

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2013 年 10 月 24 日(24.10.2013)



(10) 国際公開番号
WO 2013/157602 A1

- (51) 国際特許分類:
G10K 15/04 (2006.01) G10L 25/51 (2013.01)
G09B 15/00 (2006.01)
- (21) 国際出願番号: PCT/JP2013/061488
- (22) 国際出願日: 2013 年 4 月 18 日(18.04.2013)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2012-094853 2012 年 4 月 18 日(18.04.2012) JP
- (71) 出願人: ヤマハ株式会社(YAMAHA CORPORATION) [JP/JP]; 〒4308650 静岡県浜松市中区中沢町 1 0 番 1 号 Shizuoka (JP).
- (72) 発明者: 松本 秀一(MATSUMOTO Shuichi); 〒4308650 静岡県浜松市中区中沢町 1 0 番 1 号 ヤマハ株式会社内 Shizuoka (JP).
- (74) 代理人: 本多 弘徳(HONDA Hironori); 〒1050003 東京都港区西新橋一丁目 7 番 1 3 号 虎ノ門

イーストビルディング 1 0 階 栄光特許事務所
Tokyo (JP).

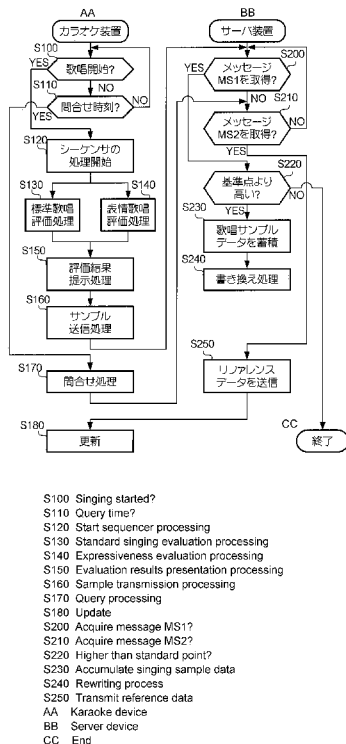
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI

[続葉有]

(54) Title: PERFORMANCE EVALUATION DEVICE, KARAOKE DEVICE, AND SERVER DEVICE

(54) 発明の名称: 演奏評価装置、カラオケ装置及びサーバ装置

[図7]



(57) Abstract: This performance evaluation device is provided with: an expressiveness reference data acquisition means that acquires expressiveness reference data that shows expressiveness that should be achieved during the performance of a piece of music and timing at which the expressiveness should be achieved for that piece of music with utterance start timing for notes or note groups included in the piece of music as a standard; a pitch and volume data generating means that generates pitch and volume data showing the pitch and volume of the performance sound from the performance sound of the piece of music by a performer; and a performance evaluation means that improves evaluation of the performance of the piece of music by the performer when at least one of the characteristics of pitch and volume shown by the pitch and volume data generated by the pitch and volume data generating means exhibits expressiveness characteristics that should be achieved according to the expressiveness reference data within a prescribed time range shown by the expressiveness reference data for the piece of music.

(57) 要約: 楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得手段と、演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成手段と、前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価手段と、を備える演奏評価装置。

WO 2013/157602 A1



(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG). 添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：演奏評価装置、カラオケ装置及びサーバ装置

技術分野

[0001] この発明は、楽曲演奏の巧拙を評価する技術に関する。

背景技術

[0002] 例えば、歌唱者の歌唱演奏の巧拙を採点する採点機能を備えた歌唱用のカラオケ装置（以下、特に断らない限り、単に「カラオケ装置」という）に関わる技術が各種提案されている。この種の技術を開示した文献として、特許文献1がある。同文献に開示されたカラオケ装置は、利用者の歌唱音から抽出したピッチとガイドメロディとしてあらかじめ準備されたデータから抽出したピッチとの差分を歌唱曲のノート毎に算出し、この差分に基づいて基本得点を算出する。また、このカラオケ装置は、ビブラートやしゃくりなどの技法を駆使した歌唱が行われた場合にはその歌唱が行われた回数に応じたボーナスポイントを算出する。このカラオケ装置は、基本得点とボーナスポイントの合計点を最終的な評価結果として利用者に提示する。この技術によると、ビブラートやしゃくりなどといった難度の高い技法を駆使した歌唱を評価結果に反映させることができる。

[0003] また、歌唱音を示す波形から、ビブラートやしゃくりなどの技法を用いた歌唱が行われたことを検出する技術を開示した文献として、例えば特許文献2乃至6がある。

先行技術文献

特許文献

- [0004] 特許文献1：日本国特開2005-107334号公報
特許文献2：日本国特開2005-107330号公報
特許文献3：日本国特開2005-107087号公報
特許文献4：日本国特開2008-268370号公報
特許文献5：日本国特開2005-107336号公報

特許文献6：日本国特開2008-225115号公報

発明の概要

発明が解決しようとする課題

- [0005] しかしながら、特許文献1の技術の場合、本来であればビブラートやしゃくりなどの技法を駆使した歌唱を行うことが好ましくない歌唱箇所についてそのような歌唱が行われた場合であっても、ボーナスポイントが加算されてしまう。このため、評価結果として提示される得点が人間の感性によるものと乖離してしまうという問題があった。
- [0006] 本発明は、このような課題に鑑みてなされたものであり、カラオケ歌唱等の楽曲演奏の評価において、人間の感性によるものにより近い評価結果を提示でき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

- [0007] 上記課題を解決するため、本発明は、楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得手段と、演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成手段と、前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価手段とを備える演奏評価装置を提供する。
- [0008] また、本発明は、上記の演奏評価装置と、楽曲の伴奏を指示する伴奏データを取得する伴奏データ取得手段と、前記伴奏データの指示に従い伴奏の楽音を示す音信号を出力する音信号出力手段とを備え、前記ピッチ音量データ生成手段は、前記音信号出力手段から出力された音信号に従いスピーカから

放音された伴奏に応じて前記演奏者により行われた前記楽曲の演奏音のピッチおよび音量を示すピッチ音量データを生成するカラオケ装置を提供する。

[0009] また、本発明は、任意数の任意の演奏者による楽曲の演奏音の各々に関し、前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準とする一のタイミングにおいて一の表情演奏が出現したことを示す表情演奏出現データを取得する表情演奏出現データ取得手段と、前記表情演奏出現データ取得手段により取得された任意数の表情演奏出現データに基づき、前記楽曲に含まれるノートまたはノート群の各々に関し、当該ノートまたはノート群の発音開始時刻を基準とするいずれのタイミングでいずれの表情演奏がいずれの頻度で出現しているかを特定し、当該特定した情報に従い、前記楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを生成する表情演奏リファレンスデータ生成手段と、前記表情演奏リファレンスデータ生成手段により生成された表情演奏リファレンスデータを演奏評価装置に送信する送信手段とを備えるサーバ装置を提供する。

また、本発明は、歌唱評価システムであって、楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す第一表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得手段と、演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成手段と、前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記第一表情演奏リファレンスデータにより示される所定時間範囲内において前記第一表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価手段と、任意数の任意の演奏者による楽曲の演奏音の各々に

関し、前記任意の演奏者による前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準とする一のタイミングにおいて一の表情演奏が出現したことを示す表情演奏出現データを取得する表情演奏出現データ取得手段と、前記表情演奏出現データ取得手段により取得された任意数の表情演奏出現データに基づき、前記任意の演奏者による楽曲に含まれるノートまたはノート群の各々に関し、当該ノートまたはノート群の発音開始時刻を基準とするいずれのタイミングでいずれの表情演奏がいずれの頻度で出現しているかを特定し、当該特定した情報に従い、前記任意の演奏者による楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記任意の演奏者による楽曲において行われるべきタイミングを前記任意の演奏者による楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す第二表情演奏リファレンスデータを生成する表情演奏リファレンスデータ生成手段と、を備える歌唱評価システムを提供する。

また、本発明は、楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得し、演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成し、前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価方法を提供する。

また、本発明は、コンピュータが実行可能なプログラムであって、楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得処理と、演奏者による前記楽曲の演奏音から当該演奏音のピ

ッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成処理と、前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価処理とを前記コンピュータに実行させるプログラムを提供する。

発明の効果

- [0010] 本発明によれば、個々の楽曲の演奏において、望ましいタイミングで望ましい表情演奏が行われると、演奏者に対し高い評価を与える演奏評価装置が実現される。その結果、演奏者により表情演奏が行われた場合、人間の感性との乖離の少ない評価がなされる。

図面の簡単な説明

- [0011] [図1]この発明の一実施形態である歌唱評価システムの構成を示す図である。
[図2]タメの歌唱音の波形を示す図である。
[図3]ビブラートの歌唱音の波形を示す図である。
[図4]コブシの歌唱音の波形を示す図である。
[図5]シャクリの歌唱音の波形を示す図である。
[図6]フォールの歌唱音の波形を示す図である。
[図7]この発明の一実施形態である歌唱評価システムの動作を示すフローチャートである。
[図8]タメについて生成した統計データの一例である。
[図9]ビブラートについて生成した統計データの一例である。
[図10]コブシについて生成した統計データの一例である。
[図11]シャクリについて生成した統計データの一例である。
[図12]フォールについて生成した統計データの一例である。
[図13]本発明の演奏評価装置を示すブロック図である。

発明を実施するための形態

[0012] 以下、図面を参照し、この発明の実施の形態を説明する。

図1は、本発明の一実施形態である歌唱評価システム1の構成を示す図である。この歌唱評価システム1は、カラオケ装置10-m ($m=1, 2 \dots M$: Mはカラオケ装置の総数)とサーバ装置30とを有する。カラオケ装置10-mは、各カラオケ店に一台または複数台ずつ設置される。サーバ装置30は、システム運営センタ内に設置される。カラオケ装置10-mとサーバ装置30はネットワーク90に接続され、互いに各種データの送受信が可能である。

[0013] カラオケ装置10-mは、利用者の歌唱を支える伴奏曲の放音と歌詞の表示とを通じた歌唱演出と、利用者の歌唱の巧拙の評価とを行う装置である。ここで、カラオケ装置10-mは、歌唱の巧拙の評価では、利用者の歌唱音のピッチ及び音量の良否を評価対象とする評価と、以下に示す5種類の表情歌唱の良否を評価対象とする評価とを行い、2つの評価の評価結果である得点をコメントメッセージとともに利用者に提示する。

a 1. タメ

これは、歌唱曲内の特定の音の歌いだしを故意に遅らせる表情歌唱である。図2に示すように、この歌唱が行われた場合、歌唱音の前の音のものから当該音のものへと音のピッチが変化する時刻が楽譜（模範的な歌唱）における両音に対応する2つのノート（音符）の遷移時刻よりも僅かな時間だけ遅れる。

b 1. ビブラート

これは、歌唱曲内の特定の音を見かけのピッチを保ちつつ細かく震わせる表情歌唱である。図3に示すように、この歌唱が行われた場合、歌唱音のピッチは楽譜におけるその音に対応するノートの高さを跨いで周期的に変化する。

c 1. コブシ

これは、歌唱曲内の特定の音の声色を発音の途中でうなるように変化させ

る表情歌唱である。図4に示すように、この歌唱が行われた場合、歌唱音のピッチは楽譜におけるその音に対応するノートの途中で一過的に上昇する。

d 1. シャクリ

これは、歌唱曲内の特定の音を本来の高さよりも低い声で発音してから本来の高さに近づけていく歌唱手法である。図5に示すように、この歌唱が行われた場合、歌唱音の発音開始時刻におけるピッチは楽譜におけるその音に対応するノートの高さよりも低くなる。そして、この歌唱音のピッチは発音開始後に緩やかに上昇してノートの高さとはほぼ同じ高さに達する。

e 1. フォール

これは、歌唱曲内の特定の音を本来の高さよりも高い声で発音してから本来の高さに近づけていく歌唱手法である。図6に示すように、この歌唱が行われた場合、歌唱音の発音始時刻におけるピッチは楽譜におけるその音に対応するノートの高さよりも高くなる。そして、この歌唱音のピッチは発音開始後に緩やかに下降してノートの高さとはほぼ同じ高さに達する。

[0014] 図1に戻り、歌唱評価システム1全体の説明を続ける。カラオケ装置10ーmは、音源11、スピーカ12、マイクロホン13、表示部14、通信インターフェース15、ボーカルアダプタ16、CPU17、RAM18、ROM19、ハードディスク20、シーケンサ21を有する。音源11は、MIDI (Musical Instrument Digital Interface) の各種メッセージに従った音信号 S_A を出力する。スピーカ12は、与えられた信号を音として放音する。マイクロホン13は、音を收音して收音信号 S_M を出力する。表示部14は、画像信号 S_I に応じた画像を表示する。通信インターフェース15は、ネットワーク90に接続された装置との間でデータを送受信する。

[0015] ボーカルアダプタ16は、音信号 S_M のピッチ及び音量を測定し、それらの時間的な変化を示すピッチ音量データを生成する。具体的には、ボーカルアダプタ16は、マイクロホン13から与えられた音信号 S_M のピッチを時間 T_s (例えば、 $T_s=30$ ミリ秒とする) 毎に検出し、この検出結果を信号 S_p として出力する。また、ボーカルアダプタ16は、マイクロホン13から与えら

れた音信号 S_M の音量を時間 T_s 毎に検出し、この検出結果を信号 S_L として出力する。

[0016] CPU 17 は、RAM 18 をワークエリアとして利用しつつ ROM 19 やハードディスク 20 に記憶されたプログラムを実行する。この CPU 17 の動作の詳細は後述する。ROM 19 には、IPL (Initial Program Loader) などが記憶されている。ハードディスク 20 には、各種歌唱曲の曲データ $MD-n$ ($n=1\sim N$) (N は、歌唱曲の種類の総数)、リファレンスデータベース DBRK、及び歌唱評価プログラム VPG が記憶されている。各歌唱曲の曲データ $MD-n$ は、歌唱曲の伴奏内容、歌唱曲の歌詞、及び歌唱曲の模範的な歌唱内容を SMF (Standard MIDI File) 形式で記録したデータである。

[0017] 具体的に説明すると、図 1 の枠内に示すように、曲データ $MD-n$ は、ヘッダ HD、伴奏トラック TR_{AC} 、歌詞トラック TR_{LY} 、模範歌唱リファレンストラック TR_{NR} を有している。ヘッダ HD には、曲番号、曲名、ジャンル、演奏時間、タイムベース (4 分音符 1 つ分の時間に相当するティック数) などの情報が記述されている。

[0018] 伴奏トラック TR_{AC} には、歌唱曲の楽譜の伴奏パートにおける各ノート $NT(i)$ (i は、楽譜の該当パートの先頭のノート $NT(1)$ から数えた順番を示す) の音の発音を指示するイベント $EV(i)_{ON}$ とその消音を指示するイベント $EV(i)_{OFF}$ 、及び相前後するイベントの実行時間差 (ティック数) を示すデルタタイム DT が時系列順に記述されている。

[0019] 歌詞トラック TR_{LY} には、歌唱曲の歌詞を示す各データ D_{LY} と、各歌詞の表示時刻 (より具体的には、各歌詞の表示時刻と各々の前の歌詞の表示時刻との間の時間差 (ティック数)) を示すデルタタイム DT が時系列順に記述されている。

[0020] 模範歌唱リファレンストラック TR_{NR} には、歌唱曲の楽譜の歌唱パートにおける各ノート $NT(i)$ の音の発音を指示するイベント $EV(i)_{ON}$ とその消音を指示するイベント $EV(i)_{OFF}$ 、及び相前後するイベントの実行時間差 (

ティック数)を示すデルタタイム DT が時系列順に記述されている。

[0021] リファレンスデータベース $DBRK$ には、5種類の表情歌唱リファレンスデータ DD_{a1} 、 DD_{a2} 、 DD_{a3} 、 DD_{a4} 、 DD_{a5} が記憶されている。表情歌唱リファレンスデータ DD_{a1} は、歌唱曲に含まれるノート $NT(i)$ の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t においてタメによる歌唱が行われた場合の評価点 $VSR(t)$ の各対を示すデータである。表情歌唱リファレンスデータ DD_{a2} は、歌唱曲に含まれるノート $NT(i)$ の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t においてビブラートによる歌唱が行われた場合の評価点 $VSR(t)$ の各対を示すデータである。表情歌唱リファレンスデータ DD_{a3} は、歌唱曲に含まれるノート $NT(i)$ の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t においてコブシによる歌唱が行われた場合の評価点 $VSR(t)$ の各対を示すデータである。表情歌唱リファレンスデータ DD_{a4} は、歌唱曲に含まれるノート $NT(i)$ の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t においてシャクリによる歌唱が行われた場合の評価点 $VSR(t)$ の各対を示すデータである。表情歌唱リファレンスデータ DD_{a5} は、歌唱曲に含まれるノート $NT(i)$ の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t においてフォールによる歌唱が行われた場合の評価点 $VSR(t)$ の各対を示すデータである。以下では、5種類の表情歌唱リファレンスデータ DD_{a1} 、 DD_{a2} 、 DD_{a3} 、 DD_{a4} 、 DD_{a5} を区別しない場合は表情歌唱リファレンスデータ DD と記す。

[0022] 歌唱評価プログラム VPG は、次の3つの機能を有する。

a 2. 標準評価機能

これは、ボーカルアダプタ16の出力信号 S_L 及び S_P が示すピッチ及び音量と模範歌唱リファレンストラック TR_{NR} 内の各イベント $EV(i)_{ON}$ 及び $EV(i)_{OFF}$ により決まる各ノート $NT(i)$ の模範ピッチ PCH_{REF} 及び模範音量 LV_{REF} とを比較し、この比較の結果に基づいて歌唱の巧拙を評価する機能である。

b 2. 表情歌唱評価機能

これは、ボーカルアダプタ 16 の出力信号 S_p が示すピッチ波形に表情歌唱の特徴波形が出現する度に、表情歌唱の対象となったノート $NT(i)$ の発音開始時刻を基準点 t_{BS} とする時間軸上における表情歌唱の特徴波形の出現時刻を求め、この出現時刻と対応する評価点 $VSR(t)$ をリファレンスデータベース $DBRK$ 内における該当の表情歌唱リファレンスデータ DD の各評価点 $VSR(t)$ の中から選択し、この評価点 $VSR(t)$ に基づいて歌唱の巧拙を評価する機能である。

c 2. 評価結果提示機能

これは、a 2 による評価の評価結果及び b 2 による評価の評価結果から得点を算出し、この得点をコメントメッセージとともに表示部 14 に表示させる機能である。

[0023] シーケンサ 21 は、リモートコントローラ（不図示）による歌唱曲の歌唱開始操作を契機として該当曲の曲データ $MD-n$ がハードディスク 20 から $RAM18$ に転送された場合に、その曲データ $MD-n$ 内のイベント $EV(i)_{ON}$ 、 $EV(i)_{OFF}$ 、およびデータ D_{LY} を装置各部に供給する。具体的には、シーケンサ 21 は、 $RAM18$ に曲データ $MD-n$ が記憶されると、この曲データ $MD-n$ のヘッダ HD に記述されたタイムベースとリモートコントローラ（不図示）により指定されたテンポとに基づいて 1 ティックの時間長を決定し、この時間長の経過に合わせてティックをカウントしつつ次の 3 つの処理を行う。

[0024] 第 1 の処理では、シーケンサ 21 は、ティックのカウント数が伴奏トラック TR_{AC} 内のデルタタイム DT と一致する度にそれに後続するイベント $EV(i)_{ON}$ （または $EV(i)_{OFF}$ ）を読み出して音源 11 に供給する。音源 11 は、シーケンサ 21 からイベント $EV(i)_{ON}$ が供給されるとそのイベント $EV(i)_{ON}$ が指定する音信号 S_A をスピーカ 12 に供給し、シーケンサ 21 からイベント $EV(i)_{OFF}$ が供給されるとスピーカ 12 への音信号 S_A の供給を止める。

- [0025] 第2の処理では、シーケンサ21は、ティックのカウント数が歌詞トラック TR_{LY} 内のデルタタイム DT と一致する度にそれに後続するデータ D_{LY} を読み出して表示部14に供給する。表示部14は、シーケンサ21からデータ D_{LY} が供給されるとそのデータ D_{LY} を歌詞テロップの画像に変換し、この画像をディスプレイ（不図示）に表示させる。
- [0026] シーケンサ21がこの第1および第2の処理を行うことにより、スピーカ12からの伴奏音の放音とディスプレイへの歌詞の表示とが進行する。利用者は、スピーカ12から放音される伴奏音を聴きつつディスプレイに表示された歌詞をマイクロホン13に向かって歌唱する。利用者がマイクロホン13に向かって歌唱している間、マイクロホン13は利用者の歌唱音の收音信号 S_M を出力し、ボーカルアダプタ16はこの信号 S_M のピッチ及び音量を示す信号 S_P 及び S_L を出力する。
- [0027] 第3の処理では、シーケンサ21は、ティックのカウント数が模範歌唱リファレンストラック TR_{NR} 内のデルタタイム DT と一致する度にそれに後続するイベント $EV(i)_{ON}$ （または $EV(i)_{OFF}$ ）を読み出してCPU17に供給する。CPU17は、シーケンサ21から供給されるイベント $EV(i)_{ON}$ 及び $EV(i)_{OFF}$ とボーカルアダプタ16の出力信号 S_P 及び S_L とを用いて利用者の歌唱の巧拙を評価する。詳しくは、後述する。
- [0028] サーバ装置30は、カラオケ店舗におけるサービスの提供を支援する役割を果たす装置である。サーバ装置30は、通信インターフェース35、CPU37、RAM38、ROM39、ハードディスク40を有する。通信インターフェース35は、ネットワーク90に接続された装置との間でデータを送受信する。CPU37は、RAM38をワークエリアとして利用しつつ、ROM39やハードディスク40に記憶された各種プログラムを実行する。このCPU37の動作の詳細は後述する。ROM39にはIPLなどが記憶されている。
- [0029] ハードディスク40には、歌唱サンプルデータベースDBS、リファレンスデータベースDBRS、および歌唱分析プログラムAPGが記憶されてい

る。歌唱サンプルデータベースDBSには、各々が1つの歌唱曲と対応する歌唱サンプルデータDS群が個別に記憶される。歌唱サンプルデータDSは、一定水準以上の歌唱力を有する者が歌唱曲を歌唱したときの歌唱音のピッチ波形及び音量波形を記録したデータである。リファレンスデータベースDBRSには、各カラオケ装置10-mのリファレンスデータベースDBRK内に格納されるべき最新の表情歌唱リファレンスデータDDが記憶される。

[0030] 歌唱分析プログラムAPGは、次の3つの機能を有する。

a 3. 蓄積機能

これは、カラオケ装置10-mから各歌唱曲の歌唱サンプルデータDSを1曲分ずつ取得し、取得した歌唱サンプルデータDSを歌唱サンプルデータベースDBSに蓄積する機能である。

b 3. 書き換え機能

これは、歌唱サンプルデータベースDBSに蓄積された歌唱サンプルデータDSの各々について、当該歌唱サンプルデータDSが示す波形内から表情歌唱の特徴波形を探索し、この探索結果から表情歌唱の対象となったノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t における表情歌唱の出現数Numとの関係を示す統計データを生成し、リファレンスデータベースDBR内の表情歌唱リファレンスデータDDにおける各時刻 t と対応する評価点VSR(t)を統計データの内容に基づいて書き換える機能である。

c 3. 送信機能

これは、書き換え機能により書き換えた表情歌唱リファレンスデータDDをカラオケ装置10-mからの要求に応じてカラオケ装置10-mに送信する機能である。

[0031] 次に、本実施形態の動作を説明する。図7は、本実施形態の動作を示すフローチャートである。図7において、カラオケ装置10-mのCPU17は、歌唱曲の歌唱開始操作が行われた場合(S100:Yes)、シーケンサ21に制御信号 S_0 を供給してシーケンサ21に処理(上述した第1～第3の

処理)を開始させる(S120)。CPU17は、シーケンサ21による処理が始まると、標準歌唱評価処理(S130)と表情歌唱評価処理(S140)の2つの処理を行う。この2つの処理の詳細は次の通りである。

[0032] a 4. 標準歌唱評価処理(S130)

この処理では、CPU17は、シーケンサ21からイベント $EV(i)_{ON}$ が供給されてから次のイベント $EV(i)_{OFF}$ が供給されるまでの時間を i 番目のノート $NT(i)$ に相当する音の発音時間 $T_{NT}(i)$ とする。CPU17は、発音時間 $T_{NT}(i)$ の間のボーカルアダプタ16の出力信号 S_P が示すピッチとイベント $EV(i)_{ON}$ のノートナンバを変換した模範ピッチ PCH_{REF} との差 PCH_{DEF} 、及びその間の信号 S_P が示す音量とイベント $EV(i)_{ON}$ のベロシティを変換した模範音量 LV_{REF} との差 LV_{DEF} を求め、この差 PCH_{DEF} 及び差 LV_{DEF} が所定範囲に収まる場合にノート $NT(i)$ の歌唱が合格であると判定する。CPU17は、利用者による歌唱の開始から終了までの間に亘ってこのノート判定を行い、歌唱の終了時点における全ノート $TN(i)$ の数を合格と判定したノート $NT(i)$ の数で除算した値に100を乗じた値を基本得点 SR_{BASE} とする。

[0033] また、この処理では、CPU17は、ボーカルアダプタ16の出力信号 S_P が示すピッチ波形内に、タメ、ビブラート、コブシ、シャクリ、フォールのいずれかの表情歌唱の特徴波形が出現した否かを判定する。ここで、タメの特徴波形の判定手法の詳細は特許文献2を、ビブラートの特徴波形の判定手法の詳細は特許文献3を、コブシの特徴波形の判定手法の詳細は特許文献4を、シャクリの特徴波形の判定手法の詳細は特許文献5を、フォールの特徴波形の判定手法の詳細は特許文献6を参照されたい。CPU17は、利用者による歌唱の開始から終了までの間に亘ってこの特徴波形判定を行い、歌唱の終了時点における表情歌唱の出現数に所定の係数を乗じた値を加算点 SR_{ADD} とする。そして、この処理では、基本得点 SR_{BASE} と加算点 SR_{ADD} の合計を標準得点 SR_{NOR} とする。

[0034] b 4. 表情歌唱評価処理(S140)

この処理では、CPU 17は、音源イベント $EV(i)_{ON}$ の出力から次のイベント $EV(i)_{OFF}$ の出力までの時間を i 番目のノート $NT(i)$ に相当する音の発音時間 $T_{NT}(i)$ とする。そして、CPU 17は、発音時間 $T_{NT}(i)$ の間のボーカルアダプタ 16の出力信号 S_p が示すピッチ波形内に表情歌唱の特徴波形が出現した場合には、発音時間 $T_{NT}(i)$ 内における表情歌唱の出現時刻と出現した表情歌唱の種類を求める。CPU 17は、そのように特定した表情歌唱の種類と出現時刻とを示す表情歌唱出現データを生成する。

[0035] そして、CPU 17は、生成した表情歌唱出現データに示される表情歌唱およびその出現時刻に応じた評価点 $VSR(t)$ を、表情歌唱リファレンスデータ DD が示す一連の評価点 $VSR(t)$ の中から選択する。CPUは、利用者による歌唱の開始から終了までの間に亘ってこのような評価点 $VSR(t)$ の選択を行い、歌唱の終了時点における評価点 $VSR(t)$ の平均値を表情得点 SR_{EX} とする。

[0036] CPU 17は、利用者による歌唱曲の歌唱が終了すると、評価結果提示処理を行う(S 150)。評価結果提示処理では、CPU 17は、標準歌唱評価処理により採点した標準得点 SR_{NOR} と表情歌唱評価処理により採点した表情得点 SR_{EX} のうち高い方の得点を選択する。そして、CPU 17は、標準得点 SR_{NOR} を選択した場合、この得点 SR_{NOR} と、例えば「クールで精緻な歌ですね」といった得点 SR_{NOR} に応じたコメントメッセージを表示部 14に表示させる。また、CPU 17は、表情得点 SR_{EX} を選択した場合、この得点 SR_{EX} と、例えば「人情味あふれていますね」といった表情得点 SR_{EX} に応じたコメントメッセージを表示部 14に表示させる。

[0037] 次に、CPU 17は、サンプル送信処理を行う(S 160)。サンプル送信処理では、CPU 17は、歌唱曲の歌唱の開始から終了までの間にボーカルアダプタ 16が出力した信号 S_p 及び S_L を当該歌唱曲の歌唱サンプルデータ DS とし、この歌唱サンプルデータ DS とステップ S 130で求めた基本得点 SR_{BASE} (歌唱評価データ)とを含むメッセージ $MS1$ をサーバ装置 30に送信する。

[0038] サーバ装置30のCPU37は、カラオケ装置10-mからメッセージMS1を取得すると(S200:Yes)、このメッセージMS1から歌唱サンプルデータDSと基本得点 SR_{BASE} を取り出し、この基本得点 SR_{BASE} を上級者とそうでない者とを分ける基準得点 SR_{TH} (たとえば、80点とする)と比較する(S220)。CPU37は、基本得点 SR_{BASE} が基準得点 SR_{TH} よりも高い場合(S220:Yes)、メッセージMS1から取り出した歌唱サンプルデータDSを歌唱サンプルデータベースDBSに蓄積する(S230)。

[0039] 続いて、CPU37は書き換え処理を行う(S240)。書き換え処理では、CPU37は、次の5つの処理を行う。第1の処理では、CPU37は、歌唱サンプルデータベースDBSに蓄積された各歌唱サンプルデータDSが示すピッチ波形内からタメの特徴波形を探索し、この探索結果を示す表情歌唱出現データ(タメが出現したノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t を示すデータ)を生成する。続いて、CPU37はタメに関し生成した表情歌唱出現データに基づき、ノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t における表情歌唱「タメ」の出現数 N_{um} との関係を示す統計データを生成し、表情歌唱リファレンスデータ DD_{a1} における各時刻 t と対応する評価点 $VS R(t)$ をこの統計データの内容に基づいて書き換える。

[0040] 図8は、タメについての統計データの一例を示す図である。この例の統計データでは、基準点 t_{BS} より時間 T_{1a1} だけ前の時刻 t_{1a1} と基準点 t_{BS} より時間 T_{4a1} だけ後の時刻 t_{4a1} との間に表情歌唱の出現数 N_{um} が分布している。そして、この例の統計データでは、基準点 t_{BS} の直後の時刻 t_{2a1} に出現数 N_{um} の最大ピークが表れており、時刻 t_{2a1} よりも遅れた時刻 t_{3a1} に出現数 N_{um} の2番目のピークが表れている。よって、この例の統計データによる書き換え後の表情歌唱リファレンスデータ DD_{a1} では、時刻 t_{2a1} の評価点 $VS R(t_{2a1})$ が最も高くなり、時刻 t_{3a1} の評価点 $VS R(t_{3a1})$ が2番目に高くなる。

[0041] 第2の処理では、CPU37は、歌唱サンプルデータベースDBSに蓄積された各歌唱サンプルデータDSが示すピッチ波形内からビブラートの特徴波形を探索し、この探索結果を示す表情歌唱出現データ（ビブラートが出現したノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t を示すデータ）を生成する。続いて、CPU37はビブラートに関し生成した表情歌唱出現データに基づき、ノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t における表情歌唱の出現数 N_{um} との関係を示す統計データを生成し、表情歌唱リファレンスデータ DD_{a2} における各時刻 t と対応する評価点 $VSR(t)$ をこの統計データの内容に基づいて書き換える。

[0042] 図9は、ビブラートについての統計データの一例を示す図である。この例の統計データでは、基準点 t_{BS} と基準点 t_{BS} より時間 T_{2a2} だけ後の時刻 t_{2a2} との間に表情歌唱の出現数 N_{um} が分布している。そして、この例の統計データでは、基準点 t_{BS} よりも時間 T_{1a2} だけ後の時刻 t_{1a2} に出現数 N_{um} の最大ピークが表れている。よって、この例の統計データによる書き換え後の表情歌唱リファレンスデータ DD_{a2} では、時刻 t_{1a2} の評価点 $VSR(t_{1a2})$ が最も高くなる。

[0043] 第3の処理では、CPU37は、歌唱サンプルデータベースDBSに蓄積された各歌唱サンプルデータDSが示すピッチ波形内からコブシの特徴波形を探索し、この探索結果を示す表情歌唱出現データ（コブシが出現したノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t を示すデータ）を生成する。続いて、CPU37はコブシに関し生成した表情歌唱出現データに基づき、ノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t における表情歌唱の出現数 N_{um} との関係を示す統計データを生成し、表情歌唱リファレンスデータ DD_{a3} における各時刻 t と対応する評価点 $VSR(t)$ をこの統計データの内容に基づいて書き換える。

[0044] 図10は、コブシについての統計データの一例を示す図である。この例の

統計データでは、基準点 t_{BS} と基準点 t_{BS} より時間 $T_{2_{a3}}$ だけ後の時刻 $t_{2_{a3}}$ との間に表情歌唱の出現数 N_{um} が分布している。そして、この例の統計データでは、基準点 t_{BS} よりも時間 $T_{1_{a3}}$ だけ後の時刻 $t_{1_{a3}}$ に出現数 N_{um} の最大ピークが表れている。よって、この例の統計データによる書き換え後の表情歌唱リファレンスデータ DD_{a3} では、時刻 $t_{1_{a3}}$ の評価点 $VSR(t_{1_{a3}})$ が最も高くなる。

[0045] 第4の処理では、CPU37は、歌唱サンプルデータベースDBSに蓄積された各歌唱サンプルデータDSが示すピッチ波形内からシャクリの特徴波形を探索し、この探索結果を示す表情歌唱出現データ（シャクリが出現したノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t を示すデータ）を生成する。続いて、CPU37はシャクリに関し生成した表情歌唱出現データに基づき、ノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t における表情歌唱の出現数 N_{um} との関係を示す統計データを生成し、表情歌唱リファレンスデータ DD_{a4} における各時刻 t と対応する評価点 $VSR(t)$ をこの統計データの内容に基づいて書き換える。

[0046] 図11は、シャクリについての統計データの一例を示す図である。この例の統計データでは、基準点 t_{BS} と基準点 t_{BS} より時間 $T_{2_{a4}}$ だけ後の時刻 $t_{2_{a4}}$ との間に表情歌唱の出現数 N_{um} が分布している。そして、この例の統計データでは、基準点 t_{BS} に出現数 N_{um} の最大ピークが表れており、基準点 t_{BS} よりも時間 $T_{1_{a4}}$ だけ遅れた時刻 $t_{1_{a4}}$ に出現数 N_{um} の2番目のピークが表れている。よって、この例の統計データによる書き換え後の表情歌唱リファレンスデータ DD_{a4} では、時刻 t_{BS} の評価点 $VSR(t_{BS})$ が最も高くなり、時刻 $t_{1_{a4}}$ の評価点 $VSR(t_{1_{a4}})$ が2番目に高くなる。

[0047] 第5の処理では、CPU37は、歌唱サンプルデータベースDBSに蓄積された各歌唱サンプルデータDSが示すピッチ波形内からフォールの特徴波形を探索し、この探索結果を示す表情歌唱出現データ（フォールが出現したノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t を示

すデータ)を生成する。続いて、CPU37はフォールに関し生成した表情歌唱出現データに基づき、ノートNT(i)の発音開始時刻を基準点 t_{BS} とする時間軸上の各時刻 t とそれらの時刻 t における表情歌唱の出現数Numとの関係を示す統計データを生成し、表情歌唱リファレンスデータDD_{a5}における各時刻と対応する評価点VSR(t)をこの統計データの内容に基づいて書き換える。

[0048] 図12は、フォールについての統計データの一例を示す図である。この例の統計データでは、基準点 t_{BS} より時間 T_{1a5} だけ後の時刻 t_{1a5} と時刻 t_{BS} から時間 T_{2a5} だけ後の時刻 t_{2a5} との間に表情歌唱の出現数Numが分布している。そして、この例の統計データでは、時刻 t_{2a5} に出現数Numの最大ピークが表れている。よって、この例の統計データによる書き換え後の表情歌唱リファレンスデータDD_{a5}では、時刻 t_{2a5} の評価点VSR(t_{2a5})が最も高くなる。

[0049] 図7において、カラオケ装置10-mのCPU17は、予め決められた問合せ時刻が到来する度に(S110:Yes)、問合せ処理を行う(S170)。この問合せ処理では、CPU17は、最新データの送信を求めるメッセージMS2をサーバ装置30に送信する(S170)。サーバ装置30のCPU37は、カラオケ装置10-mからメッセージMS2を受信すると(S210:Yes)、前回のメッセージMS2の受信時刻から今回のメッセージMS2の受信時刻までの間に内容を書き換えた表情歌唱リファレンスデータDDをメッセージMS2の送信元のカラオケ装置10-mに送信する(S250)。カラオケ装置10-mのCPU17は、サーバ装置30から表情歌唱リファレンスデータDDを受信すると、この表情歌唱リファレンスデータDDをリファレンスデータベースDBRKに上書きしてその内容を更新する(S180)。

[0050] 以上が、本実施形態の構成の詳細である。本実施形態によると、次の効果が得られる。

第1に、本実施形態の表情付け歌唱評価処理では、ボーカルアダプタ16

の出力信号の波形に表情歌唱の特徴波形が出現する度に、表情歌唱の対象となったノート $N T(i)$ の発音開始時刻を基準点とする時間軸上における表情歌唱の特徴波形の出現時刻を求め、この出現時刻と対応する評価点 $V S R(t)$ を歌唱リファレンスデータ $D D$ 内の各評価点 $V S R(t)$ の中から選択し、この選択した評価点 $V S R(t)$ に基づいて歌唱の巧拙を評価する。よって、本実施形態によると、利用者が表情歌唱を行ったとしても、そのタイミングが適切でなければ良好な評価が得られないことになる。従って、本実施形態によると、人の感性によるものにより近い評価結果を提示することができる。

[0051] 第2に、本実施形態では、歌唱サンプルデータベース $D B S$ 内に蓄積された表情歌唱リファレンスデータ $D D$ の各々について、当該データ $D D$ が示す波形内から表情歌唱の特徴波形を探索し、この探索結果から表情歌唱の対象となったノート $N T(i)$ の発音開始時刻を基準点とする時間軸上の各時刻とそれらの時刻における表情歌唱の出現数との関係を示す統計データを生成し、歌唱リファレンスデータ $D D$ における各時刻と対応する評価点 $V S R(t)$ を統計データの内容に基づいて書き換える。よって、本実施形態によると、歌唱曲を歌い込んでいる上級者らの歌い方の傾向の変化を評価結果に反映させることができる。

[0052] 以上、この発明の一実施形態について説明したが、この発明には他にも実施形態があり得る。例えば、以下の通りである。

(1) 上記実施形態では、 $C P U 17$ は、タメ、ビブラート、コブシ、シャクリ、フォールの5種類の表情歌唱をボーカルアダプタ 16 の出力信号 S_p から検出した。しかし、この5種類以外の表情歌唱を検出してもよい。たとえば、抑揚をつけた歌唱を検出してもよい。

[0053] (2) 上記実施形態では、 $C P U 17$ は、ボーカルアダプタ 16 の出力信号 S_p 及び S_L の両方を用いて標準歌唱評価処理を行い、ボーカルアダプタ 16 の出力信号 S_p 及び S_L のうちピッチを示す信号 S_p のみを用いて表情歌唱評価処理を行った。しかし、 $C P U 17$ は、信号 S_p 及び S_L の一方のみを用いて標準歌

唱評価処理を行ってもよい。また、CPU 17は、信号 S_p 及び S_L の両方を用いて表情歌唱評価処理を行ってもよい。

[0054] (3) 上記実施形態の表情歌唱評価処理では、表情歌唱の特徴波形の出現時刻に基づいて歌唱の巧拙を評価した。しかし、表情歌唱の特徴波形の出現時刻以外の要素（たとえば、タメ、ビブラート、コブシ、シャクリ、フォールの各々の長さや深さなど）を加味した評価を行ってもよい。

[0055] (4) 上記実施形態の表情歌唱評価処理では、歌唱曲に含まれるノートの各々に応じた歌唱音において出現する表情歌唱を検出する構成が採用されているが、歌唱曲に含まれる一連の複数のノート（ノート群）に応じた歌唱音において出現する表情歌唱を検出する構成が採用されてもよい。例えば、クレッシェンド・デクレッシェンドのような表情歌唱は、一連の複数のノートの歌唱において行われる表情歌唱であるため、それらの表情歌唱の検出および評価はノート群を単位として行われる方が望ましい。従って、そのような表情歌唱に関する表情歌唱リファレンスデータDDもまたノート群単位で構成されることが望ましい。

[0056] (5) 上記実施形態では、カラオケ装置10からサーバ装置30に対し、歌唱曲の歌唱の開始から終了までの間にボーカルアダプタ16が出力した信号 S_p 及び S_L を含む歌唱サンプルデータDS（ピッチ音量データ）を送信し、サーバ装置30においては歌唱サンプルデータDSから各表情歌唱の検出およびその出現のタイミングの特定処理が行われる構成が採用されている。これに代えて、カラオケ装置10からサーバ装置30に対し、マイクロホン13により收音された音を示す音信号 S_M （歌唱音を示す音声波形データ）を送信し、サーバ装置30において音信号 S_M から信号 S_p および信号 S_L を生成する処理（上記実施形態におけるボーカルアダプタ16が行う処理）が行われる構成が採用されてもよい。また、カラオケ装置10からサーバ装置30に対し、歌唱評価プログラムVPGに従い行われる表情歌唱評価処理（S140）に際し特定した表情歌唱の種別およびその出現のタイミングを示すデータ（表情歌唱出現データ）を送信し、サーバ装置30においては表情歌唱の検出

処理は行わずカラオケ装置 10 から送信されてくる表情歌唱出現データに基づき表情歌唱リファレンスデータ DD の更新処理が行われる構成が採用されてもよい。

[0057] (6) 上記実施形態では、サーバ装置 30 が統計データの生成とこれに基づく表情歌唱リファレンスデータ DD の書き換えを行った。しかし、各カラオケ装置 10-m が過去に自機により生成、もしくは他のカラオケ装置 10-m から直接またはサーバ装置 30 を介して取得した歌唱音を示す音信号 S_M や、それらの音信号 S_M から生成した信号 S_p および信号 S_L 、もしくはそれらの信号を用いて特定した表情歌唱の種別およびその出現のタイミングを示すデータ（表情歌唱出現データ）をハードディスク 20 に記憶しておき、CPU 17 がそれらを読み出して用いて、サーバ装置 30 が S 240 で行う処理と同様の処理、すなわち統計データの生成とこれに基づく表情歌唱リファレンスデータ DD の書き換えを行うようにしてもよい。

[0058] (7) 上記実施形態における歌唱の評価の方法および評価結果の歌唱者への提示の態様は様々に変更可能である。例えば、上記実施形態においては、標準歌唱評価処理 (S 130) にて表情歌唱の出現回数に基づき算出される加算点 $S R_{ADD}$ を基本得点 $S R_{BASE}$ と合計することで標準得点 $S R_{NOR}$ を算出する構成が採用されているが、標準歌唱評価処理においては表情歌唱の出現は考慮せず、基本得点 $S R_{BASE}$ のみを算出する構成が採用されてもよい。また、上記実施形態においては、歌唱者に対し、標準歌唱評価処理により採点した標準得点 $S R_{NOR}$ と表情歌唱評価処理により採点した表情得点 $S R_{EX}$ のうち高い方の得点が表示されるが、それらの両方を表示する、それらの合計点数を表示するなど、他の態様で歌唱者に対する評価結果の提示が行われてもよい。

[0059] (8) 上記実施形態では、表情歌唱リファレンスデータ DD の更新に際し、基本得点 $S R_{BASE}$ が基準得点 $S R_{TH}$ よりも高い歌唱者を上級者とし、上級者に関する歌唱サンプルデータ DS のみを用いて表情歌唱リファレンスデータ DD の更新を行う構成が採用されている。表情歌唱リファレンスデータ DD の更新に用いる歌唱サンプルデータ DS の選択方法はこれに限られない。例えば

、基本得点 SR_{BASE} に代えて、基本得点 SR_{BASE} に加算点 SR_{ADD} を合計した標準得点 SR_{NOR} を上級者の推定の基準として用いてもよい。また、全く表情歌唱を行わないために基本得点 SR_{BASE} が高得点となっている上級者を除外するために、下側の閾値（基準得点 SR_{TH} ）に加え上側の閾値を設け、上側の閾値より高い基本得点 SR_{BASE} （またはその他の得点）の歌唱者の歌唱サンプルデータ DS は表情歌唱リファレンスデータ DD の更新には用いない、という構成が採用されてもよい。また、上記のように歌唱者を上級者とそれ以外の者に 2 分する代わりに、例えば基本得点 SR_{BASE} が高い歌唱者の歌唱サンプルデータ DS に大きい重み付けを付けて表情歌唱リファレンスデータ DD の更新に用いるようにしてもよい。

[0060] (9) 上記実施形態では、楽曲演奏を評価する演奏評価装置の一例として、歌唱用のカラオケ装置に設けられ、歌唱演奏を評価する演奏評価装置を示したが、本発明にかかる演奏評価装置は歌唱演奏の評価に限定されず、各種楽器を用いた楽曲演奏の評価にも適用可能である。すなわち、上記実施形態において用いた「歌唱」という言葉は、より一般的な「演奏」という言葉で置き換えられる。なお、器楽演奏を評価する演奏評価装置においては、例えばギターにおけるチョーキングなど、個々の楽器に応じた表情演奏に関する評価が行われることになる。また、楽曲が歌唱曲でなく楽器用の楽曲である場合、楽器演奏用のカラオケ装置は、曲データ MD は歌詞トラック TR_{LY} に代えて、例えば楽譜を示すデータと、楽譜の各区間（例えば、2 小節もしくは 4 小節のブロックなど）の表示時刻を示すデルタタイムが時系列順に記述されたデータである楽譜トラックを含むように構成され、シーケンサ 21 および表示部 14 は楽譜トラックに従い、楽曲の進行に伴い伴奏箇所に応じた楽譜を示す画像信号をディスプレイに出力するように構成されることになる。なお、歌唱用のカラオケ装置および楽器演奏用のカラオケ装置において、歌詞もしくは楽譜の表示が不要な場合は、シーケンサ 21 および表示部 14 による画像信号の出力処理は行われなくてもよい。

(10) 以上の例示から理解される通り、本発明の好適な態様に係る演奏評

価装置は、図 13 に例示される通り、楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得手段 101 と、演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成手段 102 と、前記ピッチ音量データ生成手段 102 により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価手段 103 と、を備える装置として包括的に表現され、他の要素の有無や他の要素の具体的な態様は任意である。

[0061] (11) 上記実施形態では、いわゆる専用機としてのカラオケ装置に本発明にかかる演奏評価装置が設けられている例を示したが、本発明にかかる演奏評価装置は専用機に限られない。例えば、パーソナルコンピュータや携帯情報端末（例えば携帯電話機やスマートフォン）やゲーム装置等の各種装置にプログラムに従った処理を行わせることによって本発明にかかる演奏評価装置を実現する構成が採用されてもよい。また、このプログラムは、CD-ROM等の記録媒体に格納して配布したり、インターネット等の電気通信回線を利用して配布したりすることも可能である。

本出願は、2012年4月18日出願の日本特許出願、特願2012-094853に基づくものであり、その内容はここに参照として取り込まれる。

産業上の利用可能性

[0062] 本発明によれば、演奏者により表情演奏が行われた場合、人間の感性との乖離の少ない評価を行うことが可能である。

符号の説明

[0063] 1…歌唱評価システム、10…カラオケ装置、11…音源、12…スピーカ、13…マイクロホン、14…表示部、15…通信インターフェース、16…ボーカルアダプタ、17…CPU、18…RAM、19…ROM、20…ハードディスク、21…シーケンサ、30…サーバ装置、35…通信インターフェース、37…CPU、38…RAM、39…ROM、40…ハードディスク、90…ネットワーク

請求の範囲

[請求項1]

楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得手段と、

演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成手段と、

前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価手段と

を備える演奏評価装置。

[請求項2]

任意数の任意の演奏者による楽曲の演奏音の各々に関し、当該演奏音のピッチおよび音量を示すピッチ音量データを取得するピッチ音量データ取得手段と、

前記ピッチ音量データ取得手段により取得された音量ピッチデータにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における任意のタイミングにおいて予め定められた1以上の表情演奏の特性のうちの一の特性を示す場合、当該表情演奏と、前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準とする当該タイミングとの対を示す表情演奏出現データを生成する表情演奏出現データ生成手段と、

前記表情演奏出現データ生成手段により生成された任意数の表情演奏出現データに基づき、前記楽曲に含まれるノートまたはノート群の各々に関し、当該ノートまたはノート群の発音開始時刻を基準とするいずれのタイミングでいずれの表情演奏がいずれの頻度で出現してい

るかを特定し、当該特定した情報に従い表情演奏リファレンスデータを生成する表情演奏リファレンスデータ生成手段と
を備える請求項 1 に記載の演奏評価装置。

[請求項3] 表情演奏リファレンスデータを記憶する表情演奏リファレンスデータ記憶手段

を備え、

前記表情演奏リファレンスデータ生成手段によって生成された前記表情演奏リファレンスデータに基づき、前記表情演奏リファレンスデータ記憶手段に記憶される前記表情演奏リファレンスデータは書き換えられる

請求項 2 に記載の演奏評価装置。

[請求項4] 前記楽曲の模範となるピッチを示す模範演奏リファレンスデータを取得する模範演奏リファレンスデータ取得手段

を備え、

前記演奏評価手段は、前記ピッチ音量データ生成手段により生成されたピッチ音量データにより示されるピッチと、前記模範演奏リファレンスデータにより示されるピッチとの比較の結果に基づき前記演奏者による前記楽曲の演奏に対する評価を行う

請求項 1 乃至 3 のいずれかに記載の演奏評価装置。

[請求項5] 前記楽曲の模範となるピッチを示す模範演奏リファレンスデータを取得する模範演奏リファレンスデータ取得手段

を備え、

前記演奏評価手段は、前記ピッチ音量データ生成手段により生成されたピッチ音量データにより示されるピッチと、前記模範演奏リファレンスデータにより示されるピッチとの比較の結果に基づき前記演奏者による前記楽曲の演奏に対する評価を行い、

前記ピッチ音量データ取得手段により取得されるピッチ音量データは、前記演奏評価手段により前記模範演奏リファレンスデータを用い

て行われた評価の結果、もしくは前記演奏評価手段と同様の手段を備える他機により前記模範演奏リファレンスデータと同様のデータを用いて行われた評価の結果を示す演奏評価データを伴い、

前記表情演奏リファレンスデータ生成手段は、前記ピッチ音量データ取得手段により取得されるピッチ音量データのうち所定の条件を満たす演奏評価データを伴うピッチ音量データを用いて前記表情演奏出現データ生成手段により生成された表情演奏出現データに基づき、前記表情演奏リファレンスデータを生成する

請求項 2 に記載の演奏評価装置。

[請求項6]

請求項 1 乃至 5 のいずれかに記載の演奏評価装置と、

楽曲の伴奏を指示する伴奏データを取得する伴奏データ取得手段と、

前記伴奏データの指示に従い伴奏の楽音を示す音信号を出力する音信号出力手段と

を備え、

前記ピッチ音量データ生成手段は、前記音信号出力手段から出力された音信号に従いスピーカから放音された伴奏に応じて前記演奏者により行われた前記楽曲の演奏音のピッチおよび音量を示すピッチ音量データを生成する

カラオケ装置。

[請求項7]

前記楽曲は歌唱曲であり、

前記歌唱曲の歌詞を示す歌詞データを取得する歌詞データ取得手段と、

前記歌詞データにより示される歌詞であって、前記音信号出力手段により現在出力されている音信号が示す伴奏とともに歌唱されるべき歌詞を示す画像信号を出力する画像信号出力手段と

を備える請求項 6 に記載のカラオケ装置。

[請求項8]

前記楽曲は楽器により演奏される楽曲であり、

前記楽曲の楽譜を示す楽譜データを取得する楽譜データ取得手段と、
前記楽譜データにより示される楽譜であって、前記音信号出力手段により現在出力されている音信号が示す伴奏とともに行われるべき演奏を指示する楽譜を示す画像信号を出力する画像信号出力手段とを備える請求項6に記載のカラオケ装置。

[請求項9]

任意数の任意の演奏者による楽曲の演奏音の各々に関し、前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準とする一のタイミングにおいて一の表情演奏が出現したことを示す表情演奏出現データを取得する表情演奏出現データ取得手段と、

前記表情演奏出現データ取得手段により取得された任意数の表情演奏出現データに基づき、前記楽曲に含まれるノートまたはノート群の各々に関し、当該ノートまたはノート群の発音開始時刻を基準とするいずれのタイミングでいずれの表情演奏がいずれの頻度で出現しているかを特定し、当該特定した情報に従い、前記楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを生成する表情演奏リファレンスデータ生成手段と、

前記表情演奏リファレンスデータ生成手段により生成された表情演奏リファレンスデータを演奏評価装置に送信する送信手段とを備えるサーバ装置。

[請求項10]

歌唱評価システムであって、

楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す第一表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得手段と、

演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量

を示すピッチ音量データを生成するピッチ音量データ生成手段と、

前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記第一表情演奏リファレンスデータにより示される所定時間範囲内において前記第一表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価手段と、

任意数の任意の演奏者による楽曲の演奏音の各々に関し、前記任意の演奏者による前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準とする一のタイミングにおいて一の表情演奏が出現したことを示す表情演奏出現データを取得する表情演奏出現データ取得手段と、

前記表情演奏出現データ取得手段により取得された任意数の表情演奏出現データに基づき、前記任意の演奏者による楽曲に含まれるノートまたはノート群の各々に関し、当該ノートまたはノート群の発音開始時刻を基準とするいずれのタイミングでいずれの表情演奏がいずれの頻度で出現しているかを特定し、当該特定した情報に従い、前記任意の演奏者による楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記任意の演奏者による楽曲において行われるべきタイミングを前記任意の演奏者による楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す第二表情演奏リファレンスデータを生成する表情演奏リファレンスデータ生成手段と、

を備える。

[請求項11]

楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得し、

演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量

を示すピッチ音量データを生成し、

前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価方法。

[請求項12]

コンピュータが実行可能なプログラムであって、

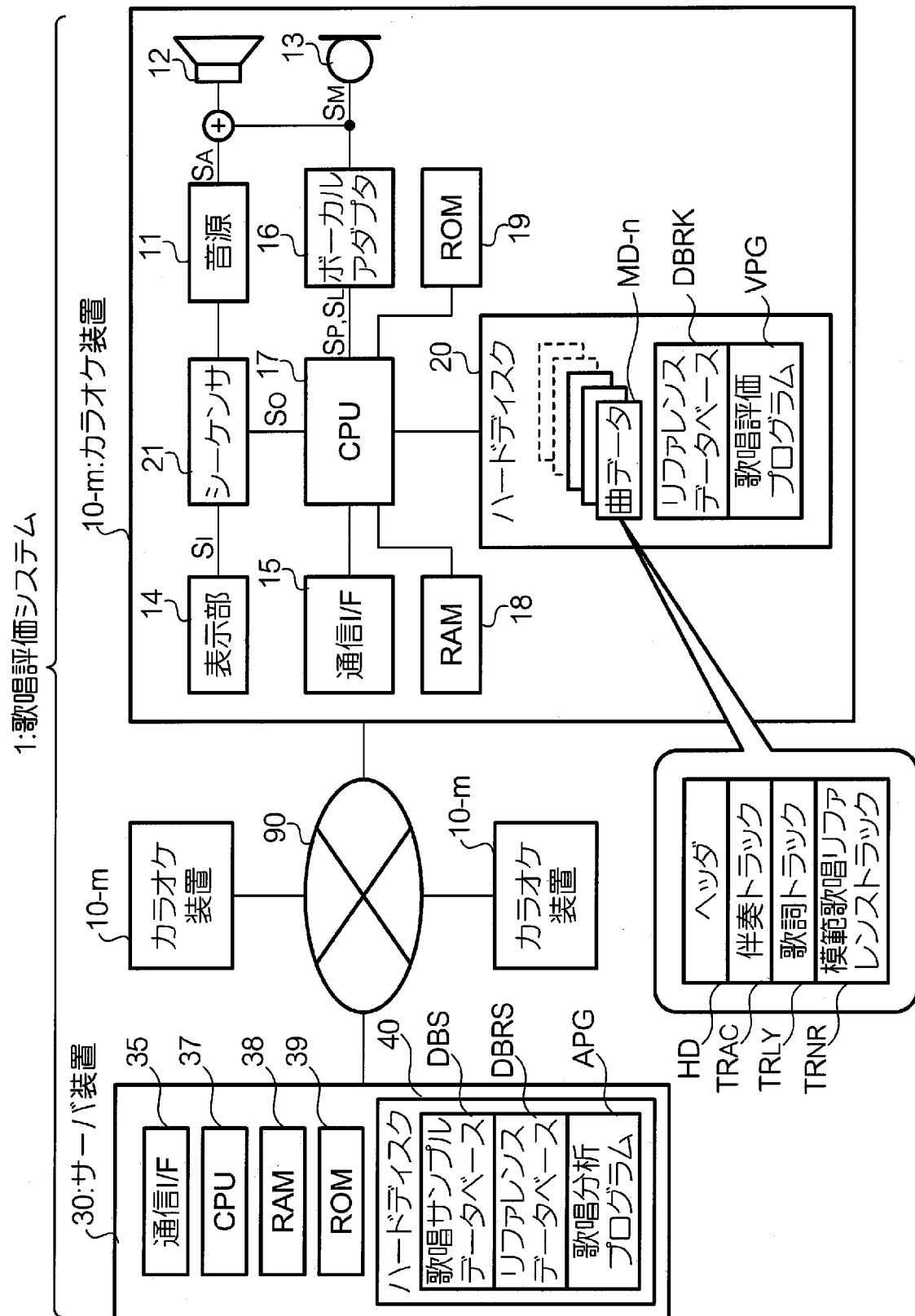
楽曲の演奏中に行われるべき表情演奏と当該表情演奏が前記楽曲において行われるべきタイミングを前記楽曲に含まれるノートまたはノート群の発音開始時刻を基準として示す表情演奏リファレンスデータを取得する表情演奏リファレンスデータ取得処理と、

演奏者による前記楽曲の演奏音から当該演奏音のピッチおよび音量を示すピッチ音量データを生成するピッチ音量データ生成処理と、

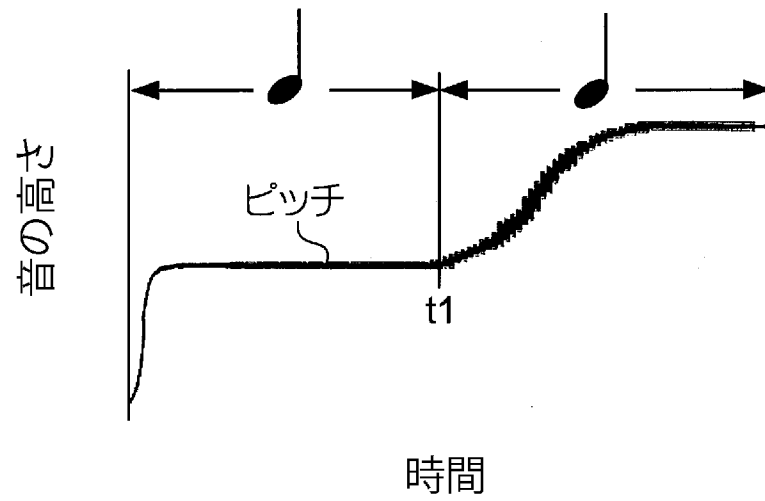
前記ピッチ音量データ生成手段により生成された前記ピッチ音量データにより示されるピッチおよび音量の少なくとも一方の特性が、前記楽曲における前記表情演奏リファレンスデータにより示される所定時間範囲内において前記表情演奏リファレンスデータにより行われるべきであるとされる表情演奏の特性を示す場合、前記演奏者による前記楽曲の演奏に対する評価を向上させる演奏評価処理と

を前記コンピュータに実行させるプログラム。

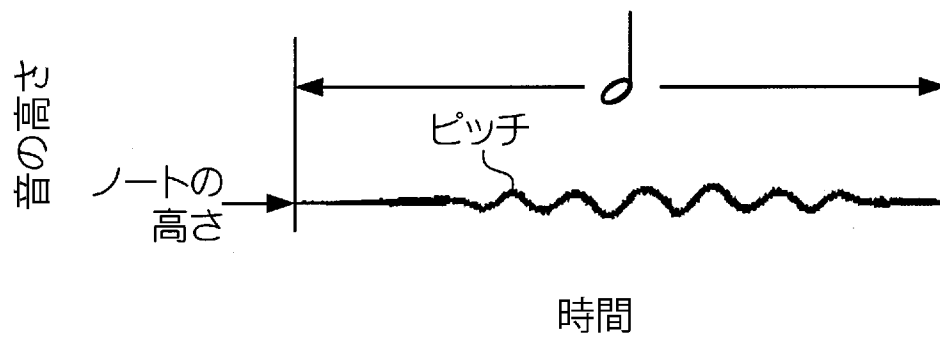
[図1]



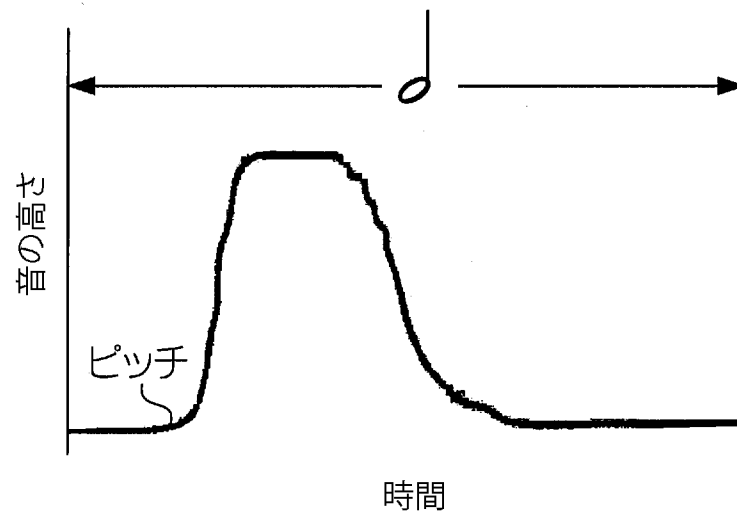
[図2]



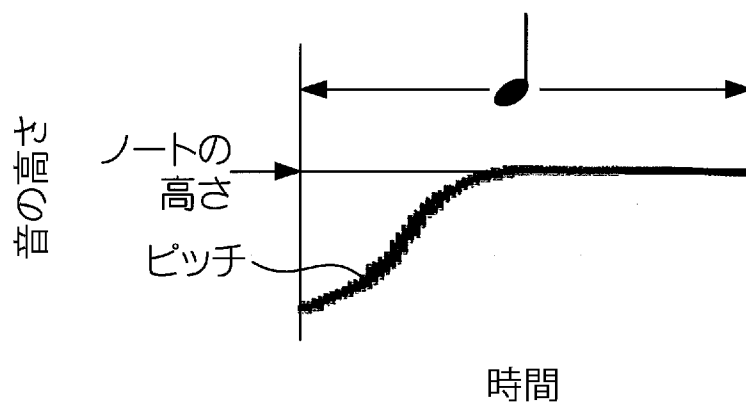
[図3]



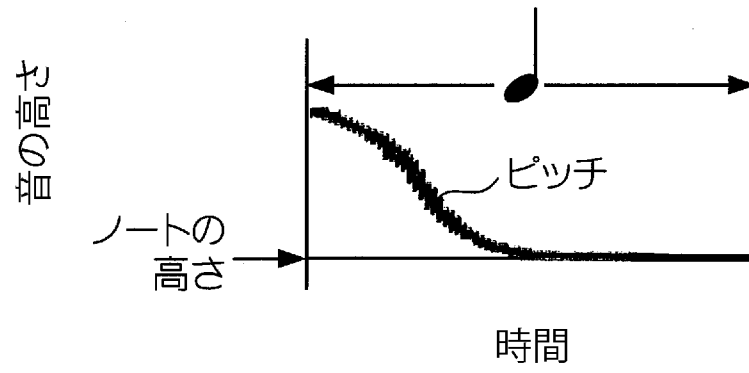
[図4]



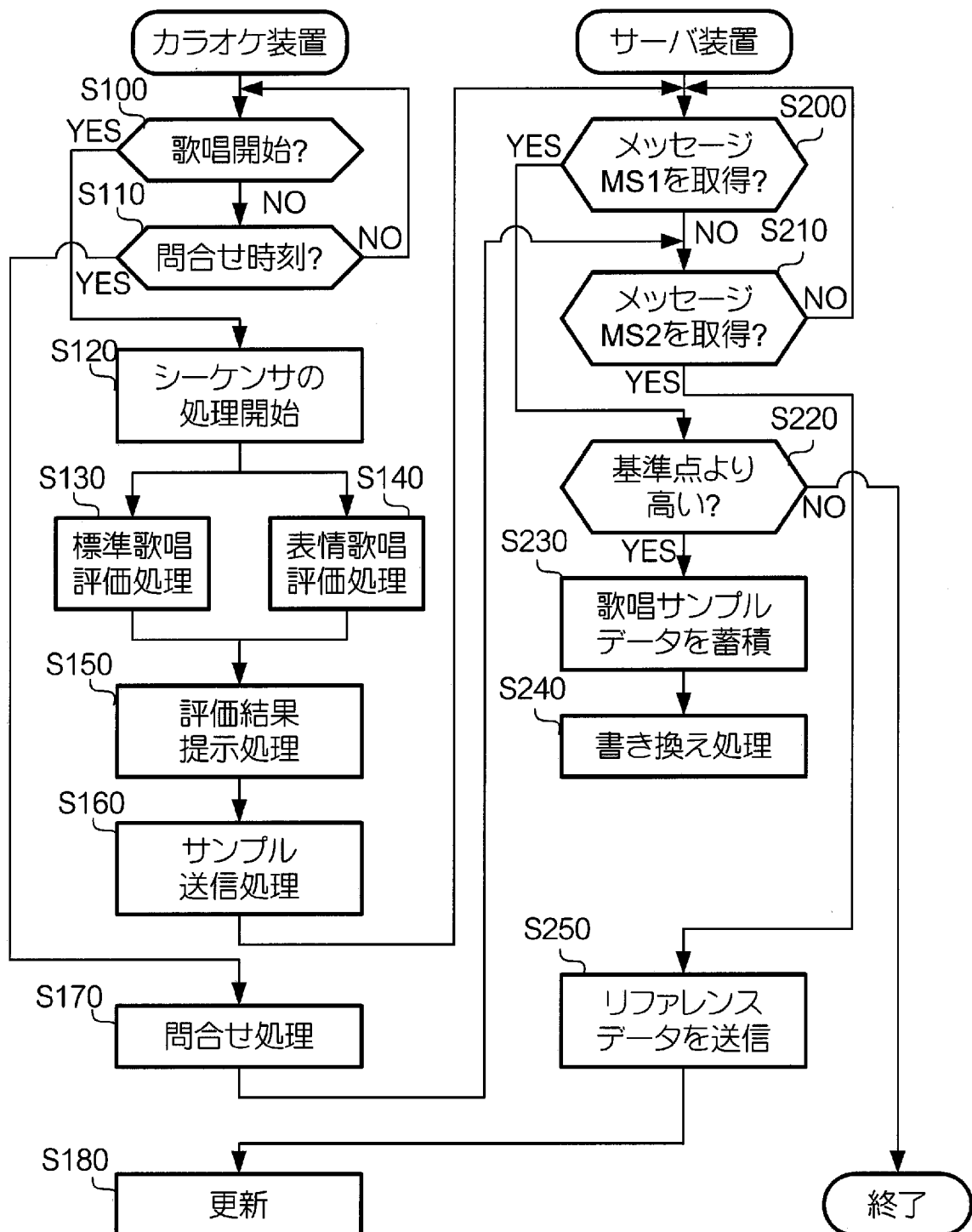
[図5]



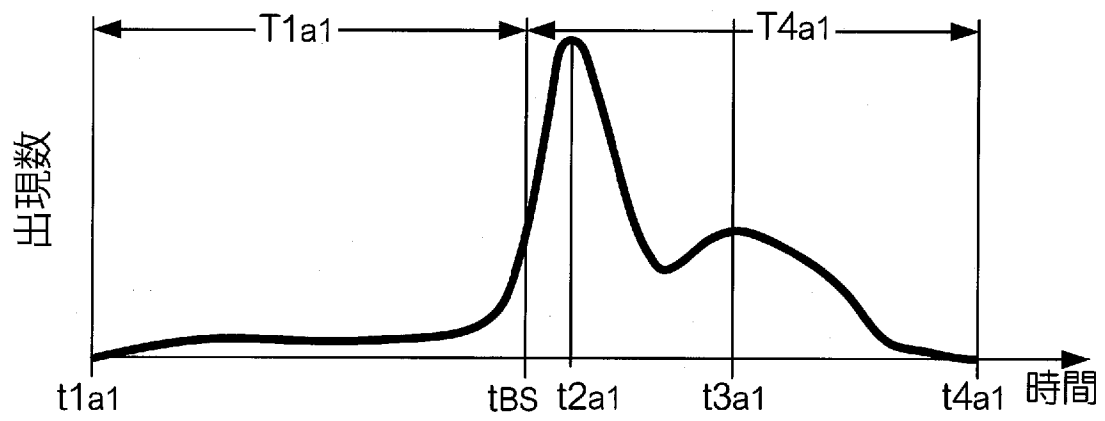
[図6]



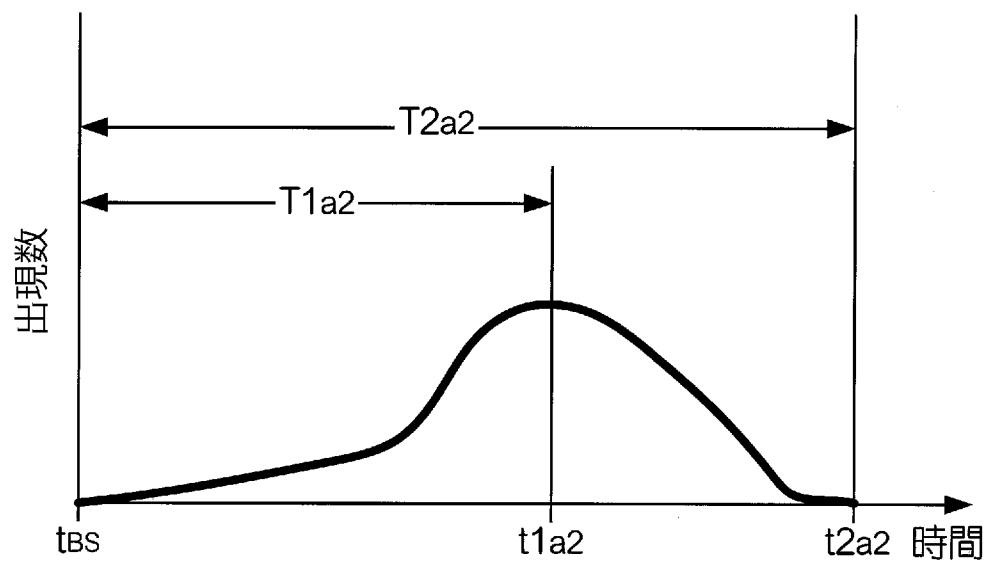
[図7]



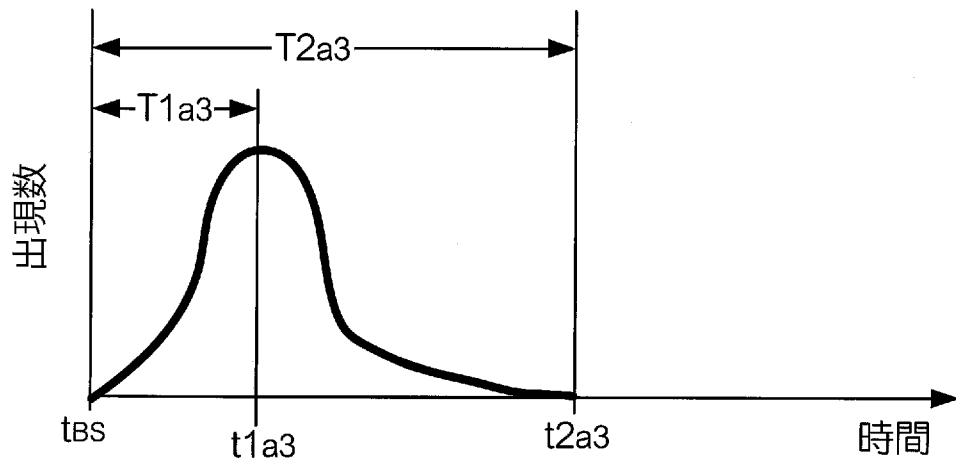
[図8]



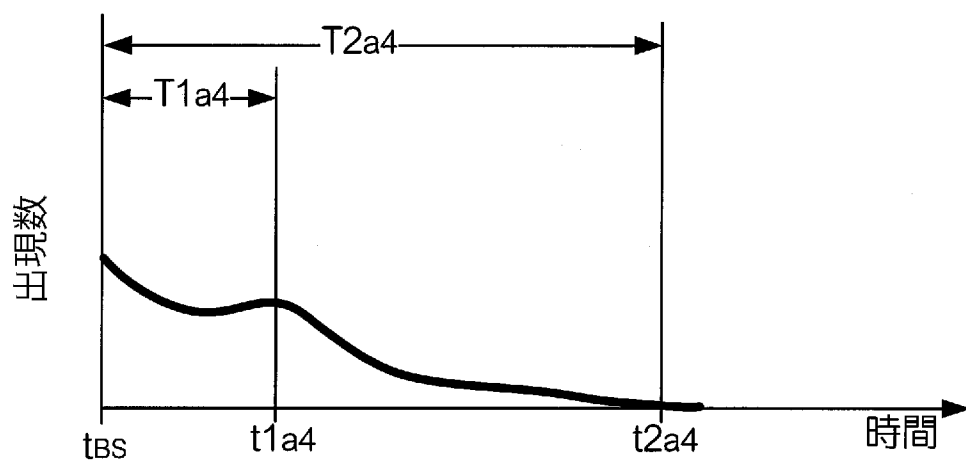
[図9]



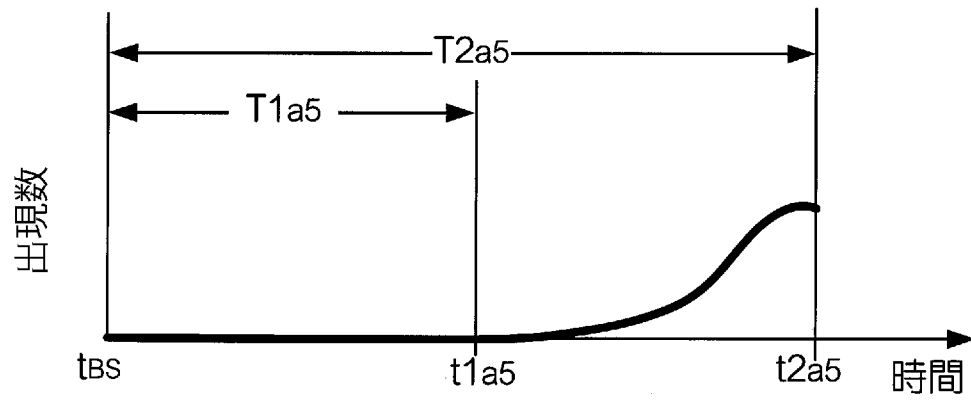
[図10]



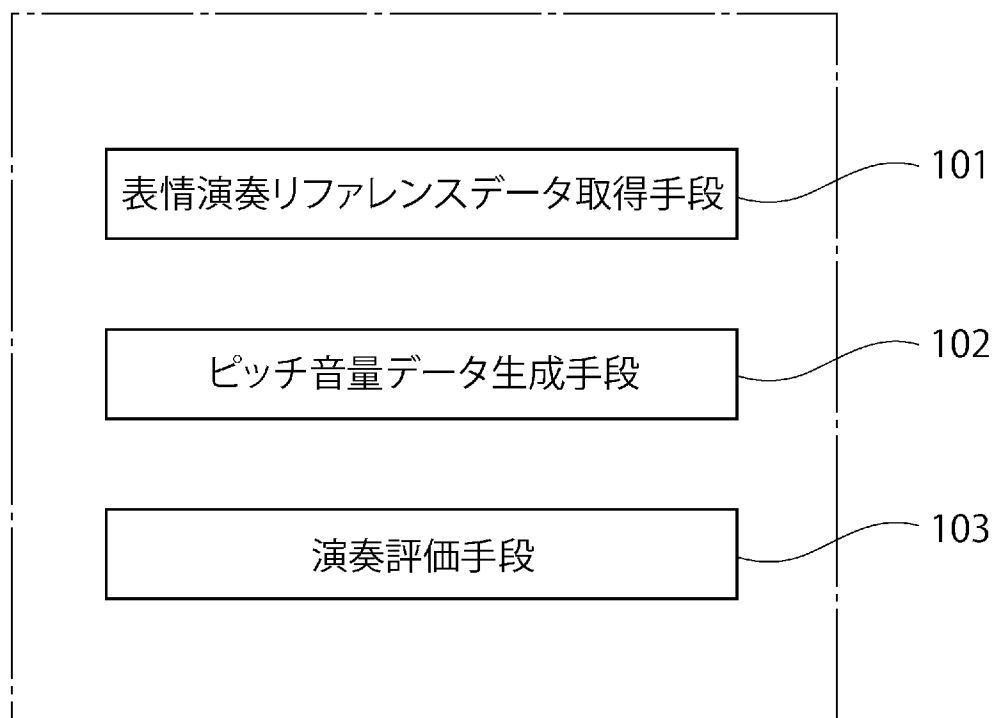
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/061488

A. CLASSIFICATION OF SUBJECT MATTER

G10K15/04(2006.01)i, G09B15/00(2006.01)i, G10L25/51(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10K15/04, G09B15/00, G10L25/51

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2013

Kokai Jitsuyo Shinan Koho 1971-2013 Toroku Jitsuyo Shinan Koho 1994-2013

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2008-139426 A (Yamaha Corp.), 19 June 2008 (19.06.2008), paragraphs [0017] to [0019], [0028], [0040] to [0042] (Family: none)	1, 4, 6-8, 11, 12
A	JP 2005-241848 A (Daiichikosho Co., Ltd.), 08 September 2005 (08.09.2005), paragraph [0026] (Family: none)	2, 3, 5, 9, 10
A	JP 2007-271977 A (Yamaha Corp.), 18 October 2007 (18.10.2007), paragraph [0040] (Family: none)	2, 3, 5, 9, 10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

17 May, 2013 (17.05.13)

Date of mailing of the international search report

28 May, 2013 (28.05.13)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/061488

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007-334364 A (Yamaha Corp., Daiichikosho Co., Ltd.), 27 December 2007 (27.12.2007), fig. 9 (Family: none)	2, 3, 5, 9, 10

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G10K15/04(2006.01)i, G09B15/00(2006.01)i, G10L25/51(2013.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G10K15/04, G09B15/00, G10L25/51

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 3 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 3 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 3 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2008-139426 A（ヤマハ株式会社）2008.06.19, 【0017】－【0019】、【0028】、【0040】－【0042】（ファミリーなし）	1, 4, 6-8, 11, 12
A	JP 2005-241848 A（株式会社第一興商）2005.09.08, 【0026】（ファミリーなし）	2, 3, 5, 9, 10
A	JP 2007-271977 A（ヤマハ株式会社）2007.10.18, 【0040】（ファミリーなし）	2, 3, 5, 9, 10
A	JP 2007-334364 A（ヤマハ株式会社、株式会社第一興商）2007.12.27, 図9（ファミリーなし）	2, 3, 5, 9, 10

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

1 7 . 0 5 . 2 0 1 3

国際調査報告の発送日

2 8 . 0 5 . 2 0 1 3

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

富澤 直樹

電話番号 03-3581-1101 内線 3591

5 Z

4 1 8 8