

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7501575号  
(P7501575)

(45)発行日 令和6年6月18日(2024.6.18)

(24)登録日 令和6年6月10日(2024.6.10)

(51)国際特許分類 F I  
H 0 4 M 3/56 (2006.01) H 0 4 M 3/56 C  
H 0 4 N 7/15 (2006.01) H 0 4 N 7/15

請求項の数 11 (全19頁)

(21)出願番号	特願2022-115486(P2022-115486)	(73)特許権者	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22)出願日	令和4年7月20日(2022.7.20)	(74)代理人	100110928 弁理士 速水 進治
(62)分割の表示	特願2018-65248(P2018-65248)の 分割	(72)発明者	合川 直志 東京都港区芝五丁目7番1号 日本電気 株式会社内
原出願日	平成30年3月29日(2018.3.29)	(72)発明者	木村 智 東京都港区芝五丁目7番1号 日本電気 株式会社内
(65)公開番号	特開2022-140529(P2022-140529 A)	(72)発明者	佐藤 伸正 東京都港区芝五丁目7番1号 日本電気 株式会社内
(43)公開日	令和4年9月26日(2022.9.26)	審査官	鈴木 順三
審査請求日	令和4年7月20日(2022.7.20)		

最終頁に続く

(54)【発明の名称】 ビデオ会議システム、ビデオ会議方法、およびプログラム

(57)【特許請求の範囲】

【請求項1】

会議の参加人物が写る画像を取得する画像取得手段と、  
前記画像の中から、人物と認識される領域を検出する人物領域検出手段と、  
前記領域に含まれる人物を特定する人物特定処理を実行する人物特定手段と、  
前記人物特定処理で前記人物が特定できなかったことを示す第1情報と、前記人物特定  
処理で特定された前記会議の参加人物の氏名を含む第2情報と、を表示装置に表示させる  
表示制御手段と、

を備え、

前記表示制御手段は、前記第1情報を前記第2情報よりも目立たせ、

前記表示制御手段は、前記人物特定処理により特定された前記会議の参加人物の数が所  
定の閾値以上である場合、前記第2情報を一覧形式で表示させるビデオ会議システム。

【請求項2】

前記表示制御手段は、前記第1情報の外形、大きさ、色、言語、およびフォントの少な  
くとも1つを前記第2情報と異ならせることにより、前記第1情報を目立たせる、  
請求項1に記載のビデオ会議システム。

【請求項3】

前記表示制御手段は、特定された前記領域以外の領域に前記第2情報を表示させる、  
請求項1または2に記載のビデオ会議システム。

【請求項4】

前記表示制御手段は、前記画像中の空き領域のサイズに応じて前記第 2 情報の大きさを  
変更する、

請求項 1 から 3 いずれか 1 項に記載のビデオ会議システム。

【請求項 5】

前記表示制御手段は、前記画像に写っている前記会議の参加人物が見る表示装置に、前  
記第 1 情報を表示させる、

請求項 1 から 4 のいずれか 1 項に記載のビデオ会議システム。

【請求項 6】

前記画像取得手段は、前記画像に基づく前記人物特定処理で特定されなかった参加人物  
が写る追加画像を、前記画像を生成した撮像装置とは異なる移動型の撮像装置から取得し、

前記人物特定手段は、前記追加画像に基づいて前記画像に基づく前記人物特定処理で特  
定されなかった参加人物を特定する、

請求項 1 から 5 のいずれか 1 項に記載のビデオ会議システム。

【請求項 7】

前記人物特定処理によって特定された人物のリストを作成するリスト作成手段を更に備  
える、

請求項 6 に記載のビデオ会議システム。

【請求項 8】

音声データを取得する音声取得手段と、

前記音声データまたは前記音声データと同期して取得された画像を解析することによっ  
て、前記音声データに関する発言者を特定する発言者特定手段と、

前記発言者の特定結果と前記音声データに基づいて生成されたテキストデータとを対応  
付けることにより、議事録データを生成する議事録作成手段と、を更に備える、

請求項 6 または 7 に記載のビデオ会議システム。

【請求項 9】

会議の参加人物が写る画像を取得する画像取得手段と、

前記画像の中から、人物と認識される領域を検出する人物領域検出手段と、

前記領域に含まれる人物を特定する人物特定処理を実行する人物特定手段と、

前記人物特定処理で前記人物が特定できなかったことを示す第 1 情報と、前記人物特定  
処理で特定された前記会議の参加人物の氏名を含む第 2 情報と、を表示装置に表示させる  
表示制御手段と、

を備え、

前記表示制御手段は、前記第 1 情報を前記第 2 情報よりも目立たなくさせ、

前記表示制御手段は、前記人物特定処理により特定された前記会議の参加人物の数が所  
定の閾値以上である場合、前記第 2 情報を一覧形式で表示させるビデオ会議システム。

【請求項 10】

コンピュータが、

会議の参加人物が写る画像を取得し、

前記画像の中から、人物と認識される領域を検出し、

前記領域に含まれる人物を特定する人物特定処理を実行し、

前記人物特定処理で前記人物が特定できなかったことを示す第 1 情報と、前記人物特定  
処理で特定された前記会議の参加人物の氏名を含む第 2 情報と、を表示装置に表示させ、

前記第 1 情報を前記第 2 情報よりも目立たせ、

前記人物特定処理により特定された前記会議の参加人物の数が所定の閾値以上である場合  
、前記第 2 情報を一覧形式で表示させる、

ことを含むビデオ会議方法。

【請求項 11】

コンピュータに、請求項 10 に記載のビデオ会議方法を実行させるプログラム。

【発明の詳細な説明】

【技術分野】

10

20

30

40

50

## 【0001】

本発明は、ビデオ会議システム、ビデオ会議方法、およびプログラムに関する。

## 【背景技術】

## 【0002】

離れた場所にいる人物と会議を行う方法の1つとして、ビデオ会議システムが利用されている。ビデオ会議システムでは、それぞれの場所で撮影された映像を互いにやり取りすることにより、互いに離れた場所にいる人物同士で会議を行うことができる。

## 【0003】

上述のビデオ会議システムに関連する技術が、例えば、下記特許文献1乃至3に開示されている。

## 【0004】

下記特許文献1には、2以上の地点間で双方向コミュニケーションを行う際、双方向コミュニケーションの参加者を容易に把握するための技術が開示されている。具体的には、(1)データベースに記憶されている識別情報を用いて各地点の参加者を認証し、(2)認証された参加者の位置を検出し、(3)各地点で撮影された映像のうち、参加者の検出位置に対応する部分に、データベースに記憶されているその参加者の属性情報を視覚的に表示するデータを付加する技術が開示されている。

## 【0005】

また、下記特許文献2には、会議の参加者が初対面の相手であっても、その参加者に関する情報を知ることが可能とする技術が開示されている。具体的には、(1)会議の参加者に関する参加者情報の入力を受け付け、(2)各参加者に対して所定の動作(口頭での返事や挙手など)を行わせ、(3)その動作を行った人物を撮影画像上で特定し、(4)特定した人物の顔認識結果と参加者情報とを対応付け、(5)その対応付けに従って各参加者が写る撮影画像上に参加者情報を合成する技術が開示されている。

## 【0006】

また、下記特許文献3には、会議の参加者として認証された人物が写る範囲を、相手側に表示する映像の範囲として自動的に調整する技術が開示されている。

## 【先行技術文献】

## 【特許文献】

## 【0007】

【文献】特開2004-129071号公報

【文献】特開2010-028715号公報

【文献】特開2015-177418号公報

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0008】

ビデオ会議用のシステムを利用して開催される会議を含め、会議の場では、誰が参加しているかをはっきりさせるべきであり、個人として特定されていない人物がそのままにいることは好ましくない。

## 【0009】

本発明は、上記の課題に鑑みてなされたものである。本発明の目的の一つは、ビデオ会議システムにおいて、個人として特定されていない人物がそのまま会議の場に残ることを抑制する技術を提供することである。

## 【課題を解決するための手段】

## 【0010】

本発明のビデオ会議システムは、  
会議の参加人物が写る画像を取得する画像取得手段と、  
前記画像の中から、人物と認識される領域を検出する人物領域検出手段と、  
前記領域に含まれる人物を特定する人物特定処理を実行する人物特定手段と、  
前記人物特定処理で前記人物が特定できなかったことを示す第1情報を表示装置に表示

10

20

30

40

50

させる表示制御手段と、  
を備える。

【0011】

本発明のビデオ会議方法は、  
コンピュータが、  
会議の参加人物が写る画像を取得し、  
前記画像の中から、人物と認識される領域を検出し、  
前記領域に含まれる人物を特定する人物特定処理を実行し、  
前記人物特定処理で前記人物が特定できなかったことを示す第1情報を表示装置に表示  
させる、  
ことを含む。

10

【0012】

本発明のプログラムは、コンピュータに、上述のビデオ会議方法を実行させる。

【発明の効果】

【0013】

本発明によれば、ビデオ会議システムにおいて、個人として特定されていない人物がその  
のまま会議の場に残ることを抑制することができる。

【図面の簡単な説明】

【0014】

【図1】第1実施形態におけるビデオ会議システムの構成例を示す図である。

20

【図2】ビデオ会議システムのハードウェア構成を例示するブロック図である。

【図3】第1実施形態のビデオ会議システムにより実行される処理の流れを例示するフロー  
ーチャートである。

【図4】表示制御部により表示される第1情報の一例を示す図である。

【図5】人物領域検出部により表示される第2情報の一例を示す図である。

【図6】表示制御部による表示の一例を示す図である。

【図7】第2情報を一覧形式で表示する例を示す図である。

【図8】第3実施形態におけるビデオ会議システムの構成例を示す図である。

【図9】第3実施形態のビデオ会議システムにより実行される処理の流れを例示するフロー  
ーチャートである。

30

【図10】第4実施形態におけるビデオ会議システムの構成例を示す図である。

【図11】第4実施形態のビデオ会議システムにより実行される処理の流れを例示するフ  
ローチャートである。

【発明を実施するための形態】

【0015】

以下、本発明の実施形態について、図面を用いて説明する。尚、すべての図面において  
、同様な構成要素には同様の符号を付し、適宜説明を省略する。また、特に説明する場合  
を除き、各ブロック図において、各ブロックは、ハードウェア単位の構成ではなく、機能  
単位の構成を表している。

【0016】

40

[第1実施形態]

[システム構成例]

図1は、第1実施形態におけるビデオ会議システム1の構成例を示す図である。図1に  
例示されるビデオ会議システム1では、サーバ装置10が、通信端末20Aおよび通信端  
末20Bと接続されている。通信端末20Aおよび通信端末20Bは、それぞれ、互いに  
離れた場所に位置する地点Aおよび地点Bに設けられている端末である。ビデオ会議は、  
これらの端末を用いて行われる。通信端末20Aには、撮像装置30Aおよび表示装置4  
0Aが接続されている。また、通信端末20Bには、撮像装置30Bおよび表示装置40  
Bが接続されている。撮像装置30Aおよび撮像装置30Bは、それぞれの地点における  
会議の参加人物を撮影するために利用される。撮像装置30Aにより生成された地点Aの

50

参加人物の画像は、サーバ装置 10 を経由して、地点 B の表示装置 40 B に表示される。また、撮像装置 30 B により生成された地点 B の参加人物の画像は、サーバ装置 10 を経由して、地点 A の表示装置 40 A に表示される。なお、地点 A の参加人物の画像は、地点 A の参加人物が撮影された画像を確認できるように、地点 A の表示装置 40 A に表示されてもよい。同様に、地点 B の参加人物の画像は、地点 B の参加人物が撮影された画像を確認できるように、地点 B の表示装置 40 B に表示されてもよい。

**【0017】**

図 1 に示されるように、ビデオ会議システム 1 は、画像取得部 110、人物領域検出部 120、人物特定部 130、および表示制御部 140 を備える。図 1 の例において、これらの処理部は、1 台のサーバ装置 10 に備えられているが、ビデオ会議システム 1 の構成は図 1 の例に制限されない。図示されていないが、これらの処理部の全部または一部は、複数のサーバ装置に分散して或いは重複して設けられていてもよい。

10

**【0018】**

画像取得部 110 は、会議の参加人物が写る画像を取得する。図 1 の例では、画像取得部 110 は、会議の参加人物が写る画像を、ネットワークを介して接続された通信端末 20 A および通信端末 20 B から取得することができる。

**【0019】**

人物領域検出部 120 は、画像取得部 110 により取得された画像の中から、人物と認識される領域を検出する。人物領域検出部 120 は、既知の一般物体検出アルゴリズムを利用して、「人物」と認識（分類）される領域を検出することができる。また、人物領域検出部 120 は、例えば、動きのある物体の領域を、人物の領域として検出してもよい。人物領域検出部 120 は、「動きのある物体」を、例えば、時系列で並ぶ複数の画像間での特徴点の移動量に基づいて判断することができる。具体的には、人物領域検出部 120 は、時系列で並ぶ複数の画像間において、基準値以上移動している特徴点が含まれる物体の領域を、人物の領域として推定することができる。

20

**【0020】**

人物特定部 130 は、人物領域検出部 120 により「人物」と認識された領域に含まれる人物を特定する処理（人物特定処理）を実行する。言い換えると、人物特定部 130 は、画像取得部 110 により取得された画像に写る参加人物を個々に特定（認証）する。人物特定部 130 は、人物領域検出部 120 により検出された領域から抽出される特徴量と、予め登録された参加人物の特徴量とを照合した結果に基づいて、各領域に含まれる人物が誰であるかを特定することができる。なお、会議の参加人物の特徴量は、その参加人物の情報（氏名、所属など）と対応付けて、サーバ装置 10 のストレージデバイスなどに予め記憶されている。ここで、人物特定部 130 は、人物領域検出部 120 で検出された領域に含まれる人物が特定できなかった場合、その領域に含まれる人物が特定できなかったことを示す情報（特定失敗情報）をその領域に関連付ける。なお、「人物が特定できなかった場合」とは、例えば、照合の結果として算出されるスコアが基準値以上となる人物が存在しなかった場合などである。

30

**【0021】**

表示制御部 140 は、各地点に設けられた撮像装置 30 により生成された画像を、その他の地点に設けられた通信端末 20 に送信する。また、表示制御部 140 は、各地点に設けられた撮像装置 30 により生成された画像を、各々の撮像装置 30 が設けられている地点の通信端末 20 に送信してもよい。各地点の通信端末 20 は、受け取った画像を、当該通信端末 20 に接続された表示装置 40 に表示させる。図 1 の例では、表示制御部 140 は、通信端末 20 B を介して取得した地点 B の画像を通信端末 20 A に送信する。また、表示制御部 140 は、通信端末 20 A を介して取得した地点 A の画像を通信端末 20 B に送信する。また、表示制御部 140 は、人物特定部 130 の人物特定処理で特定できなかった人物が存在する場合、人物特定処理で人物が特定できなかったことを示す情報（第 1 情報）を少なくともいずれかの地点に設けられた表示装置に表示させる。なお、表示制御部 140 は、例えば、特定失敗情報が関連付けられた領域が存在するか否かに基づいて、

40

50

人物特定処理で特定できなかった人物が存在するか否かを判断することができる。表示制御部 140 により表示される第 1 情報は、ビデオ会議システム 1 が、ある人物の存在を認識しているが、その人物が誰かまでは特定できていないことを示す情報と言える。

【0022】

〔ハードウェア構成例〕

ビデオ会議システム 1 の各機能構成部は、各機能構成部を実現するハードウェア（例：ハードワイヤードされた電子回路など）で実現されてもよいし、ハードウェアとソフトウェアとの組み合わせ（例：電子回路とそれを制御するプログラムの組み合わせなど）で実現されてもよい。以下、ビデオ会議システム 1 の各機能構成部が、サーバ装置 10 においてハードウェアとソフトウェアとの組み合わせによって実現される場合について、さらに説明する。

10

【0023】

図 2 は、ビデオ会議システム 1 のハードウェア構成を例示するブロック図である。図 2 の例において、サーバ装置 10 は、バス 1010、プロセッサ 1020、メモリ 1030、ストレージデバイス 1040、入出力インタフェース 1050、及びネットワークインタフェース 1060 を有する。

【0024】

バス 1010 は、プロセッサ 1020、メモリ 1030、ストレージデバイス 1040、入出力インタフェース 1050、及びネットワークインタフェース 1060 が、相互にデータを送受信するためのデータ伝送路である。ただし、プロセッサ 1020 などを互いに接続する方法は、バス接続に限定されない。

20

【0025】

プロセッサ 1020 は、CPU (Central Processing Unit) や GPU (Graphics Processing Unit) などを実現されるプロセッサである。

【0026】

メモリ 1030 は、RAM (Random Access Memory) などを実現される主記憶装置である。

【0027】

ストレージデバイス 1040 は、HDD (Hard Disk Drive)、SSD (Solid State Drive)、メモリカード、又は ROM (Read Only Memory) などを実現される補助記憶装置である。ストレージデバイス 1040 はビデオ会議システム 1 の各機能（画像取得部 110、人物領域検出部 120、人物特定部 130、および表示制御部 140 など）を実現するプログラムモジュールを記憶している。プロセッサ 1020 がこれら各プログラムモジュールをメモリ 1030 上に読み込んで実行することで、そのプログラムモジュールに対応する各機能が実現される。

30

【0028】

入出力インタフェース 1050 は、サーバ装置 10 と各種入出力デバイスとを接続するためのインタフェースである。入出力インタフェース 1050 には、キーボードやマウスといった入力装置（図示せず）、または、ディスプレイやスピーカーといった出力装置（図示せず）などが接続され得る。

40

【0029】

ネットワークインタフェース 1060 は、サーバ装置 10 をネットワークに接続するためのインタフェースである。このネットワークは、例えば LAN (Local Area Network) や WAN (Wide Area Network) である。ネットワークインタフェース 1060 がネットワークに接続する方法は、無線接続であってもよいし、有線接続であってもよい。図示されるように、サーバ装置 10 は、ネットワークインタフェース 1060 を介して、通信端末 20A および通信端末 20B と通信可能に接続されている。各通信端末 20 には、会議の参加人物を撮影するための撮像装置 30、各撮像装置 30 により生成された画像を表示させるための表示装置 40、および、会議中の音声を拾うための集音装置 50 が接続されている。また、通信端末 20 には、会議の音声を出力するための音声出力装置（図示

50

せず)が更に接続されている。

#### 【0030】

画像取得部110は、ネットワークインタフェース1060を介して各通信端末20から会議の参加人物が写る画像を取得することができる。また、表示制御部140は、ネットワークインタフェース1060を介して、各通信端末20に相手の参加人物の画像を送信することができる。また、表示制御部140は、ネットワークインタフェース1060を介して、各通信端末20にその通信端末20が備えられている地点の参加人物の画像を送信することができる。

#### 【0031】

〔処理の流れ〕

図3を用いて、第1実施形態のビデオ会議システム1により実行される処理の流れについて説明する。図3は、第1実施形態のビデオ会議システム1により実行される処理の流れを例示するフローチャートである。

#### 【0032】

画像取得部110は、通信端末20Aまたは通信端末20Bから、会議の参加人物が写る画像を取得する(S102)。そして、人物領域検出部120は、S102の処理で取得された画像の中から、人物と認識される領域を検出する(S104)。

#### 【0033】

人物特定部130は、S104の処理で検出された領域について、人物特定処理を実行する(S106)。人物特定部130は、人物が特定できなかった場合(S108:NO)、人物特定処理の対象となった領域に対して、その領域に含まれる人物が特定できなかったことを示す情報(特定失敗情報)を関連付ける(S110)。一方、人物が特定できなかった場合(S108:NO)、上述の特定失敗情報を関連付ける処理は実行されない。なお、S106からS110の処理は、S104の処理で検出された人物の領域の全てが処理されるまで繰り返される(S112:NO)。

#### 【0034】

S104の処理で検出された人物の領域の全てが処理された後(S112:YES)、表示制御部140は、各地点で撮影された参加人物の画像を、互いの相手いる地点に備えられた表示装置にそれぞれ表示させる(S114)。ここで、表示制御部140は、S108の処理で特定失敗情報が関連づけられた領域が存在する場合、その人物が特定されていないことが分かるように、例えば、その領域の人物の位置に合わせて第1情報を重畳表示させる(例:図4)。

#### 【0035】

図4は、表示制御部140により表示される第1情報の一例を示す図である。図4の例では、画像において後ろ姿が写っている2人の人物に対して、第1情報「Unknown」が重畳表示されている様子が描かれている。なお、第1情報は、その人物が特定できなかったことを示す特定の印であればよく、「Unknown」という表示に制限されない。この図では、次のようにして、「Unknown」という第1情報が表示される。まず、人物領域検出部120が、図4の画像から、4人の人物領域を検出する。検出された4人のうちの2人は後ろ向き(目、鼻、口といった個人を特定するために利用される特徴量が抽出できない状況)で写っているため、人物特定部130は、人物特定処理でこれらの人物の特定に失敗する。人物特定部130は、人物特定処理の結果に基づいて、これら2人の領域に対して特定失敗情報を関連付ける。表示制御部140は、各人物の領域に関連付けられた特定失敗情報に基づいて画像上に第1情報を重畳させて、表示装置に向けて出力する。このデータを受け取った表示装置は、例えば、図4に例示されるような画像を、その表示面に表示することができる。

#### 【0036】

本実施形態では、ビデオ会議システム1を利用して開催される会議の参加人物の中に特定できていない人物が存在する場合、その会議で利用される表示装置に第1情報が表示される。この第1情報は、人物が特定されていないことを示す情報(例えば、「Unkno

10

20

30

40

50

w n」などの表示)である。つまり、会議の参加人物は、ビデオ会議システム1で特定されていない人物がいることを、この第1情報によって一目で把握することができる。また、第1情報は画像の中で特定できなかった人物の領域の位置に合わせて表示される。これにより、会議の参加人物は、どの人物が特定されていないかを認識できる。そして、会議の参加人物は、ビデオ会議システム1がその人物を特定できるように適切な措置を取ることができる。例えば、会議の参加人物は、特定されていない人物に対して、顔(人物特定時に利用される特徴点を含む領域)がはっきりと画像に写り込むように、顔の向きや姿勢を変えるように促すことができる。

#### 【0037】

ここで、表示制御部140は、画像に写っている会議の参加人物が見る表示装置に、第1情報を表示させてもよい。すなわち、人物特定処理で特定できなかった人物がいた場合、表示制御部140は、その人物がいる地点に設けられた表示装置に第1情報を表示させてもよい。例えば、地点Aにいる会議の参加人物を人物特定処理で特定できなかったとする。この場合、表示制御部140は、撮像装置30Aで生成された画像に第1情報を重畳させたデータを、通信端末20Aに送信する。通信端末20Aは、表示装置40Aに、撮像装置30Aで生成された画像と、その画像に重畳された第1情報とを表示させる。このようにすることで、地点A側にいる会議の参加人物が、特定されていない(認証されていない)人物がいることを容易に把握することができる。その結果、例えば、特定されなかった人物は、自身の顔(人物特定時に利用される特徴点を含む領域)がはっきりと画像に写り込むように、顔の向きや姿勢を自発的に変えるといった措置を取ることができる。

#### 【0038】

##### [第2実施形態]

本実施形態は、以下の点を除き、上述の第1実施形態と同様の構成を有する。

#### 【0039】

本実施形態の表示制御部140は、人物特定部130の人物特定処理で会議の参加人物が特定された場合、その人物の氏名を含む情報(第2情報)を表示装置に更に表示させる(例:図5)。図5は、人物領域検出部120により表示される第2情報の一例を示す図である。図5の例において、画像中の奥側に座っている2人の人物については、目、鼻、口といった特徴点があきらかに写っており、人物特定部130はこれら2人の人物を特定できたとする。この場合、人物特定部130は、特定した人物の氏名を含む第2情報を取得し、その人物の領域に関連付ける。なお、第2情報は、会議の開催前に参加人物の特徴量と関連付けて取得され、ストレージデバイス1040などに事前に登録されている。第2情報は、人物の氏名のほか、その人物が所属するグループ(会社や部署など)の名称、その人物の役職名などを更にも含んでもよい。そして、表示制御部140は、人物特定部130の人物特定処理の結果に基づいて、第2情報が関連付けられた領域の人物の位置に合わせて、その第2情報を重畳表示させる。その結果、例えば、図5に例示されるような画像が、ビデオ会議システム1を利用して開催される会議の参加人物が見る表示装置40に表示される。これにより、会議の参加人物の名前や所属といった、その人物に関する情報が一目で把握できるようになる。

#### 【0040】

表示制御部140は、第1情報と第2情報とを併せて表示させる際、第1情報の表示態様を、第2情報の表示態様と異ならせてもよい。言い換えると、表示制御部140は、視覚的に第1情報と第2情報とを異ならせてもよい。一例として、表示制御部140は、第1情報を第2情報よりも目立たせてもよい。具体的には、表示制御部140は、第1情報の外形、大きさ、色、言語、およびフォントの少なくともいづれを、第2情報と異ならせることにより、第1情報を第2情報よりも目立たせることができる。例えば、表示制御部140は、第1情報を第2情報よりも大きく表示したり、第1情報の色を目立つ色に設定したり、第1情報のフォントを標準的なフォントとは異なる特殊なフォントに設定したりして、第1情報を第2情報よりも目立たせることができる。図6は、表示制御部140による表示の一例を示す図である。図6の例では、第1表示の背景色を変えることにより、

第 2 表示よりも目立たせている様子が描かれている。このようにすることで、会議の参加人物が、ビデオ会議システム 1 で特定できていない人物の存在に気づき易くなる。

【 0 0 4 1 】

また、表示制御部 1 4 0 は、人物特定処理で特定された、会議の参加人物の数が所定の閾値以上である場合、第 2 情報を一覧形式で表示させてもよい（例：図 7）。図 7 は、第 2 情報を一覧形式で表示する例を示す図である。図 7 に示されるように、表示制御部 1 4 0 は、画像中の空き領域（人物の領域以外の領域）を特定して、その領域に一覧形式の第 2 情報を表示することができる。表示制御部 1 4 0 は、空き領域のサイズに応じて、一覧の大きさを決定してもよい。また、表示制御部 1 4 0 は、一覧をスクロールバーと共に表示させて、一覧を表示する領域を節約してもよい。このようにすることで、第 2 情報が多数表示されることにより会議の参加人物の顔（画像の主となる情報）の視認性の低下を抑制できる。

10

【 0 0 4 2 】

〔 第 2 実施形態の変形例 〕

表示制御部 1 4 0 は、第 1 情報を第 2 情報よりも目立たなくさせてもよい。特に、表示制御部 1 4 0 は、ある地点の参加人物の画像を相手側の地点に送信する場合において、第 1 情報を第 2 情報よりも目立たなくさせると好ましい。例えば、表示制御部 1 4 0 は、相手側に送信する画像において、第 1 情報を小さくしたり、第 1 情報の色を薄くしたり、第 1 情報を非表示としたりして、第 1 情報を第 2 情報よりも目立たなくさせることができる。ある地点の参加人物の画像を相手側の地点に送信する場合に第 1 情報を目立たなくすることによって、その相手側にとってより重要度の高い情報（人物の顔または人物の動作など、会議で本来必要な情報）の視認性が低下することを抑制できる。

20

【 0 0 4 3 】

〔 第 3 実施形態 〕

本実施形態のビデオ会議システム 1 は、ビデオ会議システム 1 で特定できなかった人物がいる場合にその人物を特定するための構成を更に有している点を除き、上述の各実施形態と同様の構成を有する。

【 0 0 4 4 】

〔 システム構成 〕

図 8 は、第 3 実施形態におけるビデオ会議システム 1 の構成例を示す図である。図 8 に例示されるビデオ会議システム 1 では、上述の撮像装置 3 0（3 0 A、3 0 B）とは別に、移動型の撮像装置 3 2（3 2 A、3 2 B）が各地点に更に備えられている。移動型の撮像装置 3 2 は、例えば、カメラ機能付き携帯端末（スマートフォンやタブレット端末など）、或いは、カメラ付きノート P C（Personal Computer）などといった、個人が所有する装置である。

30

【 0 0 4 5 】

本実施形態において、画像取得部 1 1 0 は、撮像装置 3 0 を介して取得された画像に基づく人物特定処理で特定されなかった会議の参加人物が存在する場合、その参加人物が写る追加の画像を、移動型の撮像装置 3 2 A または撮像装置 3 2 B から取得する。そして、画像取得部 1 1 0 は、画像取得部 1 1 0 により取得された追加画像に基づいて、特定されなかった参加人物を特定する。

40

【 0 0 4 6 】

〔 処理の流れ 〕

図 9 を用いて、本実施形態のビデオ会議システム 1 により実行される処理の流れについて説明する。図 9 は、第 3 実施形態のビデオ会議システム 1 により実行される処理の流れを例示するフローチャートである。

【 0 0 4 7 】

まず、図 3 の S 1 0 6 の人物特定処理で特定されなかった人物の追加画像が、移動型の撮像装置 3 2 により生成される（S 2 0 2）。なお、会議の参加人物は、表示装置 4 0 上に表示された第 1 情報に基づいて、自身がビデオ会議システム 1 において個人として特定

50

されているか否かを把握することができる。表示装置 40 上に表示された第 1 情報から、自分が特定されていないことを認識した人物は、例えば、スマートフォンやノート PC にインストールされた専用のアプリケーションを立ち上げて、追加画像を撮影する。撮影された追加画像は、ネットワークを介してサーバ装置 10 に送信され、画像取得部 110 により取得される (S204)。ここで、撮像装置 30 により生成された画像の中に複数の未特定人物が存在する場合もある。この場合には、追加画像がどの未特定人物に対応する画像かを示す情報が必要となる。そこで、撮像装置 30 により生成された画像の中に複数の未特定人物が存在する場合、一例として、スマートフォンやノート PC 上で立ち上げたアプリケーションで、追加画像に対応する未特定人物を指定する操作を更に受け付けてもよい。例えば、スマートフォンやノート PC 上で立ち上げたアプリケーションは、追加画像の撮影前または撮影後に図 4 から図 6 などに例示される画面をスマートフォンやノート PC の表示面に表示させ、その画像の中から未特定人物を選択する操作を受け付けてもよい。この場合、画像取得部 110 は、追加画像と、その追加画像に対応する未特定人物を示す情報とを取得することができる。

10

#### 【0048】

人物特定部 130 は、スマートフォンやノート PC といった移動型の撮像装置 32 により生成された追加画像を用いて、人物特定処理を実行する (S206)。具体的には、図 3 の S106 の人物特定処理と同様に、追加画像から抽出される特徴量と、予め登録された参加人物の特徴量とを照合することによって、その追加画像の人物を特定することができる。ここで、追加画像が不鮮明であって人物が特定できない場合などには、人物特定部 130 は、その追加画像の送信元の装置に対し、画像の撮り直しを促すメッセージを出力するように構成されていてもよい。また、人物特定部 130 は、追加画像の取り直しを予め決められた回数行ったにもかかわらず未特定人物が特定できなかった場合、その未特定人物を部外者 (会議の参加人物として予め登録された人物以外の人物) と判断してもよい。この場合、人物特定部 130 は、表示装置 40 や図示しないスピーカーなどを用いて、部外者の存在を報知する処理を実行してもよい。このようにすることで、会議の参加人物が、会議の場に紛れ込んだ部外者の存在を認識することができる。

20

#### 【0049】

以上、本実施形態によれば、移動型の撮像装置 32 から、ビデオ会議システム 1 が特定できていない人物について、その人物を特定するための追加画像を取得することができる。なお、移動型の撮像装置 32 を利用することにより、その人物の特徴点がより鮮明に写る画像を取得することができる。結果として、ビデオ会議システム 1 が特定できていない人物を、精度よく特定することができる。

30

#### 【0050】

なお、人物特定部 130 は、会議の参加人物を一度特定すれば、その後取得された画像から人物の特徴量が抽出できないような場合 (例えば、俯いてしまって顔の特徴点が見えなくなった場合など) であっても、その人物が特定された状態を維持することができる。例えば、人物特定部 130 は、複数の画像間で、特定した人物の領域を画像中の位置などに基づいて追跡することにより、その人物が特定された状態を維持することができる。

#### 【0051】

40

#### 〔第 3 実施形態の変形例〕

本実施形態において、移動型の撮像装置 32 により生成された追加画像を用いる代わりに、会議の参加人物のいずれかが、特定できなかった人物を特定するための情報を直接入力してもよい。一例として、次のような動作が実行されてもよい。まず、会議の参加人物は、表示装置 40 上で「Unknown」といった第 1 情報が関連付けられている人物を確認した後、その人物の氏名といった情報を携帯型端末 (スマートフォンやノート PC など) 上で立ち上げたアプリケーションを介して入力する。なお、特定できていない人物が複数存在する場合には、それら複数の人物の中で対象とする人物を選択する入力が更に実行される。そして、人物特定部 130 は、特定できなかった人物の領域に関連付けられている特定失敗情報を、入力された情報を用いて更新する。これにより、特定できていなか

50

った人物は、ビデオ会議システム 1 で特定（認証）された状態となる。また、この結果、「Unknown」といった第 1 情報の表示は、入力された人物の氏名に置き換わることになる。

【0052】

〔第 4 実施形態〕

本実施形態では、自動的に議事録を作成する機能を更に有する点を除き、上述の各実施形態と同様の構成を有する。

【0053】

図 10 は、第 4 実施形態におけるビデオ会議システム 1 の構成例を示す図である。図 10 に例示されるビデオ会議システム 1 は、リスト作成部 150、音声取得部 160、発言者特定部 170、議事録作成部 180 を更に備える。

10

【0054】

リスト作成部 150 は、人物特定部 130 の人物特定処理によって特定された人物のリストを作成する。リスト作成部 150 は、例えば次のように動作する。まず、リスト作成部 150 は、人物特定部 130 の人物特定処理で人物が特定された場合に、人物特定部 130 からその結果を取得する。そして、リスト作成部 150 は、人物特定部 130 から取得した人物の特定結果を、メモリ 1030 などに保持されるリストに追加する。これにより、ビデオ会議システム 1 を利用して開催される会議の参加者のリストを自動的に生成することができる。

【0055】

音声取得部 160 は、図示しないマイクにより生成された、会議中の会話の音声データを取得する。発言者特定部 170 は、音声取得部 160 により取得された音声データに関する発言者を特定する。一例として、発言者特定部 170 は、例えば会議の開催前にストレージデバイス 1040 などに予め登録された各参加人物の声紋データとの照合を行うことにより、音声取得部 160 が取得した音声データに関する発言者を特定することができる。他の一例として、発言者特定部 170 は、音声データと同期して取得される画像（撮像装置 30 により生成される画像）を解析することによって、音声取得部 160 が取得した音声データに関する発言者を特定することができる。具体的には、発言者特定部 170 は、音声データと同期して取得された画像を解析した結果、口の部分が動いている人物の領域を特定する。そして、口の部分が動いている人物の領域についての人物特定処理の結果から、その発話者を特定することができる。議事録作成部 180 は、発言者特定部 170 による発言者の特定結果と、音声取得部 160 により取得された音声データに基づいて生成されたテキストデータとを対応付けることにより、議事録データを生成する。また、議事録作成部 180 は、リスト作成部 150 により生成された人物のリストを、会議の参加者として議事録データに付加することができる。

20

30

【0056】

〔ハードウェア構成例〕

本実施形態のビデオ会議システム 1 は、第 1 実施形態と同様のハードウェア構成（例：図 2）を有する。本実施形態のストレージデバイス 1040 は、上述のリスト作成部 150、音声取得部 160、発言者特定部 170 および議事録作成部 180 の機能を実現するためのプログラムモジュールを更に記憶している。プロセッサ 1020 が、これらのプログラムモジュールをメモリ 1030 上に読み出して実行することにより、上述の本実施形態の各機能が実現される。

40

【0057】

〔処理の流れ〕

図 11 を用いて、本実施形態のビデオ会議システム 1 により実行される処理の流れについて説明する。図 11 は、第 4 実施形態のビデオ会議システム 1 により実行される処理の流れを例示するフローチャートである。

【0058】

まず、音声取得部 160 は会議の音声データを取得する（S302）。会議の音声デー

50

タは、各地点に設けられている集音装置 50 により生成される。集音装置 50 は、通信端末 20 に接続されている。音声取得部 160 は、ネットワークインタフェース 1060 を介して各地点の通信端末 20 と通信して、その地点の集音装置 50 により生成された音声データを取得することができる。

#### 【0059】

そして、発言者特定部 170 は、音声取得部 160 により取得された音声データに関する発言者を特定する (S304)。一例として、発言者特定部 170 は、次のようにして、音声取得部 160 により取得された音声データに関する発言者を特定することができる。まず、発言者特定部 170 は、ストレージデバイス 1040 などに事前に登録された各参加人物の声紋データと音声データとを照合して、当該音声データの声紋との一致度が基準を満たす声紋データを特定する。そして、発言者特定部 170 は、特定した声紋データに関連付けられている参加人物の識別情報 (人物の氏名、または、人物毎に割り当てられた ID など) を取得することにより、音声取得部 160 により取得された音声データの発言者を特定することができる。他の一例として、発言者特定部 170 は、次のようにして、音声取得部 160 により取得された音声データに関する発言者を特定することができる。まず、発言者特定部 170 は、音声データと同期して画像取得部 110 により取得された画像を解析する。具体的には、発言者特定部 170 は、画像の中から人物の口の領域を検出し、その領域 (すなわち、口) が時系列で並ぶ複数の画像間で動いているか否かを判定する。そして、発言者特定部 170 は、口の領域が動いていると判定された人物の領域について、人物特定部 130 の人物特定処理の結果を取得することにより、音声取得部 160 により取得された音声データの発言者を特定することができる。なお、ここでは、例えば、第 3 実施形態で説明したような構成を利用して、全ての人物が特定されているものと仮定している。

#### 【0060】

議事録作成部 180 は、音声取得部 160 および発言者特定部 170 の処理結果に基づいて、議事録データを生成する (S306)。具体的には、議事録作成部 180 は、音声データをテキスト化する API (Application Programming Interface) などを利用して、音声取得部 160 により取得された音声データをテキストデータ化する。また、議事録作成部 180 は、発言者特定部 170 によって特定された、当該音声データの発言者の情報 (例えば、発言者の氏名など) を取得する。そして、議事録作成部 180 は、音声取得部 160 により取得された音声データから生成されたテキストデータと、その音声データに関する発言者として特定された人物の情報とを対応付けて、議事録データに追加する。また、議事録作成部 180 は、リスト作成部 150 により生成された、会議の参加人物リストを読み出し、議事録データに参加人物の情報を付加してもよい。

#### 【0061】

以上、本実施形態の構成によれば、ビデオ会議システム 1 を利用した開催される会議の議事録を、自動で作成することができる。これにより、会議の参加人物が議事録の作成する手間を削減することができる。

#### 【0062】

以上、図面を参照して本発明の実施形態について述べたが、これらは本発明の例示であり、上記以外の様々な構成を採用することもできる。

#### 【0063】

また、上述の説明で用いた複数のフローチャートでは、複数の工程 (処理) が順番に記載されているが、各実施形態で実行される工程の実行順序は、その記載の順番に制限されない。各実施形態では、図示される工程の順番を内容的に支障のない範囲で変更することができる。また、上述の各実施形態は、内容が相反しない範囲で組み合わせることができる。

#### 【0064】

上記の実施形態の一部または全部は、以下の付記のようにも記載されうるが、以下に限られない。

10

20

30

40

50

- 1 .  
会議の参加人物が写る画像を取得する画像取得手段と、  
前記画像の中から、人物と認識される領域を検出する人物領域検出手段と、  
前記領域に含まれる人物を特定する人物特定処理を実行する人物特定手段と、  
前記人物特定処理で前記人物が特定できなかったことを示す第 1 情報を表示装置に表示させる表示制御手段と、  
を備えるビデオ会議システム。
- 2 .  
前記表示制御手段は、前記人物特定処理で特定された前記会議の参加人物の氏名を含む第 2 情報を前記表示装置に更に表示させる、  
1 . に記載のビデオ会議システム。 10
- 3 .  
前記表示制御手段は、前記第 1 情報の表示態様を、前記第 2 情報の表示態様と異ならせる、  
2 . に記載のビデオ会議システム。
- 4 .  
前記表示制御手段は、前記第 1 情報を前記第 2 情報よりも目立たせる、  
3 . に記載のビデオ会議システム。
- 5 .  
前記表示制御手段は、前記第 1 情報の外形、大きさ、色、言語、およびフォントの少なくとも 1 つを前記第 2 情報と異ならせることにより、前記第 1 情報を目立たせる、  
4 . に記載のビデオ会議システム。 20
- 6 .  
前記表示制御手段は、前記人物特定処理により特定された前記会議の参加人物の数が所定の閾値以上である場合、前記第 2 情報を一覧形式で表示させる、  
2 . から 5 . のいずれか 1 つに記載のビデオ会議システム。
- 7 .  
前記表示制御手段は、前記画像に写っている前記会議の参加人物が見る表示装置に、前記第 1 情報を表示させる、  
1 . から 6 . のいずれか 1 つに記載のビデオ会議システム。 30
- 8 .  
前記画像取得手段は、前記画像に基づく前記人物特定処理で特定されなかった参加人物が写る追加画像を、前記画像を生成した撮像装置とは異なる移動型の撮像装置から取得し、  
前記人物特定手段は、前記追加画像に基づいて前記画像に基づく前記人物特定処理で特定されなかった参加人物を特定する、  
1 . から 7 . のいずれか 1 つに記載のビデオ会議システム。
- 9 .  
前記人物特定処理によって特定された人物のリストを作成するリスト作成手段を更に備える、  
8 . に記載のビデオ会議システム。 40
- 10 .  
音声データを取得する音声取得手段と、  
前記音声データまたは前記音声データと同期して取得された画像を解析することによって、前記音声データに関する発言者を特定する発言者特定手段と、  
前記発言者の特定結果と前記音声データに基づいて生成されたテキストデータとを対応付けることにより、議事録データを生成する議事録作成手段と、を更に備える、  
8 . または 9 . に記載のビデオ会議システム。
- 11 .  
前記表示制御手段は、前記第 1 情報を前記第 2 情報よりも目立たなくさせる、  
3 . に記載のビデオ会議システム。 50

- 12 .  
コンピュータが、  
会議の参加人物が写る画像を取得し、  
前記画像の中から、人物と認識される領域を検出し、  
前記領域に含まれる人物を特定する人物特定処理を実行し、  
前記人物特定処理で前記人物が特定できなかったことを示す第1情報を表示装置に表示させる、  
ことを含むビデオ会議方法。
- 13 .  
前記コンピュータが、  
前記人物特定処理で特定された前記会議の参加人物の氏名を含む第2情報を前記表示装置に更に表示させる、  
ことを含む12.に記載のビデオ会議方法。 10
- 14 .  
前記コンピュータが、  
前記第1情報の表示態様を、前記第2情報の表示態様と異ならせる処理を実行する、  
ことを含む13.に記載のビデオ会議方法。
- 15 .  
前記コンピュータが、前記第1情報を前記第2情報よりも目立たせる処理を実行する、  
ことを含む14.に記載のビデオ会議方法。 20
- 16 .  
前記コンピュータが、  
前記第1情報の外形、大きさ、色、言語、およびフォントの少なくとも1つを前記第2情報と異ならせることにより、前記第1情報を目立たせる、  
ことを含む15.に記載のビデオ会議方法。
- 17 .  
前記コンピュータが、  
前記人物特定処理により特定された前記会議の参加人物の数が所定の閾値以上である場合、前記第2情報を一覧形式で表示させる、  
ことを含む13.から16.のいずれか1つに記載のビデオ会議方法。 30
- 18 .  
前記コンピュータが、  
前記画像に写っている前記会議の参加人物が見る表示装置に、前記第1情報を表示させる、  
ことを含む12.から17.のいずれか1つに記載のビデオ会議方法。
- 19 .  
前記コンピュータが、  
前記画像に基づく前記人物特定処理で特定されなかった参加人物が写る追加画像を、前記画像を生成した撮像装置とは異なる移動型の撮像装置から取得し、  
前記追加画像に基づいて前記画像に基づく前記人物特定処理で特定されなかった参加人物を特定する、  
ことを含む12.から18.のいずれか1つに記載のビデオ会議方法。 40
- 20 .  
前記コンピュータが、  
前記人物特定処理によって特定された人物のリストを作成する、  
ことを含む19.に記載のビデオ会議方法。
- 21 .  
前記コンピュータが、  
音声データを取得し、  
前記音声データまたは前記音声データと同期して取得された画像を解析することによっ 50

て、前記音声データに関する発言者を特定し、

前記発言者の特定結果と前記音声データに基づいて生成されたテキストデータとを対応付けることにより、議事録データを生成する、

ことを含む 19 . または 20 . に記載のビデオ会議方法。

22 .

前記コンピュータが、

前記第 1 情報を前記第 2 情報よりも目立たなくさせる処理を実行する、

ことを含む 14 . に記載のビデオ会議方法。

23 .

コンピュータに、12 . から 22 . のいずれか 1 つに記載のビデオ会議方法を実行させるプログラム。

10

【符号の説明】

【0065】

1 ビデオ会議システム

10 サーバ装置

1010 バス

1020 プロセッサ

1030 メモリ

1040 ストレージデバイス

1050 入出力インタフェース

20

1060 ネットワークインタフェース

110 画像取得部

120 人物領域検出部

130 人物特定部

140 表示制御部

150 リスト作成部

160 音声取得部

170 発言者特定部

180 議事録作成部

20 通信端末

30

30 撮像装置

32 撮像装置

40 表示装置

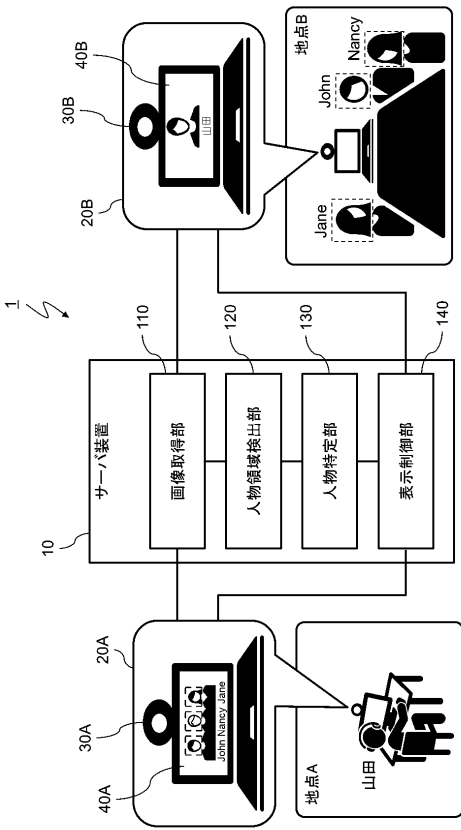
50 集音装置

40

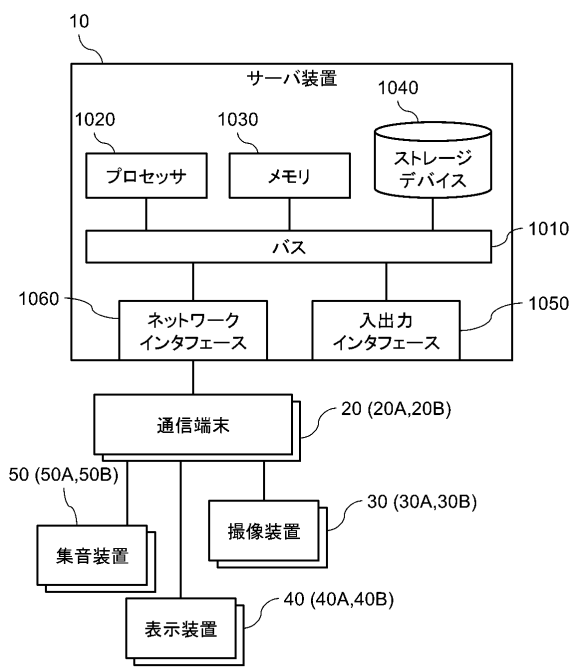
50

【図面】

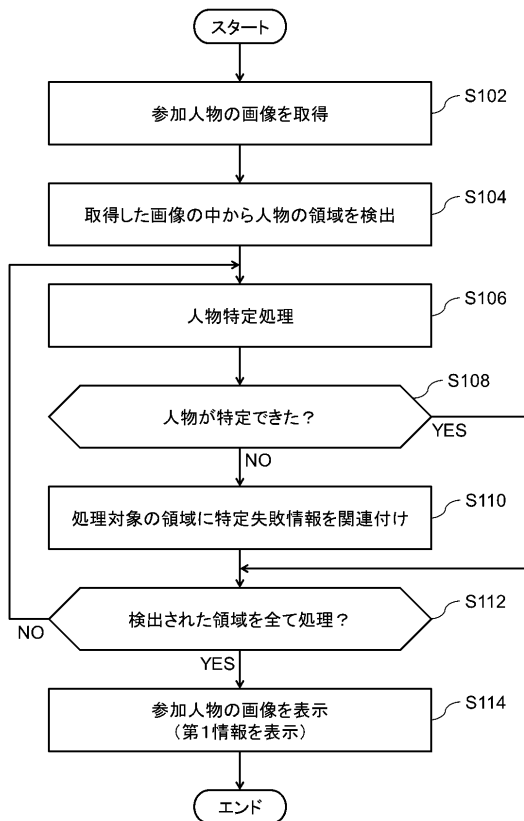
【図 1】



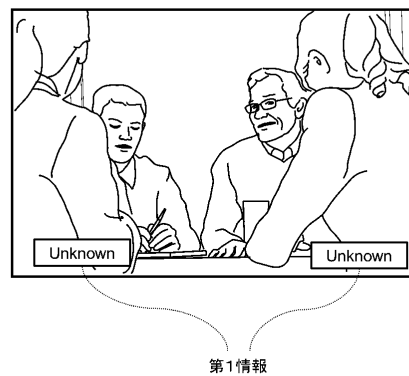
【図 2】



【図 3】



【図 4】



10

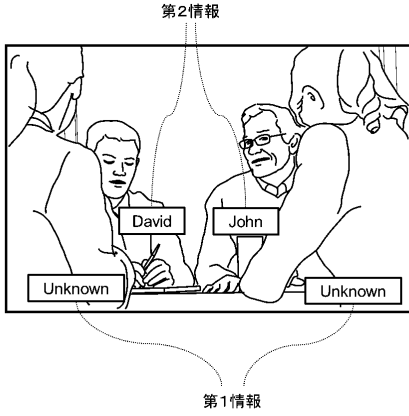
20

30

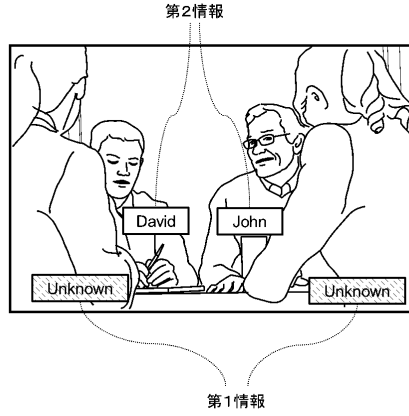
40

50

【図5】



【図6】

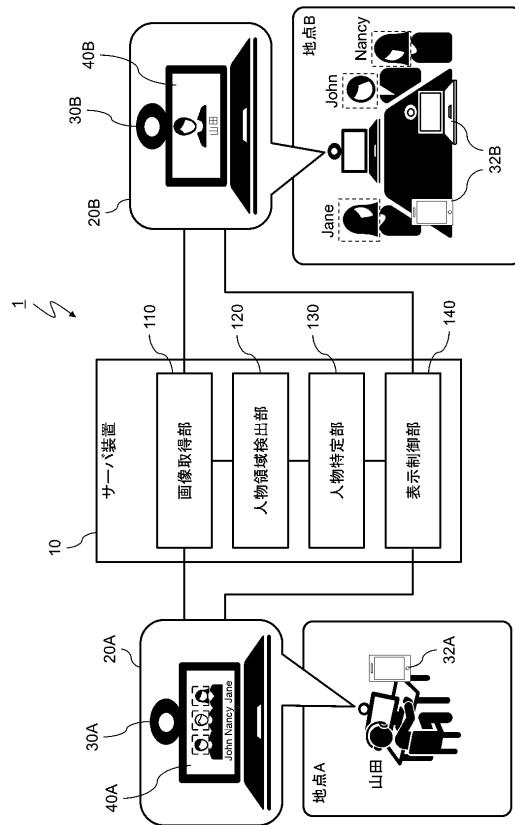


10

【図7】



【図8】



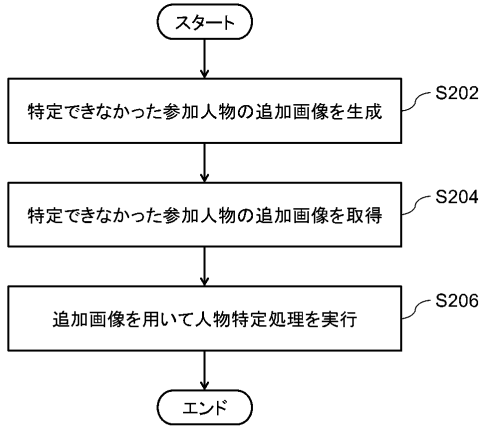
20

30

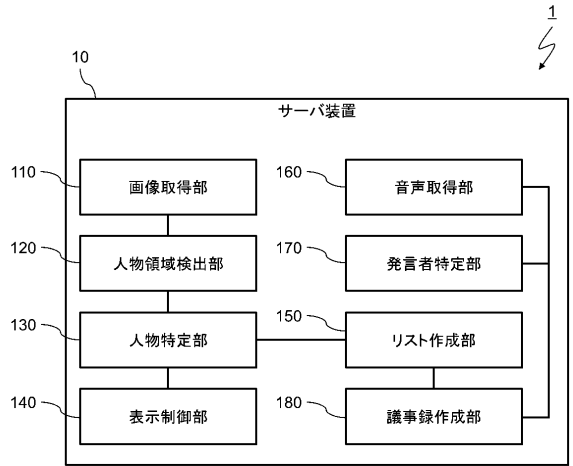
40

50

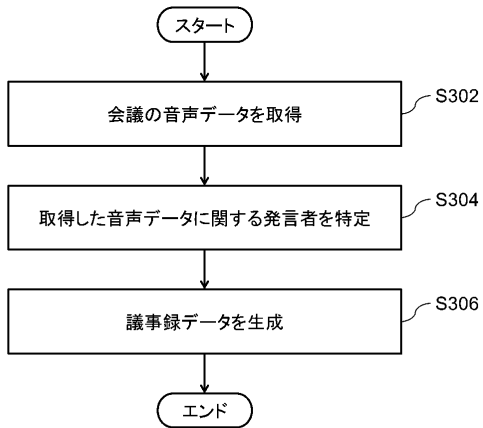
【図 9】



【図 10】



【図 11】



10

20

30

40

50

---

フロントページの続き

- (56)参考文献 特開2011-004007(JP,A)  
特開2001-195165(JP,A)  
特開2003-023612(JP,A)  
国際公開第2015/033599(WO,A1)  
特開2011-045028(JP,A)  
特開2009-206924(JP,A)
- (58)調査した分野 (Int.Cl., DB名)
- |      |      |   |       |
|------|------|---|-------|
| H04M | 3/38 | - | 3/58  |
| H04N | 7/14 | - | 7/173 |