(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2018年7月12日(12.07.2018)

WIPO | PCT

(10) 国際公開番号
WO 2018/128161 A1

(51) 国際特許分類:
G06F 3/16 (2006.01) *G06F 3/01* (2006.01)

(21) 国際出願番号： PCT/JP2017/047225

(22) 国際出願日： 2017年12月28日(28.12.2017)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2017-001267 2017年1月6日(06.01.2017) JP

(71) 出願人: 株式会社ソニー・インタラクティブエンタテインメント (**SONY INTERACTIVE ENTERTAINMENT INC.**) [JP/JP]; 〒1080075 東京都港区港南１丁目７番１号 Tokyo (JP).

(72) 発明者: 宮崎 良雄 (**MIYAZAKI, Yoshio**); 〒1080075 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 Tokyo (JP).

(74) 代理人: 特許業務法人はるか国際特許事務所 (**HARUKA PATENT & TRADEMARK ATTORNEYS**); 〒1020085 東京都千代田区六番町３ 六番町ＳＫビル５階 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,

(54) **Title:** VOICE OUTPUT DEVICE, HEAD-MOUNTED DISPLAY, AND VOICE OUTPUT METHOD AND PROGRAM

(54) 発明の名称： 音声出力装置、ヘッドマウントディスプレイ、音声出力方法及びプログラム

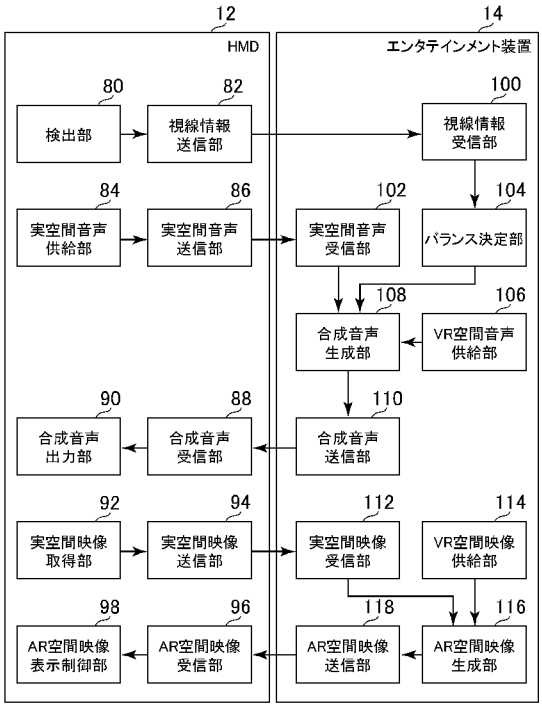


FIG. 7:
14 Entertainment device
80 Detection unit
82 Gaze information transmission unit
84 Actual space voice supply unit
86 Actual space voice transmission unit
88 Synthetic voice reception unit
90 Synthetic voice output unit
92 Actual space image acquisition unit
94 Actual space image transmission unit
96 AR space image reception unit
98 AR space image display control unit
100 Gaze information reception unit
102 Actual space voice reception unit
104 Balance determination unit
106 VR space voice supply unit
108 Synthetic voice generation unit
110 Synthetic voice transmission unit
112 Actual space image reception unit
114 VR space image supply unit
116 AR space image generation unit
118 AR space image transmission unit

(57) **Abstract:** Provided are a voice output device, a head-mounted display, and a voice output method and a program which are capable of improving the realism of an augmented reality space or mixed reality space. According to the present invention, an AR space image generation unit (116) generates an image of an augmented reality space or mixed reality space displayed on a display unit of the head-mounted display, the image including a portion occupied by an image of a virtual reality space and a portion occupied by an image of an actual space imaged by a camera for imaging the surroundings

WO 2018/128161 A1

［続葉有］

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一　国際調査報告（条約第21条(3)）

of a user. A synthetic voice generation unit (108) synthesizes a voice of the virtual reality space and a voice of the actual space around the user to generate a synthetic voice. A synthetic voice transmission unit (110) outputs a synthetic voice emitted from a speaker disposed near an ear of the user.

(57) 要約：拡張現実空間や複合現実空間の臨場感を向上できる音声出力装置、ヘッドマウントディスプレイ、音声出力方法及びプログラムを提供する。ＡＲ空間映像生成部（１１６）は、仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成する。合成音声生成部（１０８）は、仮想現実空間の音声とユーザの周辺の実空間の音声とを合成して合成音声を生成する。合成音声送信部（１１０）は、ユーザの耳の近傍に配置されるスピーカから放音される合成音声を出力する。

明　細　書

発明の名称 ：

音声出力装置、ヘッドマウントディスプレイ、音声出力方法及びプログラム

技術分野

[0001] 本発明は、音声出力装置、ヘッドマウントディスプレイ、音声出力方法及びプログラムに関する。

背景技術

[0002] 仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む画像から構成される拡張現実空間の映像を表示する技術が知られている。また近年、コンピュータグラフィックスで作られた仮想空間に実空間の情報を取り込み、実空間と仮想現実空間とを融合させる複合現実（ＭＲ）の技術が注目され始めている。実空間と仮想現実空間とが融合した複合現実空間においては、仮想現実空間に配置された仮想オブジェクトや仮想現実空間の環境と実空間に配置された物体や実空間の環境とのインタラクションがより重要となる。

発明の概要

発明が解決しようとする課題

[0003] 発明者らは、このような拡張現実空間や複合現実空間の映像を、頭部に装着して眼前に表示される映像を鑑賞できるヘッドマウントディスプレイ（ＨＭＤ）に表示させることで、拡張現実空間や複合現実空間の中にいるような感覚をユーザに体験させることを検討している。

[0004] 一般的に、ＨＭＤを装着したユーザには、例えばＢＧＭ、効果音、仮想現実空間に配置されたキャラクタ等の仮想オブジェクトの声などといった仮想現実空間の音声が、イヤホンやヘッドホン等のユーザの耳の近傍に配置されるスピーカから聞こえるようになっている。

[0005] しかしこの場合、上記スピーカから放音される音声によってユーザの周辺

の実空間の音声が遮られるため、拡張現実空間や複合現実空間の臨場感が阻害されていた。

[0006] 本発明の目的の一つは、拡張現実空間や複合現実空間の臨場感を向上できる音声出力装置、ヘッドマウントディスプレイ、音声出力方法及びプログラム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、本発明に係る音声出力装置は、仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成する画像生成部と、前記仮想現実空間の音声と前記ユーザの周辺の実空間の音声とを合成して合成音声を生成する合成音声生成部と、前記ユーザの耳の近傍に配置されるスピーカから放音される前記合成音声を出力する音声出力部と、を含む。

[0008] 本発明の一態様では、前記合成音声生成部は、前記ユーザの視線に応じた音量のバランスで前記仮想現実空間の音声と前記実空間の音声とを合成して前記合成音声を生成する。

[0009] また、本発明の一態様では、前記合成音声生成部は、前記仮想現実空間の音声と、前記ユーザの視線の方向の音声が強調された前記実空間の音声と、を合成して前記合成音声を生成する。

[0010] また、本発明の一態様では、前記仮想現実空間の音声は、記憶されている音声データ又は受信する音声ストリームが表す音声である。

[0011] また、本発明に係るヘッドマウントディスプレイは、ユーザの周辺を撮影するカメラ部と、前記ユーザの周辺の音声を取得するマイク部と、仮想現実空間の画像が占める部分と前記カメラで撮影された実空間の画像が占める部分とを含む拡張現実空間又は複合現実空間の画像を表示する表示部と、前記仮想現実空間の音声と前記マイクが取得する実空間の音声との合成音声を前記ユーザの耳の近傍に配置されるスピーカを介して出力する音声出力部と、を含む。

[0012] また、本発明に係る音声出力方法は、仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成するステップと、前記仮想現実空間の音声と前記ユーザの周辺の実空間の音声とを合成して合成音声を生成するステップと、前記ユーザの耳の近傍に配置されるスピーカから放音される前記合成音声を出力するステップと、を含む。

[0013] また、本発明に係るプログラムは、仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成する手順と、前記仮想現実空間の音声と前記ユーザの周辺の実空間の音声とを合成して合成音声を生成する手順と、前記ユーザの耳の近傍に配置されるスピーカから放音される前記合成音声を出力する手順と、をコンピュータに実行させる。

図面の簡単な説明

[0014] [図1]本発明の一実施形態に係るエンタテインメントシステムの全体構成の一例を示す図である。

[図2A]本発明の一実施形態に係るヘッドマウントディスプレイの構成の一例を示す図である。

[図2B]本発明の一実施形態に係るエンタテインメント装置の構成の一例を示す図である。

[図2C]本発明の一実施形態に係る中継装置の構成の一例を示す図である。

[図3]ＡＲ空間画像の一例を示す図である。

[図4]注視点の一例を示す図である。

[図5A]合成音声におけるＶＲ空間音声の音量と実空間音声の音量とのバランスの一例を示す図である。

[図5B]合成音声におけるＶＲ空間音声の音量と実空間音声の音量とのバランスの別の一例を示す図である。

[図6]実空間音声において強調される実空間内の領域の一例を模式的に示す図である。

[図7]本発明の一実施形態に係るヘッドマウントディスプレイ及びエンタテインメント装置で実装される機能の一例を示す機能ブロック図である。

[図8]本発明の一実施形態に係るヘッドマウントディスプレイ及びエンタテインメント装置において行われる処理の流れの一例を示すフロー図である。

発明を実施するための形態

[0015] 以下、本発明の一実施形態について、図面を参照しながら説明する。

[0016] 図１は、本発明の一実施形態に係るエンタテインメントシステム１０の全体構成の一例を示す図である。図２Ａは、本実施形態に係るヘッドマウントディスプレイ（ＨＭＤ）１２の構成の一例を示す図である。図２Ｂは、本実施形態に係るエンタテインメント装置１４の構成の一例を示す図である。図２Ｃは、本実施形態に係る中継装置１６の構成の一例を示す図である。

[0017] 図１に示すように、本実施形態に係るエンタテインメントシステム１０は、ＨＭＤ１２とエンタテインメント装置１４と中継装置１６とディスプレイ１８とカメラマイクユニット２０とコントローラ２２とを含んでいる。

[0018] 本実施形態に係るＨＭＤ１２には、例えば図２Ａに示すように、プロセッサ３０、記憶部３２、通信部３４、入出力部３６、表示部３８、センサ部４０、音声出力部４２、カメラ部４４、視線検出部４６、マイク部４８が含まれる。そして通信部３４には、第１通信部３４ａと第２通信部３４ｂとが含まれる。

[0019] プロセッサ３０は、例えばＨＭＤ１２にインストールされるプログラムに従って動作するマイクロプロセッサ等のプログラム制御デバイスである。なお、ＨＭＤ１２に、プロセッサ３０の代わりに、プロセッサ３０によって実行される処理の実装が可能な制御回路が含まれるようにしてもよい。

[0020] 記憶部３２は、例えばＲＯＭやＲＡＭ等の記憶素子などである。記憶部３２には、プロセッサ３０によって実行されるプログラムなどが記憶される。

[0021] 第１通信部３４ａ及び第２通信部３４ｂは、それぞれ例えば無線ＬＡＮモ

ジュールなどの通信インタフェースである。本実施形態では図１に示すように、第１通信部３４ａ及び第２通信部３４ｂは、ＨＭＤ１２の前側上方に配置されている。

[0022] 本実施形態では第１通信部３４ａで行われる通信の通信規格と第２通信部３４ｂで行われる通信の通信規格とは異なっている。例えば第１通信部３４ａでは６０ＧＨｚ帯での通信が行われ、第２通信部３４ｂでは５ＧＨｚ帯又は２．４ＧＨｚ帯での通信が行われる。

[0023] 入出力部３６は、例えばＨＤＭＩ（登録商標）（High-Definition Multimedia Interface）ポート、ＵＳＢポート、ＡＵＸポートなどの入出力ポートである。

[0024] 表示部３８は、ＨＭＤ１２の前側に配置されている、例えば液晶ディスプレイや有機ＥＬディスプレイ等のディスプレイであり、エンタテインメント装置１４が生成する映像などを表示させる。また表示部３８は、ＨＭＤ１２の筐体に収容される。表示部３８は、例えばエンタテインメント装置１４が出力して中継装置１６で中継される映像信号を受信して、当該映像信号が表す映像を出力するようにしてもよい。本実施形態に係る表示部３８は、例えば左目用の画像と右目用の画像を表示することによって三次元画像を表示させることができるようになっている。なお表示部３８は三次元画像の表示ができず二次元画像の表示のみができるものであっても構わない。

[0025] センサ部４０は、例えば加速度や角速度を検出可能なモーションセンサなどといったセンサである。センサ部４０は、ＨＭＤ１２の回転量、移動量などの検出結果を所定のサンプリングレートで、プロセッサ３０に出力する。

[0026] 音声出力部４２は、例えばヘッドホンやイヤホン等の、ユーザの耳の近傍に配置されるスピーカであり、エンタテインメント装置１４が生成する音声データが表す音声などを出力する。音声出力部４２は、例えばエンタテインメント装置１４が出力して中継装置１６で中継される音声信号を受信して、当該音声信号が表す音声を出力する。

[0027] カメラ部４４は、例えばデジタルカメラなどのカメラであり、所定のフレ

ームレートでＨＭＤ１２を装着するユーザの周辺の様子を撮影する。図１に示すように、本実施形態に係るカメラ部４４は、表示部３８の上方に、表示部３８の前方を撮影できるように２つ配置されている。そのため本実施形態に係るカメラ部４４は、ＨＭＤ１２を装着するユーザの前方を撮影できることとなる。また本実施形態に係るカメラ部４４は、例えば左目用の画像を撮影するためのレンズ及び右目用の画像を生成するためのレンズを備えたステレオカメラである。

[0028] 視線検出部４６は、例えばＨＭＤ１２を装着したユーザの視線を検出するセンサである。視線検出部４６は、例えば、ＨＭＤ１２の筐体の内側に配置される。また視線検出部４６は、ＨＭＤ１２を装着したユーザの視線の方向を検出する赤外線カメラを含んでいてもよい。

[0029] マイク部４８は、例えば周囲の音声を取得して当該音声を音声データに変換して出力するマイクである。またマイク部４８が、特定の方向の音声が強調された音声データを出力する機能を備えていてもよい。例えばマイク部４８が、アレイマイクロフォン等といった指向方向が可変なマイクであっても構わない。また例えばマイク部４８が、ステレオマイクであってもよい。そしてこの場合に、ステレオマイクに入力された音声を加工して特定の方向の音声を強調する処理が実行可能なプログラムが記憶部３２に記憶されており、当該プログラムがプロセッサ３０により実行されてもよい。

[0030] 本実施形態に係るエンタテインメント装置１４は、例えばゲームコンソール、ＤＶＤプレイヤ、Ｂｌｕ－ｒａｙ（登録商標）プレイヤなどといったコンピュータである。本実施形態に係るエンタテインメント装置１４は、例えば記憶されているゲームプログラムの実行や光ディスクに記録されたコンテンツの再生などによって映像や音声を生成する。そして本実施形態に係るエンタテインメント装置１４は、生成される映像を表す映像信号や生成される音声を表す音声信号を、中継装置１６を経由して、ＨＭＤ１２やディスプレイ１８に出力する。

[0031] 本実施形態に係るエンタテインメント装置１４には、例えば図２Ｂに示す

ように、プロセッサ５０、記憶部５２、通信部５４、入出力部５６が含まれる。

[0032] プロセッサ５０は、例えばエンタテインメント装置１４にインストールされるプログラムに従って動作するＣＰＵ等のプログラム制御デバイスである。本実施形態に係るプロセッサ５０には、ＣＰＵから供給されるグラフィックスコマンドやデータに基づいてフレームバッファに画像を描画するＧＰＵ（Graphics Processing Unit）も含まれている。なお、エンタテインメント装置１４に、プロセッサ５０の代わりに、プロセッサ５０によって実行される処理の実装が可能な制御回路が含まれるようにしてもよい。

[0033] 記憶部５２は、例えばＲＯＭやＲＡＭ等の記憶素子やハードディスクドライブなどである。記憶部５２には、プロセッサ５０によって実行されるプログラムなどが記憶される。また、本実施形態に係る記憶部５２には、ＧＰＵにより画像が描画されるフレームバッファの領域が確保されている。

[0034] 通信部５４は、例えば無線ＬＡＮモジュールなどの通信インタフェースである。

[0035] 入出力部５６は、ＨＤＭＩポート、ＵＳＢポートなどの入出力ポートである。

[0036] 本実施形態に係る中継装置１６は、エンタテインメント装置１４から出力される映像信号や音声信号を中継してＨＭＤ１２やディスプレイ１８に出力するコンピュータである。

[0037] 本実施形態に係る中継装置１６には、例えば図２Ｃに示すように、プロセッサ６０、記憶部６２、２つの通信部６４（第１通信部６４ａ及び第２通信部６４ｂ）、入出力部６６が含まれる。

[0038] プロセッサ６０は、例えば中継装置１６にインストールされるプログラムに従って動作するＣＰＵ等のプログラム制御デバイスである。なお、中継装置１６に、プロセッサ６０の代わりに、プロセッサ６０によって実行される処理の実装が可能な制御回路が含まれるようにしてもよい。

[0039] 記憶部６２は、例えばＲＯＭやＲＡＭ等の記憶素子などである。記憶部６

２には、プロセッサ６０によって実行されるプログラムなどが記憶される。

[0040] 第１通信部６４ａ及び第２通信部６４ｂは、例えば無線ＬＡＮモジュールなどの通信インタフェースである。本実施形態では、図１に示すように、中継装置１６の前側に、第１通信部６４ａ及び第２通信部６４ｂが含まれている。

[0041] 本実施形態では第１通信部６４ａで行われる通信の通信規格と第２通信部６４ｂで行われる通信の通信規格とは異なっている。例えば第１通信部６４ａでは６０ＧＨｚ帯での通信が行われ、第２通信部６４ｂでは５ＧＨｚ帯又は２．４ＧＨｚ帯での通信が行われる。

[0042] 入出力部６６は、ＨＤＭＩポート、ＵＳＢポートなどの入出力ポートである。

[0043] 本実施形態に係るディスプレイ１８は、例えば液晶ディスプレイ等であり、エンタテインメント装置１４から出力される映像信号が表す映像などを表示させる。

[0044] 本実施形態に係るカメラマイクユニット２０は、例えば被写体を撮像した画像などといったカメラマイクユニット２０の周辺の様子を表すデータをエンタテインメント装置１４に出力するカメラ２０ａを含んでいる。また本実施形態に係るカメラ２０ａはステレオカメラである。また本実施形態に係るカメラマイクユニット２０は、周囲の音声を取得して当該音声を音声データに変換してエンタテインメント装置１４に出力するマイク２０ｂを含んでいる。

[0045] ＨＭＤ１２と中継装置１６とは、例えば、無線通信によるデータの送受信が互いに可能になっている。エンタテインメント装置１４と中継装置１６とは、例えば、ＨＤＭＩケーブルやＵＳＢケーブルなどを介して接続されている。中継装置１６とディスプレイ１８とは、例えば、ＨＤＭＩケーブルなどを介して接続されている。エンタテインメント装置１４とカメラマイクユニット２０とは、例えば、ＡＵＸケーブルなどを介して接続されている。

[0046] 本実施形態に係るコントローラ２２は、エンタテインメント装置１４に対

する操作入力を行うための操作入力装置である。ユーザは、コントローラ２２が備える方向キーやボタンを押下したり、操作スティックを傾けたりすることで、コントローラ２２を用いて各種の操作入力を行うことができる。そして本実施形態では、コントローラ２２は、操作入力に対応付けられる入力データをエンタテインメント装置１４に出力する。また本実施形態に係るコントローラ２２は、ＵＳＢポートを備えている。そしてコントローラ２２は、ＵＳＢケーブルでエンタテインメント装置１４と接続することで、有線で入力データをエンタテインメント装置１４に出力することができる。また本実施形態に係るコントローラ２２は、無線通信モジュール等を備えており、無線で入力データをエンタテインメント装置１４に出力することができるようにもなっている。

[0047] 本実施形態では例えば、エンタテインメント装置１４で、ゲームプログラム等のプログラムが実行される。そしてエンタテインメント装置１４は、当該プログラムの実行状況に応じた３次元の仮想現実空間の映像を生成する。

[0048] また本実施形態では、カメラ部４４によりユーザの周辺の実空間の映像が撮影される。

[0049] そして本実施形態では、エンタテインメント装置１４が生成する仮想現実空間の映像と撮影される実空間の映像とに基づく拡張現実空間の映像が表示部３８に表示される。なお以下では、表示部３８に拡張現実空間の映像が表示される場合の一例について説明するが、表示部３８に複合現実空間（ＭＲ空間）の映像が表示される場合についても本実施形態が適用可能であることは言うまでもない。

[0050] 以下、仮想現実空間をＶＲ空間と呼び、拡張現実空間をＡＲ空間と呼ぶこととする。また、カメラ部４４により撮影される、ユーザが存在する実空間の映像を実空間映像と呼ぶこととする。また、エンタテインメント装置１４が生成するＶＲ空間の映像をＶＲ空間映像と呼ぶこととする。また、実空間映像とＶＲ空間映像とに基づくＡＲ空間の映像をＡＲ空間映像と呼ぶこととする。

[0051] 図３は、表示部３８に表示されるＡＲ空間映像を構成するフレーム画像の一例を示す図である。以下、図３に例示する、ＡＲ空間映像を構成するフレーム画像をＡＲ空間画像７０と呼ぶこととする。図３に示すように、ＡＲ空間画像７０には、実空間映像を構成するフレーム画像が占める部分である実空間部分７２とＶＲ空間映像を構成するフレーム画像が占める部分であるＶＲ空間部分７４とが含まれている。図３の例では、仮想現実空間内に配置された仮想オブジェクトの画像が、ＶＲ空間部分７４として示されている。ここでＡＲ空間画像７０は、実空間映像を構成するフレーム画像にＶＲ空間映像を構成するフレーム画像が重畳された画像であっても構わない。

[0052] そして本実施形態では、エンタテインメント装置１４がプログラムを実行することにより生成又は再生される仮想現実空間の音声と、ＨＭＤ１２を装着するユーザの周辺の実空間の音声と、の合成音声が、音声出力部４２から放音される。以下、プログラムを実行することにより生成又は再生される仮想現実空間の音声をＶＲ空間音声と呼び、ＨＭＤ１２を装着するユーザの周辺の実空間の音声を実空間音声と呼ぶこととする。

[0053] このように本実施形態では、ＶＲ空間音声と実空間音声とが合成された合成音声が音声出力部４２から放音される。そのため、ユーザの周辺の実空間の音声が遮られた状態で音声出力部４２からＶＲ空間音声だけが放音される場合よりも拡張現実空間の臨場感が向上することとなる。

[0054] ここでＶＲ空間音声として、例えば、ＢＧＭ、効果音、キャラクタの声などといった仮想オブジェクトに関する音声、などが挙げられる。またＶＲ空間音声は、記憶部５２に記憶された音声データやＤＶＤディスクやＢｌｕ－ｒａｙディスク等の記録媒体に記録された音声データを再生することにより得られる音声であってもよい。またＶＲ空間音声は、インターネットなどを介して受信する音声ストリームが表す音声であってもよい。

[0055] そして本実施形態では、視線検出部４６が検出するユーザの視線に応じた音量のバランスでＶＲ空間音声と実空間音声とが合成された合成音声が音声出力部４２から放音される。ここで例えば、視線検出部４６が検出するユー

ザの視線が実空間部分７２を向いているか仮想空間部分７４を向いているかに応じて合成音声におけるＶＲ空間音声の音量と実空間音声の音量とのバランスが設定されてもよい。

[0056] 図４は、ユーザの視線に対応付けられる注視点Ｐの一例を示す図である。本実施形態では、公知のアイトラッキング技術を用いることにより、視線検出部４６による視線の検出結果に基づいて、ユーザの視点７６を基準とした視線方向７８、及び、当該視線方向７８に対応付けられる注視点Ｐの位置が特定されることとする。そして本実施形態では例えば、注視点Ｐの位置に応じた音量のバランスでＶＲ空間音声と実空間音声とが合成された合成音声が音声出力部４２から放音される。

[0057] 図４には、表示部３８の中心、左上隅、右上隅、左下隅、及び、右下隅のそれぞれに対応する位置が、３次元空間内の位置Ｐ０、位置Ｐ１、位置Ｐ２、位置Ｐ３、及び、位置Ｐ４として示されている。そして本実施形態では例えば、ユーザが正面を見ている際には、位置Ｐ０が注視点Ｐとして特定される。同様に、同様にユーザが表示部３８の左上隅、右上隅、左下隅、右下隅、を見ている際には、それぞれ、位置Ｐ１、位置Ｐ２、位置Ｐ３、位置Ｐ４が、注視点Ｐとして特定される。

[0058] そして図３には、特定される注視点Ｐの例として、注視点Ｐａ、注視点Ｐｂ、及び、注視点Ｐｃが示されている。

[0059] ここで図３の例では、注視点Ｐａの位置は、ＶＲ空間部分７４内に存在している。この場合、図５Ａに示すように、ＶＲ空間音声の音量の方が実空間音声の音量よりも大きい合成音声が生成されて、音声出力部４２から放音される。この場合に例えば、図５Ａに示すように、ＶＲ空間音声の音量が実空間音声の音量の倍となるよう音量のバランスが設定されてもよい。

[0060] 一方、注視点Ｐｂ、及び、注視点Ｐｃの位置は、実空間部分７２内に存在している。この場合、図５Ｂに示すように、実空間音声の音量の方がＶＲ空間音声の音量よりも大きい合成音声が生成されて、音声出力部４２から放音される。この場合に例えば、図５Ｂに示すように、実空間音声の音量がＶＲ

空間音声の音量の倍となるようバランスが設定されてもよい。

[0061] 以上のようにして本実施形態では、ユーザは視線を変えることによりＶＲ空間音声と実空間音声との音量のバランスを制御できることとなる。例えばユーザは視線を変えることにより、ユーザの注視対象が仮想オブジェクトである場合にはＶＲ空間音声が強調され、ユーザの注視対象が実空間内のオブジェクトである場合には実空間音声が強調されるよう制御できることとなる。

[0062] また本実施形態に係る実空間音声は、注視点Ｐの位置に応じた方向からの音声が強調されるようになっている。ここで例えば、視線方向７８からの音声が強調されてもよい。また例えば、視線方向７８のベクトルを水平面に投影した方向からの音声が強調されてもよい。

[0063] 図６は、実空間音声において強調される実空間内の領域の一例を模式的に示す図である。図６には、ＨＭＤ１２を装着しているユーザを上から見た様子が描かれている。ここで例えば、注視点Ｐが図３に示す注視点Ｐａのように正面より左側である場合に、図６の領域Ｒ１に示すように、左前方向からの音声が強調されるようにしてもよい。また例えば、注視点Ｐが図３に示す注視点Ｐｂのように正面より右側である場合に、図６の領域Ｒ２に示すように、右前方向からの音声が強調されるようにしてもよい。また例えば、注視点Ｐが図３に示す注視点Ｐｃのように正面である場合に、図６の領域Ｒ３に示すように、正面方向からの音声が強調されるようにしてもよい。

[0064] また例えば、注視点Ｐａのように、注視点Ｐの位置がＶＲ空間部分７４内である場合に、図６の領域Ｒ４に示すように、実空間音声における全方向の音声の音量が等しくてもよい。すなわち、この場合は実空間音声が無指向であってもよい。

[0065] また例えば、公知の画像認識技術を用いて、実空間部分７２内において、実空間に配置されている複数の物体のそれぞれの像が示されている部分の特定ができる場合がある。この場合、注視点Ｐの位置が特定の物体（例えば壁や天井など）の像が示されている部分内に存在する際には、図６の領域Ｒ４

に示すように、実空間音声が無指向であってもよい。

[0066] 以上のようにして本実施形態では、ユーザは視線を変えることによりＶＲ空間音声と合成される実空間音声の指向方向を制御できることとなる。

[0067] なお本実施形態において、ユーザはコントローラ２２を操作することなどにより、音声出力部４２から出力されている音声の全体の音量を設定できてもよい。

[0068] 以下、本実施形態に係るＨＭＤ１２及びエンタテインメント装置１４の機能並びに本実施形態に係るＨＭＤ１２及びエンタテインメント装置１４で実行される処理についてさらに説明する。

[0069] 図７は、本実施形態に係るＨＭＤ１２及びエンタテインメント装置１４で実装される機能の一例を示す機能ブロック図である。なお、本実施形態に係るＨＭＤ１２及びエンタテインメント装置１４で、図７に示す機能のすべてが実装される必要はなく、また、図７に示す機能以外の機能が実装されていても構わない。

[0070] 図７に示すように、本実施形態に係るＨＭＤ１２は、機能的には例えば、検出部８０、視線情報送信部８２、実空間音声供給部８４、実空間音声送信部８６、合成音声受信部８８、合成音声出力部９０、実空間映像取得部９２、実空間映像送信部９４、ＡＲ空間映像受信部９６、ＡＲ空間映像表示制御部９８、を含んでいる。

[0071] 検出部８０は、視線検出部４６を主として実装される。視線情報送信部８２、実空間音声送信部８６、合成音声受信部８８、実空間映像送信部９４、ＡＲ空間映像受信部９６は、通信部３４を主として実装される。実空間音声供給部８４は、マイク部４８、又は、プロセッサ３０及びマイク部４８を主として実装される。合成音声出力部９０は、プロセッサ３０及び音声出力部４２を主として実装される。実空間映像取得部９２は、カメラ部４４を主として実装される。ＡＲ空間映像表示制御部９８は、プロセッサ３０及び表示部３８を主として実装される。

[0072] 以上の機能は、コンピュータであるＨＭＤ１２にインストールされた、以

上の機能に対応する指令を含むプログラムをプロセッサ３０で実行することにより実装されてもよい。このプログラムは、例えば、光ディスク、磁気ディスク、磁気テープ、光磁気ディスク、フラッシュメモリ等のコンピュータ読み取り可能な情報記憶媒体を介して、あるいは、インターネットなどを介してＨＭＤ１２に供給されてもよい。

[0073] また図７に示すように、本実施形態に係るエンタテインメント装置１４は、機能的には例えば、視線情報受信部１００、実空間音声受信部１０２、バランス決定部１０４、ＶＲ空間音声供給部１０６、合成音声生成部１０８、合成音声送信部１１０、実空間映像受信部１１２、ＶＲ空間映像供給部１１４、ＡＲ空間映像生成部１１６、ＡＲ空間映像送信部１１８、を含んでいる。

[0074] 視線情報受信部１００、実空間音声受信部１０２、合成音声送信部１１０、実空間映像受信部１１２、ＡＲ空間映像送信部１１８は、入出力部５６を主として実装される。バランス決定部１０４、ＶＲ空間音声供給部１０６、合成音声生成部１０８、ＶＲ空間映像供給部１１４、ＡＲ空間映像生成部１１６は、プロセッサ５０を主として実装される。

[0075] 以上の機能は、コンピュータであるエンタテインメント装置１４にインストールされた、以上の機能に対応する指令を含むプログラムをプロセッサ５０で実行することにより実装されてもよい。このプログラムは、例えば、光ディスク、磁気ディスク、磁気テープ、光磁気ディスク、フラッシュメモリ等のコンピュータ読み取り可能な情報記憶媒体を介して、あるいは、インターネットなどを介してエンタテインメント装置１４に供給されてもよい。

[0076] 検出部８０は、本実施形態では例えば、ユーザの視線を検出して、当該視線の検出結果に基づいて、合成音声のバランスの決定に用いられる視線情報を生成する。ここで検出部８０が、例えば、ＡＲ空間画像７０内における注視点Ｐの位置を示す情報や、視線方向７８のベクトルを表す情報を視線情報として生成してもよい。また検出部８０が、例えば、注視点Ｐの位置を示す目印が重畳されたＡＲ空間画像７０を視線情報として生成してもよい。

[0077] 視線情報送信部８２は、本実施形態では例えば、検出部８０が生成する視線情報を、中継装置１６経由でエンタテインメント装置１４に送信する。

[0078] 実空間音声供給部８４は、本実施形態では例えば、上述の実空間音声を実空間音声送信部８６に供給する。ここで例えば実空間音声供給部８４は、上述のように検出部８０が生成する視線情報に基づいて特定される方向の音声が強調された指向性のある音声を実空間音声送信部８６に供給してもよい。

[0079] 実空間音声送信部８６は、本実施形態では例えば、実空間音声供給部８４から供給される実空間音声を、中継装置１６経由でエンタテインメント装置１４に送信する。

[0080] 合成音声受信部８８は、本実施形態では例えば、ＶＲ空間音声と実空間音声との合成音声を受信する。ここで合成音声受信部８８は、ユーザの耳の近傍に配置されるスピーカから合成音声を放音してもよい。

[0081] 合成音声出力部９０は、本実施形態では例えば、合成音声受信部８８が受信する合成音声を出力する。

[0082] 実空間映像取得部９２は、本実施形態では例えば、カメラ部４４が撮影する実空間映像を取得する。

[0083] 実空間映像送信部９４は、本実施形態では例えば、実空間映像取得部９２が取得する実空間映像を中継装置１６経由でエンタテインメント装置１４に送信する。

[0084] ＡＲ空間映像受信部９６は、本実施形態では例えば、エンタテインメント装置１４が生成するＡＲ空間映像を受信する。

[0085] ＡＲ空間映像表示制御部９８は、本実施形態では例えば、ＡＲ空間映像受信部９６が受信するＡＲ空間映像を表示部３８に表示させる。

[0086] 視線情報受信部１００は、本実施形態では例えば、視線情報送信部８２が送信する視線情報を受信する。

[0087] 実空間音声受信部１０２は、本実施形態では例えば、実空間音声送信部８６が送信する実空間音声を受信する。

[0088] バランス決定部１０４は、本実施形態では例えば、視線情報受信部１００

が受信する視線情報に基づいて、合成音声におけるＶＲ空間音声と実空間音声とのバランスを決定する。ここで例えば、視線情報に基づいて注視点Ｐの位置がＶＲ空間部分７４内であることが特定される場合には、図５Ａに示すように、ＶＲ空間音声の音量が実空間音声の音量よりも大きくなるようバランスが決定されてもよい。一方例えば、視線情報に基づいて、注視点Ｐの位置が実空間部分７２内であることが特定される場合には、図５Ｂに示すように、実空間音声の音量の方がＶＲ空間音声の音量よりも大きくなるようバランスが決定されてもよい。

[0089] ＶＲ空間音声供給部１０６は、本実施形態では例えば、ＶＲ空間音声を合成音声生成部１０８に供給する。ここで上述のようにＶＲ空間音声は、例えば、記憶されている音声データ又は受信する音声ストリームが表す音声であってもよい。

[0090] 合成音声生成部１０８は、本実施形態では例えば、ＶＲ空間音声と実空間音声とを合成して合成音声を生成する。合成音声生成部１０８は、ＶＲ空間音声供給部１０６が供給するＶＲ空間音声と実空間音声受信部１０２が受信する実空間音声とを合成して合成音声を生成してもよい。また合成音声生成部１０８は、ユーザの視線に応じた音量のバランスでＶＲ空間音声と実空間音声とを合成して合成音声を生成してもよい。またＶＲ空間音声と合成される実空間音声は、ユーザの視線の方向の音声が強調された音声であってもよい。

[0091] 合成音声送信部１１０は、本実施形態では例えば、合成音声生成部１０８が生成する合成音声を中継装置１６経由でＨＭＤ１２に送信する。本実施形態に係る合成音声送信部１１０は、ヘッドホンやイヤホン等の、ユーザの耳の近傍に配置されるスピーカから放音される合成音声を出力する音声出力部としての役割を担うこととなる。

[0092] 実空間映像受信部１１２は、本実施形態では例えば、実空間映像送信部９４が送信する実空間映像を受信する。

[0093] ＶＲ空間映像供給部１１４は、本実施形態では例えば、ＶＲ空間映像をＡ

Ｒ空間映像生成部１１６に供給する。

[0094] ＡＲ空間映像生成部１１６は、本実施形態では例えば、実空間映像受信部１１２が受信する実空間映像とＶＲ空間映像供給部１１４が供給するＶＲ空間映像とに基づいてＡＲ空間映像を生成する。上述のようにＡＲ空間映像のフレーム画像であるＡＲ空間画像７０には、ＶＲ空間の画像が占めるＶＲ空間部分７４とユーザの周辺を撮影するカメラ部４４で撮影された実空間の画像が占める実空間部分７２とが含まれていてもよい。

[0095] ＡＲ空間映像送信部１１８は、本実施形態では例えば、ＡＲ空間映像生成部１１６が生成するＡＲ空間映像を中継装置１６経由でＨＭＤ１２に送信する。

[0096] 以下、本実施形態に係るＨＭＤ１２及びエンタテインメント装置１４で実行される処理の流れの一例を、図８に例示するフロー図を参照しながら説明する。なお本処理例では、以下に示すＳ１０１～Ｓ１１０に示す処理が繰り返し実行される。なお以下に示すＳ１０１～Ｓ１１０に示す処理が所定時間間隔で繰り返し実行されてもよい。

[0097] まず検出部８０が、検出されるユーザの視線に基づいて、表示されているＡＲ空間画像７０内における注視点Ｐの位置を示す視線情報を生成する（Ｓ１０１）。Ｓ１０１に示す処理では例えば表示されているＡＲ空間画像７０に注視点Ｐの位置を示す目印が重畳された視線情報が生成されてもよい。

[0098] そして実空間音声供給部８４が、実空間音声を表す実空間音声データを生成する（Ｓ１０２）。Ｓ１０２に示す処理では、前回のループにおける実空間音声データの生成タイミングから現在時刻までの実空間音声を表す実空間音声データが生成されてもよい。

[0099] そして実空間音声供給部８４は、Ｓ１０１に示す処理で生成された視線情報に対応付けられる方向から入力される音声が強調されるよう、マイク部４８の指向方向を変更する（Ｓ１０３）。そのため本処理例では、Ｓ１０１に示す処理で生成された視線方向に応じた方向から入力される音声が強調された実空間音声を表す実空間音声データが、次回のループにおけるＳ１０２に

示す処理で生成されることとなる。

[0100] そして実空間映像取得部９２が、カメラ部４４が撮影するフレーム画像を取得する（Ｓ１０４）。以下Ｓ１０４に示す処理で取得されるフレーム画像を実空間画像と呼ぶこととする。

[0101] そしてＨＭＤ１２が、Ｓ１０１に示す処理で生成された視線情報、Ｓ１０２に示す処理で生成された実空間音声データ、及び、Ｓ１０４に示す処理で取得された実空間画像をエンタテインメント装置１４に送信する。ここで機能的には、視線情報送信部８２が視線情報を送信し、実空間音声送信部８６が実空間音声データを送信し、実空間映像送信部９４が実空間画像を送信する。

[0102] そしてエンタテインメント装置１４がこれらの視線情報、実空間音声データ、及び、実空間画像を受信する（Ｓ１０５）。

[0103] そしてＡＲ空間映像生成部１１６が、Ｓ１０５に示す処理で受信した実空間画像を実空間部分７２として含みＶＲ空間映像供給部１１４が供給する画像をＶＲ空間部分７４として含むＡＲ空間画像７０を生成する（Ｓ１０６）。

[0104] そしてバランス決定部１０４が、Ｓ１０５に示す処理で受信した視線情報に基づいて、ＶＲ空間音声と実空間音声とのバランスを決定する（Ｓ１０７）。

[0105] そして合成音声生成部１０８が、Ｓ１０７に示す処理で決定されたバランスで、ＶＲ空間音声供給部１０６が供給するＶＲ空間音声とＳ１０５に示す処理で受信した実空間音声データが表す音声とを合成した合成音声を表す合成音声データを生成する（Ｓ１０８）。

[0106] そしてエンタテインメント装置１４は、Ｓ１０６に示す処理で生成されたＡＲ空間画像７０及びＳ１０８に示す処理で生成された合成音声データをＨＭＤ１２に送信する。ここで機能的には、合成音声送信部１１０が合成音声データを送信し、ＡＲ空間映像送信部１１８がＡＲ空間画像７０を送信する。

[0107] そしてＨＭＤ１２がこれらの合成音声データ、及び、ＡＲ空間画像７０を受信する（Ｓ１０９）。

[0108] そしてＡＲ空間映像表示制御部９８が、Ｓ１０９に示す処理で受信したＡＲ空間画像７０を表示部３８に表示させるとともに、合成音声出力部９０が、Ｓ１０９に示す処理で受信した合成音声データが表す合成音声を音声出力部４２から放音する（Ｓ１１０）。そしてＳ１０１に示す処理に戻る。

[0109] そして上述のように本処理例では所定の時間間隔でＳ１０１～Ｓ１１０に示す処理が繰り返し実行されることとなる。

[0110] なお実空間音声データの生成方法は上述の方法に限定されない。例えば上述のＳ１０２に示す処理において、実空間音声供給部８４が、マイク部４８が取得する音声に対して、直前のＳ１０１に示す処理で生成された視線情報に対応付けられる方向から入力される音声を強調する処理を実行してもよい。そして実空間音声供給部８４が、強調する処理が実行された音声を表す実空間音声データを生成してもよい。

[0111] またＶＲ空間音声の音量と実空間音声の音量とのバランスの決定方法は上述の方法に限定されない。例えばＳ１０７に示す処理において、直前のＳ１０６に示す処理で生成されたＡＲ空間画像７０と、その前のＳ１０５に示す処理で受信された視線情報と、に基づいて、合成音声におけるＶＲ空間音声と実空間音声とのバランスが決定されてもよい。例えばＳ１０６に示す処理で生成されたＡＲ空間画像７０内における、Ｓ１０５に示す処理で受信された視線情報に対応付けられる注視点Ｐの位置が特定されてもよい。そして注視点ＰがＶＲ空間部分７４に存在するか実空間部分７２に存在するかに応じて合成音声におけるＶＲ空間音声と実空間音声とのバランスが決定されてもよい。

[0112] また例えば上述のＳ１０９に示す処理で、エンタテインメント装置１４は、Ｓ１０６に示す処理で生成されたＡＲ空間画像７０及びＳ１０８に示す処理で生成された合成音声データを含むストリームデータをＨＭＤ１２に送信してもよい。そしてＡＲ空間映像表示制御部９８が、ＨＭＤ１２が受信した

当該ストリームデータからＡＲ空間画像７０を抽出してもよい。また合成音声出力部９０が、ＨＭＤ１２が受信した当該ストリームデータから合成音声データを抽出してもよい。そしてＡＲ空間映像表示制御部９８が、抽出されたＡＲ空間画像７０を表示部３８に表示させ、合成音声出力部９０が、抽出された合成音声データが表す合成音声を音声出力部４２から放音してもよい。

[0113] なお、本発明は上述の実施形態に限定されるものではない。

[0114] 例えば、注視点Ｐが含まれる部分以外に対応付けられる音声がミュートされた合成音声が生成されてもよい。例えば注視点Ｐの位置がＶＲ空間部分７４内である場合は、実空間音声がミュートされた合成音声が生成されてもよい。また例えば、注視点Ｐの位置が実空間部分７２内である場合は、ＶＲ空間音声がミュートされた合成音声が生成されてもよい。

[0115] また例えば複数の仮想オブジェクトの像がＶＲ空間部分７４としてＡＲ空間画像７０に含まれていてもよい。そしてある仮想オブジェクトの像に注視点Ｐが含まれる場合に、当該仮想オブジェクトに対応付けられる音声の音量が他の仮想オブジェクトに対応付けられる音声の音量よりも大きい合成音声が生成されてもよい。具体的には例えば、仮想オブジェクトがキャラクタであり、あるキャラクタの像に注視点Ｐが含まれる場合に、当該キャラクタの声の音量が他のキャラクタの声の音量よりも大きい合成音声が生成されてもよい。

[0116] また例えば、音声出力部４２は、ＨＭＤ１２とは別体であるヘッドホンやイヤホンであっても構わない。

[0117] また例えば、マイク部４８が取得する音声の代わりにマイク２０ｂが取得する音声が上述の実空間音声として用いられてもよい。

[0118] また以上の説明における図７においてエンタテインメント装置１４の機能として示されている示す機能の一部又は全部が、ＨＭＤ１２や中継装置１６において実装されても構わない。具体的には例えばＨＭＤ１２が、表示部３８に表示されているＡＲ空間画像７０と視線検出部４６が検出するユーザの

視線に基づいて、合成音声におけるＶＲ空間音声と実空間音声とのバランスを決定してもよい。そしてＨＭＤ１２が、エンタテインメント装置１４から受信するＶＲ空間音声とマイク部４８が取得する実空間音声とを合成して合成音声を生成してもよい。

[0119] また上述のように本発明は、表示部３８に複合現実空間（ＭＲ空間）の映像が表示される場合についても適用可能である。複合現実の技術においては、拡張現実の技術よりも、仮想現実空間に配置された仮想オブジェクトや仮想現実空間の環境と実空間に配置された物体や実空間の環境とのインタラクションがより重要となる。例えば、複合現実空間においては、仮想オブジェクトであるキャラクタと一緒にユーザが実空間に配置されたテレビを視聴することや、ユーザがキャラクタと自然な形で会話することが考えられる。また複合現実空間においては、例えば実空間音声を用いてキャラクタとユーザとが遊んだり、実空間音声に対するキャラクタの反応をユーザが楽しんだりすることも考えられる。

[0120] ここで例えば、ＶＲ空間音声とユーザの視線の方向の音声が強調された実空間音声とが合成された合成音声が生成される場合に、ＶＲ空間映像供給部１１４が、ユーザの視線の方向に存在する物体を向く仮想オブジェクトが配置された画像を供給してもよい。そしてＡＲ空間映像生成部１１６が、ユーザの視線の方向に存在する実空間内の物体を向くキャラクタ等の仮想オブジェクトの画像をＶＲ空間部分７４として含む複合空間の画像を生成してもよい。そしてこのようにして生成される複合空間の画像がＨＭＤ１２に送信され、表示部３８に表示されてもよい。

[0121] またあるいは、ＶＲ空間音声供給部１０６が、ユーザの視線の方向に存在する実空間内の物体に対応付けられる音声を供給してもよい。具体的には例えば、ユーザの視線の方向に存在する実空間内の物体を会話における話題とする音声が供給されてもよい。そして合成音声生成部１０８が、実空間内の物体を会話における話題とする音声と実空間音声とを合成して合成音声を生成してもよい。そしてこのようにして生成される合成音声がＨＭＤ１２に送

信され、音声出力部４２から出力されてもよい。

[0122] またこの場合に、ＡＲ空間映像生成部１１６が、ユーザの視線の方向に存在する実空間内の物体に関する話題の会話を行う仮想オブジェクトの画像をＶＲ空間部分７４として含む複合空間の画像を生成してもよい。またＡＲ空間映像生成部１１６が、ユーザの視線の方向に存在する実空間内の物体に応じた動作を行う仮想オブジェクトの画像をＶＲ空間部分７４として含む複合空間の画像を生成してもよい。

[0123] 例えば以上のようにすれば、複合現実空間における臨場感が向上することとなる。複合現実の技術では、上述の実空間音声が実空間の環境を構成する重要な一要素となると考えられる。そのため複合現実の技術においては、上述のようにして複合現実空間の臨場感を向上させることはより重要となる。

[0124] また、上記の具体的な文字列や数値及び図面中の具体的な文字列や数値は例示であり、これらの文字列や数値には限定されない。

請求の範囲

[請求項1] 仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成する画像生成部と、

前記仮想現実空間の音声と前記ユーザの周辺の実空間の音声とを合成して合成音声を生成する合成音声生成部と、

前記ユーザの耳の近傍に配置されるスピーカから放音される前記合成音声を出力する音声出力部と、

を含むことを特徴とする音声出力装置。

[請求項2] 前記合成音声生成部は、前記ユーザの視線に応じた音量のバランスで前記仮想現実空間の音声と前記実空間の音声とを合成して前記合成音声を生成する、

ことを特徴とする請求項１に記載の音声出力装置。

[請求項3] 前記合成音声生成部は、前記仮想現実空間の音声と、前記ユーザの視線の方向の音声が強調された前記実空間の音声と、を合成して前記合成音声を生成する、

ことを特徴とする請求項１又は２に記載の音声出力装置。

[請求項4] 前記仮想現実空間の音声は、記憶されている音声データ又は受信する音声ストリームが表す音声である、

ことを特徴とする請求項１から３のいずれか一項に記載の音声出力装置。

[請求項5] ユーザの周辺を撮影するカメラ部と、

前記ユーザの周辺の音声を取得するマイク部と、

仮想現実空間の画像が占める部分と前記カメラで撮影された実空間の画像が占める部分とを含む拡張現実空間又は複合現実空間の画像を表示する表示部と、

前記仮想現実空間の音声と前記マイクが取得する実空間の音声との

合成音声を前記ユーザの耳の近傍に配置されるスピーカを介して出力する音声出力部と、

を含むことを特徴とするヘッドマウントディスプレイ。

[請求項6] 仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成するステップと、

前記仮想現実空間の音声と前記ユーザの周辺の実空間の音声とを合成して合成音声を生成するステップと、

前記ユーザの耳の近傍に配置されるスピーカから放音される前記合成音声を出力するステップと、

を含むことを特徴とする音声出力方法。

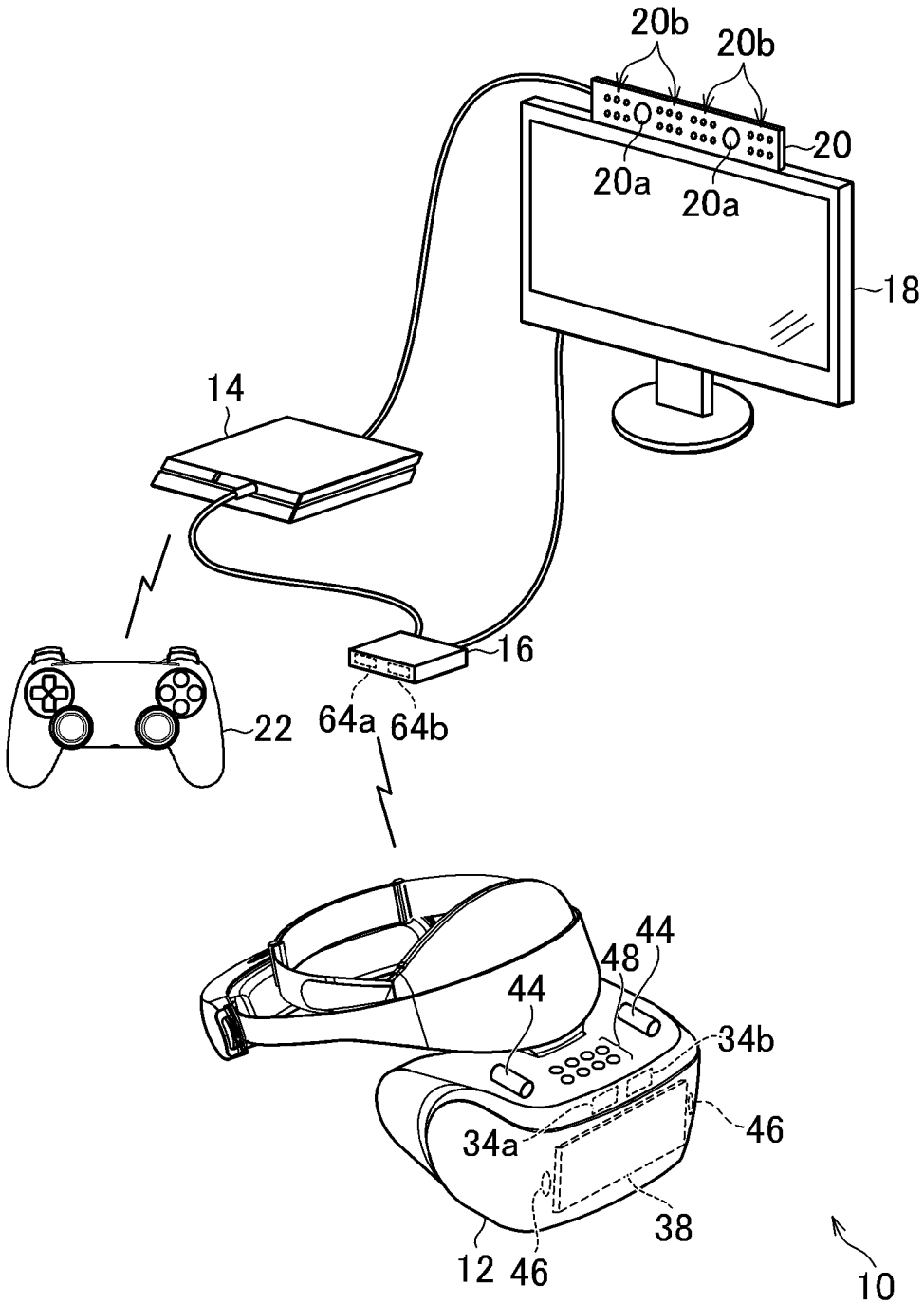
[請求項7] 仮想現実空間の画像が占める部分とユーザの周辺を撮影するカメラで撮影された実空間の画像が占める部分とを含む、ヘッドマウントディスプレイの表示部に表示される拡張現実空間又は複合現実空間の画像を生成する手順と、

前記仮想現実空間の音声と前記ユーザの周辺の実空間の音声とを合成して合成音声を生成する手順と、

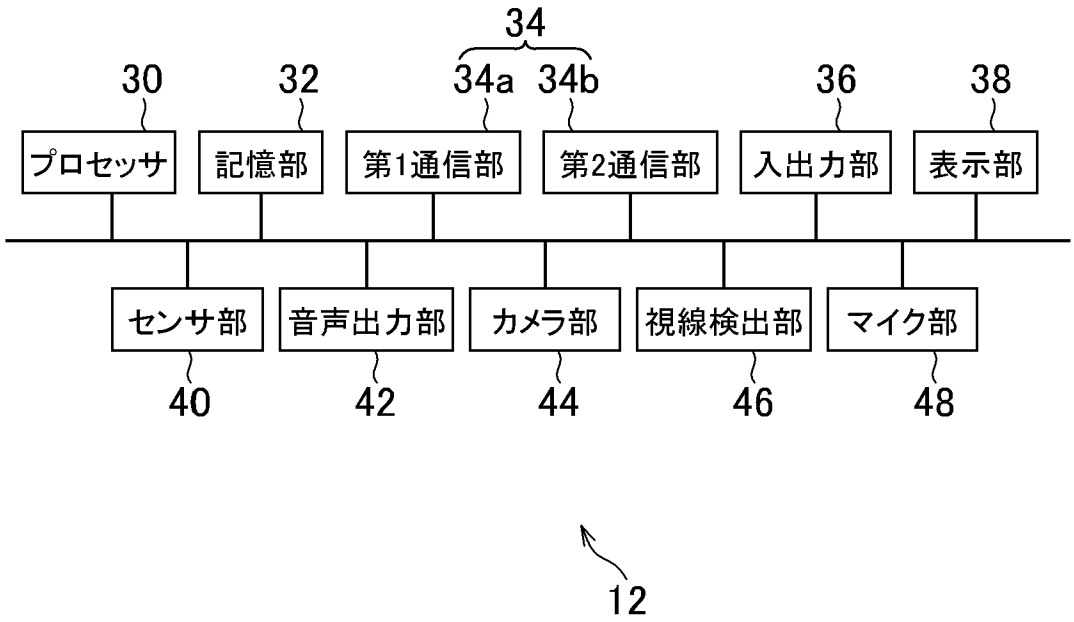
前記ユーザの耳の近傍に配置されるスピーカから放音される前記合成音声を出力する手順と、

をコンピュータに実行させることを特徴とするプログラム。

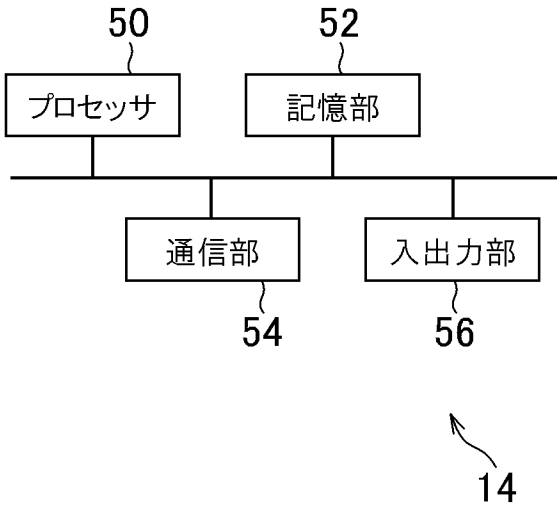
[図1]


20b
20b
20
20a
20a
18
14
16
22
64a
64b
44
48
44
34b
46
34a
38
12
46
10

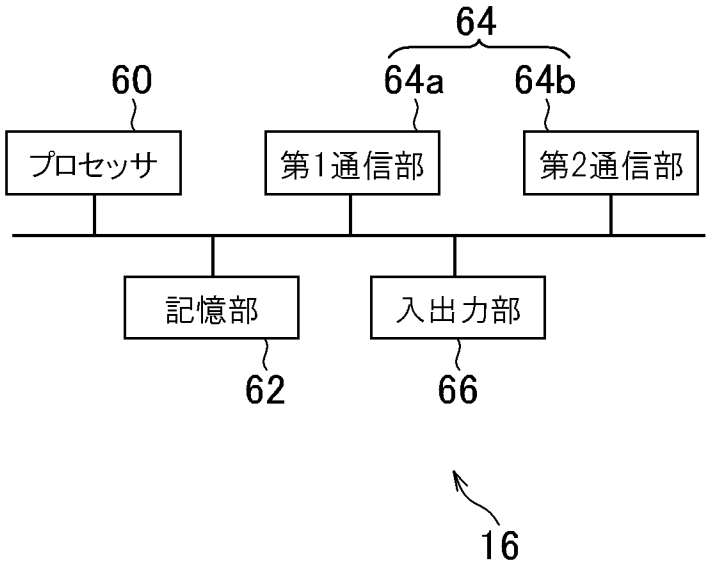
[図2A]


34
30
32
34a
34b
36
38
プロセッサ
記憶部
第1通信部
第2通信部
入出力部
表示部
センサ部
音声出力部
カメラ部
視線検出部
マイク部
40
42
44
46
48
12

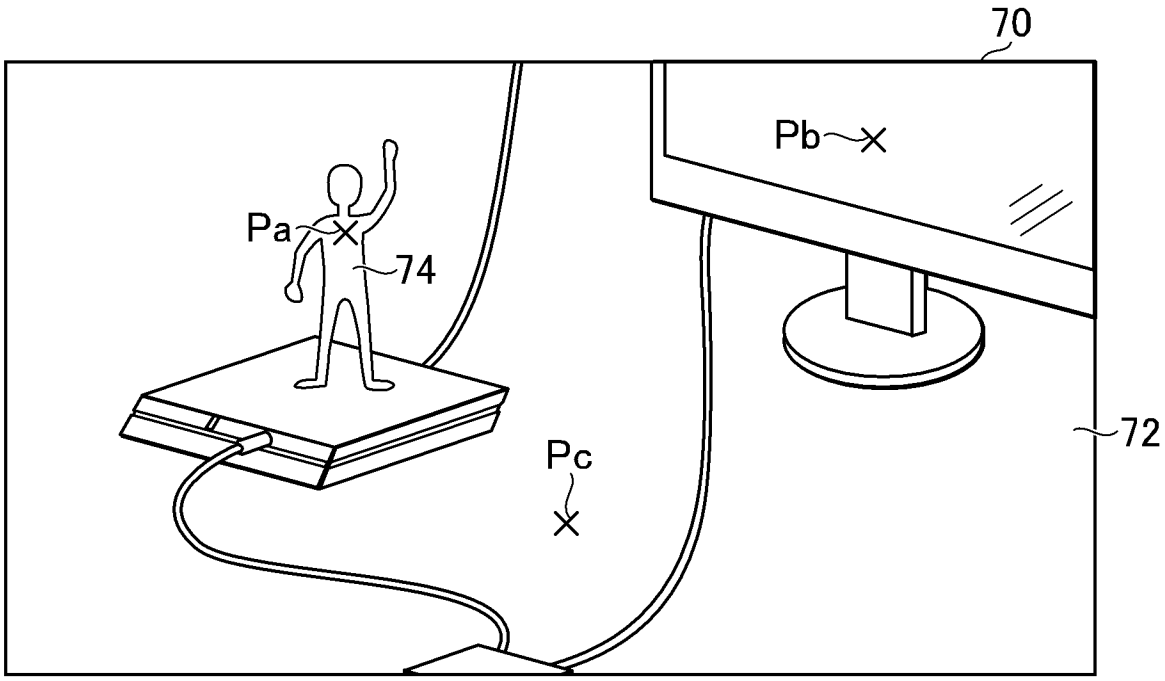
[図2B]


50
52
プロセッサ
記憶部
通信部
入出力部
54
56
14

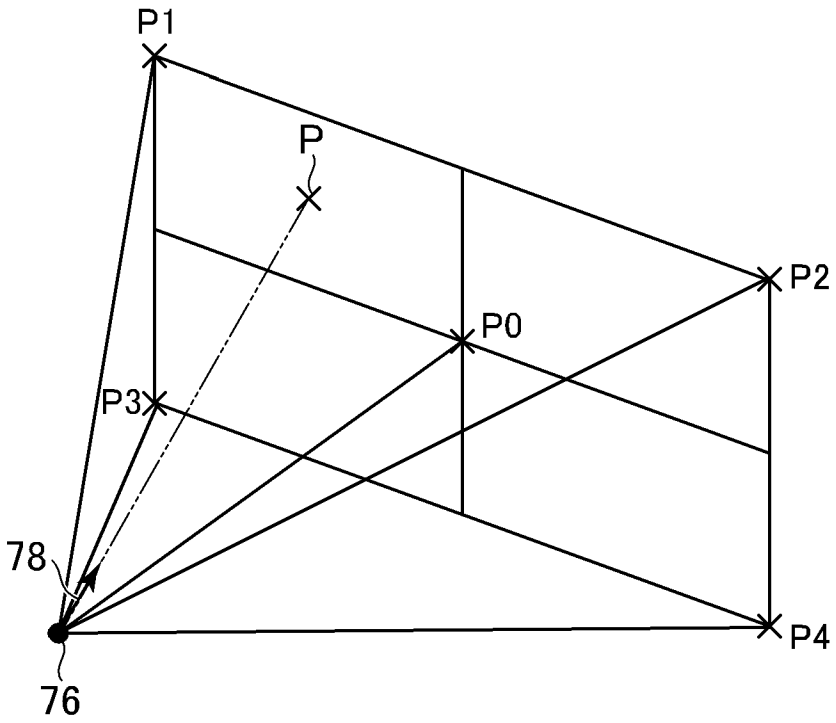
[図2C]


64
60
64a
64b
プロセッサ
第1通信部
第2通信部
記憶部
入出力部
62
66
16

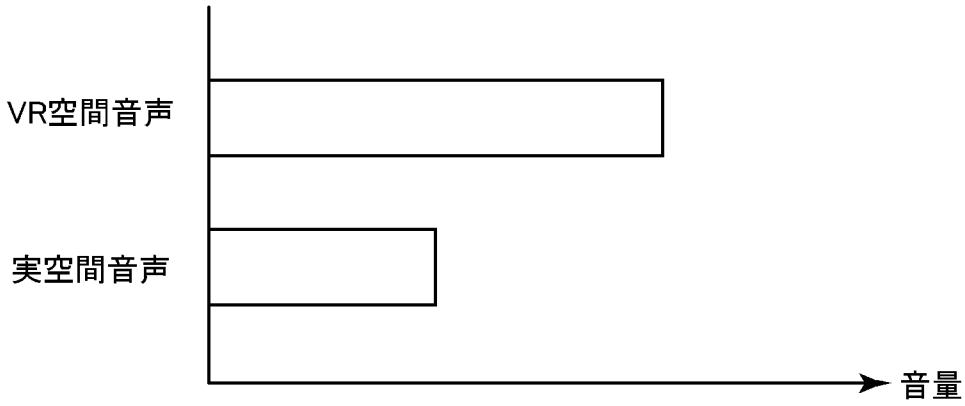
[図3]


70
Pb
Pa
74
Pc
72

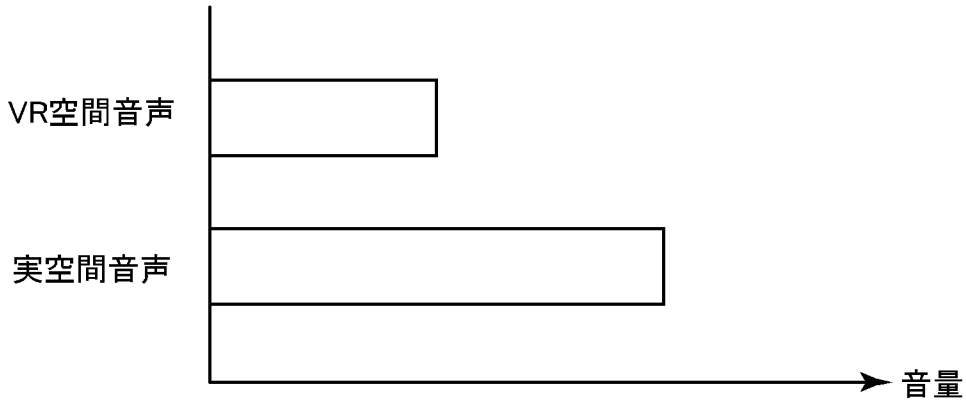
[図4]



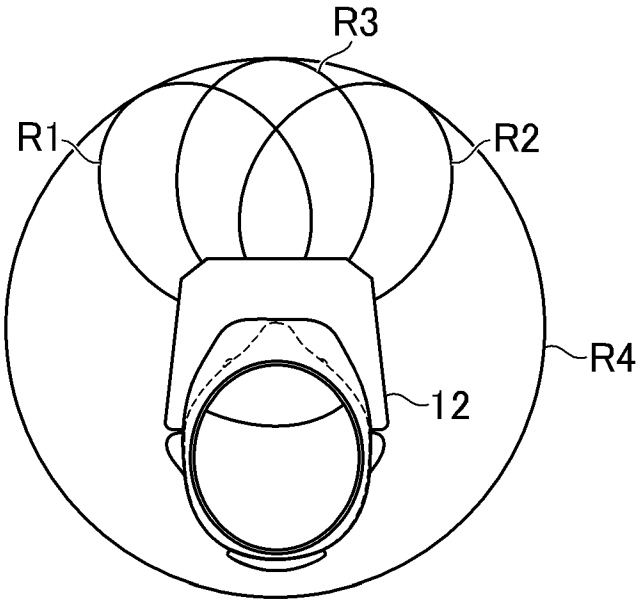
[図5A]



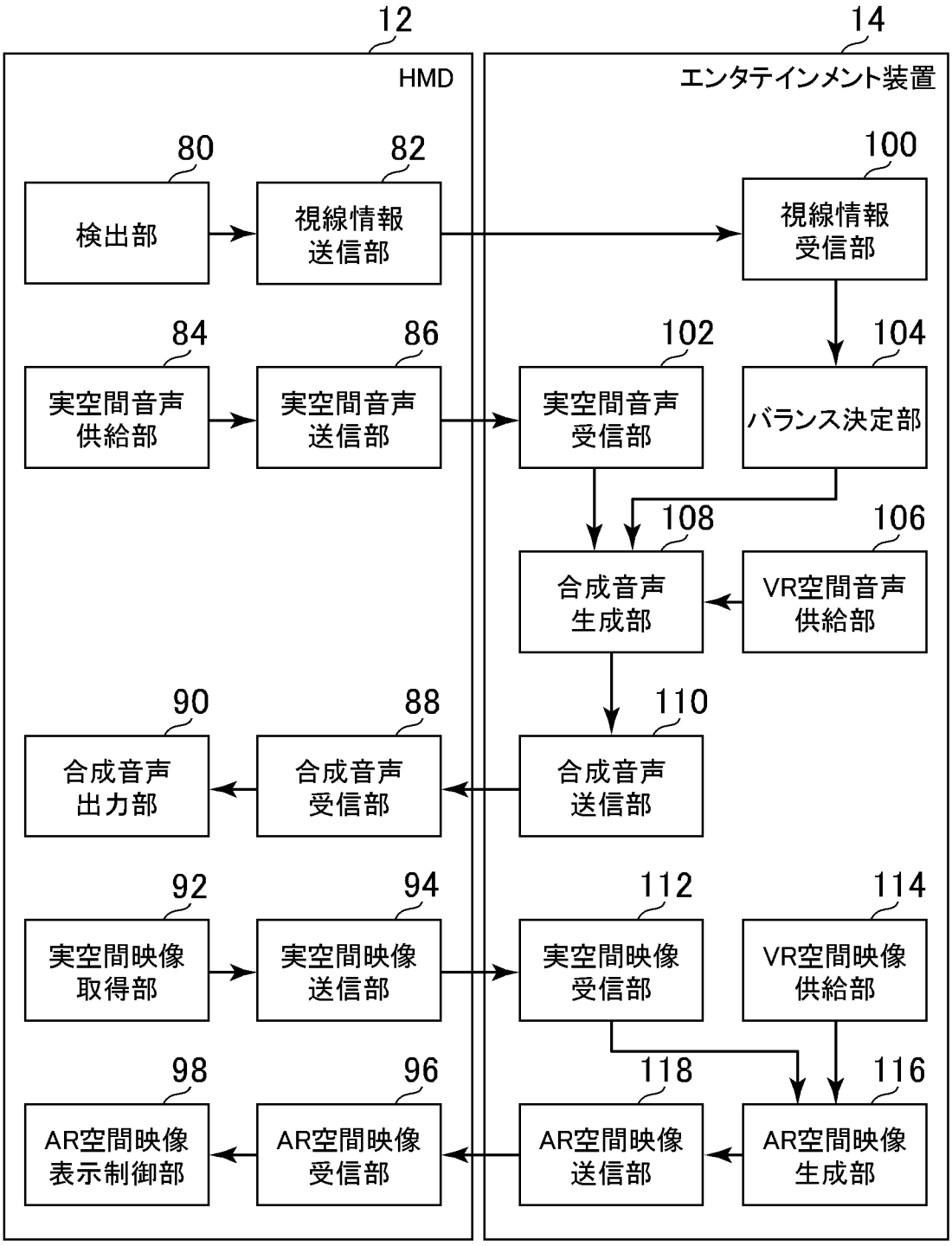
[図5B]



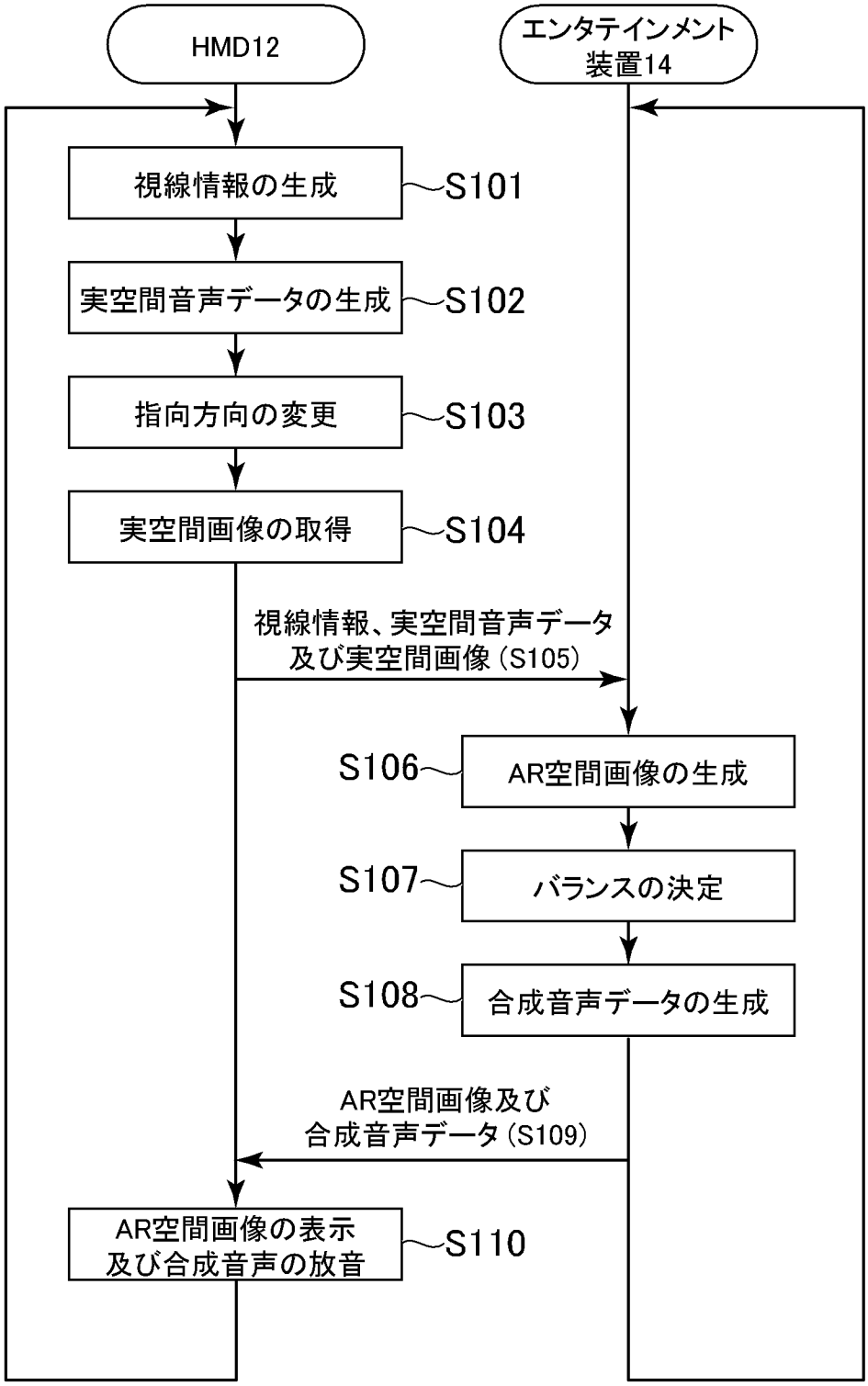
[図6]



[図7]


12
HMD
14
エンタテインメント装置
80
検出部
82
視線情報
送信部
100
視線情報
受信部
84
実空間音声
供給部
86
実空間音声
送信部
102
実空間音声
受信部
104
バランス決定部
108
合成音声
生成部
106
VR空間音声
供給部
90
合成音声
出力部
88
合成音声
受信部
110
合成音声
送信部
92
実空間映像
取得部
94
実空間映像
送信部
112
実空間映像
受信部
114
VR空間映像
供給部
98
AR空間映像
表示制御部
96
AR空間映像
受信部
118
AR空間映像
送信部
116
AR空間映像
生成部

[図8]


HMD12
エンタテインメント装置14
視線情報の生成
S101
実空間音声データの生成
S102
指向方向の変更
S103
実空間画像の取得
S104
視線情報、実空間音声データ及び実空間画像（S105）
S106
AR空間画像の生成
S107
バランスの決定
S108
合成音声データの生成
AR空間画像及び合成音声データ（S109）
AR空間画像の表示及び合成音声の放音
S110

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2017/047225

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl. G06F3/16(2006.01)i, G06F3/01(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl. G06F3/16, G06F3/01

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2018
Registered utility model specifications of Japan	1996-2018
Published registered utility model applications of Japan	1994-2018

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2016-033764 A (SEIKO EPSON CORPORATION) 10 March 2016, paragraphs [0018], [0019], [0024], [0026], [0027], [0039], [0051], [0053], [0054], [0074] & US 2016/0034251 A1, paragraphs [0037], [0038], [0044], [0045], [0047]-[0050], [0067], [0068], [0081]-[0084], [0088]-[0094], [0130]-[0132]	1-7

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search 30 January 2018	Date of mailing of the international search report 13 February 2018
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2017/047225

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2016-208348 A (SEIKO EPSON CORPORATION) 08 December 2016, paragraph [0140] & US 2016/0313973 A1, paragraph [0212]	1-7
Y	JP 10-153946 A (MITSUBISHI ELECTRIC CORPORATION) 09 June 1998, paragraphs [0047]-[0056], [0077]-[0081] (Family: none)	2-3

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１７／０４７２２５

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F3/16(2006.01)i, G06F3/01(2006.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　G06F3/16, G06F3/01

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１８年
日本国実用新案登録公報	１９９６－２０１８年
日本国登録実用新案公報	１９９４－２０１８年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2016-033764 A（セイコーエプソン株式会社）2016.03.10, 段落［００１８］－［００１９］，［００２４］，［００２６］－［００２７］，［００３９］，［００５１］，［００５３］－［００５４］，［００７４］ & US 2016/0034251 A1, 段落［００３７］－［００３８］，［００４４］－［００４５］，［００４７］－［００５０］，［００６７］－［００６８］，［００８１］－［００８４］，［００８８］－［００９４］，［０１３０］－［０１３２］	1-7

☑　Ｃ欄の続きにも文献が列挙されている。　　☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
３０．０１．２０１８	１３．０２．２０１８

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｅ　３９７９
円子　英紀
電話番号　０３－３５８１－１１０１　内線　３５２１

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2016-208348 A (セイコーエプソン株式会社) 2016.12.08, 段落［０１４０］ & US 2016/0313973 A1, 段落［０２１２］	1-7
Y	JP 10-153946 A (三菱電機株式会社) 1998.06.09, 段落［００４７］－［００５６］，［００７７］－［００８１］ (ファミリーなし)	2-3