

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5201540号
(P5201540)

(45) 発行日 平成25年6月5日(2013.6.5)

(24) 登録日 平成25年2月22日(2013.2.22)

(51) Int.Cl.

F I

G10K	15/04	(2006.01)	G10K	15/04	302D
H04N	5/225	(2006.01)	H04N	5/225	F
H04N	5/232	(2006.01)	H04N	5/232	Z
G06T	1/00	(2006.01)	H04N	5/225	A
			G06T	1/00	340A

請求項の数 6 (全 22 頁)

(21) 出願番号 特願2010-167541 (P2010-167541)
 (22) 出願日 平成22年7月26日(2010.7.26)
 (65) 公開番号 特開2012-27340 (P2012-27340A)
 (43) 公開日 平成24年2月9日(2012.2.9)
 審査請求日 平成23年9月6日(2011.9.6)

(73) 特許権者 000005267
 ブラザー工業株式会社
 愛知県名古屋市瑞穂区苗代町15番1号
 (74) 代理人 100104503
 弁理士 益田 博文
 (72) 発明者 戸田 勝巳
 名古屋市瑞穂区苗代町15番1号 ブラザー工業株式会社内
 審査官 柴垣 俊男

最終頁に続く

(54) 【発明の名称】 カラオケ装置及びカラオケ歌唱者の静止画出力方法

(57) 【特許請求の範囲】

【請求項1】

楽曲データ及び映像データを用いて、カラオケ演奏曲の再生サービスを提供するカラオケ装置であって、

歌唱者が歌唱するための前記楽曲データを再生する楽曲再生手段と、

前記楽曲再生手段により前記楽曲データの再生が行われるのにしたが、前記映像データを表示可能な表示手段と、

前記歌唱者以外の視聴者及び前記歌唱者を含む所定範囲の視野を撮影し、当該所定範囲の映像データを生成する少なくとも1台の動画撮影カメラと、

前記動画撮影カメラにより生成された前記所定範囲の映像データにより、顔が前記歌唱者の方向を向いている前記視聴者の人数を算出する算出手段と、

前記算出手段により算出される前記人数の時間的推移により、前記歌唱者及び前記視聴者を含む集団の盛り上がりタイミングを決定する決定手段と、
 を有することを特徴とするカラオケ装置。

【請求項2】

請求項1記載のカラオケ装置において、

前記歌唱者に所持され、当該歌唱者によるカラオケ歌唱の音声信号を入力するためのマイクロフォンと、

前記マイクロフォンに設けられ、標識信号を発生する標識信号発生手段と、

前記動画撮影カメラは、1台設けられるとともに、

10

20

前記マイクロフォン及び前記歌唱者を含む前記所定範囲の視野を自装置を中心とした全周を撮影し、前記標識信号発生手段から発生された前記標識信号を含む前記所定範囲の映像データを生成し、

前記算出手段は、

前記1台の動画撮影カメラにより生成された前記所定範囲の映像データに含まれる前記標識信号に基づいて前記マイクロフォン及び前記歌唱者の位置を特定するとともに、前記所定範囲のうち前記歌唱者の位置以外の範囲の映像データに所定の顔認識処理を行って前記視聴者の顔の向きを決定することにより、顔が前記歌唱者の方向を向いている前記視聴者の人数を算出する

ことを特徴とするカラオケ装置。

10

【請求項3】

請求項2記載のカラオケ装置において、

前記所定範囲の映像データから、前記決定手段により決定された前記盛り上がりタイミングにおける前記歌唱者を含む少なくとも1つの部分静止画を前記全周の中心に対して劣弧をとる扇形に切り出すとともに、その切り出した前記扇形の劣弧を直線に補正することで、前記扇形を四角形に補正する補正処理を行って、1つの静止画とする映像処理手段と、

前記静止画を、当該カラオケ装置にネットワーク接続されたサーバへ出力する静止画出力手段と

を有することを特徴とするカラオケ装置。

20

【請求項4】

請求項3記載のカラオケ装置において、

前記映像処理手段により補正処理された前記少なくとも1つの静止画に対し所定の顔認識処理を行い、前記歌唱者の顔を認識できるかどうかを判定する判定手段を有し、

前記静止画出力手段は、

前記映像処理手段により補正処理された前記少なくとも1つの静止画のうち、前記判定手段により前記歌唱者の顔を認識できると判定された静止画を、前記サーバへ出力することを特徴とするカラオケ装置。

【請求項5】

請求項4記載のカラオケ装置において、

外部から歌唱者IDを入力する歌唱者ID入力手段を有し、

前記映像処理手段はさらに、

前記所定範囲の映像データから、前記特定された前記歌唱者の位置を含む部分映像データを切り出して所定の補正処理を行い、

前記静止画出力手段は、

歌唱者ID入力手段より入力された前記歌唱者IDと、前記判定手段により前記歌唱者の顔を認識できると判定された静止画と、前記映像処理手段による補正処理後の前記部分映像データと、を対応付けて、当該カラオケ装置にネットワーク接続された前記サーバへ出力する

ことを特徴とするカラオケ装置。

40

【請求項6】

カラオケ演奏曲を再生するカラオケ装置に備えられたコンピュータが実行する、当該カラオケ演奏曲の歌唱者を含む静止画を生成して出力するためのカラオケ歌唱者の静止画出力方法であって、

少なくとも1台の動画撮影カメラにより撮影され生成された、前記歌唱者以外の視聴者及び前記歌唱者を含む所定範囲の映像データを取得する取得手順と、

前記取得手順で取得された前記所定範囲の映像データにより、顔が前記歌唱者の方向を向いている前記視聴者の人数を算出する算出手順と、

前記算出手順で算出された前記人数の時間的推移により、前記歌唱者及び前記視聴者を

50

含む集団の盛り上がりタイミングを決定する決定手順と、

前記取得手順で取得された前記所定範囲の映像データから、前記決定手順で決定された前記盛り上がりタイミングにおける前記歌唱者を含む少なくとも1つの静止画を抽出する抽出手順と、

前記抽出手順で抽出された前記静止画、又は、当該静止画を補正した後の静止画を、前記カラオケ装置に接続されたサーバへ出力する出力手順と、を有することを特徴とするカラオケ歌唱者の静止画出力方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、カラオケ演奏曲の再生サービスを提供するカラオケ装置、及び、このカラオケ装置でのカラオケ歌唱者の静止画を出力するためのカラオケ歌唱者の静止画出力方法に関する。

【背景技術】

【0002】

カラオケ装置において、カラオケ演奏曲の再生以外に、これに付帯する種々様々なサービスが既に提唱されている。例えば、歌唱者の歌唱している姿を1台のビデオカメラで撮影してディスプレイに表示するカラオケ装置が、従来既に提唱されている（例えば、特許文献1参照）。この従来のカラオケ装置には、1台のビデオカメラと、このビデオカメラの向きを所定範囲内で自在に変化させるためのサーボ式雲台と、歌唱者の持つワイヤレスマイクロフォンからの無線標識信号を受信し、その無線標識信号の発信源の位置にビデオカメラの向きを合わせるようにサーボ式雲台を駆動する駆動制御手段と、が備えられている。

【0003】

また、近年、娯楽性の向上のために、上記のようにして撮影した歌唱中の映像を録画してサーバへアップロードし、当該歌唱者や他のユーザがその歌唱映像（動作）をパソコンや携帯端末より閲覧して楽しめるようにしたサービスが、既に提供されている。通常、このようなサービスでは、閲覧したい動画を選択するためのウェブページにおいて、各動画中のある1場面の静止画が適宜に抽出選択され、サムネイルとして表示されている。

【0004】

アップロードされる動画の中からサムネイルとして選択されるべき静止画は、動画内容を代表する代表画像という意味では、本来、歌唱者を含む利用者の集団がもっとも盛り上がっているタイミングの静止画がふさわしい。このような、特定の静止画を動画から選択して抽出する技術としては、例えば、特許文献2や特許文献3に記載の手法が既に提唱されている。これらの従来技術の手法では、音量の大きいところを盛り上がりタイミングと判断し、その盛り上がりタイミングの静止画を抽出するようになっている。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開平10-240276号公報

【特許文献2】特許第4198331号

【特許文献3】特許第4435130号

【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、上記特許文献2や特許文献3に記載の技術を上記カラオケ動画における静止画の選択に適用した場合、音量を基準として盛り上がりタイミングを判定することから、録画されたカラオケ動画における、視聴者が歌唱者の歌を聴かず大声で雑談している場合も盛り上がりタイミングと判定される可能性がある。したがって、カラオケ装置の利用者の盛り上がりタイミングを正確に判定し、当該タイミングにおける最適な静止画

10

20

30

40

50

を確実に得ることは難しかった。

【0007】

本発明の目的は、カラオケ装置の利用者の盛り上がりタイミングの静止画を確実に抽出することで娯楽性を向上できる、カラオケ装置、及び、カラオケ歌唱者の静止画出力方法を提供することにある。

【課題を解決するための手段】

【0008】

上記目的を達成するために、第1の発明は、楽曲データ及び映像データを用いて、カラオケ演奏曲の再生サービスを提供するカラオケ装置であって、歌唱者が歌唱するための前記楽曲データを再生する楽曲再生手段と、前記楽曲再生手段により前記楽曲データの再生が行われるのにしたが、前記映像データを表示可能な表示手段と、前記歌唱者以外の視聴者及び前記歌唱者を含む所定範囲の視野を撮影し、当該所定範囲の映像データを生成する少なくとも1台の動画撮影カメラと、前記動画撮影カメラにより生成された前記所定範囲の映像データにより、顔が前記歌唱者の方向を向いている前記視聴者の人数を算出する算出手段と、前記算出手段により算出される前記人数の時間的推移により、前記歌唱者及び前記視聴者を含む集団の盛り上がりタイミングを決定する決定手段とを有することを特徴とする。

10

【0009】

本願第1発明においては、娯楽性の向上のために、歌唱者及び視聴者の集団の盛り上がりを検出する。まず、少なくとも1台の動画撮影カメラが、歌唱者及び視聴者を含む所定範囲の視野を撮影して、当該所定範囲の映像データを生成する。そして、生成された映像データに基づき、算出手段が、顔が歌唱者の方向を向いている視聴者の人数を算出する。

20

【0010】

ここで、集団が盛り上がっているときとは、歌唱者の歌唱によって視聴者が心より楽しんでいるときであり、その瞬間には、視聴者の視線が歌唱者のほうへ向いているのが通常である。そこで、本願第1発明においては、決定手段が、算出手段が算出した視聴者の人数の時間的推移により、盛り上がりタイミングを決定する。これにより、顔が歌唱者へ向いている視聴者の人数が最も多い瞬間を、当該集団の盛り上がりタイミングと決定することができる。この結果、例えば、上記生成された映像データのうち、当該盛り上がりタイミングにおける静止画を抽出し、当該サービス提供時間の代表画像としてサーバへアップロードしたり、カラオケ装置の表示手段に代表画像として映し出したり、等種々のサービスを行うことが可能となる。これにより、カラオケ装置の娯楽性をさらに向上することができる。

30

【0011】

第2発明は、上記第1発明において、前記歌唱者に所持され、当該歌唱者によるカラオケ歌唱の音声信号を入力するためのマイクロフォンと、前記マイクロフォンに設けられ、標識信号を発生する標識信号発生手段と、前記動画撮影カメラは、1台設けられるとともに、前記マイクロフォン及び前記歌唱者を含む前記所定範囲の視野を自装置を中心とした全周を撮影し、前記標識信号発生手段から発生された前記標識信号を含む前記所定範囲の映像データを生成し、前記算出手段は、前記1台の動画撮影カメラにより生成された前記所定範囲の映像データに含まれる前記標識信号に基づいて前記マイクロフォン及び前記歌唱者の位置を特定するとともに、前記所定範囲のうち前記歌唱者の位置以外の範囲の映像データに所定の顔認識処理を行って前記視聴者の顔の向きを決定することにより、顔が前記歌唱者の方向を向いている前記視聴者の人数を算出することを特徴とする。

40

【0012】

本願第2発明においては、カラオケ演奏時には、楽曲再生手段によって楽曲データの再生が行われるとともに表示手段により映像データが表示され、それら再生及び表示に合わせて、歌唱者がマイクロフォンにより歌唱を行う。また、動画撮影カメラは1台のみ備えられる。上記歌唱の際には、その1台の動画撮影カメラが、歌唱者及びマイクロフォンを含む所定範囲の視野を撮影し、当該所定範囲の映像データを生成する。このとき、自装置

50

を中心とした全周を撮影可能な、視野が広い動画撮影カメラ（例えば魚眼レンズを備え全周360°撮影可能なカメラ）を用いることにより、歌唱者及び視聴者の全員が、動画撮影カメラが生成した映像データの中に常に含まれる。

【0013】

ここで、マイクロフォンには、標識信号を発生する標識信号発生手段が備えられている。したがって、上記生成された所定範囲の映像データには、歌唱者の所持したマイクロフォンの位置に対応した標識信号が、歌唱者の姿と共に必ず記録されている。そこで、これに対応して、算出手段が、映像データに含まれる標識信号を用いてマイクロフォン及び歌唱者の位置を特定するとともに、歌唱者の位置以外の映像データに対し顔認識処理を行って視聴者の顔の向きを決定する。これにより、複数台のカメラを用いなくても、顔が歌唱者の方向を向いている視聴者の人数を確実に算出することができる。

10

【0014】

第3発明は、上記第2発明において、前記所定範囲の映像データから、前記決定手段により決定された前記盛り上がりタイミングにおける前記歌唱者を含む少なくとも1つの部分静止画を前記全周の中心に対して劣弧をとる扇形に切り出すとともに、その切り出した前記扇形の劣弧を直線に補正することで、前記扇形を四角形に補正する補正処理を行って、1つの静止画とする映像処理手段と、前記静止画を、当該カラオケ装置にネットワーク接続されたサーバへ出力する静止画出力手段とを有することを特徴とする。

【0015】

決定手段により決定された盛り上がりタイミングにおける静止画を生成する場合、上記のように視野が広い動画撮影カメラの映像データから得た静止画は視野の端部ほど歪んだ状態となっている場合がある。そこで本願第3発明においては、映像処理手段が、上記所定範囲の映像データから、当該盛り上がりタイミングにおける歌唱者を含む部分静止画を全周の中心に対して劣弧をとる扇形に切り出すと共に、その切り出した部分静止画に所定の補正処理を行う。補正処理としては、切り出した扇形の劣弧を直線に補正し、扇形を四角形とする。これにより、上記歪んだ状態が是正された正常な静止画を、静止画出力手段が代表画像としてサーバへ出力することができる。この結果、サーバへアクセスした各ユーザが、当該静止画を閲覧し、楽しむことができる。

20

【0016】

第4発明は、上記第3発明において、前記映像処理手段により補正処理された前記少なくとも1つの静止画に対し所定の顔認識処理を行い、前記歌唱者の顔を認識できるかどうかを判定する判定手段を有し、前記静止画出力手段は、前記映像処理手段により補正処理された前記少なくとも1つの静止画のうち、前記判定手段により前記歌唱者の顔を認識できると判定された静止画を、前記サーバへ出力することを特徴とする。

30

【0017】

歌唱者及び視聴者の集団が盛り上がっている盛り上がりタイミングにおいて、たまたま歌唱者が下を向いたり横を向いたりしている場合がある。このようなタイミングにおける静止画は、歌唱者の顔が正しく記録されていないため、代表画像としては必ずしも好ましくない。そこで本願第4発明においては、判定手段が、補正処理後の少なくとも1つの静止画それぞれに対し、歌唱者の顔を認識できるかどうかを判定する。そして、静止画出力手段は、歌唱者の顔が認識できると判定された静止画をサーバへ出力する。これにより、歌唱者の顔が正しく記録された静止画のみを、確実に代表画像としてサーバへ出力することができる。

40

【0018】

第5発明は、上記第4発明において、外部から歌唱者IDを入力する歌唱者ID入力手段を有し、前記映像処理手段はさらに、前記所定範囲の映像データから、前記特定された前記歌唱者の位置を含む部分映像データを切り出して所定の補正処理を行い、前記静止画出力手段は、歌唱者ID入力手段より入力された前記歌唱者IDと、前記判定手段により前記歌唱者の顔を認識できると判定された静止画と、前記映像処理手段による補正処理後の前記部分映像データと、を対応付けて、当該カラオケ装置にネットワーク接続された前

50

記サーバへ出力することを特徴とする。

【0019】

これにより、歌唱者の顔が写っている静止画と、当該歌唱者の歌唱者IDと、その歌唱者が歌っている映像とが、サーバへアップロードされる。この結果、当該歌唱者やその他の各ユーザが、歌唱者IDを用いて当該静止画を検索して閲覧したり、さらに対応する歌唱映像を閲覧して、楽しむことができる。

【0020】

上記目的を達成するために、第6の発明は、カラオケ演奏曲を再生するカラオケ装置に備えられたコンピュータが実行する、当該カラオケ演奏曲の歌唱者を含む静止画を生成して出力するためのカラオケ歌唱者の静止画出力方法であって、少なくとも1台の動画撮影カメラにより撮影され生成された、前記歌唱者以外の視聴者及び前記歌唱者を含む所定範囲の映像データを取得する取得手順と、前記取得手順で取得された前記所定範囲の映像データにより、顔が前記歌唱者の方向を向いている前記視聴者の人数を算出する算出手順と、前記算出手順で算出された前記人数の時間的推移により、前記歌唱者及び前記視聴者を含む集団の盛り上がりタイミングを決定する決定手順と、前記取得手順で取得された前記所定範囲の映像データから、前記決定手順で決定された前記盛り上がりタイミングにおける前記歌唱者を含む少なくとも1つの静止画を抽出する抽出手順と、前記抽出手順で抽出された前記静止画、又は、当該静止画を補正した後の静止画を、前記カラオケ装置に接続されたサーバへ出力する出力手順と、を有することを特徴とする。

【0021】

本願第6発明においては、少なくとも1台の動画撮影カメラが、歌唱者及び視聴者を含む所定範囲の視野を撮影して、当該所定範囲の映像データを生成する。そして、生成された映像データに基づき、算出手段で、顔が歌唱者の方向を向いている視聴者の人数が算出される。集団が盛り上がっているときは、歌唱者の歌唱によって視聴者が心より楽しんでいるときであるから、その瞬間には、視聴者の視線が歌唱者のほうへ向いている。そこで、決定手段で、算出手順において算出され視聴者の人数の時間的推移により、盛り上がりタイミングが決定される。これにより、顔が歌唱者へ向いている視聴者の人数が最も多い瞬間を、当該集団の盛り上がりタイミングと決定することができる。これにより、上記取得手順で取得された映像データのうち、当該盛り上がりタイミングにおける静止画が抽出手順で抽出され、当該抽出された静止画（又はそれを補正した静止画）をサービス提供時間の代表画像として出力手順でサーバへアップロードすることができる。この結果、カラオケ装置の娯楽性をさらに向上することができる。

【発明の効果】

【0022】

本発明によれば、カラオケ装置の利用者の盛り上がりタイミングの静止画を確実に抽出し、娯楽性を向上することができる。

【図面の簡単な説明】

【0023】

【図1】本発明の一実施の形態のカラオケ装置が設置されたカラオケルームを概略的に表す図である。

【図2】マイクの外観を表す側面図である。

【図3】カラオケ装置を備えたカラオケシステムの全体構成を表す機能ブロック図である。

【図4】全周カメラから入力された映像を画像処理して歌唱者の動画データを得るプロセスを表す説明図である。

【図5】全周カメラによりカラオケルーム内を撮影したときに取得される画像を模式的に表す図である。

【図6】視聴者の1人が歌唱者を見ていない状態において、全周カメラによりカラオケルーム内を撮影したときに取得される画像を模式的に表す別の図である。

【図7】視聴者の1人が歌唱者を見ていない状態でのカラオケルームを概略的に表す図で

10

20

30

40

50

ある。

【図 8】 タイミングログの一例を示す説明図である。

【図 9】 装置本体の制御部により実行される処理手順の詳細を表すフローチャートである。

【図 10】 ログ作成処理の手順の詳細を表すフローチャートである。

【図 11】 ホストサーバにアップロードされたサムネイルの表示例を表す図である。

【図 12】 歌唱者の顔が写っている静止画を抽出する変形例において、装置本体の制御部により実行される処理手順の詳細を表すフローチャートである。

【図 13】 音声レベルを考慮して抽出する変形例において、装置本体の制御部により実行される処理手順の詳細を表すフローチャートである。

10

【発明を実施するための形態】

【0024】

以下、本発明の一実施の形態を図面を参照しつつ説明する。

【0025】

図 1 は、本実施形態のカラオケ装置が設置されたカラオケルームを概略的に表す図である。

【0026】

図 1 において、カラオケ店舗等のカラオケルーム KR には、カラオケ装置 10 が設置されている。カラオケ装置 10 は、楽曲データとしての Musical Instrument Digital Interface (MIDI; 登録商標) データ及び映像データを用いて、カラオケ演奏曲の再生サービスを提供する装置である。図 1 に示すものでは、利用者 A ~ C がカラオケ歌唱を行っている。カラオケ装置 10 は、コマンドと称される装置本体 100 と、リモコン 200 と、マイクロフォン (以下、略してマイク) 300 と、1 台の全周カメラ 400 (動画撮影カメラ) とを有している。装置本体 100、リモコン 200、及びマイク 300 については、後で詳述する。

20

【0027】

全周カメラ 400 は、この例では、例えばカラオケルーム KR の天井中心に下向きに備え付けられている。全周カメラ 400 は、マイク 300 及び歌唱者 A ~ C を含む所定範囲の固定的な視野を撮影し、その範囲の映像データを生成する。具体的には、全周カメラ 400 は、魚眼レンズを有し、当該カメラ 400 を中心とした、左右 360 度、上下 90 度の半球状視野を 1 枚に収めた映像データを得る。魚眼レンズの機能によって、全周カメラ 400 の視野内の像は、近い物体ほど円の中心に写り、遠い物体ほど円の周辺部に写る性質を持つ。また、魚眼レンズにより広い視野角が得られる代償として、全ての物体は扇状に歪曲して写るようになる。従って、全周カメラ 400 によりカラオケルーム KR 内の全体を撮影すると、全体的に扇状に歪曲した画像が取得されることとなる。

30

【0028】

図 2 は、マイク 300 の外観を表す図である。図 2 において、マイク 300 は、利用者によるカラオケ歌唱の音声を音声信号に変換して入力するものである。

【0029】

マイク 300 は、マイク素子 301 が内蔵されたマイクハウジング 302 を有している。マイクハウジング 302 の上部には、電源スイッチ 303 が設けられている。マイクハウジング 302 の下部には、所定の色 (この例では、緑色) の標識信号としてのマーカ光を発光する、標識信号発生手段としての発光ダイオード (LED) 304 と、この LED 304 から発した緑色のマーカ光を均一に拡散させる半透明の光拡散球 305 とが設けられている。

40

【0030】

図 3 は、上記のカラオケ装置 10 を備えたカラオケシステムの全体構成を表す機能ブロック図である。

【0031】

図 3 において、カラオケシステム 1 は、上記カラオケルーム KR に設置された上記カラ

50

オケ装置 10 と、ホストサーバ 20 とを有している。カラオケ装置 10 とホストサーバ 20 とは、例えば通信ネットワーク等のネットワーク NW とを介し、互いに情報送受信可能に接続されている。

【0032】

カラオケ装置 10 は、上記の装置本体 100、リモコン 200、マイク 300、及び全周カメラ 400 を有している。装置本体 100 とリモコン 200 とは、例えば無線又は有線の LAN 等のネットワークを介し、互いに情報送受信可能に接続されている。装置本体 100 とマイク 300 とは、無線回線又は有線回線により接続されている。

【0033】

装置本体 100 は、制御部 101 と、大容量記憶装置 103 と、操作部 104 と、受信部 105 と、音源 106 と、音声制御部 107 と、スピーカ 108 と、表示部 109 と、通信制御部 110 とを有している。

10

【0034】

制御部 101 は、図示しない CPU や、RAM 及び ROM 等のメモリを備えている。この制御部 101 は、RAM の一時記憶機能を利用しつつ、ROM や上記大容量記憶装置 103 に予め記憶された各種プログラムを実行する。これにより、装置本体 100 全体の制御を行う。

【0035】

特に、制御部 101 は、全周カメラ 400 により得られたカラオケルーム KR 内の撮像画像に対して所定の画像処理を行い、マイク 300 を持った歌唱者の複数の画像（集合体としての動画を構成する複数の静止画。以下同様）を生成し、その画像を大容量記憶装置 103 に記憶するとともに表示部 109 に表示させる処理を行う（詳細は後述）。

20

【0036】

大容量記憶装置 103 は、例えば Hard Disk Drive (HDD) などから構成される。この大容量記憶装置 103 には、MIDI データ、背景映像データ、及び歌詞データ等の各種情報が記憶されている。また、この大容量記憶装置 103 には、利用者の歌唱時の動画データが順次記憶される。

【0037】

操作部 104 は、例えば複数のキーやスイッチなどから構成される。利用者は、この操作部 104 又は後述のリモコン 200 の操作部 204 を用いて、カラオケ演奏曲の予約操作等の各種操作を行うことができる。

30

【0038】

受信部 105 は、上記のマイク 300 から出力された歌唱者の音声信号を受信する。

【0039】

音源 106 は、上記制御部 101 によって大容量記憶装置 103 から読み出された MIDI データを再生して音声制御部 107 へ出力する。音声制御部 107 は、音源 106 から出力された MIDI データ、及び、受信部 105 を介してマイク 300 により入力された音声信号を増幅し、スピーカ 108 へ出力する。スピーカ 108 は、音声制御部 107 から出力された MIDI データ及び音声信号を音声出力する。

【0040】

なお、以下適宜、音源 106、音声出力部 107、及びスピーカ 108 を、省略して「音源 106 等」と称する。音源 106 等は、楽曲データを再生する楽曲再生手段を構成している。

40

【0041】

表示部 109 は、例えば液晶ディスプレイなどから構成され、各種映像を表示する表示手段として機能する。特に、表示部 109 は、上記音源 106 等による MIDI データの再生に同期して、言い換えれば、音源 106 等により MIDI データの再生が行われるのに従い、大容量記憶装置 103 から読み出された背景映像データ、及び歌詞データに対応したテロップ等を表示することができる。

【0042】

50

通信制御部 110 は、リモコン 200 やホストサーバ 20 との間で情報通信の制御を行う。

【0043】

リモコン 200 は、利用者がカラオケ演奏曲の予約操作等の各種操作を行うための操作端末である。このリモコン 200 は、制御部 201 と、記憶装置 203 と、操作部 204 と、表示部 209 と、通信制御部 210 とを有している。

【0044】

制御部 201 は、図示しない CPU や RAM 及び ROM 等のメモリを備えている。この制御部 201 は、RAM の一時記憶機能を利用しつつ、ROM や上記記憶装置 203 に予め記憶された各種プログラムを実行する。これにより、リモコン 200 全体の制御を行う。

10

【0045】

記憶装置 203 は、例えば不揮発性メモリなどから構成され、各種情報を記憶する。操作部 204 は、例えば複数のキーやスイッチなどから構成される。利用者は、この操作部 204 又は上記カラオケ装置 100 の操作部 104 を用いて、カラオケ演奏曲の予約操作等の各種操作を行うことができる。表示部 209 は、例えば液晶ディスプレイなどから構成され、各種表示を行う。

【0046】

通信制御部 210 は、装置本体 100 やホストサーバ 20 との間で情報通信の制御を行う。

20

【0047】

ホストサーバ 20 には、利用者の歌唱中の姿の動画データが圧縮動画ファイルとしてアップロード可能である（詳細は後述）。このホストサーバ 20 にアップロードされた動画データは、所定の Web ページにおいて特定の利用者の端末より閲覧可能となっている（後述の図 11 も参照）。

【0048】

ここで、本実施形態の特徴の 1 つとして、全周カメラ 400 により得られたカラオケルーム KR 内の映像画像に含まれるマーカ光に基づいてマイク 300 の位置が特定され、そのマイク 300 の位置を含む部分映像が切り出され、マイク 300 を持った歌唱者の動画（カラオケ投稿動画）データが取得される。このとき、全周カメラ 400 で撮像して得られた映像信号は、人間の通常の視野とは大きく異なるので、カラオケ投稿動画の用途としてそのまま使うことはできない。このため、全周カメラ 400 で撮像して得られた映像信号に対して所定の処理を施す必要がある。

30

【0049】

図 4 (a) ~ (f) は、全周カメラ 400 より入力された映像を画像処理して歌唱者の動画データを得るプロセスを表す説明図である。本処理は、装置本体 100 の制御部 101 によって実行される。

【0050】

図 4 (a) に示すように、まず全周カメラ 400 で取得したカラオケルーム KR 内の映像を入力する。ここでは、利用者 A ~ C 及びマイク 300 の映像のみを表し、テーブルや装置本体 100 等の映像は省略してある。

40

【0051】

その後、図 4 (b) に示すように、全周カメラ 400 より入力された映像において、所定の色（この例では緑）の成分以外の成分を除去するカラーフィルタ処理を行う。具体的には、色フィルタ（ここでは緑フィルタ）を通して、全周カメラ 400 より入力された映像から緑色の成分のみを抽出する。色フィルタは、RGB の G 値のみを通過させるか、又は、YUV の UV が一定範囲内にある画素値のみを通過させる、CPU 演算処理による画素データファイルである。

【0052】

その後、図 4 (c) に示すように、カラーフィルタ処理が行われた映像データを輝度フ

50

ィルタに通し、輝度が一定以上の値を示す画素値のみを通過させることで、画像データの2値化を行う。これにより、画像の中の「純粋な緑色に近く、一定以上の明るさがある」画素のみが「1」を示し、それ以外の画素は「0」を示すビットマップが得られる。

【0053】

その後、図4(d)に示すように、全周カメラ400より入力された映像についてエリア判定を行う。具体的には、予め蜘蛛の巣状に定義されたマップに従い、角度方向(人間の視覚での左右に相当)に対して8分解(A~H)、距離方向(人間の視覚での奥行きに相当)に対して3分解(1~3)又は4分解(1~4)の計28分解された各エリアについて、エリアごとにビットマップの画素値を全て加算する。この加算値が最も大きい値(図中ではエリアG2)がマーカ光を検知しており、撮影すべき歌唱者がいるエリア(方向)であると判定される。

10

【0054】

その後、図4(e)に示すように、図4(d)に示す処理で選択されたエリアについて、扇状スキャンによる画像の形状補正を行う。具体的には、エリア内にある画素を同図に示した走査線に従って並べなおす処理を行う。これによって、扇形状の直径方向はY軸、円周方向はX軸の矩形形状に変形し矯正される。ここで、中心部に近い走査線は短く、円周部に近い走査線は長い、同一値の画素で補完して拡大するか、画素を省略して縮小し、一定長の線データを得る。走査線の座標パターンは、エリアA1~H4ごとに予め用意されている。

【0055】

20

全ての走査線について変換処理が終了すると、図4(f)に示すような、最終的な出力画像(図中ではエリアG2の画像)が得られる。この画像は、全周カメラ400に写ったマーカ光の周辺領域だけを切り取った上で、湾曲した魚眼レンズの円形視野角を通常の矩形視野角へと変換補正したものである。言い換えれば、全周の中心に対して劣弧をとる扇形が切り出されると共に、その切り出された扇形の劣弧を直線に補正することで、扇形が四角形に補正されたものである。このような補正であることから、結果的にマイク300を持っている歌唱者(この例では利用者A)に対して通常のカメラを向けたのと同等の結果が得られる。本実施形態では、歌唱者(この例では利用者A)が歌唱している間の映像に対し上記変換補正が行われることで、歌唱者である利用者Aを含む連続的な映像が取得され、記憶される。

30

【0056】

なお、これらの処理は、動画を構成する各画像(静止画)に対して行われるので、毎秒30フレームの速度で処理されるが、演算能力の関係上、例えば10フレームにつき1フレームの頻度で処理を行うなど、間引きを行ってもよい。

【0057】

本実施形態では、上記のようにして歌唱者(この例では利用者A)の位置を特定した後、その歌唱者に対し、他の利用者(この例では利用者B、C)が視線を向けているかどうかで、これら利用者A、B、Cからなるグループの盛り上がりタイミングを検出する。

【0058】

すなわち、上述においては、エリアG2の画像についての処理を例にとって説明したが、それ以外の27個のエリア、すなわち、エリアA1~A3、B1~B3、C1~C4、D1~D4、E1~E3、F1~F3、G1、G3、G4、H1~H4についても、同様の手法で各エリアに対して通常のカメラを向けた場合と同等の映像を得ることができる。これにより、歌唱者である利用者Aのまわりの所定範囲(例えば通常は人の存在が考えにくいカラオケルームKRの四隅に相当するエリアC4、D4、G4、H4を除く範囲)の映像を得ることができ、当該所定範囲に他の利用者(視聴者)がいるかどうかを、公知の顔認識技術等により検出することができる。この例で言えば、利用者B及び利用者Cの存在が、上記の顔認識技術等を用いて認識される。図5は、このようにして歌唱者(利用者A)まわりの所定範囲の画像が取得された状態を概念的に表している。

40

【0059】

50

そして、さらに、本実施形態では、図5のように歌唱者（利用者A）まわりの所定範囲について取得された画像を用いて、歌唱者以外の利用者（この例では利用者B，C）が歌唱者（利用者A）に対し視線を向けているかどうかを検出する。図5の例は、破線矢印で示すように、2名の利用者B，C全員の顔が利用者Aに向き、利用者Aに対し視線を向けている状態を表しており、現実の空間における上記図1に示した状態に対応している。

【0060】

一方、例えば、利用者B，Cのうち利用者Cの顔が歌唱者である利用者Aを向いておらず、利用者Aに対し視線を向けているのは利用者Bの1名のみである状態もありうる。この場合も、上記の顔技術認識等を用いて図6に示すような歌唱者（利用者A）まわりの所定範囲の画像が取得されることで、上記のような状態であることが認識される。現実の空間では図7に示されるような状態となる。

10

【0061】

以上のようにして、本実施形態では、利用者Aが歌唱者として歌唱しているとき、所定周期（例えば数十msec等）ごとの各タイミングにおいて利用者Aの方に顔が向き視線を向けている他の利用者の人数を検出する。その検出した人数は、時系列に沿ったタイミングログに記録される。そして、歌唱者へ顔を向けている他の利用者（視聴者）が最も多いタイミングを、この集団の盛り上がりタイミングである、と決定する。

【0062】

図8は、上記盛り上がりタイミングを決定するために用いられる、上記タイミングログの例を表す説明図である。図8に示すように、タイミングログには、各データが取得された時刻（言い換えれば録画時刻）を例えばmsec単位で表す「時刻」欄と、前述のようにしてマーカ光に基づき識別された歌唱者を表す「歌唱者」欄と、上記顔認識により識別された、歌唱者以外の在室利用者すなわち視聴者を表す「視聴者」欄と、その視聴者のうち歌唱者の方を顔が向いている視聴者を表す「歌唱者の方を向いている視聴者」の欄とが、記録欄としてそれぞれ設けられている。図示のように、この例では、左から右に向かって時系列的に各データが記録されている。

20

【0063】

例えば利用者Aが歌唱している間は、「歌唱者」欄には当該時刻範囲の全タイミングにおいて「A」が記録される。利用者Aが歌唱しているカラオケルームKR内に利用者B，Cの両方が在室しているタイミングでは「視聴者」欄に「B，C」が記録される。例えば利用者Bがトイレに行くために退室し視聴者として利用者Cのみが在室しているタイミングでは「C」のみが記録される。

30

【0064】

また、利用者Aが歌唱しているときに利用者B，Cの両方が利用者Aの方を向いているタイミングでは「歌唱者の方を向いている視聴者」欄に「B，C」が記録される（図5及び図1の状態に相当）。一方、利用者Aが歌唱しているときに利用者Cは利用者Aの方を向いているが、利用者Bは例えば壁の方を向いており利用者Aの方を向いていないタイミングでは「歌唱者の方を向いている視聴者」欄に「C」のみが記録され（図6及び図7の状態に相当）、誰も利用者Aの方を向いていないタイミングでは「なし」と記録される。

【0065】

図8に示す例では、図中左右方向の中央に位置するタイミングで、歌唱者である利用者Aの方を、2人の利用者B，Cが向いている。したがって、このタイミングが、「盛り上がりタイミング」である、として決定される。

40

【0066】

そして、先に述べたように、本実施形態では、歌唱者が歌唱している間の当該歌唱者（上記の例では利用者A）を含む映像が連続的に取得され、記憶されている。このように記憶された映像すなわち動画は多数の静止画の集合体であるが、本実施形態では、上記盛り上がりタイミングとして決定されたタイミングの静止画が、上記歌唱者を撮影した映像を代表する、サムネイル画像として抽出される。

【0067】

50

図9は、上記の手法を実行するために、制御部101により実行される処理手順の詳細を表すフローチャートである。

【0068】

図9において、カラオケ演奏曲に対応したMIDIデータの再生が開始されると、このフローが開始される。すなわち、音源106等によるMIDIデータの再生と同期して、背景映像データ及び歌詞データが表示部109に表示される。すると、歌唱者によるカラオケ演奏曲の歌唱が行われ、マイク300よりカラオケ歌唱の音声が入力される。

【0069】

まずステップS10において、全周カメラ400により撮影された、カラオケルームKR内の映像データを取得する。この手順が、各請求項記載の取得手順を構成する。

10

【0070】

その後、ステップS25で、カラオケルームKR内の映像データに含まれるマーカ光に基づいて、マイク300の位置を特定する。このステップS25の処理は、前述の図4(a)～図4(c)に示した画像処理に対応するものである。そして、ステップS30において、カラオケルームKR内の映像データからマイク300の位置を含む部分映像データを切り出す。このステップS30の処理は、図4(d)に示したエリア判定処理に対応するものである。

【0071】

その後、ステップS35において、マイク300の位置を含む部分映像データの補正処理を行い、マイク300を持った歌唱者の姿が写った画像を得る。このステップS35の処理は、図4(e)、図4(f)に示した画像の形状矯正に対応する。そして、ステップS40において、補正処理後の画像を表示部109の一部領域に表示させるとともに、撮影時刻と関連づけて大容量記憶装置103に保存する。

20

【0072】

その後、ステップS100において、上記図8を用いて説明したタイミング6を作成するログ作成処理が実行される。図10は、このステップS100の詳細手順を表すフローチャートである。

【0073】

図10において、まずステップS110で、歌唱者の位置から所定範囲(前述の例ではカラオケルームKRの四隅を除く範囲)内の映像データを補正する。なお、この補正処理は、上記ステップS35での補正処理と同等のものを実行すれば足りるので、詳細な説明を省略する。

30

【0074】

その後、ステップS120において、上記ステップS110で補正された所定範囲の映像データに対し、公知の顔認識処理を実行し、視聴者の姿を検出する。

【0075】

そして、ステップS130に移り、上記ステップS120において検出した視聴者について、各視聴者の顔の向きを検出する。この検出には、上記同様、公知の適宜の顔認識処理やその他の画像解析処理により、各視聴者の顔の輪郭線や鼻・口の位置を特定し、顔の向きがカラオケルームKR内のいずれの方向を向いているかを算出すればよい。

40

【0076】

その後、ステップS140に移り、上記ステップS130での検出結果に基づき、歌唱者の方を向いている視聴者人数をカウントする。そして、ステップS120で検出されたカラオケルームKR内に在室する視聴者の数と、上記カウントされた歌唱者の方を向いている視聴者の人数と、撮影時刻とを上記タイミングログにデータとして記録する。なお、この作成されたタイミングログは、例えば上記大容量記憶装置103内に、参照可能に蓄積され格納される。なお、これらステップS120、ステップS130、及びステップS140が各請求項記載の算出手順を構成すると共に、算出手段として機能する。その後、ステップS45(図9参照)に移る。

【0077】

50

図9に戻り、ステップS45では、カラオケ演奏曲に対応したMIDIデータの再生が終了したかどうかを判定する。カラオケ演奏曲に対応したMIDIデータの再生が終了したときは、ステップS45の判定が満たされてステップS50に移る。一方、カラオケ演奏曲に対応したMIDIデータの再生が終了していないときは、ステップS45の判定が満たされず、ステップS10に戻り、同様の手順を繰り返す。これにより、カラオケ演奏曲の再生が終了しない間は、ステップS10～ステップS40及びステップS100が繰り返され、ステップS40を経るたびにステップS40において大容量記憶装置103に補正処理後の映像(動画)が順次保存されていき、また、動画を構成する各画像(静止画)に対しステップS100のログ作成処理が実行される。すなわち、各画像(静止画)ごとにステップS10～ステップS40及びステップS100が繰り返されるのである。

10

【0078】

ステップS50では、動画投稿指示操作画面を表示部109の一部領域に表示させる。なお、この動画投稿指示操作画面には、操作者(歌唱者である利用者A。但し、利用者B、Cが操作してもよい)のID、すなわち歌唱者IDの入力を促す表示が含まれている。その後、ステップS55に移る。

【0079】

ステップS55では、表示部109の動画投稿指示操作画面によって、操作者よりカラオケ動画の投稿が指示されたかどうかを判定する。すなわち、上記歌唱者IDの入力を促す表示に対応して(例えば装置本体100の操作部104又はリモコン200の操作部204により)歌唱者IDが入力されると共に、適宜の投稿指示ボタン等の操作がなされたかどうか判定される。なお、上記操作部104又は操作部204が各請求項記載の歌唱者ID入力手段として機能する。カラオケ動画の投稿が指示されたときは、ステップS55の判定が満たされてステップS200に移り、カラオケ動画の投稿が指示されないときは、ステップS55の判定が満たされず、このフローを終了する。

20

【0080】

ステップS200では、上記ステップS100のログ作成処理により作成されたタイミングログ(図8参照)の「視聴者の方を向いている視聴者」の欄を参照し、各タイミングにおける歌唱者の方を向いている視聴者の人数を取得する。そして、当該人数が最大となっているタイミングを、盛り上がりタイミングとして決定する。このステップS200が、各請求項記載の決定手順を構成すると共に、決定手段として機能する。

30

【0081】

その後、ステップS210に移り、上記ステップS40において撮影時刻と関連づけて大容量記憶装置103に記憶されていた歌唱者の姿を含む複数の画像データの中から、上記ステップS200で決定した盛り上がりタイミングに対応した画像(静止画)を取得し、その画像をサムネイル画像(代表画像)とする。このステップS210と前述のステップS30及びステップS35とが、各請求項記載の映像処理手段として機能する。その後、ステップS60に移る。

【0082】

ステップS60では、前述のステップS10～ステップS40及びステップS100の繰り返し時にステップS40で大容量記憶装置103に順次保存された補正処理後の複数の出力映像を用いた動画データと、ステップS210で取得されたサムネイル画像と、ステップS55で入力された歌唱者IDとを、互いに関連づけた態様でホストサーバ20にアップロードする。なお、このステップS60が、各請求項記載の出力手順を構成するとともに、静止画出力手段として機能する。ステップS60が完了すると、このフローを終了する。

40

【0083】

ホストサーバ20にアップロードされた上記サムネイルの、前述の所定のwebページでの表示例を図11に示す。図11に示す表示例では、歌唱者(例えば利用者A～Cのいずれか。上記の例では利用者A)が歌唱しているサムネイル画像が、カラオケ演奏曲の曲名、歌唱日時(上記盛り上がりタイミングの日時を含む)、カラオケ動画の再生回数、評

50

価等とともに表示されている。webページの画面上で例えばこのサムネイル画像（又は対応する操作部等）をクリックすることにより、上記ステップS60でアップロードされた歌唱者による歌唱時の動画をすべて再生し、閲覧することができる。この例では、このようにして行った閲覧の後の、当該カラオケ動画に対する閲覧した利用者による評価（「うまい」「おもしろい」「かわいい」「泣ける」が併せて記入され、表示される（詳細な図示は省略）。

【0084】

以上説明したように、本実施形態においては、娯楽性の向上のために、歌唱者（前述の例では利用者A）及び視聴者（前述の例では利用者B、C）の集団の盛り上がりを検出する。前述したように、集団が盛り上がっているときは、歌唱者の歌唱によって視聴者が心より楽しんでいるときであり、その瞬間には、視聴者の視線が歌唱者のほうへ向いているのが通常である。そこで、全周カメラ400の撮影結果に基づき生成されたカラオケルームKR内の映像データに対し顔認識処理を行い、その認識結果に基づき、各タイミングにおける「歌唱者の方を向いている視聴者」の人数をタイミングログとして記録する（図8、図9のステップS100参照）。そして、当該人数の時間的推移により、盛り上がりタイミングを決定する（ステップS200参照）。これにより、顔が歌唱者へ向いている視聴者の人数が最も多い瞬間を、当該集団の盛り上がりタイミングと決定することができる。この結果、全周カメラ400での撮影により生成された映像データのうち、当該盛り上がりタイミングにおける静止画を抽出し（ステップS210参照）、当該サービス提供時間の代表画像としてホストサーバ20へアップロードすることができる（ステップS60）。この結果、ホストサーバ20へアクセスした各ユーザ等が、当該静止画を閲覧し、楽しむことができる（図11参照）。なおこのサムネイルは、カラオケ装置10の表示部109に映し出すこともできる。このような種々のサービスを行うことにより、カラオケ装置10の娯楽性をさらに向上することができる。

【0085】

また、本実施形態では特に、1台の全周カメラ400が、歌唱者及びマイクロ300を含む所定範囲の視野を撮影し、当該所定範囲の映像データを生成する。このとき、カメラを中心とした全周を撮影可能な視野が広い全周カメラ400を用いることにより、歌唱者（上記の例では利用者A）及び視聴者（上記の利用者B、C）の全員が、当該カメラが生成した映像データの中に常に含まれる。そして生成された所定範囲の映像データには、歌唱者の所持したマイク300の位置に対応したマーカ光が、歌唱者の姿と共に必ず記録されている。そこで、映像データに含まれるマーカ光を用いてマイク300及び歌唱者の位置を特定する（図9のステップS25参照）とともに、歌唱者の位置以外の映像データに対し顔認識処理を行って視聴者の顔の向きを決定する（図10のステップS120、ステップS130参照）。これにより、複数台のカメラを用いなくても、全周カメラ400の1台だけで、顔が歌唱者の方向を向いている視聴者の人数を確実に算出することができる。

【0086】

ここで、上記静止画を生成する場合、上記のように視野が広い全周カメラ400の映像データから得た静止画は視野の端部ほど歪んだ状態となっている。本実施形態では特に、図9のステップS35において、上記所定範囲の映像データから、歌唱者を含む部分静止画を全周の中心に対して劣弧をとる扇形に切り出すと共に、その切り出した部分静止画に所定の補正処理を行う（前述したように、切り出した扇形の劣弧を直線に補正し、扇形を四角形とする）。これにより、上記歪んだ状態が是正された正常な静止画からなる盛り上がりタイミングのサムネイルを、ステップS60において代表画像としてホストサーバ20へ出力することができる。

【0087】

なお、本発明は、上記実施形態に限られるものではなく、その趣旨及び技術的思想を逸脱しない範囲内で種々の変形が可能である。以下、そのような変形例を説明する。

【0088】

(1) 歌唱者の顔が写っている静止画を抽出する場合

例えば、歌唱者及び視聴者の集団が盛り上がっている盛り上がりタイミングにおいて、たまたま歌唱者が下を向いたり横を向いたりしている場合がある。このようなタイミングにおける静止画は、歌唱者の顔が正しく記録されていないため、代表画像としては必ずしも好ましくない。本変形例は、このような場合に対応するためのものである。

【 0 0 8 9 】

図 1 2 は、この変形例において制御部 1 0 1 により実行される処理手順の詳細を表すフローチャートであり、上記図 9 に対応する図である。図 9 と同等の手順には同一の符号を付し、説明を省略又は簡略化する。

【 0 0 9 0 】

図 1 2 において、本変形例では、図 9 におけるステップ S 2 1 0 を省略すると共に、新たに、ステップ S 2 2 0、ステップ S 2 3 0、ステップ S 2 4 0、ステップ S 2 5 0、ステップ S 2 6 0 を設けている。

【 0 0 9 1 】

ステップ S 1 0 ~ ステップ S 4 0 及びステップ S 2 0 0 の処理は、図 9 と同様であり、詳細な説明を省略する。ステップ S 2 0 0 が終了したら、新たに設けたステップ S 2 2 0 に移る。

【 0 0 9 2 】

ステップ S 2 2 0 では、上記ステップ S 2 0 0 での盛り上がりタイミングの決定結果に基づき、盛り上がりタイミングが複数あったかどうかを判定する。すなわち、上記タイミングログにおいて、複数のタイミングにおいて、歌唱者の方を向いている視聴者の数が同人数であったかどうか、が判定される。

【 0 0 9 3 】

盛り上がりタイミングが 1 つに限られ、複数はなかった場合は、ステップ S 2 2 0 の判定が満たされず、ステップ S 2 5 0 に移る。ステップ S 2 5 0 では、上記ステップ S 4 0 において撮影時刻と関連づけて大容量記憶装置 1 0 3 に記憶されていた歌唱者の姿を含む多数の画像データの中から、上記 1 つの盛り上がりタイミングに対応した画像（静止画）を取得し、その画像をサムネイル画像とする。その後、ステップ 6 0 に移る。

【 0 0 9 4 】

一方、ステップ S 2 2 0 において、盛り上がりタイミングが複数あった場合は、ステップ S 2 2 0 の判定が満たされ、ステップ S 2 3 0 に移る。ステップ S 2 3 0 では上記ステップ S 4 0 において撮影時刻と関連づけて大容量記憶装置 1 0 3 に記憶されていた歌唱者の姿を含む多数の画像データの中から、上記複数の盛り上がりタイミングに対応した画像（静止画）を取得する。そして、それら複数の静止画の中に対し、公知の顔認識の手法を用いて、歌唱者の顔が認識できるものがあるかを判定する。なお、このステップ S 2 2 0 が、各請求項記載の判定手段として機能する。

【 0 0 9 5 】

歌唱者の顔が認識できるものがある場合は、ステップ S 2 3 0 の判定が満たされ、ステップ S 2 4 0 に移り、当該認識できる静止画をサムネイル画像とする。なお、歌唱者の顔が認識できる静止画が複数あった場合には、その中の適宜のいずれか 1 つをサムネイルとすればよい。その後、ステップ 6 0 に移る。

【 0 0 9 6 】

一方、ステップ S 2 3 0 において、歌唱者の顔が認識できる静止画が無い場合は、ステップ S 2 6 0 に移る。ステップ S 2 6 0 では、上記ステップ S 2 3 0 において取得された、複数の静止画の中の適宜のいずれか 1 つをサムネイルとする。その後、ステップ S 6 0 に移る。なお、歌唱者の顔が認識できなかったことに対応して、このステップ S 2 6 0 ではサムネイルを設定せず、表示部 1 0 9 に表示信号を出力して適宜のエラー表示を行い、このフローを終了するようにしてもよい。また盛り上がりタイミングが 1 つしかなくステップ S 2 2 0 での判定が満たされなかった場合にも、ステップ S 2 3 0 と同様の判定を行い、判定が満たされなかった場合には、上記同様にサムネイルを設定せずエラー表示とし

10

20

30

40

50

てもよい。

【0097】

なお、ステップS60の処理は、図9と同様であるので、詳細な説明を省略する。また、ステップS250、ステップS240、ステップS260と前述のステップS30及びステップS35とが、各請求項記載の映像処理手段として機能する。

【0098】

本変形例においては、サムネイルを設定してホストサーバ20へ出力する際、なるべく歌唱者の顔が正しく記録された静止画をサムネイルにすることができる。

【0099】

(2) 音声レベルを考慮して抽出する場合

例えば、歌唱者及び視聴者の集団が盛り上がっている盛り上がりタイミングにおいて、たまたま歌唱者が歌っていない場合がある。すなわち、歌唱者の歌唱ではないジェスチャーや仕草、表情等により視聴者が一斉に着目した場合等、歌唱者による歌唱以外の行動により盛り上がりが生じた場合である。このようなタイミングにおける静止画は、歌唱を行っていないため、カラオケ歌唱動画の代表画像としては必ずしも好ましくない。本変形例は、このような場合に対応するためのものである。

【0100】

図13は、この変形例において制御部101により実行される処理手順の詳細を表すフローチャートであり、上記図9や図12に対応する図である。図12と同等の手順には同一の符号を付し、説明を省略又は簡略化する。

【0101】

図12において、本変形例では、図9におけるステップS230及びステップS240に代え、ステップS300、ステップS310を設けている。

【0102】

すなわち、図13に示すように、ステップS220の判定が満たされたら、ステップS300に移る。ステップS300では、上記ステップS40において撮影時刻と関連づけて大容量記憶装置103に記憶されていた歌唱者の姿を含む多数の画像データの中から、上記複数の盛り上がりタイミングに対応した画像(静止画)を取得する。そして、それら複数の静止画の中に対し、当該タイミングにおけるマイク300からの音声レベルが所定値未満のものがあるかを判定する。すなわち、この変形例では、上記大容量記憶装置103に記憶された各画像データは、その時点でのマイク300の音声レベルとも予め対応付けられた形で、記憶されている。

【0103】

音声レベルが所定値未満のものがあった場合には、ステップS300の判定が満たされ、ステップS310に移り、音声レベルが所定値以上となっているときの静止画をサムネイル画像とする。なお、音声レベルが所定値以上となっている静止画が複数あった場合には、その中の適宜のいずれか1つをサムネイルとすればよい。その後、ステップ60に移る。

【0104】

一方、ステップS300において、音声レベルが所定値未満であるものがない場合は、ステップS260に移り、上記同様、上記ステップS230において取得された、複数の静止画の中の適宜のいずれか1つをサムネイルとする。その後、ステップS60に移る。

【0105】

なお、ステップS60の処理は、図9と同様であるので、詳細な説明を省略する。また、ステップS250、ステップS310、ステップS260と前述のステップS30及びステップS35とが、各請求項記載の映像処理手段として機能する。

【0106】

本変形例においては、サムネイルを設定してホストサーバ20へ出力する際、なるべく歌唱者が実際に歌唱している様子が記録された静止画をサムネイルにすることができる。

【0107】

10

20

30

40

50

(3) その他

以上においては、1台の全周カメラ400を用いて撮影を行った場合を例にとって説明したが、これに限られない。すなわち、カメラを複数台使い、歌唱者撮影カメラで歌唱者を撮影すると共に、視聴者を撮影するために配置した別の視聴者撮影カメラにて室内の視聴者を撮影するようにしても良い。いずれにしても、歌唱者の方向へ向かっている視聴者の人数を検出できれば足りる。

【0108】

なお、図9、図10、図12、図13等に示すフローチャートは本発明を上記フローに示す手順に限定するものではなく、発明の趣旨及び技術的思想を逸脱しない範囲内で手順の追加・削除又は順番の変更等をしてよい。

10

【0109】

また、以上既に述べた以外にも、上記実施形態や各変形例による手法を適宜組み合わせ利用しても良い。

【0110】

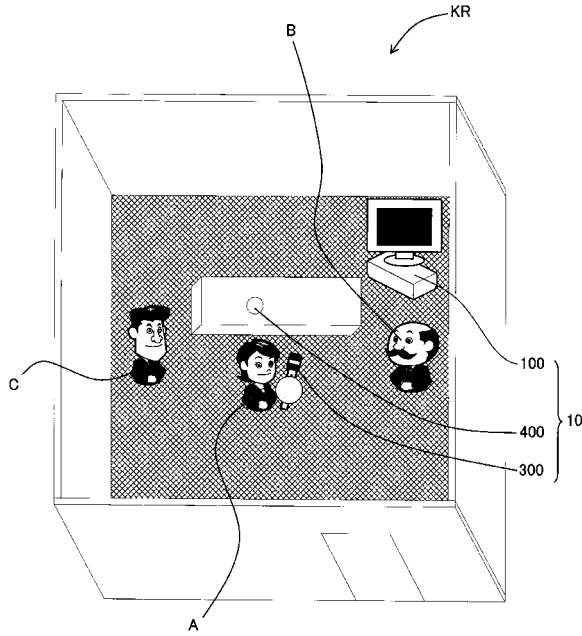
その他、一々例示はしないが、本発明は、その趣旨を逸脱しない範囲内において、種々の変更が加えられて実施されるものである。

【符号の説明】

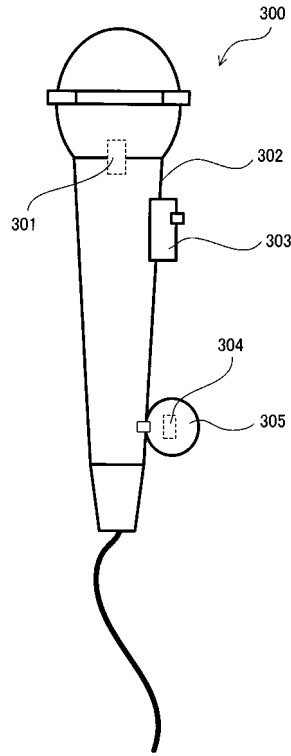
【0111】

10	カラオケ装置	
106	音源(楽曲再生手段)	20
107	音声制御部(楽曲再生手段)	
108	スピーカ(楽曲再生手段)	
109	表示部(表示手段)	
200	リモコン	
300	マイク(マイクロフォン)	
304	LED(標識信号発生手段)	
400	全周カメラ(動画撮影カメラ)	
A	利用者(歌唱者)	
B	利用者(視聴者)	
C	利用者(視聴者)	30

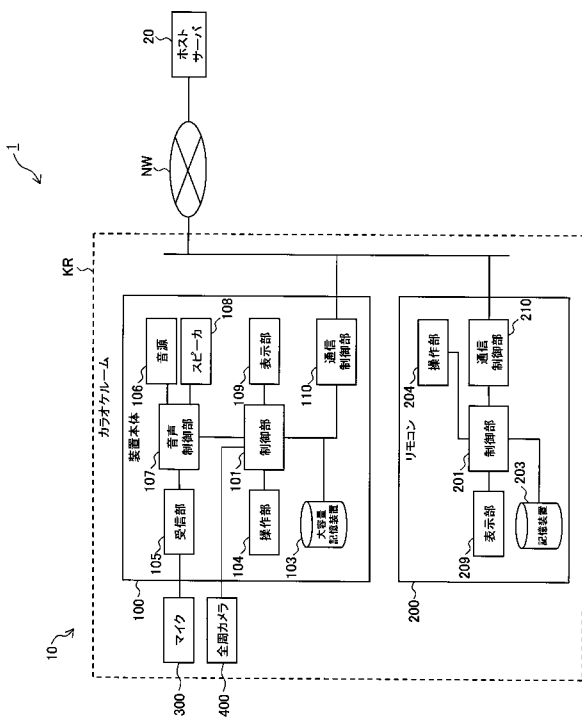
【図1】



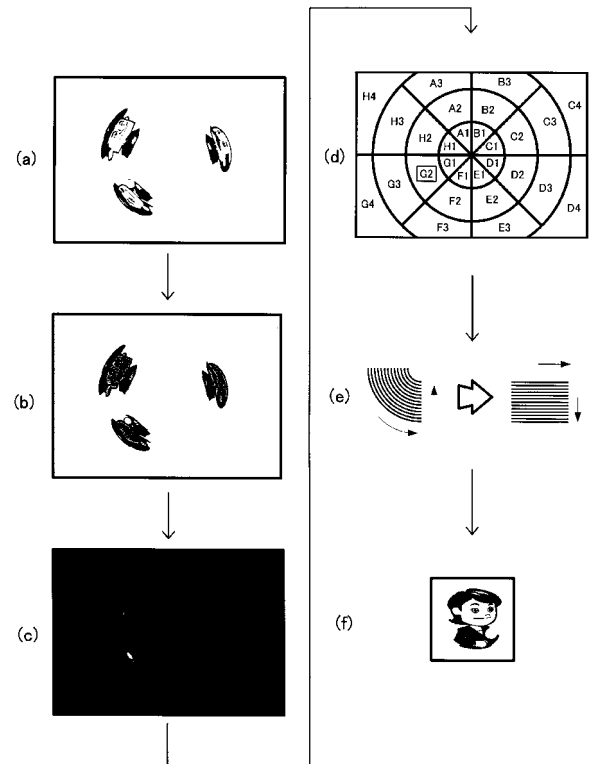
【図2】



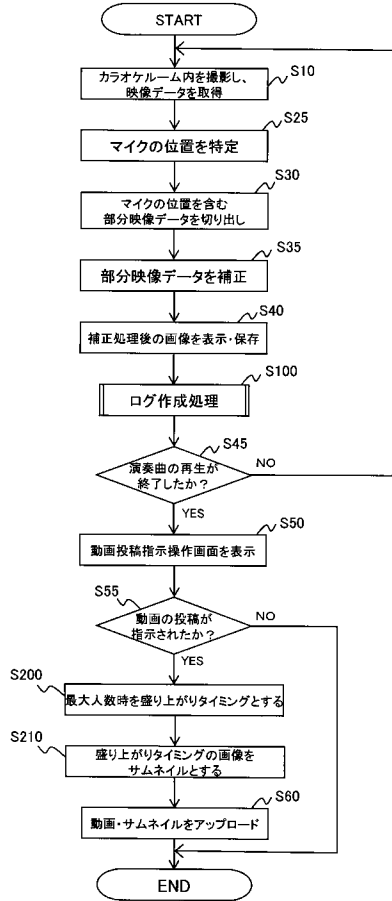
【図3】



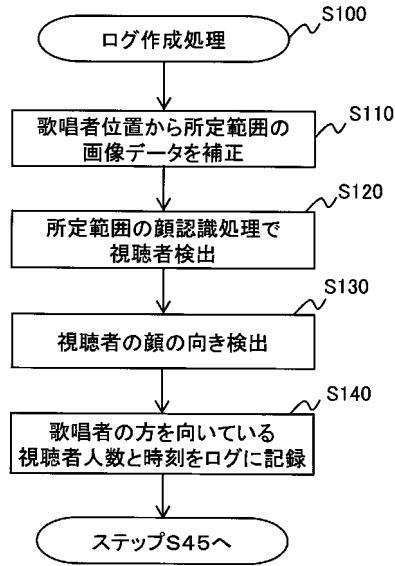
【図4】



【図9】



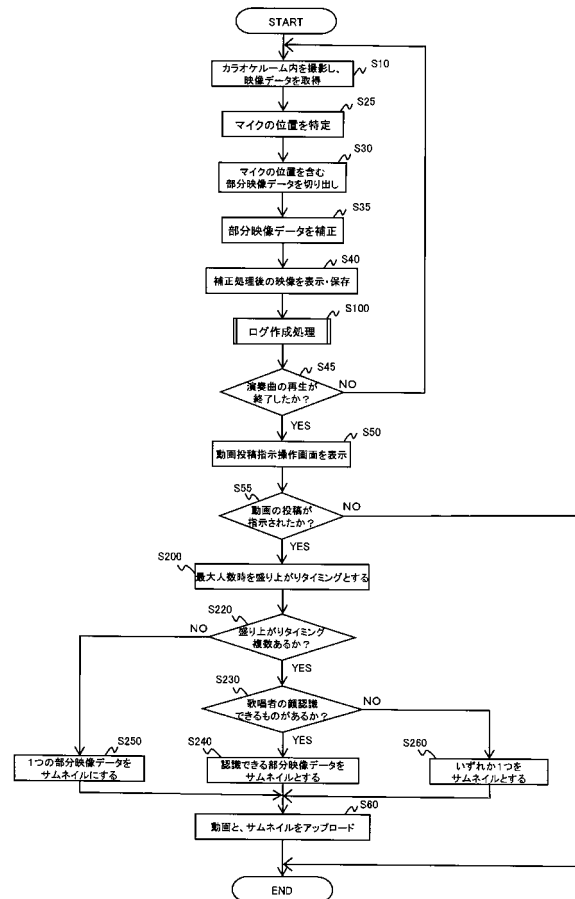
【図10】



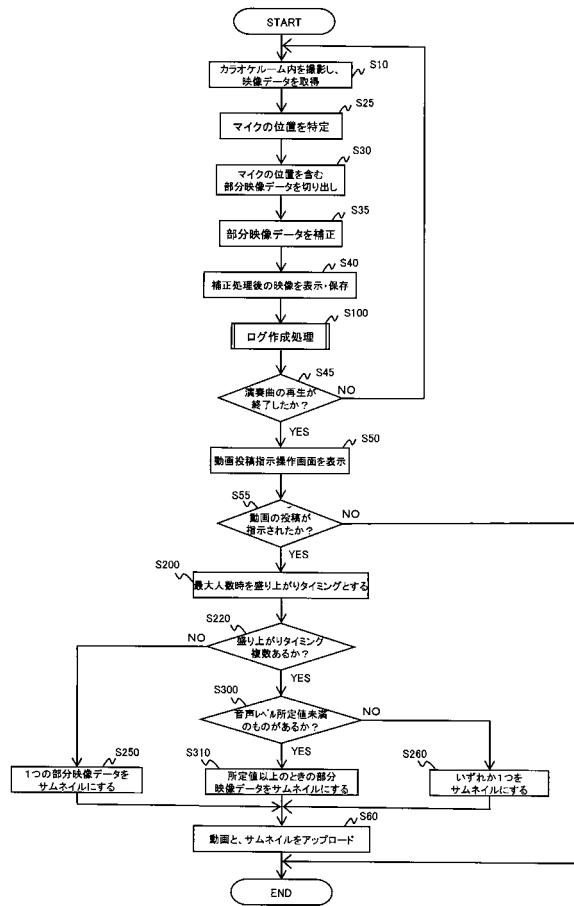
【図11】



【図12】



【図13】



フロントページの続き

- (56)参考文献 特開2006-081128(JP,A)
特開2008-061260(JP,A)
特開2002-330390(JP,A)
特開平10-207477(JP,A)
特開2007-265125(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10K	15/04
G06T	1/00
H04N	5/225
H04N	5/232