 11は、ROM (R

50

ead Only Memory) 1 2 又は記憶部 1 4 に記憶されているコンピュータプログラムを読み出して R A M (Random Access Memory) 1 3 にロードし、これを実行することにより、カラオケ装置 1 の各部を制御する。記憶部 1 4 は、例えばハードディスクなどの大容量の記憶手段であり、後述する「機械比較」の結果と「情感比較」の結果との関係から練習者の歌唱を評価するための評価アルゴリズムを記憶している。また、この記憶部 1 4 には、伴奏・歌詞データ記憶領域 1 4 a と、模範音声データ記憶領域 1 4 b と、楽譜音データ記憶領域 1 4 c と、練習者音声データ記憶領域 1 4 d という 4 つの記憶領域が設けられている。

【0013】

表示部 1 5 は、例えば液晶ディスプレイなどであり、C P U 1 1 の制御の下で、カラオケ装置 1 を操作するためのメニュー画面や、背景画像に歌詞テロップを重ねたカラオケ画面などの各種画面を表示する。操作部 1 6 は、各種のキーを備えており、押下されたキーに対応した信号を C P U 1 1 へ出力する。マイクロフォン 1 7 は、歌唱者が発音した音声を収音する収音手段である。音声処理部 1 8 は、マイクロフォン 1 7 によって収音された音声(アナログデータ)をデジタルデータに変換して C P U 1 1 に出力する。スピーカ 1 9 は、音声処理部 1 8 に接続されており、音声処理部 1 8 から出力される音声を放音する。

10

【0014】

次に、記憶部 1 4 に記憶されるデータの内容について説明する。

伴奏・歌詞データ記憶領域 1 4 a には、楽曲の伴奏を行う各種楽器の演奏音が楽曲の進行に伴って記された伴奏データと、楽曲の歌詞を示す歌詞データとが対応付けられて記憶されている。伴奏データは、例えば M I D I (Musical Instruments Digital Interface) 形式などのデータ形式であり、練習者がカラオケ歌唱する際に再生される。歌詞データは、そのカラオケ歌唱の際に歌詞テロップとして表示部 1 5 に表示される。模範音声データ記憶領域 1 4 b には、模範となる歌唱者が楽曲を歌唱した際の音声(以下、模範音声という)を表す音声データ(以下、模範音声データ)が記憶されている。この模範音声データは、例えば W A V E 形式や M P 3 (MPEG Audio Layer-3) 形式などのデータ形式で記憶されている。

20

【0015】

楽譜音データ記憶領域 1 4 c には、楽曲の楽譜によって規定された歌唱音を表す楽譜音データが記憶されている。この楽譜音データは、例えば M I D I 形式などのデータ形式であり、歌唱音のピッチとその発音タイミングとを含んでいる。練習者音声データ記憶領域 1 4 d には、マイクロフォン 1 7 から音声処理部を経て A / D 変換された音声データが、例えば W A V E 形式や M P 3 (MPEG Audio Layer-3) 形式で時系列に記憶される。この音声データは、練習者の音声を表すものであるから、以下では、練習者音声データという。なお、以下の説明においては、説明の便宜上、「模範音声データ」と「練習者音声データ」とを各々区別する必要がない場合には、これらを「音声データ」と総称する。

30

【0016】

次に、図 2 に示すブロック図を参照しながら、カラオケ装置 1 のソフトウェア構成について説明する。図 2 に示した基礎分析部 1 1 1、対応箇所検出部 1 1 2、機械比較部 1 1 3、情感比較部 1 1 4、総合評価部 1 1 5 及び報知部 1 1 6 は、C P U 1 1 が R O M 1 2 又は記憶部 1 4 に記憶されたコンピュータプログラムを実行することによって実現される。なお、図中の矢印は、データの流れを概略的に示したものである。

40

図 2 において、基礎分析部 1 1 1 は、模範音声データ記憶領域 1 4 b から読み出された模範音声データと、練習者音声データ記憶領域 1 4 d から読み出された練習者音声データとを、それぞれ所定時間長のフレーム単位に分離し、これらフレーム単位で各音声データのピッチ及びスペクトルを検出する。スペクトルの検出には F F T (Fast Fourier Transform) を用いればよい。基礎分析部 1 1 1 によって検出された各音声データのピッチは機械比較部 1 1 3 に出力されるとともに、情感比較部 1 1 4 にも出力される。また、基礎分

50

析部 1 1 1 によって検出された各音声データのスペクトルは情感比較部 1 1 4 に出力されると共に、対応箇所検出部 1 1 2 に出力される。

【 0 0 1 7 】

対応箇所検出部 1 1 2 は、模範音声のスペクトルと練習者音声のスペクトルに基づいて、両者の対応関係を求める。具体的には以下のようにして求める。

模範音声と練習者音声とは時間的に前後にずれている可能性がある。例えば図 3 (a) は、「すぎさりしひ・・・」という歌詞を練習者が歌唱した際の発音タイミングが、模範となる歌唱者の発音タイミングよりも遅れている例を示している。対応箇所検出部 1 1 2 は、このように模範音声と練習者音声とが時間的に前後にずれている場合であっても、両者を対応付けられるようにするため、練習者音声データと模範音声データのいずれか一方の時間軸を伸縮させる時間正規化 (D T W ; Dynamic Time Warping) を行い、両者の時間軸を合わせる。この D T W を行うための手法としては、D P (Dynamic Programming : 動的計画法) マッチングがあるので、これを用いればよい。これにより、例えば図 3 (a) に示すように、練習者音声の発音タイミングが模範音声の発音タイミングよりも遅れていたとしても、図 3 (b) に示すように、合わせられた時間軸上の位置を同じくする音どうしを対応付けることができる。このようにして対応箇所検出部 1 1 2 によって検出された対応箇所は、情感比較部 1 1 4 に出力される。

10

【 0 0 1 8 】

機械比較部 1 1 3 は、楽譜音データと練習者音声データとを比較し、両者の相違内容を特定する。この比較は、練習者が楽譜に規定された内容どおりに歌唱することができているかどうかを判定する意味合いがあるから、「機械比較」と称する。より具体的には、機械比較部 1 1 3 は、ピッチ比較部 1 1 3 a とタイミング比較部 1 1 3 b を備えている。ピッチ比較部 1 1 3 a は、楽譜音データが表すピッチと、基礎分析部 1 1 1 から出力されてくる練習者音声のピッチとを比較して両者の相違内容を特定する。例えば、一方の音声のピッチが他方の音声のピッチよりもどの程度高いか低いかなど、両者の差分とその高低関係を特定する。タイミング比較部 1 1 3 b は、楽譜音データが表す発音タイミングと、基礎分析部 1 1 1 から出力されてくる練習者音声のピッチが切り替わったタイミング (つまり練習者の発音タイミング) とを比較して両者の相違内容を特定する。例えば、一方の音声の発音タイミングが他方の音声の発音タイミングよりもどの程度早い遅いかなど、両者の差分とその時間的な前後関係を特定する。

20

30

【 0 0 1 9 】

情感比較部 1 1 4 は、対応箇所検出部 1 1 2 によって検出される対応箇所を参照しながら、模範音声データと練習者音声データとの間で対応する音声データどうしを比較し、両者の相違内容を特定する。この比較は、練習者が模範となる歌唱者のように情感を込めて歌唱することができているかどうかを判定する意味合いがあるから、「情感比較」と称する。より具体的には、情感比較部 1 1 4 は、スペクトル比較部 1 1 4 a と、ピッチ比較部 1 1 4 b と、タイミング比較部 1 1 4 c とを備えている。スペクトル比較部 1 1 4 a は、模範音声データと練習者音声データとの間で、それぞれ対応する箇所の音声のスペクトルを比較して両者の相違内容を特定する。例えば両者のスペクトルの差分を算出して、その差分が予め決められた閾値よりも小さいか大きいかを特定する。ピッチ比較部 1 1 4 b は、模範音声データと練習者音声データとの間で、それぞれ対応する箇所のピッチを比較して両者の相違内容を特定する。例えば、一方の音声のピッチが他方の音声のピッチよりもどの程度高いか低いかなど、両者の差分とその高低関係を特定する。タイミング比較部 1 1 4 c は、模範音声データと練習者音声データとの間で、それぞれ対応する箇所の発音タイミングを比較して両者の相違内容を特定する。例えば、一方の音声の発音タイミングが他方の音声の発音タイミングよりどの程度早い遅いかなど、両者の差分とその時間的な前後関係を特定する。

40

【 0 0 2 0 】

総合評価部 1 1 5 は、機械比較部 1 1 3 による比較結果と、情感比較部 1 1 4 による比較結果との関係から、練習者の歌唱の稚拙を総合的に評価する。報知部 1 1 6 は、その評

50

価結果を表す評価点や各種メッセージを生成し、これらを表示部15に表示することで練習者に報知する。

【0021】

2. 動作

次に、図4に示すフローチャートを参照しつつ、カラオケ装置1の動作を説明する。

練習者は、カラオケ装置1の操作部16を操作して歌唱したい曲を選定し、再生を指示する。CPU11は、この指示に応じて図4に示す処理を開始する。CPU11は、まず、指定された曲の伴奏データ及び歌詞データを伴奏・歌詞データ記憶領域14aから読み出し、伴奏データを音声処理部18に出力する(ステップS1)。音声処理部18は、出力されてくる伴奏データをアナログ信号に変換してスピーカ19に出力して放音させる。さらに、CPU11は表示部15を制御して、歌詞データが表す歌詞を表示部15に表示するとともに、「伴奏に合わせて歌唱してください」というような歌唱を促すメッセージを表示する。

10

【0022】

練習者は、表示された歌詞を参照しつつ、スピーカ19から放音される伴奏に合わせて歌唱を行う。このとき、練習者の音声はマイクロフォン17によって収音されて音声信号に変換され、音声処理部18へと出力される。そして、音声処理部18によってA/D変換された練習者音声データは、演奏開始から発音タイミングまでの経過時間を表す情報と共に、記憶部14の練習者音声データ記憶領域14dに記憶されていく(ステップS2)。

20

【0023】

伴奏データの再生が終了すると(ステップS3; YES)、CPU11は、練習者音声データ記憶領域14dに記憶された練習者音声データと、模範音声データ記憶領域14bに予め記憶されている模範音声データに対して、上述した基礎分析部111の処理を行う(ステップS4)。つまり、CPU11は、それぞれの音声データを所定時間長のフレーム単位に分離し、これらフレーム毎にピッチ及びスペクトルを検出する。

【0024】

次に、CPU11は、前述した対応箇所検出部112の処理、つまりDPマッチングによって練習者音声データの時間軸と模範音声データの時間軸とを合わせ、その時間軸上の位置を同じくする音どうしを対応付ける(ステップS5)。

30

【0025】

次に、CPU11は、前述した機械比較部113の処理を行う(ステップS6)。これにより、CPU11は、楽譜音データと練習者音声データとの間で、ピッチや発音タイミングがどの程度相違しているかということ特定することができる。例えば図5(a)に示すように、「す」という音声については楽譜音のピッチ p_1 よりも練習者音声のピッチ p_2 の方が p_a だけ高いとか、図5(b)に示すように、楽譜音の発音タイミング t_1 よりも練習者音声の発音タイミング t_2 の方が t_a だけ早い、といった具合に両者の相違内容を特定する。CPU11は、このような相違内容を、楽譜音及び練習者音声の対応箇所の全てについて特定する。

【0026】

続けて、CPU11は、前述した情感比較部114の処理を行う(ステップS7)。これにより、CPU11は、模範音声データと練習者音声データとの間で、ピッチ、発音タイミング及びスペクトルがどの程度相違しているかということ把握することができる。例えば図6(a)に示すように、「す」という音声については、模範音声のピッチ p_3 よりも練習者音声のピッチ p_2 の方が p_b だけ高いとか、図6(b)に示すように、模範音声の発音タイミング t_3 よりも練習者音声の発音タイミング t_2 の方が t_b だけ遅いとか、図6(c)に示すように、模範音声のスペクトルと練習者音声のスペクトルが類似している(フォルマント f_1 , f_2 の周波数帯域が一致している)、といった具合である。CPU11は、上記のような相違内容を、楽譜音及び練習者音声の対応箇所の全てについて特定する。

40

50

【0027】

次に、CPU11は、ステップS6において特定された相違内容と、ステップS7において特定された相違内容とに基づき、記憶部14に記憶されている評価アルゴリズムに従って練習者の歌唱を総合的に評価する(ステップS8)。

【0028】

ここで、具体的な評価方法について説明する。

例えば図7に示すように、練習者音声のピッチ p_2 が、楽譜音のピッチ p_1 よりも p_a だけ高く、且つ、模範音声のピッチ p_3 よりも p_b だけ高い($p_b < p_a$)場合を想定する。つまり、ピッチの低い順から並べると、楽譜音のピッチ p_1 、模範音声のピッチ p_3 、練習者音声のピッチ p_2 という順序となる。CPU11は、これらのピッチ p_1 、 p_2 、 p_3 の高低関係と、差分 p_a 、 p_b の大小関係に基づき、練習者音声の評価点を算出する。

10

【0029】

図8は、ピッチの評価点を算出する方法の一例を説明する図であり、模範音声及び楽譜音のピッチに対する近さと評価点との関係を示している。図8に示すように、練習者音声のピッチが模範音声のピッチと楽譜音のピッチとの間にある場合には高い評価点を獲得でき、さらに、模範音声に近づけば近づくほど高い評価点となり、また、楽譜音に近づけば近づくほど高い評価点となる。ただし、楽譜音に近づくよりも、模範音声に近づくほうがより高い評価点となる。つまり、機械比較の結果と情感比較の結果のうち、情感比較の結果を過大に参酌している。このような評価体系は、楽譜音のピッチと模範音声のピッチとの間に練習者音声のピッチがある場合にはそれを高く評価すべきであるし、また、楽譜音よりも模範音声のほうが巧い歌唱であるから、模範音声に近いピッチをより高く評価する、という考え方によるものである。

20

【0030】

また、図9は、別の評価点算出方法を説明する図である。図9に示す例であっても、楽譜音のピッチと模範音声のピッチとの間に練習者音声のピッチがある場合にはそれを高く評価する点で一致するが、異なるのは、模範音声のピッチから所定範囲にあるピッチのみを最も高く評価するという点と、楽譜音のピッチと模範音声のピッチとの間に練習者音声のピッチが無い場合には一切評価しないという点である。

【0031】

そして、図10は、さらに別の算出方法を説明する図である。図10に示す例であっても、楽譜音のピッチと模範音声のピッチとの間に練習者音声のピッチがある場合にはそれを高く評価する点で一致するが、異なるのは、楽譜音のピッチから模範音声のピッチに近づくに従って徐々に高く評価するという点である。

30

なお、図8～図10に示した例では、楽譜音のピッチよりも模範音声のピッチの方が高い場合で説明した。一方、これとは逆に、楽譜音のピッチよりも模範音声のピッチの方が低い場合もあり得るが、このような場合であっても上記と同様の考え方に従って評価すればよい。

【0032】

次に、発音タイミングの評価方法について説明する。

例えば図11に示すように、練習者音声の発音タイミング t_2 が、楽譜音の発音タイミング t_1 よりも t_a だけ早く、且つ、模範音声の発音タイミング t_3 よりも t_b だけ遅い。つまり、発音タイミングの早い順から並べると、模範音声の発音タイミング t_3 、練習者音声の発音タイミング t_2 、楽譜音の発音タイミング t_1 という順序となる。CPU11は、これらの発音タイミング t_1 、 t_2 、 t_3 の時間的な前後関係と、差分 t_a 、 t_b の大小関係に基づき、練習者音声の評価点を算出する。その具体的な算出方法は、図8～図10で説明したものと同様なものを用いればよい。つまり、図8～図10においてその横軸を発音タイミング(伴奏開始からの経過時間)として、練習者音声の発音タイミング(横軸)に対応する評価点(縦軸)を求めればよい。

40

【0033】

50

そして、CPU11は、上記のようにして対応箇所毎に求めた評価点を全て累算し、総合の評価点を算出する。そして、CPU11は、その評価点や各種メッセージを生成し、これらを表示部15に表示する(ステップS9)。例えば練習者音声模範音声と非常に近くて高い評価点を獲得したという傾向がある場合には、CPU11は、「あなたは歌手のように巧く歌えています。100点満点です。」というメッセージを表示する。また、練習者音声模範音と非常に近くて高い評価点を獲得したという傾向がある場合には、CPU11は、「あなたは楽譜どおりに歌えています。90点です。今後は情感を込めて歌う練習をすると良いでしょう。」というメッセージを表示する。また、練習者音声模範音との中間程度に位置することである程度高い評価を獲得したという傾向にある場合には、CPU11は、「もう少し情感を込めて歌えばさらに良くなるでしょう。85点です。」というメッセージを表示する。また、練習者音声模範音との間に位置しないことが多かったことで低い評価点しか得られなかった場合には、CPU11は、「まずは楽譜通りに歌う練習をしましょう。50点です。」というメッセージを表示する。さらに、CPU11は、ステップS7において模範音声のスペクトルと練習者音声のスペクトルとの相違内容を特定しているから、その内容も考慮して評価結果を求めることもできる。例えば練習者音声の音質(スペクトル)が模範音声の音質(スペクトル)に非常に近い場合、CPU11は、「あなたはこの曲の歌手の声と似ています。70点です。更に練習するともっと良くなるでしょう。」というメッセージを表示する。

10

【0034】

このように本実施形態においては、カラオケ装置1は、楽譜に規定された内容と練習者の歌唱を比較する機械比較と、模範となる歌唱者の歌唱と練習者の歌唱とを比較する情感比較とを行い、これら双方の比較結果の関係から練習者の歌唱を評価する。よって、一定の評価基準に偏ることなく、練習者の歌唱を総合的に評価することが可能となる。さらに、練習者に対して評価結果を示すことで、その練習者の歌唱の稚拙に対する的確な指導を行うことが可能となる。

20

【0035】

3. 変形例

上述した実施形態を次のように変形してもよい。

(1) 上述した実施形態においては、練習者の歌唱を評価する場合を例に挙げて説明したが、これに限らず、練習者の楽器演奏を評価するようにしてもよい。この場合、伴奏・歌詞データ記憶領域14aには、練習したい楽器(例えばギター)以外の楽器(例えばベースやドラム)の演奏データが記憶されており、模範音声データ記憶領域14bには、模範となる模範演奏データが記憶されており、楽譜音データ記憶領域14cには、楽譜に規定された楽譜音データが記憶されており、練習者音声データ記憶領域14dには、練習者の演奏データが記憶される。CPU11は、これらのデータに基づき、上記と同様の処理を経て練習者の演奏を評価する。このように、本発明は歌唱や演奏を含み得るため、本発明における「発音」という用語には、人が歌唱するとき発せられる音声のほか、楽器を演奏することで発せられる演奏音も含むものとする。

30

【0036】

(2) 機械比較と情感比較とを用いた評価方法は、図8~図10に示した例に限定されない。例えば、図8~図10では機械比較の結果よりも情感比較の結果を過大に参酌していたが、これを逆にして機械比較の結果を過大に参酌してもよいし、両者を均等に参酌してもよい。例えば練習者があまり巧く歌唱できない時期には、まずは楽譜どおりに歌唱することが大事であるから機械比較の結果を過大に参酌し、歌唱が徐々に巧くなるに従って、情感比較の結果を参酌する度合いを大きくしてもよい。つまり、図8~図10に示したような模範音声及び楽譜音のピッチに対する近さと評価点との関係は、カラオケ装置1の使用者が操作部16を操作することで適宜変更可能に構成することが望ましい。

40

また、例えば違う歌手が同じ楽曲を歌唱してもそれぞれ異なる情感が醸し出されることはよくある。そこで、それぞれの楽曲に対して模範音声をそれぞれ1つずつ用意するのはなく、複数の模範音声を用意しておいてもよい。そして、例えばある歌手の模範音声に

50

は近くなくても、別の歌手の模範音声に近い場合には、CPU 11は高い評価点を算出し、さらに「あなたは歌手の××に似ています。80点です。」といったメッセージを表示するとよい。

【0037】

(3)実施形態では、模範音声と練習者音声との間で情感比較を行う際に、音声のピッチ、発音タイミング及びスペクトルを比較対象としていたが、さらに音声のパワーを比較するようにしてもよい。つまり、CPU 11は、模範音声のパワーと練習者音声パワーとの差分や大小関係を両者の相違内容として特定し、その差分や大小関係に応じた評価結果を求めるようにしてもよい。

【0038】

(4)CPU 11が、過去の評価結果を記憶しておき、その評価結果をも考慮したメッセージを報知するようにしてもよい。例えば過去の評価では機械比較の結果が良好で情感比較の結果が芳しくなかったのに、その後、情感比較の結果が向上してきた場合には、CPU 11が「情感を込めて歌うのが巧くなりましたね」などというようなメッセージを表示すると、練習者の練習の励みになる。このため、カラオケ装置1で歌唱を行う際には練習者に割り当てられたIDを入力するようにし、記憶部14にはそのIDに対応付けてその練習者の過去の評価結果を記憶する。そして、CPU 11は、練習者のIDに対応付けて記憶された過去の評価結果を参照しながら、最新の評価に関するメッセージを生成する。

【0039】

(5)報知部116によって評価結果を報知する形態は、表示に限らず、音声メッセージを出力するような形態であってもよい。また、メッセージを電子メール形式で練習者のメール端末に送信するという形態であってもよい。また、メッセージを記憶媒体に出力して記憶させるようにしてもよく、この場合、練習者はコンピュータを用いてこの記憶媒体から評価結果を読み出させることで、それを参照することができる。要は、練習者に対して何らかの手段でメッセージ乃至情報を伝えられるように、評価結果を出力するものであればよい。

【0040】

(6)実施形態では、ハードディスク等の記憶部14に練習者音声データを記憶するようしていたが、歌唱の評価を終えた後に練習者音声をすぐに破棄する場合には、練習者音声データをRAM 13に記憶するようにしてもよい。

【0041】

(7)実施形態では、練習者音声データを記憶する際には歌詞を表示し、さらに伴奏データを再生しながら練習者に歌唱させる、所謂カラオケ歌唱を行うようしていたが、これは必ずしも必要ではない。つまり、練習者が歌詞の表示や伴奏データの再生が無いままで歌唱し、それを録音して評価するようにしてもよい。歌唱能力が相当に高い練習者であっても、歌詞の表示や伴奏が無い状態でうまく歌唱することは容易ではないから、練習者の歌唱能力をより厳密に評価することが可能となる。

【0042】

(8)実施形態では、CPU 11が図4に示す処理を実行するたびに、模範音声データに対して基礎分析を行っていたが(ステップS4)、これに限らず、模範音声データに対して予め基礎分析を行った結果を記憶部14に記憶しておいてもよいし、過去に一度でも基礎分析を行ったことがあれば、その結果を記憶部14に記憶しておいてもよい。なお、模範音声データや練習者音声データはWAVE形式やMP3形式のデータとしたが、データの形式はこれに限定されるものではなく、音声を示すデータであればどのような形式のデータであってもよい。

【0043】

(9)さらに、実施形態においては、模範音声データを記憶部14に記憶させて、カラオケ装置1のCPU 11が記憶部14から模範音声データを読み出すようにしたが、これに代えて、通信ネットワークを介して音声データを受信するようにしてもよい。

【0044】

10

20

30

40

50

(10) 実施形態では、カラオケ装置1が、図2に示した機能の全てを実現するようになっていた。これに対し、通信ネットワークで接続された2以上の装置が上記機能を分担するようにし、それら複数の装置を備えるシステムが同実施形態のカラオケ装置1を実現するようにしてもよい。例えば、マイクロフォンやスピーカ、表示装置及び入力装置等を備え、報知部116を実現するコンピュータ装置と、基礎分析部111、対応箇所検出部112、機械比較部113、情感比較部114及び総合評価部115を実現するサーバ装置とが通信ネットワークで接続されたシステムとして構成されていてもよい。この場合は、コンピュータ装置が、マイクロフォンから入力された音声を音声データに変換してサーバ装置に送信し、サーバ装置が、受信した音声データと模範音声データ及び楽譜音データとの比較処理を行い、その比較に基づく評価結果をコンピュータ装置に送信するようにすればよい。

【0045】

(11) 上述した実施形態における評価装置としてのカラオケ装置1のCPU11によって実行されるプログラムは、磁気テープ、磁気ディスク、フロッピー（登録商標）ディスク、光記録媒体、光磁気記録媒体、CD（Compact Disk）-ROM、DVD（Digital Versatile Disk）、RAMなどの記録媒体に記憶した状態で提供し得る。また、インターネットのようなネットワーク経由でカラオケ装置1にダウンロードさせることも可能である。

【図面の簡単な説明】

【0046】

【図1】カラオケ装置のハードウェア構成の一例を示すブロック図である。

【図2】カラオケ装置のソフトウェア構造の一例を示すブロック図である。

【図3】DPマッチングによる時間軸の調整の一例を示す図である。

【図4】カラオケ装置のCPUが行う処理の流れを示すフローチャートである。

【図5】練習者音声と楽譜音との相違内容の一例を示す図である。

【図6】練習者音声と模範音声との相違内容の一例を示す図である。

【図7】練習者音声のピッチ、楽譜音のピッチ及び模範音声のピッチの関係の一例を示す図である。

【図8】CPUによる評価方法を示す図である。

【図9】CPUによる評価方法を示す図である。

【図10】CPUによる評価方法を示す図である。

【図11】練習者音声の発音タイミング、楽譜音の発音タイミング及び模範音声の発音タイミングの関係の一例を示す図である。

【符号の説明】

【0047】

1...カラオケ装置、11...CPU、12...ROM、13...RAM、14...記憶部、15...表示部、16...操作部、17...マイクロフォン、18...音声処理部、19...スピーカ、111...基礎分析部、112...対応箇所検出部、113...機械比較部、114...情感比較部、115...総合評価部、116...報知部。

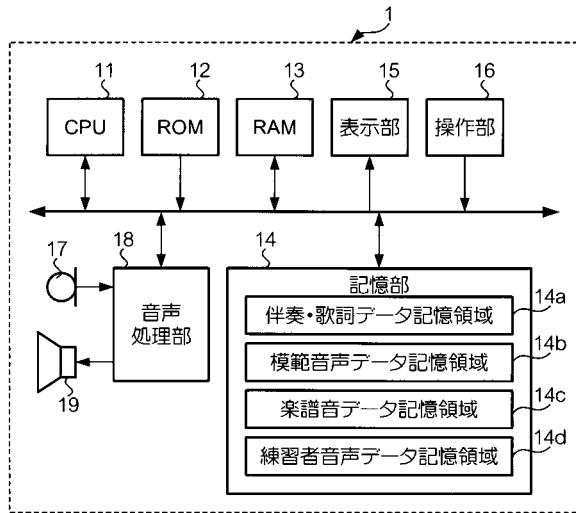
10

20

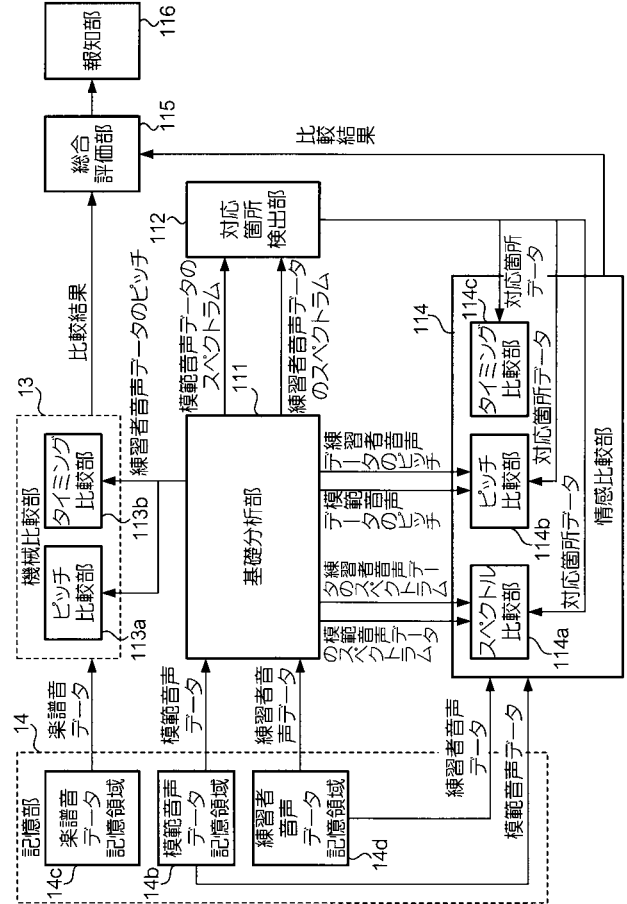
30

40

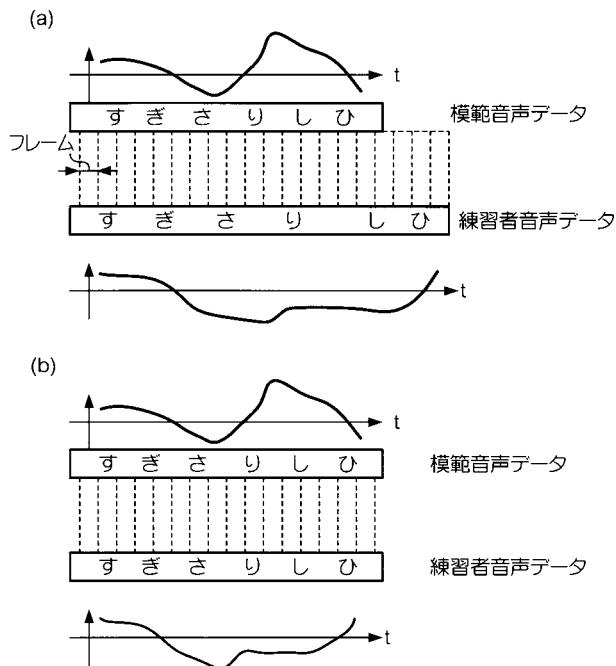
【図1】



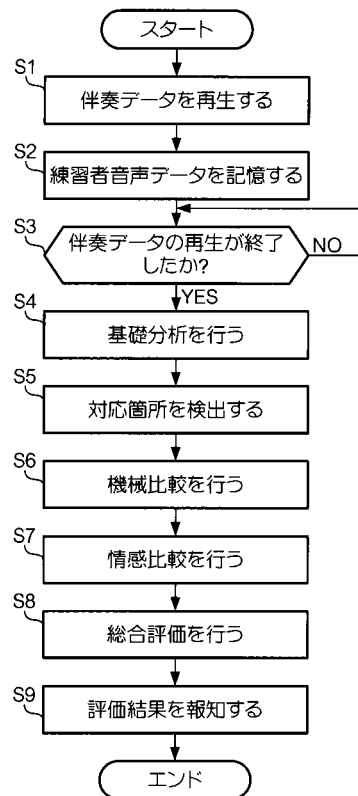
【図2】



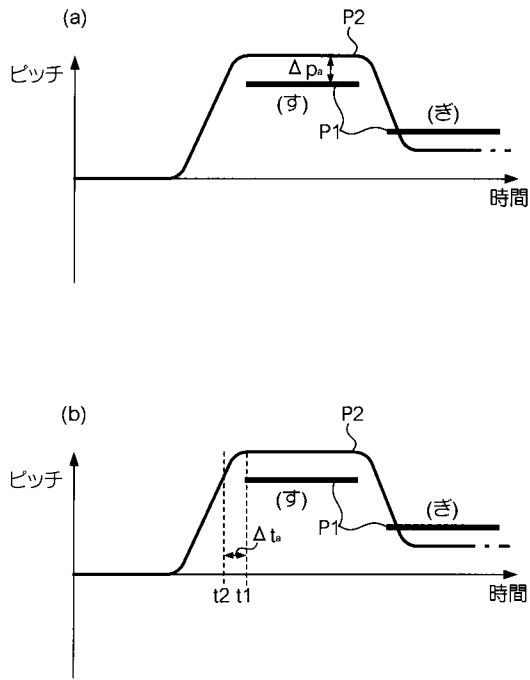
【図3】



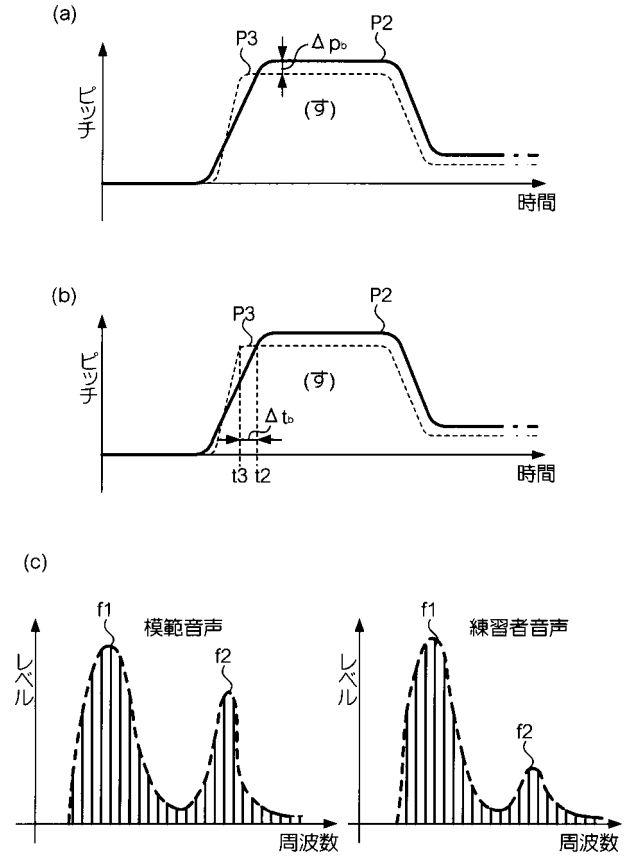
【図4】



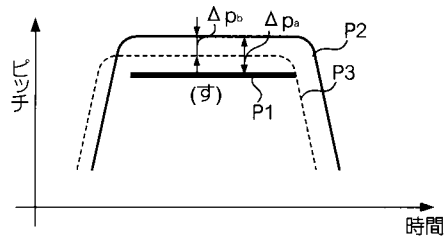
【 図 5 】



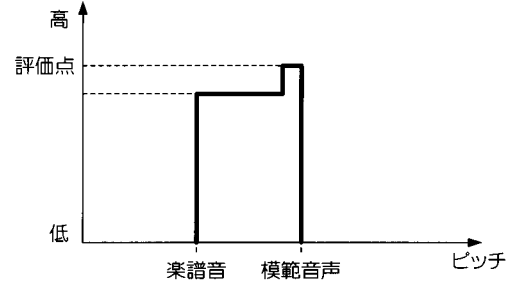
【 図 6 】



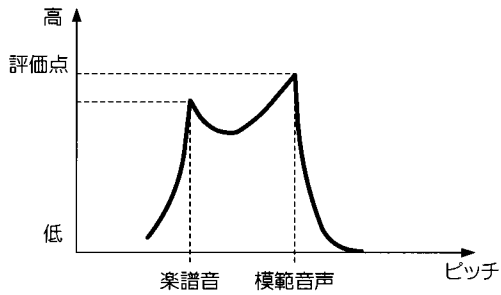
【 図 7 】



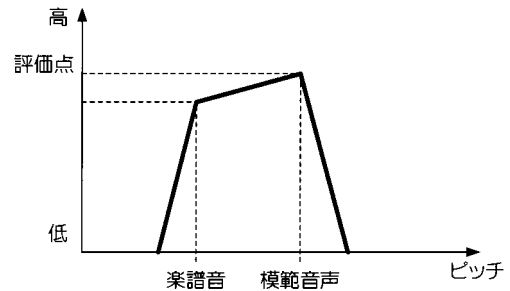
【 図 9 】



【 図 8 】



【 図 10 】



【 図 1 1 】

