

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6110731号

(P6110731)

(45) 発行日 平成29年4月5日(2017.4.5)

(24) 登録日 平成29年3月17日(2017.3.17)

(51) Int.Cl.			F I		
G10K	15/04	(2006.01)	G10K	15/04	302D
G06F	3/01	(2006.01)	G06F	3/01	570
G06F	3/0481	(2013.01)	G06F	3/0481	150

請求項の数 4 (全 13 頁)

(21) 出願番号	特願2013-116493 (P2013-116493)	(73) 特許権者	390004710
(22) 出願日	平成25年5月31日(2013.5.31)		株式会社第一興商
(65) 公開番号	特開2014-235301 (P2014-235301A)		東京都品川区北品川5丁目5番26号
(43) 公開日	平成26年12月15日(2014.12.15)	(74) 代理人	100130362
審査請求日	平成28年3月14日(2016.3.14)		弁理士 小川 嘉英
		(72) 発明者	橘 聡
			東京都品川区北品川5丁目5番26号 株
			株式会社第一興商内
		(72) 発明者	執行 里恵
			東京都品川区北品川5丁目5番26号 株
			株式会社第一興商内
		審査官	千本 潤介

最終頁に続く

(54) 【発明の名称】 ジェスチャーによるコマンド入力識別システム

(57) 【特許請求の範囲】

【請求項1】

歌唱者の歌唱姿態を録画する機能と、歌唱者の動作を認識してコマンド入力を行う機能とを有するカラオケシステムにおいて、

歌唱者の歌唱姿態を撮影する撮影手段と、

歌唱者の歌唱音声を入力する音声入力手段と、

前記撮影手段で撮影した歌唱者の歌唱姿態及び入力された歌唱音声をを用いて動画データを作成して保存する動画データ作成手段と、

予め定められた歌唱者の動作と、当該歌唱者の動作に対応する所定のコマンドとを紐付けして管理するコマンドデータベースと、

歌唱者の動作映像を解析し、当該歌唱者の動作が前記コマンドデータベースに存在する場合に、当該歌唱者の動作をコマンド入力のためのジェスチャーであると判定するコマンド判定手段と、

前記歌唱者の動作がコマンド入力のためのジェスチャーであると判定された場合に、前記動画撮影された歌唱者の歌唱姿態の中から当該コマンド入力が行われている撮影区間を特定するコマンド入力区間特定手段と、

前記コマンド入力が行われていると特定された撮影区間を含む所定区間の動画データを編集して、当該コマンド入力のためのジェスチャーが含まれない一連の動画データとして編集し、あるいは当該コマンド入力のためのジェスチャーを識別可能とする動画データ編集手段と、

10

20

を備えたことを特徴とするジェスチャーによるコマンド入力識別システム。

【請求項 2】

前記コマンド判定手段で判定の対象となる前記動作映像とは、前記撮影手段で撮影している歌唱者の歌唱姿態と、前記動画データ作成手段で作成された動画データに含まれる歌唱者の歌唱姿態のいずれか一方であることを特徴とする請求項 1 に記載のジェスチャーによるコマンド入力識別システム。

【請求項 3】

前記動画データ編集手段は、前記コマンドデータベースに基づいて、前記コマンド入力が行われていると特定された撮影区間の動画データを、当該コマンド入力のためのジェスチャーを含まない歌唱者の顔を中心とした映像にデジタルズームアップした映像データに置き換えることを特徴とする請求項 1 又は 2 に記載のジェスチャーによるコマンド入力識別システム。

10

【請求項 4】

前記コマンドデータベースは、コマンドの内容に応じた映像データを含み、
前記動画データ編集手段は、前記コマンドデータベースに基づいて、前記コマンド入力が行われていると特定された撮影区間の動画データに、当該コマンドの内容に応じた映像データを合成して、前記コマンド入力のためのジェスチャーを識別可能とすることを特徴とする請求項 1 又は 2 に記載のジェスチャーによるコマンド入力識別システム。

【発明の詳細な説明】

【技術分野】

20

【0001】

本発明は、ジェスチャーによるコマンド入力識別システムに関するものであり、例えば、ジェスチャーにより種々のコマンドを実行可能なカラオケシステムにおいて、コマンド入力のためのジェスチャーと歌唱時の振り付けとを区別可能なシステムに関するものである。

【背景技術】

【0002】

近年、ナチュラルユーザーインターフェイスとして、利用者の動作（ジェスチャー）を撮像して 3 次元解析等を行うことにより、種々のコマンドを認識することが可能な技術が開発されている。この技術を用いて、利用者の動作（ジェスチャー）とコマンドとを紐付けしておくことにより、利用者はカラオケリモコン装置等を操作することなく、動作（ジェスチャー）によりコマンド入力を行うことができる。

30

【0003】

例えば、利用者の動作（ジェスチャー）により自動演奏、歌唱音声に対する効果等を制御することが可能な技術が開示されている（特許文献 1 参照）。特許文献 1 に記載された技術は、撮像手段から得られる人物の姿形と、三次元位置計測装置から得られる人物の三次元位置情報とに基づき、画像処理装置において人物の三次元モデルとのマッチングを行い人物のジェスチャーを検出する。そして、画像処理装置は、検出されたジェスチャーに基づき、楽音処理装置および映像表示装置に制御イベントを伝達する。楽音処理装置及び映像表示装置は、受け取った制御イベントに基づき、自動演奏、歌唱音声に対する効果等を制御するようになっている。

40

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開平 11 - 175061 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

ところで、カラオケを楽しむ際に、歌唱だけではなくアーティストのダンスの振り付けを真似て歌唱することがあり、このような歌唱者の歌唱姿態を他人に見て貰いたいという

50

要望がある。そこで、現在普及しているカラオケシステムでは、歌唱者の歌唱姿態を動画撮影してDVD等の記録媒体に記録したり、ウェブサイトで公開したりする機能を有しているものがある。このような動画公開機能を有するカラオケシステムに、上述した利用者の動作（ジェスチャー）によりコマンド入力を行う機能を採用することが考えられる。

【0006】

しかし、歌唱者の歌唱姿態を動画撮影する場合に、ダンスの振り付けだけでなく、コマンド入力のための動作（ジェスチャー）が含まれていると、振り付けとは関係のない動作が録画されてしまい、動画を公開する利用者や動画の視聴者にとって、動画公開の興趣を削いでしまうおそれがあった。

【0007】

本発明は、上述した事情に鑑み提案されたもので、歌唱者の歌唱姿態を動画撮影する機能と、歌唱者の動作を認識してコマンド入力を行う機能とを有するカラオケシステムにおいて、両者を識別して、コマンド入力のための動作が行われた場合に、当該動作を他の映像に置き換え、あるいは、コマンド入力のための動作であることを識別できるような情報を映像に付加して、動画公開の興趣を高めることが可能なジェスチャーによるコマンド入力識別システムを提供することを目的とする。

【課題を解決するための手段】

【0008】

本発明のジェスチャーによるコマンド入力識別システムは、上述した事情に鑑み提案されたもので、以下の特徴点を有している。すなわち、本発明のジェスチャーによるコマンド入力識別システムは、歌唱者の歌唱姿態を録画する機能と、歌唱者の動作を認識してコマンド入力を行う機能とを有するカラオケシステムにおいて、撮影手段と、音声入力手段と、動画データ作成手段と、コマンドデータベースと、コマンド判定手段と、コマンド入力区間特定手段と、動画データ編集手段とを備えたことを特徴とするものである。

【0009】

撮影手段は、歌唱者の歌唱姿態を撮影するための手段である。音声入力手段は歌唱者の歌唱音声を入力するための手段である。動画データ作成手段は、撮影手段で撮影した歌唱者の歌唱姿態及び音声入力手段により入力された歌唱音声を用いて動画データを作成して保存するための手段である。コマンドデータベースは、予め定められた歌唱者の動作と、当該歌唱者の動作に対応する所定のコマンドとを紐付けして管理するためのデータベースである。

【0010】

コマンド判定手段は、歌唱者の動作映像を解析し、当該歌唱者の動作がコマンドデータベースに存在する場合に、当該歌唱者の動作をコマンド入力のためのジェスチャーであると判定するための手段である。コマンド入力区間特定手段は、歌唱者の動作がコマンド入力のためのジェスチャーであると判定された場合に、動画撮影された歌唱者の歌唱姿態の中から当該コマンド入力が行われている撮影区間を特定するための手段である。動画データ編集手段は、コマンド入力が行われていると特定された撮影区間の動画データを編集して、当該コマンド入力のためのジェスチャーが含まれない一連の動画データとして編集し、あるいは当該コマンド入力のためのジェスチャーを識別可能とするための手段である。

【0011】

また、コマンド判定手段で判定の対象となる動作映像とは、撮影手段で撮影している歌唱者の歌唱姿態と、動画データ作成手段で作成された動画データに含まれる歌唱者の歌唱姿態のいずれか一方とすることが可能である。

【0012】

また、動画データ編集手段は、コマンドデータベースに基づいて、コマンド入力が行われていると特定された撮影区間を含む所定区間の動画データを、当該コマンド入力のためのジェスチャーを含まない歌唱者の顔を中心とした映像にデジタルズームアップした映像データに置き換えることが可能である。

【0013】

10

20

30

40

50

また、コマンドデータベースに、コマンドの内容に応じた映像データを含ませることにより、動画データ編集手段において、コマンドデータベースに基づいて、コマンド入力が行われていると特定された撮影区間の動画データに、当該コマンドの内容に応じた映像データを合成して、コマンド入力のためのジェスチャーを識別可能とすることが可能である。

【0014】

このような構成からなるジェスチャーによるコマンド入力識別システムでは、音声入力手段（例えば、マイクロホン及びその付帯機器）の機能により歌唱者の歌唱音声を入力し、撮影手段（例えば、ビデオカメラ及びその付帯機器）の機能により、歌唱者の歌唱姿態を撮影する。そして、動画データ作成手段の機能により、撮影した歌唱姿態及び入力された歌唱音声を用いて動画データを作成して、所定の記録装置（例えば、HDD）に保存する。保存した動画データは、DVD等の記録媒体に記録したり、動画公開を目的としたウェブサイト等にアップロードしたりすることができる。

10

【0015】

また、本発明では、予め定められた歌唱者の動作と、当該歌唱者の動作に対応する所定のコマンドとを紐付けして管理するためのデータベースを備えており、コマンド判定手段の機能により、歌唱者の動作映像を解析し、当該歌唱者の動作がコマンドデータベースに存在する場合に、当該歌唱者の動作をコマンド入力のためのジェスチャーであると判定する。歌唱者の動作がコマンド入力のためのジェスチャーであると判定された場合には、当該コマンドに基づくカラオケ演奏装置の操作が行われる。このようなコマンドには、例えば、音量の増減、効果音（拍手や歓声等）の発生、エコーの強弱の設定、演奏キーの変更、演奏テンポの変更等がある。

20

【0016】

さらに、歌唱者の動作がコマンド入力のためのジェスチャーであると判定された場合には、コマンド入力区間特定手段の機能により、コマンド入力が行われている撮影区間を特定する。そして、動画データ編集手段の機能により、コマンド入力が行われていると特定された撮影区間を含む所定区間の動画データについて、当該コマンド入力のためのジェスチャーが含まれない一連の動画データとして編集し、あるいは当該コマンド入力のためのジェスチャーを識別可能とする処理を実施する。これにより、コマンド入力のためのジェスチャーが含まれない一連の動画データ、あるいは歌唱者の動作がコマンド入力のためのジェスチャーであることを識別可能な動画データが作成される。コマンド入力のためのジェスチャーが含まれない一連の動画データとは、例えば、歌唱者の顔を中心とした映像にデジタルズームアップした映像データ、歌唱者のストップモーション画像、予め用意されたムード映像、カラオケを一緒に楽しんでいる他の利用者の映像等である。

30

【0017】

また、コマンドデータベースに、コマンドの内容に応じた映像データ（コマンドの説明文字データ、コマンドの内容表示データ）を含ませた場合には、コマンド入力が行われていると特定された撮影区間の映像において、ジェスチャーがコマンド入力のための動作であることを識別可能とする。

【発明の効果】

40

【0018】

本発明のジェスチャーによるコマンド入力識別システムによれば、歌唱者の歌唱姿態を録画する機能と、歌唱者の動作を認識してコマンド入力を行う機能とを有するカラオケシステムにおいて、現に撮影している歌唱者の映像、あるいは保存されている動画データにおいて、コマンド入力のための動作が映っている場合に、当該動作をコマンド入力のためのジェスチャーであると識別して、当該ジェスチャーを他の映像に置き換え、あるいはジェスチャーがコマンド入力のための動作であることを識別可能とする。

【0019】

したがって、歌唱動画を録画して後に再生したり、歌唱動画の公開を行ったりした場合に、歌唱動画の中に余分な動作が含まれることがなく、また、余分な動作が含まれたまま

50

でも、その動作自体がコマンド入力のための動作（ジェスチャー）であると識別することができるので、歌唱者及び視聴者の興味を高めることが可能となる。

【図面の簡単な説明】

【0020】

【図1】本発明の実施形態に係るジェスチャーによるコマンド入力識別システムの構成を示すブロック図。

【図2】動画データの編集処理の手順を示すフローチャート。

【図3】映像の置き換えを説明する説明図。

【図4】コマンド入力のためのジェスチャーを識別可能とする映像処理の模式図。

【発明を実施するための形態】

【0021】

図面を参照して、本発明のジェスチャーによるコマンド入力識別システム（以下、コマンド入力識別システムと略記する）の実施形態について説明する。図1～図4は本発明の実施形態に係るコマンド入力識別システムを示すもので、図1はコマンド入力識別システムの構成を示すブロック図、図2は動画データの編集処理の手順を示すフローチャート、図3は映像の置き換えを説明する説明図、図4はコマンド入力のためのジェスチャーを識別可能とする映像処理の模式図である。

【0022】

<コマンド入力識別システムの概要>

本発明の実施形態に係るコマンド入力識別システム10は、歌唱者の歌唱姿態を録画する機能と、歌唱者の動作を認識してコマンド入力を行う機能とを有するカラオケシステムに適用する技術であり、主要な構成要素として、図1に示すように、撮影手段（ビデオカメラ32及びその付帯機器）と、音声入力手段（マイクロホン34及びA/Dコンバータ53）と、動画データ作成手段51と、コマンドデータベース45cと、コマンド判定手段48と、コマンド入力区間特定手段49と、動画データ編集手段50とを備えている。各手段等は、カラオケ演奏端末30及びその付帯装置の機能手段として実現される。

【0023】

なお、以下の説明において、プログラムとは、RAM等に記憶され、CPU等のハードウェアで実行されることにより、その機能を発揮するソフトウェアだけではなく、同等の機能を発揮することが可能な論理回路も含む概念である。

【0024】

<カラオケ演奏端末>

本発明の実施形態に係るコマンド入力識別システム10を適用するカラオケ演奏端末30は、図1に示すように、カラオケ本体40、カラオケリモコン装置31、ビデオカメラ32、スピーカ33、マイクロホン34、表示装置35、ミキシングアンプ36を備えている。また、本実施形態のカラオケ演奏端末30は、ルータ20及びデータ送受信回線70を介して、管理サーバ60にネットワーク接続されている。

【0025】

<管理サーバ>

管理サーバ60は、会員情報の管理、カラオケ演奏端末30に対する楽曲データ等の配信、利用者がアップロードしたカラオケ動画の公開等を行うためのサーバである。データ送受信回線70は、例えば、インターネット回線、デジタル通信回線、アナログ通信回線、無線通信回線、LAN等、環境に応じてどのような回線を用いてもよいが、ネットワークに対する第三者の侵入やデータの傍聴及び改竄が困難であるとともに、帯域を独占せずに安価な通信網であるという点で、インターネットにより構成されるVPNを利用することが好ましい。

【0026】

なお、単独の管理サーバ60により、上述した複数の機能を実現するのではなく、各機能に特化したサーバを設け、各サーバにより各機能を実現してもよい。この際、仮想化技術により、1つのサーバに複数の機能を持たせることもできる。

10

20

30

40

50

【 0 0 2 7 】

<カラオケリモコン装置>

カラオケリモコン装置 3 1 は、ユーザインタフェース機能を備えており、カラオケ本体 4 0 のローカル送受信手段 4 6 との間で有線方式又は無線方式によりデータの送受信を行うようになっている。このカラオケリモコン装置 3 1 は、楽曲検索手段 3 1 a として機能するプログラム、楽曲索引データベース 3 1 b、種々のデータを記憶するためのデータ記憶部 3 1 c、データの入出力を行うための入出力表示部 3 1 d 等を備えている。このカラオケリモコン装置 3 1 に付帯するスイッチ類や、入出力表示部 3 1 d に表示される各種のアイコン等を操作することにより、選曲操作等が行われる。

【 0 0 2 8 】

<楽曲検索手段 / 楽曲索引データベース>

楽曲検索手段 3 1 a は、利用者の指示に基づき、楽曲索引データベース 3 1 b を参照して楽曲を検索するためのプログラムからなる。楽曲索引データベース 3 1 b は、カラオケ演奏端末 3 0 で演奏に供されるカラオケ楽曲について、その属性情報を記述したデータベースであり、例えば、楽曲番号・曲名・歌手名・歌い出し部分の歌詞・流行時期・音楽ジャンル区分・デュエット曲か否かなど、種々の属性情報がこれに含まれている。

【 0 0 2 9 】

<マイクロホン>

マイクロホン 3 4 は、歌唱音声の入力を行うための装置である。マイクロホン 3 4 から入力された歌唱音声信号は、ミキシングアンプ 3 6 により、音楽再生制御手段 5 2 から送出される演奏音声信号とミキシングされると共に増幅され、スピーカ 3 3 へ出力される。なお、マイクロホン 3 4 からの音声入力信号は、A / D コンバータ 5 3 によりデジタル変換され、動画データ作成手段 5 1 における動画の作成や歌唱採点手段（図示せず）における採点等に使用される。本実施形態では、マイクロホン 3 4 及び A / D コンバータ 5 3 が音声入力手段として機能する。

【 0 0 3 0 】

<ビデオカメラ>

ビデオカメラ 3 2 は、利用者の歌唱姿態を撮影するための撮影手段として機能する装置であり、撮像レンズ及び撮像素子を主要な構成要素とし、フォーカシング機構、ズーム機構、パン・チルト機構等を備えていてもよい。ビデオカメラ 3 2 で撮影が行われると、映像信号（ビデオ信号）が入力される。この映像信号は動画作成手段 5 1 における動画の作成やコマンド判定手段 4 8 における歌唱者の動作映像の解析等に使用される。なお、図 1 に示す例では、ビデオカメラ 3 2 を 2 台設けている。これは、利用者の歌唱姿態だけではなく、置換映像データ 4 4 c として、観客の様子を撮影する場合を考慮したためである。

【 0 0 3 1 】

<表示装置>

表示装置 3 5 は、カラオケ楽曲に関連した背景映像や歌詞テロップ等を表示するための装置で、例えば、液晶ディスプレイ等により構成される。

【 0 0 3 2 】

<カラオケ本体>

カラオケ本体 4 0 は、ネットワーク送受信手段 4 1、中央制御手段 4 2、ROM 4 3、RAM 4 4、HDD 4 5、ローカル送受信手段 4 6、予約管理手段 4 7、コマンド判定手段 4 8、コマンド入力区間特定手段 4 9、動画データ編集手段 5 0、動画データ作成手段 5 1、音楽再生制御手段 5 2、A / D コンバータ 5 3、映像再生制御手段 5 4 を備えている。

【 0 0 3 3 】

<中央制御手段>

中央制御手段 4 2 は、カラオケ本体 4 0 を総合的に制御するための手段であり、例えば CPU 及びその周辺機器により構成されており、CPU 等が ROM 4 3 等に記憶されたプログラムに従って動作することにより、制御機能を発揮することができるようになってい

10

20

30

40

50

る。

【 0 0 3 4 】

< R O M / R A M >

R O M 4 3 は、カラオケ本体 4 0 を構成する各機器を制御するためのプログラムデータや数値データを記憶するための機器で、例えば半導体メモリ等で構成される。また、R A M 4 4 は、プログラムや各種データを一時的に記憶する一時記憶領域として機能するもので、例えば半導体メモリ等で構成される。なお、物理的な半導体メモリにより R A M 4 4 を構成するのではなく、ハードディスク記憶装置等を用いて仮想的な R A M 4 4 を構成してもよい。

【 0 0 3 5 】

本実施形態では、R A M 4 4 に、予約待ち行列 4 4 a、動画データ 4 4 b、置換映像データ 4 4 c が記憶されるようになっている。なお、予約待ち行列 4 4 a は、選曲予約されたカラオケ楽曲について、演奏順に楽曲 I D を並べて構成されたデータテーブルであり、選曲者の利用者 I D、動画撮影を行う旨のフラグ等、他の識別データが紐付けられている場合もある。

10

【 0 0 3 6 】

< 動画データ >

動画データ 4 4 b は、ビデオカメラ 3 2 により撮像され、ビデオ信号として入力された画像データと、マイクロホン 3 4 から入力され、A / D コンバータ 5 3 によりデジタル変換された歌唱音声データとを動画データ作成手段 5 1 において同期させたデータである。この動画データ 4 4 b は、D V D 等の記録媒体に記録したり、管理サーバ 6 0 等にアップロードして公開したりすることができる。

20

【 0 0 3 7 】

< 置換映像データ >

置換映像データ 4 4 c は、コマンド入力のためのジェスチャーであると特定された撮影区間における動画データ 4 4 b と置き換えるための映像データ素材であり、例えば、予め用意されたムード映像、カラオケを一緒に楽しんでいる他の利用者の映像等からなる。

【 0 0 3 8 】

< H D D >

H D D 4 5 は、大容量記憶装置として機能するもので、少なくとも、楽曲データベース 4 5 a、映像データベース 4 5 b、コマンドデータベース 4 5 c が格納されている。なお、H D D 4 5 に替えて、あるいは H D D 4 5 と共に、データを書き替え可能な D V D 等の大容量記憶装置を用いてもよい。

30

【 0 0 3 9 】

< 楽曲データベース / 映像データベース >

楽曲データベース 4 5 a は、演奏制御データ (M I D I 規格のデータ) 及び歌詞描出データが同期されて構成される楽曲データについて、楽曲 I D と対応付けてそれぞれ構成されたデータベースである。演奏制御データは、各楽曲の演奏を制御するためのデジタルデータであり、歌詞描出データは演奏に同期した歌詞文字の表示タイミングデータ及び色変わりデータを含んでいる。映像データベース 4 5 b は、演奏されるカラオケ楽曲に対応した背景映像を、当該カラオケ楽曲の楽曲 I D に対応させた映像ファイルとして所定数格納したデータベースである。

40

【 0 0 4 0 】

< コマンドデータベース >

コマンドデータベース 4 5 c は、予め定められた歌唱者の動作と、当該歌唱者の動作に対応する所定のコマンドとを紐付けして管理するためのデータベースである。コマンドデータベース 4 5 c で管理されるコマンドは、カラオケ演奏端末 3 0 の操作に関するものであり、例えば、音量の増減、効果音 (拍手や歓声等) の発生、エコーの強弱の設定、演奏キーの変更、演奏テンポの変更等である。これらのコマンドと紐付けされるジェスチャーは、例えば、左手又は右手を横に振り出す、左手又は右手を上挙げる、両手を上に挙げ

50

る、これらの動作を連続して組み合わせる等、種々の動作とすることができる。また、コマンドデータベース45cに、コマンドの内容に応じた映像データを含ませてもよい。

【0041】

後に詳述するが、撮影手段で撮影した動作が、コマンド入力のためのジェスチャーである場合には、当該ジェスチャーに紐付けられたコマンド入力が行われるとともに、動画データ編集手段50の機能により、動画編集が行われる。ジェスチャーによるコマンド入力については、例えば、上述した特許文献1（特開平11-175061号公報）に記載された技術を応用することができる。

【0042】

<送受信手段>

ローカル送受信手段46は、カラオケ本体40とカラオケリモコン装置31との間で、データの送受信を行うための電子回路及びプログラムからなる。本実施形態では、赤外線通信により、カラオケ本体40とカラオケリモコン装置31との間でデータの送受信が行われる。また、本実施形態では、ルータ20との間でデータの送受信を行うためのネットワーク送受信手段41を備えている。

10

【0043】

<予約管理手段>

予約管理手段47は、利用者により楽曲検索手段31aの機能を用いて選曲された楽曲IDを演奏順に並べて予約待ち行列44aを作成し、この予約待ち行列44aをRAM44に格納して管理するためのプログラムからなる。

20

【0044】

<音楽再生制御手段>

音楽再生制御手段52は、楽曲IDに基づいて楽曲データベース45aから抽出された演奏制御データに基づいて、音源データをデジタル再生すると共にアナログ変換してミキシングアンプ36に出力するための電子回路である。上述したように、ミキシングアンプ36は、マイクロホン34から入力された歌唱者の歌唱音声信号と、音楽再生制御手段52から送られる演奏音声信号とをミキシングすると共に、アンプ機能により増幅してスピーカ33より出力させるための装置である。

【0045】

<映像再生制御手段>

映像再生制御手段54は、カラオケ楽曲の演奏中に、映像データベース45bから抽出した映像データと、楽曲データベース45aに含まれる歌詞描出データに基づいて作成される歌詞文字とを、当該カラオケ楽曲の楽曲データに同期させて表示装置35に出力するためのプログラムからなる。

30

【0046】

<動画データ作成手段>

動画データ作成手段51は、撮影手段（ビデオカメラ32及びその付帯機器）で撮影した歌唱者の歌唱姿勢及びマイクロホン34から入力され、A/Dコンバータ53でデジタル変換された歌唱音声を用いて動画データ44bを作成して保存するための電子機器及びプログラムからなる。上述したように、動画データ作成手段51で作成した動画データ44bは、HDD45等に保存されて、動画データ編集手段50における動画の編集、DVD等の記録媒体への記録、管理サーバ60における歌唱動画の公開等に使用される。

40

【0047】

<コマンド判定手段>

コマンド判定手段48は、歌唱者の動作映像を解析し、当該歌唱者の動作がコマンドデータベース45cに存在する場合に、当該歌唱者の動作をコマンド入力のためのジェスチャーであると判定するためのプログラムからなる。上述したように、歌唱者の動作映像は、ビデオカメラ32で撮影され、デジタルデータとしてコマンド判定手段48に入力される。コマンド判定手段48では、3次元解析等、周知の解析技術により解析し、コマンドデータベース45cに存在する動作（ジェスチャー）と比較することにより、歌唱者の動

50

作がコマンド入力のためのジェスチャーであるか否かを判定する。歌唱者の動作がコマンド入力のためのジェスチャーであると判定された場合には、中央制御手段33が音楽再生制御手段等に対し、音量の増減、効果音の発生、エコーの強弱の設定などの指示を出す。

【0048】

<コマンド入力区間特定手段>

コマンド入力区間特定手段49は、コマンド判定手段48の機能により、歌唱者の動作がコマンド入力のためのジェスチャーであると判定された場合に、動画撮影された歌唱者の歌唱姿態の中から当該コマンド入力が行われている撮影区間を特定するためのプログラムからなる。すなわち、コマンド入力区間特定手段49は、現に撮影されている歌唱者の歌唱姿態、あるいはHDD45等に保存されている動画データ44bについて、コマンド

10

【0049】

<動画データ編集手段>

動画データ編集手段50は、コマンド入力が行われていると特定された撮影区間の動画データを編集して、コマンド入力のためのジェスチャーが含まれない一連の動画データとして編集し、あるいは当該コマンド入力のためのジェスチャーを識別可能とするためのプログラムからなる。

【0050】

例えば、コマンド入力区間特定手段49の機能により、現に撮影されている歌唱者の歌唱姿態、あるいはHDD45等に保存されている動画データ44bに、コマンド入力が行

20

【0051】

また、現に撮影されている歌唱者の歌唱姿態、あるいはHDD45等に保存されている動画データ44bに、コマンド入力が行われていると特定された撮影区間が存在すると、当該撮影区間の映像において、他の映像に置き換える代わりに、ジェスチャーがコマンド入力のための動作であることを識別するための説明表示を動画データ44bに付加してもよい。

【0052】

本実施形態では、コマンド入力のためのジェスチャーと置き換える置換映像データ44cはRAM44に格納されている。この置換映像データ44cは、例えば、予め用意されたムード映像、カラオケを一緒に楽しんでいる他の利用者の映像等からなる。また、動画データ44bを直接加工して、歌唱者の顔を中心とした映像にデジタルズームアップした映像データ、歌唱者のストップモーション画像などに置き換えてもよい。

30

【0053】

動画データ44bの編集(置換データ44cへの置き換え)は、コマンド入力が行われていると特定された撮影区間を含む前後所定の撮影区間について行うことが好ましい。すなわち、コマンド入力が行われていると特定された撮影区間のみについて、他の映像データと置き換えた場合には、動画の繋がりが不自然なものとなる可能性がある。したがって、歌唱フレーズの区切り等を、動画編集を行う撮影区間の開始点及び終了点とすることにより、動画として不必要なコマンド入力のためのジェスチャーが他の映像に置き換えられるだけでなく、動画全体の繋がりが自然なものとなり、さらに一層、歌唱者及び視聴者の興味を高めることが可能となる。

40

【0054】

<ジェスチャー動画の編集/ジェスチャーの識別可能表示>

次に、図2~図4を参照して、コマンド入力のためのジェスチャーを他の映像に置き換え、あるいはジェスチャーを識別可能とする手順について説明する。

【0055】

ジェスチャー動画の編集処理では、図2に示すように、現に撮影している歌唱者の動作

50

映像、あるいは保存されている動画データ44bにおける歌唱者の動作映像を解析する(S1)。続いて、歌唱者の動作がコマンドデータベース45cに存在するか否かを判定し(S2)、歌唱者の動作がコマンドデータベース45cに存在する場合に、当該歌唱者の動作をコマンド入力のためのジェスチャーであると判定する(S3)。

【0056】

ここで、歌唱者の動作がコマンド入力のためのジェスチャーであると判定されると、現に撮影している歌唱者の動作映像、あるいは保存されている動画データ44bの中から当該コマンド入力が行われている撮影区間を特定する(S4)。一方、コマンド入力のためのジェスチャーが存在しなければ、撮影区間の特定は行わない。

【0057】

続いて、特定された撮影区間の動画データを他の映像データ(動画データ44bを直接加工した映像あるいは置換映像データ44c)に置き換え、あるいはコマンド入力のためのジェスチャーであることを識別可能とする(S5)。そして、特定された撮影区間のすべてにおいて、動画データの置き換え又は識別可能とする処理が終了したか否か(特定された撮影区間が、さらに存在するか否か)を判断し(S6)、特定された撮影区間がさらに存在すれば、順次、画像の置き換え、あるいはコマンド入力のためのジェスチャーであることを識別可能とする処理を行う(S5)。そして、すべての撮影区間に画像について画像の置き換え、あるいはコマンド入力のためのジェスチャーであることを識別可能とする処理が終了したら、動画データ44bの編集処理を終了する。

【0058】

画像の置き換え処理は、図3に示すように、コマンド入力のためのジェスチャーが行われていると特定された撮影区間を含む所定区間について、当該ジェスチャーの動画データ44bを他の映像データ(置換映像データ44c)に置き換えて、歌唱者の顔画像のズームアップを表示する等、画像データの編集を行う。

【0059】

また、コマンド入力のためのジェスチャーであることを識別可能とする処理は、図4に示すように、コマンド入力のためのジェスチャーが行われている映像に対して、当該コマンドの内容を説明する説明映像を合成すればよい。

【0060】

<他の実施形態>

本発明のシステム及びその周辺装置を構成する機器や手段は上述したものに限定されず、その利用目的に応じて、必要な機器や手段のみの構成としたり、適宜他の機器や手段を付加したりすることができる。また、各手段をそれぞれ別個のものとして構成するのではなく、複数の機能を統合した手段として構成してもよい。

【符号の説明】

【0061】

- 10 コマンド入力識別システム
- 20 ルータ
- 30 カラオケ演奏端末
- 31 カラオケリモコン装置
- 31a 楽曲検索手段
- 31b 楽曲索引データベース
- 31c データ記憶部
- 31d 入出力表示部
- 32 ビデオカメラ
- 33 スピーカ
- 34 マイクホン
- 35 表示装置
- 36 ミキシングアンプ
- 40 カラオケ本体

10

20

30

40

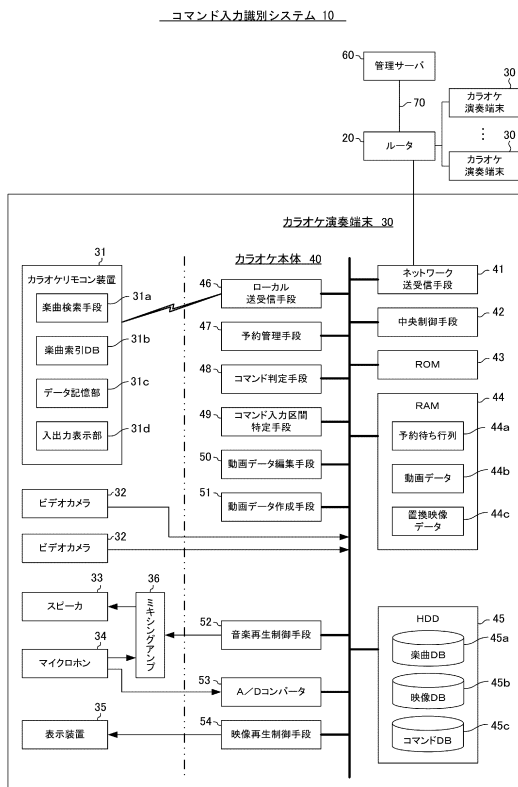
50

- 4 1 ネットワーク送受信手段
- 4 2 中央制御手段
- 4 3 R O M
- 4 4 R A M
- 4 4 a 予約待ち行列
- 4 4 b 動画データ
- 4 4 c 置換映像データ
- 4 5 H D D
- 4 5 a 楽曲データベース
- 4 5 b 映像データベース
- 4 5 c コマンドデータベース
- 4 6 ローカル送受信手段
- 4 7 予約管理手段
- 4 8 コマンド判定手段
- 4 9 コマンド入力区間特定手段
- 5 0 動画データ編集手段
- 5 1 動画データ作成手段
- 5 2 音楽再生制御手段
- 5 3 A / Dコンバータ
- 5 4 映像再生制御手段
- 6 0 管理サーバ
- 7 0 データ送受信回線

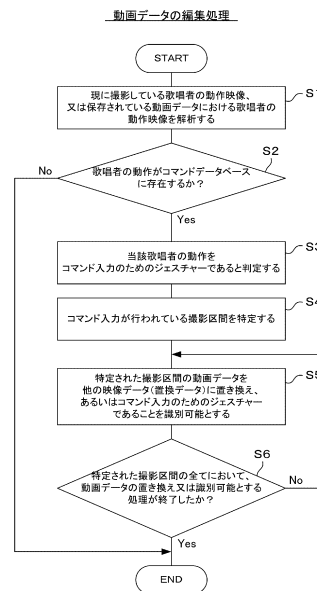
10

20

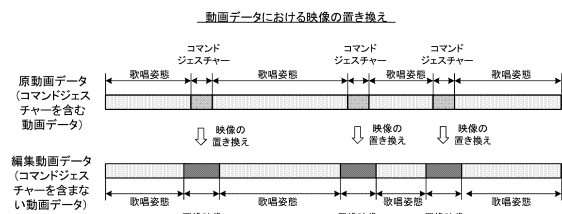
【 図 1 】



【 図 2 】



【 図 3 】



【 図 4 】



フロントページの続き

- (56)参考文献 特開平11-175061(JP,A)
特開2009-198714(JP,A)
国際公開第2009/044525(WO,A1)
国際公開第2012/147960(WO,A1)
特開2002-318588(JP,A)
特開2008-192004(JP,A)
米国特許出願公開第2012/0268374(US,A1)

(58)調査した分野(Int.Cl., DB名)

G10K 15/04
G06F 3/01
G06F 3/0481