

(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
2022年3月3日(03.03.2022)



(10) 国際公開番号

WO 2022/044342 A1

- (51) 国際特許分類:
H04R 1/40 (2006.01) *H04R 3/00* (2006.01)
- (21) 国際出願番号: PCT/JP2020/032962
- (22) 国際出願日: 2020年8月31日(31.08.2020)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: マクセル株式会社 (MAXELL, LTD.)
[JP/JP]; 〒6188525 京都府乙訓郡大山崎町大山崎小泉1番地 Kyoto (JP).
- (72) 発明者: 秋山 仁 (AKIYAMA Hitoshi); 〒6188525
京都府乙訓郡大山崎町大山崎小泉1番地
マクセル株式会社内 Kyoto (JP). 川前 治
(KAWAMAE Osamu); 〒6188525 京都府乙訓

郡大山崎町大山崎小泉1番地 マクセル株式会社内 Kyoto (JP). 伊藤 保 (ITO Tamotsu);
〒6188525 京都府乙訓郡大山崎町大山崎小泉1
番地 マクセル株式会社内 Kyoto (JP).

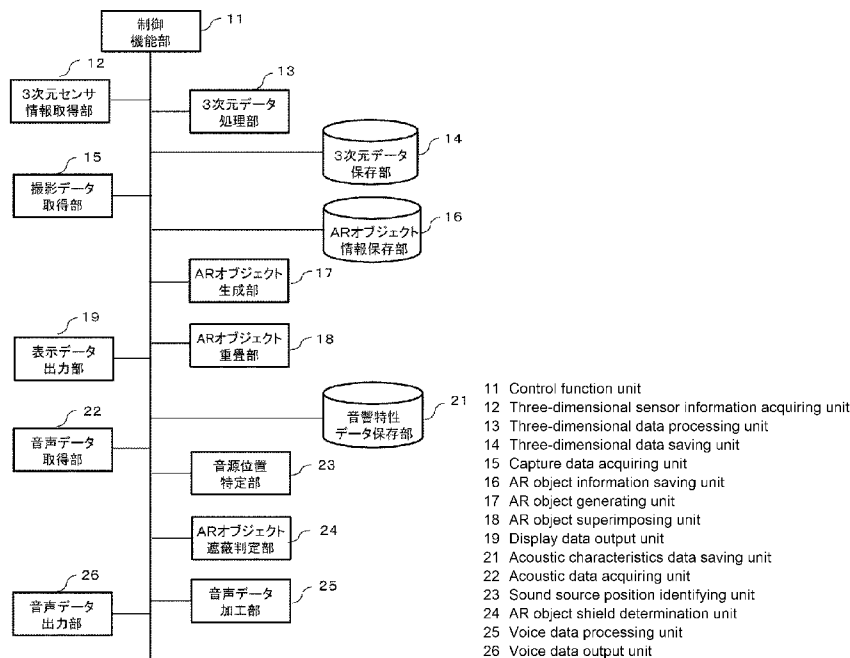
(74) 代理人: 青稜特許業務法人 (SEIRYO I.P.C.);
〒2200004 神奈川県横浜市西区北幸二丁目
10番27号 Kanagawa (JP).

(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, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

(54) Title: HEAD-MOUNTED DISPLAY AND VOICE PROCESSING METHOD THEREFOR

(54) 発明の名称: ヘッドマウントディスプレイ及びその音声処理方法

図2B



(57) Abstract: This HMD displays a video of a real space captured by an imaging unit 71 and an AR object generated by an AR object generating unit 17 superimposed on each other on a display unit 72. Correspondingly, voice generated in the real space or voice generated by the AR object is output from a voice output unit 82. At this time, a voice data processing unit 25, in accordance with the positional relationship between an object present in the real space and an AR object disposed in superimposition therewith, subjects voice data of at least one of the voice from the real space and the

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,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, 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)）

voice generated by the AR object to processing/treatment, taking the acoustic characteristics of a surrounding object or a surrounding AR object into consideration. An acoustic characteristics data saving unit 21 saves the acoustic characteristics of the AR object in patterned form using the shape of each constituent element as a parameter. In this way, it is possible to perform in real-time a realistic display of an AR object and voice output taking the arrangement of the AR object into consideration.

(57) 要約：HMDは、撮像部71で撮影した実空間の映像と、ARオブジェクト生成部17で生成したARオブジェクトとを重畳して表示部72に表示する。これに合わせ、実空間にて発せられた音声またはARオブジェクトが発する音声を音声出力部82から出力する。その際、音声データ加工部25は、実空間に存在する物体とこれに重畳して表示されるARオブジェクトの位置関係に応じて、実空間からの音声とARオブジェクトが発する音声の少なくとも一方の音声データを、周囲の物体または周囲のARオブジェクトの音響特性を加味して加工処理する。音響特性データ保存部21は、ARオブジェクトの音響特性を、構成要素毎にその形状をパラメータとしてパターン化して保存する。これにより、現実感のあるARオブジェクトの表示と、ARオブジェクトの配置を考慮した音声出力を、リアルタイムに実現することができる。

明 細 書

発明の名称：ヘッドマウントディスプレイ及びその音声処理方法
技術分野

[0001] 本発明は、ヘッドマウントディスプレイ及びその音声処理方法に関する。

背景技術

[0002] 近年、ヘッドマウントディスプレイ（以下、HMD：Head Mounted Display）が、実用化されている。HMDは眼鏡形式の表示画面上に、コンピュータなどで作成した仮想現実（VR：Virtual Reality）の画像（VRオブジェクト）を表示し、ユーザが、まるで仮想世界にいるような感覚を体験することができる。さらには、現実空間の映像に、コンピュータで生成した拡張現実（AR：Augmented Reality）の画像（ARオブジェクト）を重畳して表示し、ARオブジェクトを、あたかも現実空間に存在するが如く提示することも可能となっている。

[0003] 上記VRオブジェクトを用いたゲームにおいて、特許文献1には、仮想空間のエンターテインメント性を向上させるため、音反射オブジェクトという球形のオブジェクトを規定し、この音反射オブジェクトによる反射が発生するとしてマイクに入力したユーザ音声を加工し、ヘッドフォンから出力させる構成が開示される。

先行技術文献

特許文献

[0004] 特許文献1：特開2018-11193号公報

発明の概要

発明が解決しようとする課題

[0005] 従来の一般的なHMDは、撮影手段（カメラ）により実空間の映像を取り込めるビデオ透過型（ビデオシースルー型）の場合、HMDの表示画面上に、撮影した実空間の映像と、ARオブジェクトを重畳して表示している。その際、ユーザに提供する音声（効果音を含む音響空間）は、ARオブジェク

トの存在有無にかかわらず、HMDに搭載した音声入力手段（マイク）で取得した実空間の音声をそのまま出力している。このため、ユーザが聞く音声には、ユーザが視認するARオブジェクトの影響が反映されず、ARオブジェクトの配置状況によっては、ユーザに対し違和感を与えることがあった。

[0006] これに関し、ARオブジェクトの存在を仮定し、そのARオブジェクトの存在による音響特性の影響を、コンピュータにより音場シミュレーションをする有限差分時間領域（FDTD: Finite Difference Time Domain）法などの手法を使用することもできる。しかしながら、このようなシミュレーションは演算負荷が大きく、処理時間が増大して、リアルタイムで処理することは困難である。

[0007] 前記特許文献1には、非透過型のHMDを用い、仮想空間内において、VRオブジェクトを配置し、音反射オブジェクトの存在により、音響効果を変化させる手法が記載されている。しかしながら、音反射オブジェクト以外のVRオブジェクトやARオブジェクトを配置したことによる、実空間の音源で発生した音声への影響は考慮されておらず、ユーザに与える違和感を解消することにはならない。また、特許文献1では、音声加工対象である音源は音反射オブジェクトの中心に存在すると規定されているので、音源位置とARオブジェクトの位置関係が変化するような場合には適用が困難である。

[0008] 本発明の目的は、実空間内にARオブジェクトを配置するとともに、実空間で発生した音声をARオブジェクトの配置を考慮してリアルタイムに加工し、ユーザに提示するヘッドマウントディスプレイ及びその音声処理方法を提供することである。

課題を解決するための手段

[0009] 前記課題を解決するため、本発明のヘッドマウントディスプレイは、映像処理として、実空間の映像を撮影する撮像部と、ARオブジェクトを生成するARオブジェクト生成部と、前記撮影した実空間の映像と、前記生成したARオブジェクトとを重畳するARオブジェクト重畳部と、前記重畳された実空間の映像とARオブジェクトとを表示する表示部と、を備える。また

音声処理として、実空間にて発せられた音声を入力する音声入力部と、前記 A R オブジェクト重畳部により配置された実空間に存在する物体とこれに重畳して表示される A R オブジェクトの位置関係に応じて、実空間からの音声と A R オブジェクトが発する音声の少なくとも一方の音声データを、周囲の物体または周囲の A R オブジェクトの音響特性を加味して加工処理する音声データ加工部と、前記音声データ加工部にて加工処理された実空間からの音声または A R オブジェクトの音声を出力する音声出力部と、を備える。

発明の効果

[0010] 本発明によれば、現実感のある A R オブジェクトの表示と、A R オブジェクトの配置を考慮した音声出力を、リアルタイムに実現することができる。

図面の簡単な説明

- [0011] [図1A]本発明で扱う音声処理の一例を示す模式図。
[図1B]本発明で扱う音声処理の一例を示す模式図。
[図1C]本発明で扱う音声処理の他の例を示す模式図。
[図1D]本発明で扱う音声処理の他の例を示す模式図。
[図2A]実施例 1 における H M D のハードウェア構成を示す図。
[図2B]実施例 1 における H M D の機能ブロック構成を示す図。
[図3]音響特性を表記法の例を示す図。
[図4]A R オブジェクトの構成要素のインパルス応答を記述したテーブル。
[図5]インパルス応答の波形を模式的に示す図。
[図6A]A R オブジェクト重畳表示時の映像・音声処理の例を示す図。
[図6B]A R オブジェクト重畳表示時の映像・音声処理の例を示す図。
[図6C]A R オブジェクト重畳表示時の映像・音声処理の例を示す図。
[図7A]A R オブジェクト重畳表示時の映像・音声処理の他の例を示す図。
[図7B]A R オブジェクト重畳表示時の映像・音声処理の他の例を示す図。
[図8]A R オブジェクト重畳表示時の映像処理手順を示すフローチャート。
[図9]A R オブジェクト重畳表示時の音声処理手順を示すフローチャート。
[図10]実施例 2 における H M D の機能ブロック構成を示す図。

[図11A] A Rオブジェクト重畳表示の例を示す図。

[図11B] A Rオブジェクト重畳表示の例を示す図。

[図12] 重畳表示時の映像処理手順を示すフローチャート。

[図13] 重畳表示時の音声処理手順を示すフローチャート。

[図14] 実施例3のHMDをユーザが装着した状態を示す図。

[図15] 実施例4におけるHMDの機能ブロック構成を示す図。

[図16] 重畳表示時の音声処理手順を示すフローチャート。

[図17] 実施例5におけるHMDシステムの全体構成を示す図。

発明を実施するための形態

[0012] 以下、本発明の実施形態について、図面を用いて説明する。

実施例 1

[0013] 図1 Aおよび図1 Bは、本発明で扱う音声処理の一例を示す模式図である。ユーザ10は、ヘッドマウントディスプレイ（HMD）1を装着して、人物30の口301から発する音声を聞いている状態を示している。HMD1は、ユーザ10の頭部に装着され、ユーザ10の視界はHMD1の表示画面で覆われている。HMD1の正面中央にはカメラ70が配置されており、ユーザ10の正面方向における実空間の映像を撮影している。HMD1はビデオ透過型HMDであり、撮影した映像は、HMD1の表示画面に表示される。

[0014] さらに、図示していないが、カメラ70の近傍には実空間の音声を取得するマイクロフォン（以下、「マイク」）が配置されている。マイクで取得した音声は、必要に応じて後述するように加工され、ヘッドフォン80からユーザ10に出力される。HMD1とヘッドフォン80は電氣的に接続されているが、構造的に一体化させてもよい。

[0015] 図1 Aは、ユーザ10と人物30の口301とを結ぶ線分上に、A Rオブジェクトが無い状態を示す。この状態では、HMD1のカメラ70で撮影した映像を表示することにより、ユーザ10は前方の人物30を視覚的に認識できる。また、人物30の口301から発する音声をHMD1のマイクで取

得し、そのまま加工せずにヘッドフォン８０から出力することにより、ユーザ１０は人物３０の音声として違和感なく聞くことができる。

[0016] 図１Ｂは、ユーザ１０と人物３０の口３０１とを結ぶ線分上に、ＡＲオブジェクト３５が配置された状態を示す。ここでのＡＲオブジェクト３５は、例えば衝立のように、音声の伝達を減衰させるような音響的障害物であることを想定する。この状態では、ＨＭＤ１のカメラ７０で人物３０を撮影できても、人物３０はＡＲオブジェクト３５の陰に隠れているので、ＨＭＤ１の表示画面に表示されず、ユーザ１０は人物３０を視覚的に認識することができない。一方、人物３０の口３０１から発する音声はＨＭＤ１のマイクで取得し、ヘッドフォン８０から出力することができる。しかし、図１Ａと同じように、取得した音声データをそのまま加工せずにヘッドフォン８０から出力すると、目に見えない人物３０の音声はＡＲオブジェクト３５（音響的障害物）の影響を受けずに聞こえることで、ユーザ１０は違和感を覚えることになる。

[0017] そこで本実施例では、実空間に存在しないＡＲオブジェクト３５を表示する場合、ＡＲオブジェクト３５の音響特性を加味して、人物３０の口３０１から発する音声を、あたかも実空間にＡＲオブジェクト３５が存在するが如く加工する。具体的には、ＡＲオブジェクト３５の形状や材質に応じて、音声データを減衰させる加工処理を施して出力する。その結果、ユーザ１０に与える音声の違和感を軽減させることができる。

[0018] また図１Ｃおよび図１Ｄは、本発明で扱う音声処理の他の例を示す模式図である。この例では、ＨＭＤ１を装着したユーザ１０は、人物３０の後方に位置し、口３０１から発する音声を聞いている状態である。

[0019] 図１Ｃは、ユーザ１０と人物３０の口３０１とを結ぶ直線上に、ＡＲオブジェクトが無い状態を示す。人物３０の口３０１から発する音声をＨＭＤ１のマイクで取得し、そのまま加工せずにヘッドフォン８０から出力することにより、ユーザ１０は人物３０の音声として違和感なく聞くことができる。ただしこの場合は、ユーザ１０は人物３０の後方に位置するので、図１Ａの

場合と比較して聞こえる音量は小さくなる。

[0020] 図1Dは、ユーザ10と人物30の口301とを結ぶ延長線上に、ARオブジェクト35'が配置された状態を示す。ここでのARオブジェクト35'は、例えば鏡面のように、音声の伝達を反射させるような音響的反射物であることを想定する。この状態では、人物30の口301から発する音声は、ARオブジェクト35'で反射されてユーザ10に届く分が存在する（図中、鎖線で示す）。よって、図1Cと同じように、マイクで取得した音声データをそのまま加工せずにヘッドフォン80から出力すると、ユーザ10は違和感を覚えることになる。

[0021] そこで本実施例では、実空間に存在しないARオブジェクト35'を表示する場合、人物30の口301から発する音声を、ARオブジェクト35'の音響特性（反射特性）を加味して加工する。具体的には、ARオブジェクト35'の形状や材質に応じて、音声データを増幅させる加工処理を施して出力する。その結果、ユーザ10に与える音声の違和感を軽減させることができる。

[0022] このように本実施例では、実空間に存在しないARオブジェクトを表示するとき、ARオブジェクトの音響特性（減衰特性／反射特性）を加味して音声データを加工して出力するものであるが、以下では、図1Bのように、ARオブジェクトが音響的障害物（減衰特性）である場合を中心に説明を行う。

以下、本実施例のヘッドマウントディスプレイ（HMD）の構成と動作を説明する。

[0023] [HMDのハードウェア構成]

図2Aは、ヘッドマウントディスプレイ（HMD）1のハードウェア構成を示す図である。HMD1は、主制御部2、システムバス3、記憶部4、センサ部5、通信処理部6、映像処理部7、音声処理部8、操作入力部9を含んで構成される。

[0024] 主制御部2は、所定の動作プログラムに従ってHMD1全体を制御するマ

イクロプロセッサユニットである。システムバス 3 は、主制御部 2 と HMD 1 内の各構成ブロックとの間で各種コマンドやデータなどの送受信を行うためのデータ通信路である。

[0025] 記憶部 4 は、HMD 1 の動作を制御するためのプログラムなどを記憶するプログラム部 4 1、動作設定値やセンサ部からの検出値やコンテンツを含むオブジェクトやライブラリからダウンロードしたライブラリ情報などの各種データを記憶する各種データ部 4 2、各種プログラム動作で使用するワークエリアなどの書き替え可能なプログラム機能部 4 3 を有する。

[0026] 記憶部 4 は、ネットワーク上からダウンロードした動作プログラムや前記動作プログラムで作成した各種データ等を記憶可能である。また、ネットワーク上からダウンロードした動画や静止画や音声等のコンテンツを記憶可能である。また、カメラによる撮影機能を使用して撮影した動画や静止画等のデータを記憶可能である。ここに記憶部 4 は、HMD 1 に外部から電源が供給されていない状態であっても記憶している情報を保持する必要がある。従って、例えば、フラッシュ ROM や SSD (Solid State Drive) などの半導体素子メモリ、HDD (Hard Disc Drive) などの磁気ディスクドライブ、等のデバイスが用いられる。なお、記憶部 4 に記憶された各動作プログラムは、ネットワーク上の各サーバ装置からのダウンロード処理により更新及び機能拡張することが可能である。

[0027] センサ部 5 は、HMD 1 の状態を検出するための各種のセンサ群である。センサ群には、GPS (Global Positioning System) 受信部 5 1、地磁気センサ部 5 2、3次元センサ部 5 3、加速度センサ部 5 4、ジャイロセンサ部 5 5 が含まれる。これらのセンサ群により、HMD 1 の位置、傾き、方角、動き、等を検出する。また、これ以外に、照度センサ、高度センサ、近接センサ等、をさらに備えてもよい。

[0028] このうち 3次元センサ部 5 3 は、対象物の各ポイントまでの距離を計測するものである。3次元センサ (3D スキャナー) には種々の方式があり、例えば、非接触式の光 (格子パターン) 投影法、非接触式のレーザ光切断方式

、レーザによる非接触式の位相差方式（フェイズシフト方式）などが可能である。本実施例ではいずれの方式でも構わない。

[0029] 通信処理部6は、L A N（Local Area Network）通信部61、電話網通信部62を有する。L A N通信部61は、アクセスポイント等を介してインターネット等のネットワーク網と接続され、ネットワーク網上の各ネットワークサーバ装置とデータの送受信を行う。アクセスポイント等との接続は、W i - F i（登録商標）等の無線接続で行われてよい。

[0030] 電話網通信部62は、移動体電話通信網の基地局等との無線通信により、電話通信（通話）及びデータの送受信を行う。基地局等との通信はW - C D M A（Wideband Code Division Multiple Access）（登録商標）方式やG S M（Global System for Mobile communications）方式、L T E（Long Term Evolution）方式、或いはその他の通信方式によって行われてよい。

[0031] L A N通信部61、電話網通信部62、は、それぞれ符号化回路や復号回路やアンテナ等を備える。また、通信処理部6が、赤外線通信部やB l u e t o o t h（登録商標）通信部など、他の通信部をさらに備えていてもよい。

[0032] 映像処理部7は、撮像部71、表示部72、映像演算部73を有する。撮像部71は、C C D（Charge Coupled Device）やC M O S（Complementary Metal Oxide Semiconductor）センサ等の電子デバイスを用いて、レンズから入力した光を電気信号に変換することにより、外景や対象物の映像データを取得するカメラ70（図1A～1D）である。

[0033] 表示部72は、画像データ（A Rオブジェクト）や映像データ（カメラ撮影映像）を、H M D 1を装着したユーザ10に提供する。ここでH M D 1は実空間を透視しないビデオ透過型なので、表示部72は、例えばバックライト型液晶ディスプレイや、自己発光型である有機E Lディスプレイなどの表示デバイスである。

[0034] 映像演算部73は、コンピュータによる拡張現実（A Rオブジェクト）の画像を生成する処理、撮像部71で取得した実空間の映像とA Rオブジェク

トとを重畳する処理等、映像に関する演算処理を行う。

[0035] 音声処理部 8 は、音声入力部 8 1、音声出力部 8 2、音声演算部 8 3 を有する。音声入力部 8 1 は、実空間の音やユーザの音声などを音声データに変換して入力するマイクである。音声出力部 8 2 は、ユーザに必要な音声情報等を出力するヘッドフォン 8 0（図 1 A～1 D）である。

[0036] 音声演算部 8 3 は、音源の位置検出や、AR オブジェクトの音響特性（材質・厚さ・断面積等による音響変化）に基づく音声データの加工処理など、音声に関する演算処理を行う。

[0037] 操作入力部 9 は、HMD 1 に対する操作指示の入力を受け付ける。操作入力部 9 は、ボタンスイッチ等を並べた操作キー、等で構成されるが、その他の操作デバイスをさらに備えてもよい。また、通信処理部 6 を利用し、有線通信または無線通信により接続された別体の携帯端末機器を用いて HMD 1 の操作を行ってもよい。さらには、映像処理部 7 の撮像部 7 1 の撮影映像を解析し、ジェスチャなどの動作で、HMD 1 の操作を行なってもよい。

[0038] なお、図 2 A に示した HMD 1 のハードウェア構成例は、本実施例に必須ではない構成も多数含んでいるが、これらが備えられていない構成であっても本実施例の効果を損なうことはない。

[0039] [HMD の機能ブロック構成]

図 2 B は、ヘッドマウントディスプレイ（HMD）1 の機能ブロック構成を示す図である。制御機能部 1 1 に含まれる各機能ブロックは、主に、図 2 A における主制御部 2 と、記憶部 4 のプログラム部 4 1 及びプログラム機能部 4 3 により実現される。各機能ブロックの機能について説明する。

[0040] 3 次元センサ情報取得部 1 2 は、センサ部 5 の 3 次元センサ部 5 3 からの情報を取得する。3 次元センサ部 5 3 からの情報には、HMD 1 から実空間における対象物の各ポイントまでの距離情報が含まれる。

[0041] 3 次元データ処理部 1 3 は、3 次元センサ情報取得部 1 2 が取得した情報（対象物の各ポイントまでの距離情報）に基づいて、対象物の形状を把握する。3 次元データ保存部 1 4 は、3 次元データ処理部 1 3 で得られた実空間

の３次元データを、記憶部４の各種データ部４２に保存する。本実施例では、この実空間の３次元データを予め取得して、３次元データ保存部１４により保存していることを前提としている。

[0042] 撮影データ取得部１５は、映像処理部７の撮像部７１（カメラ７０）により実空間を撮影し、撮影データを取得する。この機能により、直接実空間を視認できないビデオ透過型ＨＭＤ１を装着するユーザ１０に対して、実空間の映像を提供することができる。

[0043] ＡＲオブジェクト情報保存部１６は、ＡＲオブジェクトに関する情報を、記憶部４の各種データ部４２に保存する。ＡＲオブジェクトに関する情報には、ＡＲオブジェクトの画像に関する情報（形状・大きさ・配置等）や、ＡＲオブジェクトを構成する仮想の要素（以下、構成要素）に関する付帯情報（各構成要素の材質・厚さ・断面積等）が含まれる。また、通信処理部６のＬＡＮ通信部６１を介して、インターネット網に接続された外部のサーバより、ＡＲオブジェクトの情報及びＡＲオブジェクトの付帯情報を取得することもできる。

[0044] ＡＲオブジェクト生成部１７は、ＡＲオブジェクト情報保存部１６に保存されているＡＲオブジェクトの画像に関する情報に基づき、映像処理部７の映像演算部７３を用いて、ＡＲオブジェクトを生成する。生成するＡＲオブジェクトは、３次元データ保存部１４に保存されている実空間の形状情報に合わせて、その形状・大きさ・配置といった表示仕様を決定する。

[0045] ＡＲオブジェクト重畳部１８は、撮影データ取得部１５で取得した実空間の映像と、ＡＲオブジェクト生成部１７で生成したＡＲオブジェクトとを、映像処理部７の映像演算部７３を用いて重畳する。具体的には、ＡＲオブジェクトの前面（手前）に位置する実空間の映像は表示するが、ＡＲオブジェクトの背面（奥）に位置する実空間の映像は表示しない等のオクルージョン処理を行い、重畳処理を行う。

[0046] 表示データ出力部１９は、ＡＲオブジェクト重畳部１８により重畳された実空間の映像とＡＲオブジェクトの画像を、ＨＭＤ１の表示画面に表示する

。

- [0047] 音響特性データ保存部 21 は、AR オブジェクトの構成要素毎に、材質、厚さ、断面積等をパラメータに、音響特性（減衰特性／反射特性）にかかわる音響特性データを保存する。勿論、音響特性データ保存部 21 には、採用する AR オブジェクトの構成要素以外の材料に関する音響特性データも保存していることは言うまでもない。さらに、通信処理部 6 の LAN 通信部 61 を介して、インターネット網に接続された外部のサーバより、AR オブジェクトの構成要素の材質に関する音響特性を取得することもできる。
- [0048] 音声データ取得部 22 は、音声処理部 7 の音声入力部 81（マイク）で、実空間の音声を取得する。音源位置特定部 23 は、音源の位置を検出する。これには、3次元データ保存部 14 に保存されている実空間の配置情報と、撮影データ取得部 15 にて取得した実空間の映像から、音声処理部 8 の音声演算部 83 を用いて、音源の位置を特定する。
- [0049] AR オブジェクト遮蔽判定部 24 は、音源位置特定部 23 で取得した音声データの音源位置と、表示データ出力部 19 で表示する AR オブジェクトの表示位置との位置関係を用いて、前記音声データの音源位置とユーザ 10 の位置（すなわち HMD 1 の位置）とを結ぶ線分上に、前記 AR オブジェクトの表示位置が存在するかどうかを判定する。すなわち、ユーザ 10（HMD 1）から見て、AR オブジェクトにより音声データの音源位置が遮蔽されているかどうかを判定する。
- [0050] 音声データ加工部 25 は、AR オブジェクト遮蔽判定部 24 の判定結果に応じて、音声データ取得部 22 にて取得した音声データを解析・加工する。具体的には、AR オブジェクトにより音声データの音源位置が遮蔽されていると判定した場合は、音響特性データ保存部 21 で保存されている材質毎の音響特性の中から、AR オブジェクト構成要素の情報に基づき対応する音響特性を選択する。さらに、選択した音響特性を用いて、音声処理部 8 の音声演算部 83 により、前記音声データを解析・加工する。
- [0051] 音声データ出力部 26 は、音声データ加工部 25 にて加工した音声を、音

声処理部 7 の音声出力部 8 2（ヘッドフォン 8 0）から、ユーザ 1 0 に向けて出力する。

[0052] [ARオブジェクトの音響特性]

次に、音響特性データ保存部 2 1 に格納する AR オブジェクトの構成要素に対する音響特性について説明する。

[0053] 図 3 は、音響特性を表記法の例を示す図であり、構成要素が異なる 2 つの例（a）、（b）について示している。（a）は低域通過型の特性、（b）は帯域通過型の特性の一例であり、横軸に時間、または周波数をとって、その振幅の変化（応答特性）を表記している。AR オブジェクトの現物がある場合は、その現物の音響特性を、定常解析手法や、インパルス応答などの非定常解析手法等で測定して、音響特性データ保存部 2 1 に格納しておくこともできる。しかし、本実施例では、演算処理の簡略化を図るため、AR オブジェクトの構成要素の材質・厚さ・断面積等の付帯情報に応じたインパルス応答を、パターン化して格納している。インパルス応答とは、インパルス（時間的に非常に短い信号）を入力したときの出力振幅波形の時間変化である。

[0054] 図 4 は、AR オブジェクトの構成要素に対するインパルス応答を記述したテーブルである。このテーブルは、音響特性データ保存部 2 1 に格納されている。テーブル 4 5 0 の項目は、AR オブジェクトの構成要素を識別する番号 4 5 1 と、構成要素の材質 4 5 2、構成要素の厚さ 4 5 3、構成要素の断面積 4 5 4 のパラメータに対し、構成要素の音響特性（インパルス応答）4 5 5 をパターン化して記述している。材質 4 5 2 のパラメータは、金属、ガラス、プラスチック等で区別する。厚さ 4 5 3 は、厚い／薄いの 2 パターン、断面積 4 5 4 は、大／小の 2 パターンで区別する。各パラメータのインパルス応答 4 5 5 は、応答値 1 ～ 応答値 4 等でパターン化して与えている。

[0055] 図 5 は、インパルス応答の波形を模式的に示す図であり、図 4 の応答値 1 ～ 4 を例に示している。インパルス応答は、インパルスを入力したときの出力振幅波形の時間変化であり、構成要素の材質、厚さ、断面積といったパラ

メータの組み合わせで変化する。

[0056] 例えば応答値 1 は、材質が金属で厚さが厚く、断面積が大きい場合であり、応答波形は振幅の減衰度が大きくなる。これに対し応答値 2 は、材質が金属で厚さが厚く、断面積が小さい場合であり、応答波形は振幅の減衰度が小さくなる。また、応答値 3 や応答値 4 は、材質が金属で厚さが薄い場合であり、周波数特性の違いが応答波形の形状の違いとなって表れている。このように、構成要素の各種条件に対するインパルス応答をパターン化して表現している。金属以外の材質に対するインパルス応答値も、同様にパターン化することができる。

[0057] 本実施例では、AR オブジェクトの構成要素の材質、厚さ、断面積に応じて、テーブルを参照して対応するインパルス応答値を選択する。なお、AR オブジェクト情報保存部 16 にて、AR オブジェクトの構成要素に関する情報（材質・厚さ・断面積等）がない場合は、AR オブジェクトの画像に関する情報（形状・大きさ・表面テクスチャ）等から推定することも可能である。このように、インパルス応答値がパターン化されているので、演算処理の簡略化と高速化を図ることができる。

[0058] この構成要素の音響特性をパターン化したテーブル、および、インパルス応答値は、HMD 1 の記憶部 4 の各種データ部 42 に保存していてもよいが、AR オブジェクトの材質が分かった段階で、通信処理部 6 の LAN 通信部 61 を介して接続された外部のサーバから、必要なインパルス応答値のみをダウンロードして利用することもできる。

[0059] なお、本実施例で利用する音響特性としてインパルス応答値を採用したが、図 3 で示したように、周波数応答を規定するフィルタ特性（例えば、ローパスフィルタ）を採用することもできる。

[0060] また、本実施例で利用する音響特性は、図 1 B の状態を想定し、音声 AR オブジェクト 35 を通過するときの減衰特性に注目して説明した。さらには、図 1 D の状態に対応するために、音声 AR オブジェクト 35' により反射するときの反射特性も考慮することで、より現実味を帯びた音声処理が

可能となる。この場合、図4のテーブルには、ARオブジェクトと実空間の音源との距離等をパラメータとして追加すればよい。

[0061] [HMDによる重畳表示の例]

図6A～図6Cは、ARオブジェクト重畳表示時の映像・音声処理の例を示す図である。まず、映像処理から説明する。

[0062] 図6Aは、ユーザの前方に見える実空間の風景であり、HMD1の撮像部71（カメラ70）により撮影される。本例における実空間には、部屋の背景31と、部屋の隅に座った状態の人物32（人物32が発する音源である口321）が存在している。

[0063] 図6Bは、HMD1のARオブジェクト生成部17により生成したARオブジェクトである。本例では、ARオブジェクトとして自動車のARオブジェクト36の例を示す。ARオブジェクト情報保存部16には、様々なARオブジェクトの画像情報が保存されており、それから選択したものである。

[0064] 図6Cは、図6Aの実空間の撮影映像（部屋の背景31と人物32の映像）と、図6BのARオブジェクト36とを重畳した映像であり、HMD1の表示データ出力部19により、表示画面（表示部72）に表示される。HMD1を装着したユーザ10は、HMD1を介しこのような重畳映像を視認することができる。

[0065] 本実施例のビデオ透過型HMDでは、実空間の映像を撮像部71で撮影し、その映像データを処理してARオブジェクト36とともに表示部72に重畳表示する。その際、ARオブジェクト重畳部18は、3次元データ保存部14に保存されている実空間の状態（物体の配置や大きさ等）を考慮し、ARオブジェクトの大きさや配置などが、実空間の状態と矛盾しないように処理する。つまり、ARオブジェクト36で遮蔽されない実空間の映像（背景31の一部）を表示し、ARオブジェクト36で遮蔽される実空間の映像データ（背景31の一部と人物32）は表示しないように処理を行なう。

[0066] 図6Cの場合は、ARオブジェクト36（自動車）は背景31や人物32の手前に存在し、人物32はARオブジェクト36（自動車）の背後に隠れ

ている状態を表示している。なお、この例では、人物32はARオブジェクト36に隠れて見えないが、その位置を示すため、人物32の輪郭を破線32'で示している。

[0067] 次に、音声処理について説明する。図6Aの状態では、人物32が口321から音声を発すると、音声入力部81（マイク）で集音し、そのまま音声出力部82（ヘッドフォン80）から出力される。HMD1を装着したユーザ10は、その音声を聞いても何ら違和感がない。この場合の音源位置は、音源位置特定部23により、人物32の口321であることが特定される。

[0068] 一方図6Cの状態では、ARオブジェクト36（自動車）により隠れて見えない人物32'の音声を図6Aと同様の音量で出力すると、ユーザ10にとって対応する音源が分からず、違和感が生じる。そこで本実施例では、音源位置がARオブジェクト36により遮蔽される位置関係にある場合、ARオブジェクト36の構成要素の音響特性に基づき音声データを加工して出力する。

[0069] そのためARオブジェクト遮蔽判定部24では、ユーザ10から見て、ARオブジェクトにより音声データの音源位置が遮蔽されているかどうかを判定する。音源位置が遮蔽されていると判定した場合は、音声データ加工部25により、音声データを少し籠もった（こもった）音声に加工し、その加工した音声を音声出力部82（ヘッドフォン80）から出力する。これによりユーザ10は、人物32'の発する音声は物陰から聞こえてくる感覚になり、ARオブジェクト36（自動車）の背後にいる人物32'からの音声であることを直感的に把握することができる。

[0070] 音声データ加工部25では、音響特性データ保存部21に保存されているARオブジェクト（自動車）の構成要素（例えば、材質が金属、厚さが薄い、断面積が大きい）に対応する音響特性であるインパルス応答値3（図4、図5）を用いることで、現実感の増した音声データに加工することができる。

[0071] 図7Aと図7Bは、ARオブジェクト重畳表示時の映像・音声処理の他の

例を示す図である。ここでは、ARオブジェクト（自動車）に人物が乗車した場合を想定し、ウィンドウの開閉状態に応じて、人物からの音声データの処理を異ならせる場合を示す。

[0072] 図7Aは、ARオブジェクト37と実空間の人物33を重畳して表示した状態である。人物33は、ARオブジェクト37（自動車）に乗車しているように、ARオブジェクト37の配置および大きさを調整している。また、ARオブジェクト37のウィンドウ371は閉じている状態で、人物33はウィンドウ371越しに半透明で表示されている。

[0073] この状態で、人物33が発する音声は、ARオブジェクト37の音響特性に応じた加工を行い、HMD1を装着したユーザ10に提示される。人物33の口331とユーザ10とを結ぶ線分上には、ARオブジェクトの構成要素として材質がガラスであるウィンドウ371が存在するので、音響特性にはガラスのインパルス応答値を用いる。

[0074] 図7Bは、ARオブジェクト37の助手席のウィンドウ371が開いている状態を示している。人物33の口331とユーザ10とを結ぶ線分上には、ARオブジェクトの構成要素が存在しないので、図6Aの状態と同様に、人物33が発する音声には何ら加工せずユーザ10に出力する。このように、ユーザ10に提示する音声を、ARオブジェクトの変化（ウィンドウの開閉状態）に応じて異ならせることで、ユーザ10の聞く音声はより現実味を帯びたものとなる。

[0075] 本例では、ARオブジェクトの変化としてウィンドウの開閉について説明したが、ARオブジェクトの変化であればどのような形態にも適用可能である。例えば、ARオブジェクトとして表示した住宅におけるドアの開閉状態や、人物の口に装着するマスクの着脱状態などにも適用できることは言うまでもない。

[0076] また、上記の例では、ARオブジェクトが静止している場合としたが、ARオブジェクトが移動する場合についても上記の音声処理を適用できることは言うまでもない。その場合には、音源の前を通過するARオブジェクトの

位置によって、ユーザ 10 が聞く音声、時間的に変化するようにすればよい。

[0077] [重畳表示の映像・音声処理手順]

図 8 は、実空間の映像と A R オブジェクトとを重畳表示する映像処理手順を示すフローチャートである。ここでは、図 6 A の実空間の映像と図 6 B の A R オブジェクトを重畳して、図 6 C のように重畳して表示する場合について説明する。また、以下の処理は、図 2 B に示した各機能ブロックにより実行される。

[0078] 重畳表示の処理が開始 (S 4 1 1) されると、撮影データ取得部 1 5 により、実空間の撮影データを取得する (S 4 1 2)。

[0079] A R オブジェクト情報保存部 1 6 に保存されている A R オブジェクトの中から、表示すべき A R オブジェクト 3 6 の画像に関する情報 (形状・大きさ・配置等の情報) を入手する (S 4 1 3)。

[0080] A R オブジェクト生成部 1 7 は、入手した A R オブジェクト 3 6 の画像に関する情報を解析し、A R オブジェクト 3 6 を生成する (S 4 1 4)。A R オブジェクトの生成にあたっては、3 次元データ保存部 1 4 に保存されている実空間の情報 (予め取得した実空間の 3 次元データ) を入手し、その実空間の形状を考慮して、表示すべき A R オブジェクト 3 6 の配置や大きさなどを決定する。

[0081] A R オブジェクト重畳部 1 8 は、撮影データ取得処理 (S 4 1 2) で取得した撮影データ (実空間の映像) と、A R オブジェクトの生成処理 (S 4 1 4) で生成した A R オブジェクト 3 6 とを重畳する (S 4 1 5)。その際、ユーザ 1 0 (HMD 1) から見る実空間の映像と A R オブジェクトの奥行方向の位置関係に応じて、陰に隠れる映像 (画像) 部分は表示させないオクルージョン処理を行う。

[0082] 表示データ出力部 1 9 は、A R オブジェクトの重畳処理 (S 4 1 5) がなされた表示データを、HMD 1 の表示画面に出力する (S 4 1 6)。

[0083] 以上で、実空間の映像 (図 6 A の背景 3 1 と人物 3 2) と、A R オブジェ

クト36（図6B）とを重畳して図6Cのように表示する映像処理を終了する（S417）。

[0084] 図9は、重畳表示の際の音声データの処理手順を示すフローチャートである。図6Aの実空間において人物32が発した音声を、図6CのARオブジェクト36が重畳された状態でユーザ10が聞こえる音声データに変換して出力する処理である。

[0085] 音声データの処理が開始（S431）されると、ARオブジェクト情報保存部16に保存されているARオブジェクト36の構成要素に関する付帯情報（構成要素の材質・厚さ・断面積等）を入手する（S432）。

[0086] 音声データ取得部22は、音声入力部81に音声入力があるかどうかを判定する（S433）。判定の結果、音声入力があった場合はS434へ進み、音声データ取得部22は入力した音声データを取得する。音声入力がない場合は音声入力があるまで待機する。

[0087] 音源位置特定部23は、音源位置（人物32の口321の位置）を特定する（S435）。音源位置である人物32の口321の位置は、3次元データ保存部14に保存されている実空間の配置情報と、撮影データ取得部15が取得した実空間の映像とを解析することにより、特定することができる。

[0088] ARオブジェクト遮蔽判定部24は、ARオブジェクトにより音源位置が遮蔽されているかどうかを判定する。すなわち、音源（人物32の口321）とHMD1を装着したユーザ10とを結ぶ線分上に、ARオブジェクト36が存在しているかどうかを判定する（S436）。

[0089] 判定の結果、ARオブジェクトにより音源位置が遮蔽されている場合は、S437へ進み、ARオブジェクトにより音源位置が遮蔽されていない場合は、S439へ進む。

[0090] S437では、音源を遮蔽しているARオブジェクト36の構成要素を特定する。そして、音響特性データ保存部21に保存されている音響特性データの中から、ARオブジェクト36の構成要素に対応する音響特性データを選択する。図6Cの場合には、ARオブジェクト36の音響特性データは、

材質が金属、厚さが薄い、断面積が大きい、に対応するインパルス応答値 3 を採用している。

[0091] 音声データ加工部 25 は、S 437 で選択した音響特性（応答値 3）に応じて音声データの加工処理を行う（S 438）。

[0092] S 439 では、音声データ出力部 26 は、音声出力部 81（ヘッドフォン 80）から音声データを出力する。

[0093] 以上で、図 6C のような重畳表示において、実空間における音声を HMD 1 を装着したユーザ 10 が聞こえる音声データに変換する処理を終了する（S 440）。

[0094] なお、上記の AR オブジェクト遮蔽判定処理（S 436）では、図 1B の状態を想定し、AR オブジェクト 35 により音源位置 301 が遮蔽されているかどうかを判定した。さらに、図 1D の状態に対応するためには、音源位置 301 の前方にある AR オブジェクト 35' により音声反射されているかどうかを判定し、AR オブジェクト 35' の音響特性（反射特性）に応じて音声データの加工処理（S 438）を行うようにすればよい。

[0095] また上記の説明では、音源として実空間で発せられた音声（人物 32, 33 の発する音声）のみを取り上げたが、さらに、AR オブジェクト（自動車）36, 37 が発する音声（例えばエンジン音やクラクション音）も含めて処理することができる。その場合は、音声データ出力処理（S 439）において、実空間の音声データと、AR オブジェクトから発生する音声データとを重畳して音声出力を行う。

[0096] また音声データ加工処理（S 438）では、音源である AR オブジェクトが実空間の物体により遮蔽／反射される場合、あるいは音源である AR オブジェクトが他の AR オブジェクトにより遮蔽／反射される場合には、上記説明と同様に、遮蔽／反射する物体の音響特性データに基づいて AR オブジェクトの音声データの加工処理を行うことは言うまでもない。

[0097] 上記実施例では、重畳表示における映像処理と音声処理は、ハードウェア要素である映像演算部 73 と音声演算部 83で行っているが、これらはソフ

トウェア処理によっても実現できることは言うまでもない。

[0098] 実施例 1 によれば、AR オブジェクトの重畳表示と、AR オブジェクトの配置を考慮した音声出力を実現し、ユーザは違和感なく映像と音声を視聴することができる。また、音声演算部による音声データ加工処理ではパターン化された音響特性データを用いるので、処理時間が短縮しリアルタイムに加工することができる。

実施例 2

[0099] 前記実施例 1 では、AR オブジェクトを重畳表示するとき、実空間における配置に矛盾が生じないように AR オブジェクトの配置や大きさを決定した。これに対し実施例 2 では、AR オブジェクトの配置や大きさを先に決定し、これに合わせて実空間の映像を編集して重畳させるようにしている。

[0100] 実施例 2 における HMD の基本構成は実施例 1 で述べた構成と同様であるが、機能ブロックについては、実空間の映像を編集するための機能を追加している。

[0101] 図 10 は、HMD の機能ブロック構成を示す図である。図 2B の機能ブロック構成に対し、撮影データ編集部 27 と音源位置移動部 28 とを追加している。

[0102] 撮影データ編集部 27 では、撮影データを編集し、撮影対象物の大きさや配置を編集する。すなわち、AR オブジェクト生成部 17 により生成した AR オブジェクトに合わせて撮影対象物を編集し、AR オブジェクト重畳部 18 により両者を重畳して表示する。

[0103] 音源位置移動部 28 では、撮影データ編集部 27 により編集された映像において、音源の移動が伴う場合、移動した音源位置に合致するように、音声データを加工する。音源位置は、左右ヘッドフォンの音量差や、左右ヘッドフォンの到達時間差等を調整することにより、任意に設定することができる。

[0104] 図 11A と図 11B は、AR オブジェクト重畳表示の例を示す図である。実空間の撮影データである人物 34 の映像を編集して AR オブジェクト 37

（自動車）と重畳し、人物３４がＡＲオブジェクト３７（自動車）に乗車している状態を表示する場合である。

[0105] 図１１Ａは、人物３４とＡＲオブジェクト３７とを、いずれも何ら映像加工を施さずに重畳した場合を示す。この場合、破線で示す人物３４は、ＡＲオブジェクト３７（自動車）のフロント部（エンジン部）に存在し、現実にはあり得ない状態となっている。前記実施例１（図７）では、ＡＲオブジェクト３７（自動車）の配置を変えて人物３４が乗車している状態を実現したが、本実施例では、人物３４の配置を変えて自動車に乗車している状態を実現する。

[0106] 図１１Ｂは、人物３４の映像を編集してＡＲオブジェクト３７と重畳した場合を示す。この例では、人物３４の位置を矢印３９の方向に移動させ、ＡＲオブジェクト３７のウィンドウ３７１の位置に配置させている。配置後の人物３４’はウィンドウ３７１越しに半透明に見えるよう、また人物３４’の大きさもＡＲオブジェクト３７の大きさに合わせて調整している。このように図１１Ｂでは、撮影データ編集部２７により実空間に存在する人物３４の撮影データ（大きさ及び配置）を編集し、ＡＲオブジェクト３７と重畳しても何ら不自然さがないような映像を実現している。

[0107] 図１１Ｂの映像処理では、人物３４の位置を矢印３９の方向に移動させて人物３４’として表示している。その際、人物３４から発する音声の音源位置は、移動前の口３４１から移動後の口３４１’に変化する。これに合わせて音源位置移動部２８は、人物３４の音声、移動した音源位置３４１’から発せられたように音声データを加工する。これにより、人物３４’の音声についても、不自然さを感じさせないように提供できる。

[0108] 図１２は、本実施例における重畳表示時の映像処理手順を示すフローチャートである。これは、実施例１（図８）のフローチャートをベースとし、実空間の映像を編集するためのステップＳ４２１～Ｓ４２２を追加している。ここでは、図８のフローチャートとの相違点のみ説明する。

[0109] Ｓ４１５にて撮影データにＡＲオブジェクトを重畳する処理を行った後、

新たに追加したS 4 2 1では、重畳表示において、撮影データとARオブジェクトとの位置関係に、何か不自然さがあるかどうかを判定する。例えば図1 1 Aの重畳表示では、実空間の人物3 4が、現実には存在できないARオブジェクト3 7（自動車）のフロント部（エンジン部）に存在し、不自然な状態となっている。

[0110] 判定の結果、不自然な状態である場合は、S 4 2 2へ進み撮影データの編集を行う。不自然でない場合は、S 4 1 6へ進み、そのまま表示データを出力する。

[0111] S 4 2 2では、撮影データ編集部2 7により撮影データを編集する。例えば図1 1 Bの場合は、人物3 4の撮影データを編集し、人物3 4がARオブジェクト3 7のウィンドウ3 7 1の位置に配置させ、また人物3 4の大きさも、ARオブジェクト3 7の大きさに合わせて調整する。その後、S 4 1 6にて表示データとして出力する。

[0112] 図1 3は、重畳表示の際の音声データの処理手順を示すフローチャートである。これは、実施例1（図9）のフローチャートをベースとし、音源位置の移動に伴う音声処理のステップS 4 4 1～S 4 4 3を追加している。ここでは、図9のフローチャートとの相違点のみ説明する。

[0113] S 4 3 5にて、音源位置特定部2 3により音源位置（例えば図1 1 Aにおける人物3 4の口3 4 1）を特定した後、新たに追加したS 4 4 1にて、撮影データ編集部2 7により撮影データを編集したかどうかを判定する。つまり、図1 2の映像処理フローチャートにおいて、S 4 2 2の撮影データ編集処理を実行したかどうかを判定する。

[0114] 判定の結果、撮影データを編集した場合は、追加したS 4 4 2へ進み、撮影データを編集しなかった場合は、S 4 3 6へ進む。

[0115] S 4 4 2では、撮影データの編集に伴い、音源位置の移動が必要かどうかを判断する。例えば図1 1 Aから図1 1 Bへの撮影データの編集において、音源位置（人物3 4の口3 4 1）の移動が必要かどうかを判定する。音源位置の移動が必要である場合は、S 4 4 3に進み、音源位置の移動が必要でな

い場合は、S 4 3 6 に進む。

[0116] S 4 4 3 では、音源位置移動部 2 8 により、移動した音源位置に合致するように音声データを加工する。

[0117] その後 S 4 3 6 にて、A R オブジェクト遮蔽判定部 2 4 は、A R オブジェクトにより音源位置が遮蔽されているかどうかを判定する。もちろんこのときの音源位置は、S 4 4 3 で音源位置が移動されていれば移動後の音源位置となる。

[0118] さらに S 4 3 7 以降では、遮蔽する A R オブジェクトの音響特性に応じて音声データを加工し、出力する。

[0119] その結果、図 1 1 A の人物 3 4 の口 3 4 1 から発せられた音声は、図 1 1 B の口 3 4 1' から発せられたように、かつウィンドウ 3 7 1 の遮蔽の影響を加味して出力される。

[0120] 実施例 2 によれば、実空間の映像と、A R オブジェクトの配置および大きさの関係が不自然であった場合でも、実空間の映像を編集することや、実空間の音声を加工することにより、その不自然さを解消して提示することができる。

[0121] さらに実施例 2 では、実空間の映像を編集することで、現実には実現不可能な表示も可能となる。例えば A R オブジェクトとして空中に浮かばせた自動車に、実空間の人物が乗車している仮想の状態をリアリティに表示できる。実空間では人物はそのような姿勢をとることは困難であるが、実空間の映像を編集することで容易に実現できる。

実施例 3

[0122] 実施例 3 では、HMD がカメラを 2 個搭載した場合について説明する。

図 1 4 は、本実施例の HMD 1 をユーザ 1 0 が装着した状態を示す図である。HMD 1 の撮像部 7 1 として左右端に 2 個のカメラ 7 1 1, 7 1 2 を搭載しており、2 個のカメラで撮影した実空間の映像をステレオ映像として表示することで、ユーザ 1 0 は実空間の奥行き方向を直観的に認識できるようになる。なお、HMD 1 の内部構成や映像・音声処理は実施例 1 の場合と同

様であり、重複する説明を省略する。

[0123] また、図 1 4 では図示していないが、音声入力部 8 1（マイク）についてもカメラと同様に、HMD 1 の左右端に 2 個のマイクを搭載することにより、実空間の音声を音場空間が広がるステレオ音声として取得することができる。そして、HMD 1 の音声出力部 8 2（左右のヘッドフォン 8 2 1, 8 2 2）からステレオ音声を出力することで、ユーザ 1 0 に対し奥行き感のある音声を提供することができる。

[0124] 実施例 3 によれば、HMD 1 を装着したユーザ 1 0 は、実空間の映像の奥行き方向を直観的に認識することや、実空間の音声を音場空間が広がるステレオ音声として聞くことが可能となる。

実施例 4

[0125] 前記実施例では音源が 1 個の場合を想定していたが、実施例 4 では、音源が複数個存在する場合について説明する。実施例 4 における HMD の基本構成は実施例 1 で述べた構成と同様であるが、機能ブロックについては、複数の音源を分離するための機能を追加している。

[0126] 図 1 5 は、HMD の機能ブロック構成を示す図である。図 2 B の機能ブロック構成に対し、音源分離部 2 9 を追加している。

[0127] 音源分離部 2 9 は、音声データ取得部 2 2 で取得した音声データを解析し、音源毎の個々の音源に分離する処理を行う。音源を分離する一般的な手法として、外部に複数のマイクを設置し、複数マイクの個々の位置における音声データを位相差や音圧差などを検出し、音源の位置を正確に求める方法がある。これに対し本実施例では、実施例 3 で述べたように、HMD 1 の左右端に 2 個のマイクを搭載し、1 個の音源が発する音が左右のマイクに到達する位相差や音量差により音源の方向を検出して、簡易的に音源を分離する。また、3 次元データ保存部 1 4 に保存されている実空間の配置情報と、撮影データ取得部 1 5 で取得した実空間の映像から、音源の位置を特定し分離することもできる。

[0128] 図 1 6 は、音声データの処理手順を示すフローチャートである。これは、

実施例 1（図 9）のフローチャートをベースとし、音源分離に伴うステップ S 4 4 5、S 4 4 6 を追加している。ここでは、図 9 のフローチャートとの相違点のみ説明する。

[0129] S 4 3 5 にて、音源位置特定部 2 3 により音源位置を特定した後、新たに追加した S 4 4 5 にて、音源分離部 2 9 により、複数の音源を個々に分離する。そして次の S 4 3 6 以降で、分離した個々の音源に対して、遮蔽される A R オブジェクトの構成要素の音響特性に基づいて音声データの加工処理を行う。

[0130] なお、複数の音源が存在する場合、音量（音圧）や周波数などで閾値を設け、閾値以下の小さな音量の音源および閾値以上の高周波の音源については、S 4 3 6 ～ S 4 3 8 の音声データ加工処理を行わないようにして、音声加工の対象とする音源の数を絞るようにしてもよい。S 4 4 6 では、分離した全ての音源について処理が終了したかどうかを判定し、終了していれば S 4 3 9 にて音声データを出力する。

[0131] 本実施例においても、音声データ出力処理（S 4 3 9）において、A R オブジェクトからの音声（例えば自動車のエンジン音やクラクション音）を重畳して音声出力を行えることは言うまでもない。

[0132] 実施例 4 によれば、音源が複数存在した場合においても、個々の音源について A R オブジェクトの配置を考慮して音声を出力することができる。

実施例 5

[0133] 実施例 5 では、外部のサーバに接続して処理に必要な情報を入手する H M D システムについて説明する。

図 1 7 は、H M D システムの全体構成を示す図である。ユーザ 1 0 が装着している H M D 1 は、通信部 6 の L A N 通信部 6 1 を介して、無線ルータ 6 5 およびネットワーク網 6 6 に接続されている。ネットワーク網 6 6 には、複数のサーバ 6 7 ～ 6 9 が接続されている。

[0134] 第 1 のサーバ 6 7 には、A R オブジェクトの画像に関する情報（形状・大きさ・配置等の情報）や、A R オブジェクトの構成要素に関する付帯情報（

各構成要素の材質・厚さ・断面積等)、ARオブジェクトに関する情報が保存・管理されている。この第1のサーバ67を利用することにより、ARオブジェクト情報保存部16で必要とする記憶部4の各種データ部42の負荷を軽減することができる。

[0135] 第2のサーバ68には、材料毎に材質、厚さ、断面積をパラメータに音響特性データがパターン化されて、保存・管理されている。第2のサーバ68で保存・管理されている材料は、ARオブジェクトの構成要素の材料のみならず、ARオブジェクトの構成要素以外の材料に関する音響特性データも含まれている。この第2サーバを利用することにより、音響特性データ保存部21で必要とする記憶部4の各種データ部42の負荷を軽減することができる。

[0136] 第3のサーバ69は、映像処理部7の映像演算部73や、音声処理部8の音声演算部83等で行っている演算処理を高速に行うサーバである。この第3のサーバを利用することにより、ARオブジェクト生成部17、ARオブジェクト重畳部18、音源位置特定部23、音声データ加工部25の負荷を軽減することができる。

[0137] 本実施例では、HMD1の負荷を軽減するため、全てのサーバ67～69を利用しているが、必要に応じて、適宜サーバを選択して利用できることは言うまでもない。なお、これらのサーバは、個々に独立することなく、集約することも可能である。

[0138] 以上、本発明の実施形態について実施例1～5を用いて説明したが、本発明の技術を実現する構成は前記実施例に限られるものではなく、様々な変形例に対しても適用できる。例えば、ある実施例の構成の一部を他の実施例の構成と置き換えることが可能であり、また、ある実施例の構成に他の実施例の構成を加えることも可能である。すなわち、HMDにより実空間の映像にARオブジェクトを重畳して表示するとき、実空間の物体とARオブジェクトの少なくとも一方が音源となる場合に適用できる。そして、音源からの音声データを、周囲の物体または周囲のARオブジェクトの音響特性(位置関

係により減衰特性または反射特性を用いる)を加味して加工処理して出力するものである。

[0139] 各実施例の説明においては、ビデオ透過型のHMDを前提としたが、ユーザが使用する端末はHMDの形態でなくともよい。例えば、スマートフォンのカメラとディスプレイを利用してARオブジェクトを表示し、ユーザが装着したヘッドフォンに本実施例に示した音響処理を施した音声を出力するようにしてもよい。

[0140] 各実施例で説明した機能等は、それらの一部または全部を、例えば集積回路で設計する等によりハードウェアで実現してもよい。また、マイクロプロセッサユニット等がそれぞれの機能等を実現するプログラムを解釈して実行することによりソフトウェアで実現してもよい。ハードウェアとソフトウェアを併用してもよい。その際のソフトウェアは、製品出荷の時点で、予めHMDのプログラム部41等に格納された状態であってもよいし、製品出荷後に、インターネット上の各種サーバ等から取得するものであってもよい。また、メモリカードや光ディスク等で提供されるソフトウェアを取得するものであってもよい。

[0141] また、図中に示した制御線や情報線は説明上必要と考えられるものを示しており、必ずしも製品上の全ての制御線や情報線を示しているとは限らない。実際には殆ど全ての構成が相互に接続されていると考えてもよい。

符号の説明

[0142] 1…ヘッドマウントディスプレイ(HMD)、2…主制御部、4…記憶部、5…センサ部、6…通信処理部、7…映像処理部、8…音声処理部、9…操作入力部、10…ユーザ、12…3次元センサ情報取得部、13…3次元データ処理部、14…3次元データ保存部、15…撮影データ取得部、16…ARオブジェクト情報保存部、17…ARオブジェクト生成部、18…ARオブジェクト重畳部、19…表示データ出力部、21…音響特性データ保存部、22…音声データ取得部、23…音源位置特定部、24…ARオブジェクト遮蔽判定部、25…音声データ加工部、26…音声データ出力部、2

7…撮影データ編集部、28…音源位置移動部、29…音源分離部、32～34…人物、35～37…ARオブジェクト、42…各種データ部、53…3次元センサ部、67～69…サーバ、70…カメラ、71…撮像部、72…表示部、73…映像演算部、80…ヘッドフォン、81…音声入力部、82…音声出力部、83…音声演算部。

請求の範囲

- [請求項1] ユーザが頭部に装着し拡張現実のオブジェクト（以下、ARオブジェクト）を表示するヘッドマウントディスプレイにおいて、
実空間の映像を撮影する撮像部と、
ARオブジェクトを生成するARオブジェクト生成部と、
前記撮影した実空間の映像と、前記生成したARオブジェクトとを重畳するARオブジェクト重畳部と、
前記重畳された実空間の映像とARオブジェクトとを表示する表示部と、
実空間にて発せられた音声を入力する音声入力部と、
前記ARオブジェクト重畳部により配置された実空間に存在する物体とこれに重畳して表示されるARオブジェクトの位置関係に応じて、実空間からの音声とARオブジェクトが発する音声の少なくとも一方の音声データを、周囲の物体または周囲のARオブジェクトの音響特性を加味して加工処理する音声データ加工部と、
前記音声データ加工部にて加工処理された実空間からの音声またはARオブジェクトの音声を入力する音声出力部と、
を備えることを特徴とするヘッドマウントディスプレイ。
- [請求項2] 請求項1に記載のヘッドマウントディスプレイにおいて、
当該ヘッドマウントディスプレイから見て、前記ARオブジェクト重畳部により配置されたARオブジェクトにより実空間の音源位置が遮蔽されているかどうかを判定するARオブジェクト遮蔽判定部を備え、
前記ARオブジェクト遮蔽判定部の判定により実空間の音源位置が遮蔽されていると判定された場合、前記音声データ加工部は、実空間の音声データを、遮蔽しているARオブジェクトの音響特性を加味して加工処理することを特徴とするヘッドマウントディスプレイ。
- [請求項3] 請求項2に記載のヘッドマウントディスプレイにおいて、

A R オブジェクトの構成要素毎に、その形状をパラメータとして、音響特性にかかわる音響特性データをパターン化して保存する音響特性データ保存部を備え、

前記音声データ加工部は、前記音響特性データ保存部から該当するA R オブジェクトの音響特性データを選択して、実空間の音声データの加工処理を行うことを特徴とするヘッドマウントディスプレイ。

[請求項4]

請求項1に記載のヘッドマウントディスプレイにおいて、

前記生成したA R オブジェクトに合わせて前記撮影した実空間の映像の大きさや配置を編集する撮影データ編集部と、

前記撮影データ編集部により編集された実空間の映像において音源位置の移動が伴う場合、移動した音源位置に合致するように実空間の音声データを加工する音源位置移動部と、

を備えることを特徴とするヘッドマウントディスプレイ。

[請求項5]

請求項1に記載のヘッドマウントディスプレイにおいて、

前記撮像部として複数のカメラを有し、

前記表示部には、前記複数のカメラで撮影した実空間の映像をステレオ映像として表示するとともに、

前記音声入力部として複数のマイクを有し、

前記音声出力部から、前記複数のマイクに入力した実空間の音声をステレオ音声で出力する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]

請求項1に記載のヘッドマウントディスプレイにおいて、

前記音声入力部として複数のマイクを有し、

前記実空間にて発せられた複数の音声に対して、前記複数のマイクを用いて、それぞれの音源位置を特定する音源分離部を備え、

前記音声データ加工部は、複数の音声に対し、前記特定したそれぞれの音源位置に応じて音声データの加工処理を行うことを特徴とするヘッドマウントディスプレイ。

[請求項7]

請求項1に記載のヘッドマウントディスプレイにおいて、

通信部を介して少なくとも1つの外部サーバに接続し、

前記外部サーバから、前記ARオブジェクト生成部にて生成するARオブジェクトに関する情報を入手し、

または前記外部サーバから、前記音声データ加工部にて音声データを加工処理するために使用するARオブジェクトの音響特性データを入手し、

または前記外部サーバを用いて、前記ARオブジェクト生成部、前記ARオブジェクト重畳部、前記音声データ加工部の少なくとも1つの処理を行うことを特徴とするヘッドマウントディスプレイ。

[請求項8] ユーザが頭部に装着したヘッドマウントディスプレイで拡張現実のオブジェクト（以下、ARオブジェクト）を表示する際の音声処理方法において、

撮影した実空間の映像に、生成したARオブジェクトを重畳して表示する映像処理ステップに合わせて、

実空間にて発せられた音声を入力する音声入力ステップと、

前記映像処理ステップにより配置された実空間に存在する物体とこれに重畳して表示されるARオブジェクトの位置関係に応じて、実空間からの音声とARオブジェクトが発する音声の少なくとも一方の音声データを、周囲の物体または周囲のARオブジェクトの音響特性を加味して加工処理する音声データ加工ステップと、

前記加工処理された実空間からの音声またはARオブジェクトの音声を出力する音声出力ステップと、

を備えることを特徴とする音声処理方法。

[請求項9] 請求項8に記載の音声処理方法において、
前記ヘッドマウントディスプレイから見て、前記映像処理ステップにより配置されたARオブジェクトにより実空間の音源位置が遮蔽されているかどうかを判定するARオブジェクト遮蔽判定ステップを備

え、

前記A Rオブジェクト遮蔽判定ステップの判定により実空間の音源位置が遮蔽されていると判定された場合、前記音声データ加工ステップでは、実空間の音声データを、遮蔽しているA Rオブジェクトの音響特性を加味して加工処理することを特徴とする音声処理方法。

[請求項10]

請求項9に記載の音声処理方法において、

予め、A Rオブジェクトの構成要素毎に、その形状をパラメータとして、音響特性にかかわる音響特性データをパターン化して保存しておく、

前記音声データ加工ステップでは、前記保存している音響特性データから、該当するA Rオブジェクトの音響特性データを選択して、実空間の音声データの加工処理を行うことを特徴とする音声処理方法。

[請求項11]

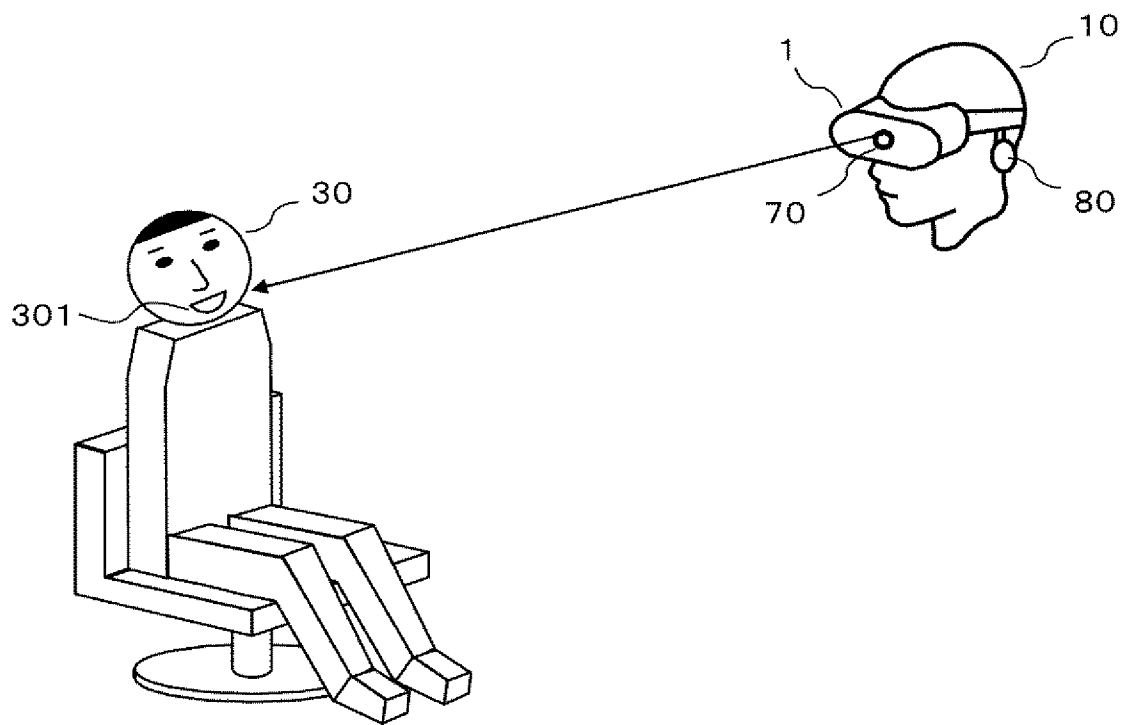
請求項8に記載の音声処理方法において、

前記映像処理ステップにて、前記生成したA Rオブジェクトに合わせて前記撮影した実空間の映像の大きさや配置を編集し、編集された実空間の映像において音源位置の移動が伴う場合、移動した音源位置に合致するように実空間の音声データを加工する音源位置移動ステップ、

を備えることを特徴とする音声処理方法。

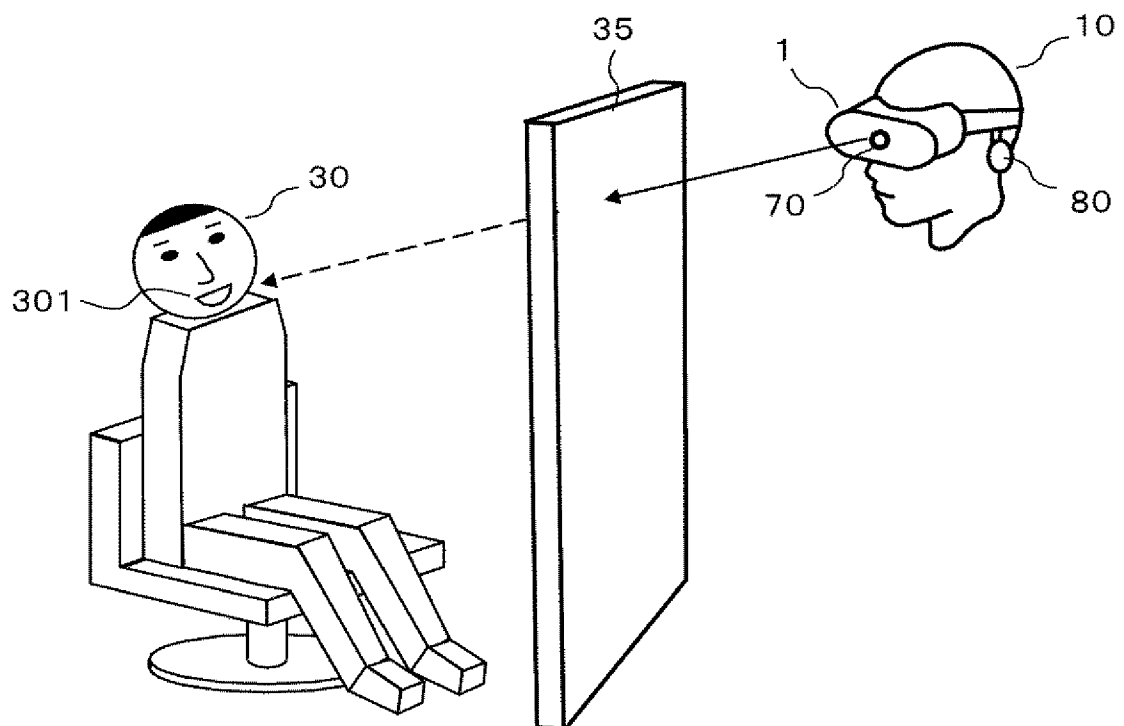
[図1A]

図 1 A



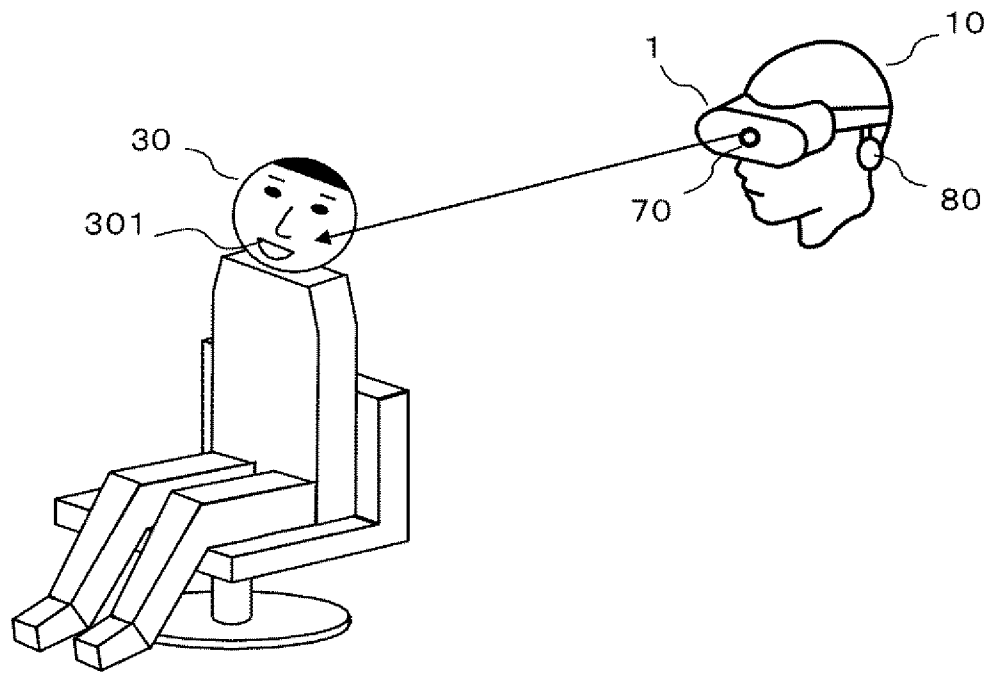
[図1B]

図 1 B



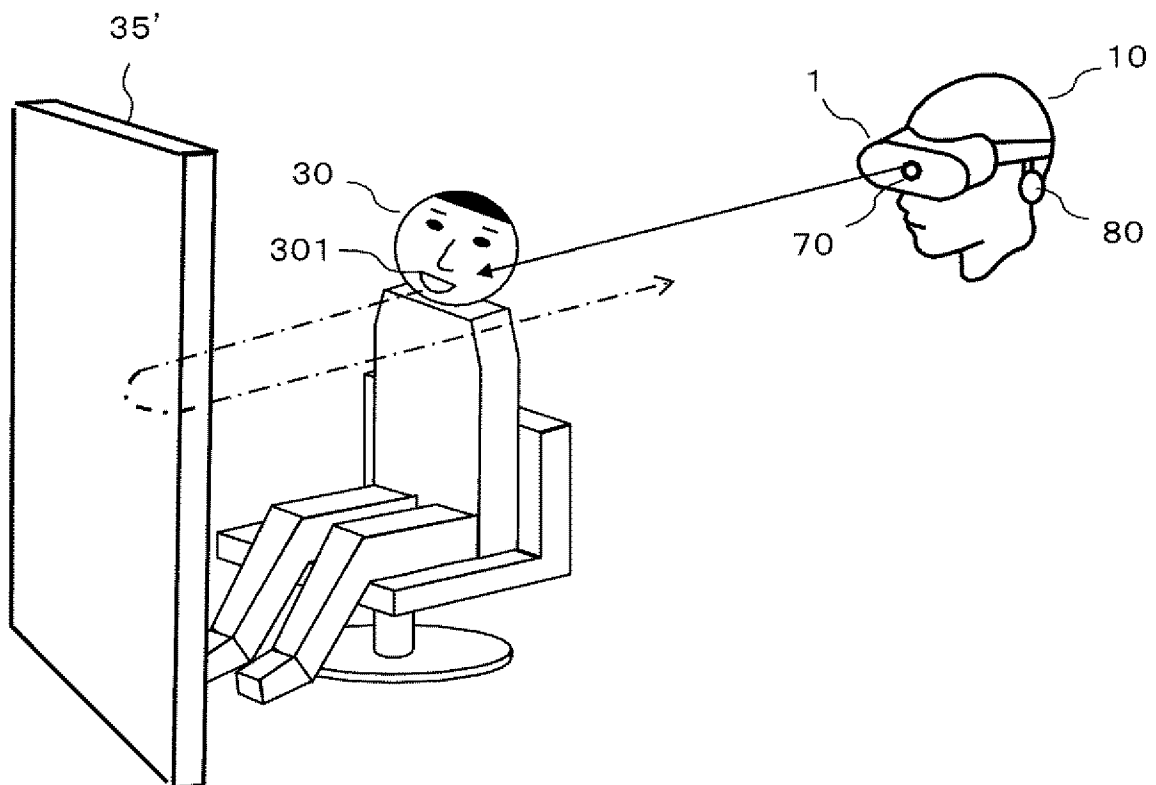
[図1C]

図 1 C



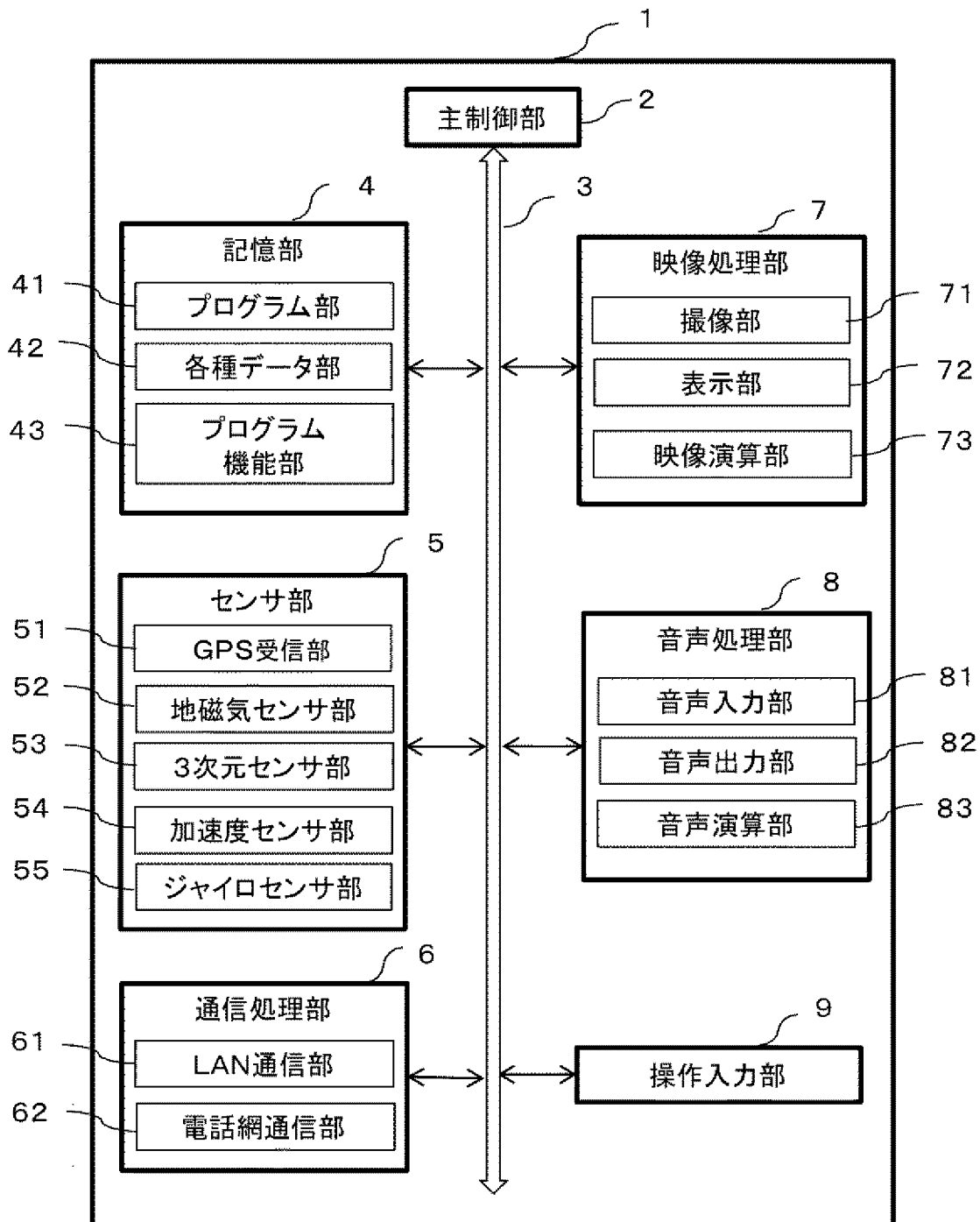
[図1D]

図 1 D



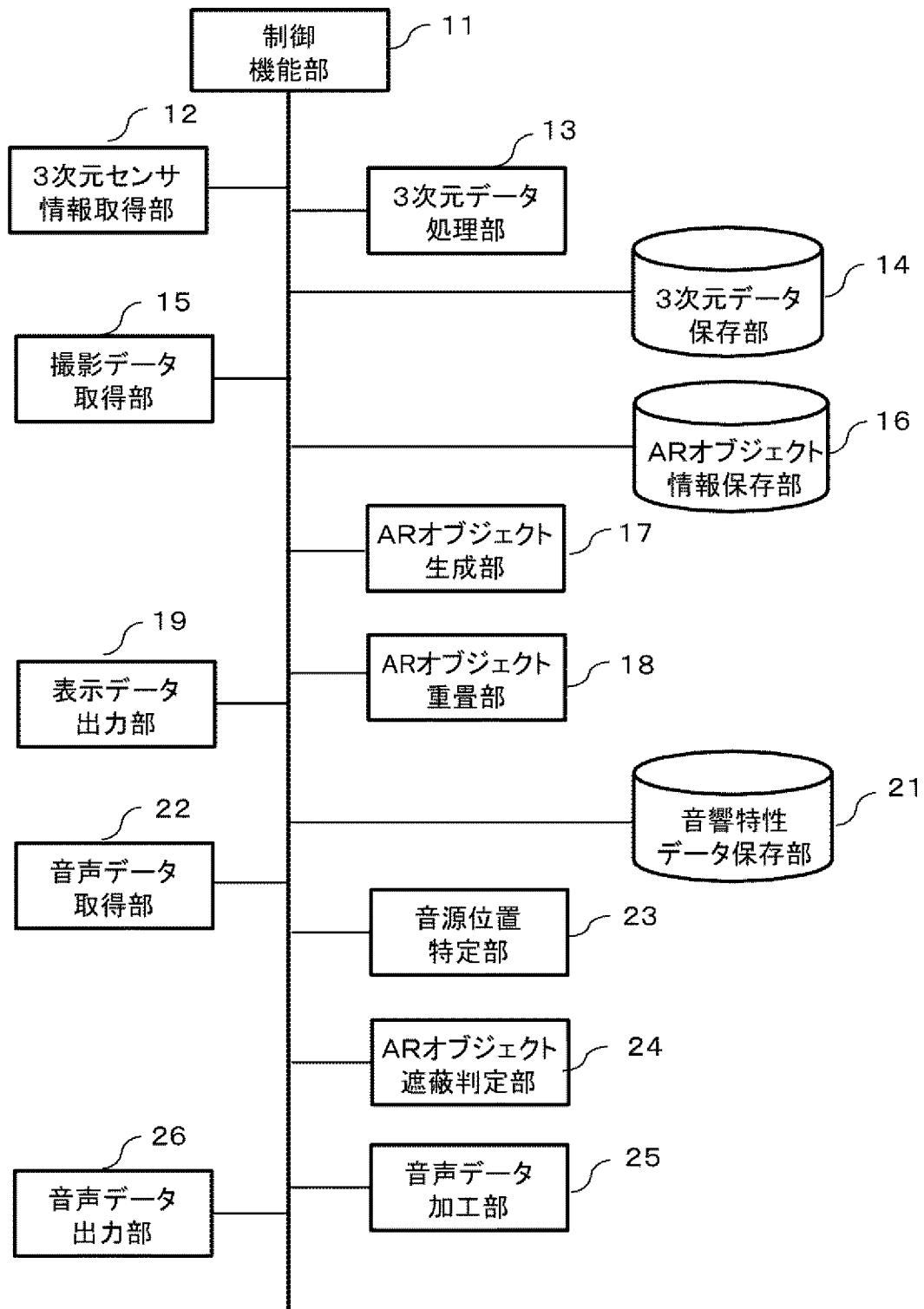
[図2A]

図 2 A



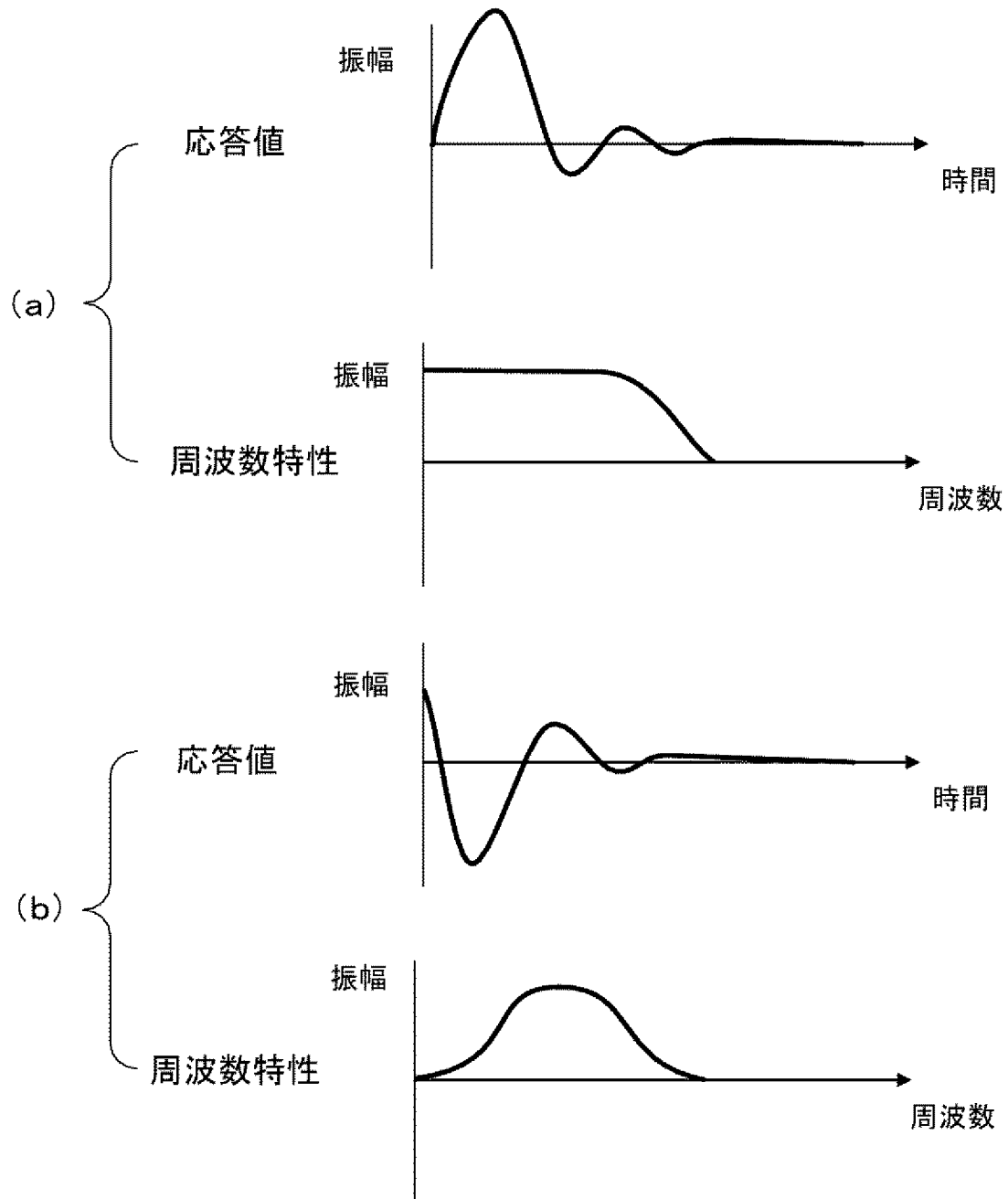
[図2B]

図 2 B



[図3]

図 3



[図4]

図 4

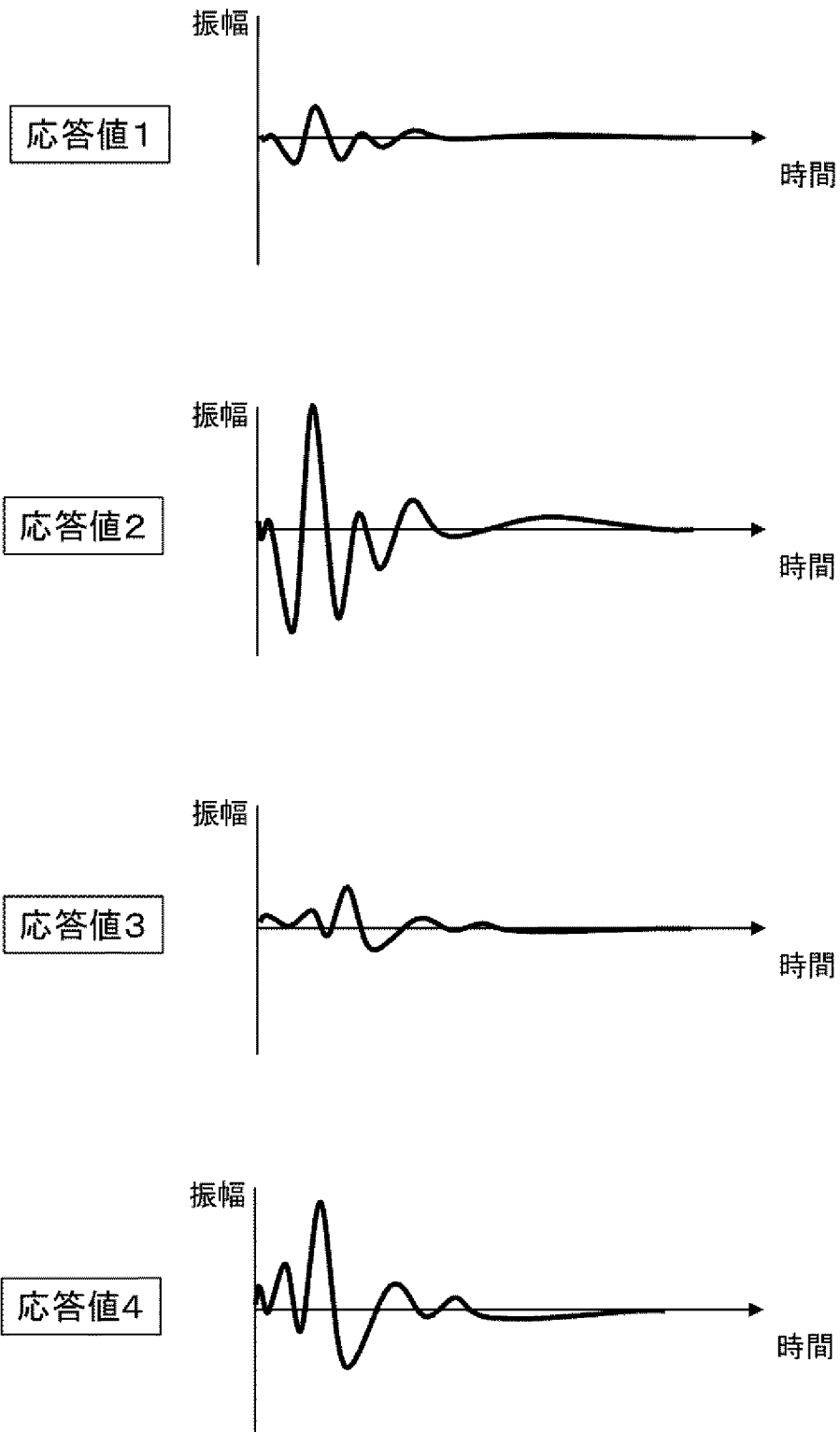
450

451 452 453 454 455

No.	材質	厚さ	断面積	インパルス 応答
1	金属	厚い	大	応答値1
			小	応答値2
		薄い	大	応答値3
			小	応答値4
2	ガラス	⋮	⋮	⋮
3	プラスチック	⋮	⋮	⋮
4	木材	⋮	⋮	⋮
5	段ボール	⋮	⋮	⋮
6	コンクリート	⋮	⋮	⋮
7	人体	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮

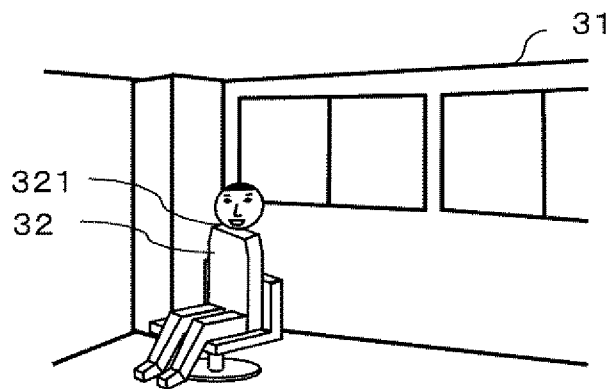
[図5]

図 5



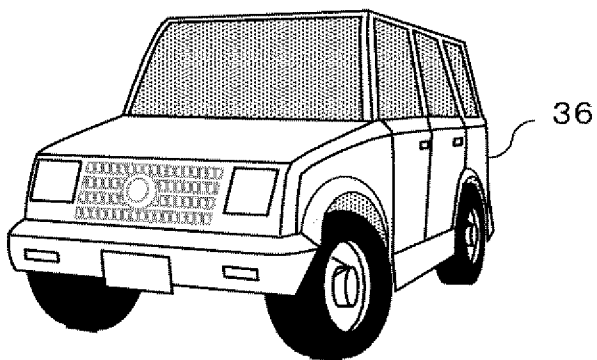
[図6A]

図 6 A



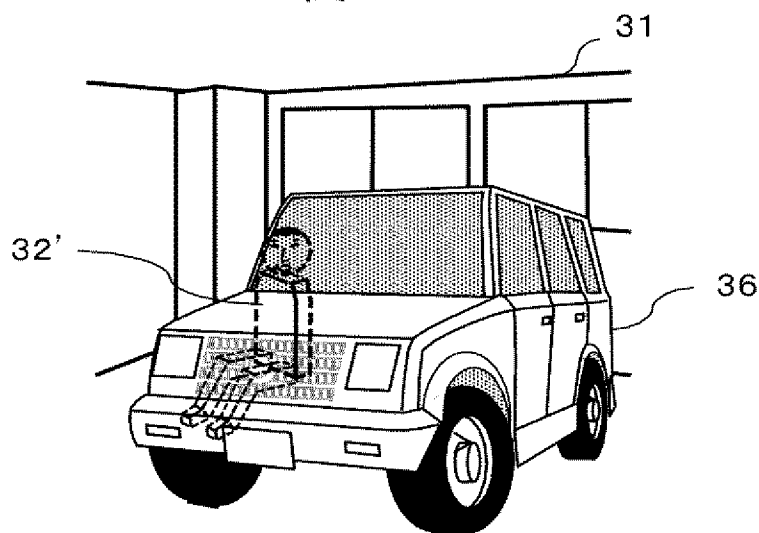
[図6B]

図 6 B



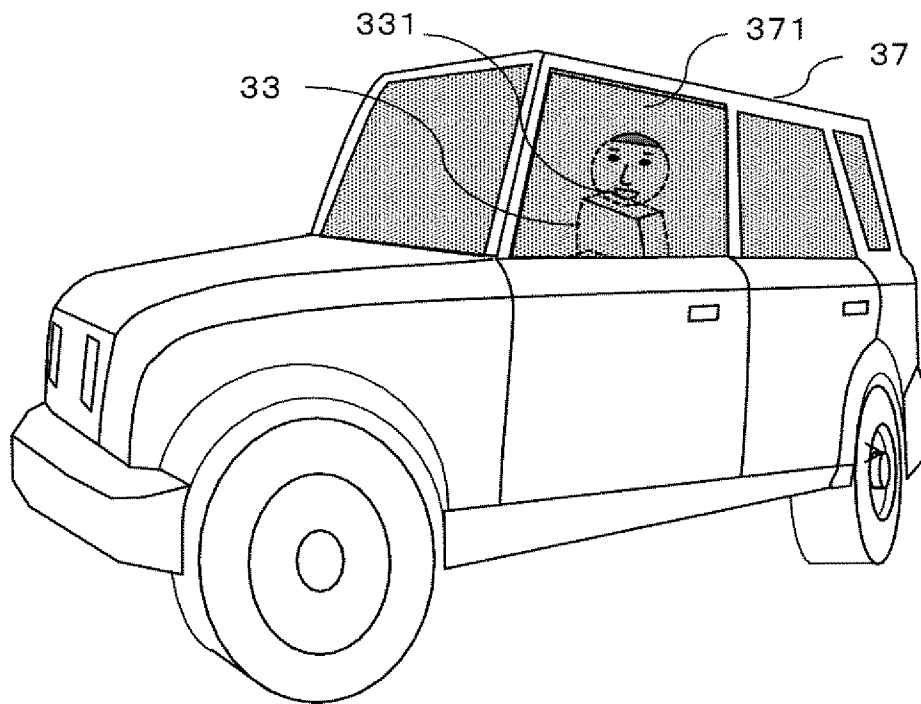
[図6C]

図 6 C



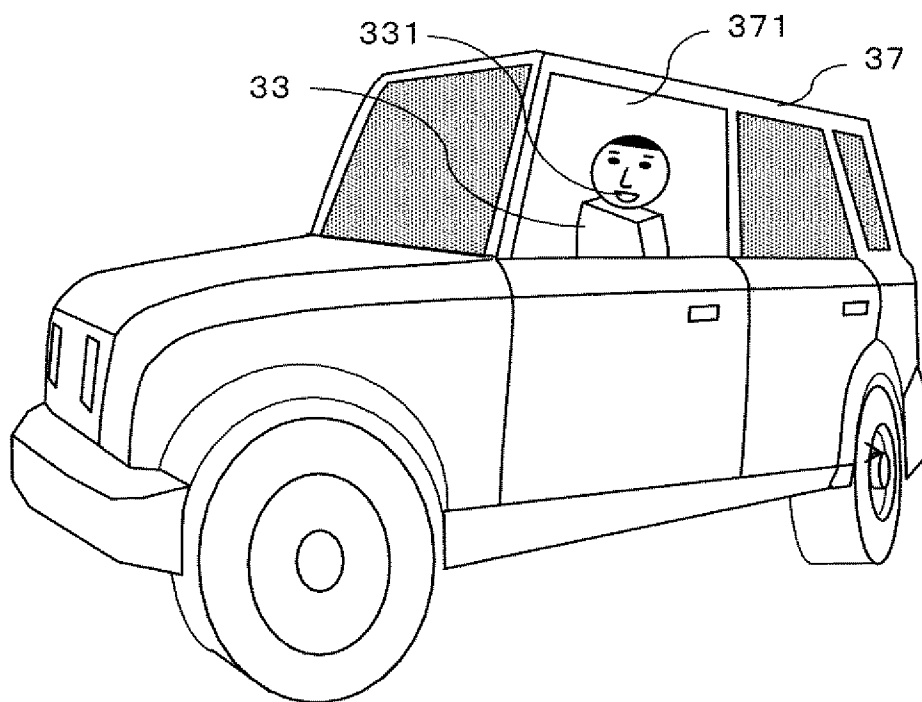
[図7A]

図 7 A



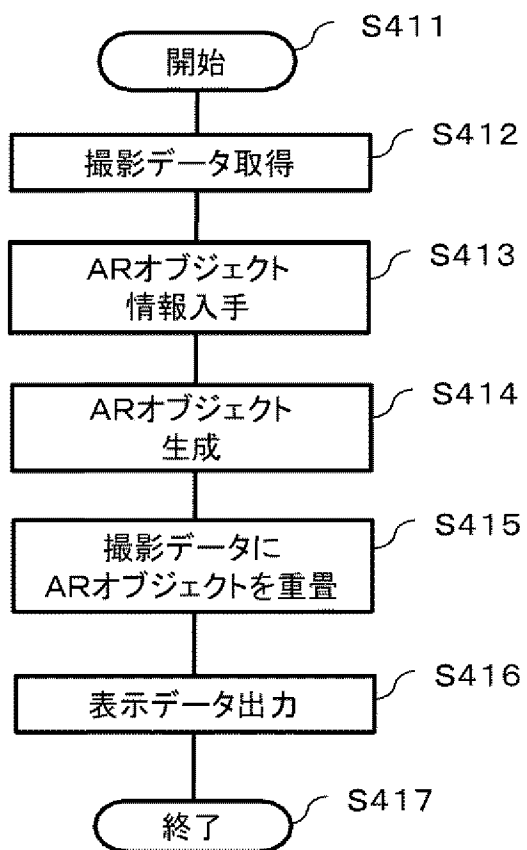
[図7B]

図 7 B

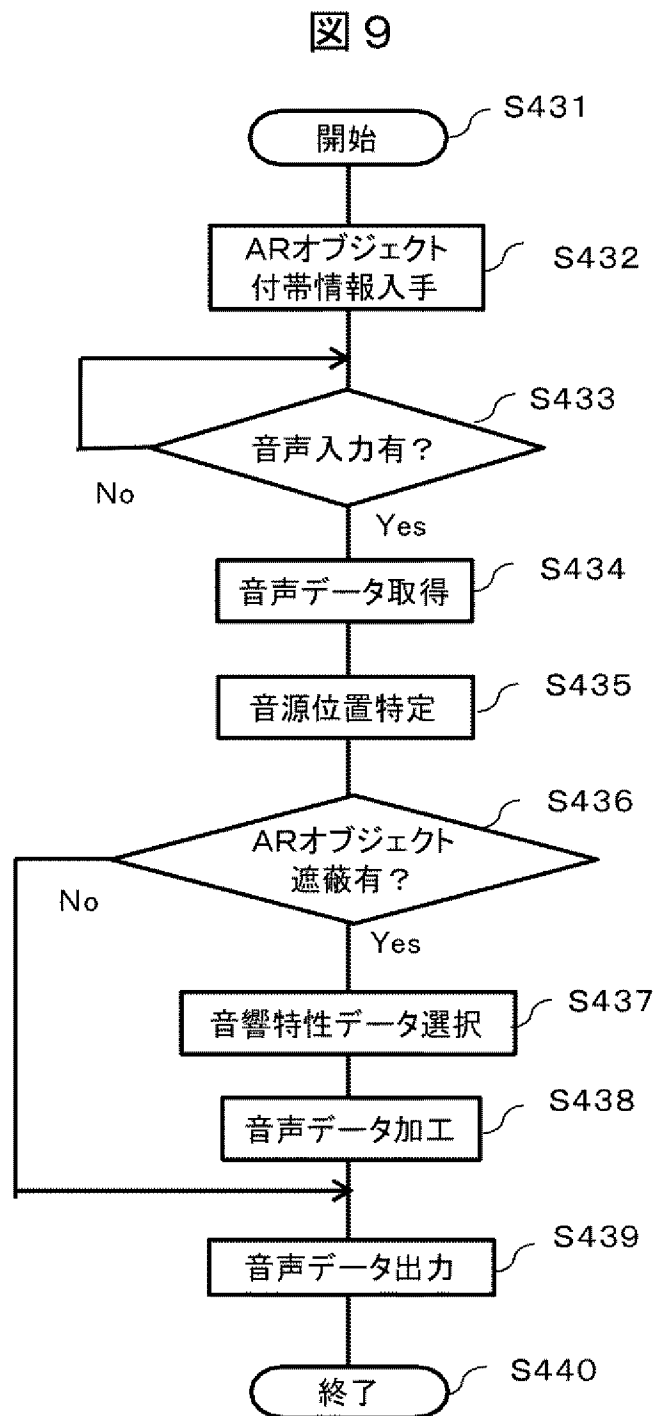


[図8]

図 8

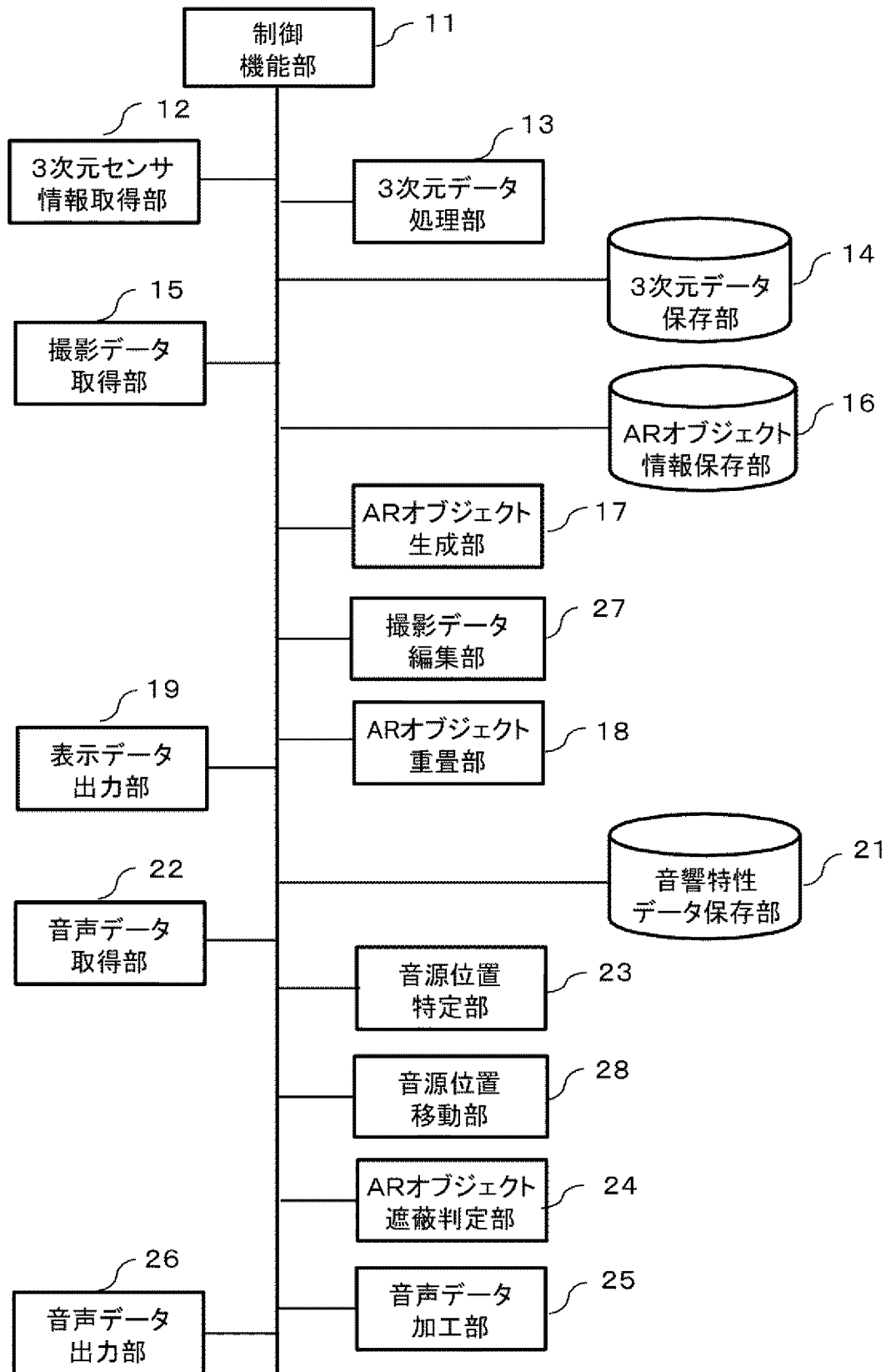


[図9]



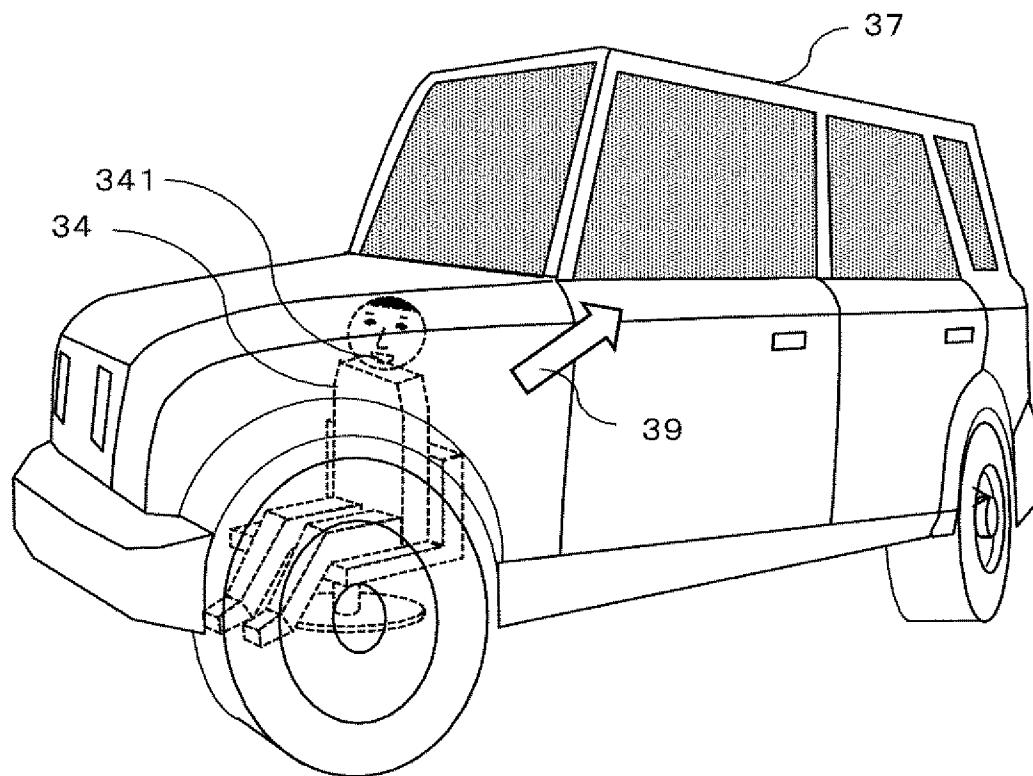
[図10]

図 10



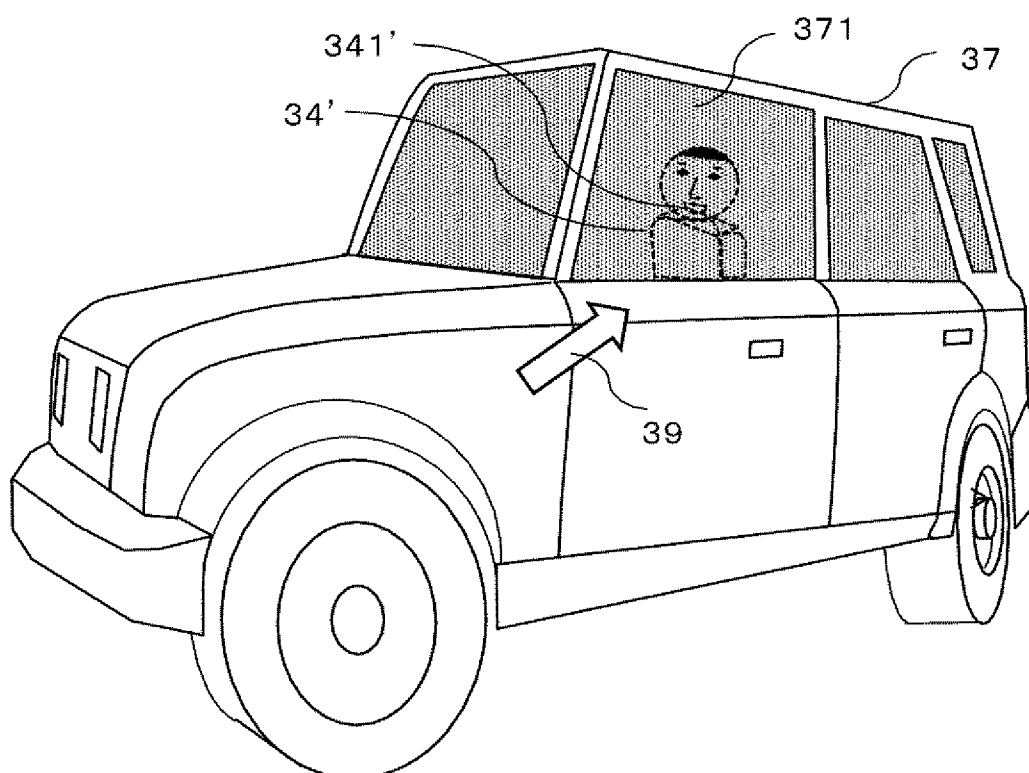
[図11A]

図 1 1 A



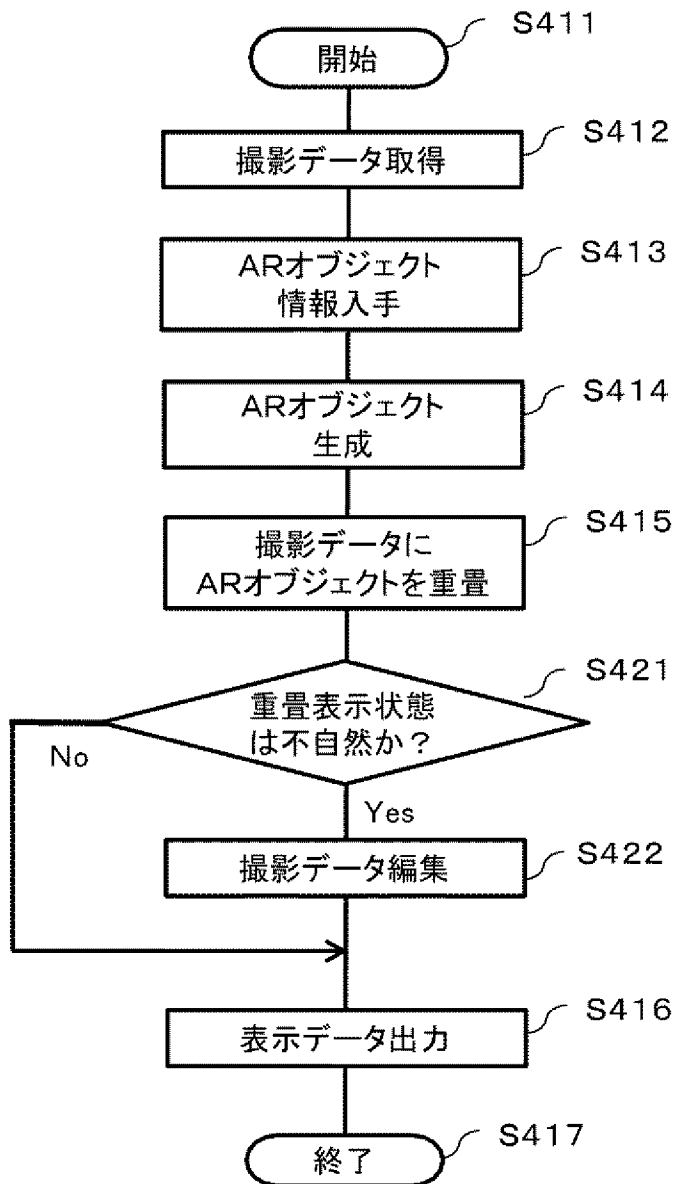
[図11B]

図 1 1 B



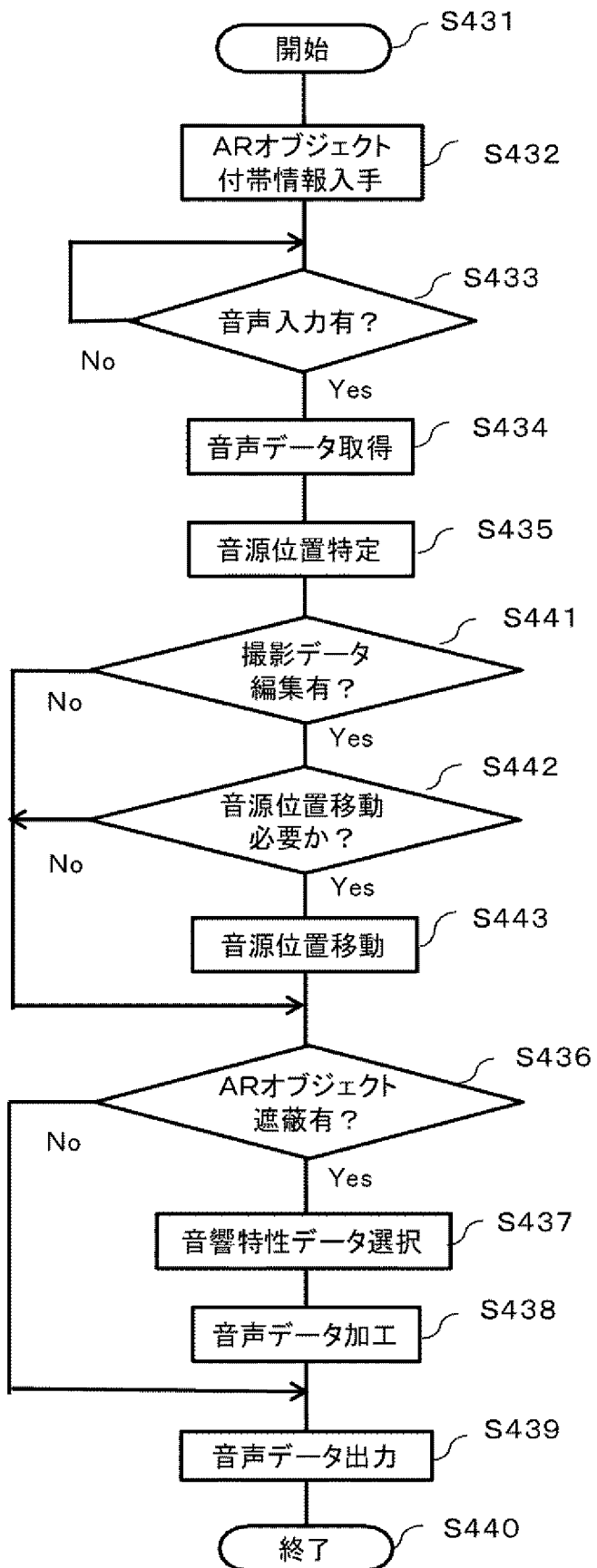
[図12]

図 1 2



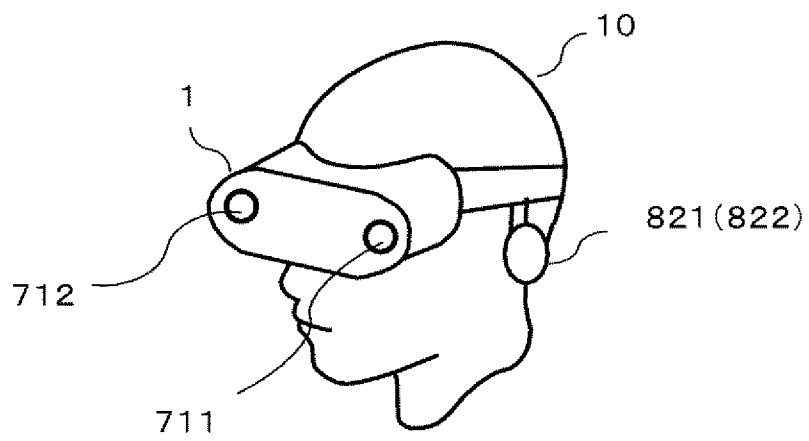
[図13]

図 1 3



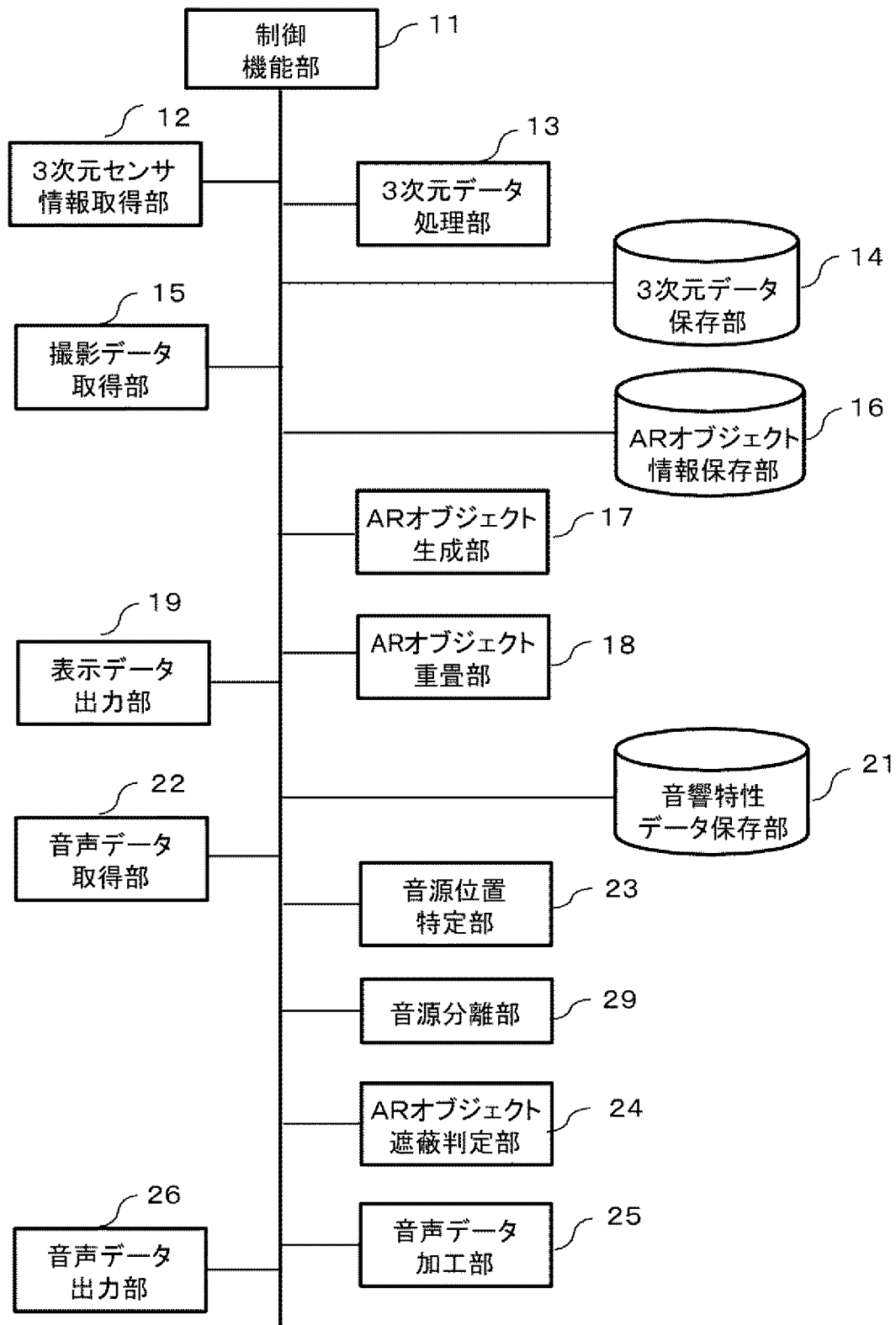
[図14]

図 1 4



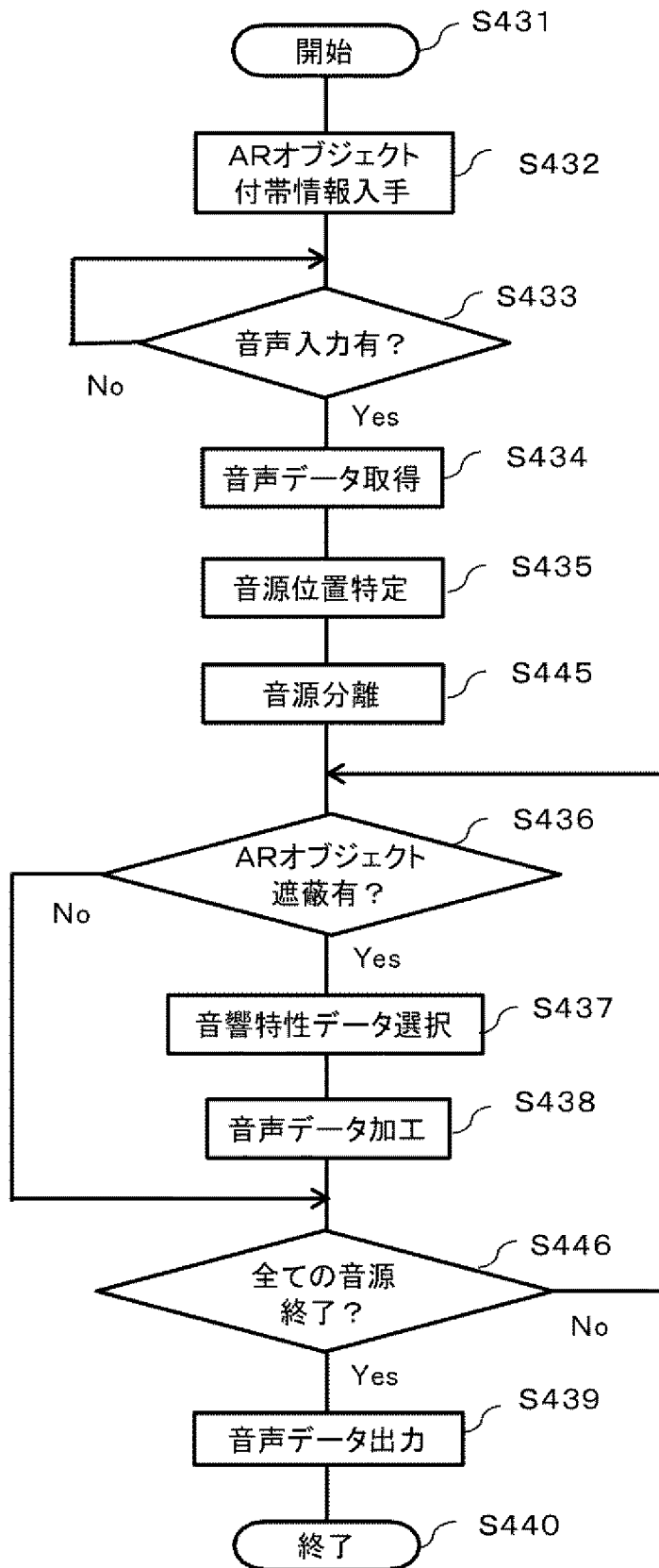
[図15]

図 15



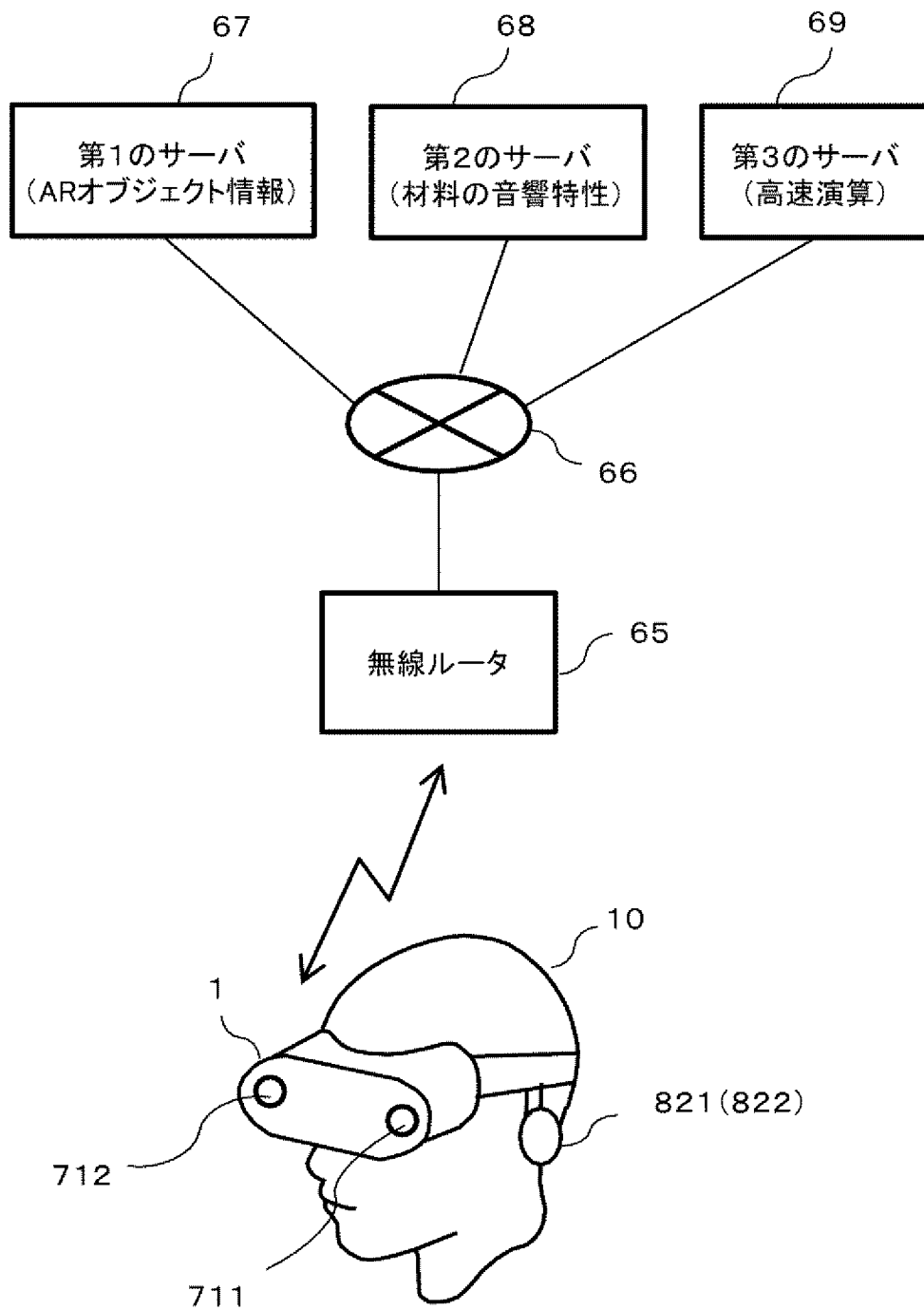
[図16]

図 1 6



[図17]

図 17



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/032962

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04R1/40 (2006.01) i, H04R3/00 (2006.01) i
FI: H04R3/00310, H04R3/00320, H04R1/40320A

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. H04R1/40, H04R3/00

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	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

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.
X	JP 2009-116690 A (CANON INC.) 28 May 2009 (2009-05-28), paragraphs [0018], [0037], [0040], [0044], [0045], [0049], [0051], [0054]-[0071], [0076]-[0078], fig. 4, 5	1-11
A	JP 2017-182460 A (SEIKO EPSON CORPORATION) 05 October 2017 (2017-10-05), paragraphs [0027], [0039], [0091]	5, 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

26 October 2020

Date of mailing of the international search report

17 November 2020

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
Information on patent family members

International application No.

PCT/JP2020/032962

JP 2009-116690 A 28 May 2009

US 2010/0265164 A1
paragraphs [0049], [0068], [0071],
[0075], [0076], [0080],
[0082], [0085]-[0102],
[0107]-[0109], fig. 4, 5

JP 2017-182460 A 05 October 2017

US 2017/0287222 A1
paragraphs [0073], [0085],
[0139]

A. 発明の属する分野の分類（国際特許分類（IPC）） H04R 1/40(2006.01)i; H04R 3/00(2006.01)i FI: H04R3/00 310; H04R3/00 320; H04R1/40 320A		
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）） H04R1/40; H04R3/00		
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 1971-2020年 日本国実用新案登録公報 1996-2020年 日本国登録実用新案公報 1994-2020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
X	JP 2009-116690 A (キヤノン株式会社) 28.05.2009 (2009 - 05 - 28) [0018]、[0037]、[0040]、[0044] - [0045]、[0049]、 [0051]、[0054] - [0071]、[0076] - [0078]、図4、図5	1-11
A	JP 2017-182460 A (セイコーエプソン株式会社) 05.10.2017 (2017 - 10 - 05) [0027]、[0039]、[0091]	5, 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抵触するものではなく、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 26.10.2020		国際調査報告の発送日 17.11.2020
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号		権限のある職員（特許庁審査官） 富澤 直樹 5Z 4188 電話番号 03-3581-1101 内線 3591

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2020/032962

引用文献			公表日	パテントファミリー文献	公表日
JP	2009-116690	A	28.05.2009	US 2010/0265164 A1 [0049]、[0068]、 [0071]、[0075]－ [0076]、[0080]、 [0082]、[0085]－ [0102]、[0107]－ [0109]、図4、図5	
JP	2017-182460	A	05.10.2017	US 2017/0287222 A1 [0073]、[0085]、 [0139]	