

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-195521

(P2017-195521A)

(43) 公開日 平成29年10月26日(2017.10.26)

(51) Int.Cl.	F I	テーマコード(参考)
HO4R 3/00 (2006.01)	HO4R 3/00 320	2H199
GO2B 27/02 (2006.01)	GO2B 27/02 Z	5D220

審査請求 未請求 請求項の数 2 O L (全 26 頁)

(21) 出願番号 特願2016-84824 (P2016-84824)
 (22) 出願日 平成28年4月20日(2016.4.20)

(71) 出願人 310021766
 株式会社ソニー・インタラクティブエンタ
 テイメント
 東京都港区港南1丁目7番1号
 (74) 代理人 100105924
 弁理士 森下 賢樹
 (74) 代理人 100109047
 弁理士 村田 雄祐
 (74) 代理人 100109081
 弁理士 三木 友由
 (74) 代理人 100134256
 弁理士 青木 武司

最終頁に続く

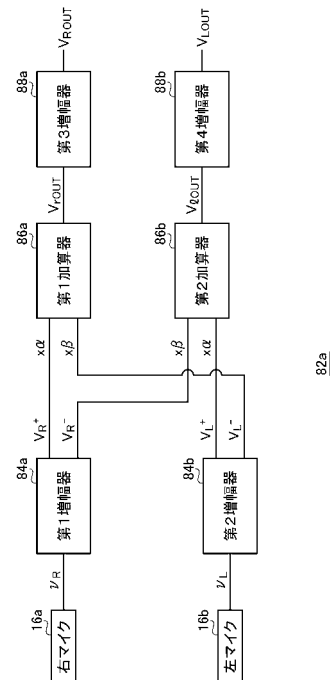
(54) 【発明の名称】 位相差増幅装置

(57) 【要約】

【課題】ロボットが取得した音声信号を加工する技術を提供する。

【解決手段】右マイク16aは、第1音声信号を生成し、第1増幅器84aは、第1音声信号を増幅した第1正相信号と、第1音声信号を反転増幅した第1逆相信号とを出力する。左マイク16bは、右マイク16aとは異なる位置に配置されて第2音声信号を生成し、第2増幅器84bは、第2音声信号を増幅した第2正相信号と、第2音声信号を反転増幅した第2逆相信号とを出力する。第1加算器86aは、第1正相信号を第1係数倍した信号と、第2逆相信号を第2係数倍した信号とを加算する。第2加算器86bは、第2正相信号を第1係数倍した信号と、第1逆相信号を第2係数倍した信号とを加算する。第3増幅器88aは、第1加算器86aの出力信号を第3係数倍して出力する。第4増幅器88bは、第2加算器86bの出力信号を第3係数倍して出力する。

【選択図】 図11



82a

【特許請求の範囲】

【請求項 1】

第 1 音声信号を生成する第 1 マイクと、
 第 1 音声信号を増幅した第 1 正相信号と、第 1 音声信号を反転増幅した第 1 逆相信号と
 を出力する第 1 増幅器と、
 前記第 1 マイクとは異なる位置に配置され、第 2 音声信号を生成する第 2 マイクと、
 第 2 音声信号を増幅した第 2 正相信号と、第 2 音声信号を反転増幅した第 2 逆相信号と
 を出力する第 2 増幅器と、
 第 1 正相信号を第 1 係数倍した信号と、第 2 逆相信号を第 2 係数倍した信号とを加算す
 る第 1 加算器と、
 第 2 正相信号を第 1 係数倍した信号と、第 1 逆相信号を第 2 係数倍した信号とを加算す
 る第 2 加算器と、
 を備えることを特徴とする位相差増幅装置。

10

【請求項 2】

第 1 加算器の出力信号を第 3 係数倍して出力する第 3 増幅器と、
 第 2 加算器の出力信号を第 3 係数倍して出力する第 4 増幅器と、
 を備えることを特徴とする請求項 1 に記載の位相差増幅装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ユーザの動きに応じてロボットを動かし、ロボットが生成した視聴データを利用する技術に関する。

20

【背景技術】

【0002】

ヘッドマウントディスプレイ (HMD) が様々な分野で利用されている。HMD にヘッドトラッキング機能をもたせ、ユーザの頭部の姿勢と連動して表示画面を更新することで、映像世界への没入感を高められる。

【先行技術文献】

【特許文献】

【0003】

30

【特許文献 1】特開 2015 - 95045 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

近年、遠隔地に配置したロボットを自分の分身として利用するトレイグジスタンスと呼ばれる技術が登場している。遠隔地にいるロボットが周囲の画像データや音声データをユーザに送信し、ユーザ側で再生することで、ユーザは、ロボットの場所にいるような臨場感をもって、周囲の人達とコミュニケーションをとることが可能となる。

【0005】

本発明者はトレイグジスタンスと HMD の連携による可能性に注目し、トレイグレジスタンスシステムの利便性および有用性を高める技術を開発するに至った。

40

【0006】

本発明はこうした課題に鑑みてなされたものであり、その目的は、遠隔操作するロボットの構造や、ロボットが取得した視聴データを加工する技術、またロボットが取得した視聴データを有用に活用するための技術を提供することにある。

【課題を解決するための手段】

【0007】

上記課題を解決するために、本発明のある態様の位相差増幅装置は、第 1 音声信号を生成する第 1 マイクと、第 1 音声信号を増幅した第 1 正相信号と、第 1 音声信号を反転増幅した第 1 逆相信号とを出力する第 1 増幅器と、第 1 マイクとは異なる位置に配置され、第

50

2音声信号を生成する第2マイクと、第2音声信号を増幅した第2正相信号と、第2音声信号を反転増幅した第2逆相信号とを出力する第2増幅器と、第1正相信号を第1係数倍した信号と、第2逆相信号を第2係数倍した信号とを加算する第1加算器と、第2正相信号を第1係数倍した信号と、第1逆相信号を第2係数倍した信号とを加算する第2加算器と、を備える。

【0008】

なお、以上の構成要素の任意の組合せ、本発明の表現を方法、装置、システム、コンピュータプログラム、コンピュータプログラムを読み取り可能に記録した記録媒体、データ構造などの間で変換したものもまた、本発明の態様として有効である。

【発明の効果】

10

【0009】

本発明によれば、遠隔操作するロボットの構造や、ロボットが取得した視聴データを加工する技術、またロボットが取得した視聴データを有用に活用するための技術を提供できる。

【図面の簡単な説明】

【0010】

【図1】実施例における情報処理システムの構成例を示す図である。

【図2】ロボットの利用場面の例を示す図である。

【図3】HMDの外観形状の例を示す図である。

【図4】HMDの機能ブロックを示す図である。

20

【図5】ロボットの外観構成を示す図である。

【図6】挿通部材の構成を示す図である。

【図7】ロボットの断面を示す図である。

【図8】ロボットにおける筐体の姿勢の例を示す図である。

【図9】ロボットにおける筐体の姿勢の例を示す図である。

【図10】ロボットの機能ブロックを示す図である。

【図11】音声処理部が備える位相差増幅装置の回路構成を示す図である。

【図12】信号波形の位相差を説明するための図である。

【図13】入力信号波形の位相差を増幅する原理を説明するための図である。

【図14】応用技術を実現するためのロボットの機能ブロックを示す図である。

30

【図15】処理装置の機能ブロックを示す図である。

【図16】全天球パノラマ画像を説明するための図である。

【図17】画像記録部に記録されている撮影画像データを説明するための図である。

【図18】画像生成部が生成するフレーム画像と、画像データとの関係を示す図である。

【発明を実施するための形態】

【0011】

図1は、実施例における情報処理システム1の構成例を示す。情報処理システム1は、ロボット10と、ユーザAが頭部に装着するヘッドマウントディスプレイ装置(HMD)100とを備える。HMD100は、両眼用の表示パネル102と、両耳用のイヤホン104と、マイク106とを備える。ここでは音声出力手段としてイヤホン104を採用しているが、耳に当てる形状のヘッドホンを採用してもよい。HMD100はアクセスポイント(AP)2を介して、ネットワーク4に接続される。AP2は無線アクセスポイントおよびルータの機能を有し、HMD100は、AP2と既知の無線通信プロトコルで接続するが、ケーブルで接続してもよい。

40

【0012】

ロボット10は、アクチュエータ装置12と、アクチュエータ装置12により姿勢を変更可能に駆動される筐体20とを備える。筐体20には、右カメラ14a、左カメラ14b、右マイク16a、左マイク16bおよびスピーカ18が搭載される。以下、右カメラ14aおよび左カメラ14bを特に区別しない場合には「カメラ14」と呼び、右マイク16aおよび左マイク16bを特に区別しない場合には「マイク16」と呼ぶ。実施例に

50

においてカメラ 14 およびマイク 16 は、アクチュエータ装置 12 により駆動される筐体 20 に設けられるが、スピーカ 18 は、たとえばアクチュエータ装置 12 の半球状のハウジング 36 に設けられてもよい。ロボット 10 はアクセスポイント (AP) 3 を介して、ネットワーク 4 に接続される。ロボット 10 は、AP 3 と既知の無線通信プロトコルで接続するが、ケーブルで接続してもよい。

【0013】

情報処理システム 1 において、HMD 100 とロボット 10 はネットワーク 4 を介して通信可能に接続し、ロボット 10 は、ユーザ A のいわば分身として動作する。ユーザ A が装着している HMD 100 の動きはロボット 10 に伝達され、アクチュエータ装置 12 が、HMD 100 の動きに連動して筐体 20 を動かす。たとえばユーザ A が首を前後に振ると、アクチュエータ装置 12 が筐体 20 を前後に振るように動かし、ユーザ A が首を左右に振ると、アクチュエータ装置 12 が筐体 20 を左右に振るように動かす。これによりロボット 10 の周囲にいる人は、ユーザ A がその場にいるかのような感覚をもって、ユーザ A とコミュニケーションをとることができる。

10

【0014】

右カメラ 14 a および左カメラ 14 b は、筐体 20 の前面にて横方向に所定の間隔を空けて配置される。右カメラ 14 a および左カメラ 14 b はステレオカメラを構成し、右カメラ 14 a は右目用画像を所定の周期で撮影し、左カメラ 14 b は左目用画像を所定の周期で撮影する。撮影された右目用画像および左目用画像は、リアルタイムでユーザ A の HMD 100 に送信される。HMD 100 は、受信した右目用画像を右目用表示パネルに表示し、受信した左目用画像を左目用表示パネルに表示する。これによりユーザ A は、ロボット 10 の筐体 20 が向いている方向の映像をリアルタイムで見ることができる。

20

【0015】

右マイク 16 a および左マイク 16 b は、筐体 20 において横方向に所定の間隔を空けて配置される。右マイク 16 a および左マイク 16 b はステレオマイクを構成し、横方向に所定の間隔を空けて配置されることで、音源の位置に応じて音声が届く時間が異なるようにされる。音声の到達時間の差は、右マイク 16 a および左マイク 16 b が生成する音声信号の位相差として表れる。なお右マイク 16 a および左マイク 16 b の音声信号の位相差を大きくするために、右マイク 16 a および左マイク 16 b は可能な限り離して、具体的には筐体 20 の両側面に配置されることが好ましい。

30

【0016】

右マイク 16 a および左マイク 16 b で生成された音声信号は、後述するように加工されて、右耳用音声データおよび左耳用音声データとしてリアルタイムでユーザ A の HMD 100 に送信される。HMD 100 は、受信した右耳用音声データを右耳用のイヤホン 104 から出力し、受信した左耳用音声データを左耳用のイヤホン 104 から出力する。これによりユーザ A は、ロボット 10 の周囲の音声をリアルタイムで聞くことができる。

【0017】

人間が左右方向の音源の位置を音波の両耳への到達時間の差によって知覚することは知られているが、実際には到達時間の差だけでなく、音波を集める耳介の形状、音波を中耳に伝える外耳道の形状等にも依存して音源の位置を知覚している。また人間の正面に対して右側または左側に音源がある場合、距離の近い側の耳介と比べると、距離の遠い側の耳介に音波が到達するためには経路中に顔が位置するため、音波の到達時間差は音源からの距離差以上に大きくなる。

40

【0018】

一方で筐体 20 の前面は平坦な形状を有し、またマイク 16 は耳介や外耳道に相当する形状を有していないため、音声到達時間差は、実質的に音源と両マイクとの距離差に対応することになる。実施例では、筐体 20 の両側面に右マイク 16 a および左マイク 16 b を配置して両者を最大限離れた位置に設けているが、右マイク 16 a で生成した音声信号と左マイク 16 b で生成した音声信号を増幅して右耳用イヤホンと左耳用イヤホンから出力しても、音源の左右方向の位置をよく知覚できないことが本発明者の実験により明らか

50

になった。

【0019】

つまり人間が普段聞き慣れている音と比べると、右マイク16aおよび左マイク16bが生成する音声信号の位相差は、左右方向を知覚するには小さいことが実験により判明した。そこでロボット10は、右マイク16aおよび左マイク16bの音声信号の位相差を増幅して、より人間の両耳で聞こえる音に近づけた音声データをHMD100に提供する仕組みを備えている。この仕組みについては後述する。

【0020】

HMD100において、マイク106は、ユーザAが発した音声信号を生成する。ユーザAによる音声データは、リアルタイムでロボット10に送信され、ロボット10は、受信した音声データをスピーカ18から出力する。これによりロボット10の周辺にいる人は、ユーザAが発した音声をリアルタイムで聞くことができる。

10

【0021】

このように情報処理システム1では、ロボット10がユーザAにより遠隔操作されてユーザAの顔の動きや音声を再現し、またユーザAがHMD100を通じて、ロボット周辺の画像や音声を視聴でき、ユーザAとロボット10周辺の人とが、リアルタイムでコミュニケーションをとることができる。このような情報処理システム1は、様々な環境において有用に利用される。

【0022】

図2は、ロボット10の利用場面の例を示す。この例では、部屋で会議をしており、ユーザAの分身であるロボット10が、テーブル上に配置されている。この例でロボット10は正面の4人の方向を向いており、カメラ14は、正面の4人を画角内で撮影している。ロボット10は、カメラ14の撮影画像をリアルタイムでユーザAのHMD100に送信する。ユーザAはHMD100の表示パネル102を通じて部屋の状況を見ながら会議に参加し、発言するとユーザAの音声がロボット10にリアルタイムで送信され、ロボット10は、スピーカ18からユーザAの音声を出力する。

20

【0023】

また上記したように、ロボット10は、左右のマイク16で生成した音声信号の位相差を増幅した音声データを、リアルタイムでHMD100に送信する。これによりユーザAは、部屋内で声を出した人が、筐体20が向いている方向に対して右側に位置するのか、または左側に位置するのか、または正面に位置するのかを知覚できる。ユーザAは、自分の右側の人が発言したと感じると、首を右に回して右側を向く。このときロボット10の筐体20もユーザAの首の動きに連動して右側に向くため、カメラ14は、右側に座っている参加者を撮影することになる。

30

【0024】

このようにユーザAは、分身であるロボット10がユーザAの動きに連動することで、遠隔地にいながら、あたかも部屋にいるような感覚で会議に参加できる。また実際に部屋にいる参加者も、ユーザAの声や、また筐体20の動きから、ユーザAと違和感なくコミュニケーションをとることができる。なお図2に示す利用場面は一例であり、他の利用場面においてもユーザAは遠隔地にいながら、ロボット10から視聴データを得ることができる。

40

【0025】

図3は、HMD100の外観形状の例を示す。この例においてHMD100は、出力機構部110および装着機構部112から構成される。装着機構部112は、ユーザが被ることにより頭部を一周してHMD100を頭部に固定する装着バンド108を含む。装着バンド108はユーザの頭囲に合わせて長さの調節が可能な素材または構造とする。

【0026】

出力機構部110は、HMD100をユーザが装着した状態において左右の目を覆う形状の筐体114を含み、内部には目に正対する位置に表示パネル102を備える。表示パネル102は液晶パネルや有機ELパネルなどであってよい。筐体114内部には、表示

50

パネル 102 とユーザの目との間に位置し、ユーザの視野角を拡大する左右一对の光学レンズが備えられる。

【0027】

HMD 100 はさらに、装着時にユーザの耳に差し込まれるイヤホン 104 を備える。なおイヤホン 104 は、音声出力手段の一例であり、HMD 100 はヘッドホンも備えてもよい。このとき HMD 100 とヘッドホンとは、一体に構成されてもよいが、別体であってもよい。

【0028】

HMD 100 は、姿勢センサが検出したセンサ情報、およびマイク 106 からの音声信号を符号化した音声データをロボット 10 に送信し、またロボット 10 で生成された画像データおよび音声データを受信して、表示パネル 102 およびイヤホン 104 から出力する。

10

【0029】

なお図 3 に示す HMD 100 は、両目を完全に覆う没入型（非透過型）のディスプレイ装置を示すが、透過型のディスプレイ装置であってもよい。また形状としては、図示されるような帽子型であってもよいが、眼鏡型であってもよい。

【0030】

図 4 は、HMD 100 の機能ブロックを示す。制御部 120 は、画像信号、音声信号、センサ情報などの各種信号およびデータや、命令を処理して出力するメインプロセッサである。記憶部 122 は、制御部 120 が処理するデータや命令などを一時的に記憶する。姿勢センサ 124 は、HMD 100 の回転角度や傾きなどの姿勢情報を所定の周期で検出する。姿勢センサ 124 は、少なくとも 3 軸の加速度センサおよび 3 軸のジャイロセンサを含む。マイク 106 は、ユーザの声を電気信号に変換して音声信号を生成する。

20

【0031】

通信制御部 126 は、ネットワークアダプタまたはアンテナを介して、有線または無線通信により、ロボット 10 との間で信号やデータを送受信する。通信制御部 126 は、制御部 120 から、姿勢センサ 124 で検出された姿勢情報、およびマイク 106 からの音声信号を符号化した音声データを受け取り、ロボット 10 に送信する。また通信制御部 126 は、ロボット 10 から、画像データおよび音声データを受け取り、制御部 120 に供給する。制御部 120 は、画像データおよび音声データをロボット 10 から受け取ると、画像データを表示パネル 102 に供給して表示させ、また音声データをイヤホン 104 に供給して音声出力させる。

30

【0032】

図 5 は、ロボット 10 の外観構成を示す。筐体 20 は、カメラ 14、マイク 16 およびスピーカ 18 を收容する。カメラ 14 およびスピーカ 18 は筐体前面に設けられ、マイク 16 は筐体側面に設けられる。筐体 20 は保護カバー 19 を有し、ロボット 10 を使用しない状態では、保護カバー 19 が筐体前面を覆う閉位置に配置されて、カメラ 14 およびスピーカ 18 を保護する。図 5 に示す状態は、保護カバー 19 が閉位置から略 180 度回転した開位置に配置され、カメラ 14 が露出して、周囲を撮影可能となっている。保護カバー 19 は開位置で固定されるストッパ機構を有することが好ましい。

40

【0033】

筐体 20 はアクチュエータ装置 12 によって姿勢を変更可能に支持されている。アクチュエータ装置 12 は、脚部 40 と、脚部 40 の上部に支持される半球状のハウジング 36 と、筐体 20 を駆動するための駆動機構 50 とを備える。駆動機構 50 は、長尺方向に第 1 貫通長孔 32 a を形成された第 1 円弧状アーム 32 と、長尺方向に第 2 貫通長孔 34 a を形成された第 2 円弧状アーム 34 と、第 1 円弧状アーム 32 と第 2 円弧状アーム 34 とを交差させた状態で、第 1 円弧状アーム 32 と第 2 円弧状アーム 34 とを回動可能に支持する台座 30 とを備える。台座 30 の上側は、カバー 38 により覆われており、カバー 38 で覆われた空間には、第 1 円弧状アーム 32 および第 2 円弧状アーム 34 をそれぞれ回転させるモータが配置されている。なお台座 30 は、ハウジング 36 に対して回動可能に

50

支持されており、ハウジング 36 内には、台座 30 を回転させるモータが配置されている。

【0034】

第 1 円弧状アーム 32 および第 2 円弧状アーム 34 は半円状に形成され、同じ回転中心を有するように両端部が台座 30 に支持される。半円状の第 1 円弧状アーム 32 の径は、半円状の第 2 円弧状アーム 34 の径よりも僅かに大きく、第 1 円弧状アーム 32 は、第 2 円弧状アーム 34 の外周側に配置される。第 1 円弧状アーム 32 と第 2 円弧状アーム 34 は、台座 30 において直交するように配置されてよい。実施例では、第 1 円弧状アーム 32 が台座 30 に支持された両端部を結ぶラインと、第 2 円弧状アーム 34 が台座 30 に支持された両端部を結ぶラインとが直交する。挿通部材 42 は、第 1 貫通長孔 32a および第 2 貫通長孔 34a に挿通されて、第 1 貫通長孔 32a および第 2 貫通長孔 34a の交差位置に配置される。挿通部材 42 は、第 1 円弧状アーム 32 および第 2 円弧状アーム 34 の回転により、第 1 貫通長孔 32a 内および第 2 貫通長孔 34a 内を摺動する。

10

【0035】

図 6 は、挿通部材 42 の構成を示す。挿通部材 42 は、第 1 貫通長孔 32a および第 2 貫通長孔 34a の挿通状態を維持するように、第 1 貫通長孔 32a よりも幅広の第 1 規制部 42a と、第 2 貫通長孔 34a よりも幅広の第 2 規制部 42b とを備える。第 1 規制部 42a は第 1 貫通長孔 32a よりも上側に配置され、第 2 規制部 42b は第 2 貫通長孔 34a よりも下側に配置されて、挿通部材 42 が第 1 貫通長孔 32a および第 2 貫通長孔 34a から脱落することを防止する。挿通部材 42 を第 1 貫通長孔 32a および第 2 貫通長孔 34a に取り付けの際は、第 1 規制部 42a または第 2 規制部 42b のいずれか一方が軸部 42c とは別体に形成され、軸部 42c を第 1 貫通長孔 32a および第 2 貫通長孔 34a に挿入した状態で、軸部 42c の端部に固定する構造をとってもよい。

20

【0036】

軸部 42c は、第 1 貫通長孔 32a および第 2 貫通長孔 34a に挿入される部分であり、第 1 貫通長孔 32a および第 2 貫通長孔 34a の交差箇所に常時位置する。軸部 42c は、第 1 貫通長孔 32a 内および第 2 貫通長孔 34a 内において回転を規制される。実施例では軸部 42c が、第 1 貫通長孔 32a および第 2 貫通長孔 34a の幅よりも僅かに狭い幅をもつ矩形断面を有し、第 1 貫通長孔 32a 内および第 2 貫通長孔 34a 内で回転を規制されるが、それ以外の手段により軸部 42c の回転が規制されてよい。たとえば第 2 円弧状アーム 34 の内周面にレールが設けられ、第 2 規制部 42b にレール溝が設けられて、レールとレール溝とが嵌合することで軸部 42c の回転が規制されてもよい。第 1 規制部 42a には筐体 20 が取り付けられ、軸部 42c の回転が規制されることで、筐体 20 を所望の姿勢に維持することが可能となる。

30

【0037】

なお軸部 42c は、第 1 貫通長孔 32a および第 2 貫通長孔 34a の幅よりも狭い幅を有することで、第 1 貫通長孔 32a 内および第 2 貫通長孔 34a 内を摺動可能とする。これにより挿通部材 42 は、第 1 円弧状アーム 32 および第 2 円弧状アーム 34 の回転により、第 1 貫通長孔 32a に沿って移動でき、また第 2 貫通長孔 34a に沿って移動できる。

40

【0038】

図 7 は、ロボット 10 の断面を示す。図 7(a) は、第 1 円弧状アーム 32 と第 2 円弧状アーム 34 とが台座 30 に対して 90 度起立した状態で第 2 円弧状アーム 34 に沿って切断した断面を示し、図 7(b) は、第 1 円弧状アーム 32 と第 2 円弧状アーム 34 とが台座 30 に対して 90 度起立した状態で第 1 円弧状アーム 32 に沿って切断した断面を示す。

【0039】

第 1 モータ 52 は、第 1 円弧状アーム 32 を回転させるために設けられ、第 2 モータ 54 は、第 2 円弧状アーム 34 を回転させるために設けられる。第 1 モータ 52 および第 2 モータ 54 は、台座 30 上に配置されて、台座 30 が回転すると、第 1 モータ 52 および

50

第2モータ54も台座30とともに回転する。第3モータ56は、台座30を回転させるために設けられ、ハウジング36内に配置される。第1モータ52、第2モータ54および第3モータ56は、図示しない電源装置から電力を供給されて回転する。

【0040】

第1モータ52が第1円弧状アーム32を回転し、第2モータ54が第2円弧状アーム34を回転し、第3モータ56が台座30を回転することで、アクチュエータ装置12は、挿通部材42に取り付けられた筐体20の向きおよび姿勢を変化させられる。

【0041】

図8および図9は、ロボット10における筐体20の姿勢の例を示す図である。

図8(a)および(b)は、筐体20を左右方向に傾けた例を示す。図9(a)および(b)は、筐体20を前後方向に傾けた例を示す。このようにロボット10の駆動機構50は、筐体20に任意の姿勢をとらせることが可能となる。筐体20の姿勢は、第1モータ52および第2モータ54の駆動量を調整することで制御され、また筐体20の向きは、第3モータ56の駆動量を調整することで制御される。

10

【0042】

図10は、ロボット10の機能ブロックを示す。ロボット10は、外部からの入力を受け付けて処理する入力システム22と、外部への出力を処理する出力システム24とを備える。入力システム22は、受信部60、センサ情報取得部62、動き検出部64、視線方向決定部66、アクチュエータ制御部68、音声データ取得部70および音声処理部72を備える。また出力システム24は、画像処理部80、音声処理部82および送信部90を備える。

20

【0043】

図10において、さまざまな処理を行う機能ブロックとして記載される各要素は、ハードウェア的には、回路ブロック、メモリ、その他のLSIで構成することができ、ソフトウェア的には、メモリにロードされたプログラムなどによって実現される。したがって、これらの機能ブロックがハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現できることは当業者には理解されるところであり、いずれかに限定されるものではない。

【0044】

上記したようにHMD100は、姿勢センサ124が検出したセンサ情報およびマイク106が生成した音声信号を符号化した音声データをロボット10に送信し、受信部60は、センサ情報および音声データを受信する。音声データ取得部70は、受信した音声データを取得し、音声処理部72が音声処理を実施して、スピーカ18から出力する。これによりロボット10は、ユーザAの音声をリアルタイムで再生し、ロボット10の周囲の人が、ユーザAの声を聞くことができる。

30

【0045】

センサ情報取得部62は、HMD100の姿勢センサ124が検出した姿勢情報を取得する。動き検出部64は、ユーザAの頭部に装着されたHMD100の姿勢を検出する。視線方向決定部66は、動き検出部64により検出されたHMD100の姿勢に応じて筐体20のカメラ14の視線方向を定める。

【0046】

動き検出部64は、HMD100を装着したユーザの頭部の姿勢を検出するヘッドトラッキング処理を行う。ヘッドトラッキング処理は、ユーザの頭部の姿勢に、HMD100の表示パネル102に表示する視野を連動させるために行われ、実施例のヘッドトラッキング処理では、HMD100の水平基準方向に対する回転角度と、水平面に対する傾き角度とが検出される。水平基準方向は、たとえばHMD100の電源がオンされたときに向いている方向として設定されてよい。

40

【0047】

視線方向決定部66は、動き検出部64により検出されたHMD100の姿勢に応じて、視線方向を定める。この視線方向は、ユーザAの視線方向であり、ひいては分身であるロボット10のカメラ14の視線方向(光軸方向)である。

50

【0048】

カメラ14の視線方向(光軸方向)をユーザAの視線方向に連動させるために、ロボット10の基準姿勢を事前に設定しておく必要がある。図5には、第1円弧状アーム32と第2円弧状アーム34とが台座30に対して90度起立した状態を示しているが、この状態を水平方向として設定し、またロボット10の電源がオンされたときに筐体20の前面が向いている方向を、水平基準方向として設定してよい。なおロボット10は、HMD100と同様に姿勢センサを有して、水平方向を自律的に設定できるようにしてもよい。

【0049】

HMD100およびロボット10の基準姿勢を設定した状態で、視線方向決定部66は、動き検出部64により検出された回転角度および傾き角度を、そのままカメラ14の視線方向(光軸方向)として決定してよい。動き検出部64が、HMD100の回転角度および傾き角度を検出すると、視線方向決定部66は、HMD100の視線方向を3次元座標のベクトル(x, y, z)として決定し、このときロボット10のカメラ14の視線方向を同じ(x, y, z)と決定してもよく、また何らかの補正を加えた(x', y', z')として決定してもよい。

10

【0050】

アクチュエータ制御部68は、視線方向決定部66で決定された視線方向となるようにカメラ14の向きを制御する。具体的にアクチュエータ制御部68は、第1モータ52、第2モータ54、第3モータ56に供給する電力を調整して、HMD100の動きに、筐体20の動きを追従させる。アクチュエータ制御部68によるモータ駆動制御は、リアルタイムに実施され、したがって筐体20の向きは、ユーザAの視線の向きと同じように動かされる。

20

【0051】

実施例のアクチュエータ装置12によれば、筐体20は、第1円弧状アーム32および第2円弧状アーム34の回転中心を基準として駆動されるが、この動きは人の首と同じ動きを示す。アクチュエータ装置12は、2本の半円アームを交差させた簡易な構造でユーザAの首の動きを再現する。

【0052】

人は首の動きによって意思を伝達する。たとえば日本では首を縦に振ると肯定を、横に振ると否定を表現するが、アクチュエータ装置12は、ユーザAの首の動きと同じように筐体20を動かすため、ロボット10の周囲の人は、筐体20の動きによってもユーザAの意思を感じ取ることができる。そのためユーザAの首の動きを簡易な構造で再現できることは、トレイグジステンズ技術において非常に有用である。

30

【0053】

次に出力系統24について説明する。

出力系統24において、右カメラ14aおよび左カメラ14bは、アクチュエータ装置12により制御された方向に向けられて、それぞれの画角内を撮影する。右カメラ14aおよび左カメラ14bは、たとえば大人の平均的な両目の間隔となるように離れて配置されてよい。右カメラ14aが撮影した右目用画像データおよび左カメラ14bが撮影した左目用画像データは、送信部90からHMD100に送信されて、それぞれ表示パネル102の右半分および左半分に表示される。これらの画像は、右目および左目から見た視差画像を形成し、表示パネル102を2分割してなる領域にそれぞれ表示させることで、画像を立体視させることができる。なおユーザAは光学レンズを通して表示パネル102を見るために、画像処理部80は、予めレンズによる光学歪みを補正した画像データを生成して、HMD100に供給してもよい。

40

【0054】

右カメラ14aおよび左カメラ14bは、所定の周期(たとえば1/60秒)で撮影を行い、送信部90は遅延なく画像データをHMD100に送信する。これによりユーザAはロボット10の周囲の状況をリアルタイムで見ることができ、また顔の向きを変えることで、見たい方向を見ることができる。

50

【 0 0 5 5 】

右マイク 1 6 a および左マイク 1 6 b は、ロボット 1 0 の周囲の音を電気信号に変換して音声信号を生成する。以下、右マイク 1 6 a が生成する音声信号を「第 1 音声信号」と呼び、左マイク 1 6 b が生成する音声信号を「第 2 音声信号」と呼ぶ。上記したように、右マイク 1 6 a および左マイク 1 6 b は、筐体 2 0 において横方向に離れて配置されているため、右マイク 1 6 a が生成する第 1 音声信号と左マイク 1 6 b が生成する第 2 音声信号には、位相差が生じる。

【 0 0 5 6 】

本発明者は、第 1 音声信号および第 2 音声信号を、そのままの位相差で符号化して H M D 1 0 0 に提供した場合に、ユーザが、音源の方向を認識できない、つまり音声は右側から聞こえてくるのか、または左側から聞こえてくるのかを判別しにくいという知見を実験により得た。実験では筐体 2 0 の横方向の幅を大人の人間の顔幅程度 (1 6 c m) に設定しているが、人間の耳における音波伝達構造をマイク 1 6 では再現できないために、第 1 音声信号および第 2 音声信号の位相差のみでは、人間が音源の方向を知覚するには足りないという結論が得られた。

10

【 0 0 5 7 】

これを解決する手段として、筐体 2 0 の横方向の幅を大きくして、第 1 音声信号と第 2 音声信号の位相差を大きくすることが考えられるが、その場合は筐体 2 0 の重量が重くなり、アクチュエータ装置 1 2 で使用するモータの出力を高める必要が生じる。また筐体 2 0 の横方向の幅を大きくすると、右マイク 1 6 a と左マイク 1 6 b の間隔が、人間の両耳の間隔よりも広くなるため、実際に人が音を聞く感覚とは異なる音声信号が取得されることになる。

20

【 0 0 5 8 】

そこで本発明者は、第 1 音声信号と第 2 音声信号の位相差を増幅することで、この問題を解決することを考え出した。音声処理部 8 2 は、以下に説明するように、右マイク 1 6 a が生成する第 1 音声信号および左マイク 1 6 b が生成する第 2 音声信号の位相差を増幅する機能を有する。なおロボット 1 0 は、リアルタイムでマイク音声を H M D 1 0 0 に伝達する必要があるため、音声処理部 8 2 は、位相差増幅機能を、ハードウェア回路によって実現する。

【 0 0 5 9 】

図 1 1 は、音声処理部 8 2 が備える位相差増幅装置 8 2 a の回路構成を示す。位相差増幅装置 8 2 a は、右マイク 1 6 a が生成した第 1 音声信号 v_R と左マイク 1 6 b が生成した第 2 音声信号 v_L の位相差を増幅して出力するアナログ回路装置である。

30

【 0 0 6 0 】

第 1 増幅器 8 4 a は、右マイク 1 6 a から第 1 音声信号 v_R を入力されると、第 1 音声信号 v_R を増幅した第 1 正相信号 V_{R+} と、第 1 音声信号 v_R を反転増幅した第 1 逆相信号 V_{R-} とを出力する。第 1 増幅器 8 4 a は、入力信号の正相成分を増幅して出力するオペアンプと、入力信号の逆相成分を増幅して出力するオペアンプとから構成されてもよいが、正相成分および逆相成分を出力する 2 つの出力端子を有するオペアンプから構成されてもよい。

40

【 0 0 6 1 】

また第 2 増幅器 8 4 b は、左マイク 1 6 b から第 2 音声信号 v_L を入力されると、第 2 音声信号 v_L を増幅した第 2 正相信号 V_{L+} と、第 2 音声信号 v_L を反転増幅した第 2 逆相信号 V_{L-} とを出力する。第 2 増幅器 8 4 b も、第 1 増幅器 8 4 a と同様に、それぞれ正相成分および逆相成分を出力する 2 つのオペアンプから構成されてもよく、また正相成分および逆相成分の双方を出力する 1 つのオペアンプから構成されてもよい。

【 0 0 6 2 】

第 1 加算器 8 6 a は、第 1 正相信号 V_{R+} を第 1 係数倍 (倍) した信号と、第 2 逆相信号 V_{L-} を第 2 係数倍 (倍) した信号とを加算した出力信号 V_{ROUT} を出力する。ここで、 α は、0 より大きく、1 以下の値を示す。なお β は異なるように設定され

50

、この例では α である。出力信号 V_{ROUT} は、以下の式で表現される。

$$V_{ROUT} = \alpha \times V_{R^+} + \beta \times V_{L^-}$$

【0063】

第1加算器86aは、第1正相信号 V_{R^+} を α 倍に分圧する分圧回路の出力と、第2逆相信号 V_{L^-} を β 倍に分圧する分圧回路の出力とを加算する加算回路であってもよいが、第1正相信号 V_{R^+} を α 倍した電圧信号と第2逆相信号 V_{L^-} を β 倍した電圧信号とを加算するオペアンプであってもよい。

【0064】

第2加算器86bは、第2正相信号 V_{L^+} を第1係数倍 (γ 倍) した信号と、第1逆相信号 V_{R^-} を第2係数倍 (δ 倍) した信号とを加算した出力信号 V_{LOUT} を出力する。出力信号 V_{LOUT} は、以下の式で表現される。

$$V_{LOUT} = \gamma \times V_{L^+} + \delta \times V_{R^-}$$

【0065】

第2加算器86bは、第2正相信号 V_{L^+} を γ 倍に分圧する分圧回路の出力と、第1逆相信号 V_{R^-} を δ 倍に分圧する分圧回路の出力とを加算する加算回路であってもよいが、第2正相信号 V_{L^+} を γ 倍した電圧信号と第1逆相信号 V_{R^-} を δ 倍した電圧信号とを加算するオペアンプであってもよい。

【0066】

第3増幅器88aは、第1加算器86aの出力信号 V_{ROUT} を第3係数倍 (ϵ 倍) して V_{ROUT} を出力し、第4増幅器88bは、第2加算器86bの出力信号 V_{LOUT} を第3係数倍 (ζ 倍) して V_{LOUT} を出力する。音声処理部82において、位相差増幅装置82aからの出力信号 V_{ROUT} 、 V_{LOUT} は、それぞれ音声符号化されて、右耳用音声データおよび左耳用音声データとして送信部90からHMD100に送信される。

【0067】

図12は、信号波形の位相差を説明するための図である。図12(a)は、右マイク16aが生成する第1音声信号 v_R と左マイク16bが生成する第2音声信号 v_L の波形の関係を示す。ここでは説明の便宜上、第1音声信号 v_R と第2音声信号 v_L とをそれぞれ同倍に増幅した第1正相信号 V_{R^+} と第2正相信号 V_{L^+} との関係を示している。この入力波形では、ロボット10の筐体20から見て音源が右側に配置されており、第1正相信号 V_{R^+} の位相の方が、第2正相信号 V_{L^+} よりも僅かに進んでおり、また振幅は第1正相信号 V_{R^+} の方が高い。

【0068】

図12(b)は、第1加算器86aの出力信号 V_{ROUT} と第2加算器86bの出力信号 V_{LOUT} の波形の関係を示す。図12(a)に示す入力波形の位相差と比較すると、図12(b)に示す加算器の出力波形の位相差が広がっている(増幅している)ことが分かる。

【0069】

図13は、入力信号波形の位相差を増幅する原理を説明するための図である。図13(a)は、第1正相信号 V_{R^+} および第1逆相信号 V_{R^-} と、第2正相信号 V_{L^+} および第2逆相信号 V_{L^-} を2次元座標系で表現している。第1正相信号 V_{R^+} と第2正相信号 V_{L^+} の位相差は θ である。

【0070】

図13(b)は、第1加算器86aの出力信号 V_{ROUT} と第2加算器86bの出力信号 V_{LOUT} を示す。上記したように、 V_{ROUT} 、 V_{LOUT} は、

$$V_{ROUT} = \alpha \times V_{R^+} + \beta \times V_{L^-}$$

$$V_{LOUT} = \gamma \times V_{L^+} + \delta \times V_{R^-}$$

と表現される。図13(b)では、 $\alpha = 1.0$ 、 $\beta = 0.6$ を設定する。

【0071】

図13(b)に示すように、 V_{ROUT} と V_{LOUT} の位相差は θ' となり、図13(a)に示す位相差 θ よりも大きくなっている。このように位相差増幅装置82aは、入力

10

20

30

40

50

された2つの音声信号の位相差を増幅する。

【0072】

本発明者によるシミュレーションの結果、入力信号の位相差が15度のとき、出力信号の位相差は4倍の60度となり、入力信号の位相差が30度のとき、出力信号の位相差は3倍の90度となり、入力信号の位相差が45度のとき、出力信号の位相差は約2.7倍の120度となることが分かった。

【0073】

このシミュレーション結果によると、位相差が小さいほど増幅率が大きくなっている。実際の筐体20では、入力信号の位相差は5度~20度程度であり、位相差増幅装置82aは、この範囲における増幅率を大きくできることで、出力信号の位相差を、ユーザが音源の方向を聞き分けられる程度に広げられる。位相差増幅装置82aからの出力信号 $V_{R\text{OUT}}$ 、 $V_{L\text{OUT}}$ は、それぞれ音声符号化されて、右耳用音声データおよび左耳用音声データとして送信部90からHMD100に送信される。

【0074】

HMD100において、右耳用音声データは、右耳用のイヤホン104から音声として出力され、左耳用音声データは、左耳用のイヤホン104から音声として出力される。ユーザAは、位相差を増幅された音声を両耳から聞くことで、音源の方向を認識する。ユーザAは、右側から声が聞こえてきたと感じれば、顔を右側に向ける。このときユーザAの顔の動きに連動してロボット10の筐体20が右側を向くため(図2参照)、ロボット10のカメラ14は、右側の環境を撮影して、撮影画像データをリアルタイムでHMD100に送信する。これによりユーザAは、発声した人の顔を見ながら話すことができ、従来にはない優れたユーザインタフェースを実現できる。

【0075】

なお上記した例では、 $\alpha = 1.0$ 、 $\beta = 0.6$ と設定したが、 α 、 β の値は、実験により適切に設定されることが好ましい。図5に示すように、右マイク16aおよび左マイク16bは、筐体20の側面を窪ませた位置であって、前面からみて奥側の位置に設けている。マイク16における音波の伝達構造は、筐体側面の形状に依存するため、 α 、 β の比は、実験により最適に求められることが好ましい。

【0076】

なお図5において、マイク16は、後板17の横方向の内側に配置されている。これは後板17に、前方からの音波と後方からの音波の周波数特性を異ならせ、後方からの高域成分を低減させる役割をもたせるためである。つまり後板17は、マイク16に対して人の耳介のような機能をもち、後方からの音波が後板17を回り込んでマイク16に到達するようにしている。なお前方からの音波と後方からの音波の周波数特性を異ならせるために、後板17は、さらに上下方向および横方向に広げられて形成されてもよい。マイク16の後方に後板17のような音波遮蔽体を形成することで、ユーザAは、音源の前後方向の位置を聞き分けることも可能となる。

【0077】

このように情報処理システム1では、ユーザAが、自分の分身であるロボット10を用いて、リアルタイムでロボット10の周囲にいる人達と自由にコミュニケーションをとることができる。以下では、情報処理システム1の利用可能性をさらに高める技術について提案する。

【0078】

従来より、カメラの傾きを変えながら撮影した画像をステッチ(縫い合わせ)して全天球パノラマ画像を生成する技術が知られている。最近では、専用のパンチルトカメラも販売されており、個人でも全天球パノラマ画像を撮影できるようになっている。

【0079】

情報処理システム1において、ロボット10は、ユーザAの頭部の動きに応じた視線方向にカメラ14を向けて、周囲を撮影する。ユーザAが様々な方向を向くことで、カメラ14が様々な方向を撮影する。この撮影画像に、視線方向を表現する3次元ベクトルを付

10

20

30

40

50

加して記録しておくことで、仮想的な全天球パノラマ画像を生成することが可能となる。

【0080】

図14は、ロボット10の機能ブロックの変形例を示す。この機能ブロックは、図10に示す機能ブロックを前提としており、その中で視線方向決定部66から画像処理部80に対して、決定した視線方向が供給されることを示している。

【0081】

ユーザAによるロボット10の使用時、送信部90は、両眼用の画像データおよび両耳用の音声データ（以下、まとめて「視聴データ」と呼ぶこともある）を、ネットワーク4経由でユーザAのHMD100に送信している。このとき送信部90は、同じ視聴データをネットワーク4経由でルータ5を介して処理装置200にも送信し、処理装置200はユーザAの視聴データを記録する。

10

【0082】

処理装置200は、ユーザAの視聴データを記録しつつ、ユーザAの画像データをもとに全天球パノラマ画像をリアルタイム生成し、ユーザAとは異なるユーザBの視線方向に応じた画像をユーザBのHMD100aに提供する機能をもつ。なおHMD100aは、これまで説明したHMD100と同じ構成を備える。処理装置200は、たとえば単一のサーバにより構成されてもよいが、クラウドサービスを提供するサーバ群により構成されてもよい。

【0083】

処理装置200が全天球パノラマ画像を生成できるようにするために、画像処理部80は、フレーム画像データのそれぞれに、視線方向決定部66から供給される視線方向を示すベクトル情報と、撮影開始点からの経過時間を示す撮影時間情報とを付加する。ベクトル情報は、ロボット10のカメラ14の視線方向を示す。撮影時間情報は、撮影開始点からの時間を表現するものであればよく、たとえば撮影された順番を示すフレーム番号であってもよい。

20

【0084】

この技術では、ユーザAによるロボット10の使用時に、ユーザBがHMD100aを装着し、ロボット10から供給されるユーザAの視聴データをもとに生成される画像データおよび音声データを、HMD100aに提供する。ユーザAの視聴データをそのまま再生するだけであれば、処理装置200は、受信した視聴データをそのままユーザBのHMD100aにストリーミング配信するだけでよいが、この技術では、処理装置200が、ユーザAの画像データをもとに構成される全天球パノラマ画像から、ユーザBの視線方向にもとづいた画像を再構成して、ユーザBのHMD100aに提供できるようにする。なお音声データは、ユーザBのHMD100aにストリーミング配信される。

30

【0085】

図15は、処理装置200の機能ブロックを示す。処理装置200は、受信部202、センサ情報取得部204、動き検出部206、視線方向決定部208、画像決定部210、音声決定部212、視聴データ提供部214、送信部216および記録部218を備える。記録部218は画像記録部220および音声記録部222を含む。受信部202が、ロボット10から送信された視聴データを受信すると、画像記録部220は、受信した画像データを順次記録し、音声記録部222は、受信した音声データを順次記録する。なお画像データは、フレーム画像ごとに、撮影時のベクトル情報および撮影時間情報を付加されている。

40

【0086】

ユーザBは、HMD100aを通じて、処理装置200に、ユーザAの視聴データの再生指示を送信する。処理装置200は、再生指示を受け付けると、視聴データの再生処理を開始する。音声決定部212はユーザBに提供する音声データを決定し、音声記録部222に記録された音声データを、音声記録部222からただちに読み出し視聴データ提供部214に提供する。つまり音声決定部212は、ロボット10から提供される音声データを、HMD100aにストリーミング配信する。したがってユーザBは、ユーザAが聞

50

いている音声と同じ音声をHMD 100 aのイヤホン104から聞くことができる。

【0087】

処理装置200による再生処理中、受信部202は、ユーザBが装着したHMD 100 aから送信されるセンサ情報を受信し、センサ情報取得部204は、受信したセンサ情報を取得する。このセンサ情報は、姿勢センサ124がHMD 100 aの姿勢を検出した姿勢情報である。動き検出部206は、ユーザBの頭部に装着されたHMD 100 aの姿勢を検出する。視線方向決定部208は、動き検出部206により検出されたHMD 100 aの姿勢に応じて、全天球パノラマ画像における仮想カメラの視線方向を定める。画像決定部210はユーザBに提供する画像データを決定し、画像記録部220に記録された複数の画像データを用いて、決定された視線方向に向けた仮想カメラにより撮影される画像を合成して画像データを生成する。

10

【0088】

視聴データ提供部214は、画像決定部210で決定された画像データと、音声決定部212で決定された音声データとを合わせた視聴データを、送信部216からユーザBのHMD 100 aに提供する。

【0089】

図15において、さまざまな処理を行う機能ブロックとして記載される各要素は、ハードウェア的には、回路ブロック、メモリ、その他のLSIで構成することができ、ソフトウェア的には、メモリにロードされたプログラムなどによって実現される。したがって、これらの機能ブロックがハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現できることは当業者には理解されるところであり、いずれかに限定されるものではない。

20

【0090】

処理装置200は、全方位のパノラマ画像を生成する。したがってユーザBが首を左または右に回して水平方向の視線を左または右に回転させることで、左方向または右方向のパノラマ画像がHMD 100 aの表示パネル102に表示され、またユーザBが首を上または下に傾けて、垂直方向に視線を傾けることで、上方向または下方向のパノラマ画像がHMD 100 aの表示パネル102に表示されるようになる。

【0091】

図16は、処理装置200が生成する全天球パノラマ画像を説明するための図である。この技術では、ユーザBが球体の中心に位置し、視線の方向を変更することで、見える画像が変更される仮想環境を実現する。画像決定部210は、画像記録部220に記録されている画像データをステッチ（縫い合わせ）して、全天球パノラマ画像を生成する。

30

【0092】

実施例では、説明を単純にするためにロボット10がカメラ14をズームせず、一定の拡大率で画像データを取得する。そのため画像決定部210は、画像データに付加されたベクトル情報にもとづいて、画像データを全天球の内周面に張り合わせることで、全天球パノラマ画像を構成する。なお複数の画像データが重複する箇所については、最新の画像データで上書きし、これによりリアルタイムの状況に近い全天球パノラマ画像を構成できるようになる。

40

【0093】

なお実際の画像決定部210の画像生成処理は、処理負荷を軽減するために、常に全天球パノラマ画像を再構成するのではなく、ユーザBが位置する中心点9から撮影されるフレーム画像7を動的に生成する処理となる。このとき画像決定部210は、仮想カメラ8の撮影範囲（画角）を、実際のロボット10のカメラ14の撮影範囲（画角）に対応するように設定することが好ましい。これによりユーザAの視線方向とユーザBの視線方向とが一致しているタイミングでは、ユーザBは、ユーザAと同じ画像を見られるようになる。

【0094】

このように画像決定部210は、画像データにメタデータとして設定されているベクト

50

ル情報を用いて、イメージスティッチング処理を実施し、ユーザBの視線方向から定まる撮影範囲のフレーム画像7を生成する。動き検出部206は、ユーザBのヘッドトラッキング処理を行うことで、ユーザBの頭部（実際にはHMD100a）の回転角度および傾きを検出する。ここでHMD100aの回転角度は、水平面の基準方向に対する回転角度であり、基準方向は、たとえばHMD100aの電源がオンされたときに向いている方向として設定されてよい。またHMD100aの傾きは、水平面に対する傾斜角度である。ヘッドトラッキング処理として既知の技術が利用されてよく、動き検出部206は、HMD100aの姿勢センサが検出したセンサ情報から、HMD100aの回転角度および傾きを検出する。

【0095】

視線方向決定部208は、検出したHMD100aの回転角度および傾きにしたがって、仮想球体における仮想カメラ8の姿勢を定める。仮想カメラ8は、仮想球体の中心点9から仮想球体の内周面を撮影するように配置されており、視線方向決定部208は、仮想カメラ8の光軸の向きを、ロボット10のカメラ14の光軸方向と一致するように決定してもよい。

【0096】

ロボット10において、視線方向決定部66は、ユーザAのHMD100の視線方向を3次元座標のベクトル (x, y, z) として決定すると、ロボット10のカメラ14の視線方向を同じ (x, y, z) と決定してもよいことを説明した。処理装置200においても、視線方向決定部208は、ユーザBのHMD100aの視線方向を3次元座標のベクトル (x, y, z) として決定すると、仮想カメラ8の視線方向を同じ (x, y, z) と決定してもよい。また視線方向決定部66において、HMD100の視線方向を所定の変換式で補正してカメラ14の視線方向を求める場合には、視線方向決定部208においても、HMD100aの視線方向を同じ変換式で補正して仮想カメラ8の視線方向を求めてもよい。このように、それぞれの3次元座標系を取り扱うことで、ユーザAの視線方向とユーザBの視線方向とが一致しているタイミングでは、ユーザBは、ユーザAと同じ画像を見られるようになる。

【0097】

画像決定部210は、仮想カメラ8のフレーム画像7を生成すると、光学レンズ用の光学歪み補正を施し、視聴データ提供部214に画像データを供給する。なお図16においては1つの仮想カメラ8が示されているが、実際には左目用と右目用の2つの仮想カメラ8が配置されて、それぞれの画像データが、ロボット10から提供される左目用画像データおよび右目用画像データにもとづいて生成される。

【0098】

図17は、画像記録部220に記録されている撮影画像データを説明するための図である。ここでは説明の便宜上、片目用の複数の画像データを示し、またユーザBの視線方向に対して適切なアフィン変換を施した状態の画像データを2次元平面上に配置している。なおユーザBの視線方向については後述する。

【0099】

画像決定部210は、各撮影画像の重複する部分をつなぎ合わせて、全天球パノラマ画像を生成する機能をもつ。撮影画像をつなぎ合わせる技術については、たとえば同一出願人による特許第5865388号にも記載されているように既知の技術を利用してよい。以下では、画像記録部220に記録された複数の撮影画像データのうち、いずれの撮影画像データを用いるかを選択する手法について説明する。

【0100】

図17に、5つの画像データI1~I5を示す。各画像データに含まれる (x, y, z) は、撮影時のカメラ14の視線方向（ベクトル情報）を表現し、“t”は、撮影時間情報を表現する。ここで画像データI1は、ベクトル情報 (x_1, y_1, z_1) および撮影時間情報t1を付加情報として有する。同様に画像データI2は、ベクトル情報 (x_2, y_2, z_2) および撮影時間情報t2を付加情報として有し、画像データI3は、ベクトル

10

20

30

40

50

ル情報 (x_3, y_3, z_3) および撮影時間情報 t_3 を付加情報として有し、画像データ I 4 は、ベクトル情報 (x_4, y_4, z_4) および撮影時間情報 t_4 を付加情報として有し、画像データ I 5 は、ベクトル情報 (x_5, y_5, z_5) および撮影時間情報 t_5 を付加情報として有する。

【 0 1 0 1 】

なお付加情報である撮影時間情報 $t_1 \sim t_5$ は、撮影開始点 (時間 0) からの経過時間を表現し、 $t_1 < t_2 < t_3 < t_4 < t_5$ の関係にある。したがって画像データ I 1 ~ I 5 の中では、画像データ I 1 が最初に撮影され、画像データ I 5 が最後に撮影されている。画像決定部 2 1 0 は、撮影時間情報と、視線方向決定部 2 0 8 が決定した仮想カメラ 8 の視線方向にもとづいて、合成画像を生成するための画像データを選択する。

10

【 0 1 0 2 】

具体的に画像決定部 2 1 0 は、視線方向決定部 2 0 8 が決定した仮想カメラ 8 の視線方向、つまり HMD 1 0 0 a を装着したユーザ B が向いている方向から、全天球パノラマ画像から切り取る撮影範囲 (仮想カメラ 8 の画角) を定め、撮影範囲内に含まれる画像を含む画像データを、画像データに付加されたベクトル情報にもとづいて抽出する。

【 0 1 0 3 】

図 1 8 は、画像決定部 2 1 0 が生成すべきフレーム画像 7 と、画像データとの関係を示す図である。図 1 7 および図 1 8 において、各画像データ I 1 ~ I 5 は、各ベクトル情報にもとづいて、仮想カメラ 8 の視線方向 (X, Y, Z) に直交する 2 次元平面上に写像されており、各画像データ I 1 ~ I 5 の位置は、2 次元平面において 4 つの頂点座標で定義される。画像決定部 2 1 0 は、仮想カメラ 8 の視線方向 (X, Y, Z) により、全天球パノラマ画像における仮想カメラ 8 の画角の位置 (撮影範囲) を定め、視線方向に直交する 2 次元平面のフレーム画像 7 の 4 つの頂点座標を定める。画像決定部 2 1 0 は、画像記録部 2 2 0 に記録された画像データのうち、フレーム画像 7 に含まれる画像データを抽出する。図示されるように画像データ I 1 ~ I 5 は、フレーム画像 7 に含まれる画像を含んでいるため、仮想カメラ 8 の撮影範囲内に含まれる画像データとして抽出される。

20

【 0 1 0 4 】

画像決定部 2 1 0 は、複数の画像データが重複する領域については、遅い撮影時間情報をもつ画像データを優先して用いて合成画像を生成する。図 1 8 に示す例では、フレーム画像 7 に、撮影時間の早い画像データから順に、つまり画像データ I 1 から順に書き込み、新しい画像データで順次上書きしていくようにフレーム画像 7 を合成する。

30

【 0 1 0 5 】

このように画像決定部 2 1 0 は、複数の画像データのうち重複する領域については、より現在時刻に近い撮影時間情報をもつ画像データを用いて、合成画像を生成するようにする。たとえば撮影範囲内に含まれる画像で、画像データ I 4 と画像データ I 5 に重複する部分が存在した場合には、より遅い撮影時間をもつ画像データ I 5 を重複部分に埋め込む。これにより、現在時刻に近い画像データを用いて合成画像を生成できるようになり、ユーザ B に対して、現在時刻に近い合成画像を提供できるようになる。

【 0 1 0 6 】

この画像再生アプリケーションでは、ユーザ B が向く方向によっては、画像データが不足して、フレーム画像 7 の生成が困難な場合がある。特にロボット 1 0 が撮影を開始した直後は、画像データ数がそもそも少ないため、画像決定部 2 1 0 が、ユーザ B の視線方向に応じたフレーム画像 7 を生成できないことも生じる。実際には起こりえないが、撮影期間中、ユーザ A が HMD 1 0 0 を全く動かさなかった場合には、画像記録部 2 2 0 に記録されている画像データのベクトル情報は全て同じとなるため、たとえばユーザ B が、ユーザ A と真逆の方向を向くと、当該視線方向の仮想カメラ 8 の撮影範囲内に含まれる画像データは存在しない。

40

【 0 1 0 7 】

そのような場合、画像決定部 2 1 0 は、受信したユーザ A の画像データに、ユーザ B の視線方向の画像を生成できない旨のメッセージを重畳した画像データを生成して、視聴デ

50

ータ提供部 214 から HMD 100a に提供してもよい。たとえばユーザ B の視線方向の画像を所定割合（たとえば 30%）以上合成できない場合に、画像決定部 210 は合成画像の生成を行わず、上記メッセージとともに、ユーザ A が見た画像データを視聴データ提供部 214 に供給してもよい。

【0108】

また画像決定部 210 は、複数の画像データからフレーム画像 7 を合成するため、生成されたフレーム画像 7 は、つぎはぎ画像となり、視認性が悪くなる場合もある。そのため、たとえば撮影範囲内の所定割合（たとえば 50%）の画像を 1 つの画像データから形成できない場合には、上記したように、画像決定部 210 は、ユーザ B の視線方向の画像を生成できない旨のメッセージを、ユーザ A の画像データに重畳した画像データを生成して

10

【0109】

上記した例では、画像決定部 210 が、遅い撮影時間情報をもつ画像データを優先して用いて合成画像を生成することを説明したが、より早い撮影時間情報をもつ画像データを用いることでフレーム画像 7 の所定割合以上を構成できる場合には、より早い撮影時間情報をもつ画像データを用いてもよい。

【0110】

また時間が経過すると、ロボット 10 の撮影している環境に変化が生じるため、過去の画像データを用いた合成画像をユーザ B に提供することが好ましくないことも考えられる。そのため画像決定部 210 は、所定時間以上前の画像データについては、合成画像に含めないように、画像抽出処理を行ってもよい。

20

【0111】

以上は、ユーザ B が、ユーザ A の視聴データをリアルタイムで利用する例を示した。以下は、その応用技術について説明する。応用技術では、処理装置 200 が、ユーザ A の視聴データのリアルタイム再生を目的とするのではなく、二次利用を目的として視聴データを記録する。

【0112】

視聴データの二次利用のために、ロボット 10 において画像処理部 80 は、フレーム画像データのそれぞれに、撮影時間情報とベクトル情報を付加し、また音声処理部 82 は、音声データに、録音開始点からの経過時間を示す録音時間情報を付加する。なおカメラ 14 による撮影（録画）およびマイク 16 による録音は同じタイミングで開始されるため、撮影開始点および録音開始点は同じタイミングを示す。撮影時間情報および録音時間情報は、ロボット 10 におけるクロック生成部で生成された時刻情報であってよい。画像データおよび音声データに付加情報を付加する形式は何であってもよく、処理装置 200 が、再生用の視聴データを生成する際に参照できる形式であればよい。

30

【0113】

この応用技術では、ユーザ A がロボット 10 の使用を終了した後、別のユーザ B（ユーザ A であってもよい）が HMD 100a を装着して、処理装置 200 に記録されたユーザ A の視聴データをもとに生成される画像データおよび音声データを、HMD 100a に提供する。このとき実施例で説明したように、処理装置 200 は、ユーザ A の視聴データをもとに全天球パノラマ画像を構成し、全天球パノラマ画像からユーザ B の視線方向にもとづいた画像を再構成して、ユーザ B の HMD 100a に提供できるようにする。この利用環境では、ロボット 10 は使用しない。

40

【0114】

図 15 を参照して、画像記録部 220 は、ロボット 10 から送信された画像データを記録しており、音声記録部 222 は、ロボット 10 から送信された音声データを記録している。この応用技術においては、画像記録部 220 および音声記録部 222 は、ロボット 10 からユーザ A に対して送信された視聴データの全てが記録済みの状態にある。なお画像データは、撮影時間情報と、撮影時のベクトル情報とを付加されており、音声データは、録音時間情報を付加されている。

50

【0115】

ユーザ B は、HMD 100 a を通じて、処理装置 200 に、ユーザ A の視聴データの再生指示を送信する。処理装置 200 は、再生指示を受け付けると、視聴データの再生処理を開始する。なお記録部 218 が 1 時間分の視聴データを記録している場合、ユーザ B は、1 時間の範囲内で、任意の時間から再生を開始できるようにしてもよい。この場合、受信部 202 は、ユーザ B から時間指定を受け付け、画像決定部 210 および音声決定部 212 に供給する。

【0116】

音声決定部 212 は、再生開始点からの経過時間を示す再生時間情報に対応する録音時間情報をもつ音声データを、音声記録部 222 から読み出し、視聴データ提供部 214 に提供する。再生開始点は、視聴データの再生開始点を意味し、したがって撮影開始点および録音開始点と同じタイミングを示す。音声決定部 212 は再生時間情報に録音時間情報が一致する音声データを音声記録部 222 から読み出し、視聴データ提供部 214 に提供する。

10

【0117】

処理装置 200 による再生処理中、受信部 202 は、ユーザ B が装着した HMD 100 a から送信されるセンサ情報を受信し、センサ情報取得部 204 は、受信したセンサ情報を取得する。このセンサ情報は、姿勢センサ 124 が HMD 100 a の姿勢を検出した姿勢情報である。動き検出部 206 は、ユーザ B の頭部に装着された HMD 100 a の姿勢を検出する。視線方向決定部 208 は、動き検出部 206 により検出された HMD 100 a の姿勢に応じて、仮想カメラの視線方向を定める。画像決定部 210 は、画像記録部 220 に記録された複数の画像データを用いて、決定された視線方向に向けた仮想カメラにより撮影される画像を合成する。視聴データ提供部 214 は、画像決定部 210 で合成された画像データと、音声決定部 212 で読み出された音声データとを合わせた視聴データを、送信部 216 から HMD 100 a に提供する。

20

【0118】

画像決定部 210 は、ユーザ B による視聴データの再生時間以前にユーザ A が見た画像をステッチ（縫い合わせ）して、ユーザ B が位置する中心点 9 から撮影されるフレーム画像 7 を動的に生成する。

【0119】

ユーザ B による視聴データの再生時間以前にユーザ A が見た画像について説明する。画像記録部 220 に、撮影開始点から 1 時間分の画像データが記録されている場合、ユーザ B による再生開始点からの再生時間は、1 時間以内のどこかのタイミングで特定される。たとえば再生時間が再生開始から 15 分のタイミングである場合、15 分以内の撮影時間情報が付加された画像、つまり撮影開始点から 15 分が経過するまでに撮影された画像が、再生時間以前にユーザ A が見た画像となる。つまり再生開始から 15 分の時点を再生しているのであれば、画像決定部 210 は、撮影開始から 15 分以内の撮影時間情報が付加された画像データを用いてフレーム画像 7 を生成し、再生開始から 45 分の時点を再生しているのであれば、画像決定部 210 は、撮影開始から 45 分以内の撮影時間情報が付加された画像データを用いてフレーム画像 7 を生成する。

30

40

【0120】

図 18 を参照して、画像決定部 210 は、再生時間情報以前の撮影時間情報を付加された画像データを抽出するようにし、再生時間情報よりも後の撮影時間情報を付加された画像データを抽出しないようにする。たとえば、フレーム画像 7 を再生する時間情報が時間 t_3 より後であって、時間 t_4 より前であれば、画像決定部 210 は、画像データ I1 ~ I3 を抽出し、画像データ I4、I5 は抽出しない。このように、再生時間情報以前の撮影時間情報を付加された画像データを用いて合成画像を生成することで、画像決定部 210 は、再生時間よりも後に撮影された画像をユーザ B に見せないようにする。

【0121】

視聴データ提供部 214 は、再生時間に対応する録音時間情報をもつ音声データを H M

50

D 1 0 0 a に送信しているため、ユーザ B は、再生時間に同期した音声を聞いている。そのため再生時間以前の状況については概ね承知しており、提供される画像データが再生時間以前の画像データから合成されたものであれば、どのような状況が表示されているかを把握できる。しかしながら提供される画像データが、再生時間より後の画像データから合成されていれば、ユーザ B は承知していない画像を見せられることになり、ユーザ B を混乱させることが予想される。そこで画像決定部 2 1 0 は、再生時間よりも後に撮影された画像をユーザ B に見せないようにする。

【 0 1 2 2 】

なお画像決定部 2 1 0 は、複数の画像データのうち重複する部分については、再生時間情報に近い撮影時間情報をもつ画像データを用いて、合成画像を生成するようにする。たとえば撮影範囲内に含まれる画像で、画像データ I 1 と画像データ I 2 に重複する部分が存在した場合には、より後に撮影された画像データ I 2 を重複部分に埋め込む。これにより、再生時間情報に近い画像データを用いて合成画像を生成できるようになり、ユーザ B には、再生時間の直近の画像データから合成した画像を提供できるようになる。

10

【 0 1 2 3 】

以上、本発明を実施例をもとに説明した。実施例は例示であり、それらの各構成要素や各処理プロセスの組合せにいろいろな変形例が可能なこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

【 0 1 2 4 】

実施例では、画像決定部 2 1 0 が、イメージスティッチング処理を実施し、ユーザ B の視線方向から定まる撮影範囲のフレーム画像 7 を生成することを説明した。変形例では、イメージスティッチング処理を実施することなく、画像決定部 2 1 0 が、仮想カメラ 8 の視線方向と、画像記録部 2 2 0 に記録された画像データに付加されたベクトル情報とにもとづいて、ユーザ B に提供する画像データを決定する。

20

【 0 1 2 5 】

この変形例で画像決定部 2 1 0 は、仮想カメラ 8 の視線方向に対応するベクトル情報を付加された画像データを、ユーザ B に提供する画像データとして決定する。仮想カメラ 8 の視線方向に対応するベクトル情報とは、仮想カメラ 8 の視線方向と一致するベクトル情報、および仮想カメラ 8 の視線方向と実質的に一致するとみなせるベクトル情報を含む。具体的には、仮想カメラ 8 の視線方向とベクトル情報とが所定の角度以内（たとえば 1 0 度）にある場合に、画像決定部 2 1 0 は、仮想カメラ 8 の視線方向とベクトル情報とが実質的に一致することを判定してもよい。

30

【 0 1 2 6 】

ユーザ A の視聴データを同期再生する場合、画像決定部 2 1 0 は、仮想カメラ 8 の視線方向に対応するベクトル情報を付加された画像データのうち、最も遅い撮影時間情報をもつ画像データを、ユーザ B に提供する画像データとして決定する。これによりユーザ B に対して、現在時刻に近い画像を提供できるようになる。

【 0 1 2 7 】

仮想カメラ 8 の視線方向に対応するベクトル情報を付加された画像データが画像記録部 2 2 0 に記録されていない場合、画像決定部 2 1 0 は、高さ方向（z 軸方向）の成分を除いた（x、y）成分が実質的に一致するとみなせるベクトル情報を付加された画像データを、ユーザ B に提供する画像データとして決定してもよい。一致するとみなせるベクトル情報とは、（x、y）成分が所定角度以内（たとえば 7 度）の範囲にあるベクトル情報である。（x、y）成分のみの近似をみることで、画像決定部 2 1 0 は、仮想カメラ 8 の視線方向に対応するベクトル情報を付加された画像データを見つけやすくし、これによりユーザ B に画像データを提供できない事態を回避できるようになる。

40

【 0 1 2 8 】

ユーザ A の視聴データを二次利用する場合、画像決定部 2 1 0 は、ユーザ B による視聴データの再生時間以前にユーザ A が見た画像から、ユーザ B に提供する画像データを決定する。つまり画像決定部 2 1 0 は、再生時間情報以前の撮影時間情報を付加された画像デ

50

ータの中から、仮想カメラ 8 の視線方向に対応するベクトル情報を付加された画像データを決定する。このとき、該当する画像データが複数存在すれば、画像決定部 210 は、再生時間情報に近い撮影時間情報をもつ画像データを選択することが好ましい。

【0129】

なおユーザ A が横方向に首を回したときの画像データが画像記録部 220 に記録されているケースを検討する。ユーザ B がユーザ A から少し遅れて、ユーザ A と逆方向に首を回すと、HMD 100a に、ユーザ A が見た画像が逆再生されることが生じうる。この場合、画像データの時系列が逆転することになり、ユーザ B に違和感を生じさせる可能性がある。そのためユーザ B が連続的に視線方向を変化させる際に、ユーザ A の画像データを逆再生させることになる場合には、画像決定部 210 は、画像データを固定して、提供する画像データを変化させないようにしてもよい。

10

【0130】

情報処理システム 1 の有用性を高めるために、ロボット 10 は、触覚センサや振動センサなど、外部からの入力を受け付ける入力センサをさらに備えてもよい。図 10 に示す機能ブロックにおいて、入力センサは、出力系統 24 に設けられ、入力センサのセンサ情報は、送信部 90 から HMD 100 に送信される。HMD 100 は、センサ情報を出力する出力手段を備え、センサ情報を振動などに変換してユーザ A に伝達してもよい。

【0131】

また情報処理システム 1 では、ロボット 10 が、ユーザ A の首の動きに筐体 20 を連動させることを説明したが、さらにユーザ A の表情などの状態を伝達する手段を有してもよい。たとえば HMD 100 は、装着したユーザ A の目や眉の動きを検出するセンサや、声の調子を解析する手段などを備える。目や眉の動きは、ユーザの表情を表現するものであり、また声の調子は、ユーザの心理状態を表現する。目や眉の動きおよび / または声の調子に関する情報は、HMD 100 からロボット 10 に送信され、ロボット 10 は、筐体 20 に設けた表情ユニットを駆動して、ユーザ A の表情、心理状態などを再現してもよい。表情ユニットは、筐体 20 の前面においてカメラ 14 の上部に形成した駆動部（たとえば眉の形状を模したもの）であってよく、HMD 100 から送信された情報をもとに、駆動部が駆動される。また表情ユニットは、ユーザ A の表情や心理状態を色で表現するディスプレイであってよく、表示色を変化させることで、ユーザ A の表情や心理状態を表現してもよい。

20

30

【符号の説明】

【0132】

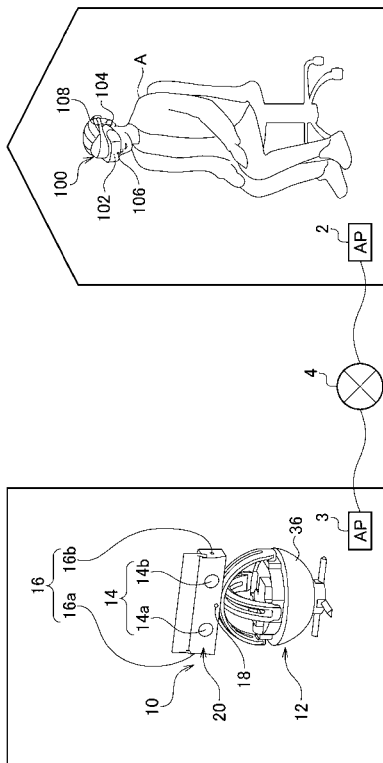
1・・・情報処理システム、10・・・ロボット、12・・・アクチュエータ装置、14a・・・右カメラ、14b・・・左カメラ、16a・・・右マイク、16b・・・左マイク、20・・・筐体、22・・・入力系統、24・・・出力系統、30・・・台座、32・・・第 1 円弧状アーム、32a・・・第 1 貫通長孔、34・・・第 2 円弧状アーム、34a・・・第 2 貫通長孔、36・・・ハウジング、38・・・カバー、40・・・脚部、42・・・挿通部材、42a・・・第 1 規制部、42b・・・第 2 規制部、42c・・・軸部、50・・・駆動機構、52・・・第 1 モータ、54・・・第 2 モータ、56・・・第 3 モータ、60・・・受信部、62・・・センサ情報取得部、64・・・動き検出部、66・・・視線方向決定部、68・・・アクチュエータ制御部、70・・・音声データ取得部、72・・・音声処理部、80・・・画像処理部、82・・・音声処理部、82a・・・位相差増幅装置、84a・・・第 1 増幅器、84b・・・第 2 増幅器、86a・・・第 1 加算器、86b・・・第 2 加算器、88a・・・第 3 増幅器、88b・・・第 4 増幅器、90・・・送信部、92・・・画像記録装置、100・・・HMD、102・・・表示パネル、104・・・イヤホン、106・・・マイク、108・・・装着バンド、110・・・出力機構部、112・・・装着機構部、114・・・筐体、120・・・制御部、122・・・記憶部、124・・・姿勢センサ、126・・・通信制御部、200・・・処理装置、202・・・受信部、204・・・センサ情報取得部、206・・・動き検出部、208・・・視線方向決定部、210・・・画像決定部、212・・・音声決定部

40

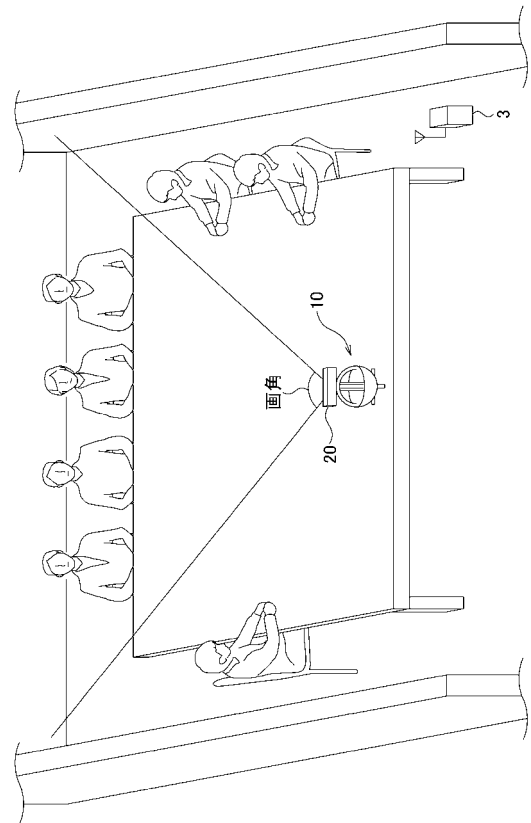
50

、 2 1 4 . . . 視聴データ提供部、 2 1 6 . . . 送信部、 2 1 8 . . . 記録部、 2 2 0 . . . 画像記録部、 2 2 2 . . . 音声記録部。

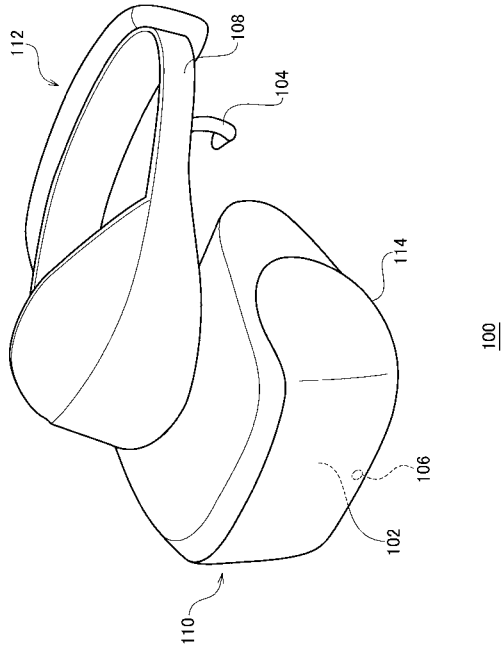
【 図 1 】



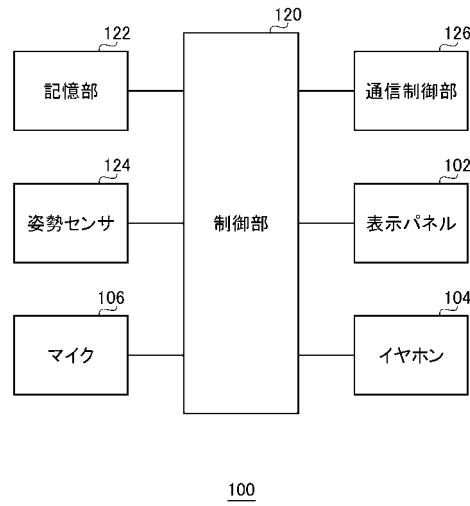
【 図 2 】



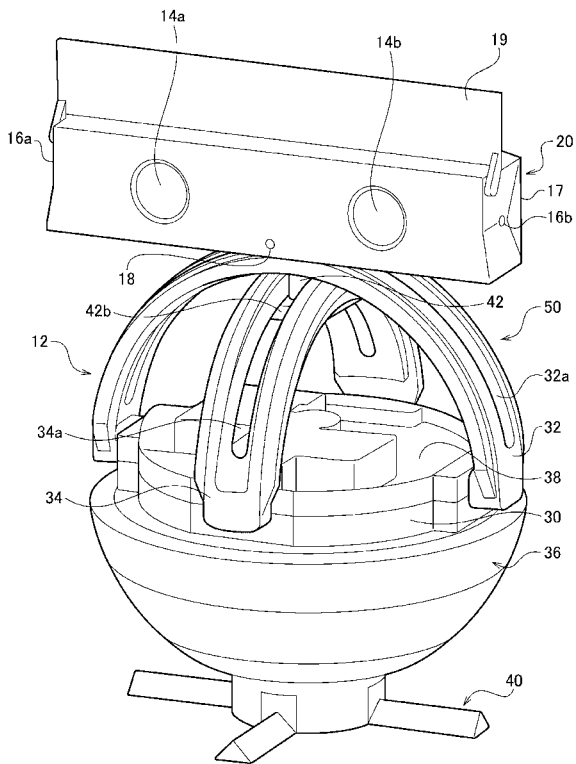
【 図 3 】



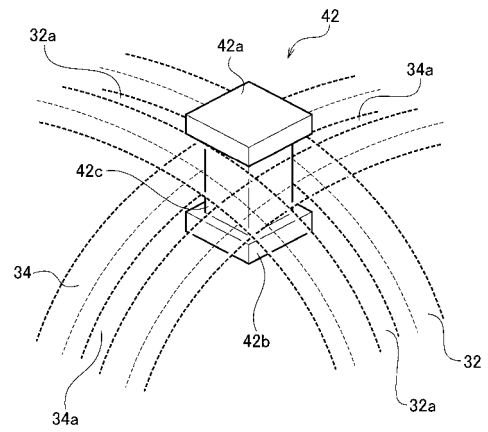
【 図 4 】



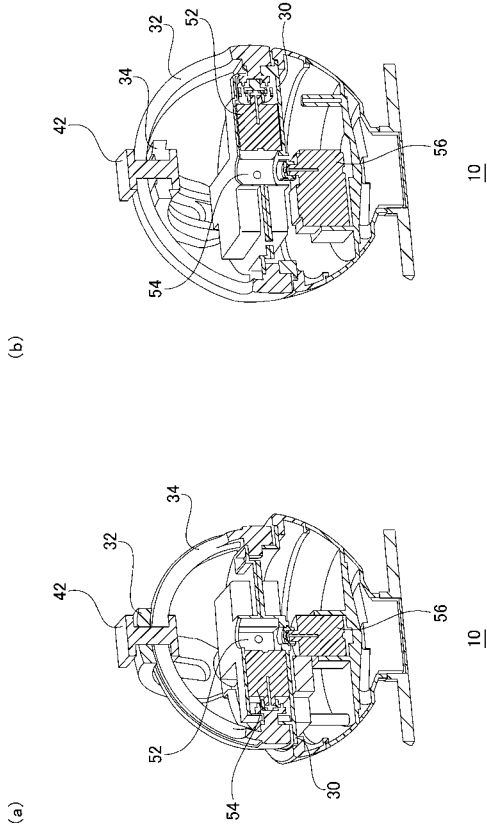
【 図 5 】



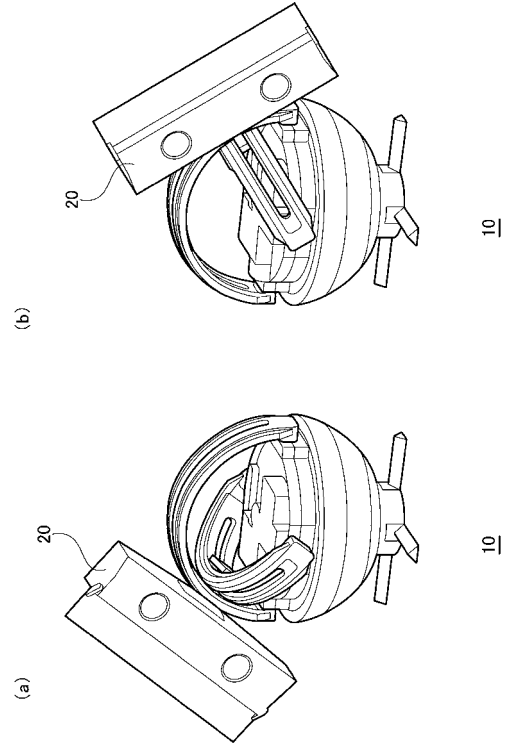
【 図 6 】



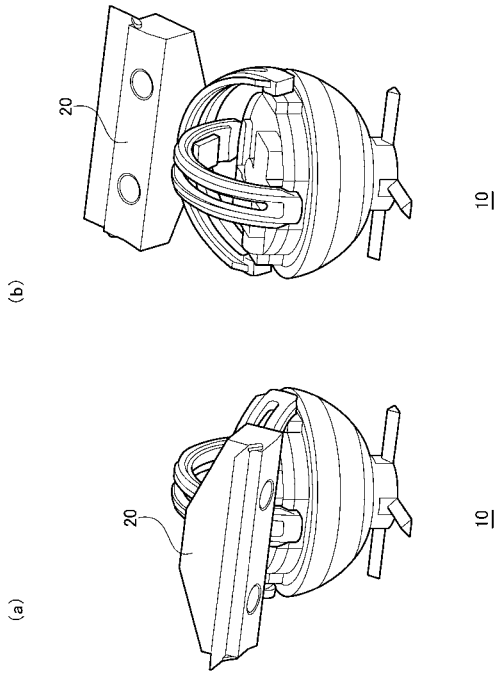
【 図 7 】



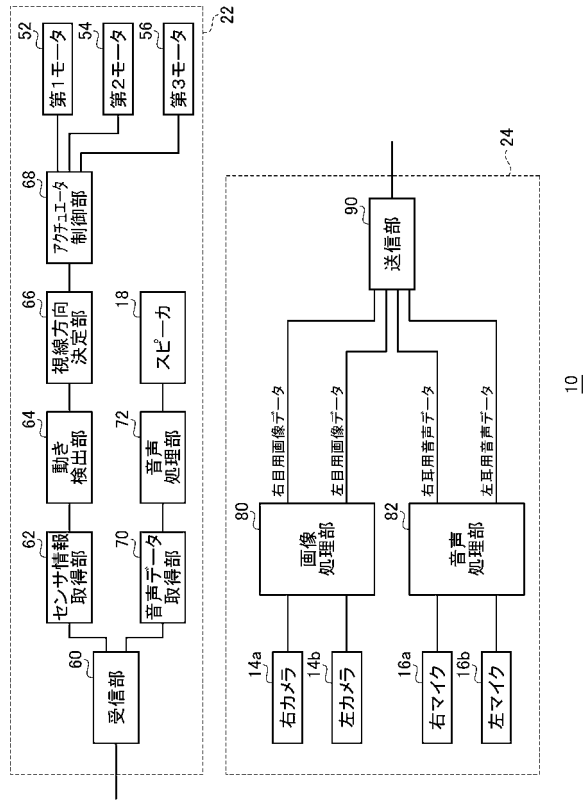
【 図 8 】



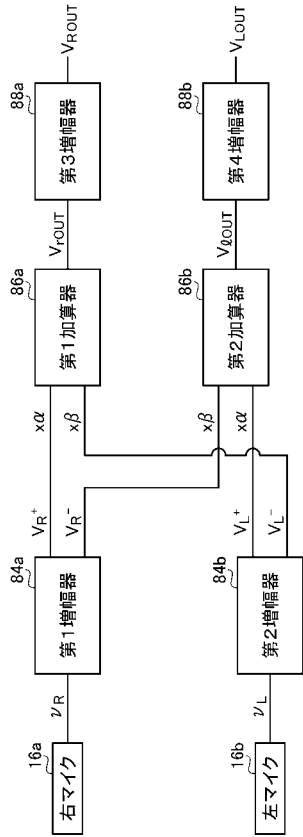
【 図 9 】



【 図 10 】

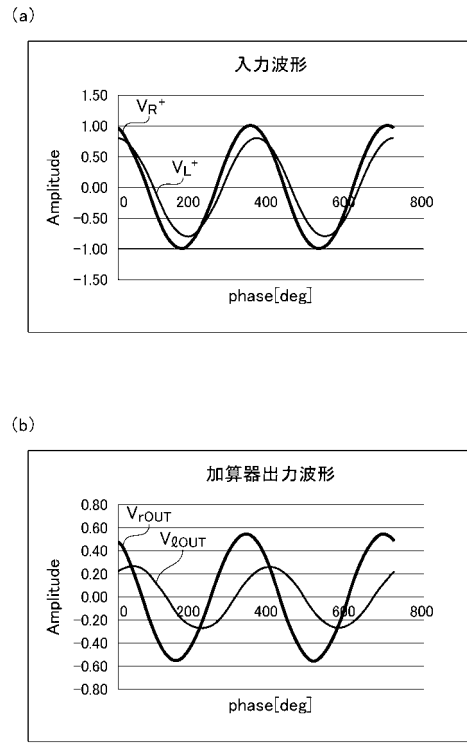


【図 1 1】

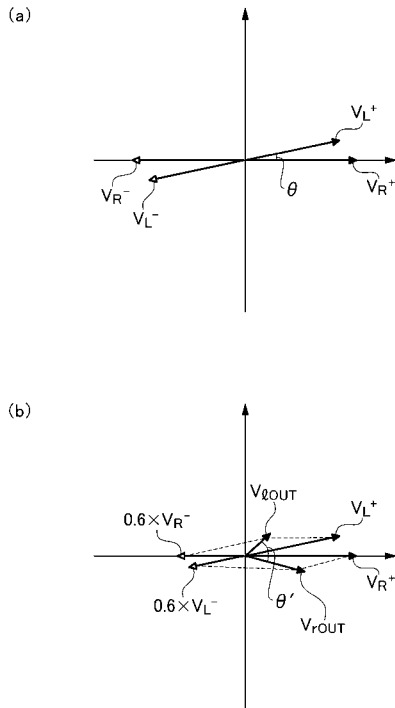


82a

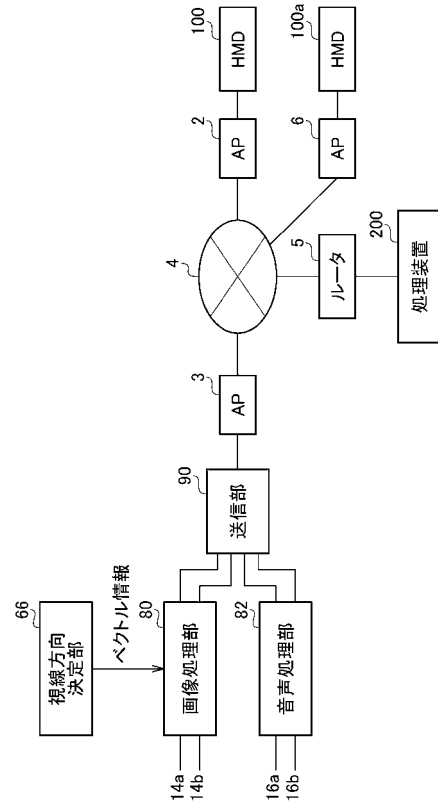
【図 1 2】



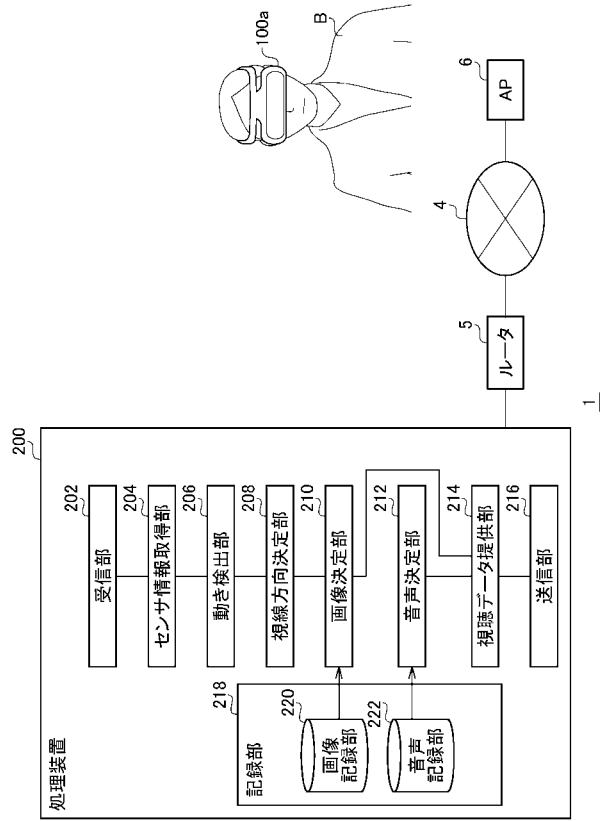
【図 1 3】



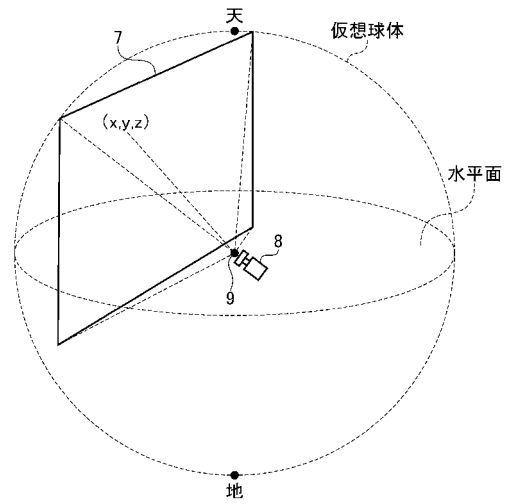
【図 1 4】



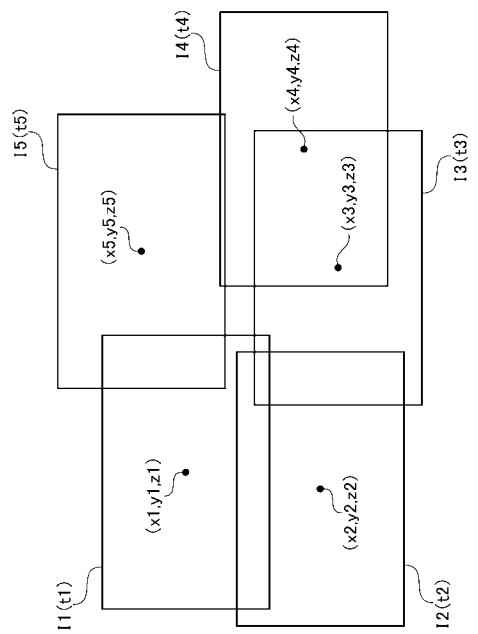
【図 15】



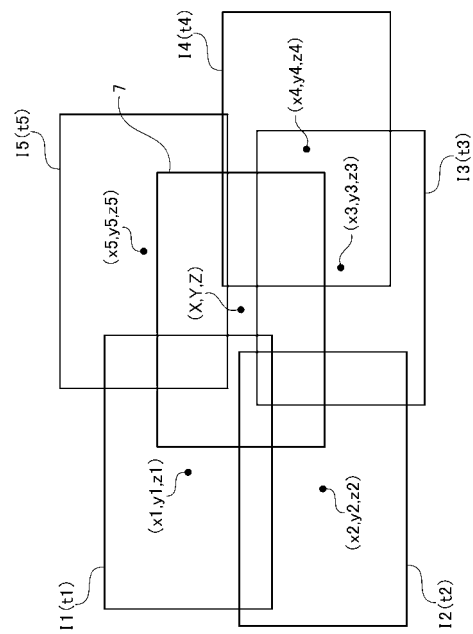
【図 16】



【図 17】



【図 18】



フロントページの続き

(72)発明者 高橋 佑介

東京都港区港南1丁目7番1号 株式会社ソニー・インタラクティブエンタテインメント内

Fターム(参考) 2H199 CA04 CA92 CA93 CA94 CA95 CA96 CA97

5D220 BA08 BC08