

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7525598号
(P7525598)

(45)発行日 令和6年7月30日(2024.7.30)

(24)登録日 令和6年7月22日(2024.7.22)

(51)国際特許分類

F I

G O 6 T 19/00 (2011.01)

G O 6 T 19/00 3 0 0 B

H O 4 N 7/15 (2006.01)

H O 4 N 7/15

請求項の数 29 (全35頁)

(21)出願番号	特願2022-518567(P2022-518567)	(73)特許権者	519155642
(86)(22)出願日	令和2年4月30日(2020.4.30)		佐藤 壘
(86)国際出願番号	PCT/JP2020/018345		東京都千代田区東神田3-3-8 TK
(87)国際公開番号	WO2021/220494		S東神田ビル地下1階
(87)国際公開日	令和3年11月4日(2021.11.4)	(74)代理人	110000279
審査請求日	令和5年5月1日(2023.5.1)		弁理士法人ウィルフォート国際特許事務所
		(72)発明者	佐藤 壘
			東京都千代田区東神田3-3-8 TK
			S東神田ビル地下1階
		審査官	鈴木 明

最終頁に続く

(54)【発明の名称】 通信端末装置、通信方法、およびソフトウェアプログラム

(57)【特許請求の範囲】

【請求項1】

通信回線を介して複数で相互接続する通信端末装置であって、
自装置を用いる自ユーザの声を含む音声を取得する音声取得部と、
所定のユーザ空間上にある人物の三次元データを形成可能にする視覚情報を取得する視覚情報取得部と、
前記ユーザ空間の視覚情報に基づく画像伝達情報と、前記音声に基づく音声伝達情報とを他装置に伝達する情報伝達部と、
前記他装置から伝達された画像伝達情報に基づく、前記自装置のユーザ空間と前記他装置のユーザ空間とを配置した共有仮想空間上にある人物の三次元データを、前記共有仮想空間に対して相対位置が固定された表示面に投影するように射影変換をして生成される、前記自ユーザの視点位置から前記表示面を介して共有仮想空間上の前記人物を見たときのような見え方となる表示画像を、前記表示面に表示する画像表示部と、
前記他装置から伝達された音声伝達情報に基づいて音声を出力する音声出力部と、
を有する通信端末装置。

【請求項2】

継続的に前記ユーザ空間の視覚情報に基づいて前記自ユーザの視点位置を特定する人物捕捉部を更に有し、
前記画像表示部は、前記特定される視点位置を追跡して前記表示画像を生成し、前記表示面に表示する、

請求項 1 に記載の通信端末装置。

【請求項 3】

前記人物捕捉部は、前記ユーザ空間の視覚情報に基づいて、前記自ユーザを検出し、
前記情報伝達部は、前記自ユーザが検出されると、前記画像伝達情報として、前記自ユーザの顔部分を表す情報を前記他装置に送信し、

前記画像表示部は、前記他装置から受信した顔部分を表す情報に基づく顔画像を含む画像を前記表示面に表示する、

請求項 2 に記載の通信端末装置。

【請求項 4】

前記情報伝達部は、前記自ユーザが検出されると、前記画像伝達情報として、更に、前記自ユーザの骨格を表す情報を前記他装置に送信し、

前記画像表示部は、前記他装置から受信した顔部分を表す情報に基づく顔画像と、前記他装置から受信した骨格を表す情報に基づく画像とを含む前記共有仮想空間の画像を前記表示面に表示する、

請求項 3 に記載の通信端末装置。

【請求項 5】

前記画像表示部は、前記共有仮想空間の人物を加工処理が可能な点群および／もしくはポリゴンメッシュにより表示する、

請求項 1 に記載の通信端末装置。

【請求項 6】

前記音声出力部は、前記他装置から伝達された音声伝達情報に基づいて、前記共有仮想空間における前記他装置のユーザ空間が配置された方向から聞こえる音声を再生する、

請求項 1 に記載の通信端末装置。

【請求項 7】

前記ユーザ空間における所定の位置をユーザ定位置とし、

前記視覚情報取得部は、それぞれの撮像方向の視覚情報を取得する複数の撮像器を有し、前記複数の撮像器は、それぞれが、互いに隔離して前記ユーザ定位置との相対位置を固定し、前記ユーザ空間の少なくとも一部が撮像される方向に前記撮像方向を向けて配置され、

前記情報伝達部は、前記他装置の複数の撮像器のうち、前記撮像器の撮像方向と、前記他装置のユーザ定位置と前記自装置のユーザ定位置を結ぶ直線との相対角度に基づいて選択した撮像器のみについて、当該撮像器で取得された視覚情報に基づく画像伝達情報を、前記他装置から取得する、

請求項 1 に記載の通信端末装置。

【請求項 8】

前記ユーザ空間における所定の位置をユーザ定位置とし、

前記視覚情報取得部は、前記ユーザ定位置から所定の方向をユーザ定方向とし、前記ユーザ定方向の右側に配置された右側撮像器と、前記ユーザ定方向の左側に配置された左側撮像器とを含み、

前記情報伝達部は、前記共有仮想空間において前記自装置のユーザ定位置から前記ユーザ定方向の右側に前記他装置のユーザ定位置が配置される場合には前記他装置の右側撮像器による情報を含み前記他装置の左側撮像器による情報を含まない画像伝達情報を前記他装置から取得し、前記共有仮想空間において前記自装置のユーザ定位置から前記ユーザ定方向の左側に前記他装置のユーザ定位置が配置される場合には前記他装置の左側撮像器による情報を含み前記他装置の右側撮像器による情報を含まない画像伝達情報を前記他装置から受信する、

請求項 1 に記載の通信端末装置。

【請求項 9】

前記ユーザ空間における所定の位置をユーザ定位置とし前記ユーザ定位置から所定の方向をユーザ定方向とし、

10

20

30

40

50

前記画像表示部は、互いに並んだ第1線と第2線のそれぞれの上に前記自装置および前記他装置のうち少なくとも1つの装置のユーザ定位置を配置し、前記第1線の上にユーザ定位置が配置された装置のユーザ定方向を前記第2線に向かう方向とし、前記第2線の上にユーザ定位置が配置された装置のユーザ定方向を前記第1線に向かう方向とするように、前記共有仮想空間を形成する、

請求項1に記載の通信端末装置。

【請求項10】

前記ユーザ空間における所定の位置をユーザ定位置とし前記ユーザ定位置から所定の方向をユーザ定方向とし、

前記画像表示部は、前記自装置および前記他装置のユーザ定位置をリングもしくは多角形状に配置し、前記自装置および前記他装置のユーザ定方向を前記リングもしくは多角形の内側に向かう方向とするように、前記共有仮想空間を形成する、

請求項1に記載の通信端末装置。

【請求項11】

前記ユーザ空間における所定の位置をユーザ定位置とし前記ユーザ定位置から所定の方向をユーザ定方向とし、

前記画像表示部は、所定の線の上に前記自装置および前記他装置のユーザ定位置を配置し、前記自装置および前記他装置のユーザ定方向を前記線に対して同一方向とするように、前記共有仮想空間を形成する、

請求項1に記載の通信端末装置。

【請求項12】

前記ユーザ空間における所定の位置をユーザ定位置とし前記ユーザ定位置から所定の方向をユーザ定方向とし、

前記画像表示部は、前記ユーザ定位置から見て前記ユーザ定方向に配置された矩形で平面の表示面を有する正面表示器と、前記ユーザ定位置から見て前記ユーザ定方向の右側に配置された矩形で平面の表示面を有する右側表示器と、前記ユーザ定位置から見て前記ユーザ定方向の左側に配置された矩形で平面の表示面を有する左側表示器とを有し、前記右側表示器の表示面は、前記正面表示器の表示面の右辺に近接し、前記左側表示器の表示面は、前記正面表示器の表示面の左辺と近接する、

請求項1に記載の通信端末装置。

【請求項13】

前記画像表示部は、前記正面表示器、前記左側表示器、前記右側表示器からなる3つの表示器の底辺に近接して配置され矩形の上面に表示面を有する下側表示器を更に有する、および／または、前記正面表示器、前記左側表示器、前記右側表示器からなる3つの表示器の天井側の辺に近接して配置され矩形の下面に表示面を有する上側表示器を更に有する、

請求項12に記載の通信端末装置。

【請求項14】

前記ユーザ空間における所定の位置をユーザ定位置とし、

前記画像表示部は、前記ユーザ定位置から見て所定の方向であるユーザ定方向に配置された矩形で平面の表示面を有する正面表示器と、前記正面表示器の底辺に近接して配置され矩形の上面に表示面を有する下側表示器を更に有する、および／または、前記正面表示器の上辺に近接して配置され矩形の下面に表示面を有する上側表示器を更に有する、

請求項1に記載の通信端末装置。

【請求項15】

前記ユーザ空間における所定の位置をユーザ定位置とし、

前記画像表示部は、前記表示面が前記ユーザ定位置を水平方向に囲む内面である、

請求項1に記載の通信端末装置。

【請求項16】

前記画像表示部は、前記ユーザ定位置を水平方向に囲む内面である表示面を有する表示器と、前記表示器の底辺に近接して配置され上面に表示面を有する下側表示器を更に有す

10

20

30

40

50

る、および／または、前記表示器の上辺に近接して配置され下面に表示面を有する上側表示器を更に有する、

請求項 15 に記載の通信端末装置。

【請求項 17】

前記ユーザ空間における所定の位置をユーザ定位置とし、

前記画像表示部は、前記表示面の少なくとも一部の外縁近傍が前記ユーザ定位置に向かって湾曲している、

請求項 1 に記載の通信端末装置。

【請求項 18】

前記画像表示部は、前記表示面の少なくとも一部の外縁近傍が前記ユーザ定位置に向かって湾曲している表示面を有する表示器と、前記表示器の底辺に近接して配置され上面に表示面を有する下側表示器を更に有する、および／または、前記表示器の上辺に近接して配置され下面に表示面を有する上側表示器を更に有する、

請求項 17 に記載の通信端末装置。

【請求項 19】

前記画像表示部は、前記表示面が柱形状の外周面である、

請求項 1 に記載の通信端末装置。

【請求項 20】

前記自装置に対応するオブジェクトと前記他装置に対応するオブジェクトを移動可能に配置することができる画面を表示し、前記画面に対する前記オブジェクトの配置により、前記共有仮想空間に対する前記自装置および前記他装置のユーザ空間の配置の指定を受け付ける配置指定受付部を更に有する、

請求項 1 に記載の通信端末装置。

【請求項 21】

前記音声取得部と前記視覚情報取得部と前記情報伝達部と前記画像表示部と前記音声出力部とが筐体に内蔵され、前記視覚情報取得部の撮像器が前記筐体の所定の位置に配設され、平面の表示面を備える少なくとも 1 つ以上の表示器に接続され前記画像表示部による画像のデータを前記表示器に出力する少なくとも 1 つ以上の外部出力端子が前記筐体に設けられ、

前記画像表示部は、前記表示面のサイズと前記筐体に対する前記表示面の相対位置および相対姿勢の指定を受け、前記指定に基づいて、前期共有仮想空間上にある人物の三次元データを前記表示面に投影することにより、前記自ユーザの視点位置から前記表示面を介して共有仮想空間上の前記人物を見たときのような見え方となる表示画像を生成し、前記外部出力端子から出力する、

請求項 1 に記載の通信端末装置。

【請求項 22】

前記情報伝達部は、前記画像伝達情報に含まれている人物の該画像伝達情報に含まれていない部分の情報を補完する、

請求項 1 に記載の通信端末装置。

【請求項 23】

外部接続端子を更に有し、

前記音声取得部、前記視覚情報取得部、人物捕捉部、前記情報伝達部、前記画像表示部、前記音声出力部の少なくともひとつに対する制御を受け付けるための制御情報を前記外部接続端子から入力する、

請求項 1 に記載の通信端末装置。

【請求項 24】

前記人物捕捉部は、前記自ユーザが所定の動作を行ったことを検知し、

前記音声取得部、前記視覚情報取得部、前記情報伝達部、前記画像表示部、前記音声出力部の少なくともひとつが前記動作に対応する処理を実行する、

請求項 2 に記載の通信端末装置。

10

20

30

40

50

【請求項 25】

前記人物捕捉部は、前記自ユーザが前記共有仮想空間上の物体に手をかざす第1動作と、前記物体に向けた手を握って移動させる第2動作とを検知し、

前記画像表示部は、前記第1動作が検知された後に前記第2動作が検知されたら前記物体を移動させ、前記第1動作が検知されずに前記第2動作が検知されても前記物体を移動させない、

請求項2に記載の通信端末装置。

【請求項 26】

前記画像表示部は、前記画像を前記表示面に表示する表示装置と、前記ユーザが着用する三次元眼鏡装置とを有し、

前記表示装置は、前記視点位置に応じて、視差を有する左眼用画像および右眼用画像を生成し、前記表示面に表示し、

前記三次元眼鏡装置は、前記左眼用画像を前記自ユーザの左眼に見せ、前記右眼用画像を前記自ユーザの右眼に見せる、

請求項1に記載の通信端末装置。

【請求項 27】

前記人物捕捉部は、所望の心理状態の指定を受けると、前記視覚情報における前記自ユーザの顔部分を前記心理状態を表す表情に加工し、

前記情報伝達部は、前記自ユーザの顔部分が加工された視覚情報に基づく画像伝達情報を他装置に伝達する、

請求項2に記載の通信端末装置。

【請求項 28】

通信回線を介して複数で相互接続する通信端末装置による通信方法であって、

前記通信端末装置が、

自装置を用いる自ユーザの声を含む音声を取得し、

所定のユーザ空間上にある人物の三次元データを形成可能にする視覚情報を取得し、

前記ユーザ空間の視覚情報に基づく画像伝達情報と、前記音声に基づく音声伝達情報とを他装置に伝達し、

前記他装置から伝達された画像伝達情報に基づく、前記自装置のユーザ空間と前記他装置のユーザ空間とを配置した共有仮想空間上にある人物の三次元データを、前記共有仮想空間に対して相対位置が固定された表示面に投影するように射影変換をして生成される、前記自ユーザの視点位置から前記表示面を介して共有仮想空間上の前記人物を見たときのような見え方となる表示画像を、前記表示面に表示し、

前記他装置から伝達された音声伝達情報に基づいて音声を出力する、
通信方法。

【請求項 29】

コンピュータを通信回線を介して複数で相互接続する通信端末装置の制御装置として動作させるためのソフトウェアプログラムであって、

自装置を用いる自ユーザの声を含む音声を取得し、

所定のユーザ空間上にある人物の三次元データを形成可能にする視覚情報を取得し、

前記ユーザ空間の視覚情報に基づく画像伝達情報と、前記音声に基づく音声伝達情報とを他装置に伝達し、

前記他装置から伝達された画像伝達情報に基づく、前記自装置のユーザ空間と前記他装置のユーザ空間とを配置した共有仮想空間上にある人物の三次元データを、前記共有仮想空間に対して相対位置が固定された表示面に投影するように射影変換をして生成される、前記自ユーザの視点位置から前記表示面を介して共有仮想空間上の前記人物を見たときのような見え方となる表示画像を、前記表示面に表示し、

前記他装置から伝達された音声伝達情報に基づいて音声を出力する、
手順をコンピュータに実行させるためのソフトウェアプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、遠隔地にいるユーザ同士の良い対話を実現する通信端末装置に関する。

【背景技術】

【0002】

遠隔でのコミュニケーションの手段として、汎用的なパーソナルコンピュータ等で行うウェブ会議、専用の機器を用いるテレビ会議、スマートフォン等で利用可能なビデオ通話がある。いずれも遠隔地にいる相手を主に正面から撮影した映像を平面の画面に表示し、音声で通話を行うというものである。そのため、実際に同じ空間にいて対面で会話をしているような感覚は得にくい。これに対して、仮想空間を共有してコミュニケーションの質を高める様々な技術が提案されている。

10

【0003】

特許文献1には、複数のメンバーが同一の空間を共有する遠隔会議システムが開示されている。特許文献1の遠隔会議システムは、サーバ装置と複数のクライアント装置から構成される。クライアント装置は、ユーザが装着する、加速度センサおよびディスプレイ装置を有するヘッドセットと、赤外線センサと撮像手段と収音手段とを有する3次元形状取得センサと、を有する。サーバ装置は、クライアント装置から送信された深度信号を用いてユーザの動作解析を行い、動作解析に基づくCG（コンピュータグラフィック）を生成し、実物体データを表示し、拡張現実空間を合成し、合成した拡張現実空間の情報をクライアント装置に送信する。クライアント装置は、サーバ装置から受信した拡張現実空間情報に基づいて、ヘッドセットのディスプレイ装置に画像を表示し、スピーカに音声信号を出力する。また、特許文献1には、CGの形状情報に、予め取得しておいた顔のテクスチャを張り付けることも開示されている。

20

【0004】

特許文献2には、仮想空間を複数のユーザで共有し、各ユーザの実空間における位置と仮想空間における位置を対応させる技術が開示されている。特許文献2に開示された仮想空間共有システムは、離れた場所にいる複数のユーザのそれぞれの場所における位置情報と視界情報を検出する手段と、その位置情報および視界情報を送信し、各ユーザ宛てに送信された仮想空間情報をそれぞれ受信するユーザ制御手段と、各ユーザの位置情報および視界情報を受信し、複数のユーザで共有する1つの仮想空間内で各ユーザの位置情報に応じた位置に各ユーザを表す仮想オブジェクトを配置し、各ユーザの位置情報および視界情報に応じた各ユーザの視点からそれぞれ対応する仮想空間を表示する仮想空間情報を生成し、各ユーザ宛てに送信する情報処理手段と、各ユーザごとにそれぞれのユーザ制御手段に受信した仮想空間情報を入力し、各ユーザの視点から見た仮想空間を表示する表示手段とを備える。

30

【0005】

特許文献3には、仮想現実空間共有システムにおいて通信負荷を低減する技術が開示されている。特許文献3に開示された技術は、第1ユーザによって使用される第1端末と第2ユーザによって使用される第2端末との間で仮想現実空間を共有する仮想現実空間共有システムにおいて、前記第1端末に表示するための仮想現実空間画像を生成する画像生成装置が、前記第2端末から仮想現実空間における前記第2ユーザの位置と視線方向とを含む視線情報を取得する視線情報取得部と、前記第2端末からの前記視線情報に基づいて前記第1端末に表示するための仮想現実空間画像を生成する画像生成部と、前記画像生成部によって生成された前記仮想現実空間画像を前記第1端末へ供給する画像供給部と、を備える。また、特許文献3には、仮想現実空間に相手ユーザの代わりにアバター等のオブジェクトを表示し、そのオブジェクトの顔の向きを相手ユーザの顔の向きと一致させることが開示されている。

40

【先行技術文献】

【特許文献】

【0006】

50

【文献】特開 2015-184986 号公報

【文献】特開 2002-149581 号公報

【文献】特開 2017-078891 号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

人と人が対面で行うコミュニケーションによる意思疎通や信頼関係構築などにおいて、お互いの動作、姿勢、表情、顔色、視線などが非言語（non-verbal）の情報伝達手段として重要な役割を果たすと言われている。仮想空間を共有して行うコミュニケーションにおいてもこれら非言語情報の伝達が重要となる。

10

【0008】

しかし、特許文献 1 の技術では、複数のユーザが共有する仮想空間に現れる相手ユーザはアバター等の CG の画像、あるいは予め取得しておいた顔のテクスチャを相手ユーザの形状に張り付けた画像となる。特許文献 2 の技術では、複数のユーザが共有する仮想空間に現れる相手ユーザは仮想オブジェクトとなる。特許文献 3 の技術では、複数のユーザが共有する仮想空間に現れる相手ユーザはアバターとなる。そのため、特許文献 1～3 の技術を用いたコミュニケーションは、非言語の情報伝達が十分とは言えなかった。

【0009】

また、特許文献 1～3 のいずれにおいても、ヘッドマウントディスプレイを頭に装着することは、ユーザにとって不快であると共に、対面での会話と異なる違和感を与えるものである。

20

【0010】

本発明の目的は、遠隔地にいるユーザ同士の良好な対話を実現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】

【0011】

本開示のひとつの態様による通信端末装置は、自装置を用いる自ユーザの声を含む音声を取得する音声取得部と、所定のユーザ空間上にある人物の三次元データを形成可能にする視覚情報を取得する視覚情報取得部と、ユーザ空間の視覚情報に基づいて自ユーザの視点位置を特定する人物捕捉部と、ユーザ空間の視覚情報に基づく画像伝達情報と、音声に基づく音声伝達情報とを他装置に伝達する情報伝達部と、ユーザ空間を固定的に配置した第 1 仮想空間と他装置のユーザ空間が固定的に配置された第 2 仮想空間とを所定の相対位置および相対角度で配置した共有仮想空間に対して相対位置が固定された表示面に、他装置から伝達された画像伝達情報に基づいて、自ユーザの視点位置から見える共有仮想空間の画像を表示する画像表示部と、他装置から伝達された音声伝達情報に基づいて音声出力する音声出力部と、を有する。

30

【発明の効果】

【0012】

本開示のひとつの態様によれば、遠隔地にいるユーザ同士の良好な対話を実現する通信端末装置を提供できる。

40

【図面の簡単な説明】

【0013】

【図 1】実施形態による通信システムのブロック図である。

【図 2】実施形態による通信端末装置のブロック図である。

【図 3】画像を表示する処理について説明するための概念図である。

【図 4】表示面に共有仮想空間の画像が表示された様子を示す図である。

【図 5】表示面に表示される画像の変化の様子を示す図である。

【図 6】実施例 1 による通信端末装置のブロック図である。

【図 7】実施例 1 による通信端末装置の概略平面図である。

【図 8】実施例 1 の通信端末装置の概略斜視図である。

50

【図 9】実施例 1 における画像表示処理のフローチャートである。

【図 10】実施例 1 における配置指定画面の一例を示す図である。

【図 11】複数のユーザが同じ会議室で対面で行う会議が模範的に実現された遠隔会議の様子を示す概念図である。

【図 12】他の構成による遠隔会議の様子を示す概念図である。

【図 13】更に他の構成による遠隔会議の様子を示す概念図である。

【図 14】点群を加工した画面が表示面に表示された様子を示す概念図である。

【図 15】実施例 3 による通信端末装置のブロック図である。

【図 16】実施例 3 による通信端末装置の概略平面図である。

【図 17】実施例 3 の通信端末装置の概略斜視図である。

10

【図 18】実施例 4 による通信端末装置のブロック図である。

【図 19】実施例 4 による通信端末装置の概略平面図である。

【図 20】実施例 4 の通信端末装置の概略斜視図である。

【図 21】実施例 4 における画像表示処理のフローチャートである。

【図 22】実施例 5 による通信端末装置のブロック図である。

【図 23】実施例 5 による通信端末装置の概略平面図である。

【図 24】実施例 5 の通信端末装置の概略斜視図である。

【図 25】実施例 5 における配置指定画面の一例を示す図である。

【図 26】複数のユーザが共有スペースでの対面による対話が模範的に実現された様子
を示す概念図である。

20

【図 27】実施例 6 による通信端末装置のブロック図である。

【図 28】実施例 6 による通信端末装置の概略平面図である。

【図 29】実施例 6 の通信端末装置の概略斜視図である。

【図 30】実施例 6 のユーザから得られる視界を示す概念図である。

【発明を実施するための形態】

【0014】

以下、本発明の実施形態について図面を参照して説明する。

【0015】

図 1 は、本実施形態による通信システムのブロック図である。通信システム 1 は、複数の通信端末装置 10 を備えている。通信システム 1 は、遠隔にいる複数のユーザ 91 があ
たかも同じ空間で対面に対話しているような体験を再現するシステムである。複数の通信
端末装置 10 は、基本的に同様の構成および機能を備え、通信ネットワーク 90 を介して
相互に接続可能である。

30

【0016】

なお、ここでは、ある通信端末装置 10 に着目し、その通信端末装置 10 を自装置と称
し、その自装置を用いるユーザ 91 を自ユーザと称し、自装置と異なる通信端末装置 10
を他装置と称し、他装置を用いるユーザ 91 を他ユーザと称する場合がある。他装置は 1
つ以上であればよく、複数であってもよい。

【0017】

図 2 は、本実施形態による通信端末装置のブロック図である。通信端末装置 10 は、音
声取得部 11 と、視覚情報取得部 12 と、人物捕捉部 13 と、情報伝達部 14 と、画像表
示部 15 と、音声出力部 16 とを有する。

40

【0018】

本実施形態では、通信端末装置 10 を用いているユーザ 91 が存在するであろう実空間
の所定の範囲をユーザ空間とし、そのユーザ空間内でユーザ 91 が動くことを想定してい
る。

【0019】

音声取得部 11 は、ユーザ空間にいるユーザ 91 の声を含む音声を取得する。例えば、
音声取得部 11 は、不図示のマイクロフォンを備え、マイクロフォンにより音声の情報を
取得する。

50

【0020】

視覚情報取得部12は、ユーザ空間にある物体の三次元データを生成可能にする視覚情報を取得する。三次元データは、三次元空間上の三次元オブジェクトを表現するデータであればよく、特に具体的手法は限定されない。ユーザ空間にある物体には、通信端末装置10を用いているユーザ91が含まれる。視覚情報は、例えば画像の情報および深度の情報を含む。また、視覚情報は、例えば30～60fps(frames per second)といった所定のフレームレートで取得される。視覚情報取得部12は、不図示のイメージセンサおよび不図示の深度センサを備え、イメージセンサにより画像をリアルタイムで取得し、深度センサにより深度をリアルタイムで取得してもよい。

【0021】

人物捕捉部13は、視覚情報取得部12で取得されたユーザ空間の画像および／または深度に基づいて、ユーザ91の視点位置を特定する。ここで視点とは眼球がある位置のことである。視点位置は、三次元空間上の位置であり、例えば、三軸の直交座標系の座標で表すことができる。例えば、人物捕捉部13は、画像および／または深度の情報に基づいてユーザ空間内の人物をリアルタイムで検出し、更に、検出された人物の頭部を特定し、その頭部の中央をユーザ91の視点位置とする。なお、視点位置は、ユーザ91の視点の位置あるいはユーザ91の視点とみなすことができる位置であればよく、本実施形態の例に限定されない。顔認識処理または人物認識処理により認識された顔の領域の中央、あるいは認識された両目の中央をユーザ91の視点位置としてもよい。

【0022】

情報伝達部14は、視覚情報取得部12で取得されたユーザ空間の画像および深度に基づく画像伝達情報と、音声取得部11で取得された音声に基づく音声伝達情報とを、他装置に伝達する。

【0023】

画像伝達情報は、他装置の画像表示部15による画像の再生を可能にする情報である。画像伝達情報は、例えば、視覚情報取得部12で取得されたままの画像および深度の情報を含んでいてもよいし、それらを圧縮した情報を含んでいてもよいし、それらの一部が削減された情報を含んでいてもよいし、人物の特徴点のみの画像および深度の情報を含んでいてもよい。特徴点は、ユーザ91の顔および身体の特徴点であり、例えば、腕の関節の位置、眼の位置、口の位置などを示す。これらの特徴点の情報から人物の姿勢、動作、および表情の再現が可能となる。

【0024】

音声伝達情報は、他装置の音声出力部16による音声の出力を可能にする情報である。音声伝達情報は、例えば、音声取得部11で取得されたままの音声の情報を含んでいてもよいし、それらを圧縮した情報を含んでいてもよいし、それらの一部が削減された情報を含んでいてもよいし、それらを加工した情報をふくんでいてもよい。

【0025】

画像表示部15は、自装置のユーザ空間を固定的に配置した第1仮想空間と、他装置のユーザ空間が固定的に配置された第2仮想空間とを所定の相対位置に配置した共有仮想空間を形成し、その仮想空間に対して相対位置が固定された表示面に、他装置から伝達された画像伝達情報に基づいて、自ユーザの視点位置から見える共有仮想空間の画像を表示する。なお、自装置で取得された視覚情報に基づき他装置の表示面に表示する画像を生成する処理を行う物理的実体が存在する場所は特に限定されない。当該処理を、自装置にて行ってもよいし、他装置にて行ってもよいし、自装置と他装置の間でクラウドコンピューティングにより行ってもよい。

【0026】

自ユーザの視点位置が人物捕捉部13によりリアルタイムで更新され、画像表示部15は、リアルタイムで更新される自ユーザの視点位置の情報を用い、リアルタイムの視点位置から見える共有仮想空間の映像を表示面に表示する。その結果、表示面には運動視差を有する映像が映し出される。

10

20

30

40

50

【0027】

画像が表示される表示面は、特に限定されず、ディスプレイ装置の画像を表示する表示面や、プロジェクタの画像が投射される表示面などを含む。表示面は、1つまたは複数の平面で構成されてもよい。表示面が矩形で平面であれば、表示面を、その表示面に含まれる点毎に三次元空間における位置により表すことをしなくても、表示面全体としての位置（例えば中心の位置）、姿勢（例えばロール角、ピッチ角、おおよびヨー角）、おおよびサイズ（例えば縦および横それぞれの長さ）により表すことができる。また、表示面は、1つまたは複数の曲面で構成されてもよい。例えば、円筒形の内面の周方向の所定角度範囲に渡る曲面であれば、円筒の断面の半径と、円筒の高さと、円の中心から見て曲面がひろがる角度範囲とにより、一意に決まる。また、他の例として、球形の内面の所定角度範囲に渡る曲面であれば、球の半径と、球の中心から見た角度範囲（ロール角、ピッチ角、おおよびヨー角）により、一意に決まる。曲面の表示面は、一般に、その表示面に含まれる各点の三次元空間における位置で表すことができる。更に、表示面は、1つまたは複数の曲面と、1つまたは複数の平面とを含んで構成されてもよい。

10

【0028】

このように、自ユーザにとって共有仮想空間上の他ユーザがそこに存在するかのように錯覚させるべく確からしく自然な画像に見えるような処理として、本実施形態では、他装置から伝達された画像および深度の情報に基づいて共有仮想空間の三次元オブジェクトのデータ（以下、三次元データ）を生成し、表示面に表示する画像を生成するとき、三次元データに定義された共有仮想空間における三次元オブジェクトを表示面、すなわち二次元の面に投影させるような射影変換を行う。ただし、同様な効果を得る他の処理手法を用いてもよい。他の処理手法の例として、共有仮想空間を基準座標空間とし、その基準座標空間において三次元データの各点をその各点と視点位置とを結ぶ直線が表示面と交わる点に投影するという手法がある。また、表示面に表示する画像を生成する他の処理方法として、経験則に従った特定の行列や数値の四則演算処理を、画像や画像がもつ三次元パラメータに対して行うことにしてもよい。

20

【0029】

図3は、画像を表示する処理について説明するための概念図である。図3には、第1仮想空間21と、第2仮想空間31と、共有仮想空間41とが上方から見た平面図により示されている。図3には、図に表す都合で、空間が平面で示されているが、高さ方向にも広がっている。

30

【0030】

画像表示部15は、自装置のユーザ空間22を固定的に配置した第1仮想空間21を定義する。図3の例では、視覚情報取得部12が、画像を取得するイメージセンサと深度を取得する深度センサを一体的に構成した撮像器40を有するものとし、ユーザ空間22は撮像器40の画角範囲に広がる空間としている。ただし、ユーザ空間22の大きさおよび形状は特に限定されない。第1仮想空間21の大きさおよび形状は特に限定されない。第1仮想空間21の大きさは無限大であってもよい。自装置のユーザ空間22には、人物捕捉部13により特定された自ユーザ23の視点位置24が存在する。

【0031】

また、画像表示部15は、他装置のユーザ空間32を固定的に配置した第2仮想空間31を定義する。第2仮想空間31の大きさおよび形状も特に限定されない。第2仮想空間31の大きさは無限大であってもよい。図3の例では、第2仮想空間31と第1仮想空間21は大きさおよび形状が等しい例を示している。しかし、それらは必ずしも等しくなくてもよい。他装置のユーザ空間32には他ユーザ33が存在し、他ユーザ33の画像および深度は他装置により取得され、その画像および深度の情報に基づく画像伝達情報が自装置に伝達される。

40

【0032】

画像表示部15は、第1仮想空間21と第2仮想空間31とを所定の相対位置に配置した共有仮想空間41を定義する。ここでは、共有仮想空間41は第1仮想空間21と第2

50

仮想空間 3 1 とを重ね合わせたものとなっている。共有仮想空間 4 1 には自ユーザ 2 3 の視点位置 2 4 と他ユーザ 3 3 とが存在することになる。

【0033】

さらに、画像表示部 1 5 は、共有仮想空間 4 1 に対する相対位置が固定された表示面 4 2 に、他装置から伝達された画像伝達情報に基づいて、自ユーザ 2 3 の視点位置 2 4 から見える共有仮想空間の画像を表示する。

【0034】

図 4 は、表示面に共有仮想空間の画像が表示された様子を示す図である。

【0035】

撮像器 4 0 で取得された画像および／または深度に基づいて特定された自ユーザ 2 3 の視点位置 2 4 から見える共有仮想空間 4 1 の画像が表示面 4 2 に表示されている。自ユーザ 2 3 からは自身の前に他ユーザ 3 3 が居るように見える。

10

【0036】

なお、画像表示部 1 5 は、互いに接続された通信端末装置 1 0 のユーザ空間に実際に存在する物体だけでなく、複数の通信端末装置 1 0 が共通する仮想的な物体を共有仮想空間 4 1 に構成し、表示面にその画像を表示可能としてもよい。例えば、ある通信端末装置 1 0 にてユーザ 9 1 が視覚情報取得部 1 2 の撮像器に、ある実在の物体をかざすと、視覚情報取得部 1 2 がその物体の三次元データを取得し、複数の通信端末装置 1 0 にて、画像表示部 1 5 がその三次元データに基づく仮想的な物体を共有仮想空間 4 1 上に構成することを可能にしてもよい。

20

【0037】

音声出力部 1 6 は、他装置から伝達された音声伝達情報に基づいて音声を出力する。音声出力部 1 6 は、例えば、不図示のスピーカを備え、スピーカにより音声を出力する。その際、音声出力部 1 6 は、ステレオ方式、サラウンド方式、三次元音響方式などにより、他装置から伝達された音声伝達情報に基づく音声を、共有仮想空間 4 1 における他ユーザのユーザ定位置が配置された方向から聞こえる音声として再現してもよい。遠隔地の複数人が仮想空間において互いの実写画像と立体音響によるコミュニケーションを行うことができる。

【0038】

図 5 は、表示面に表示される画像の変化の様子を示す図である。

30

【0039】

上述したように、人物捕捉部 1 3 が自ユーザ 2 3 の視点位置 2 4 を特定し、画像表示部 1 5 は、表示面 4 2 に、他装置から伝達された画像伝達情報に基づいて、自ユーザ 2 3 の視点位置 2 4 から見える共有仮想空間の画像を表示する。したがって、図 5 に示すように、仮に他ユーザ 3 3 が動かなかったとしても、自ユーザ 2 3 の視点位置 2 4 が変化すれば、運動視差により、表示面 4 2 に表示される画像は変化する。例えば、自ユーザ 2 3 は視点位置 2 4 を正面から右方向にずらせば、他ユーザ 3 3 の斜め左の横顔を覗き込むことも可能である。

【0040】

以上説明したように、本実施形態によれば、通信端末装置 1 0 は他の通信端末装置 1 0 と仮想空間を共有し、その仮想空間に相対位置が固定された表示面 4 2 に、その仮想空間を共有している他ユーザ 3 3 の実写の姿を自ユーザ 2 3 の視点位置 2 4 から見えるように表示するので、遠隔地の複数人が共有する仮想空間において互いの実写の画像を見ながら行う良好なコミュニケーションが実現される。

40

【0041】

なお、図 2 に示した音声取得部 1 1、視覚情報取得部 1 2、情報伝達部 1 4、人物捕捉部 1 3、音声出力部 1 6、および画像表示部 1 5 が実行する処理をソフトウェアプログラムにより規定し、メモリおよびプロセッサを有するコンピュータがそのソフトウェアプログラムを実行することにしてもよい。

【0042】

50

以下、より具体的な実施例について説明する。

【実施例 1】

【0043】

実施例 1 では、複数のユーザが同じ会議室で対面で行う会議を模擬的に実現する遠隔会議システムを例示する。実施例 1 の遠隔会議システムは基本的な構成は図 1 ～図 4 に示した実施形態による通信システム 1 と同じである。

【0044】

図 6 は、実施例 1 による通信端末装置のブロック図である。本実施例による通信端末装置 10 は、音声取得部 11 と、視覚情報取得部 12 と、人物捕捉部 13 と、情報伝達部 14 と、画像表示部 15 と、音声出力部 16 と、配置指定受付部 17 とを有する。音声取得部 11、視覚情報取得部 12、人物捕捉部 13、情報伝達部 14、画像表示部 15、および音声出力部 16 は、基本的に、図 2 に示した実施形態のものと同様である。配置指定受付部 17 は、共有仮想空間 41 に対する自装置のユーザ空間 22 および他装置のユーザ空間 32 の配置をユーザ 91 からの指定により受け付け、他装置と共有する。本実施例では、視覚情報取得部 12 が 3 つの撮像器 40A、40B、40C を備え、画像表示部 15 が 3 つの表示器 42A、42B、42C を備えている。本実施例では、視覚情報取得部 12 および画像表示部 15 は、配置指定受付部 17 により受け付けられた配置に基づき、後述する動作を行う。

【0045】

図 7 は、実施例 1 による通信端末装置の概略平面図である。図 8 は、実施例 1 の通信端末装置の概略斜視図である。

【0046】

図 7 に示すように、実施例 1 では、ユーザ空間 22 における所定の位置をユーザ定位置 51 とし、そのユーザ定位置 51 から所定の方向をユーザ定方向 52 とする。例えば、ユーザ定位置 51 は通信端末装置 10 を用いているユーザ 91 が主に存在するであろう位置である。例えば、ユーザ定方向 52 は、通信端末装置 10 を用いているユーザ 91 が主に向いているであろう方向である。

【0047】

図 8 に示すように、表示器 42A ～ 42C はいずれも矩形で平面の表示面を有し、表示面をユーザ定位置 51 に向けて配置された表示装置である。表示器 42B は、ユーザ定位置 51 から見てユーザ定方向 52 に配置される。表示器 42C は、ユーザ定位置 51 から見てユーザ定方向 52 の右側に配置される。表示器 42A は、ユーザ定位置 51 から見てユーザ定方向 52 の左側に配置される。表示器 42A ～ 42C はいずれも表示面をユーザ定位置 51 に向けて配置される。表示器 42C は、表示器 42B と互いの表示面同士で 90° 以上 180° 未満の角度が好ましく、本実施例では 90° で、表示器 42B の右辺に近接しているが、ユーザが快適に本装置を利用できる限りにおいて 90° 未満の角度で配置されていてもよい。表示器 42A は、表示器 42B と互いの表示面同士で 90° 以上 180° 未満の角度が好ましく、本実施例では 90° で、表示器 42B の左辺と近接しているが、ユーザが快適に本装置を利用できる限りにおいて 90° 未満の角度で配置されていてもよい。また、表示器 42A ～ 42C はいずれも水平面に対して 90° で立たせた例を示しているが、水平面に対して 0° ～ 180° の間で角度をつけてもよい。

【0048】

撮像器 40A ～ 40C は、それぞれがユーザ定位置 51 との相対位置が固定され、ユーザ定方向 52 に対する相対方向が固定されて、互いに隔離して、撮像方向をユーザ定位置 51 に向けユーザ空間 22 の少なくとも一部が撮像されるように配置されている。好ましくは、撮像器 40B は、表示器 42B の表示面の右辺の中央近傍に配置される。撮像器 40A は、表示器 42A の表示面の右辺、あるいは表示器 42A の表示面の左辺に配置される。撮像器 40C は、表示器 42C の表示面の右辺、あるいは表示器 42C の表示面の左辺に配置される。本実施例では、図 8 に示すように、撮像器 40A は表示器 42A の表示面の左辺に配置され、撮像器 40C は表示器 42C の表示面の右辺に配置されている。

【0049】

図9は、実施例1における画像表示処理のフローチャートである。画像表示処理は、自装置が他装置との相互通信による遠隔会議を実現する処理であり、主に情報伝達部14、画像表示部15、および配置指定受付部17により実行される。

【0050】

まず、ステップS101にて、配置指定受付部17は、共有仮想空間41に対する自装置および他装置の配置の指定を受け付ける。自装置の表示器42A～42Cおよび撮像器40A～40Cに対してユーザ空間22およびユーザ定位置51が固定的に定義され、他装置の表示器42A～42Cおよび撮像器40A～40Cに対してユーザ空間32および他装置のユーザ定位置が固定的に定義されているので、共有仮想空間41に対する自装置および他装置の配置を受け付けることにより、共有仮想空間41に対するユーザ空間22、32と、ユーザ定位置51と、他装置のユーザ定位置との配置を決定できる。

10

【0051】

このとき、配置指定受付部17は、例えば、自装置に対応するオブジェクトと他装置に対応するオブジェクトを移動可能に配置できる共有空間を上方から見た平面図による配置指定画面を表示器42Bの表示面に表示し、平面図に対して指定されたオブジェクトの位置に基づいて、共有仮想空間41に対するユーザ空間22、32の配置の指定を受け付けることにしてもよい。

【0052】

図10は、実施例1における配置指定画面の一例を示す図である。配置指定画面60には会議室領域61と決定ボタン63が含まれている。会議室領域61は、仮想的な会議室に相当する共有仮想空間41を模した平面図である。会議室領域61には、会議に参加する各メンバーが利用する通信端末装置10に相当するオブジェクトである机アイコン62が配置されている。机アイコン62は、例えばマウス操作あるいは表示器42Bへのタッチ操作や後述するジェスチャ操作により移動させたり固定させたりできる。会議室領域61の形状および大きさを変更できるようにしてもよい。ここでは一例として会議室領域61が共有仮想空間41を模したものであるとしているので、会議室領域61の形状および大きさに応じて共有仮想空間41の形状および大きさを決定してもよい。決定ボタン63が押下されると、そのときの会議室領域61に対する机アイコン62の配置に基づいて、共有仮想空間41における各通信端末装置10のユーザ空間22、32、ユーザ定位置51、および他装置のユーザ定位置の配置が決定される。

20

30

【0053】

複数人が向かい合う会議の場面がある。交渉などで複数人対複数人で向かい合う会議の典型的な形態がある。1人対1人、1人対複数人など様々な配置が可能であるが、図10の例では、6人のユーザが3人対3人対向する配置が例示されている。また、ここでは、X社に所属するAさん、Bさん、Cさんと、Y社に所属するDさん、Eさん、Fさんが向かい合って配置されているが、この所属と配置の関係は例示である。

【0054】

また、ここでは、配置指定受付部17は、ユーザ91が机アイコン62を自由に移動させて配置を指定する例を示したが、他の例も可能である。例えば、配置指定受付部17が配置の指定を補助してもよいし、配置指定受付部17が配置を自動で設定あるいは提案してもよい。

40

【0055】

配置指定受付部17が配置の指定を補助する場合、典型的な複数の配置構成のパターンを予め用意しておき、ユーザ91にいずれかを選択させることにしてもよい。配置指定受付部17は、例えば、最初に会議に参加するユーザの人数が入力されると、その人数に対応するパターンをユーザに提示し、いずれかを選択させることにしてもよい。

【0056】

また、配置指定受付部17が配置を自動で設定あるいは提案する場合、例えば、この会議が何らかの発表を行う場である場合に、メインプレゼンターは、会議参加者全員からよ

50

く見える位置（図10の例では、X社Bさん、Y社Eさんの位置）に配置されるようにしてもよい。また、例えば、会議での発言量の多寡でメインプレゼンターを特定し、配置を設定することにしてもよい。

【0057】

本実施例では、互いに平行な2つの直線のそれぞれの上に複数の通信端末装置10のユーザ定位置51を配置し、各直線上の通信端末装置10のユーザ定方向52を他方の直線に垂直に向かう方向とすることで実現できる。なお、ここでユーザ定位置51の直線上への配置は厳密な位置決めを意味するものではなく、またユーザ定方向52の直線への垂直な方向づけは厳密な方向決めを意味するものではない。どちらも複数の通信端末装置10のユーザが会議において机を横に並べている状態が模擬できる程度でよい。

10

【0058】

ステップS102にて、情報伝達部14は、共有仮想空間における自装置の撮像器40A～40Cの撮像方向と、他装置のユーザ定位置と自装置のユーザ定位置とを結ぶ直線との相対角度に基づいて、他装置のユーザ定位置の所定範囲内（例えばユーザ空間）から自装置のユーザ定位置の所定範囲内（例えばユーザ空間）にある物体を見たときに見える部分の画像が取得されるように撮像器を選択し、選択した撮像器で取得される画像および深度の情報を含む画像伝達情報をその他装置に送信する。

【0059】

情報伝達部14は、共有仮想空間41において自装置のユーザ定位置51からユーザ定方向52の右側に他装置のユーザ定位置が配置される場合には、右側にある撮像器40Cによる情報を含み、左側にある撮像器40Aによる情報を含まない画像伝達情報を他装置に送信すればよい。また、情報伝達部14は、共有仮想空間41において自装置のユーザ定位置51からユーザ定方向52の左側に他装置のユーザ定位置が配置される場合には、左側にある撮像器40Aによる情報を含み右側にある撮像器40Cによる情報を含まない画像伝達情報を他装置に送信すればよい。

20

【0060】

例えば、自ユーザ23がX社のAさんであり、他ユーザ33がY社のFさんであるという組合せに着目すると、自装置の左側にある撮像器40Aで取得されるX社のAさんの画像および深度のデータがなくても、他装置においてY社のFさんの視点位置から見える範囲のX社のAさんの画像を生成することができる。そのため、X社のAさんが用いる自装置の情報伝達部14は、撮像器14Bと撮像器14Cで取得される画像および深度のデータに基づく画像伝達情報を、Y社のFさんが用いる他装置に伝達する。また、自ユーザ23がX社のAさんであり、他ユーザ33がX社のBさんであるという組合せに着目すると、自装置の左側にある撮像器40Aで取得されるX社のAさんの画像および深度のデータがなくても、他装置においてX社のBさんの視点位置から見える範囲のX社のAさんの画像を生成することができる。X社のAさんが用いる自装置の情報伝達部14は、撮像器14Bと撮像器14Cで取得される画像および深度のデータに基づく画像伝達情報を、X社のBさんが用いる他装置に伝達する。

30

【0061】

なお、ここでは、情報伝達部14は、複数の撮像器40A～40Cのうち、その撮像器の撮像方向と、他装置のユーザ定位置と自装置のユーザ定位置を結ぶ直線との相対角度に基づいて撮像器を選択する処理を例示したが、他の例も可能である。例えば、情報伝達部14は、複数の撮像器40A～40Cのうち、その撮像器の撮像方向と、他装置のユーザ定位置と自装置のユーザ定位置を結ぶ直線との相対角度に加え、他装置の表示面の位置、角度、および大きさに基づいて、撮像器を選択することにしてもよい。その場合、情報伝達部14は、他装置のユーザ定位置の所定範囲内（例えばユーザ空間）から他装置の表示面を介して自装置のユーザ定位置の所定範囲内（例えばユーザ空間）にある物体を見たときに見える部分の画像が取得されるように撮像器を選択してもよい。なお、ここでは、自ユーザの情報伝達部14が通信ネットワーク90へ送る画像伝達情報をあらかじめ選択してから送る例を示したが、この方法に限らない。情報伝達部14は一度すべての撮像器か

40

50

ら得られた画像伝達情報を通信ネットワーク 90 上のサーバに送信した後に、このサーバ側で画像伝達情報を取捨選択する処理を行って他ユーザに送る方式をとってもよい。また、ステップ S 102 における送信データの選択は必ず行うものではなく、データ通信帯域が十分確保でき、ネットワーク上のデータ量の削減が不要な環境下においてはこのステップをスキップしてもよい。

【0062】

次に、ステップ S 103 にて、画像表示部 15 は、表示器 42A、42B、42C の表示面に、他装置から伝達された画像伝達情報に基づいて、自ユーザ 23 の視点位置 24 から見える共有仮想空間 41 の画像を表示する。共有仮想空間 41 には各ユーザのユーザ空間 22、32 が固定的に配置されているので、自ユーザ 23 の視点位置 24 からは、その配置に応じた位置にいる他ユーザが見える。

10

【0063】

図 11 は、複数のユーザが同じ会議室で対面で行う会議が模擬的に実現された遠隔会議の様子を示す概念図である。図 11 には、図 10 に示した配置が模擬的に実現された例が示されている。画像表示部 15 は、互いに平行な第 1 直線と第 2 直線のそれぞれの上に自装置および他装置のうち少なくとも 1 つの装置のユーザ定位置 51 を配置し、第 1 直線上にユーザ定位置が配置された装置のユーザ定方向 52 を第 1 直線に対して垂直に第 2 直線に向かう方向とし、第 2 直線上にユーザ定位置 51 が配置された装置のユーザ定方向 52 を第 2 直線に対して垂直に第 1 直線に向かう方向とするように、共有仮想空間を形成すればよい。

20

【0064】

図 11 における手前側には、X 社の A さんと B さんと C さんが直線上に横に並び奥を向いている。奥側には、Y 社の D さんと E さんと F さんが直線上に横に並び手前を向いている。A さん、B さん、C さん、D さん、E さん、F さんは、実際にはそれぞれ異なる場所で通信端末装置 10 を用いている。それぞれのユーザの通信端末装置 10 の表示器 42A ~ 42C には、図 11 に示した配置で、それぞれのユーザの視点位置から見える共有仮想空間 41 の画像が表示される。それにより、全てのユーザは、図 11 に示したような配置で会議室に机を並べた 6 人で会議しているときのそれぞれの座席からの視界が得られる。

【0065】

次に、ステップ S 104 にて、配置指定受付部 17 は、共有仮想空間 41 に対する自装置および他装置の配置を変更する要求があるか否か判定する。配置の変更の要求があれば、配置指定受付部 17 は、ステップ S 101 に戻り配置の指定を受け付ける。

30

【0066】

自ユーザあるいは他ユーザは、図 10 に例示した配置指定画面 60 に対する操作により、会議室に相当する共有仮想空間 41 の形状および大きさと、会議室における各ユーザの配置に相当する共有仮想空間 41 に対するユーザ空間の配置を変更することができる。

【0067】

図 12 は、他の構成による遠隔会議の様子を示す概念図である。図 12 の例では、実際にはそれぞれ異なる場所で通信端末装置 10 を用いている X 社の A さん、B さん、C さんが、共有仮想空間 41 において円卓を囲むように対面している。画像表示部 15 は、所定の円上もしくは多角形状に自装置および他装置のユーザ定位置 51 を配置し、自装置および他装置のユーザ定方向を円もしくは多角形の中心に向かう方向とするように、共有仮想空間 41 を形成すればよい。

40

【0068】

図 13 は、更に他の構成による遠隔会議の様子を示す概念図である。図 13 の例では、実際にはそれぞれ異なる場所で通信端末装置 10 を用いている X 社の A さん、B さん、C さんが、共有仮想空間 41 において、互いに横に並んで、画像の共有が可能な画像画面 65 を向いている。画像表示部 15 は、所定の直線上に自装置および他装置のユーザ定位置 51 を配置し、自装置および他装置のユーザ定方向を直線に対して垂直な同一方向とするように、共有仮想空間 41 を形成すればよい。複数のユーザが横並びになって、一緒に仮

50

想空間上の巨大スクリーンに投影された映像を見ることができる。例えば、通信端末装置 10 に接続したパーソナルコンピュータ上の映像や、撮影現場で取得された映像をリアルタイムでストリーミング再生された映像などである。また、共有仮想空間をユーザ全員が内側に包含されるような広さの球体として定義し、その球体の内側にストリーミング再生中もしくは撮影済みの 360° 映像を貼り付けることで空間全体を共有しながら一緒に見るという利用が可能である。なお、図 13 の画像画面 65 は、複数の通信端末装置 10 が共通する仮想的な物体として共有仮想空間 41 に構成し、表示面にその画像を表示したものである。図 13 の例では、画像画面 65 には、C さんのパーソナルコンピュータ 68 に表示されている画像が表示されている。

【0069】

ステップ S104 にて配置の変更の要求がなければ、ステップ S105 にて、画像表示部 15 は、遠隔会議の終了の要求があるか否か判定する。遠隔会議の終了の要求がなければ、画像表示部 15 はステップ S103 に戻る。遠隔会議の終了の要求があれば、通信端末装置 10 は遠隔会議を終了する。

【0070】

本実施例によれば、自装置と他装置の仮想空間を重ね合わせて共有し、その仮想空間に相対位置が固定された表示面に、その仮想空間にいる他ユーザの実写の姿を自ユーザの視点位置から見えるように表示するので、遠隔地の複数人が共有する仮想空間において互いの実写の画像を見ながら行う良好なコミュニケーションが実現される。

【0071】

また、本実施例によれば、3つの撮像器 40A～40C のうち、撮像方向と、他装置のユーザ位置と自装置のユーザ位置 51 を結ぶ直線との相対角度に基づいて選択した撮像器のみについて、その撮像器で取得された視覚情報に基づく画像伝達情報を、その他装置に送信することもできるので、自装置から他装置に送信する画像伝達情報が削減され、ネットワーク上のデータ量を低減することもできる。

【0072】

また、本実施例によれば、図 7、図 8 に示したように、表示面がユーザ位置 51 を囲むように配置されるので、共有仮想空間 41 を広い角度範囲で表示する構成をコンパクトな占有スペースにより実現することができる。

【0073】

また、本実施例によれば、画面上で平面図の中で会議室の机を動かす感覚で自装置と他装置の位置関係を容易に設定することができる。

【0074】

また、本実施例において、画像表示部 15 は、他装置から伝達された画像および深度の情報に基づいて、第 2 仮想空間の物体を点群もしくはポリゴンメッシュで表す共有仮想空間 41 の三次元データを生成し、生成した三次元データに基づいて、自ユーザ 23 の視点位置 24 から見える共有仮想空間 41 の画像を表示面に表示することにしてもよい。その際、画像表示部 15 は、点群の点を加工可能にしてもよい。例えば、画像表示部 15 は、三次元データにおける自ユーザの手前側にある三次元オブジェクトが表現された点の表示サイズを小さくすることにより、その三次元オブジェクトを半透明に表示し、後ろにある三次元オブジェクトも見えるようにしてもよい。点群により表示される人物を半透明に加工することにより、人物の顔あるいは体による非言語情報を残しつつ人物の背後の情報も伝達することができる。例えば、文字を書いている手を半透明にし、書いているときに手で隠れている文字を見えるようにすることで、手の動きによる非言語情報と書かれた文字の情報を同時に伝えることができる。また、例えば、文字を書いている人物を半透明にし、書いているとき人物の姿に隠れている文字を見えるようにすることで、人物の動きによる非言語情報と書かれた文字の情報を同時に伝えることができる。図 14 は、点群を加工した画面が表示面に表示された様子を示す概念図である。図 14 には、表示器 42A、42B、42C の表示面に、点群を加工して半透明に表示された他ユーザ 33 の背景にある仮想空間上のホワイトボード 35 の文字が見えている様子が示されている。ここでは、点

10

20

30

40

50

群の点の表示サイズを変化させる加工を例示したが、他の様々な加工も可能である。例えば、点群を構成する各点の色を変えたり、表示する点をまびいたり、各点を特定の形状をもった立体オブジェクトに置換するなどしたうえで描画させる加工も可能である。こうした加工により、例えば、外殻形状だけを維持した単色の人物に変換し、化粧をしていない女性ユーザの使用や、表情や着衣物を特定されない目的での使用にあわせた人物表現が可能となる。また、伝達するデータ量の削減および実写画像によるリアルな表現を避けるために人体の顔以外の部分について、取得された画像および／または深度情報から計算によって求められた骨格を表す情報に対し、仮想的な3DCGオブジェクトを用いて肉付けすることで、人物の姿勢や動作を表現するような加工も可能である。あるいは、顔以外、もしくは顔と片手以外は描画させないなど、骨格を表す情報をもとに表示するユーザの体の部位を選択的に表示することも可能である。加えて、あるユーザが注目されるように、ユーザの周囲に光をまとわせて目立たせたり、一方で、あるユーザ以外が注目されないように、各点を視点位置に応じた背景色に近い色で適切に着色することで背景に溶け込むようにさせて目立たせなくする加工、いわゆる光学迷彩を実装することも可能である。

【0075】

また、本実施例では、人物捕捉部13が常時継続的にユーザ91の視点位置24を特定し、画像表示部15が、その視点位置24を追跡して、その視点位置24から見える共有仮想空間41の画像を表示器42A～42Cの表示面に表示する例を示した。この構成は、ユーザ91が動いたときに運動視差によりユーザ91に視覚的に臨場感を与えるものである。しかし、会議などの場面では、ユーザ91の視点位置24の動きが大きいことがある。そのような場合を考慮すると、ユーザ91の視点位置24を常時追跡しない構成も可能である。

【0076】

他の例として、画像表示部15は、人物捕捉部13でユーザ91の視点位置24を一旦特定したら、視点位置24を固定してそこから見える共有仮想空間41の画像を表示器42A～42Cの表示面に表示することにしてもよい。例えば、事前のセンシング動作において、ユーザ91に自然な姿勢を採らせ、その間に、人物捕捉部13が視点位置24を特定することにしてもよい。あるいは、人物捕捉部13が一定時間の間に視点位置24の特定を複数回実行し、その平均値を平均視点位置とし、画像表示部15は、平均視点位置から見える共有仮想空間41の画像を表示器42A～42Cの表示面に表示することにしてもよい。あるいは、人物捕捉部13は一定時間の間はユーザ91の視点位置24を継続的に算出し、画像表示部15が、その視点位置24を追跡して、その視点位置24から見える共有仮想空間41の画像を表示器42A～42Cの表示面に表示し、その間に、人物捕捉部13が平均視点位置を算出し、画像表示部15が人物捕捉部13により特定される視点位置24から平均視点位置に所定時間をかけて収束させることにしてもよい。

【0077】

さらに他の例として、画像表示部15は、予め基準視点位置を定めておき、基準指定位置から見える共有仮想空間41の画像を表示器42A～42Cの表示面に表示することにしてもよい。その場合、人物捕捉部13がユーザ91の視点位置24を特定する処理は不要となる。基準視点位置として、例えば、標準的な体格の人物が通信端末装置10を用いるとき視点位置を用いてもよい。あるいは、通信端末装置10を用いた複数のユーザの視点位置を測定し、その平均値を基準指定位置として用いることにしてもよい。このような計算や定義により固定された視点位置は、各ユーザが通信ネットワーク90に接続した際に通信ネットワーク90上のサーバに保存された情報を読み込むこと、もしくは、自ユーザの通信端末装置に保存された情報を読み込むことで、以前に本装置を使用した際に決定した視点位置を再度使うことにしてもよい。

【0078】

なお、本実施例の通信端末装置10は、例えば、1人用の会議ブースを構成する。この会議ブースは同じ部屋に複数配置される可能性がある。また、在宅ワークのために会議ブースが個人の住宅に配置される可能性もある。そのような場合には、ユーザ91の声が周

囲に漏れないように騒音対策が採られることが好ましい。そこで、音声出力部 16 は、イヤホン、ネックスピーカーなどにより音声を出力できるものであってもよい。また、音声出力部 16 は、ユーザ定位置 51 に向かう指向性を有するスピーカーを有するものであってもよい。さらに、通信端末装置 10 は、消音スピーカーにより周囲に漏れるユーザ 91 の音声を低減するものであってもよい。さらに、音声出力部 16 からの出力を完全に切り、他ユーザからの声は音声出力部 16 からの出力の代わりに、音声認識技術により文字化されて画像表示部 15 に出力されてもよい。

【0079】

また、本実施例では、画像表示部 15 が 3 つの表示器 42A~42C を備える例を示したが、他の構成も可能である。例えば、画像表示部 15 は、表示器 42A~42C に加え、表示器 42A、42B、42C の底辺に近接して配置され矩形の上面に表示面を有する表示器をさらに備えてもよい。さらに、画像表示部 15 は、表示器 42A、42B、42C の天井方向の辺に近接して配置され矩形の下面に表示面を有する表示器を備えてもよい。そのいずれの場合も、画像表示部 15 は、リアルタイムで更新される自ユーザの視点位置の情報をを用い、リアルタイムの視点位置から見える共有仮想空間の映像を 4 つの表示器の表示面に表示する。図 13 に示したように各装置のユーザ定位置 51 およびユーザ定方向 52 を配置して複数のユーザが一緒にスクリーン映像を見るという利用形態において、共有仮想空間の映像を下面や上面まで視野が広がる視野でみるができる。また、通信端末装置 10 はアプリケーションソフトウェアを実行可能なコンピュータの機能を持ち、例えば、上面に表示面を有する表示器の表示面にアプリケーションソフトウェアの画面を表示してもよい。その場合に、その表示器をタッチパネル式ディスプレイとし、ユーザ 91 によるアプリケーションソフトウェアの操作を受け付けてもよい。なお、ここで記載した上面に表示面を有する表示器および下面に表示面を有する表示器の配置は、水平面に対して 0° から 180° までの角度で配置されていてもよい。

【0080】

また、本実施例では、撮像器 40A が表示器 42A の表示面の上辺あるいは左辺に配置され、撮像器 40C が表示器 42C の表示面の上辺あるいは表示面の右辺に配置される例を示した。しかし、本構成に限定されることはない。他の例として、撮像器 40A が表示器 42A の表示面の上辺と表示器 42B の表示面の上辺が互いに隣接する位置に配置され、撮像器 40C が表示器 42C の表示面の上辺と表示器 42B の表示面の上辺が互いに隣接する位置に配置されてもよいし、それら表示器の辺の直上や近接した位置でなくともよい。撮像器 40A~40C を配置する目的は自ユーザ 23 のイメージ画像および深度情報を自ユーザの全体像を取得できるように広く取得するため、もしくは、自ユーザの顔の表情や掌など特定部位をより高解像度で取得するためであり、その目的に合わせて、撮像器 40A~40C はそれぞれユーザ定位置 51 に対し遠ざけるように配置されていてもよいし、近づけるように配置されていてもよい。また、各撮像器と各表示器とは直接もしくは接続具などを介して間接的に接続されていなくとも、それらの相対位置や相対角度が固定される方法であれば、例えば、画像表示部 15 が固定的に配置されている室内の床に立たせた固定金具や天井からつるされた固定金具などに各表示器が固定されていてもよい。

【0081】

また、本実施例の通信端末装置 10 は、外部接続端子を有し、その外部接続端子にパーソナルコンピュータを接続し、パーソナルコンピュータから制御できてもよい。通信端末装置 10 は、音声取得部 11、視覚情報取得部 12、人物捕捉部 13、情報伝達部 14、画像表示部 15、音声出力部 16、および配置指定受付部 17 に関する制御を受け付けるための制御用データを外部接続端子から出力し、それに対する制御情報を外部接続端子から入力する。例えば、配置指定受付部 17 は、図 10 に示した配置指定画面 60 のデータを外部接続端子を介してパーソナルコンピュータに出力し、机アイコン 62 を移動させるなどの制御情報をパーソナルコンピュータから外部接続端子を介して取得し、取得した制御情報に基づいて共有仮想空間 41 に対するユーザ空間 22、32 の配置の指定を受け付けてもよい。また、画像表示部 15 は、パーソナルコンピュータから外部接続端子を介し

10

20

30

40

50

て取得される制御情報に基づいて、図 13 に例示した画像画面 65 を共有仮想空間 41 に構成し、表示面にその画像を表示することにしてもよい。外部接続端子は、一例として、有線 LAN 端子や USB 端子である。制御情報はパーソナルコンピュータと通信端末装置 10 との間で直接やりとりされる、もしくは、通信ネットワーク 90 上のサーバを介してやりとりされる。このとき物理的な接続端子でなくとも無線 LAN や Bluetooth 等による無線接続であってもよい。またほかの例として HDMI (High-Definition Multimedia Interface) (HDMI は登録商標) の通信規格に従う端子である。制御情報は、HDMI キャプチャ情報の中に挿入して伝送される。また、いずれの場合においても、パーソナルコンピュータから通信端末装置 10 への一方的な制御信号の送信、すなわち通信端末装置 10 からパーソナルコンピュータへの制御データが送られない状況での制御であってもよい。

10

【0082】

また、本実施例の通信端末装置 10 は、ユーザ 91 がジェスチャにより制御できてもよい。通信端末装置 10 は、手を用いた所定の動作（以下、制御動作という）と、音声取得部 11、視覚情報取得部 12、情報伝達部 14、画像表示部 15、音声出力部 16、および配置指定受付部 17 に関する所定の制御とを予め対応付けたジェスチャ設定情報を格納しておく。人物捕捉部 13 が捕捉した人物の動作を監視し、人物が制御動作を行ったことを検知すると、ジェスチャ設定情報にて制御動作に対応付けられている制御を、音声取得部 11、視覚情報取得部 12、情報伝達部 14、画像表示部 15、音声出力部 16、または配置指定受付部 17 に指示する。

20

一例として、複数の通信端末装置 10 が共有する仮想的な物体として共有仮想空間 41 に構成し、表示面にその画像を表示した物体を移動させる操作について述べる。ジェスチャ設定情報には、物体に手をかざすという第 1 制御動作とその物体を選択された状態にするという第 2 制御とが対応づけられ、物体に向けて握った手を移動するという第 2 制御動作とその物体を共有仮想空間 41 内で移動させるという第 2 制御とが対応付けられている。視覚情報取得部 12 で取得される視覚情報に基づいて人物捕捉部 13 が第 1 制御動作を検知する。人物捕捉部 13 で第 1 制御動作が検知されると、画像表示部 15 は物体を選択された状態にする。このとき、人物捕捉部 13 は、共有仮想空間 41 上での視点位置 24 と開いた手の中央位置とを結ぶ直線上あるいは、肘や肩関節など体の特定部位の関節位置と手の中央位置とを結ぶ直線上、あるいはそのどちらかの直線の近傍にある物体をユーザ 41 が選択した物体であると判断し、画像表示部 15 は、その物体を選択された状態にする。例えば、物体を例えば色、光、影などによる強調表示することにより、その物体が選択されたことを表してもよい。第 1 制御動作が検知されたことあるいは物体が選択状態になったことは、情報伝達部 14 から他装置に通知されてもよい。その状態で、人物捕捉部 13 が第 2 制御動作を検知すると、画像表示部 15 は共有仮想空間 41 上でその物体を移動させる。このとき、画像表示部 15 は、第 1 制御動作が検知された後に第 2 制御動作が検知されたら物体を移動させるが、第 1 制御動作が検知されずに第 2 制御動作が検知されたときには物体を移動させない。共有仮想空間 41 で物体が移動したことは、画像伝達情報により他装置にも伝わり、他装置の表示においても物体が移動する。なお、ここでは仮想的な物体の移動を例示したが、ジェスチャにより移動する物体は特に限定されない。例えば、図 13 に示した画像画面 65 を移動させたり、サイズ、位置、角度を変化させたりするものとしてもよい。

30

40

【0083】

また、本実施例では、通信端末装置 10 は、1 つの視点位置 24 から見た共有仮想空間 41 の画像を表示面に表示するものとしたが、この例に限定されることはない。変形例として、通信端末装置 10 は、左右の眼の視差を与えた共有仮想空間 41 の 2 つの画像（以下、それぞれを左眼用画像、右眼用画像という）からなる画像を表示することにしてもよい。画像表示部 15 は、画像を表示面に表示する表示装置と、ユーザ 91 が着用する三次元眼鏡装置とを有する。画像表示部 15 は、左眼用画像および右眼用画像を生成し、表示面に表示する。三次元眼鏡装置は、左眼用画像をユーザ 91 の左眼に見せ、右眼用画像を

50

ユーザ 9 1 の右眼に見せる。左右の眼に視差を持たせることで、画像に立体感を与え、より高い臨場感および没入感をユーザ 9 1 に与えることができる。なお、左右の眼に視差を有する画像を見せる手法は特に限定されない。例えば、アナグリフ式であってもよいし、偏光式であってもよいし、液晶シャッター式であってもよい。なお、一般に三次元眼鏡はヘッドマウントディスプレイのように不安感や不快感を与えることはない。

【0084】

なお、本実施例では、自装置と他装置とで全く同じ共有仮想空間 4 1 を共有する例を示したが、これに限定されることはない。他の例として、共有仮想空間 4 1 に対する自装置のユーザ空間 2 2 および他装置のユーザ空間 3 2 の配置を自装置と他装置とで異ならせてもよい。例えば、各装置の表示器の配置やサイズを考慮してユーザ 9 1 にとって違和感ない範囲で、自ユーザから他ユーザの表情や動作が見えやすいように、それぞれに配置を調整してもよい。自装置と他装置とで異なる配置の共有仮想空間 4 1 を用いる場合には、各装置が用いる配置の情報すべての装置で共有し、それぞれの装置が他のそれぞれの装置に視覚情報を送信すべき撮像器を選択することとしてもよい。あるいは、各装置が他の各装置にどの方向から撮像された視覚情報が欲しいかを指定し、各装置は各装置へ当該装置が指定した方向から撮像した視覚情報を送ることにしてもよい。

【0085】

また、本実施例では、通信端末装置 1 0 が、顔部分についてはリアルタイムの実写画像を表示する例を示したが、この例に限定されることはない。他の例として、送信側の装置からはユーザ 9 1 の心理状態を受信側の装置に通知し、受信側の装置では通知された心理状態に基づいて表示するユーザの表情を加工することにしてもよい。自装置の情報伝達部 1 4 は、自装置の人物捕捉部 1 3 が自ユーザの表情から自ユーザの心理状態を推定する。例えば、自ユーザが、喜んでいる状態、怒っている状態、悲しんでいる状態、楽しんでいる状態、会話内容に合意している状態、合意していない状態、理解できていない状態のいずれであるかを推定する。自装置の情報伝達部 1 4 は、人物捕捉部 1 3 で推定された心理状態の情報を、顔部分の画像の代わりに、画像伝達情報として受信側の装置に送信する。受信側の装置では、通知された心理状態の情報に基づいて、送信側のユーザの顔部分の画像を生成し、表示する。これによれば、相手の心理という可視化しづらい定性的な情報を可視化でき、より戦略的に有効なコミュニケーションを実現するための補助となりうる。また、装置間で伝達するデータ量を削減することができる。

【0086】

また、さらに他の例として、送信側の装置がユーザ 9 1 の所望により、ユーザ 9 1 の視覚情報における顔の表情をユーザ 9 1 が所望する心理状態を表す表情に加工可能にしてもよい。自装置の人物捕捉部 1 3 は、ユーザ 9 1 から所望の心理状態の指定を受けると、ユーザ 9 1 の顔の表情を指定された心理状態を表す表情に加工する。情報伝達部 1 4 は、人物捕捉部 1 3 で顔の表情が加工された画像伝達情報を他装置に送信する。ユーザ 9 1 の所望により表情を加工することができるので、表情から伝わる情報をユーザ 9 1 が意図的に利用して有効なコミュニケーションを図ることができる。

【実施例 2】

【0087】

実施例 1 では、通信端末装置 1 0 は、他装置に送信すべき画像を取得する撮像器を選択し、選択した撮像器で取得される画像および深度の情報を含む画像伝達情報をその他装置に送信する。これにより、自装置から他装置へ送信する画像伝達情報のデータ量を削減している。実施例 2 では、更に自装置から他装置へ送信する画像伝達情報のデータ量を削減する例を示す。

【0088】

実施例 2 の通信システム 1 の構成は図 1 に示した実施例 1 のものと基本的に同じ構成である。また、実施例 2 の通信端末装置 1 0 の構成は、図 6 ～図 8 に示した実施例 1 のものと同様である。実施例 2 の通信端末装置 1 0 が実行する画像表示処理は、図 9 に示した実施例 1 のものと基本的に同様である。ただし、実施例 2 では、ステップ S 1 0 2 およびス

10

20

30

40

50

ステップ S 1 0 3 における具体的処理が実施例 1 の処理と異なる。

【 0 0 8 9 】

ステップ S 1 0 2 にて、実施例 2 では、人物捕捉部 1 3 が、視覚情報取得部 1 2 で取得されるユーザ空間 2 2 の画像および／または深度に基づいて、ユーザ空間 2 2 内に自ユーザを検出し、情報伝達部 1 4 が、視覚情報取得部 1 2 で取得されるユーザ空間 2 2 の画像および／または深度の情報から検出された自ユーザの顔部分を表す情報と、自ユーザの骨格を表す情報とを生成し、それらの情報を画像伝達情報として他装置に送信する。ここで骨格を表す情報とは、人物の体勢の特定を可能にする情報であり、例えば、頭、胴、尻、肩、肘、手、膝、踵など身体の特徴的な箇所の位置を示す情報であってもよい。ここでいう人物の体勢は、ポーズや動作などを人間が目で見てもコミュニケーションに役立つ非言語情報が得られる程度の人体の形態である。どの程度まで詳細で正確な体勢の特定を要するかは特に限定されない。

10

【 0 0 9 0 】

ステップ S 1 0 3 にて、実施例 2 では、画像表示部 1 5 は、他装置から受信した顔部分を表す情報に基づく顔画像と、他装置から受信した骨格を表す情報に基づく人物の体勢の画像とを合成した人物を含む、共有仮想空間の 3 次元オブジェクトを、ユーザの視点位置に合わせて適切に 2 次元の画像に変換し表示器 4 2 A ～ 4 2 C の表示面に表示する。骨格を表す情報に基づく人物の体勢の画像は、例えば、骨格を表す情報に対し、仮想的な 3 D C G オブジェクトを用いて肉付けすることで、人物の姿勢や動作を表現するような加工をした画像であってもよい。

20

【 0 0 9 1 】

本実施例によれば、送信側の通信端末装置 1 0 から受信側の通信端末装置 1 0 に撮像された全体の情報を送らなくてよいので、通信端末装置 1 0 間で伝達するデータ量を削減することができる。また、人体の顔以外の部分は仮想的なオブジェクトで表現されるので、実写画像のリアルな表示を避けることが可能となる。

【実施例 3】

【 0 0 9 2 】

実施例 1 では、図 6 ～ 図 8 に示したように、通信端末装置が、それぞれに矩形で平面の表示面を有し表示面をユーザ定位置に向けて配置された 3 つの表示器を有する構成を例示した。これに対して、実施例 3 では、通信端末装置 1 0 が、ユーザ定位置を囲むように湾曲した曲面の表示面を有する 1 つの表示器を有する構成を例示する。

30

【 0 0 9 3 】

図 1 5 は、実施例 3 による通信端末装置のブロック図である。本実施例による通信端末装置 1 0 は、音声取得部 1 1 と、視覚情報取得部 1 2 と、人物捕捉部 1 3 と、情報伝達部 1 4 と、画像表示部 1 5 と、音声出力部 1 6 と、配置指定受付部 1 7 とを有する。音声取得部 1 1、視覚情報取得部 1 2、人物捕捉部 1 3、情報伝達部 1 4、画像表示部 1 5、および音声出力部 1 6 は、基本的に、図 2 に示した実施形態のものと同様である。配置指定受付部 1 7 は、共有仮想空間 4 1 に対する自装置のユーザ空間 2 2 および他装置のユーザ空間 3 2 の配置のユーザによる指定を受け付ける。

【 0 0 9 4 】

本実施例では、実施例 1 と異なり、画像表示部 1 5 が 1 つの表示器 4 2 B を備えている。

40

【 0 0 9 5 】

図 1 6 は、実施例 3 による通信端末装置の概略平面図である。図 1 7 は、実施例 3 の通信端末装置の概略斜視図である。

【 0 0 9 6 】

図 1 6 に示すように、実施例 4 では、実施例 1 と同様に、ユーザ空間 2 2 における所定の位置をユーザ定位置 5 1 とし、そのユーザ定位置 5 1 から所定の方角をユーザ定方向 5 2 とする。例えば、ユーザ定位置 5 1 は通信端末装置 1 0 を用いている自ユーザ 2 3 が存在するであろう位置である。例えば、ユーザ定方向 5 2 は、通信端末装置 1 0 を用いている自ユーザ 2 3 が向いているであろう方向である。

50

【0097】

図17に示すように、表示器42Bはユーザ定位置51を囲むように湾曲した曲面の表示面を有する表示装置である。より具体的には、表示器42Bの表示面は、円筒形の内面の周方向の所定角度範囲（一例として180°）に渡る曲面である。表示器42Bは、曲面のスクリーンに対して複数のプロジェクタで映像を投射する構成でもよいし、有機ELディスプレイを湾曲させて配置したものであってもよい。表示器42Bの表示面の左右方向の中央は、ユーザ定位置51から見てユーザ定方向52に配置される。

【実施例4】

【0098】

実施例1では、図6～図8に示したように、通信端末装置が、それぞれに矩形で平面の表示面を有し表示面をユーザ定位置に向けて配置された3つの表示器と、互いに水平方向に隔離して撮像方向をユーザ定位置に向けて配置された3つの撮像器とを有する構成を例示した。これに対して、実施例4では、通信端末装置10が、矩形で平面の表示面を有し表示面をユーザ定位置に向けた1つの表示器と、撮像方向をユーザ定位置に向けて配置された1つの撮像器とを有する構成を例示する。

【0099】

図18は、実施例4による通信端末装置のブロック図である。本実施例による通信端末装置10は、音声取得部11と、視覚情報取得部12と、人物捕捉部13と、情報伝達部14と、画像表示部15と、音声出力部16と、配置指定受付部17とを有する。音声取得部11、視覚情報取得部12、人物捕捉部13、情報伝達部14、画像表示部15、および音声出力部16は、基本的に、図2に示した実施形態のものと同様である。配置指定受付部17は、共有仮想空間41に対する自装置のユーザ空間22および他装置のユーザ空間32の配置のユーザによる指定を受け付ける。

【0100】

本実施例では、実施例1と異なり、視覚情報取得部12が1つの撮像器40Bを備え、画像表示部15が1つの表示器42Bを備えている。

【0101】

図19は、実施例4による通信端末装置の概略平面図である。図20は、実施例4の通信端末装置の概略斜視図である。

【0102】

図19に示すように、実施例4では、実施例1と同様に、ユーザ空間22における所定の位置をユーザ定位置51とし、そのユーザ定位置51から所定の方向をユーザ定方向52とする。例えば、ユーザ定位置51は通信端末装置10を用いている自ユーザ23が存在するであろう位置である。例えば、ユーザ定方向52は、通信端末装置10を用いている自ユーザ23が向いているであろう方向である。

【0103】

図20に示すように、表示器42Bは矩形で平面の表示面を有し、表示面をユーザ定位置51に向けて配置された表示装置である。表示器42Bは、ユーザ定位置51から見てユーザ定方向52に配置される。撮像器40Bは、ユーザ定位置51との相対位置が固定され、ユーザ定方向52に対する相対方向が固定されて、撮像方向をユーザ定位置51に向けて配置されている。好ましくは、撮像器40Bは、表示器42Bの表示面の上辺もしくは下辺の中央近傍に配置される。

【0104】

図21は、実施例4における画像表示処理のフローチャートである。実施例4における画像表示処理は、図9に示した実施例1のものとは、送信データを選択するステップS102の処理が無い点で異なるが、それ以外は基本的に実施例1のものと同様である。上述したように、実施例4では、視覚情報取得部12は1つの撮像器40Bを有し、情報伝達部14は、その撮像器40Bで取得された画像および深度の情報を含む画像伝達情報を他装置に送信する。

【0105】

なお、本実施例の通信端末装置 10 は、音声取得部 11 と視覚情報取得部 12 と人物捕捉部 13 と情報伝達部 14 と画像表示部 15 と音声出力部 16 とが筐体に内蔵され、視覚情報取得部の撮像器 40B が筐体の所定の位置に配設され、平面の表示面を備える表示器に接続され画像表示部 15 による画像のデータを表示器に出力する外部出力端子が筐体に設けられた装置として構成できる。その構成において、画像表示部 15 は、表示面のサイズと筐体に対する表示面の相対位置および相対姿勢の指定を受け、その指定に基づいて、表示面に自ユーザの視点位置から見える共有仮想空間の画像のデータを生成し、外部出力端子から出力する。外部に汎用的な平面の表示器を接続することにより、空間を共有したコミュニケーションを実現する通信端末装置 10 を容易に構成することができる。なお、本実施例においては表示器 42B を 1 つだけ備える構成を示したが、筐体からの外部出力端子を複数とし、複数の表示器を備える構成としてもよい。この場合、画像表示部 15 が受ける、表示面のサイズと筐体に対する表示面の相対位置および相対姿勢の指定は、接続される表示器の数だけ必要となる。

【0106】

また、上述したように、本実施例の通信端末装置 10 は、視覚情報取得部 12 が有する撮像器 40B がひとつであるため、他装置に伝達できる画像伝達情報から生成できる三次元データが限定される。また図 8、図 16 に示す構成に代表される実施例 1、3 の構成においても、設置される視覚情報取得部 12 の各撮像器の配置によっては生成できる三次元データは限定され完全なものにならないことがある。したがって、例えば、自装置あるいは他装置において、本実施例の通信端末装置 10 のユーザ 91 の視覚情報あるいは画像伝達情報に含まれていない部分を補完することにしてもよい。

自装置にて補完する場合、人物捕捉部 13 が視覚情報取得部 12 からの視覚情報から人物を捕捉したら、情報伝達部 14 が、その捕捉された人物の視覚情報に含まれていない部分を補完して画像伝達情報を生成する。具体的には、情報伝達部 14 は、予め視覚情報取得部 12 により取得されたユーザ 91 の側面や背面の補完用視覚情報を保存しておき、視覚情報取得部 12 で取得される視覚情報に含まれていない人物の側面や背面の部分を補完視覚情報により補完してもよい。あるいは、人物捕捉部 13 は、視覚情報取得部 12 で取得される視覚情報に含まれていない人物の側面や背面の部分を、その視覚情報に含まれている部分の情報をもとに補完してもよい。例えば、ユーザ 91 が黒色の服を着ていれば、ユーザ 91 の背中の視覚情報を黒色で作成してもよい。

他装置にて補完する場合、情報伝達部 14 及び／もしくは通信ネットワーク 90 上のサーバが、自装置（他装置にとっては他装置）から伝達された画像伝達情報に含まれる人物についての、その画像伝達情報に含まれていない部分を補完する。具体的には、情報伝達部 14 は、予め自装置（他装置にとっては他装置）からユーザ 91 の側面や背面の補完用視覚情報を取得して保存しておき、画像伝達情報に含まれていない人物の側面や背面の部分を補完視覚情報により補完してもよい。あるいは、人物捕捉部 13 は、画像伝達情報に含まれていない人物の側面や背面の部分を、その画像伝達情報に含まれている部分の情報をもとに補完してもよい。

【実施例 5】

【0107】

実施例 1 は、図 7 ～図 8 に示したように、通信端末装置が矩形で平面の表示面を有する表示器を備える構成を例示した。実施例 5 では、通信端末装置が曲面の表示面を有する表示器を備える構成を例示する。

【0108】

図 22 は、実施例 5 による通信端末装置のブロック図である。本実施例による通信端末装置 10 は、音声取得部 11 と、視覚情報取得部 12 と、人物捕捉部 13 と、情報伝達部 14 と、画像表示部 15 と、音声出力部 16 と、配置指定受付部 17 とを有する。音声取得部 11、視覚情報取得部 12、人物捕捉部 13、情報伝達部 14、画像表示部 15、音声出力部 16、および配置指定受付部 17 は、基本的に、図 6 に示した実施例 1 のものと同様である。

【0109】

本実施例では、視覚情報取得部12が3つの撮像器40A、40B、40Cを備え、画像表示部15が1つの表示器42Dを備えている。

【0110】

図23は、実施例5による通信端末装置の概略平面図である。図24は、実施例5の通信端末装置の概略斜視図である。

【0111】

図23、24に示すように、実施例5では、画像表示部15の表示器42Dは垂直な円筒形状の内周面を表示面とする構成である。表示器42Dの円筒の内部がユーザ空間22となっている。ユーザ空間22における所定の位置をユーザ定位置51とする。例えば、ユーザ定位置51は通信端末装置10を用いている自ユーザ23が存在するであろう位置である。表示器42Dが円筒形状なのでその中心軸の近傍にユーザ定位置51を設定してもよい。

【0112】

表示器42Dはユーザ定位置51を囲む円筒形状の内周面を表示面とした表示装置である。

【0113】

撮像器40A～40Cは、それぞれが表示器42Dの表示面に近接し、ユーザ定位置51との相対位置が固定され、撮像方向をユーザ定位置51に向けて配置されている。撮像器40A～40Cは、例えば、撮像器40Aは表示器42Dの中心軸を中心として撮像器40Bとの間に110°の角度をなす方向に、撮像器40Cは表示器42Dの中心軸を中心として撮像器40Bとの間に-110°の角度をなす方向に配置されていてもよい。また、撮像器40A～40Cの高さは、例えば、自ユーザ23として標準的な身長の人物を想定し、その頭上、足元、あるいは顔の高さとしてもよい。

【0114】

実施例5における画像表示処理のフローチャートは、図9に示した実施例1のものと同様である。

【0115】

図25は、実施例5における配置指定画面の一例を示す図である。配置指定画面60には共有スペース領域64と決定ボタン63が含まれている。共有スペース領域64は、自ユーザ23と他ユーザ33とが共有する仮想的なスペースに相当する共有仮想空間41を模した平面図である。共有スペース領域64には、対話に参加する各メンバーが利用する通信端末装置10に相当するオブジェクトであるブースアイコン66が配置されている。ブースアイコン66は、例えばマウス操作、表示器42Dの表示面へのタッチ操作、撮像器40A～40Cで取得される視覚情報に基づき検出されるジェスチャー操作により移動させたり固定させたりできる。共有スペース領域64の形状および大きさを変更できるようにしてもよい。ここでは一例として共有スペース領域64が共有仮想空間41を模したものである。共有スペース領域64の形状および大きさに応じて共有仮想空間41の形状および大きさを決定してもよい。決定ボタン63が押下されると、そのときの共有スペース領域64に対するブースアイコン66の配置に基づいて、共有仮想空間41における各通信端末装置10のユーザ空間と、ユーザ定位置51と、他装置のユーザ定位置との配置が決定される。図25の例では、3人のユーザが三角形の頂点をなす配置が例示されている。

【0116】

図26は、複数のユーザが共有スペースでの対面による対話が擬似的に実現された様子を示す概念図である。図26には、図25に示した配置が模擬的に実現された例が示されている。図26には、X社のAさんとBさんとCさんが三角形の頂点をなす位置に立ち、対話を行っている。それぞれのユーザの通信端末装置10の表示器42Dには、図26に示した配置で、それぞれのユーザの視点位置から見える共有仮想空間41の画像が表示される。それにより、全てのユーザは、図26に示したような配置で対話をしている視界が

得られる。

【0117】

本実施例によれば、自ユーザ23から見て360°度どの方向にも他のユーザと共有する仮想空間の画像が表示され、自ユーザ23に高い臨場感を与えることができる。例えば、共有仮想空間としてCGで構成された仮想空間あるいは実写映像で構成される三次元空間を用い、その共有仮想空間の中で複数のユーザが一緒に居て対話をする状態を再現することができる。実写映像は、いずれの通信端末装置10の設置場所と異なる地点で、例えば、360度カメラで撮影されるリアルタイムの映像あるいは録画された映像であってもよい。

【0118】

なお、本実施例では、視覚情報取得部12が3つの撮像器40A～40Cを備える例を示したが、視覚情報取得部12が備える撮像器の個数は3つに限定されない。他の例として、2つであってもよいし、4つ以上であってもよい。また、撮像器が配置される位置についても本実施例に記載された配置に限定されない。

【実施例6】

【0119】

実施例5では、円筒形状の内周面に表示面を有する通信端末装置を例示した。これに対して、実施例6では、円柱形状の外周面に表示面を有する通信端末装置を例示する。

【0120】

図27は、実施例6による通信端末装置のブロック図である。本実施例による通信端末装置10は、音声取得部11と、視覚情報取得部12と、人物捕捉部13と、情報伝達部14と、画像表示部15と、音声出力部16と、配置指定受付部17とを有する。音声取得部11、視覚情報取得部12、人物捕捉部13、情報伝達部14、画像表示部15、音声出力部16、および配置指定受付部17は、基本的に、図6に示した実施例1のものと同様である。

【0121】

本実施例では、視覚情報取得部12が8つの撮像器40A～40Hを備え、画像表示部15が1つの表示器42Eを備えている。

【0122】

図28は、実施例6による通信端末装置の概略平面図である。図29は、実施例6の通信端末装置の概略斜視図である。

【0123】

図28、29に示すように、実施例6では、画像表示部15の表示器42Eは垂直な円柱形状の外周面を表示面とする構成である。

【0124】

表示器42Eは円柱形状の外周面を表示面とした表示装置である。表示器42Eの円柱の外部がユーザ空間22となっている。撮像器40A～40Hは、それぞれが表示器42Eの表示面に近接し、円柱の外側に向けて固定され、撮像方向をユーザ定位置51に向けて配置されている。撮像器40A～40Hは、例えば、表示器42Eの中心軸を中心として互いに45°の角度をなす方向に配置されていてもよい。また、撮像器40A～40Hの高さは、例えば、自ユーザ23として標準的な身長的人物を想定し、その頭上、足元、あるいは顔の高さとしてもよい。

【0125】

実施例6の通信端末装置10は、実施例5の通信端末装置10との組合せにより対話を行うのに好適である。以下、実施例6の通信端末装置10のユーザをユーザ91(5)と称し、実施例5の通信端末装置10のユーザをユーザ91(4)と称することにする。

【0126】

実施例6の通信端末装置10が設置された実空間の広さをもとに共有仮想空間41の大きさを設定する。例えば、実施例6の通信端末装置10における共有仮想空間41を無限大としてもよい。実施例6の通信端末装置10の円柱形状の中心軸と、実施例5の通信端

10

20

30

40

50

末装置 10 の円筒形状の中心軸との相対位置を共有仮想空間 41 において一致させる。そのうえで、例えば、撮像器 40A ~ 40H から得られるイメージ映像をリアルタイムに合成して作成した、通信端末装置 10 が設置されたユーザ 91 (5) がいる実空間の 360° 映像をユーザ 91 (4) の通信端末装置 10 の表示器に投影することで、ユーザ 91 (4) には、ユーザ 91 (5) がいる実空間に自身が入り込んだような視界が得られる。ユーザ 91 (5) には、自身がいる実空間にユーザ 91 (4) が現れたような視界が得られる。図 30 は、実施例 6 のユーザから得られる視界を示す概念図である。ユーザ 91 (5) が実際にいる室内の実空間 71 に、実際には遠隔にいるユーザ 91 (4) が現れ、ユーザ 91 (5) とユーザ 91 (4) が直接対面して対話している状態が模擬されている。

【0127】

本実施例によれば、自ユーザは、自身が実際に居る実空間と表示面により表示される共有仮想空間とを同時に見ることができるので、自ユーザにとって自身の居る場所に他ユーザが現れたような感覚での他ユーザとのコミュニケーションが可能となる。また、実施例 5 の通信端末装置 10 を用いる他ユーザにとっては、実施例 6 の通信端末装置 10 が設置された場所を自ユーザとともに訪問している感覚でのコミュニケーションが可能となる。例えば、遠隔地の視察といったユースケースが想定される。

【0128】

なお、本実施例では、円柱形状の外周面が表示面となっている構成を例示したが、表示面の形状が本構成に限定されることはない。例えば、多角柱や球体形状の外周面が表示面となる構成であってもよいし、卵型形状の外周面が表示面となる構成であってもよい。

【符号の説明】

【0129】

1…通信システム、10…通信端末装置、11…音声取得部、12…視覚情報取得部、13…人物捕捉部、14…情報伝達部、14C…撮像器、14B…撮像器、15…画像表示部、16…音声出力部、17…配置指定受付部、21…仮想空間、22…ユーザ空間、23…自ユーザ、24…視点位置、31…仮想空間、32…ユーザ空間、33…他ユーザ、35…ホワイトボード、40…撮像器、41…共有仮想空間、42…表示面、42A~42E…表示器、51…ユーザ定位置、52…ユーザ定方向、60…配置指定画面、61…会議室領域、62…机アイコン、63…決定ボタン、64…共有スペース領域、65…画像画面、66…ブースアイコン、71…実空間、90…通信ネットワーク、91…ユーザ

10

20

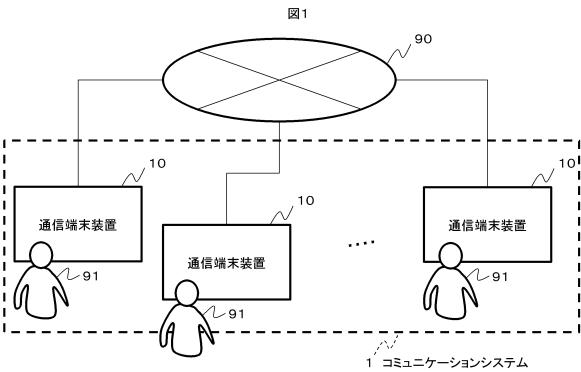
30

40

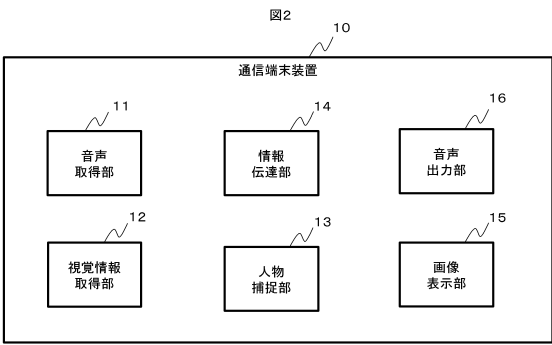
50

【図面】

【図 1】

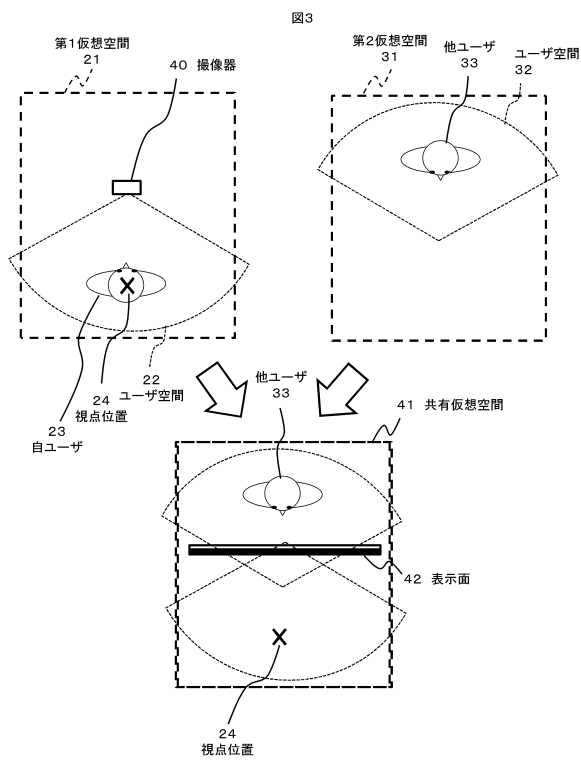


【図 2】

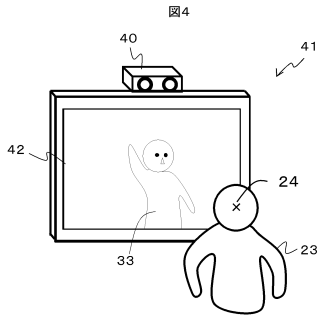


10

【図 3】



【図 4】



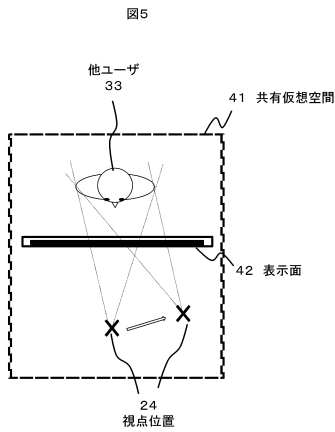
20

30

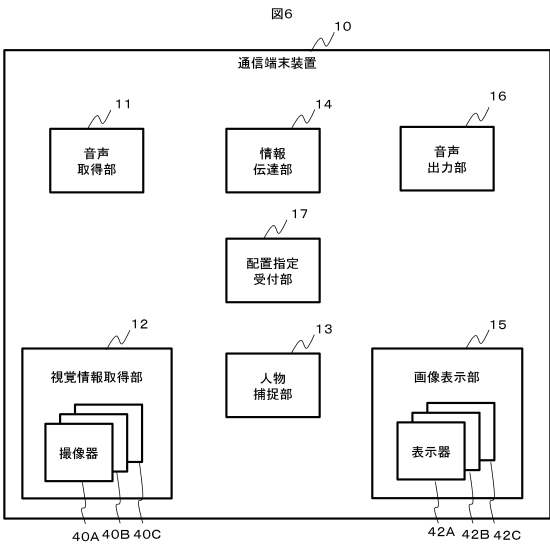
40

50

【図 5】

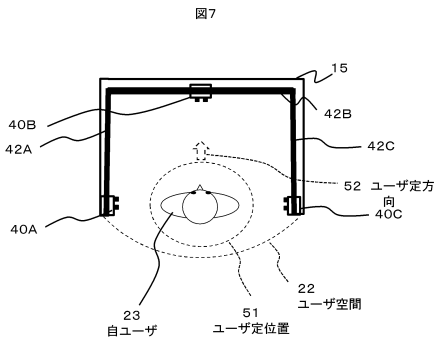


【図 6】

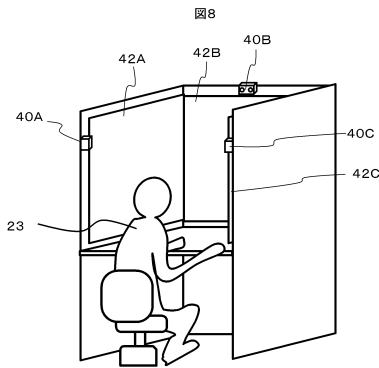


10

【図 7】



【図 8】



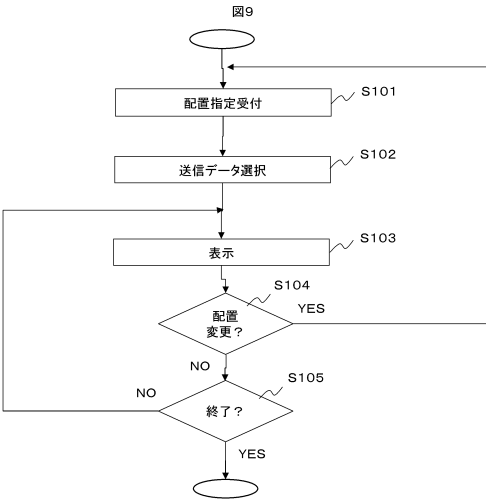
20

30

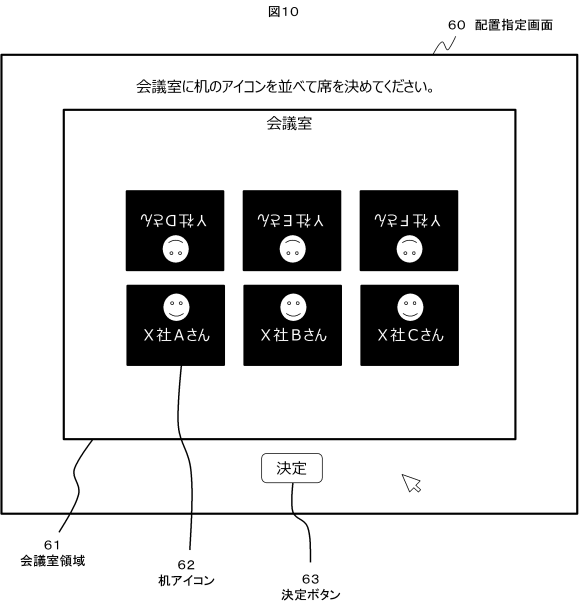
40

50

【図 9】

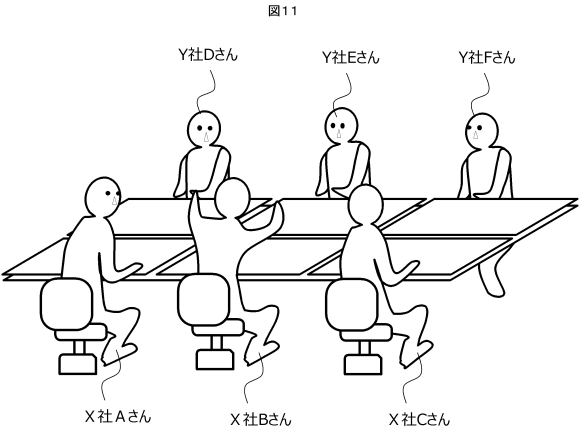


【図 10】

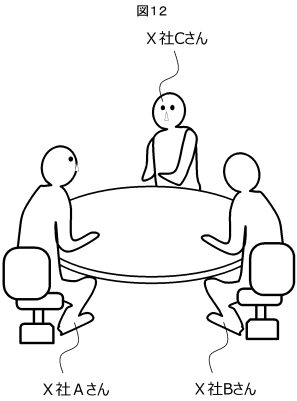


10

【図 11】



【図 12】



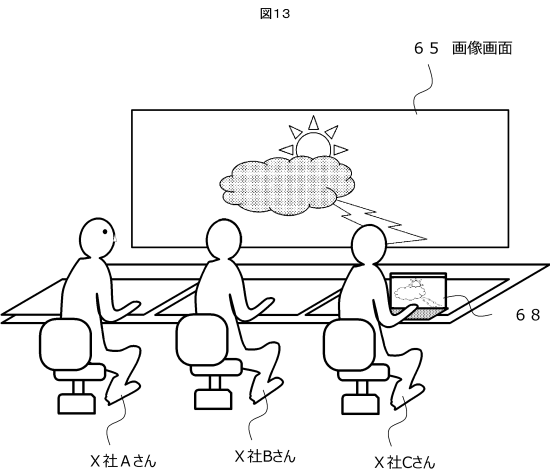
20

30

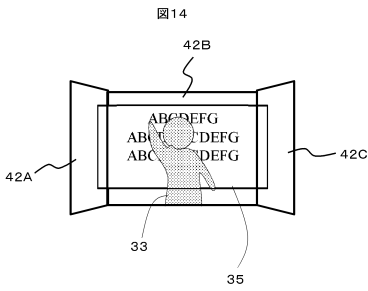
40

50

【図 1 3】

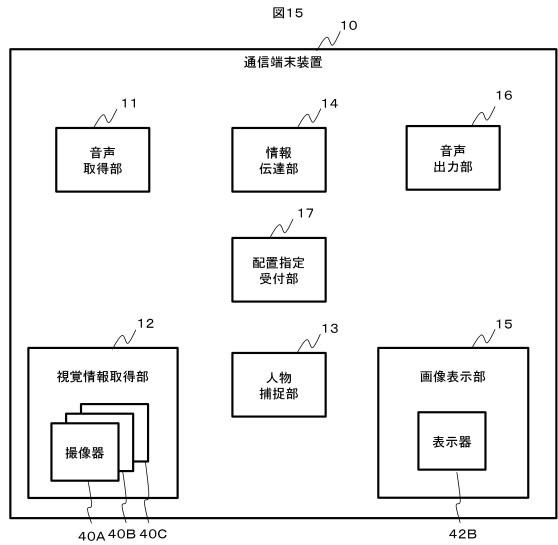


【図 1 4】

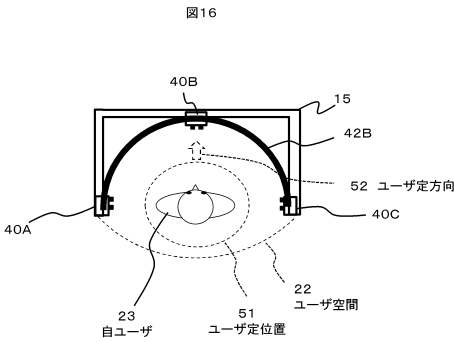


10

【図 1 5】



【図 1 6】



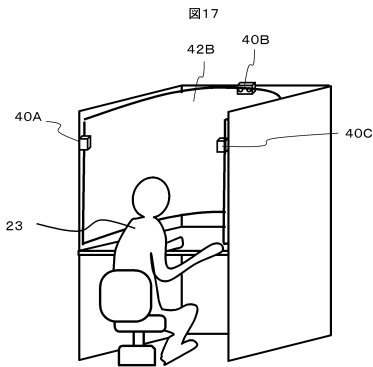
20

30

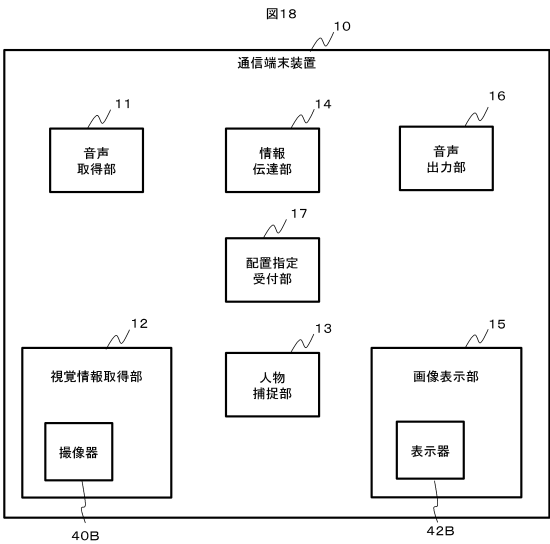
40

50

【図 17】

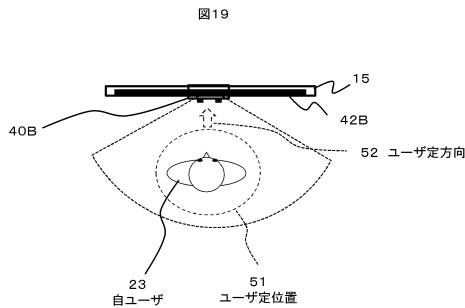


【図 18】

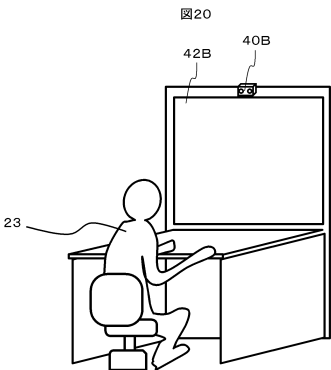


10

【図 19】



【図 20】



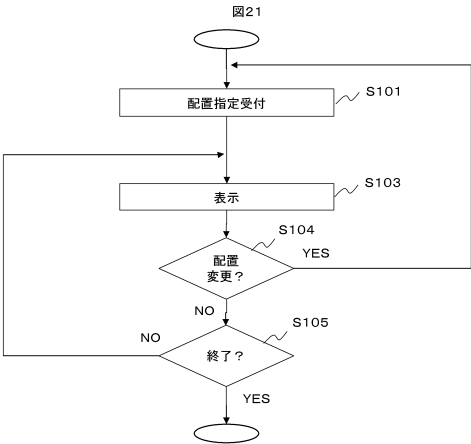
20

30

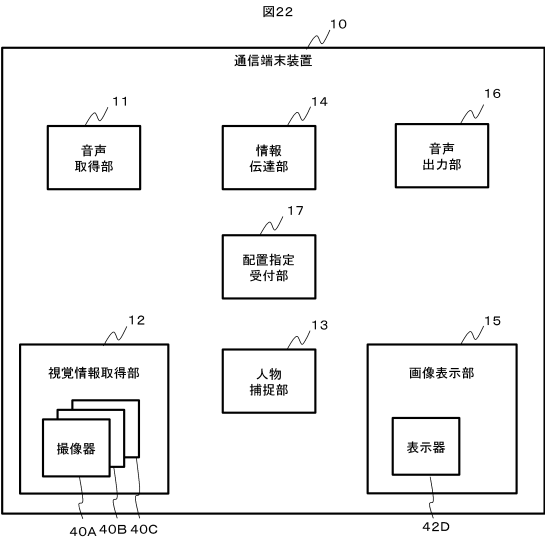
40

50

【図 2 1】

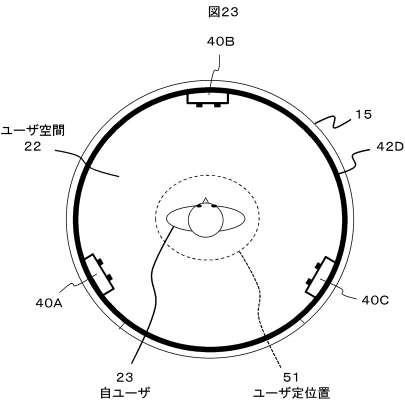


【図 2 2】

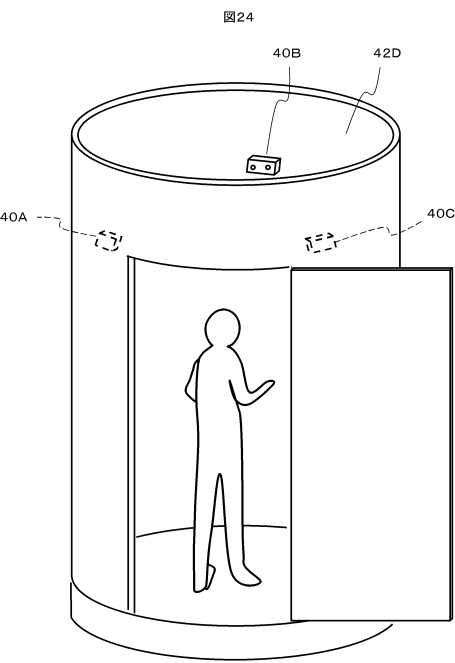


10

【図 2 3】



【図 2 4】



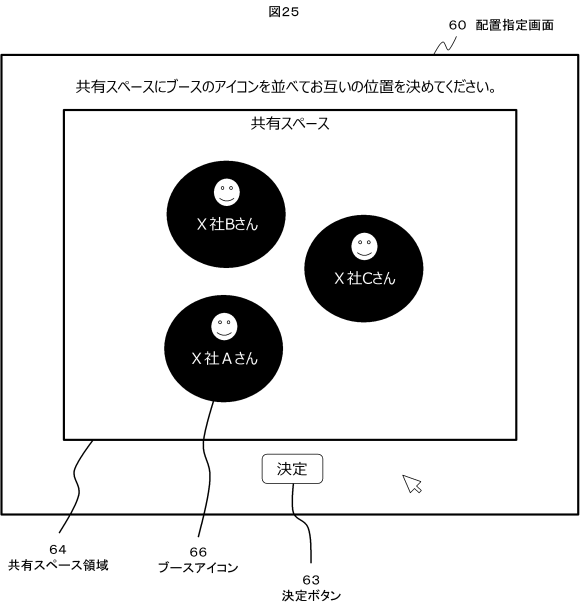
20

30

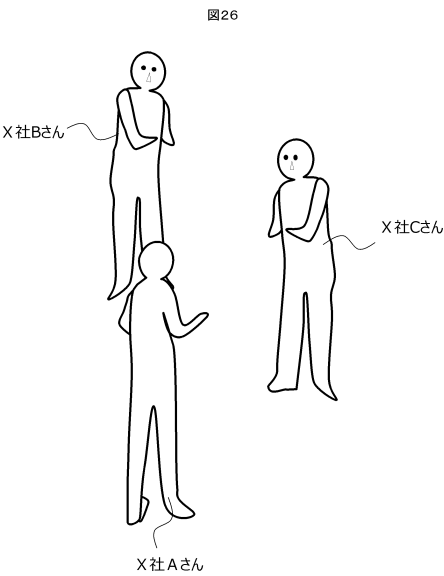
40

50

【図 2 5】



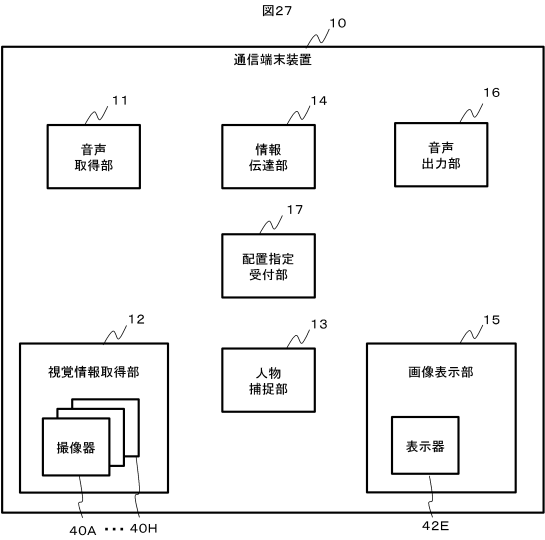
【図 2 6】



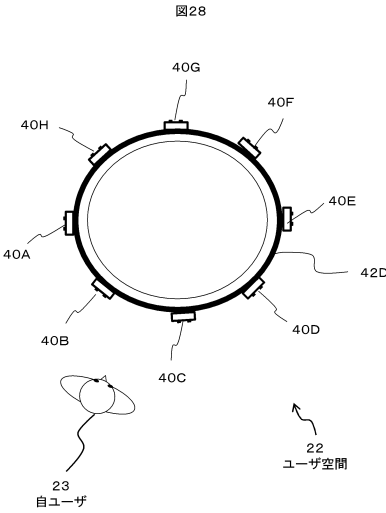
10

20

【図 2 7】



【図 2 8】

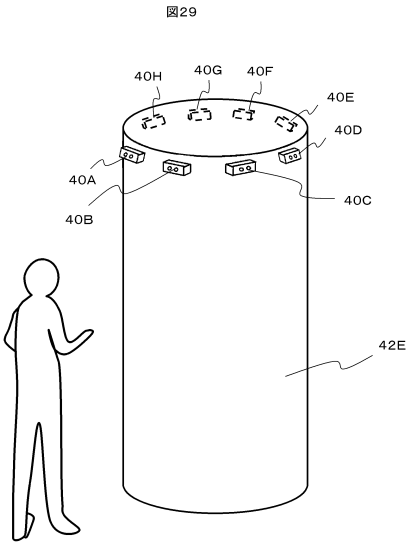


30

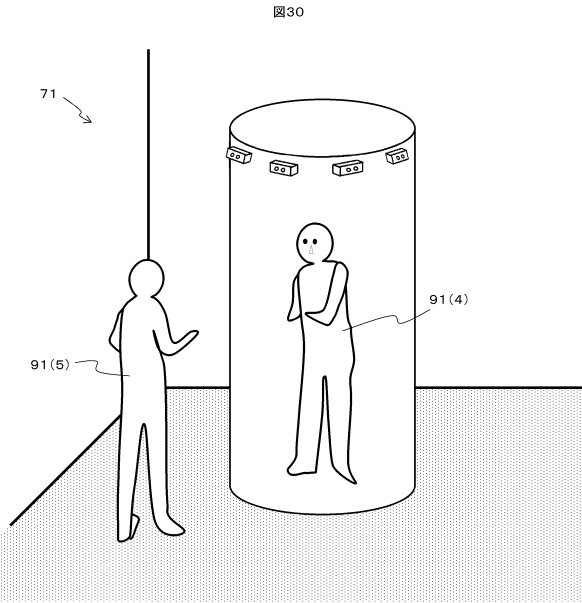
40

50

【図 29】



【図 30】



10

20

30

40

50

フロントページの続き

(56)参考文献 国際公開第2019/230852 (WO, A1)
特開2020-065229 (JP, A)
特開2000-165831 (JP, A)
特開2017-017522 (JP, A)
特開2012-169948 (JP, A)
国際公開第2019/187747 (WO, A1)
特開2006-287297 (JP, A)
国際公開第2018/182063 (WO, A1)
特開平06-274094 (JP, A)
特開平08-167048 (JP, A)

(58)調査した分野 (Int.Cl., DB名)
G06T 19/00
H04N 7/15