

明 細 書

発明の名称： 楽曲制作支援プログラム、楽曲制作支援装置、楽曲制作支援方法、及び音楽再生装置

技術分野

[0001] 本発明は、楽曲制作支援プログラム、楽曲制作支援装置、楽曲制作支援方法、音楽再生装置、及び演技音声制作支援プログラムに関する。

背景技術

[0002] 従来より、新人歌手が歌手デビューするに際し、ボイストレーニング等をしっかり積むなど、相当な歌唱力訓練が必要である。ただ歌唱力訓練をいくら積んだとしても、オーディション等の審査でその実力がプロ並に評価されるなど、優れた歌唱力の持ち主でない限り、歌唱力が未熟な歌手にとって、現実的な歌手デビューは極めて困難であった。

[0003] また、歌手の楽曲（例えば新人歌手のデビュー曲）を制作する場合にあたっては、制作の度にとても多くの時間、労力、費用を要しているという現状がある。例えば、1曲のレコーディングに何時間もの時間がかかる上に、その後のピッチ調整や歌声にメロディ・伴奏のmix編集作業にも数日を要する。さらにレコーディングスタジオや収録機器、高価なパソコン及び編集ソフトウェアといったものの経費、ならびに、関係者（歌手、音楽プロデューサー、技術スタッフ等）の労力及び人件費も少なくない。

[0004] ところで近年、AI技術の急速な発達により、歌唱声質変換にかかる技術が知られている。例えば特許文献1に記載される発明においては、カラオケ等において、歌唱ユーザによって指定されたターゲットのボーカルモデルを用いて、当該歌唱ユーザによって予約された楽曲に基づいて歌唱を行うユーザの声質をターゲットの声質に変換して出力する歌唱声質変換プログラム及び歌唱声質変換装置が記載されている。

先行技術文献

特許文献

[0005] 特許文献1：特開2023-013684号公報

発明の概要

発明が解決しようとする課題

[0006] しかしながら、特許文献1に記載される発明においては、歌唱ユーザの声質をターゲットと称される特定の歌手の声質に変換するものであるため、歌唱ユーザの歌唱力を特定の歌手のようなプロ並みの歌唱に飛躍させるものではない。このため、歌唱ユーザの歌唱力が未熟な場合、特定の歌手の声質に変換されたとしても、その歌唱力は未熟なもののままである。

[0007] 本発明者は、上述したように、例えば新人歌手が歌手デビューするにあたり、いくら本人が希望しても歌唱力が未熟な歌手にとって現実的な歌手デビューは極めて困難であること、楽曲を収録する場合にあたっては、収録の度にとっても多くの時間、労力、費用を要しているという現状に鑑みて、AI技術による歌唱声質変換技術によって、プロ歌手のボーカル歌声を、新人歌手の歌声に変換し出力することにより、歌唱力如何にかかわらず、勿論ボイストレーニングなどの歌唱力訓練は全く必要なしに、誰でも自分の声質でプロ並みの「歌声」を手に入れることができると考えた。また従来と比較し、本発明によれば、楽曲制作にかかる時間、労力、費用といったコストも大幅に削減できることから、誰でも短期間に少ないコストで以って歌手デビューが可能であると考えた。

[0008] 本発明は、上記の点に鑑み提案されたものであり、一つの側面において、歌唱技量や演技技量等の技量未熟者に対するデビュー機会の拡大を目的とする。

課題を解決するための手段

[0009] 上記の課題を解決するため、本発明に係る楽曲制作支援プログラムは、第1歌手の歌声データを収録する第1歌声収録手段、前記第1歌手の歌声データに基づいて、前記第1歌手の声質変換モデルを作成する声質変換モデル作成手段、第2歌手により第1言語で歌唱された楽曲の第1言語歌声データを収録し、前記第2歌手又は第3歌手により第2言語で歌唱された前記楽曲の

第2言語歌声データを収録する第2歌声収録手段、前記第1言語歌声データ及び前記第2言語歌声データを、前記第1歌手の声質変換モデルにより前記第1歌手の声質にそれぞれ変換する声質変換手段、として機能させる。

[0010] 上記の課題を解決するため、本発明に係る音楽再生装置は、第1歌手の歌声データを収録する第1歌声収録手段と、前記第1歌手の歌声に基づいて、前記第1歌手の声質変換モデルを作成する声質変換モデル作成手段と、第2歌手により第1言語で歌唱された楽曲の第1言語歌声データを収録し、前記第2歌手又は第3歌手により第2言語で歌唱された前記楽曲の第2言語歌声データを収録する第2歌声収録手段と、前記第1言語歌声データ及び前記第2言語歌声データを、前記第1歌手の声質変換モデルにより前記第1歌手の声質にそれぞれ変換する声質変換手段と、を有する。

[0011] 上記の課題を解決するため、本発明に係る演技音声制作支援方法は、楽曲制作支援装置が、第1歌手の歌声データを収録する第1歌声収録ステップと、前記第1歌手の歌声に基づいて、前記第1歌手の声質変換モデルを作成する声質変換モデル作成ステップと、第2歌手により第1言語で歌唱された楽曲の第1言語歌声データを収録し、前記第2歌手又は第3歌手により第2言語で歌唱された前記楽曲の第2言語歌声データを収録する第2歌声収録ステップと、前記第1言語歌声データ及び前記第2言語歌声データを、前記第1歌手の声質変換モデルにより前記第1歌手の声質にそれぞれ変換する声質変換ステップと、を実行する。

発明の効果

[0012] 本発明の実施の形態によれば、一つの側面において、歌唱技量や演技技量等の技量未熟者に対するデビュー機会の拡大することができる。

図面の簡単な説明

[0013] [図1]本実施形態に係る楽曲制作支援システムのシステム構成例を示す図である。

[図2]本実施形態に係る楽曲制作支援サーバのハードウェア構成例を示す図である。

[図3]本実施形態に係る楽曲制作支援サーバのソフトウェア構成例を示す図である。

[図4]本実施形態に係るDBのデータ構成例を示す図である。

[図5]本実施形態に係るDBのデータ構成例を示す図である。

[図6]本実施形態に係る楽曲制作支援システムに係るフローチャート1である。

[図7]本実施形態に係る楽曲制作支援システムに係るフローチャート2である。

[図8]本実施形態に係る楽曲制作支援システムに係るフローチャート3である。

[図9]本実施形態に係る音楽再生装置の外観例を示す図である。

[図10]本実施形態に係る演技音声制作支援サーバのソフトウェア構成例を示す図である。

[図11]本実施形態に係る演技音声映像制作支援システムに係るフローチャート1である。

[図12]本実施形態に係る演技音声映像制作支援システムに係るフローチャート2である。

[図13]本実施形態に係る演技音声映像制作支援システムに係るフローチャート3である。

[図14]本実施形態に係る楽曲制作支援システムのシステム構成例を示す図である。

[図15]本実施形態に係る楽曲制作支援サーバのソフトウェア構成例を示す図である。

発明を実施するための形態

[0014] 本発明の実施形態について、図面を参照しつつ詳細に説明する。

[実施形態1]

＜システム構成＞

図１は、本実施形態に係る楽曲制作支援システムのシステム構成例を示す図である。図１の楽曲制作支援システム１００は、楽曲制作支援サーバ１０、ＰＣ２０、プロ歌手のマイク３０、新人歌手のマイク４０、及びＣＤライター５０を含み構成されている。

[0015] 楽曲制作支援サーバ１０は、予め歌唱技能の高い歌手（以下例えばプロ歌手と呼ぶ）のボーカル歌声を用意しておくとともに、声質変換技術により、ターゲット歌手（以下例えば歌手デビューする新人歌手と呼ぶ）の声質特徴を抽出し、プロ歌手のボーカル歌声をターゲットの新人歌手の声質に変換することにより、ターゲット歌手のための楽曲制作支援を行うサーバ装置である。楽曲制作支援サーバ１０は、声質変換後の歌声をターゲット歌手の歌声の基本周波数と合わせるためのピッチ調節、ピッチ調節した声質変換後のボーカル（及びハモリボーカル）にリバーブかけたり、ステレオ変換、ノイズ除去など編集し、メロディや伴奏とＭｉｘし、最終的な楽曲（例えば新人歌手のデビュー曲）を完成させる。

[0016] ＰＣ（Personal Computer）２０は、楽曲制作支援サーバ１０にインターネットを含むネットワークを介してアクセスし、楽曲制作にかかる操作を行うための操作端末である。ＰＣ２０は、例えばパーソナルコンピュータ、タブレット端末又はスマートフォンなどにより実現され、所定のアプリケーションプログラム（例えば汎用ウェブブラウザや専用アプリ等）がインストールされる。なお、ＰＣ２０に楽曲制作支援サーバ１０の有する機能を実現するプログラムをインストールすることで、ＰＣ２０と楽曲制作支援サーバ１０を一体装置として実現してもよい。

[0017] プロ歌手は、当該プロ歌手の歌唱力の高いボーカル歌声を予め用意する目的で、レコーディング、即ちマイク３０を用いて、新人歌手の楽曲として発売させたい楽曲（例えば新人歌手のデビュー曲）そのものを歌唱することで、プロ歌手の歌声を収録保存する。また、プロ歌手の歌声は、従来通りの方法でレコーディングスタジオで収録したものを使用してもよい。

[0018] 新人歌手は、歌手デビューする新人歌手の声質変換モデルを作成する目的で、マイク４０を用いて新人歌手の歌声（サンプル）を収録保存する。声質変換モデルの元となる歌唱曲は、所定の楽曲であって１０曲程度分が望ましいが、新人歌手の声質特徴を抽出可能である限りにおいて１曲でもよい。所定の楽曲は、新人歌手のデビュー曲と同一曲を含んでもよいし、新人歌手の声質特徴を抽出可能である限りにおいて必ずしも含まなくともよい。また新人歌手の声質特徴を抽出可能である限りにおいて、必ずしも新人歌手の歌声でなくともよく、例えば会話音声や朗読音声であってもよい。

[0019] ＣＤライター５０は、完成した最終的な楽曲（例えば新人歌手のデビュー曲）を、記憶媒体の一例としてのＣＤ（Compact Disc）に焼き付ける装置である。こうして制作された楽曲のＣＤが新人歌手のデビュー曲として販売される。

[0020] 本実施形態に係る楽曲制作支援システム１００によれば、音程が不安定であったり、ビブラートなどの歌い方ができないなど、歌唱技能の低い新人歌手であっても、プロの歌手の歌声をその新人歌手による声質の歌声に変換できるので、変換後の歌声をＣＤにして販売したり、音楽データファイルとして音楽配信することで、誰でも短期間に低コストで歌手デビューすることが可能となる。

[0021] なお、楽曲制作支援サーバ１０とＰＣ２０とがインターネット等で接続されるなど互いに遠隔地に所在する場合、プロ歌手のマイク３０、新人歌手のマイク４０、及びＣＤライター５０は、ＰＣ２０側に接続してもよい。

[0022] （ハードウェア構成）

図２は、本実施形態に係る楽曲制作支援サーバのハードウェア構成例を示す図である。楽曲制作支援サーバ１０は、ＣＰＵ（Central Processing Unit）１１、ＲＯＭ（Read Only Memory）１２、ＲＡＭ（Random Access Memory）１３、ＨＤＤ（Hard Disk Drive）１４、通信装置１５、及びＧＰＵ（Graphics Processing Unit）１６を有する。

[0023] CPU 11及びGPU 16は、各種プログラムの実行や演算処理を行う。ROM 12は、起動時に必要なプログラムなどが記憶されている。RAM 13は、CPU 11での処理を一時的に記憶したり、データを記憶したりする作業エリアである。HDD 14は、各種データ及びプログラムを格納する。通信装置 15は、通信ネットワークや通信ケーブルを介して他装置とのデータ通信を行う。

[0024] (ソフトウェア構成)

図3は、本実施形態に係る楽曲制作支援サーバのソフトウェア構成例を示す図である。楽曲制作支援サーバ10は、主な機能部として、歌声収録部101、声質変換モデル作成部102、入力受付部103、声質変換部104、歌声編集部105、Mix編集部106、完成楽曲出力部107、及び記憶部109（非図示）を有する。

[0025] 歌声収録部101は、第1歌手（新人歌手等）の歌声データを収録し、第2歌手（日本人プロ歌手等）及び第3歌手（外国人プロ歌手等）により歌唱された楽曲の歌声データを収録する機能を有している。

[0026] 声質変換モデル作成部102は、第1歌手の歌声データに基づいて、第1歌手の声質変換モデル109bを作成する機能を有している。

[0027] 入力受付部103は、声質変換処理の実行に際し、新人歌手、プロ歌手、楽曲、言語など声質変換に関する情報の入力をユーザから受け付ける。

[0028] 声質変換部104は、第2歌手及び第3歌手により歌唱された楽曲の歌声データを、第1歌手の声質変換モデル109bにより第1歌手の声質に変換する機能を有している。

[0029] 歌声編集部105は、声質変換歌声データ（第1歌手の声質に変換された楽曲の歌声データ）を対象に、例えばターゲットとなる新人歌手の歌声の基本周波数と合わせるためのピッチ調節、ピッチ調節した声質変換後のボーカル（及びハモリボーカル）歌声にリバーブかけたり、ステレオ変換、ノイズ除去など、各種編集処理を行う機能を有している。

- [0030] M i x 編集部 1 0 6 は、各種編集処理された声質変換歌声データを対象に、メロディ又は伴奏データを M i x する機能を有している。
- [0031] 完成楽曲出力部 1 0 7 は、各種編集処理、M i x された声質変換歌声データ（楽曲の歌声データ）を、第 1 歌手により歌唱された楽曲データとして出力する機能を有している。出力は、第 1 歌手により歌唱された楽曲データを、外部記憶媒体（C D 等）に記憶したり、インターネットを含むネットワークを介して外部サーバに音楽配信する。
- [0032] 記憶部 1 0 9 は、事前学習済みベースモデル 1 0 9 a、新人歌手の声質変換モデル 1 0 9 b、新人歌手情報 D B 1 0 9 c、プロ歌手歌声保存 D B 1 0 9 d、声質変換歌声保存 D B 1 0 9 e、楽曲 D B 1 0 9 f、完成楽曲 D B 1 0 9 g を記憶する機能を有している。
- [0033] なお、これら各機能部は、本発明に係る機能部であって、楽曲制作支援サーバ 1 0 が楽曲制作支援に必要な各種機能部を他にも備えていることは言うまでもない。また各機能部は、楽曲制作支援サーバ 1 0 を構成するコンピュータの C P U、G P U、R O M、R A M 等のハードウェア資源上で実行されるプログラムによって実現されるものである。これらの機能部は、「手段」、「モジュール」、「ユニット」、又は「回路」に読替えてもよい。また、楽曲制作支援サーバ 1 0 における記憶部 1 0 9 の各 D B は、ネットワーク上の外部記憶装置に配置することも可能である。また、上記コンピュータプログラム及びアプリケーションプログラムは、コンピュータが読み取り可能な記憶媒体に格納されていてもよい。
- [0034] （各種 D B）

図 4（a）は、本実施形態に係る新人歌手情報 D B のデータ構成例を示す図である。新人歌手情報 D B 1 0 9 c は、声質変換モデルに声質を学習させたい学習ターゲットとしての、例えば新人歌手の歌手情報が予め登録された D B である。本実施形態に係る新人歌手情報 D B 1 0 9 c は、例えば新人歌手の「新人歌手 I D」、「氏名」、「性別」、「所属」などのデータ項目を

有する。なお、新人歌手情報DB109cに登録される新人歌手は、その声質の声質変換モデル109bが作成済みの歌手とする。

[0035] 図4(b)は、本実施形態に係るプロ歌手歌声保存DBのデータ構成例を示す図である。プロ歌手歌声保存DB109dは、プロ歌手の歌手情報と、当該プロ歌手による楽曲（例えば新人歌手のデビュー曲）の歌声データが予め登録されたDBである。本実施形態に係るプロ歌手歌声保存DB109dは、例えば「プロ歌手歌声ID」、プロ歌手の「氏名」、「楽曲ID」、「プロ歌手歌声データ」、「言語」などのデータ項目を有する。

[0036] このうち「楽曲ID」は、当該プロ歌手が歌唱した楽曲を特定するためのIDである。楽曲DB109fの楽曲IDと対応する。「プロ歌手歌声データ」は、当該プロ歌手が予め当該楽曲（例えば新人歌手のデビュー曲）を歌唱することで収録された歌声データファイルである。「言語」は、当該楽曲の歌詞言語である。同一曲であっても、歌詞が外国語等の他言語化される場合がある。このような場合、当該プロ歌手が日本語歌詞の当該楽曲を歌唱するとともに、他言語歌詞の当該楽曲を歌唱することで、他言語歌詞分のプロ歌手歌声データが得られる。

[0037] また仮に、当該プロ歌手が他言語での当該楽曲の歌唱が困難の場合、他のプロ歌手が他言語で当該楽曲の歌唱を行うことにより、他言語歌詞分のプロ歌手歌声データを得てもよい。例えば図4(b)に示すプロ歌手歌声保存DB109dには、同一の楽曲ID「M0001」は、プロ歌手Mの歌唱力でもって当該楽曲が日本語歌詞で歌い上げたプロ歌手歌声データ（プロ歌手歌声ID「P0001」）と、英語ネイティブであるプロ歌手Nの歌唱力でもって当該楽曲が英語歌詞で歌い上げたプロ歌手歌声データ（プロ歌手歌声ID「P0002」）とが保存されている。

[0038] 図5(a)は、本実施形態に係る声質変換歌声保存DBのデータ構成例を示す図である。声質変換歌声保存DB109eは、プロ歌手によるプロ歌手歌声データを、ターゲットの新人歌手の声質に変換して保存したDBである。本実施形態に係る声質変換歌声保存DB109eは、例えば「声質変換

歌声 I D」、「新人歌手 I D」、「プロ歌手歌声 I D」、「楽曲 I D」、「声質変換歌声データ」などのデータ項目を有する。

[0039] このうち「新人歌手 I D」は、声質変換後の歌声（声質）の持ち主であるターゲットの新人歌手の新人歌手 I Dを示す。「プロ歌手歌声 I D」は、声質変換前の歌声の持ち主であるプロ歌手のプロ歌手歌声 I Dを示す。「楽曲 I D」は、当該プロ歌手が歌唱した楽曲（例えば新人歌手のデビュー曲）の楽曲 I Dを示す。「声質変換歌声データ」は、プロ歌手によるプロ歌手歌声データが、ターゲットの新人歌手の声質に変換された後の歌声データである。

[0040] 図 5（b）は、本実施形態に係る楽曲 DB 109 f のデータ構成例を示す図である。楽曲 DB は、新人歌手の楽曲として発売させたい楽曲（例えば新人歌手のデビュー曲）の情報が予め登録された DB である。本実施形態に係る楽曲 DB 109 f は、例えば「楽曲 I D」、「楽曲タイトル」、「メロディ・伴奏データ」などのデータ項目を有する。

[0041] このうち「メロディ・伴奏データ」は、当該楽曲のメロディ・伴奏などの音声データである。ターゲットの新人歌手の声質に変換された後の歌声質変換歌声データが、メロディ・伴奏データと M i x 編集されることで、最終的な楽曲が完成する。

[0042] 図 5（c）は、本実施形態に係る完成楽曲 DB のデータ構成例を示す図である。完成楽曲 DB 109 g は、新人歌手の楽曲として発売する楽曲（例えば新人歌手のデビュー完成曲）の情報が保存された DB である。本実施形態に係る完成楽曲 DB 109 g は、例えば「完成楽曲 I D」、「完成楽曲データ」、「楽曲タイトル」、「新人歌手 I D」、「プロ歌手歌声 I D」などのデータ項目を有する。

[0043] このうち「完成楽曲データ」は、ターゲットの新人歌手の声質に変換された後の歌声質変換歌声データが、メロディ・伴奏データと M i x 編集されることで完成した最終的な楽曲（例えば新人歌手のデビュー完成曲）の楽曲データである。「楽曲タイトル」は、発売する当該楽曲のタイトルを示す。

「新人歌手 I D」は、完成楽曲の歌い手とする新人歌手の新人歌手 I Dを示す。「プロ歌手歌声 I D」は、声質変換前の楽曲を実際に歌唱したプロ歌手歌声のプロ歌手歌声 I Dを示す。

[0044] なお、本実施形態に係る上記 D B のデータ構成はあくまで一例であり、他のデータ項目があってもよい。

[0045] <情報処理>

(声質変換モデル作成処理)

図 6 は、本実施形態に係る楽曲制作支援システムに係るフローチャート 1 である。楽曲制作支援サーバ 10 の C P U 又は G P U が本フローチャート処理を実現可能なプログラムを読み込んで実行させることで、各ステップ（以下、「S」と表記する）を実現することができる。

[0046] S 1 : 楽曲制作支援サーバ 10 は、マイク 40 を介して、声質変換モデルを作成するための元となる新人歌手の歌声を収録（取得）する。具体的に、新人歌手は所定の楽曲を例えば 10 曲程度分歌唱する。なお説明上、当該新人歌手のその歌唱力は低いものとする。

[0047] S 2 : 楽曲制作支援サーバ 10 は、収録した新人歌手の歌声の特徴に基づいて、当該新人歌手の声質変換モデル 109 b を作成する。声質変換モデル作成について具体的には、H u B E R T (Hidden Unit Bidirectional Encoder Representationd form Transformers) をベースにするとよい。声質変換モデル作成時には、特定の人声を一から作成するのではなく、まず一般的な人声の事前学習済みベースモデル 109 a を作成しておく。このモデルは、多くの言語圏のたくさんの人の発話データを学習し、人間の一般的な声の特性を持つものである。この事前学習済みベースモデル 109 a に、学習したい人（例えば新人歌手）の声の情報を上から重ねるようにして個々人の声質変換モデルを作成すると、学習対象の人の大量のデータを用意する必要がなく、少量の声で学習し、新人歌手の声質変換モデルを生成することが可能である。

[0048] このように、学習時は変換先である新人歌手の声質特徴量（音声特徴量）を事前学習済みベースモデル 109 a に入力することで精度の高い声質変換モデル 109 b を迅速に作成することができる。声質変換（推論）時は変換元であるプロ歌手の声質特徴量（音声特徴量）を、作成した声質変換モデル 109 b によって新人歌手の歌声に声質変換することになる。

[0049] なお、本実施形態に係る事前学習済みベースモデル 109 a は、日本語のみならず多言語における多数の人間の歌声や音声を事前学習したものであり、様々な言語の母音子音を学習したベースモデルを利用することができる。また事前学習済みベースモデル 109 a は、J-POP、洋楽、アニメソング、演歌、ゲームソングといったように、楽曲ジャンル毎に特化したベースモデルを用意してもよい。例えばアニメソングジャンル毎に特化したベースモデルの場合は、アニソングジャンルにおけるプロ歌手の歌声データがさらに追加学習・移転学習されうる。

[0050] （プロ歌手歌声保存処理）

上述したように、予めプロ歌手が新人歌手の楽曲として発売させたい楽曲（例えば新人歌手のデビュー曲）そのものを歌唱することで、プロ歌手による当該楽曲の歌声データが収録されてプロ歌手歌声保存 DB 109 d に保存される。

[0051] また本実施形態に係る楽曲は英語歌詞版が存在する。このような場合、当該プロ歌手が日本語歌詞の当該楽曲を歌唱するとともに、他言語歌詞の当該楽曲を歌唱することで、他言語歌詞分のプロ歌手歌声データが得られる。当該プロ歌手が他言語での当該楽曲の歌唱が困難の場合、他のプロ歌手が他言語で当該楽曲の歌唱を行うことにより、他言語化分のプロ歌手歌声データを得てもよい。例えば図 4（b）に示されるように、本実施形態にプロ歌手歌声保存 DB 109 d においては、異なるプロ歌手 M 及びプロ歌手 N による、異なる言語歌詞（日本語及び英語）により歌唱された同一楽曲（楽曲 ID : M0001）のプロ歌手歌声データがそれぞれ保存されている。例えばプロ

歌手Mは日本語ネイティブの歌手であり、プロ歌手Nは英語ネイティブの歌手である。

[0052] また楽曲がハモリボーカルを含むような楽曲である場合には、プロ歌手の歌声データとして、当該プロ歌手によるメインボーカルの歌声データのみならず、当該プロ歌手又は他のプロ歌手によるハモリボーカルの歌声データが別に、プロ歌手歌声保存DB109dに保存される。

[0053] (声質変換処理)

図7は、本実施形態に係る楽曲制作支援システムに係るフローチャート2である。

[0054] S21: 楽曲制作支援サーバ10は、PC20を介して、新人歌手、プロ歌手、楽曲、及び言語の入力を受け付ける。これら入力については、新人歌手情報DB109cに登録された新人歌手の中から、プロ歌手歌声保存DB109dに保存されたプロ歌手の氏名、楽曲、言語の中から選択的に入力することも可能である。

[0055] S22: 楽曲制作支援サーバ10は、S21で上記入力を受け付けると、プロ歌手歌声保存DB109dに保存された該当のプロ歌手歌声データを、当該新人歌手の声質変換モデル109bを用いて、該当の新人歌手の声質に変換する。より具体的に、プロ歌手歌声データに基づいて抽出されたプロ歌手の歌声特徴量(音声特徴量)を、当該新人歌手の声質変換モデル109bに入力し、出力として新人歌手の声質に変換されたプロ歌手歌声データを得る。

[0056] S23: 楽曲制作支援サーバ10は、S22で得られた新人歌手の声質に変換後のプロ歌手歌声データを、声質変換歌声データとして、例えば図5(a)に示されるように、当該新人歌手の新人歌手ID、当該プロ歌手のプロ歌手ID、当該楽曲の楽曲IDなどとともに、声質変換歌声保存DB109eに保存する。ここで保存された声質変換後のプロ歌手歌声データは、その声質が当該新人歌手のものであるため、あたかも当該新人歌手が当該プロ

歌手相当の歌唱力でもって当該楽曲を歌唱しているように聞こえるボーカル音声データである。

[0057] また、上述したように当該楽曲がハモリボーカルを含むような楽曲である場合、プロ歌手の歌声データとして、当該プロ歌手によるメインボーカルの歌声データと、ハモリボーカルの歌声データとがプロ歌手歌声保存DB109dにセットで保存される。この場合、楽曲制作支援サーバ10は、S21で上記入力を受け付けると、当該新人歌手の声質変換モデル109bを用いて、プロ歌手歌声保存DB109dに保存されたメインボーカルの歌声データを当該新人歌手の声質に変換するとともに、ハモリボーカルの歌声データを当該新人歌手の声質に変換することができる。また楽曲制作支援サーバ10は、新人歌手の声質に変換後のメインボーカルのプロ歌手歌声データとハモリボーカルのプロ歌手歌声データとを、声質変換歌声データとして、当該新人歌手の新人歌手ID、当該プロ歌手のプロ歌手ID、当該楽曲の楽曲IDなどとともに、声質変換歌声保存DB109eに保存する。

[0058] (編集処理)

図8は、本実施形態に係る楽曲制作支援システムに係るフローチャート3である。

S31：楽曲制作支援サーバ10は、声質変換歌声保存DB109eに保存された声質変換歌声データを対象に各種編集処理を行う。各種編集処理は、例えばターゲットとなる新人歌手の歌声の基本周波数と合わせるためのピッチ調節、ピッチ調節した声質変換後のボーカル（及びハモリボーカル）歌声にリバーブかけたり、ステレオ変換、ノイズ除去などがある。

[0059] S32：楽曲制作支援サーバ10は、各種編集処理された声質変換歌声データを、例えば図5(b)に示されるように、楽曲DB109fに当該楽曲に対応して保存されたメロディ・伴奏データとMixし、最終的な楽曲（例えば新人歌手のデビュー曲）を完成させる。

[0060] S33：楽曲制作支援サーバ10は、完成させた最終的な楽曲（例えば新人歌手のデビュー曲）を、例えば図5(c)に示されるように、完成楽曲デ

ータとして完成楽曲保存DB109gに保存する。その後、完成楽曲データは、例えば新人歌手のデビュー曲として、CDライター50によりCDに焼き付け・包装されたり、音楽データファイルとして音楽配信される。

[0061] 以下に完成楽曲データにかかる属性情報例を示す。

例1：

- ・ 完成楽曲ID：M0001
- ・ 新人歌手氏名：A
- ・ 新人歌手ID：S0001
- ・ 楽曲タイトル：「ルビーめんたい」
- ・ 楽曲ID：M0001
- ・ プロ歌手氏名：M
- ・ プロ歌手歌声ID：P0001
- ・ 言語：日本語

新人歌手Aのデビュー曲の楽曲タイトルは「ルビーめんたい」である。実際のところは、プロ歌手Mの歌唱力でもって当該楽曲を日本語で歌い上げた歌声が、当該新人歌手のその声質に変換されたものをボーカル音声として制作されたものである。

[0062] 例1が示すように、完成楽曲データは、プロ歌手の歌唱力でもって当該楽曲を歌い上げた歌声が、当該新人歌手のその声質に変換されたものをボーカル音声として制作されたものである。このため、音程が不安定であったり、ビブラートなどの歌い方ができないなど、歌唱技能の低い新人歌手（歌の下手な新人歌手）であっても、プロの歌手レベルの歌声で、誰でも短期間に低コストで歌手デビューすることが可能となる。

[0063] 例2：

- ・ 完成楽曲ID：M0002
- ・ 新人歌手氏名：A
- ・ 新人歌手ID：S0001
- ・ 楽曲タイトル：「Ruby Mental」

- ・ 楽曲 I D : M O O O 1
- ・ プロ歌手氏名 : N
- ・ プロ歌手歌声 I D : P O O O 2
- ・ 言語 : 英語

新人歌手 A の海外デビュー曲の楽曲タイトルは「Ruby Mental」(英語歌詞)である。実際のところは、英語ネイティブのプロ歌手 N の歌唱力でもって当該楽曲を英語で歌い上げた歌声が、当該新人歌手のその声質に変換されたものをボーカル音声として制作されたものである。

[0064] 例 2 が示すように、プロ歌手の歌声はどんな言語の歌声であっても、そのプロ歌手歌声データが当該言語で歌い上げられた歌声のデータである限り、新人歌手の声質に変換できるため、新人歌手が当該言語を全く話せなくともましてはその言語歌詞を歌い上げられなくとも、新人歌手の声質でプロ歌手の歌声による当該言語歌詞の歌声を制作することが可能である。これにより、当該外国語のプロ歌手から収録した歌声を用いて、新人歌手の声質に変換し、当該外国で C D を販売したり、配信することにより誰でも容易に世界中の国で歌手デビューすることができる。

[0065] また、既に日本で知名度のある歌唱力の高い有名歌手であっても、日本以外の外国語が苦手な場合がある。しかしながら、同様に有名歌手の歌声を、当該外国語のプロ歌手から収録した歌声を用いて、外国語歌詞の歌声に変換できるので、日本の有名歌手が世界中の国で容易に歌手デビューすることができる。例えば、某有名プロアニメソング歌手による世界的にも有名となった日本発アニメソングに適用すれば、世界中の人々は同じ有名な歌手の歌声で自国歌詞のアニメソングを聞くことができる。

[0066] <応用例>

図 9 は、本実施形態に係る音楽再生装置の外観例を示す図である。本実施形態に係る音楽再生装置 60 は、C P U (非図示)、R A M (非図示)、記憶部 (非図示)、表示部 61、操作部 62 - 64、音楽出力部 (非図示)、U S B インターフェース 65 等を有する。

[0067] CPUは、操作部から入力される制御信号に基づいて、記憶部に記憶されている各種プログラムを読み出し、RAMのワークエリアに展開し、プログラムを実行する。記憶部には、各種プログラム（制御プログラムや音楽再生プログラム等）、本実施形態に係る楽曲制作支援システム100により作成された「完成楽曲データ」が記憶される。表示部61は、CPUから出力される表示信号に従って、装置の動作状態、再生楽曲及び設定等の各種画面を表示する。

[0068] 操作部は、電源ON/OFFボタン62、「完成楽曲データ」の再生/停止を切り替える再生ボタン63、再生音量を大小調整する音量調整ボタン64等を備えている。

[0069] 音楽出力部は、DAコンバータ、アンプ、スピーカー等により構成され、再生操作された音楽をスピーカーから出力する。USBインターフェース65は、USB規格に従った通信を行う。「完成楽曲データ」は例えばUSBインターフェースを介して記憶部に記憶される。

[0070] ここで、図9に示されるように本実施形態に係る再生ボタン63は、2つつのボタンが対を形成しており、具体的に、再生ボタン63-1a及び再生ボタン63-1b、再生ボタン63-2a及び再生ボタン63-2b、再生ボタン63-3a及び再生ボタン63-3b、再生ボタン63-4a及び再生ボタン63-4bの、合計8つのボタンを有している。

[0071] そして各々の再生ボタンにはそれぞれ「完成楽曲データ」が1つつ対応付けられており、ユーザは再生ボタンが押下操作されると、音楽再生装置60は、押下操作された再生ボタンに対応付けられた「完成楽曲データ」を再生する。再生ボタン63-1a及び再生ボタン63-1bに対応付けられた「完成楽曲データ」にかかる属性情報例を示す。

再生ボタン63-1a：

- ・ 完成楽曲ID：M0001
- ・ 新人歌手氏名：A
- ・ 新人歌手ID：S0001

- ・ 楽曲タイトル：「ルビーめんたい」
- ・ 楽曲ID：M0001
- ・ プロ歌手氏名：M
- ・ プロ歌手歌声ID：P0001
- ・ 言語：日本語
- 再生ボタン63-1b：
- ・ 完成楽曲ID：M0002
- ・ 新人歌手氏名：A
- ・ 新人歌手ID：S0001
- ・ 楽曲タイトル：「Ruby Mentai」
- ・ 楽曲ID：M0001
- ・ プロ歌手氏名：N
- ・ プロ歌手歌声ID：P0002
- ・ 言語：英語

[0072] 再生ボタン63-1a及び再生ボタン63-1bに対応付けられた「完成楽曲データ」は、いずれも新人歌手Aのデビュー曲の楽曲タイトルは「ルビーめんたい」である。但し、再生ボタン63-1aに対応付けられた「完成楽曲データ」は、プロ歌手Mの歌唱力でもって当該楽曲（日本語歌詞）を日本語で歌い上げた歌声が、当該新人歌手Aのその声質に変換されたものをボーカル音声として制作されたものである。これに対し、再生ボタン63-1bに対応付けられた「完成楽曲データ」は、英語ネイティブのプロ歌手Nの歌唱力でもって当該楽曲（英語歌詞）を英語で歌い上げた歌声が、当該新人歌手Aのその声質に変換されたものをボーカル音声として制作されたものである。

[0073] 即ち、ユーザが再生ボタン63-1aを押下操作した場合、新人歌手Aの歌声による楽曲タイトル「ルビーめんたい」（日本語歌詞）が再生されるのに対し、ユーザが再生ボタン63-1bを押下操作した場合、新人歌手Aの歌声による楽曲タイトル「Ruby Mentai」（英語歌詞）が再生される。

[0074] 本実施形態に係る音楽再生装置60によれば、複数（例えば2つ）ずつのボタンが対なした再生ボタン操作により、同一歌手歌声による同一楽曲ではあるもののその歌詞言語を容易に切り替えて選曲・再生し、これを聴くことができる。例えば、同じアニメソングを同じ歌手の歌声で、日本語歌詞で聴いたり、英語などの他言語歌詞で聴くことができる。また繰り返し言語を切り替えて、様々な言語歌詞で聴き比べ、歌詞の意味を理解しながら、真似て歌うことで、様々な言語の発音学習に役立てることも可能である。

[0075] なお、本実施形態に係る音楽再生装置60は合計8つのボタンを有しているため、「完成楽曲データ」との対応付けはユーザが任意に設定可能である。8つのボタンの場合における各々のボタンと「完成楽曲データ」との対応付け一例を以下に示す。

- ・再生ボタン63-1a：楽曲A（日本語歌詞）
- ・再生ボタン63-1b：楽曲A（英語歌詞）
- ・再生ボタン63-2a：楽曲B（日本語歌詞）
- ・再生ボタン63-2b：楽曲B（中国語歌詞）
- ・再生ボタン63-3a：楽曲C（日本語歌詞）
- ・再生ボタン63-3b：楽曲C（英語歌詞）
- ・再生ボタン63-4a：楽曲C（中国語歌詞）
- ・再生ボタン63-4b：楽曲C（タイ語歌詞）

[0076] [実施形態2]

上記実施形態1においては、楽曲制作支援システム100を用いて、歌唱技能の低い新人歌手であっても、プロの歌手レベルの歌声で、誰でも短期間に低コストで歌手デビューする仕組みについて説明した。実施形態2においては、同様の仕組みを用いて、音声演技技能の低い新人声優であっても、プロの声優レベルの演技音声で、誰でも短期間に低コストで声優デビューする仕組みについて説明する。声優とは、アニメ等の映像作品に自身の姿を見せずに声だけで出演し演技する者をいう。広義の意味でナレータも含む。

[0077] なお例えば、実施形態１に係る楽曲制作支援システム１００、楽曲制作支援サーバ１０、新人歌手、プロ歌手などは、実施形態２においては、それぞれ演技音声映像制作支援システム２００、演技音声制作支援サーバ７０、新人声優、プロ声優に読み替えることができる。

[0078] (ソフトウェア構成)

図１０は、本実施形態に係る演技音声制作支援サーバのソフトウェア構成例を示す図である。演技音声制作支援サーバ７０は、主な機能部として、音声収録部７０１、声質変換モデル作成部７０２、入力受付部７０３、声質変換部７０４、演技音声編集部７０５、Ｍｉｘ編集部７０６、完成映像作品出力部７０７、及び記憶部７０９（非図示）を有する。

[0079] 音声収録部７０１は、第１声優（新人声優等）の音声データを収録し、第２声優（日本人プロ声優等）及び第３声優（外国人プロ声優等）により歌唱された楽曲の演技音声データを収録する機能を有している。

[0080] 声質変換モデル作成部７０２は、第１声優の音声データに基づいて、第１声優の声質変換モデル７０９ｂを作成する機能を有している。

[0081] 入力受付部７０３は、声質変換処理の実行に際し、新人声優、プロ声優、映像作品、言語など声質変換に関する情報の入力をユーザから受け付ける。

[0082] 声質変換部７０４は、第２声優及び第３声優により演技された台本の演技音声データを、第１声優の声質変換モデル７０９ｂにより第１声優の声質に変換する機能を有している。

[0083] 演技音声編集部７０５は、声質変換演技音声データ（第１声優の声質に変換された台本の演技音声データ）を対象に、各種編集処理を行う機能を有している。

[0084] Ｍｉｘ編集部７０６は、各種編集処理された声質変換演技音声データを対象に、映像データをＭｉｘする機能を有している。

[0085] 完成映像作品出力部７０７は、各種編集処理、Ｍｉｘされた声質変換演技音声データ（台本の演技音声データ）を、第１声優により演技された完成映像作品データとして出力する機能を有している。

[0086] 記憶部 709 は、事前学習済みベースモデル 709 a、新人声優の声質変換モデル 709 b、新人声優情報 DB 709 c、プロ声優演技音声保存 DB 709 d、声質変換歌声保存 DB 709 e、映像作品 DB 709 f、完成映像作品 DB 709 g を記憶する機能を有している。

[0087] (声質変換モデル作成処理)

図 11 は、本実施形態に係る演技音声映像制作支援システムに係るフローチャート 1 である。演技音声制作支援サーバ 70 の CPU 又は GPU が本フローチャート処理を実現可能なプログラムを読み込んで実行させることで、各ステップを実現することができる。

[0088] S51：演技音声制作支援サーバ 70 は、マイク 40 を介して、声質変換モデルを作成するための元となる新人声優の音声を収録（取得）する。具体的に、新人声優は所定の台本を例えば数冊程度分朗読する。なお説明上、当該新人声優のその演技力は低いものとする。

[0089] S52：演技音声制作支援サーバ 70 は、収録した新人声優の音声の特徴に基づいて、当該新人声優の声質変換モデル 709 b を作成する。声質変換モデル作成について具体的には、例えば公知の声質変換による学習と声質変換を適用することができる。

[0090] 事前学習モデルのベースモデルが存在することで、学習対象の人の大量のデータを用意する必要がなく、数分の音声データがあれば、ターゲットとする新人声優の声質特徴量を精度よく学習することが可能である。なお、本実施形態に係る事前学習済みベースモデル 709 a は、日本語のみならず多言語における多数の人間の音声を事前学習したものであり、様々な言語の母音子音を学習したベースモデルを利用する。

[0091] (プロ声優演技音声保存処理)

予めプロ声優が新人声優の演技音声として制作したい映像作品の台本の台詞を演技することで、プロ声優の演技音声データが、プロ声優演技音声保存 DB 709 d に保存される。

[0092] なお、当該プロ声優が日本語台詞の当該台本を演技するとともに、他言語台詞の当該台本を演技することで、他言語化分のプロ声優演技音声データが得られる。なお、当該プロ声優が他言語での当該映像作品の台本演技が困難である場合、他のプロ声優が他言語で当該台本演技を行うことにより、他言語化分のプロ声優演技音声データを得てもよい。

[0093] (声質変換処理)

図12は、本実施形態に係る演技音声映像制作支援システムに係るフローチャート2である。

S61：演技音声制作支援サーバ70は、PC20を介して、新人声優、プロ声優、台本、及び言語の入力を受け付ける。なおこれら入力については、新人声優情報DB709cに登録された新人声優の中から、プロ声優演技音声保存DB709dに保存されたプロ声優の氏名、台本、言語の中から選択的に入力することが可能である。

[0094] S62：演技音声制作支援サーバ70は、S61で上記入力を受け付けると、プロ声優演技音声保存DB709dに保存されたプロ声優演技音声データを、当該新人声優の声質変換モデル709bを用いて、新人声優の声質に変換する。

[0095] S63：演技音声制作支援サーバ70は、S62で得られた新人声優の声質に変換後のプロ声優演技音声データを、声質変換演技音声データとして、当該新人声優の新人声優ID、当該プロ声優のプロ声優ID、当該台本の台本IDなどとともに、声質変換演技音声保存DB709eに保存する。ここで保存された声質変換後の演技音声は、その声質が当該新人声優のものであるため、あたかも当該新人声優が当該プロ声優相当の演技力でもって当該台本を演技しているように聞こえる演技音声である。

[0096] (編集処理)

図13は、本実施形態に係る演技音声映像制作支援システムに係るフローチャート3である。

S 7 1 : 演技音声制作支援サーバ 7 0 は、声質変換演技音声保存 D B 7 0 9 e に保存された声質変換演技音声データを対象に各種編集処理を行う。各種編集処理は、例えばノイズ除去などがある。

[0097] S 7 2 : 演技音声制作支援サーバ 7 0 は、各種編集処理された声質変換演技音声データを、映像作品 D B 7 0 9 f に当該台本に対応して保存された映像データと効果音や B G M と M i x し、最終的な映像作品を完成させる。

[0098] 本実施形態によれば、プロ声優の演技力でもって当該台本を読み上げた演技音声で、当該新人声優のその声質に変換されたものを演技音声とした完成映像作品データが制作される。このため、抑揚が不安定であったり、滑舌がよくないなど、演技技能の低い新人声優であっても、プロの声優レベルの演技音声で、誰でも短期間に低コストで声優デビューすることが可能となる。

[0099] また、プロ声優の演技音声はどんな言語台詞の演技音声であっても、そのプロ声優演技音声データが当該言語台詞で読み上げられた演技音声のデータである限り、新人声優の声質に変換できるため、新人声優が当該言語台詞の台本を全く読めなくとも、新人声優の声質でプロ歌声優レベルの当該言語の演技音声を制作することが可能である。これにより、外国語のプロ声優から収録した外国語による演技音声を用いて、新人声優の声質に変換し、当該外国で映像作品を販売したり、配信することにより誰でも容易に世界中の国で声優デビューすることができる。

[0100] また、既に日本で知名度のある歌唱力の高い有名声優であっても、日本以外の外国語が苦手な場合がある。しかしながら、同様に有名声優の音声を、当該外国語が堪能なプロ声優から収録した演技音声を用いて、外国語発話による演技音声に変換できるので、日本の有名声優が世界中の国で容易に声優デビューすることができる。

[0101] 【実施形態 3】

実施形態 3 においては、上記実施形態 1 に係る楽曲制作支援システム 1 0 0 を変形して、実在しない歌手（以下、仮想歌手と呼ぶ）の声質変換モデルを作成し、バーチャルな実在しない人物を歌手デビューさせるものである。

[0102] 図14は、本実施形態に係る楽曲制作支援システムのシステム構成例を示す図である。図14の楽曲制作支援システム300は、楽曲制作支援サーバ80、PC20、プロ歌手のマイク30、複数人の歌手らのマイク40、及びCDライター50を含み構成されている。

[0103] 楽曲制作支援サーバ10は、予め歌唱技能の高い歌手（以下例えばプロ歌手と呼ぶ）のボーカル歌声を用意しておくとともに、声質変換技術により、ターゲット歌手（実在しない仮想歌手）の声質特徴を抽出し、プロ歌手のボーカル歌声をターゲットの仮想歌手の声質に変換することにより、ターゲット歌手のための楽曲制作支援を行うサーバ装置である。また楽曲制作支援サーバ80は、声質変換後の歌声をターゲット歌手の歌声の基本周波数と合わせるためのピッチ調節、ピッチ調節した声質変換後のボーカル（及びハモリボーカル）にリバーブかけたり、ステレオ変換、ノイズ除去など編集し、メロディや伴奏とMixし、最終的な楽曲（例えば仮想歌手のデビュー曲）を完成させる。

[0104] プロ歌手は、当該プロ歌手のボーカル歌声を予め用意する目的で、レコーディング、即ちマイク30を用いて、仮想歌手の楽曲として発売させたい楽曲（例えば仮想歌手のデビュー曲）そのものを歌唱することで、プロ歌手の歌声を収録保存する。なお、プロ歌手のボーカル歌声について、プロ歌手が実際に歌唱した歌声に代えて、従来の歌声生成ソフト等で生成されたボーカル歌声でもよい。この場合、例えばプロ歌手であっても人間が歌えない音域や歌唱テクニックのボーカル歌声を作成することができる。

[0105] 歌手A、B、C、D・・・は、仮想歌手の声質変換モデルを作成する目的で、マイク40を用いてこれら歌手らの歌声（サンプル）を収録保存する。

[0106] 図15は、本実施形態に係る楽曲制作支援サーバのソフトウェア構成例を示す図である。楽曲制作支援サーバ80は、主な機能部として、歌声収録部101、仮想声質変換モデル作成部102-2、入力受付部103、声質変換部104、歌声編集部105、Mix編集部106、完成楽曲出力部107、及び記憶部109（非図示）を有する。図15に示す楽曲制作支援サー

バのソフトウェア構成例は、図3と比べると、仮想声質変換モデル作成部102-2及び仮想歌手の声質変換モデル109b-2の点で異なっている。また新人歌手情報DB109cが仮想歌手情報DB109c-2となっている。

[0107] 実在しない仮想歌手の仮想歌手の声質変換モデル109b-2を作成する方法は、例えば以下の通りである。

・楽曲制作支援サーバ10（歌声収録部101）は、マイク40を介して、声質変換モデルを作成するための元となる複数人分の実在する歌手A、B、C、D・・・の歌声を収録（取得）する。その後楽曲制作支援サーバ10（仮想声質変換モデル作成部102-2）は、収録した複数人分の歌手の歌声の特徴に基づいて、当該歌手らの声質変換モデルをそれぞれ人数分作成する。そして複数人分歌手の声質変換モデルを、所定割合（例えば4：3：2：1等）を決めてマージ（混ぜる）し、仮想歌手の声質変換モデル109b-2を作成すると、実在しない声質変換モデルを作成することができる。

[0108] ・楽曲制作支援サーバ10（歌声収録部101）は、マイク40を介して、声質変換モデルを作成するための元となる複数人分の実在する歌手A、B、C、D・・・の歌声を収録（取得）する。その後楽曲制作支援サーバ10（仮想声質変換モデル作成部102-2）は、収録した複数人分の歌手の歌声データを混ぜて学習し、仮想歌手の声質変換モデル109b-2を作成することができる。

[0109] ・また、上記実在する歌手A、B、C、D・・・の歌声に代えて、従来の歌声生成ソフト等で生成された仮想歌手A'、B'、C'、D'・・・における複数人分の合成歌声を使用してもよい。

[0110] ・また、実在する歌手の歌声と従来の歌声生成ソフト等で生成された合成歌声とを混ぜて、上記3例のように声質変換モデルを作成してもよい。

[0111] 仮想歌手の声質変換モデル109b-2の作成以降、実施形態3に係る楽曲制作支援システム300においては、実施形態1に係る新人歌手を仮想歌

手に読み替えることにより、図6-8に示したS2～S33のフローチャートを同様に実施可能であるため、繰り返しの説明は省略する。

[0112] 本実施形態によれば、プロ歌手の歌唱力でもって当該楽曲を歌い上げた歌声が、実在しない仮想歌手のその声質に変換されたものをボーカル音声とした完成楽曲データが制作される。プロの歌手レベルの歌声で、短期間に低コストで実在しない歌手をデビューさせることが可能となる。こうしてデビューした仮想歌手（アンドロイド歌手ともいえる）及び楽曲は、特定歌手に依存しないため、楽曲視聴における趣向性向上は勿論のこと、著作権等の権利関係調整のハードルが低下しその用途可能性を広げることが可能である。

[0113] 以上、本実施形態によれば、一つの側面において、歌唱技量や演技技量等の未熟者に対するデビュー機会を拡大することが可能である。なお、本発明の好適な実施の形態により、特定の具体例を示して本発明を説明したが、特許請求の範囲に定義された本発明の広範な趣旨および範囲から逸脱することなく、これら具体例に様々な修正および変更を加えることができることは明らかである。すなわち、具体例の詳細および添付の図面により本発明が限定されるものと解釈してはならない。

[0114] 本発明によって制作された声質変換された歌声を、事前学習モデル生成に加えたり、声質変換モデルの学習データに加えたり、プロ歌手の歌声に加えてもよい。また、本発明により実在しない人間の声質変換モデルを作成して、演技音声制作支援サーバを用いれば、実在しない人間を仮想声優としてデビューさせることもできる。また、歌唱力が未熟な俳優や声優が、映画、テレビドラマ、アニメの中で歌うシーンがある場合、本発明により当該俳優や声優の声質変換モデルを作成し、プロ歌手の歌声を当該俳優や声優の声質に歌声を変換して、その歌声を映像制作に使用することもできる。

符号の説明

[0115] 10、80 楽曲制作支援サーバ
 20 PC
 30、40 マイク

5 0	C D ライター
7 0	演技音声制作支援サーバ
1 0 0、3 0 0	楽曲制作支援システム
2 0 0	演技音声映像制作支援システム
1 0 1	歌声収録部
1 0 2	声質変換モデル作成部
1 0 3	入力受付部
1 0 4	声質変換部
1 0 5	歌声編集部
1 0 6	M i x 編集部
1 0 7	完成楽曲出力部
1 0 8	歌声合成部
1 0 9	記憶部
7 0 1	音声収録部
7 0 2	声質変換モデル作成部
7 0 3	入力受付部
7 0 4	声質変換部
7 0 5	演技音声編集部
7 0 6	M i x 編集部
7 0 7	完成映像作品出力部
7 0 9	記憶部

請求の範囲

- [請求項1] 第1歌手の歌声データを収録する第1歌声収録手段、
前記第1歌手の歌声データに基づいて、前記第1歌手の声質変換モデルを作成する声質変換モデル作成手段、
第2歌手により第1言語で歌唱された楽曲の第1言語歌声データを収録し、前記第2歌手又は第3歌手により第2言語で歌唱された前記楽曲の第2言語歌声データを収録する第2歌声収録手段、
前記第1言語歌声データ及び前記第2言語歌声データを、前記第1歌手の声質変換モデルにより前記第1歌手の声質にそれぞれ変換する声質変換手段、
として機能させるための楽曲制作支援プログラム。
- [請求項2] 前記声質変換手段により前記第1歌手の声質に変換された前記第1言語歌声データ及び前記第2言語歌声データに、メロディ又は伴奏データをM i xする編集手段、
前記編集手段によりM i xされた前記第1言語歌声データ及び前記第2言語歌声データを、前記第1歌手により第1言語で歌唱された第1言語楽曲データ及び前記第1歌手により第2言語で歌唱された第2言語楽曲データとして出力する出力手段、
として機能させるための請求項1に記載の楽曲制作支援プログラム。
- [請求項3] 前記出力手段は、前記第1言語歌声データ及び前記第2言語歌声データを、外部記憶媒体に記憶する記憶装置又は外部サーバに配信する通信装置を含む、
請求項2に記載の楽曲制作支援プログラム。
- [請求項4] 前記楽曲は、アニメソングである、
請求項1に記載の楽曲制作支援プログラム。
- [請求項5] 前記第1歌声収録手段は、
複数の歌声データを収録し、

前記声質変換モデル作成手段は

前記複数の歌声データに基づいて、実在しない前記第1歌手の声質変換モデルを作成する、
請求項1に記載の楽曲制作支援プログラム。

[請求項6]

請求項2に記載の楽曲制作支援プログラムにより出力された前記第1言語楽曲データ及び前記第2言語楽曲データを記憶した記憶部と、

前記第1言語楽曲データと対応付けられた第1言語再生操作部と、

前記第2言語楽曲データと対応付けられた第2言語再生操作部と、

前記第1言語再生操作部が操作された場合に前記第1言語楽曲を再生し、前記第2言語再生操作部が操作された場合に前記第2言語楽曲を再生する音楽出力部と、

を備える音楽再生装置。

[請求項7]

第1歌手の歌声データを収録する第1歌声収録手段と、

前記第1歌手の歌声に基づいて、前記第1歌手の声質変換モデルを作成する声質変換モデル作成手段と、

第2歌手により第1言語で歌唱された楽曲の第1言語歌声データを収録し、前記第2歌手又は第3歌手により第2言語で歌唱された前記楽曲の第2言語歌声データを収録する第2歌声収録手段と、

前記第1言語歌声データ及び前記第2言語歌声データを、前記第1歌手の声質変換モデルにより前記第1歌手の声質にそれぞれ変換する声質変換手段と、

を有することを特徴とする楽曲制作支援装置。

[請求項8]

楽曲制作支援装置が、

第1歌手の歌声データを収録する第1歌声収録ステップと、

前記第 1 歌手の歌声に基づいて、前記第 1 歌手の声質変換モデルを作成する声質変換モデル作成ステップと、

第 2 歌手により第 1 言語で歌唱された楽曲の第 1 言語歌声データを収録し、前記第 2 歌手又は第 3 歌手により第 2 言語で歌唱された前記楽曲の第 2 言語歌声データを収録する第 2 歌声収録ステップと、

前記第 1 言語歌声データ及び前記第 2 言語歌声データを、前記第 1 歌手の声質変換モデルにより前記第 1 歌手の声質にそれぞれ変換する声質変換ステップと、

を実行することを特徴とする楽曲制作支援方法。