

(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
2024年3月14日(14.03.2024)



(10) 国際公開番号

WO 2024/053271 A1

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

(21) 国際出願番号: PCT/JP2023/027195

(22) 国際出願日: 2023年7月25日(25.07.2023)

(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 2022-143559 2022年9月9日(09.09.2022) JP

(71) 出願人: ソニーセミコンダクタソリューションズ株式会社(SONY SEMICONDUCTOR SOLUTIONS CORPORATION) [JP/JP]; 〒2430014

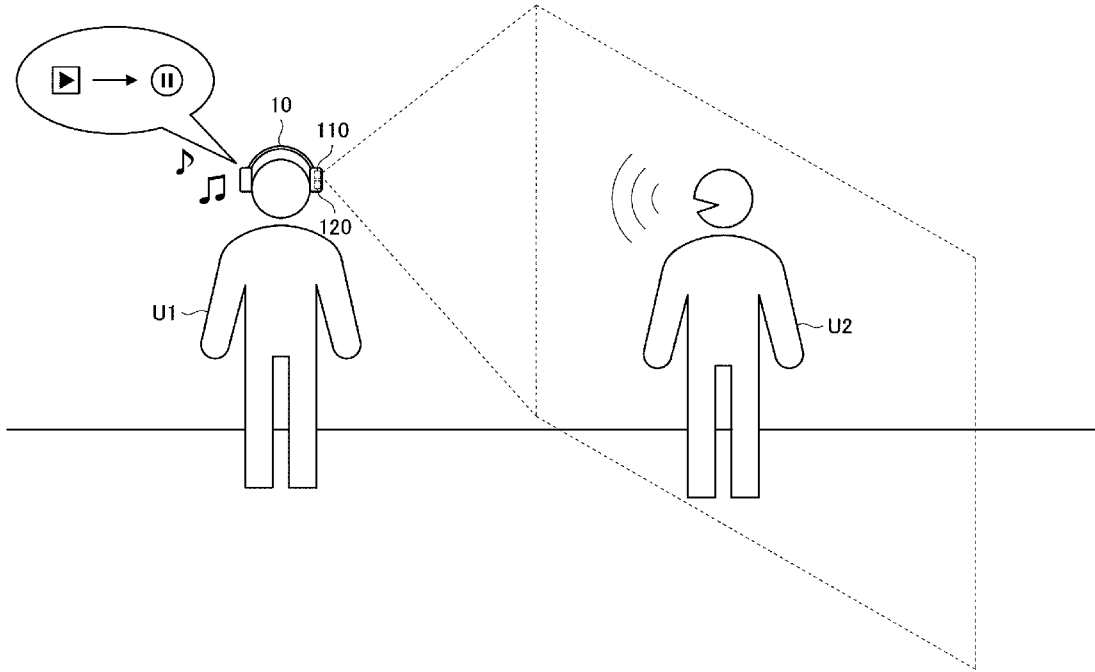
神奈川県厚木市旭町四丁目14番1号 Kanagawa (JP).

(72) 発明者: 桑原 彬任 (KUWABARA, Akito); 〒2430014 神奈川県厚木市旭町四丁目14番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 吉村 真一 (YOSHIMURA, Shinichi); 〒2430014 神奈川県厚木市旭町四丁目14番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP).

(74) 代理人: 伊藤 学, 外 (ITO, Manabu et al.); 〒1600023 東京都新宿区西新宿7-22-37 ストック西新宿413 協学国際特許事務所 Tokyo (JP).

(54) Title: INFORMATION PROCESSING DEVICE, PROGRAM, AND INFORMATION PROCESSING METHOD

(54) 発明の名称: 情報処理装置、プログラム、および情報処理方法



(57) Abstract: [Problem] To further improve convenience for a user in detecting a change in the surrounding environment based on sensing data. [Solution] Provided is an information processing device comprising: a determination result acquisition unit which acquires a determination result of whether a change in the surrounding environment of a sensor unit indicated by sensing data acquired from the sensor unit satisfies a predetermined condition; and a recognition unit which starts a recognition process for recognizing the situation of the surrounding environment of the sensor unit on the

[続葉有]

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))

basis of acquisition of the determination result which is acquired from the determination acquisition unit and indicates that the predetermined condition is satisfied.

(57) 要約: 【課題】 センシングデータに基づく周辺環境の変化の検出において、ユーザの利便性をさらに向上させる。 【解決手段】 センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、を備える、情報処理装置が提供される。

明 細 書

発明の名称： 情報処理装置、プログラム、および情報処理方法
技術分野

[0001] 本開示は、情報処理装置、プログラム、および情報処理方法に関する。

背景技術

[0002] ユーザがヘッドホン等の音再生装置を用いて音楽等のコンテンツを楽しむ際に、センシングデータに基づいて、センサ部またはセンサ部を含み構成された音再生装置を装着しているユーザの周辺状況の変化を検出し、検出結果に応じてコンテンツの出力を制御する技術が知られている。例えば、特許文献1には、ユーザがヘッドホンで音楽を聴いている際に、マイクロフォンで当該ユーザの発話を検出されると、再生音の音像定位を任意の位置に移動させる技術が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2011-097268号公報

発明の概要

発明が解決しようとする課題

[0004] 上記技術のように、センシングデータに基づき周辺環境の変化を検出する技術において、ユーザの利便性のさらなる向上が望まれる。

[0005] そこで、本開示では、センシングデータに基づく周辺環境の変化の検出において、ユーザの利便性をさらに向上させることが可能な、新規かつ改良された技術を提案する。

課題を解決するための手段

[0006] 上記課題を解決するために、本開示のある観点によれば、センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得された

ことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、を備える、情報処理装置が提供される。

[0007] また、本開示によれば、コンピュータを、センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、として機能させるための、プログラムが提供される。

[0008] また、本開示によれば、センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得することと、前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始することと、を含む、コンピュータにより実行される情報処理方法が提供される。

図面の簡単な説明

[0009] [図1]本開示の一実施形態による情報処理システムの概要を説明するための説明図である。

[図2]第一の実施形態による情報処理装置10の構成例を説明するためのブロック図である。

[図3]イメージセンサ部110の機能構成例を説明するブロック図である。

[図4]音差分処理部1501による、情報処理装置10の周辺環境の音の変化の判断結果取得処理を説明するための流れ図である。

[図5]音差分処理部1501による差分の算出について説明するための概念図である。

[図6]第一の実施形態による情報処理装置10の動作例を説明するためのフローチャートである。

[図7]第一の実施形態によるイメージセンサ部110の動作例を説明するためのフローチャートである。

[図8]第二の実施形態による情報処理装置 1 1 の構成例を説明するためのブロック図である。

[図9]第二の実施形態によるイメージセンサ部 1 1 1 の構成例を説明するためのブロック図である。

[図10]第二の実施形態による情報処理装置 1 1 の動作例を説明するためのフローチャートである。

[図11]第二の実施形態によるイメージセンサ部 1 1 1 の動作例を説明するためのフローチャートである。

[図12]第三の実施形態による情報処理装置 1 2 の構成例を説明するためのブロック図である。

[図13]第三の実施形態によるイメージセンサ部 1 1 2 の構成例を説明するためのブロック図である。

[図14]第三の実施形態による情報処理装置 1 2 の動作例を説明するためのフローチャートである。

[図15]第三の実施形態によるイメージセンサ部 1 2 2 の動作例を説明するためのフローチャートである。

[図16]第四の実施形態の概要を説明するための説明図である。

[図17]第四の実施形態による情報処理装置 1 3 の構成例を説明するためのブロック図である。

[図18]第四の実施形態によるイメージセンサ部 1 1 0 の設置位置の一例を説明するための説明図である。

[図19]第四の実施形態によるイメージセンサ部 1 1 0 の設置位置の他の一例を説明するための説明図である。

[図20]第四の実施形態によるイメージセンサ部 1 1 0 の設置位置の他の一例を説明するための説明図である。

[図21]第四の実施形態による情報処理装置 1 3 の動作例を説明するためのフローチャートである。

[図22]第四の実施形態によるイメージセンサ部 1 1 0 の動作例を説明するた

めのフローチャートである。

[図23]ハードウェア構成90の一例を示したブロック図である。

発明を実施するための形態

[0010] 以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

[0011] また、本明細書および図面において、実質的に同一の機能構成を有する複数の構成要素を、同一の符号の後に異なる数字を付して区別する場合もある。ただし、実質的に同一の機能構成を有する複数の構成要素の各々を特に区別する必要がない場合、複数の構成要素の各々に同一符号のみを付する。

[0012] なお、説明は以下の順序で行うものとする。

1. 概要
2. 第一の実施形態
 - 2-1. 構成例
 - 2-2. 動作
3. 第二の実施形態
 - 3-1. 構成例
 - 3-2. 動作
4. 第三の実施形態
 - 4-1. 構成例
 - 4-2. 動作
5. 第四の実施形態
 - 5-1. 構成例
 - 5-2. 動作
6. ハードウェア構成
7. 補足

[0013] <1. 概要>

本開示は、センシングデータに基づきユーザの周辺環境の変化を検出する技術に関する。以下では、本開示の好適な適用先として、ヘッドホンまたはHMD（Head Mounted Display）等の、ユーザが装着して音楽等のコンテンツを聴取することが可能な音再生装置を利用しているユーザの周辺環境の変化を、センシングデータに基づいて検出する例を主に説明する。

[0014] （課題の整理）

ヘッドホンまたはHMD等の頭部装着型の情報処理装置を装着しているユーザは、当該情報処理装置で音楽等のコンテンツ音を聴取していると、ユーザの周辺で発生した音に気付きにくい。また、ヘッドホンを装着して音楽等のコンテンツ再生音を聴いているユーザが他のユーザに話しかけようとすると、コンテンツ再生音が再生されたままでは、自身の音声と他のユーザの音声聞こえづらい。そのため、ユーザには、音楽等のコンテンツの一時停止操作、または、ヘッドホンを外す等の動作を行う手間が生じる。そこで、このような不都合を解消するための様々な技術が検討されている。

[0015] 例えば、上記で説明したように、特許文献1には、マイクロフォンで検出された音に基づき、周辺環境の音の変化を検出し、コンテンツ再生音の音像定位を任意の位置に移動させる技術が開示されている。

[0016] また、他の一例として、イヤホンの左右のそれぞれが有する人感センサによって、他のユーザの接近を検知し、ユーザへ通知する技術も行われている。または、他の一例として、超音波センサを用いて、イヤホンから所定の距離内に入った動体を検知し、動体を検知されるとコンテンツ再生音の消音等の方法でユーザに報知する技術も行われている。

[0017] しかし、上述のように、人感センサまたは超音波センサを用いて周辺環境の変化を検出する方法では、検出された物体または動体が、ユーザに話しかけるために接近してきた他のユーザであるのか、または、単にユーザの側を通り過ぎただけの物体または他のユーザであるのかが区別されない。

[0018] また、上述のように、周辺環境の音の変化に基づいてコンテンツ再生音の

出力を制御する方法では、ユーザUが咳をした音、または、ユーザUの独り言の音等の、他のユーザへの話しかけまたは会話ではない音が発生した場合にも、コンテンツ再生音の音像定位の移動の処理が行われる場合があった。

[0019] そこで、本開示による情報処理システムは、センシングデータに基づく周辺環境の変化の検出において、ユーザの利便性をさらに向上させることが可能な技術を提供することを目的とする。より具体的に、情報処理装置10は、情報処理装置10の周辺の音および周辺の画像の2種類のセンシングデータを用いて、情報処理装置10の周辺に位置するユーザと、他のユーザとの間での話しかけまたは会話の有無を認識する。

[0020] 図1は、本開示の一実施形態による情報処理システムの概要を説明するための説明図である。図1に示したように、本開示による情報処理システムは、情報処理装置10を含む。

[0021] 情報処理装置10は、音を再生することが可能な情報処理装置である。図1に示した例では、情報処理装置10はヘッドホンである。情報処理装置10は、イメージセンサ部110およびマイクロフォン120を備える。

[0022] 情報処理装置10は、音楽等のユーザの所望のコンテンツを出力する機能を有する。図1に示した例では、情報処理装置10は、音楽等の音コンテンツを出力する。情報処理装置10は、図1に図示しない、スマートフォン、タブレット型端末、またはPC (Personal Computer) 等の任意のユーザ利用端末と有線または無線接続により通信を行い、コンテンツの再生、停止、または早送り等の操作指示を受信してもよい。

[0023] イメージセンサ部110は、情報処理装置10の周辺環境の画像を取得する機能を有する。例えば、イメージセンサ部110は、情報処理装置10に接近している他のユーザU2の画像を取得する。情報処理装置10は、当該画像を解析することにより、例えば情報処理装置10に接近している他のユーザU2を検出することが出来る。

[0024] マイクロフォン120は、情報処理装置10の周辺の音を検出する機能を有する。例えば、マイクロフォン120は、ユーザUに話しかけている、ま

たは、ユーザUと会話している他のユーザU 2の音声を取得する。または、マイクロフォン120は、他のユーザU 2に話かけている、または、他のユーザU 2と会話しているユーザUの音声を取得することが出来る。

[0025] 情報処理装置10は、マイクロフォン120により取得される周辺の音に基づいて、情報処理装置10の周辺環境の変化を検出する。さらに、情報処理装置10は、検出された周辺環境の変化が、所定の条件を満たすか否かの判断結果を取得する。所定の条件は、例えば、情報処理装置10の周辺の音が、所定の時間前と比較して、所定の閾値が示す差分以上に変化したことであってもよい。

[0026] より具体的には、例えば、他のユーザU 2がユーザU 1に話しかけたとする。情報処理装置10は、他のユーザU 2がユーザU 1に話しかけている音声をマイクロフォン120により取得する。情報処理装置10は、当該音声と、他のユーザU 2がユーザU 1に話しかける前の周辺の音との差分をとることにより、周辺の音が所定の閾値が示す差分以上に変化したことを検出する。すると、情報処理装置10は、周辺環境の変化が所定の条件を満たすと判断する。

[0027] また、情報処理装置10は、周辺環境の変化が所定の条件を満たす場合に、イメージセンサ部110により取得される情報処理装置10の周辺環境の画像、および、マイクロフォン120により取得される情報処理装置10の周辺の音を解析して、情報処理装置10の周辺環境の状況を認識する認識処理を行う。

[0028] 例えば、情報処理装置10は、情報処理装置10の周辺環境の画像および情報処理装置10の周辺の音に基づいて、ユーザUと他のユーザとの間での話しかけ、または、会話の有無を認識してもよい。さらに、情報処理装置10は、認識結果に応じて、コンテンツの出力を制御してもよい。

[0029] 例えば、情報処理装置10は、ユーザUと他のユーザとの間での話しかけが有ると認識すると、情報処理装置10が有するスピーカにおいて出力されるコンテンツ再生音を一時停止させてもよい。さらに、情報処理装置10は

、話しかけが有ると認識すると、マイクロフォン１２０により取得される情報処理装置１０の周辺の音（以下、外部音と称する）を、上記スピーカから出力させてもよい。

[0030] これにより、ユーザＵが情報処理装置１０を装着して音楽等のコンテンツを聞いている間に、ユーザＵ１または他のユーザＵ２がどちらかに話しかけたり、会話を継続したりしていることを認識することが出来る。さらに、ユーザＵと他のユーザとの間での話しかけまたは会話が有ると認識されると、情報処理装置１０から出力されるコンテンツ再生音が一時停止されるとともに、外部音が再生される。従って、ユーザＵが、他のユーザから話しかけられていることに気付きやすくなる。また、ユーザＵが、情報処理装置１０を装着したままでも、他のユーザに話しかけたり、会話を行ったりしやすくなる。従って、ユーザの利便性が向上する。

[0031] また、情報処理装置１０は、情報処理装置１０の周辺の音が示す周辺環境の変化が、所定の条件を満たす場合にのみ、周辺環境の状況の認識処理を行う。これにより、情報処理装置１０が、常に認識処理を行うような構成と比較して、消費電力を低減することが出来る。

[0032] 以上、図１を参照して、本開示の一実施形態による情報処理システムの概要を説明した。続いて、図２～図７を参照して、本開示の情報処理システムの第一の実施形態を説明する。

[0033] ＜２．第一の実施形態＞

＜２－１．構成例＞

まず、図２を参照して、第一の実施形態による情報処理装置１０の構成例を説明する。図２は、第一の実施形態による情報処理装置１０の構成例を説明するためのブロック図である。図２に示したように、情報処理装置１０は、イメージセンサ部１１０、マイクロフォン１２０、スピーカ１３０、記憶部１４０、および、制御部１５０を有する。

[0034] （イメージセンサ部１１０）

イメージセンサ部１１０は、情報処理装置１０の周辺環境を撮像する機能

を有する。情報処理装置 10 は、複数のイメージセンサ部 110 を有していてもよい。ここで、図 3 は、イメージセンサ部 110 の機能構成例を説明するブロック図である。図 3 に示したように、イメージセンサ部 110 は、画素アレイ 210 およびイメージセンサ制御部 220 を有する。

[0035] 画素アレイ 210 は、フォトダイオード等の受光素子によって構成され、光を検出すると電気信号を生成する機能を有する。画素アレイ 210 は、アレイ状に配置された複数の受光素子によって構成される。画素アレイ 210 は、情報処理装置 10 の周辺環境において反射した光を検出し、電気信号に変換することにより、情報処理装置 10 の周辺環境の画像を撮像する。

[0036] イメージセンサ制御部 220 は、イメージセンサ部 110 の動作全般を制御する機能を有する。このようなイメージセンサ制御部 220 は、画像縮小処理部 2201、および、画像出力部 2202 としての機能を有する。

[0037] 画像縮小処理部 2201 は、画素アレイ 210 から出力された電気信号に信号処理を行うことにより、画素アレイ 210 により取得された情報処理装置 10 の周辺環境の画像データの縮小処理を行う機能を有する。画像縮小処理部 2201 は、後述する情報処理装置 10 の認識部 1503 により行われる認識処理の入力データとして設定された画像サイズにあわせて、上記画像の縮小処理を行う。これにより、認識部 1503 の処理能力に応じて、認識部 1503 の認識処理の負荷を調整することが出来る。

[0038] なお、認識部 1503 により行われる認識処理の入力データとして設定された画像サイズに応じて、画素アレイ 210 により取得された画像データの縮小が不要である場合には、画像縮小処理部 2201 は上記画像の縮小処理を行わなくてもよい。この場合、イメージセンサ制御部 220 は、画像縮小処理部 2201 としての機能を有していなくてもよい。

[0039] 画像出力部 2202 は、画像縮小処理部 2201 による縮小処理後の周辺環境の画像データを制御部 150 に出力する。または、画像出力部 2202 は、画像縮小処理部 2201 による縮小処理が行われない場合には、画素アレイ 210 により取得された周辺環境の画像データを、制御部 150 へ出力

する。

[0040] 以上、図3を参照して、イメージセンサ部110の構成例を説明した。

[0041] (マイクロフォン120)

マイクロフォン120は、情報処理装置10の周辺の音を取得する機能を有する。例えば、マイクロフォン120は、ユーザU1および他のユーザU2の音声の音データを取得する。情報処理装置10は、取得した周辺の音データを、制御部150へ出力する。

[0042] また、マイクロフォン120は、制御部150の制御に従い、所定時間長（例えば、数秒間）を有する時間区間において取得した音を、記憶部140の音メモリに記憶させる機能を有する。例えば、マイクロフォン120は、過去数秒分の周辺の音のデータを、音メモリに記憶させる。

[0043] (スピーカ130)

スピーカ130は、音を出力する機能を有する。スピーカ130は、骨伝導スピーカとして構成されてもよい。

[0044] (記憶部140)

記憶部140は制御部150の処理に用いられるプログラムおよび演算パラメータ等を記憶するROM (Read Only Memory)、および適宜変化するパラメータ等を一時記憶するRAM (Random Access Memory) により実現される。

[0045] (制御部150)

制御部150は、情報処理装置10の動作全般を制御する。このような制御部150は、音差分処理部1501、認識部1503、および、出力制御部1505としての機能を有する。

[0046] 音差分処理部1501は、マイクロフォン120により所定時間長を有する第一の時間区間において取得された第一の音データと、マイクロフォン120により第一の時間区間と異なり上記所定時間長を有する第二の時間区間において取得された第二の音データと、に基づいて、情報処理装置10の周辺の音の変化が所定の条件を満たすか否かの判断結果を取得する。より詳細

には、音差分処理部 1501 は、上記第一の音データと、第二の音データの差分が、所定の閾値を超えるかどうかを判断する。

[0047] ここで、図 4 および図 5 を参照して、音差分処理部 1501 による、音の変化が所定の条件を満たすか否かの判断結果取得の処理についてより具体的に説明する。

[0048] 図 4 は、音差分処理部 1501 による、情報処理装置 10 の周辺環境の音の変化の判断結果取得処理を説明するための流れ図である。

[0049] 図 4 に示したように、まず、マイクロフォン 120 により、情報処理装置 10 の周辺の音データが取得される。マイクロフォン 120 は、制御部 150 の制御に従って、現時点より過去に取得された、所定の時間長（例えば、数秒分）の情報処理装置 10 の周辺の音を、過去の音データとして記憶部 140 の音メモリに記憶させる。過去の音データは、第二の音データの一例である。

[0050] また、マイクロフォン 120 は、上記過去の音データと同じ所定の時間長の、現在の情報処理装置 10 の周辺の音データ（例えば、現時点から過去数秒分の音データ）を、現在の音データとして音差分処理部 1501 に出力する。

[0051] 音差分処理部 1501 は、マイクロフォン 120 および記憶部 140 のそれぞれから、過去の音データおよび現在の音データを取得して、過去の音データと現在の音データの差分を算出する。

[0052] 図 5 は、音差分処理部 1501 による差分の算出について説明するための概念図である。図 5 に示した波形 W は、各音データの波形のイメージ図である。

[0053] 波形 W1 は、過去の音データを表す。また、波形 W2 は、現在の音データを表す。音差分処理部 1501 は、波形 W1 と波形 W2 の音データの差分を算出する。波形 W3 は、音データの差分を表す。

[0054] また、音差分処理部 1501 は、波形 W3 が示す差分の絶対値を算出する。波形 W4 は、算出された差分の絶対値を表す。

- [0055] さらに、音差分処理部 1501 は、差分の絶対値を積分し、積分値を算出する。算出された積分値は、前記差分を示す差分値の一例である。
- [0056] 音差分処理部 1501 は、算出された上記差分値が、所定の閾値を越えた場合に、情報処理装置 10 の周辺環境の変化が所定の条件を満たすと判断する。音差分処理部 1501 は、判断結果を制御部 150 に出力する。このような音差分処理部 1501 は、情報処理装置 10 の判断結果取得部の一例である。
- [0057] または、音差分処理部 1501 は、算出された差分値を、制御部 150 に出力してもよい。この場合、制御部 150 が、当該差分値が所定の閾値を超えているかどうかを判断してもよい。
- [0058] 図 4 に戻って説明を続ける。制御部 150 は、音差分処理部 1501 から取得した判断結果に基づき、後述する認識部 1503 の認識処理を開始させる指示を出力するか否かを判定する。
- [0059] 図 2 に戻って情報処理装置 10 の構成例の説明を続ける。認識部 1503 は、音差分処理部 1501 または制御部 150 により、情報処理装置 10 の周辺環境の変化が所定の条件を満たすことを示す判断結果が取得されると、情報処理装置 10 の周辺環境の状況を認識する認識処理を開始する。
- [0060] より具体的に、認識部 1503 は、マイクロフォン 120 により取得される情報処理装置 10 の周辺の音データと、イメージセンサ部 110 により取得される情報処理装置 10 の周辺環境の画像データを解析する。また、認識部 1503 は、解析結果に基づき、情報処理装置 10 の周辺環境の状況として、情報処理装置 10 の周辺に位置するユーザと他のユーザの間での話しかけまたは会話の有無を認識する。
- [0061] 認識部 1503 は、例えば、情報処理装置 10 の周辺の音データおよび情報処理装置 10 の周辺環境の画像データと、情報処理装置 10 の周辺に位置するユーザと他のユーザとの間での話しかけまたは会話の有無の正解データと、を入力データとするニューラルネットワークを用いて、情報処理装置 10 の周辺の音データおよび周辺環境の画像データと、上記ユーザと他のユーザ

ザとの間での話しかけまたは会話の有無の関係を学習してもよい。また、認識部 1503 は、学習の結果得られたモデルを用いて、情報処理装置 10 の周辺に位置するユーザと他のユーザとの間での話しかけまたは会話の有無を認識してもよい。

[0062] または、認識部 1503 は、情報処理装置 10 の周辺に位置するユーザと他のユーザとの間での話しかけが無い状態での、情報処理装置 10 の周辺の音データおよび情報処理装置 10 の周辺環境の画像データのみを入力データとするニューラルネットワークを用いて、情報処理装置 10 の周辺に位置するユーザと他のユーザとの間での話しかけが無い状態の情報処理装置 10 の周辺の音データおよび周辺環境の画像データを学習してもよい。この場合、認識部 1503 は、情報処理装置 10 の周辺に位置するユーザと他のユーザとの間での話しかけまたは会話がある状態の情報処理装置 10 の周辺の音データおよび情報処理装置 10 の周辺環境の画像データが入力された場合に、情報処理装置 10 の周辺に位置するユーザと他のユーザとの間での話しかけまたは会話を検出してもよい。

[0063] このとき、認識部 1503 は、上記入力データとして、画像縮小処理が行われた状態の周辺環境の画像データを用いてもよい。これにより、認識部 1503 による認識処理の負荷を低減することが可能である。

[0064] 出力制御部 1505 は、認識部 1503 により情報処理装置 10 の周辺のユーザと他のユーザとの間での話しかけまたは会話があると認識されると、スピーカ 130 により出力されるコンテンツの出力を制御する機能を有する。

[0065] 例えば、出力制御部 1505 は、認識部 1503 により話しかけまたは会話があると認識されると、スピーカ 130 により出力されていたコンテンツ再生音を一時停止させる処理を行ってもよい。または、コンテンツの出力の制御の他の一例として、出力制御部 1505 は、情報処理装置 10 の外部音をスピーカ 130 に出力させてもよい。または、コンテンツの出力の制御の他の一例として、出力制御部 1505 は、スピーカ 130 により出力される

コンテンツ再生音の音量を低減させる処理、または、コンテンツ音の音像定位を変化させる処理を行ってもよい。

[0066] さらに、情報処理装置 10 は、図 2 に図示しない、連続または断続した振動を発生させることが可能なバイブレータ（振動機）を備えていてもよい。この場合、出力制御部 1505 は、コンテンツの出力の制御の他の一例として、当該バイブレータに振動を発生させる制御を行ってもよい。または、情報処理装置 10 が任意のユーザ利用端末と通信接続されており、かつ、当該ユーザ利用端末にバイブレータ機能が有る場合には、出力制御部 1505 は、コンテンツの出力の制御の他の一例として、当該ユーザ利用端末を振動させる制御を行ってもよい。

[0067] このように、認識部 1503 によりユーザ間での話しかけまたは会話が有ると認識された場合に出力制御部 1505 がコンテンツの出力の制御を行うことで、ユーザ U1 は、情報処理装置 10 でコンテンツ再生音を聞いている場合でも、周囲から話しかけられていることに気づきやすくなる。また、ユーザ U1 が周囲に話しかける、または、他のユーザと会話を行う間に、コンテンツ再生音の停止操作または情報処理装置 10 を頭部から外す動作を行う手間が省略される。従って、ユーザの利便性が向上する。

[0068] 以上、図 2 ～図 6 を参照して、第一の実施形態による情報処理装置 10 の構成例を説明した。続いて、図 6 を参照して、第一の実施形態による情報処理装置 10 の動作例を説明する。

[0069] <2-2. 動作>

図 6 は、第一の実施形態による情報処理装置 10 の動作例を説明するためのフローチャートである。まず、情報処理装置 10 の音差分処理部 1501 が、マイクロフォン 120 および記憶部 140 のそれぞれから、過去の音データおよび現在の音データを取得する（S101）。

[0070] 次いで、音差分処理部 1501 が、過去の音データと現在の音データの差分値を算出する（S103）。

[0071] 算出された差分値が所定の閾値を越えていない場合（S105／NO）、

情報処理装置１０は、Ｓ１０１に戻って一連の処理を繰り返す。

[0072] 算出された差分値が所定の閾値を越えている場合（Ｓ１０５／ＹＥＳ）、制御部１５０は、認識部１５０３に、情報処理装置１０の周辺環境の状況の認識処理を開始する指示を出力する（Ｓ１０５）。認識部１５０３は、認識処理を行う（Ｓ１１１）。

[0073] 認識部１５０３により、情報処理装置１０の周辺のユーザと他のユーザの間での話しかけ、または、会話が有ると認識されなかった場合（Ｓ１１３／ＮＯ）、出力制御部１５０５は、スピーカ１３０に外部音を出力させる処理である、外部音取り込み機能をＯＦＦにし（Ｓ１１９）、Ｓ１０１に戻って一連の処理を繰り返す。

[0074] 認識部１５０３により、情報処理装置１０の周辺のユーザと他のユーザの間での話しかけ、または、会話が有ると認識された場合（Ｓ１１３／ＹＥＳ）、出力制御部１５０５は、外部音取り込み機能をＯＮにする（Ｓ１１５）。

[0075] 情報処理装置１０の一連の処理を終了する終了操作が行われた場合（Ｓ１１７／ＹＥＳ）、情報処理装置１０は処理を終了する。情報処理装置１０は、一連の処理を終了する終了操作が行われるまで、Ｓ１０１～Ｓ１１９の処理を繰り返す（Ｓ１１７／ＮＯ）。

[0076] 以上、図６を参照して、第一の実施形態による情報処理装置１０の動作例を説明した。続いて、図７を参照して、第一の実施形態によるイメージセンサ部１１０の動作例を説明する。

[0077] 図７は、第一の実施形態によるイメージセンサ部１１０の動作例を説明するためのフローチャートである。図７のフローチャートに示した動作は、図６のフローチャートに示した動作の開始と同時に開始され、並列して実行される。

[0078] まず、イメージセンサ部１１０の画素アレイ２１０は、情報処理装置１０の周辺環境の画像データの取得を開始する（Ｓ２０１）。画素アレイ２１０は、制御部１５０から、認識部１５０３による周辺環境の状況の認識処理を

開始する指示が出力されるまで、周辺環境の画像データの取得を継続する（S203／NO）。

[0079] 制御部150から、認識部1503による周辺環境の状況の認識処理を開始する指示が出力されると（S203／YES）、イメージセンサ部110の画像縮小処理部2201は、画素アレイ210により取得された周辺環境の画像データを、所定の解像度に縮小する処理を行う（S205）。

[0080] 画像出力部2202は、画像縮小処理部2201により縮小処理が行われた周辺環境の画像データを、認識部1503に出力する（S207）。

[0081] イメージセンサ部110の一連の処理を終了する終了操作が行われた場合（S209／YES）、イメージセンサ部110は処理を終了する。イメージセンサ部110は、一連の処理を終了する終了操作が行われるまで、S201～S209の処理を繰り返す（S209／NO）。

[0082] 以上、図7を参照して、第一の実施形態によるイメージセンサ部110の動作例を説明した。続いて、図8～図11を参照して、本開示の第二の実施形態による情報処理システムについて説明する。

[0083] <3. 第二の実施形態>

第二の実施形態では、イメージセンサ部111のイメージセンサ制御部221が、情報処理装置11の周辺環境の状況を認識する認識処理を行う。まず、図8および図9を参照して、第二の実施形態による情報処理装置11の構成例を説明する。

[0084] <3-1. 構成例>

図8は、第二の実施形態による情報処理装置11の構成例を説明するためのブロック図である。図8に示したように、情報処理装置11は、イメージセンサ部111、マイクロフォン120、スピーカ130、記憶部140、および、制御部151を有する。なお、マイクロフォン120、スピーカ130、および、記憶部140は、上記で図2を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0085] イメージセンサ部111は、第一の実施形態によるイメージセンサ部11

0と同様に、情報処理装置10の周辺環境の画像を取得する機能を有する。さらに、イメージセンサ部111は、第一の実施形態における認識部1503、および、出力制御部1505に相当する機能を有する。

[0086] 図9は、第二の実施形態によるイメージセンサ部111の構成例を説明するためのブロック図である。図9に示したように、イメージセンサ部111は、画素アレイ210およびイメージセンサ制御部221を有する。

[0087] 画素アレイ210は、第一の実施形態による画素アレイ210に相当する機能を有する。

[0088] イメージセンサ制御部221は、第一の実施形態によるイメージセンサ制御部220と同様に、画像縮小処理部2201、および、画像出力部2202としての機能を有する。さらに、第二の実施形態によるイメージセンサ制御部221は、認識部2203および出力制御部2204としての機能を有する。

[0089] 認識部2203は、第一の実施形態による認識部1503に相当する機能を有する。画像出力部2202は、情報処理装置10の音差分処理部1501により、認識処理を開始する指示が出力されると、画像出力部2202から出力される情報処理装置10の周辺環境の画像データと、情報処理装置11の120により取得される情報処理装置10の周辺の音データに基づいて、情報処理装置10の周辺環境の状況を認識する認識処理を行う。認識部2203は、認識処理の結果を出力制御部2204に出力する。

[0090] 出力制御部2204は、第一の実施形態による出力制御部1505に相当する機能を有する。出力制御部2204は、認識部2203による周辺環境の状況の認識結果に応じて、スピーカ130から出力されるコンテンツを制御する。

[0091] (制御部151)

図8に戻って、情報処理装置11の構成例の説明を続ける。制御部151は、情報処理装置11の動作全般を制御する機能を有する。このような情報処理装置11は、音差分処理部1501としての機能を有する。音差分処理

部１５０１は、上記で図２を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0092] 以上、図８および図９を参照して、第二の実施形態による情報処理装置１１の構成例を説明した。

<３－２．動作>

図１０は、第二の実施形態による情報処理装置１１の動作例を説明するためのフローチャートである。なお、図１０において、Ｓ１０１、Ｓ１０３、Ｓ１０７、およびＳ１１７は、上記で図６を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0093] Ｓ１０１～Ｓ１０３の処理が行われたのち、情報処理装置１１の音差分処理部１５０１は、算出された差分値が所定の閾値を超えているか否かを判定する。差分値が所定の閾値を超えていない場合（Ｓ３０１／ＮＯ）、情報処理装置１１は、Ｓ１０１～Ｓ３０１の処理を繰り返す。

[0094] 差分値が所定の閾値を超えている場合（Ｓ３０１／ＹＥＳ）、制御部１５１は、イメージセンサ部１１１へ、認識処理を開始する指示を出力する（Ｓ１０７）。次いで、Ｓ１１７の処理が行われる（Ｓ１１７）。

[0095] 以上、図１０を参照して、第二の実施形態による情報処理装置１１の動作例を説明した。続いて、図１１を参照して、第二の実施形態によるイメージセンサ部１１１の動作例を説明する。

[0096] 図１１は、第二の実施形態によるイメージセンサ部１１１の動作例を説明するためのフローチャートである。なお、図１１において、Ｓ２０１、Ｓ２０５、Ｓ２０７、およびＳ２０９は、上記で図７を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0097] Ｓ２０１の処理に次いで、イメージセンサ部１１１の画像縮小処理部２２０１は、情報処理装置１０の制御部１５１により、認識処理を開始する指示が出力されたか否かを判定する。認識処理を開始する指示が出力されていない場合（Ｓ４０３／ＮＯ）、Ｓ２０１～Ｓ４０３の処理を繰り返す。

[0098] 認識処理を開始する指示が出力された場合（Ｓ４０３／ＹＥＳ）、Ｓ２０

5～S207の処理が行われる。次いで、イメージセンサ部111の認識部2203が、周辺環境の状況を認識する認識処理を行う（S409）。

[0099] 認識部2203により、情報処理装置10の周辺のユーザと他のユーザの間での話しかけ、または、会話が有ると認識されなかった場合（S413／NO）、出力制御部2204は、外部音取り込み機能をOFFにするための制御信号を出力し（S417）、S201に戻って一連の処理を繰り返す。

[0100] 認識部2203により、情報処理装置10の周辺のユーザと他のユーザの間での話しかけ、または、会話が有ると認識された場合（S413／YES）、出力制御部2204は、外部音取り込み機能をONにするための制御信号を出力する（S415）。次いで、S209の処理が行われる。

[0101] 以上、図11を参照して、第二の実施形態によるイメージセンサ部111の動作例を説明した。続いて、図12～図15を参照して、本開示の第三の実施形態による情報処理システムについて説明する。

[0102] <4. 第三の実施形態>

第三の実施形態では、第一の実施形態において認識部1503が行っていた、情報処理装置10の周辺環境の状況の認識処理が、後述する情報処理装置12の制御部152と、イメージセンサ部112のイメージセンサ制御部222とに分担して実施される。

[0103] <4-1. 構成例>

図12は、第三の実施形態による情報処理装置12の構成例を説明するためのブロック図である。図12に示したように、情報処理装置12は、イメージセンサ部112、マイクロフォン120、スピーカ130、記憶部140、および、制御部152を有する。なお、マイクロフォン120、スピーカ130、および、記憶部140は、上記で図2を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0104] イメージセンサ部112は、第一の実施形態によるイメージセンサ部110と同様に、情報処理装置12の周辺環境の画像を取得する機能を有する。さらに、イメージセンサ部112は、情報処理装置12の周辺環境の画像を

解析して画像認識を行う機能を有する。

[0105] また、イメージセンサ部 112 は、後述する音認識処理部 1506 により出力される情報処理装置 12 の周辺の音の解析結果と、イメージセンサ部 112 による周辺環境の画像の解析結果に基づいて、情報処理装置 12 の周辺環境の状況を推定する機能を有する。情報処理装置 12 は、例えば、情報処理装置 12 に障害物が接近していることを推定し、スピーカ 130 に、ユーザへの通知等を出力させる。

[0106] 図 13 は、第三の実施形態によるイメージセンサ部 112 の構成例を説明するためのブロック図である。図 13 に示したように、イメージセンサ部 112 は、画素アレイ 210 およびイメージセンサ制御部 222 を有する。

[0107] 画素アレイ 210 は、第一の実施形態による画素アレイ 210 に相当する機能を有する。

[0108] イメージセンサ制御部 222 は、第一の実施形態によるイメージセンサ制御部 220 と同様に、画像縮小処理部 2201、および、画像出力部 2202 としての機能を有する。さらに、第三の実施形態によるイメージセンサ制御部 222 は、画像認識処理部 2205 としての機能を新たに有する。

[0109] 画像認識処理部 2205 は、情報処理装置 12 の音認識処理部 1506 により認識処理を開始する指示が出力されると、画像出力部 2202 から出力される情報処理装置 12 の周辺環境の画像データを解析することにより、画像認識を行う機能を有する。例えば、画像出力部 2202 は、機械学習により予め学習済みのモデルを用いて、情報処理装置 12 に接近する障害物と推定される移動体を認識してもよい。画像認識処理部 2205 は、画像の解析結果（画像認識処理の結果）を後述する周辺環境推定部 1507 に出力する。

[0110] （制御部 152）

図 12 に戻って、情報処理装置 12 の構成例の説明を続ける。制御部 152 は、情報処理装置 12 の動作全般を制御する機能を有する。このような情報処理装置 12 は、音差分処理部 1501、および、音認識処理部 1506

としての機能を有する。音差分処理部 1501 は、上記で図 2 を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0111] 音認識処理部 1506 は、情報処理装置 12 の周辺の音を解析することにより、音認識を行う機能を有する。例えば、音認識処理部 1506 は、機械学習により予め学習済みのモデルを用いて、情報処理装置 12 の周辺の音から、情報処理装置 12 に接近する障害物と推定される移動体を認識してもよい。音認識処理部 1506 は、音の解析結果（音認識処理の結果）を周辺環境推定部 1507 に出力する。

[0112] 周辺環境推定部 1507 は、画像出力部 2202 による情報処理装置 12 の周辺環境の画像データの解析結果と、音認識処理部 1506 による情報処理装置 12 の周辺の音データの解析結果に基づいて、情報処理装置 12 の周辺環境の状況を推定する、周辺環境推定処理を行う機能を有する。より詳細には、周辺環境推定部 1507 は、情報処理装置 12 に障害物と推定される移動物体が接近していることを推定してもよい。このような周辺環境推定部 1507 は、情報処理装置 12 の認識部の一例である。

[0113] 例えば、周辺環境推定部 1507 は、人または車などの、音を発する移動物体が情報処理装置 12 に接近していることを推定してもよい。

[0114] 周辺環境推定部 1507 は、音認識処理部 1506 による音の解析結果と、画像認識処理部 2205 による画像解析結果を入力データとするニューラルネットワークを用いて、上記音の解析結果、上記画像解析結果、および、周辺環境の状況を示す正解データの関係を学習してもよい。周辺環境推定部 1507 は、学習済みのモデルを用いて、周辺環境の状況を推定してもよい。

[0115] または、周辺環境の状況の推定処理として、周辺環境推定部 1507 は、音認識処理部 1506 による音の解析結果と、画像認識処理部 2205 による画像解析結果に応じて、ルールベースの推定処理を行ってもよい。例えば、周辺環境推定部 1507 は、音の解析結果および画像解析結果のいずれもが同一の周辺環境の状況を示す認識結果である場合に、当該認識結果を情報

処理装置 12 の周辺環境の状況の推定結果としてもよい。

[0116] また、周辺環境推定部 1507 は、周辺環境推定処理の結果に応じて、スピーカ 130 から出力されるコンテンツを制御する機能を有する。例えば、周辺環境推定部 1507 は、スピーカ 130 から、障害物の接近をユーザに知らせる通知メッセージを出力する制御を行ってもよい。または、周辺環境推定部 1507 は、スピーカ 130 から出力されるコンテンツ再生音の一時停止、コンテンツ再生音の音量の低減、または、コンテンツ再生音の音像定位の変化の、いずれか一つ以上の制御を行ってもよい。

[0117] 以上、図 12 および図 13 を参照して、第三の実施形態による情報処理装置 12 の構成例を説明した。

[0118] <4-2. 動作>

図 14 は、第三の実施形態による情報処理装置 12 の動作例を説明するためのフローチャートである。なお、図 14 において、S101、S103、S105、および S107 は、上記で図 6 を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0119] S101～S105 の処理が行われたのち、情報処理装置 12 の音差分処理部 1501 により、算出された差分値が、所定の閾値を超えている場合（S507／YES）、制御部 151 は、画像認識処理部 2205 および音認識処理部 1506 へ、認識処理を開始する指示を出力する（S507）。次いで、音認識処理部 1506 が、音認識処理を行う（S511）。

[0120] 次いで、周辺環境推定部 1507 は、イメージセンサ部 112 の画像認識処理部 2205 による画像認識処理結果を取得する（S513）。周辺環境推定部 1507 は、周辺環境推定処理を行う（S515）。

[0121] 周辺環境推定部 1507 は、推定した周辺環境の状況から、外部音取り込み機能を起動するか否かを決定する（S517）。周辺環境推定部 1507 は、外部音取り込み機能を起動すると決定しなかった場合（S517／NO）、周辺環境推定部 1507 は、外部音取り込み機能を OFF にするための制御信号を生成する（S525）。また、周辺環境推定部 1507 は、スピ

ーカ 130 に、音楽等のコンテンツを再生させる（S527）。

[0122] 周辺環境推定部 1507 は、外部音取り込み機能を起動すると決定した場合（S517／YES）、周辺環境推定部 1507 は、外部音取り込み機能を ON にする（S519）。また、周辺環境推定部 1507 は、ユーザに対する音声メッセージをスピーカ 130 に再生させる（S521）。

[0123] 次いで、情報処理装置 12 の一連の処理を終了する終了操作が行われた場合（S523／YES）、情報処理装置 12 は処理を終了する。12 は、一連の処理を終了する終了操作が行われるまで、S101～S523 の処理を繰り返す（S523／NO）。

[0124] 以上、図 14 を参照して、第三の実施形態による情報処理装置 12 の動作例を説明した。続いて、図 15 を参照して、第三の実施形態によるイメージセンサ部 122 の動作例を説明する。

[0125] 図 15 は、第三の実施形態によるイメージセンサ部 122 の動作例を説明するためのフローチャートである。なお、図 15 において、S201～S207 は、上記で図 7 を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0126] S207 の処理に次いで、イメージセンサ部 122 の画像認識処理部 2205 は、画像認識処理を行う（S609）。画像認識処理部 2205 は、画像認識処理結果を、周辺環境推定部 1507 に出力する（S611）。

[0127] 次いで、イメージセンサ部 112 の一連の処理を終了する終了操作が行われた場合（S613／YES）、イメージセンサ部 112 は処理を終了する。情報処理装置 12 は、一連の処理を終了する終了操作が行われるまで、S201～S613 の処理を繰り返す（S613／NO）。

[0128] 以上、図 15 を参照して、第三の実施形態によるイメージセンサ部 122 の動作例を説明した。続いて、図 16～図 22 を参照して、本開示の第四の実施形態による情報処理システムについて説明する。

[0129] <5. 第四の実施形態>

第四の実施形態では、本開示の適用先が、HMD である例を説明する。図

16は、第四の実施形態の概要を説明するための説明図である。

[0130] 図16に示した情報処理装置13は、ユーザU1の頭部に装着されているHMDである。画面Dは、情報処理装置13のディスプレイに表示される画面例を示す。第四の実施形態による情報処理装置13は、ユーザU1の周辺環境の状況の認識結果に応じて、情報処理装置13のコンテンツ再生音およびコンテンツ画像の表示を一時停止する。このとき、情報処理装置13は、情報処理装置13の周辺の音を出力する。また、情報処理装置13は、情報処理装置13の周辺の画像データをディスプレイに表示させる。

[0131] 図16の左側に示したように、情報処理装置13は、VR (Virtual Reality) 技術、またはAR (Augmented Reality) 技術を利用した、3次元の仮想空間の画像である画面D1を表示することが出来る。また、情報処理装置13は、画面D1の表示と同期されたコンテンツ音を出力する。

[0132] ユーザU1は、情報処理装置13を装着してVRまたはARのコンテンツを体験している間、目と耳がふさがれているような状態になるため、ユーザU1の周囲の状況を把握しづらい。情報処理装置13を実現するHMDによっては、ユーザの操作に基づき、HMDの周囲を撮像する外向きカメラで取得された周囲の映像を表示する機能を備えている場合があるが、周囲の映像が表示されるためにはユーザによる能動的な動作が求められる。このため、情報処理装置13を装着してVRまたはARのコンテンツを体験している間、ユーザU1は、他のユーザからの自身への話しかけ、または、他のユーザまたはペット等の接近などの周囲の状況を察知し難い。そのため、ユーザU1が、他のユーザおよびペット等に衝突してしまうなどの状況が発生し得る。

[0133] そこで、本開示の第四の実施形態による情報処理装置13によれば、情報処理装置13の周辺環境の状況として、ユーザU1と他のユーザとの間での話しかけまたは会話の有無、および、ユーザU1の周囲の人またはペット等の接近が認識される。情報処理装置13は、認識結果に応じて、情報処理装

置 1 3 に出力されるコンテンツを制御する。これにより、情報処理装置 1 3 を装着しているユーザ U 1 の利便性が向上する。

[0134] <5-1. 構成例>

図 1 7 は、第四の実施形態による情報処理装置 1 3 の構成例を説明するためのブロック図である。図 1 7 に示したように、情報処理装置 1 3 は、イメージセンサ部 1 1 0、マイクロフォン 1 2 0、スピーカ 1 3 0、記憶部 1 4 0、制御部 1 5 0、および、映像出力部 1 6 2 を有する。なお、イメージセンサ部 1 1 0、マイクロフォン 1 2 0、スピーカ 1 3 0、記憶部 1 4 0、および制御部 1 5 0 は、上記で図 2 を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0135] (映像出力部 1 6 2)

映像出力部 1 6 2 は、ユーザの左右の眼にそれぞれ固定された左右の画面を備え、左眼用画像および右眼用画像を表示する機能を有する。映像出力部 1 6 2 の画面は、例えば液晶ディスプレイ (LCD: Liquid Crystal Display)、有機 EL (Electro Luminescence) ディスプレイなどの表示パネル、または、網膜直描ディスプレイなどのレーザー走査方式ディスプレイで構成される。また、映像出力部 1 6 2 は、表示画面を拡大投影して、ユーザの瞳に所定の画角からなる拡大虚像を結像する結像光学系を備えてもよい。

[0136] 映像出力部 1 6 2 は、制御部 1 5 0 の制御に従って、イメージセンサ部 1 1 0 により取得される情報処理装置 1 3 の周辺環境の画像データを表示する。

[0137] なお、第四の実施形態によるイメージセンサ部 1 1 0 は、情報処理装置 1 3 の周囲 3 6 0 度方向の画像を取得できる位置に配置される。また、情報処理装置 1 3 は、複数のイメージセンサ部 1 1 0 を備えていてよい。

[0138] 図 1 8 は、第四の実施形態によるイメージセンサ部 1 1 0 の設置位置の一例を説明するための説明図である。図 1 8 の外観図 V 1 は、情報処理装置 1 3 を斜め後方からみた斜視図である。イメージセンサ部 1 1 0 は、例えば、

情報処理装置 13 A の後方部分の位置 P 1 に設置されてもよい。この場合、イメージセンサ部 110 は、魚眼レンズを備え、情報処理装置 13 A を装着するユーザ U 1 の後方全体を撮像してもよい。

[0139] また、図 19 は、第四の実施形態によるイメージセンサ部 110 の設置位置の他の一例を説明するための説明図である。図 19 の外観図 V 2 は、情報処理装置 13 A を前方からみた正面図である。イメージセンサ部 110 は、位置 P 1 に加え、位置 P 2 の位置に設置されてもよい。この場合、イメージセンサ部 110 は、魚眼レンズを含んで構成され、情報処理装置 13 A を装着するユーザ U 1 の前方全体を撮像してもよい。

[0140] また、図 20 は、第四の実施形態によるイメージセンサ部 110 の設置位置の他の一例を説明するための説明図である。図 20 の外観図 V 3 は、情報処理装置 13 B 前方からみた正面図である。図 20 に示した例では、情報処理装置 13 B は、情報処理装置 13 B の前面の四方の位置 P 4 ～位置 P 6 の各々に 1 ずつ設置されてもよい。この場合、情報処理装置 13 B は、計 4 つのイメージセンサ部 110 を備えていてもよい。これにより、情報処理装置 13 A を装着するユーザ U 1 の前方全体が撮像され得る。

[0141] 以上、図 17 ～図 20 を参照して、第四の実施形態による情報処理装置 13 の構成例を説明した。続いて、図 21 および図 22 を参照して、第四の実施形態による情報処理装置 13 の動作例を説明する。

[0142] <5-2. 動作>

図 21 は、第四の実施形態による情報処理装置 13 の動作例を説明するためのフローチャートである。なお、図 21 において、S101～S107 の処理は、上記で図 6 を参照して説明した通りであるので、ここでの重複する説明を省略する。

[0143] まず、S101～S107 の処理が行われる。S107 において、情報処理装置 13 の制御部 150 により認識処理の開始指示が出力されると、情報処理装置 13 の認識部 1503 は、周辺環境の状況の認識処理を行う（S711）。例えば、認識部 1503 は、周辺環境の状況として、他のユーザま

たはペット等が情報処理装置 13 に接近していることを認識してもよい。さらに、認識部 1503 は、認識結果に基づいて、情報処理装置 13 のスピーカ 130 および映像出力部 162 のそれぞれに、外部音取り込みとコンテンツ音声の停止、および、周辺映像の表示を行わせる制御が必要か否かを推定してもよい（S713）。

[0144] S713 において、上記制御が必要と判定されなかった場合（S713／NO）、出力制御部 1505 は、外部音取り込み機能を OFF にするための制御信号を生成する（S723）。また、出力制御部 1505 は、映像出力部 162 に、仮想空間の画像等のコンテンツ映像を表示させる（S725）。さらに、出力制御部 1505 は、スピーカ 130 に、コンテンツ音声を再生させる（S727）。次いで、情報処理装置 13 は、S101 に戻って、S101～S727 の処理を繰り返す。

[0145] S713 において、上記制御が必要と判定された場合（S713／YES）、出力制御部 1505 は、外部音取り込み機能を ON にするための制御信号を生成する（S715）。また、出力制御部 1505 は、映像出力部 162 に、情報処理装置 13 の周辺環境の映像を表示させる（S717）。さらに、出力制御部 1505 は、スピーカ 130 に、コンテンツ音声の再生を停止させる（S719）。

[0146] 次いで、情報処理装置 13 の一連の処理を終了する終了操作が行われた場合（S721／YES）、情報処理装置 13 は処理を終了する。情報処理装置 13 は、一連の処理を終了する終了操作が行われるまで、S101～S727 の処理を繰り返す（S721／NO）。

[0147] 以上、図 21 を参照して、第四の実施形態による情報処理装置 13 の動作例を説明した。続いて、図 22 を参照して、第四の実施形態によるイメージセンサ部 110 の動作例を説明する。

[0148] 図 22 は、第四の実施形態によるイメージセンサ部 110 の動作例を説明するためのフローチャートである。なお、図 15 において、S201～S207 は、上記で図 7 を参照して説明した通りであるので、ここでの重複する

説明を省略する。

[0149] S 2 0 7 の処理に次いで、イメージセンサ部 1 1 0 のイメージセンサ制御部 2 2 1 は、情報処理装置 1 3 の認識部 1 5 0 3 により、周辺映像の表示が必要と推定されたか否かの判定を行う (S 8 0 8)。周辺映像の表示が必要と推定されなかった場合 (S 8 0 8 / N O)、情報処理装置 1 3 は、S 2 0 1 ~ S 8 0 8 の処理を繰り返す。

[0150] 周辺映像の表示が必要と推定された場合 (S 8 0 8 / Y E S)、画像出力部 2 2 0 2 は、イメージセンサ制御部 2 2 1 の制御に従って、画素アレイ 2 1 0 により取得された情報処理装置 1 3 の周辺環境の画像データを映像出力部 1 6 2 に出力する (S 8 0 9)。

[0151] 次いで、イメージセンサ部 1 1 1 の一連の処理を終了する終了操作が行われた場合 (S 8 1 1 / Y E S)、イメージセンサ部 1 1 1 は処理を終了する。イメージセンサ部 1 1 1 は、一連の処理を終了する終了操作が行われるまで、S 2 0 1 ~ S 8 1 1 の処理を繰り返す (S 8 1 1 / N O)。

[0152] 以上、図 2 2 を参照して、第四の実施形態によるイメージセンサ部 1 1 0 の動作例を説明した。

[0153] < 6. ハードウェア構成 >

以上、本開示の実施形態を説明した。上述した、情報処理装置 1 0 の周辺環境の変化が所定の条件を満たすか否かの判断結果の取得、情報処理装置 1 0 の周辺の音および周辺環境の画像データに基づく周辺環境の状況の認識処理、認識結果に基づくコンテンツ出力の制御、などの情報処理は、ソフトウェアと、ハードウェアとの協働により実現される。以下、情報処理装置 1 0 に適用され得るハードウェア構成例を説明する。

[0154] 図 2 3 は、ハードウェア構成 9 0 の一例を示したブロック図である。なお、以下に説明するハードウェア構成 9 0 のハードウェア構成例は、情報処理装置 1 0 のハードウェア構成の一例に過ぎない。したがって、情報処理装置 1 0 は、それぞれ、必ずしも図 2 3 に示したハードウェア構成の全部を有している必要はない。また、情報処理装置 1 0 の中に、図 2 3 に示したハード

ウェア構成の一部が存在しなくてもよい。さらに、情報処理装置 11、情報処理装置 12、および情報処理装置 13 についても、以下に説明するハードウェア構成 90 が適用され得る。

[0155] 図 23 に示すように、ハードウェア構成 90 は、CPU 901、ROM (Read Only Memory) 903、および RAM 905 を含む。また、ハードウェア構成 90 は、ホストバス 907、ブリッジ 909、外部バス 911、インターフェース 913、入力装置 915、出力装置 917、ストレージ装置 919、ドライブ 921、接続ポート 923、通信装置 925 を含んでもよい。ハードウェア構成 90 は、CPU 901 に代えて、またはこれとともに、GPU (Graphics Processing Unit)、DSP (Digital Signal Processor) または ASIC (Application Specific Integrated Circuit) と呼ばれるような処理回路を有してもよい。

[0156] CPU 901 は、演算処理装置および制御装置として機能し、ROM 903、RAM 905、ストレージ装置 919、またはリムーバブル記録媒体 927 に記録された各種プログラムに従って、ハードウェア構成 90 内の動作全般またはその一部を制御する。ROM 903 は、CPU 901 が使用するプログラムおよび演算パラメータなどを記憶する。RAM 905 は、CPU 901 の実行において使用するプログラムまたは／および、その実行において適宜変化するパラメータなどを一時的に記憶する。CPU 901、ROM 903、および RAM 905 は、CPU バスなどの内部バスにより構成されるホストバス 907 により相互に接続されている。さらに、ホストバス 907 は、ブリッジ 909 を介して、PCI (Peripheral Component Interconnect / Interface) バスなどの外部バス 911 に接続されている。

[0157] CPU 901 が ROM 903、RAM 905、およびソフトウェアと協働することにより、例えば、制御部 150、および、イメージセンサ制御部 220 の機能が実現され得る。

- [0158] 入力装置 915 は、例えば、ボタンなど、ユーザによって操作される装置である。入力装置 915 は、マウス、キーボード、タッチパネル、スイッチおよびレバーなどを含んでもよい。また、入力装置 915 は、ユーザの音声を検出するマイクロフォンを含んでもよい。入力装置 915 は、例えば、赤外線またはその他の電波を利用したリモートコントロール装置であってもよいし、ハードウェア構成 90 の操作に対応した携帯電話などの外部接続機器 929 であってもよい。入力装置 915 は、ユーザが入力した情報に基づいて入力信号を生成して CPU 901 に出力する入力制御回路を含む。ユーザは、この入力装置 915 を操作することによって、ハードウェア構成 90 に対して各種のデータを入力したり処理動作を指示したりする。
- [0159] また、入力装置 915 は、撮像装置、およびセンサを含んでもよい。撮像装置は、例えば、CCD (Charge Coupled Device) または CMOS (Complementary Metal Oxide Semiconductor) などの撮像素子、および撮像素子への被写体像の結像を制御するためのレンズなどの各種の部材を用いて実空間を撮像し、撮像画像を生成する装置である。撮像装置は、静止画を撮像するものであってもよいし、また動画を撮像するものであってもよい。
- [0160] センサは、例えば、測距センサ、加速度センサ、ジャイロセンサ、地磁気センサ、振動センサ、光センサ、音センサなどの各種のセンサである。センサは、例えばハードウェア構成 90 の筐体の姿勢など、ハードウェア構成 90 自体の状態に関する情報、または、ハードウェア構成 90 の周辺の明るさまたは騒音など、ハードウェア構成 90 の周辺環境に関する情報を取得する。また、センサは、GPS (Global Positioning System) 信号を受信して装置の緯度、経度および高度を測定する GPS センサを含んでもよい。
- [0161] 出力装置 917 は、取得した情報をユーザに対して視覚的または聴覚的に通知することが可能な装置で構成される。出力装置 917 は、例えば、LCD (Liquid Crystal Display)、有機 EL ディスプレイ

レイなどの表示装置、スピーカおよびヘッドホンなどの音出力装置などであり得る。また、出力装置 917 は、PDP (Plasma Display Panel)、プロジェクタ、ホログラム、プリンタ装置などを含んでもよい。出力装置 917 は、ハードウェア構成 90 の処理により得られた結果を、テキストまたは画像などの映像として出力したり、音声または音響などの音として出力したりする。また、出力装置 917 は、周囲を明るくする照明装置などを含んでもよい。

[0162] ストレージ装置 919 は、ハードウェア構成 90 の記憶部の一例として構成されたデータ格納用の装置である。ストレージ装置 919 は、例えば、HDD (Hard Disk Drive) などの磁気記憶デバイス、半導体記憶デバイス、光記憶デバイス、または光磁気記憶デバイスなどにより構成される。このストレージ装置 919 は、CPU 901 が実行するプログラムまたは各種データ、および外部から取得した各種のデータなどを格納する。

[0163] ドライブ 921 は、磁気ディスク、光ディスク、光磁気ディスク、または半導体メモリなどのリムーバブル記録媒体 927 のためのリーダライタであり、ハードウェア構成 90 に内蔵、あるいは外付けされる。ドライブ 921 は、装着されているリムーバブル記録媒体 927 に記録されている情報を読み出して、RAM 905 に出力する。また、ドライブ 921 は、装着されているリムーバブル記録媒体 927 に記録を書き込む。

[0164] 接続ポート 923 は、機器をハードウェア構成 90 に直接接続するためのポートである。接続ポート 923 は、例えば、USB (Universal Serial Bus) ポート、IEEE1394 ポート、SCSI (Small Computer System Interface) ポートなどであり得る。また、接続ポート 923 は、RS-232C ポート、光オーディオ端子、HDMI (登録商標) (High-Definition Multimedia Interface) ポートなどであってもよい。接続ポート 923 に外部接続機器 929 を接続することで、ハードウェア構成 90 と外部接続機器 929 との間で各種のデータが交換され得る。

[0165] 通信装置 925 は、例えば、ローカルネットワーク、または、無線通信の基地局との通信ネットワークに接続するための通信デバイスなどで構成された通信インターフェースである。通信装置 925 は、例えば、有線または無線 LAN (Local Area Network)、Bluetooth (登録商標)、Wi-Fi (登録商標)、または WUSB (Wireless USB) 用の通信カードなどであり得る。また、通信装置 925 は、光通信用のルータ、ADSL (Asymmetric Digital Subscriber Line) 用のルータ、または、各種通信用のモデムなどであってもよい。通信装置 925 は、例えば、インターネットまたは他の通信機器との間で、TCP/IP などの所定のプロトコルを用いて信号などを送受信する。また、通信装置 925 に接続されるローカルネットワークまたは基地局との通信ネットワークは、有線または無線によって接続されたネットワークであり、例えば、インターネット、家庭内 LAN、赤外線通信、ラジオ波通信または衛星通信などである。

[0166] <7. 補足>

以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0167] 例えば、上記実施形態では、判断結果取得部の一例としての音差分処理部 1501 が、判断結果取得処理として、周辺の音の変化が所定の閾値が示す差分以上に変化したことを検出した場合に、情報処理装置 10 の周辺環境の変化が所定の条件を満たすと判断するとした。より具体的には、音差分処理部 1501 が、現在の音データと過去の音データとの差分を算出し、算出された差分を示す差分値が所定の閾値を超える場合に、所定の条件が満たされたと判断するとした。しかし、本技術は係る例に限定されない。例えば、情

報処理装置 10 の制御部 150 が、イメージセンサ部 110 により取得された情報処理装置 10 の周辺環境の画像データを用いて、周辺環境の変化が所定の条件を満たすか否かの判断結果を取得してもよい。

[0168] 例えば、制御部 150 は、現在の周辺環境の画像データと、過去の周辺環境の画像データの差分を検出してもよい。制御部 150 は、画像データ上で差分が検出された領域が、所定の閾値が示す範囲を超える場合に、周辺環境の状況が所定の条件を満たすと判断してもよい。これにより、例えば情報処理装置 10 に他のユーザが接近している場合等の周辺環境の変化が検出され得る。

[0169] なお、上記のように、制御部 150 が周辺環境の画像データを用いて判断結果取得処理を行う場合には、情報処理装置 10 を装着しているユーザの頭部の動きによる周辺環境の画像データの変化を考慮することが望ましい。例えば、情報処理装置 10 は、情報処理装置 10 の加速度および角速度を検出することが可能な加速度センサおよび角速度センサ、または IMU を備えていてもよい。制御部 150 は、情報処理装置 10 の加速度および角速度に基づきユーザの頭部の位置および姿勢を算出するヘッドトラッキングを行ってもよい。制御部 150 は、ヘッドトラッキングの結果に基づいて、判断結果取得処理によって検出される周辺環境の画像データの変化がユーザの頭部の動きによって生じたものである場合には、周辺環境の変化が所定の条件を満たさないと判断してもよい。これにより、イメージセンサ部 110 により取得される周辺環境の画像データによっても、情報処理装置 10 の周辺環境の変化が所定の条件を満たすか否かの判断が行われ得る。

[0170] また、上記のように制御部 150 が周辺環境の画像データを用いて判断結果取得処理を行う場合には、画像縮小処理部 2201 および認識部 1503 は次のような処理を行ってもよい。具体的には、画像縮小処理部 2201 は、制御部 150 から認識処理を開始する指示が出力された場合、画素アレイ 210 により取得される周辺環境の画像データから、制御部 150 により上記差分が検出された領域を切り出す処理を行ってもよい。さらに、画像縮小

処理部 2201 は、切り出された領域の画像データを所定の解像度に縮小する処理を行ってもよい。認識部 1503 は、縮小処理が行われた、切り出された領域の画像データを入力データとして、認識処理を行ってもよい。これにより、認識部 1503 の認識処理において、情報処理装置 10 の周辺のユーザと他のユーザの間での話しかけまたは会話の有無の検出精度が向上し得る。

[0171] また、上記実施形態では、イメージセンサ部 110 およびマイクロフォン 120 が、情報処理装置 10 と同一体として構成される例を説明した。しかし、本技術は係る例に限定されない。例えば、イメージセンサ部 110 およびマイクロフォン 120 は、情報処理装置 10 と別体として構成されてもよい。この場合、イメージセンサ部 110 は、ユーザの周辺環境を俯瞰的に撮像可能な位置に設置されることが望ましい。

[0172] また、上記ではユーザと他のユーザが同一空間に存在することが想定されたが、ユーザと他のユーザは別空間に存在していてもよい。例えば、ユーザが存在する空間には、他のユーザの映像および音声を出力する出力装置が配置されており、情報処理装置 10 は、当該出力装置を介する話しかけ、および会話の有無を認識してもよい。当該出力装置は、据え置き式の装置であってもよいし、他のユーザによる遠隔操作に従って移動可能な装置であってもよい。

[0173] また、本実施形態による情報処理装置 10、およびイメージセンサ部 111 の動作の処理におけるステップは、必ずしも説明図として記載された順序に沿って時系列に処理する必要はない。例えば、情報処理装置 10 およびイメージセンサ部 111 の動作の処理における各ステップは、説明図として記載した順序と異なる順序で処理されてもよく、並列的に処理されてもよい。

[0174] また、上述した情報処理装置 10 およびイメージセンサ部 111 に内蔵される CPU、ROM および RAM などのハードウェアに、本実施形態による情報処理システムの機能を発揮させるための 1 以上のコンピュータプログラムも作成可能である。また、当該 1 以上のコンピュータプログラムを記憶さ

せた、コンピュータが読み取り可能な記憶媒体も提供される。

[0175] また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

[0176] なお、以下のような構成も本開示の技術的範囲に属する。

(1)

センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、

前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、

を備える、情報処理装置。

(2)

前記センサ部は、当該センサ部の周辺の音を検出し、

前記判断結果取得部は、前記音の変化が前記所定の条件を満たすか否かの判断結果を取得する、

前記(1)に記載の情報処理装置。

(3)

前記判断結果取得部は、前記センサ部により所定時間長を有する第一の時間区間において取得された第一の音データと、前記センサ部により前記第一の時間区間と異なり前記所定時間長を有する第二の時間区間において取得された第二の音データと、に基づいて、前記音の変化が前記所定の条件を満たすか否かの判断結果を取得する、

前記(2)に記載の情報処理装置。

(4)

前記判断結果取得部は、前記第一の音データと、前記第二の音データとの

差分を算出し、

前記所定の条件は、算出された前記差分を示す差分値が、所定の閾値を超えることである、

前記（３）に記載の情報処理装置。

（５）

前記認識部は、前記センシングデータとして、前記センサ部から前記音データを取得し、

前記音データを解析して前記周辺環境の状況を認識する処理を行う、

前記（１）～（４）のいずれか一項に記載の情報処理装置。

（６）

前記センサ部は、前記センシングデータとして、当該センサ部の周辺環境の画像データを取得し、

前記認識部は、前記音データおよび前記画像データを解析して前記周辺環境の状況を認識する処理を行う、

前記（１）～（５）のいずれか一項に記載の情報処理装置。

（７）

前記認識部は、前記音データおよび前記画像データの各々の解析結果に基づき、前記周辺環境の状況として、前記センサ部の周辺に位置するユーザと他のユーザの間での話しかけまたは会話の有無を認識する、

前記（６）に記載の情報処理装置。

（８）

前記認識部は、前記センサ部により取得される前記音データまたは前記センサ部により取得される前記画像データの少なくともいずれかまたは両方と、前記センサ部の周辺に位置するユーザと他のユーザとの間での話しかけの有無の正解データと、を入力データとするニューラルネットワークを用いて、前記音データまたは前記画像データの少なくともいずれかまたは両方および前記ユーザと前記他のユーザとの間での話しかけまたは会話の有無の関係を学習し、

学習の結果得られたモデルを用いて、前記センサ部の周辺に位置するユーザと他のユーザとの間での話しかけまたは会話の有無を認識する、

前記（６）または（７）に記載の情報処理装置。

（９）

前記認識部により前記ユーザと前記他のユーザとの間での話しかけまたは会話が有ると認識されると、音出力部により出力されるコンテンツ再生音を一時停止させる処理を行う、出力制御部をさらに備える、前記（７）または（８）に記載の情報処理装置。

（１０）

前記出力制御部は、前記認識部により前記ユーザと前記他のユーザとの間での話しかけまたは会話が有ると認識されると、前記センサ部により取得される前記センサ部の周辺の前記音データを前記音出力部に出力させる、

前記（７）～（９）のいずれか一項に記載の情報処理装置。

（１１）

前記認識部により前記ユーザと前記他のユーザとの間での話しかけまたは会話が有ると認識されると、音出力部により出力されるコンテンツ再生音の音量を低減させる処理、または、前記コンテンツ再生音の音像定位を変化させる処理、のいずれかまたは両方を行う、出力制御部をさらに備える、前記（７）または（８）に記載の情報処理装置。

（１２）

前記認識部は、前記周辺環境の状況として、前記センサ部に接近する障害物の有無を認識する、

前記（６）に記載の情報処理装置。

（１３）

前記認識部により前記センサ部に接近する障害物が有ると認識されると、音出力部により出力されるコンテンツ再生音を一時停止させるとともに、前記音出力部にユーザへの通知音を出力させる制御を行う出力制御部をさらに備える、

前記（１２）に記載の情報処理装置。

（１４）

前記出力制御部は、前記認識部により前記センサ部に接近する障害物が有ると認識されると、前記センサ部により取得される前記センサ部の周辺の前記音データを前記音出力部に出力させる、

前記（１３）に記載の情報処理装置。

（１５）

前記出力制御部は、前記認識部により前記センサ部に接近する障害物が有ると認識されると、画像出力部により出力されるコンテンツ画像の表示を一時停止させるとともに、前記センサ部により取得される当該センサ部の周辺の画像データを前記画像出力部に表示させる、

前記（１４）に記載の情報処理装置。

（１６）

前記センサ部により取得された前記第二の音データを記憶する記憶部をさらに備える、前記（３）～（１５）のいずれか一項に記載の情報処理装置。

（１７）

コンピュータを、

センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、

前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、

として機能させるための、プログラム。

（１８）

センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得することと、

前記所定の条件が満たされたことを示す判断結果が取得されたことに基づ

き、前記センサ部の前記周辺環境の状況を認識する認識処理を開始することと、

を含む、コンピュータにより実行される情報処理方法。

符号の説明

- [0177] 1 0 情報処理装置
 - 1 1 0 イメージセンサ部
 - 1 2 0 マイクロフォン
 - 1 3 0 スピーカ
 - 1 4 0 記憶部
 - 1 5 0 制御部
 - 1 5 0 1 音差分処理部
 - 1 5 0 3 認識部
 - 1 5 0 5 出力制御部
 - 1 5 0 6 音認識処理部
 - 1 5 0 7 周辺環境推定部
 - 2 1 0 画素アレイ
 - 2 2 0 イメージセンサ制御部
 - 2 2 0 1 画像縮小処理部
 - 2 2 0 2 画像出力部
 - 2 2 0 3 認識部
 - 2 2 0 4 出力制御部
 - 2 2 0 5 画像認識処理部

請求の範囲

- [請求項1] センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、
- 前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、
- を備える、情報処理装置。
- [請求項2] 前記センサ部は、当該センサ部の周辺の音を検出し、
- 前記判断結果取得部は、前記音の変化が前記所定の条件を満たすか否かの判断結果を取得する、
- 請求項 1 に記載の情報処理装置。
- [請求項3] 前記判断結果取得部は、前記センサ部により所定時間長を有する第一の時間区間において取得された第一の音データと、前記センサ部により前記第一の時間区間と異なり前記所定時間長を有する第二の時間区間において取得された第二の音データと、に基づいて、前記音の変化が前記所定の条件を満たすか否かの判断結果を取得する、
- 請求項 2 に記載の情報処理装置。
- [請求項4] 前記判断結果取得部は、前記第一の音データと、前記第二の音データとの差分を算出し、
- 前記所定の条件は、算出された前記差分を示す差分値が、所定の閾値を超えることである、
- 請求項 3 に記載の情報処理装置。
- [請求項5] 前記認識部は、前記センシングデータとして、前記センサ部から前記音データを取得し、
- 前記音データを解析して前記周辺環境の状況を認識する処理を行う、
- 請求項 4 に記載の情報処理装置。

- [請求項6] 前記センサ部は、前記センシングデータとして、当該センサ部の周辺環境の画像データを取得し、
- 前記認識部は、前記音データおよび前記画像データを解析して前記周辺環境の状況を認識する処理を行う、
- 請求項5に記載の情報処理装置。
- [請求項7] 前記認識部は、前記音データおよび前記画像データの各々の解析結果に基づき、前記周辺環境の状況として、前記センサ部の周辺に位置するユーザと他のユーザの間での話しかけまたは会話の有無を認識する、
- 請求項6に記載の情報処理装置。
- [請求項8] 前記認識部は、前記センサ部により取得される前記音データまたは前記センサ部により取得される前記画像データの少なくともいずれかまたは両方と、前記センサ部の周辺に位置するユーザと他のユーザとの間での話しかけの有無の正解データと、を入力データとするニューラルネットワークを用いて、前記音データまたは前記画像データの少なくともいずれかまたは両方および前記ユーザと前記他のユーザとの間での話しかけまたは会話の有無の関係を学習し、
- 学習の結果得られたモデルを用いて、前記センサ部の周辺に位置するユーザと他のユーザとの間での話しかけまたは会話の有無を認識する、
- 請求項7に記載の情報処理装置。
- [請求項9] 前記認識部により前記ユーザと前記他のユーザとの間での話しかけまたは会話があると認識されると、音出力部により出力されるコンテンツ再生音を一時停止させる処理を行う、出力制御部をさらに備える、請求項7に記載の情報処理装置。
- [請求項10] 前記出力制御部は、前記認識部により前記ユーザと前記他のユーザとの間での話しかけまたは会話があると認識されると、前記センサ部により取得される前記センサ部の周辺の前記音データを前記音出力部

に出力させる、

請求項 9 に記載の情報処理装置。

[請求項11] 前記認識部により前記ユーザと前記他のユーザとの間での話しかけまたは会話が有ると認識されると、音出力部により出力されるコンテンツ再生音の音量を低減させる処理、または、前記コンテンツ再生音の音像定位を変化させる処理、のいずれかまたは両方を行う、出力制御部をさらに備える、請求項 7 に記載の情報処理装置。

[請求項12] 前記認識部は、前記周辺環境の状況として、前記センサ部に接近する障害物の有無を認識する、

請求項 6 に記載の情報処理装置。

[請求項13] 前記認識部により前記センサ部に接近する障害物が有ると認識されると、音出力部により出力されるコンテンツ再生音を一時停止させるとともに、前記音出力部にユーザへの通知音を出力させる制御を行う出力制御部をさらに備える、

請求項 1 2 に記載の情報処理装置。

[請求項14] 前記出力制御部は、前記認識部により前記センサ部に接近する障害物が有ると認識されると、前記センサ部により取得される前記センサ部の周辺の前記音データを前記音出力部に出力させる、

請求項 1 3 に記載の情報処理装置。

[請求項15] 前記出力制御部は、前記認識部により前記センサ部に接近する障害物が有ると認識されると、画像出力部により出力されるコンテンツ画像の表示を一時停止させるとともに、前記センサ部により取得される当該センサ部の周辺の画像データを前記画像出力部に表示させる、

請求項 1 4 に記載の情報処理装置。

[請求項16] 前記センサ部により取得された前記第二の音データを記憶する記憶部をさらに備える、請求項 3 に記載の情報処理装置。

[請求項17] コンピュータを、

センサ部により取得されたセンシングデータが示す前記センサ部の

周辺環境の変化が所定の条件を満たすか否かの判断結果を取得する判断結果取得部と、

前記判断結果取得部により前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始する認識部と、
として機能させるための、プログラム。

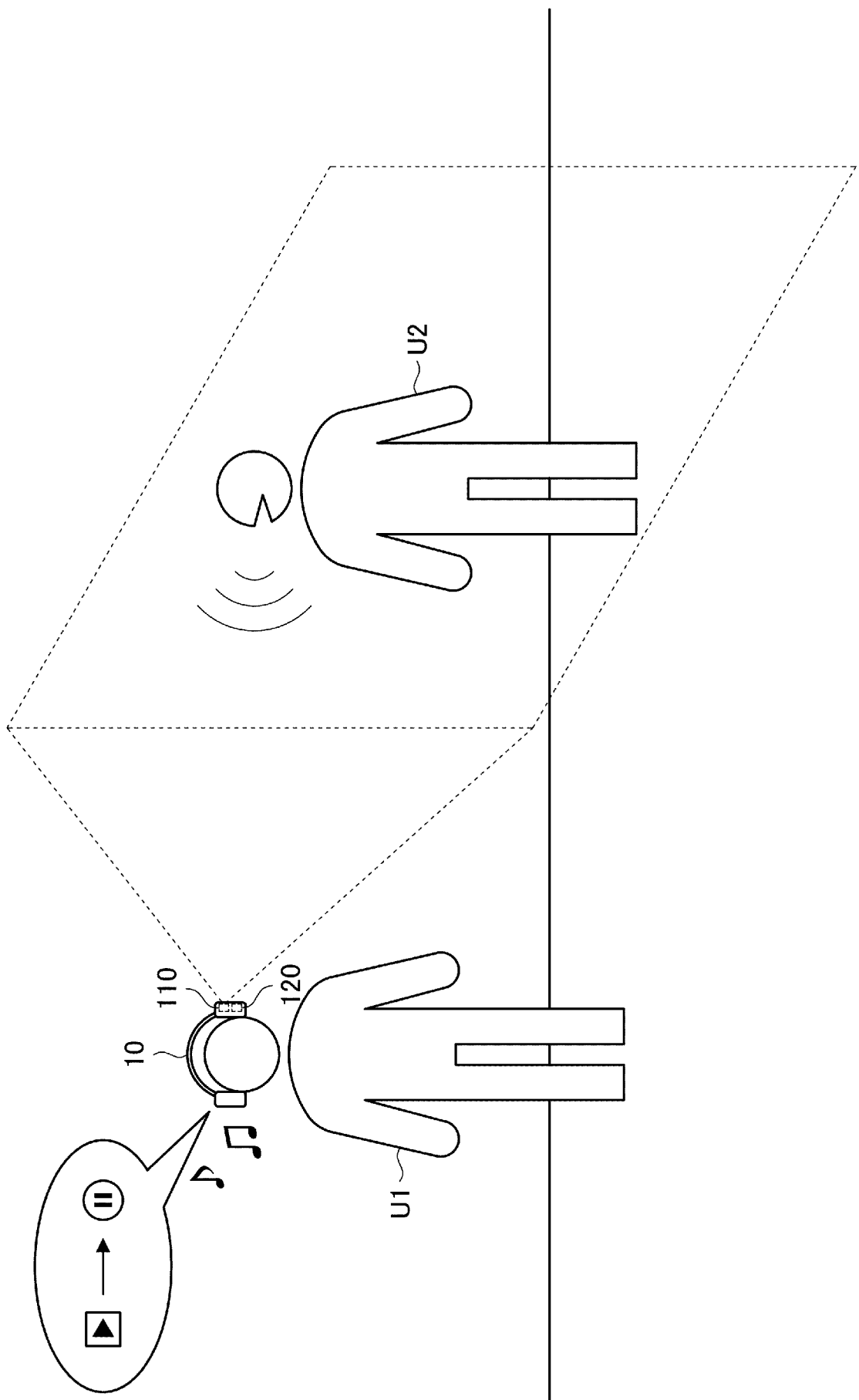
[請求項18]

センサ部により取得されたセンシングデータが示す前記センサ部の周辺環境の変化が所定の条件を満たすか否かの判断結果を取得することと、

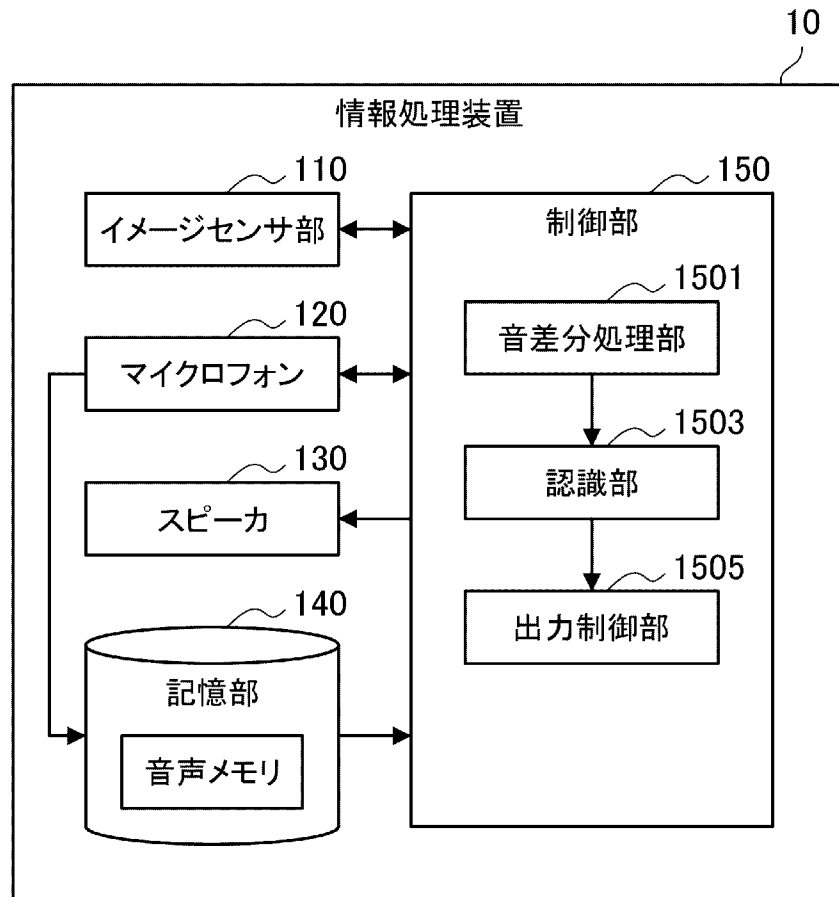
前記所定の条件が満たされたことを示す判断結果が取得されたことに基づき、前記センサ部の前記周辺環境の状況を認識する認識処理を開始することと、

を含む、コンピュータにより実行される情報処理方法。

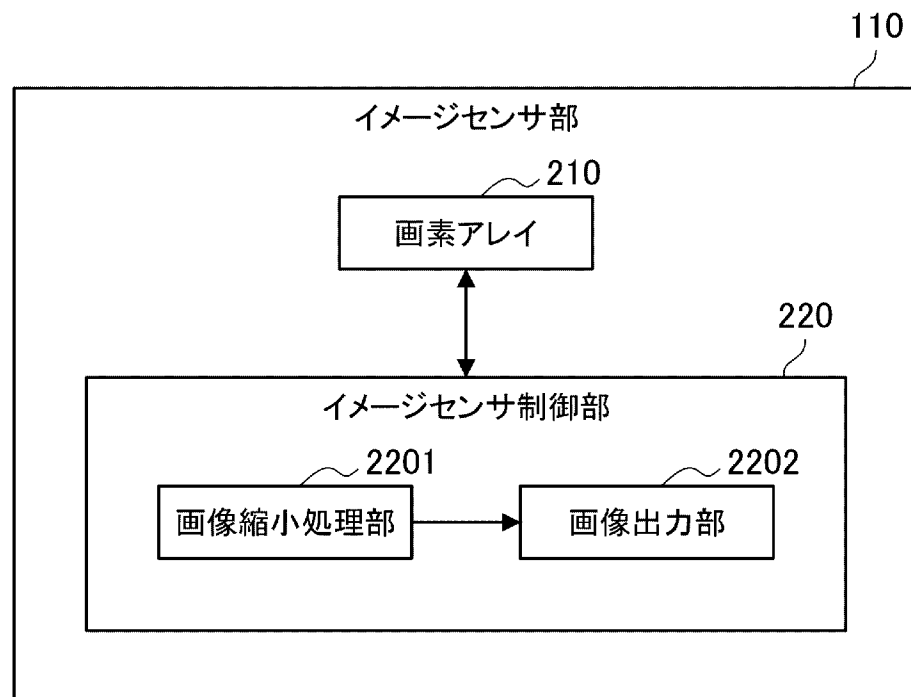
[図1]



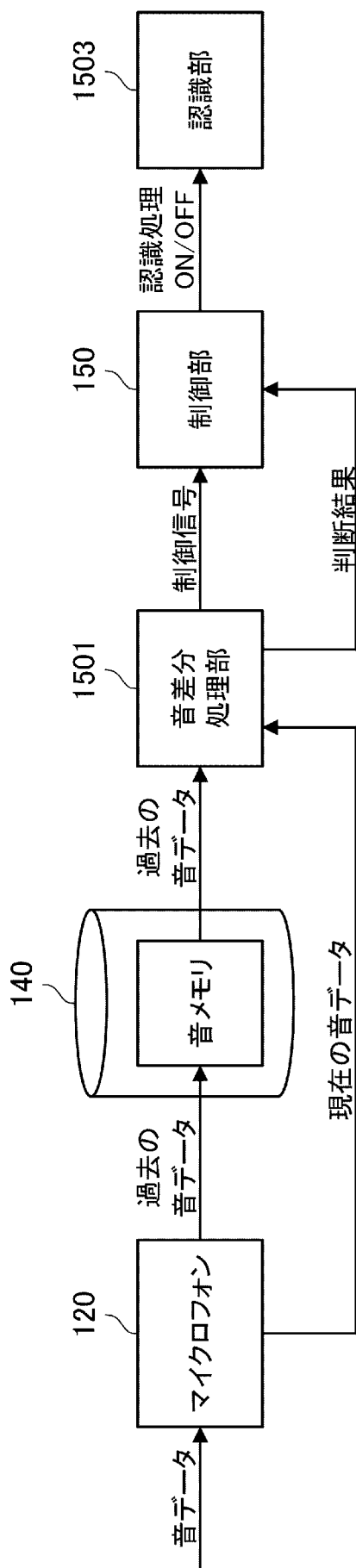
[図2]



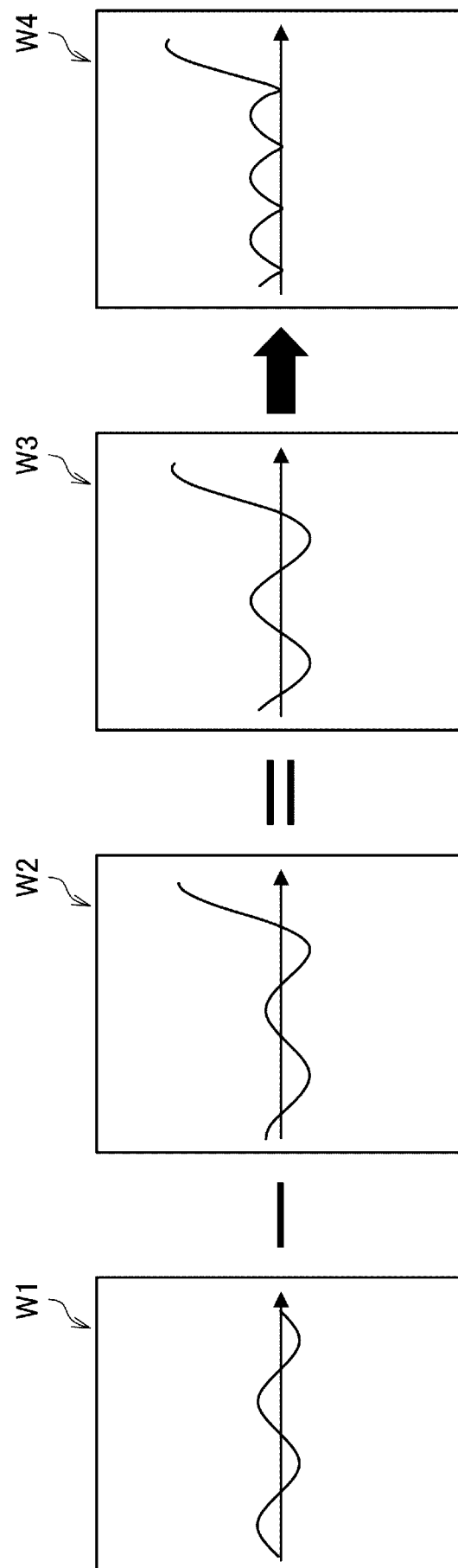
[図3]



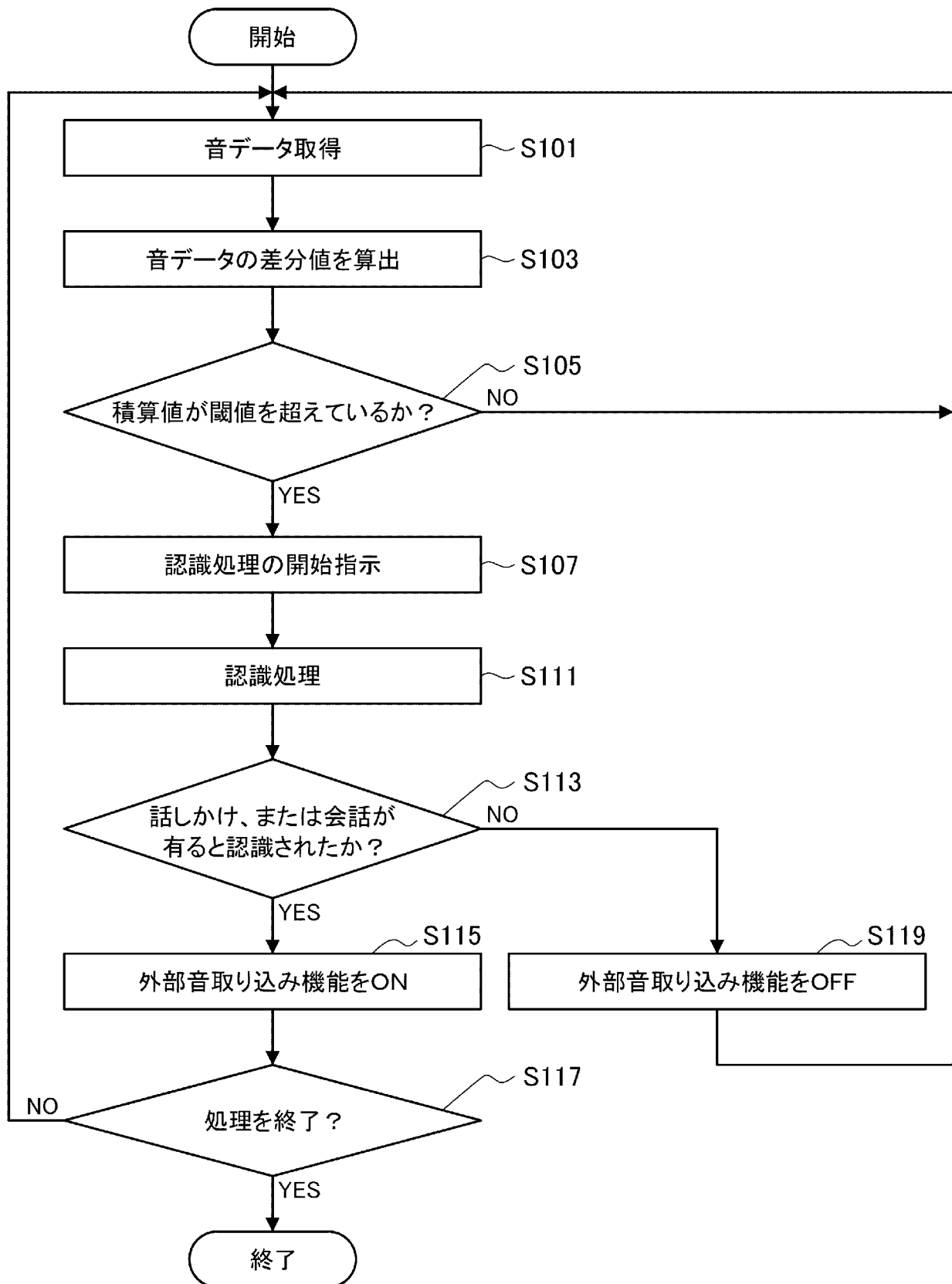
[図4]



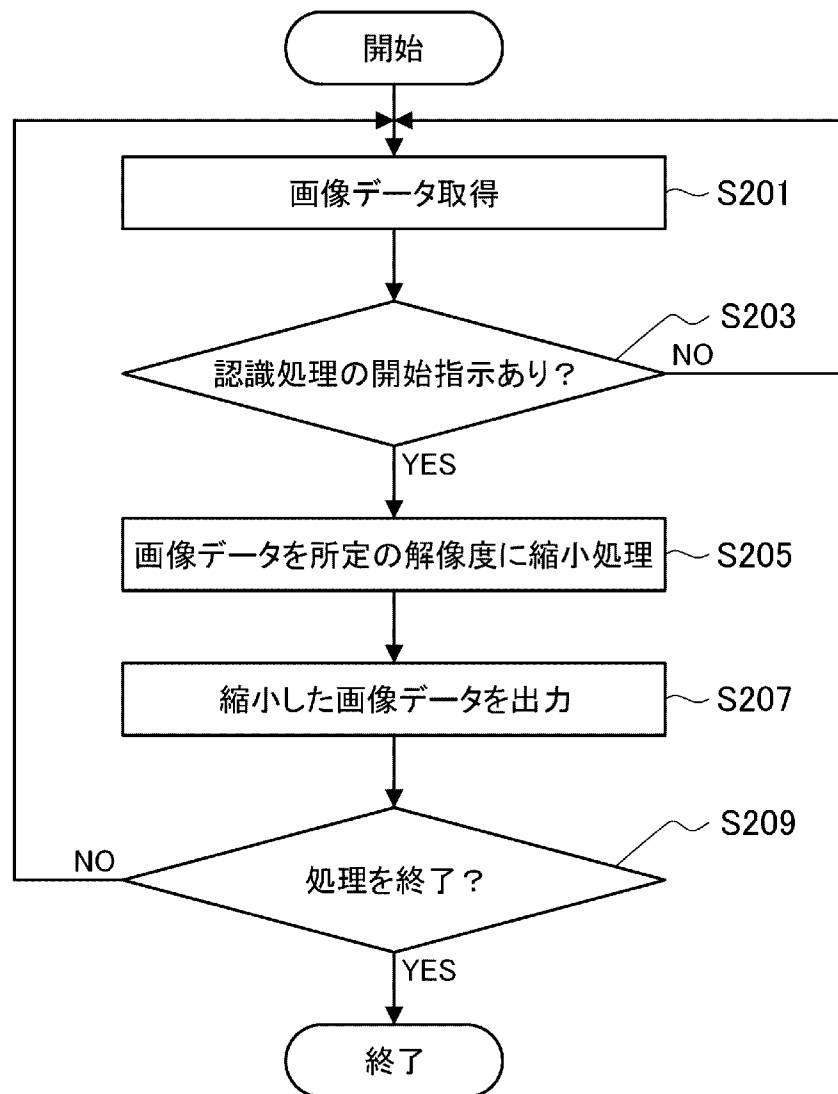
[図5]



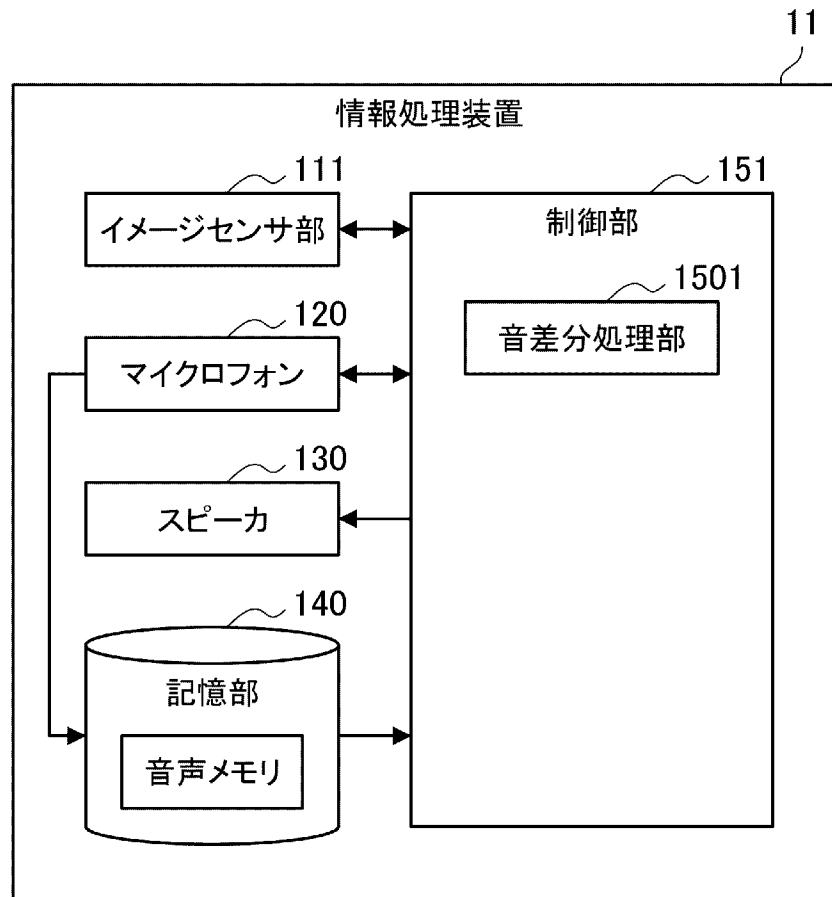
[図6]



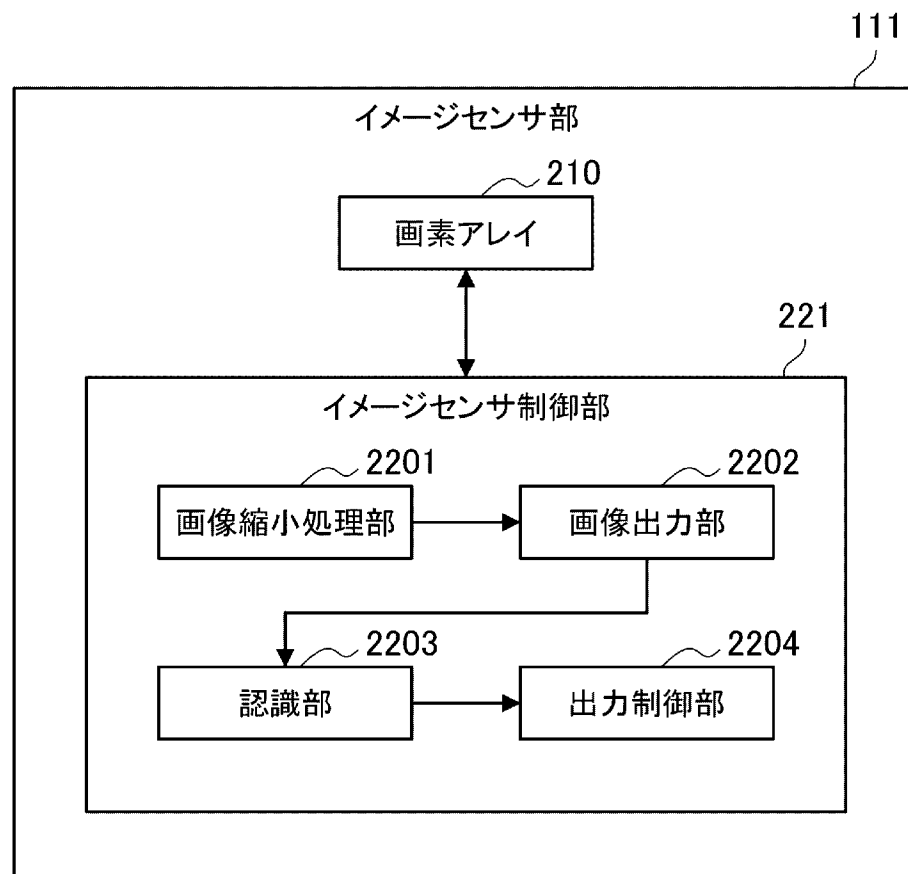
[図7]



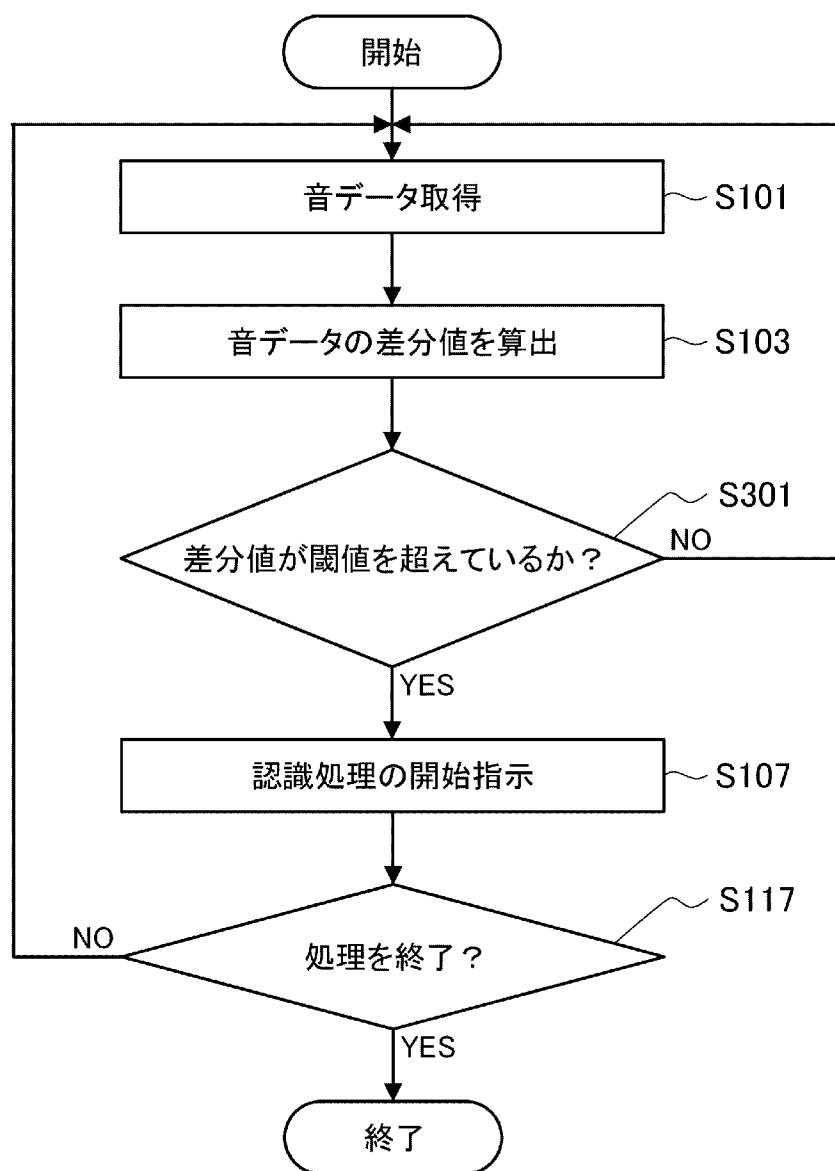
[図8]



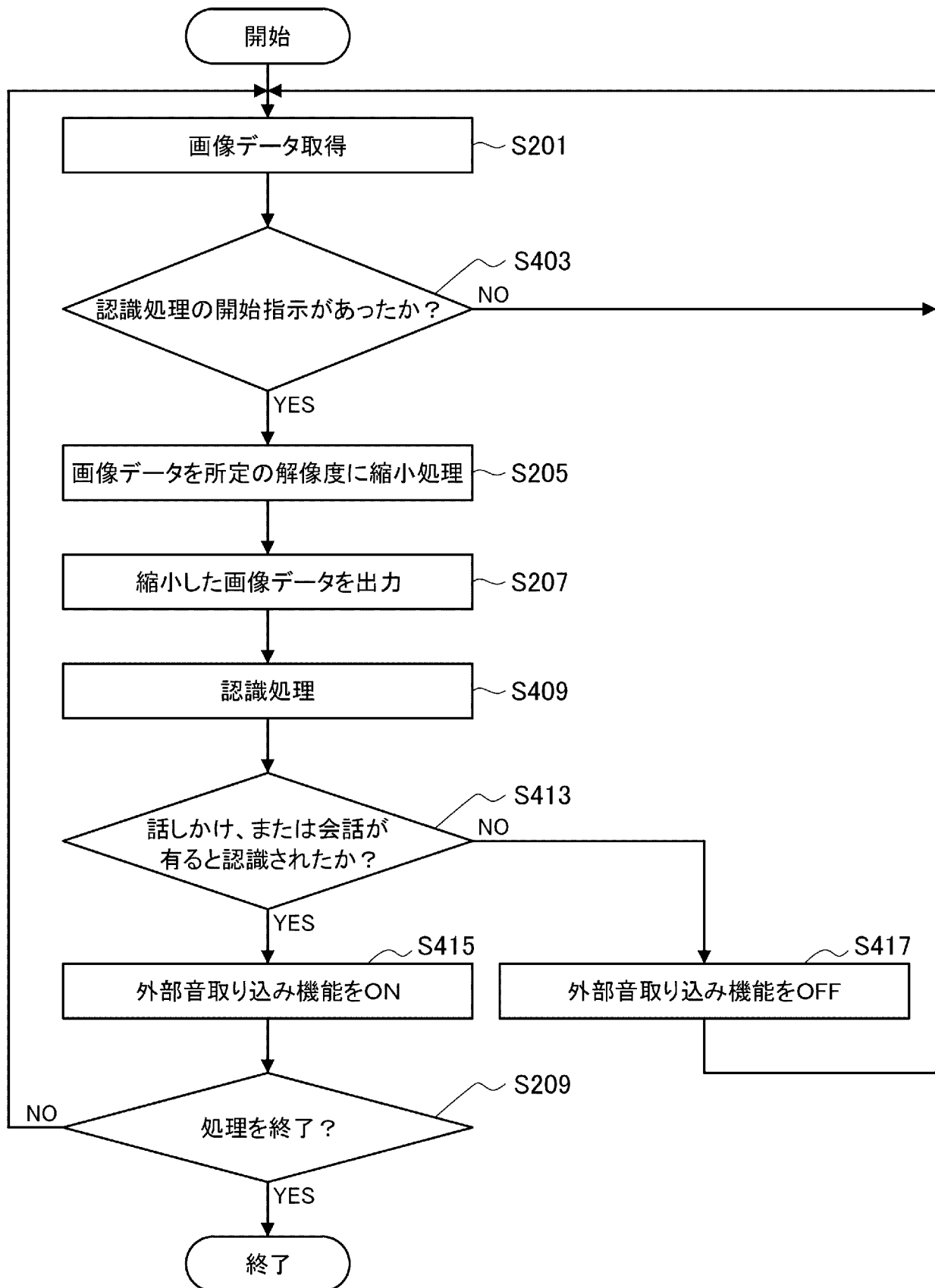
[図9]



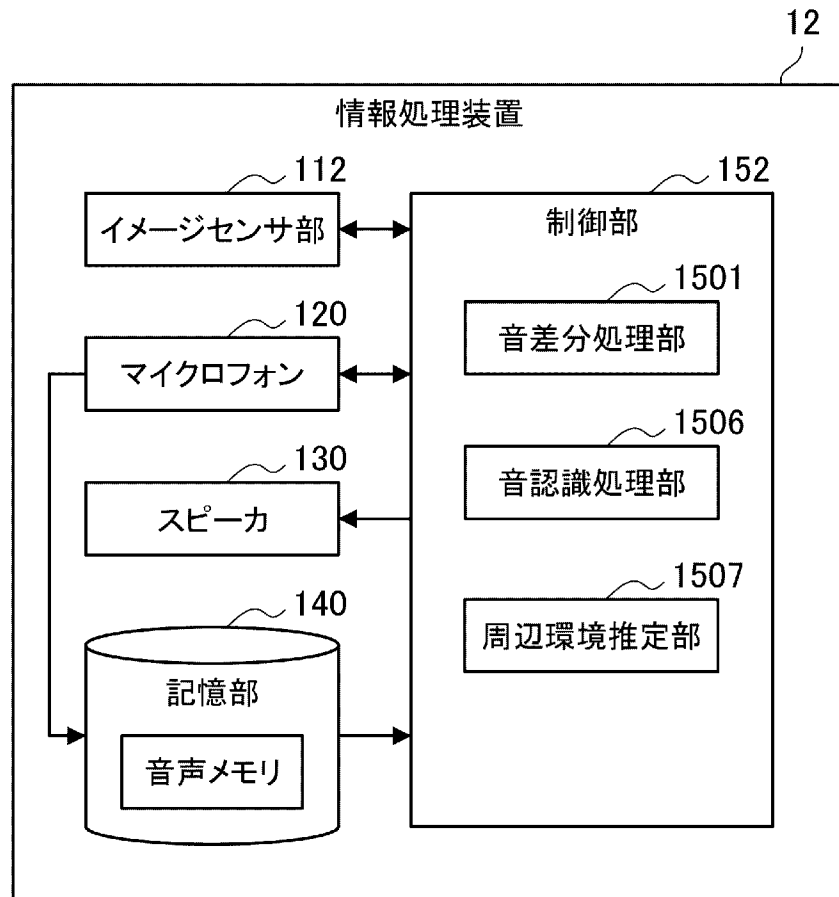
[図10]



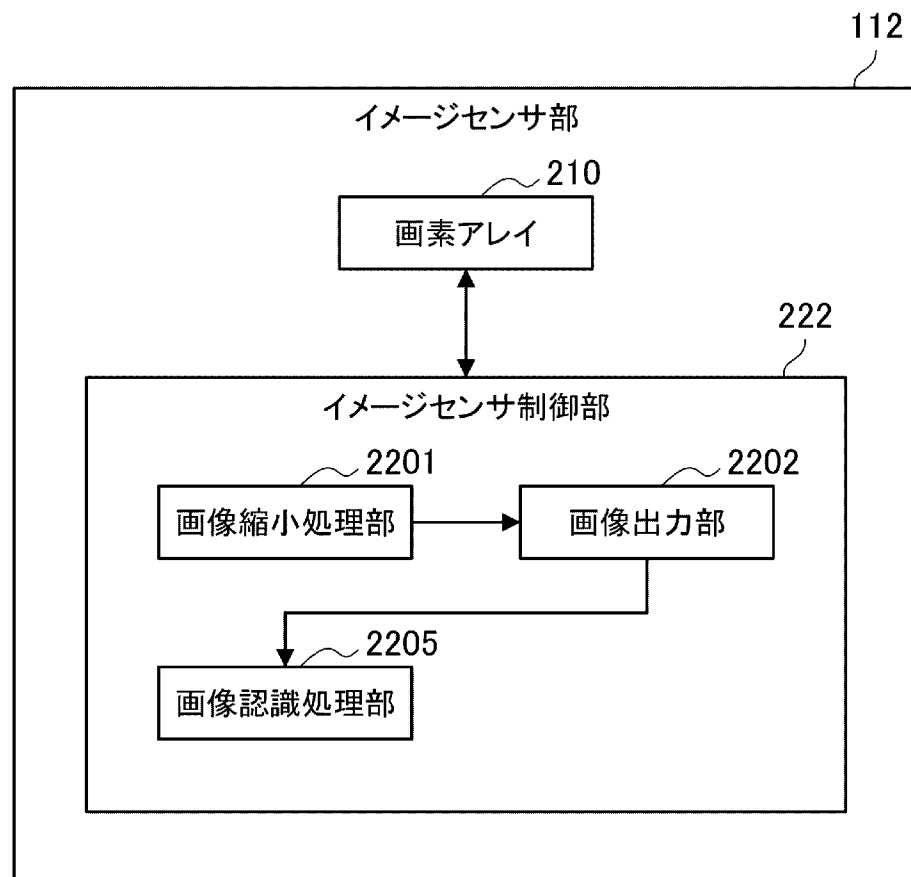
[図11]



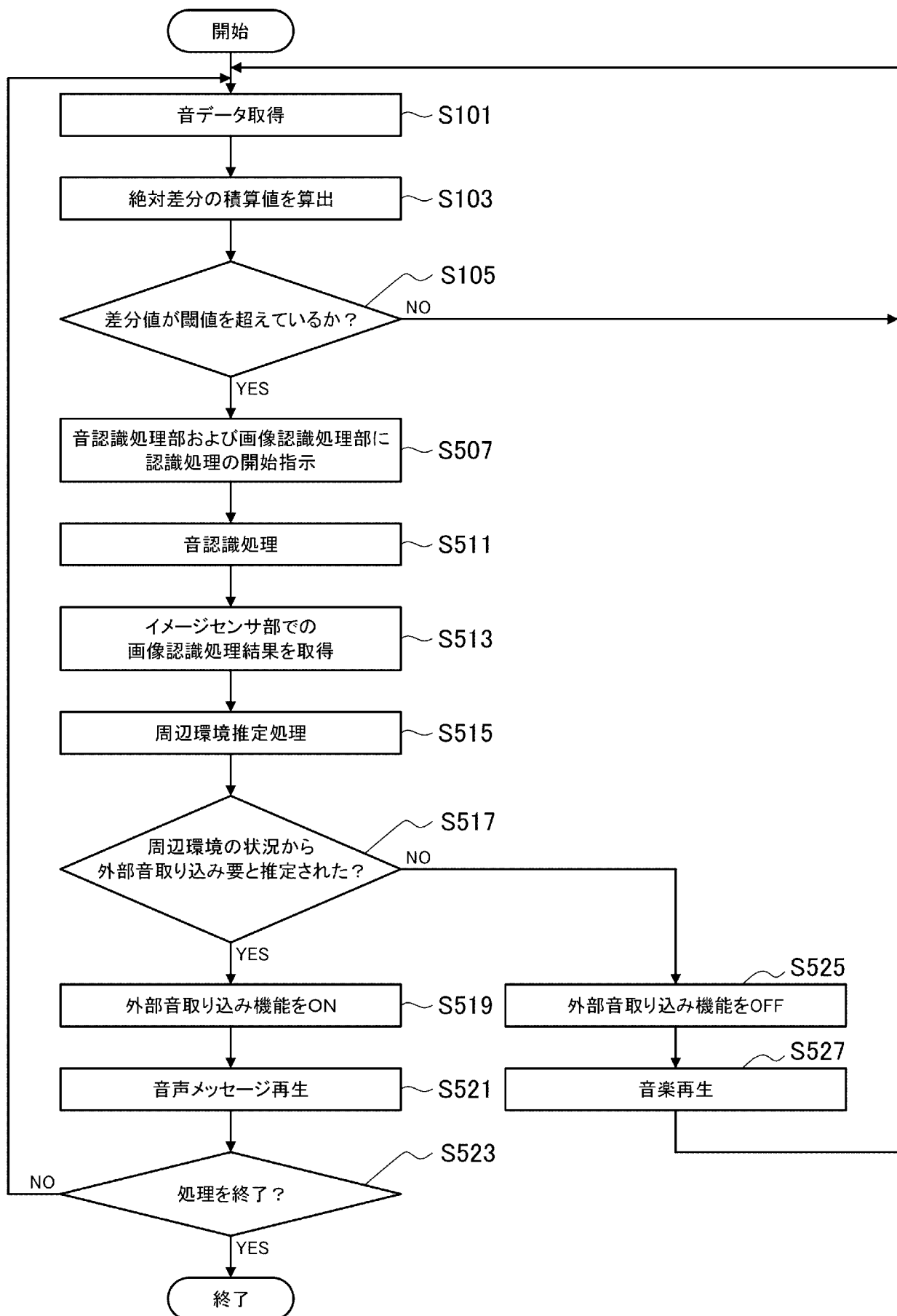
[図12]



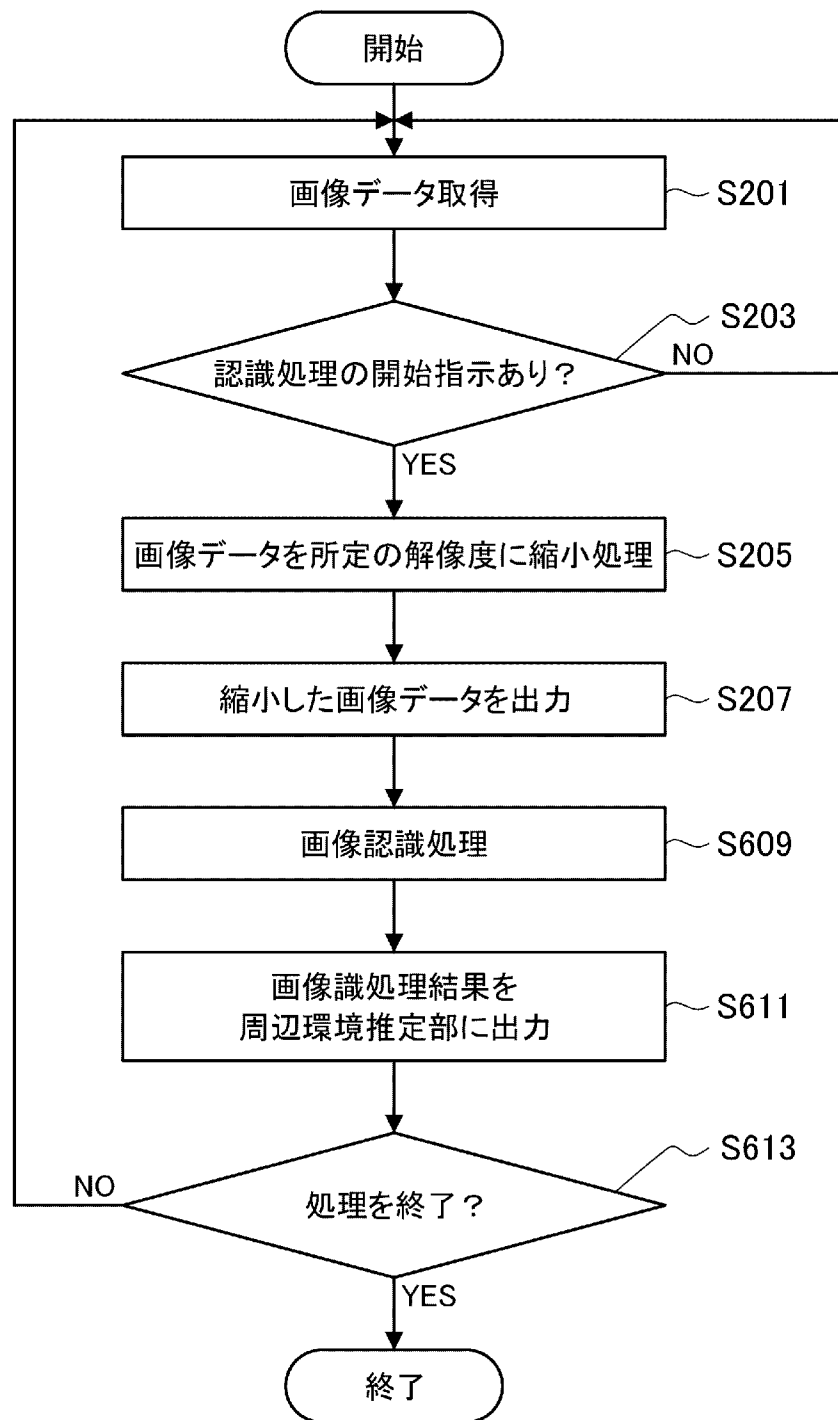
[図13]



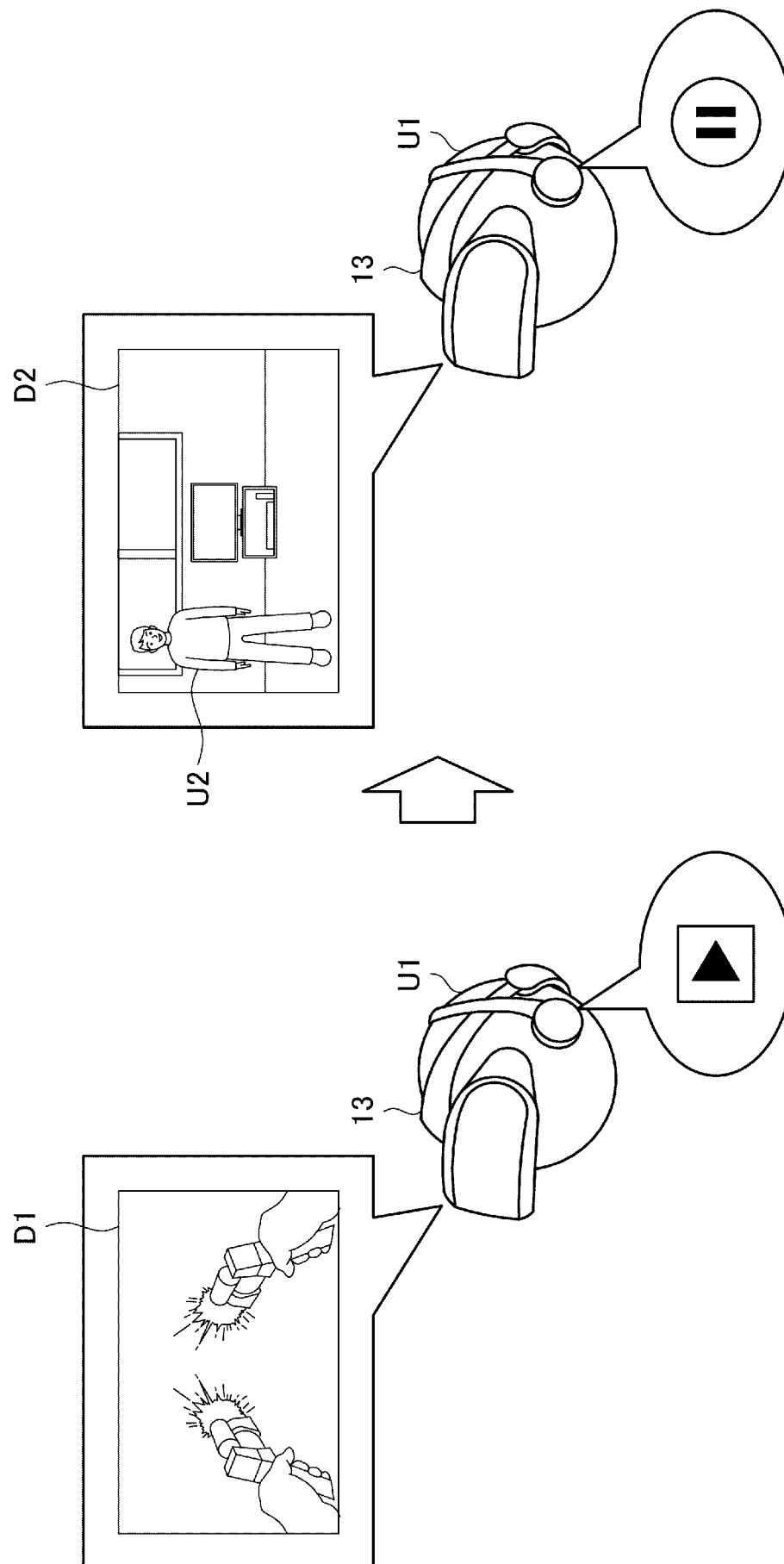
[図14]



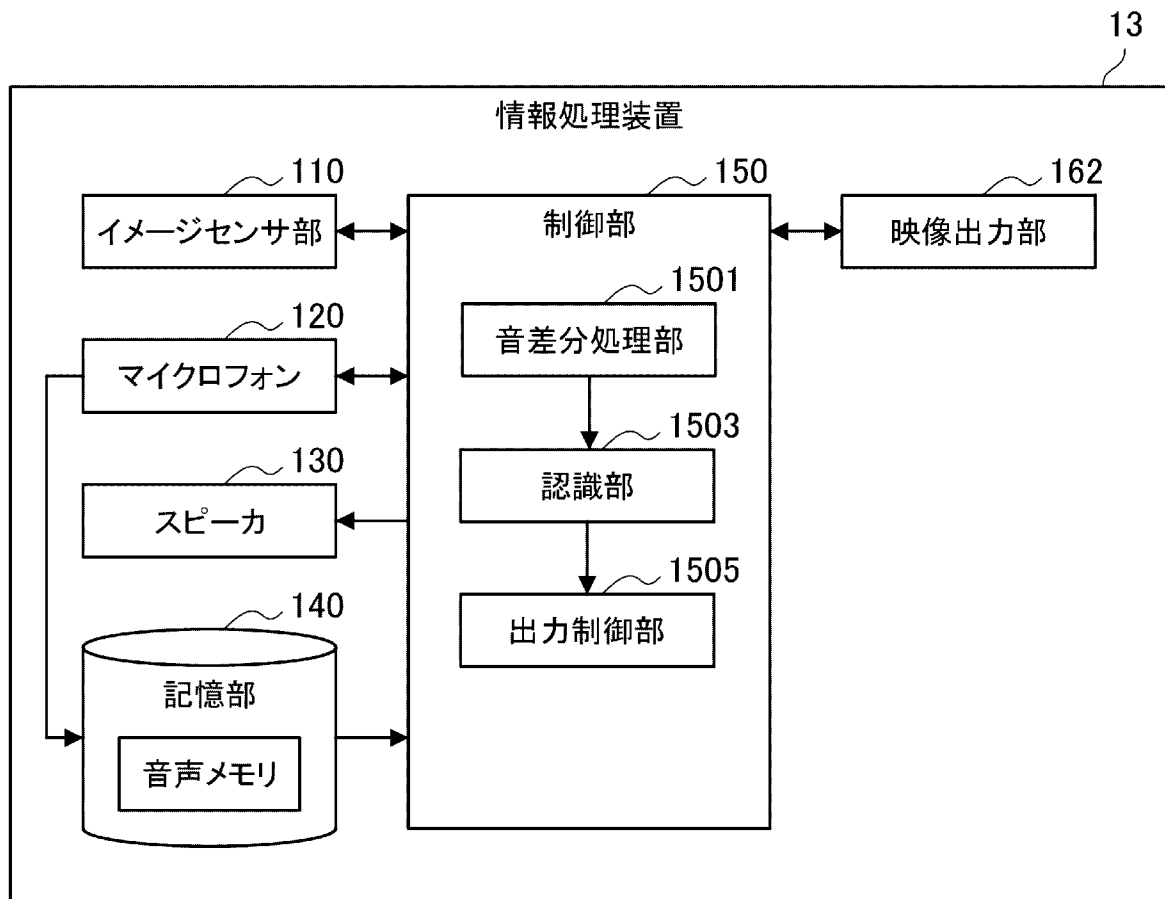
[図15]



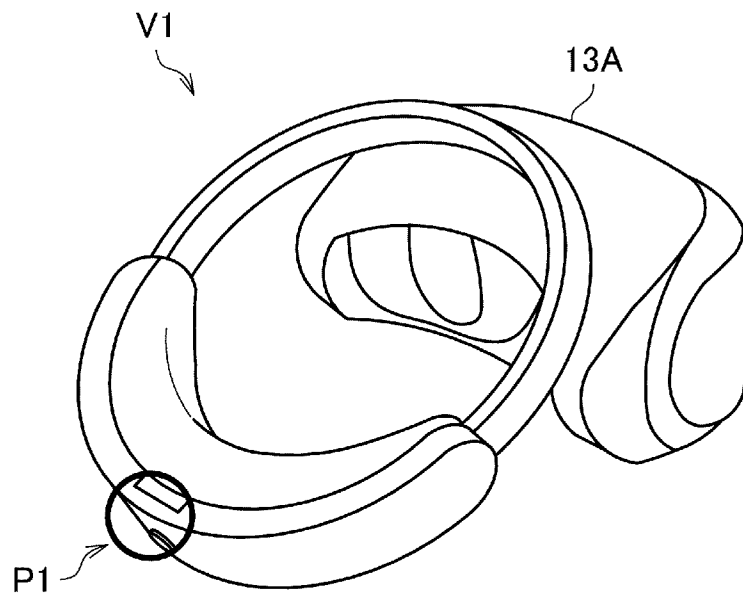
[図16]



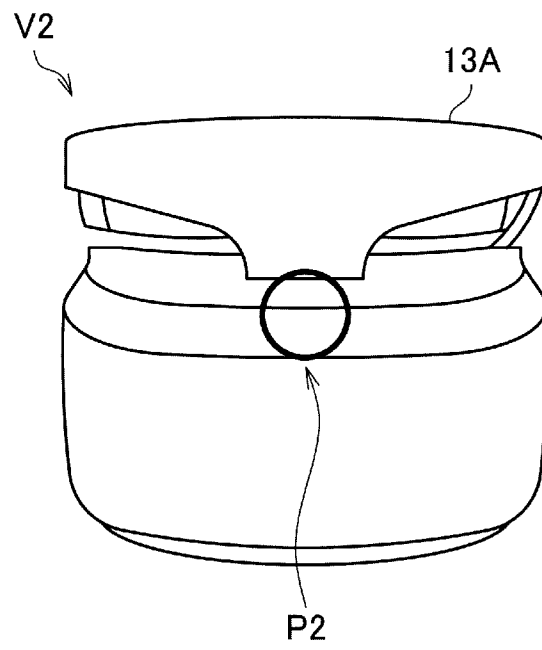
[図17]



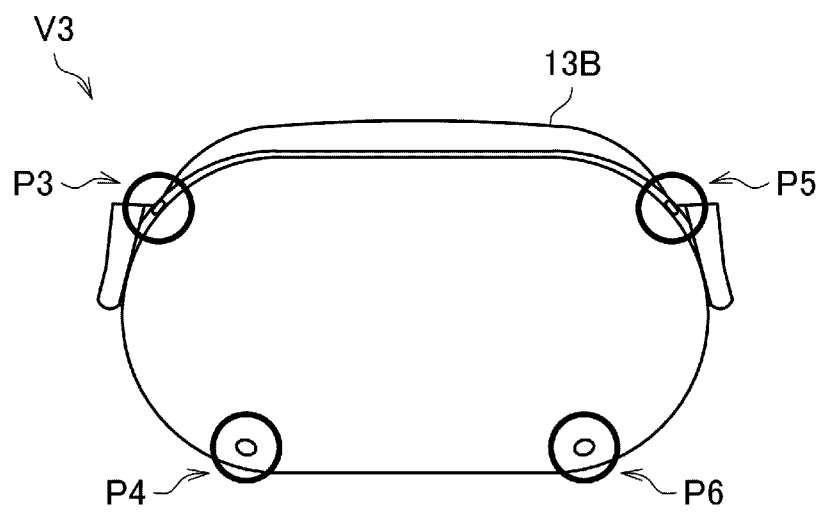
[図18]



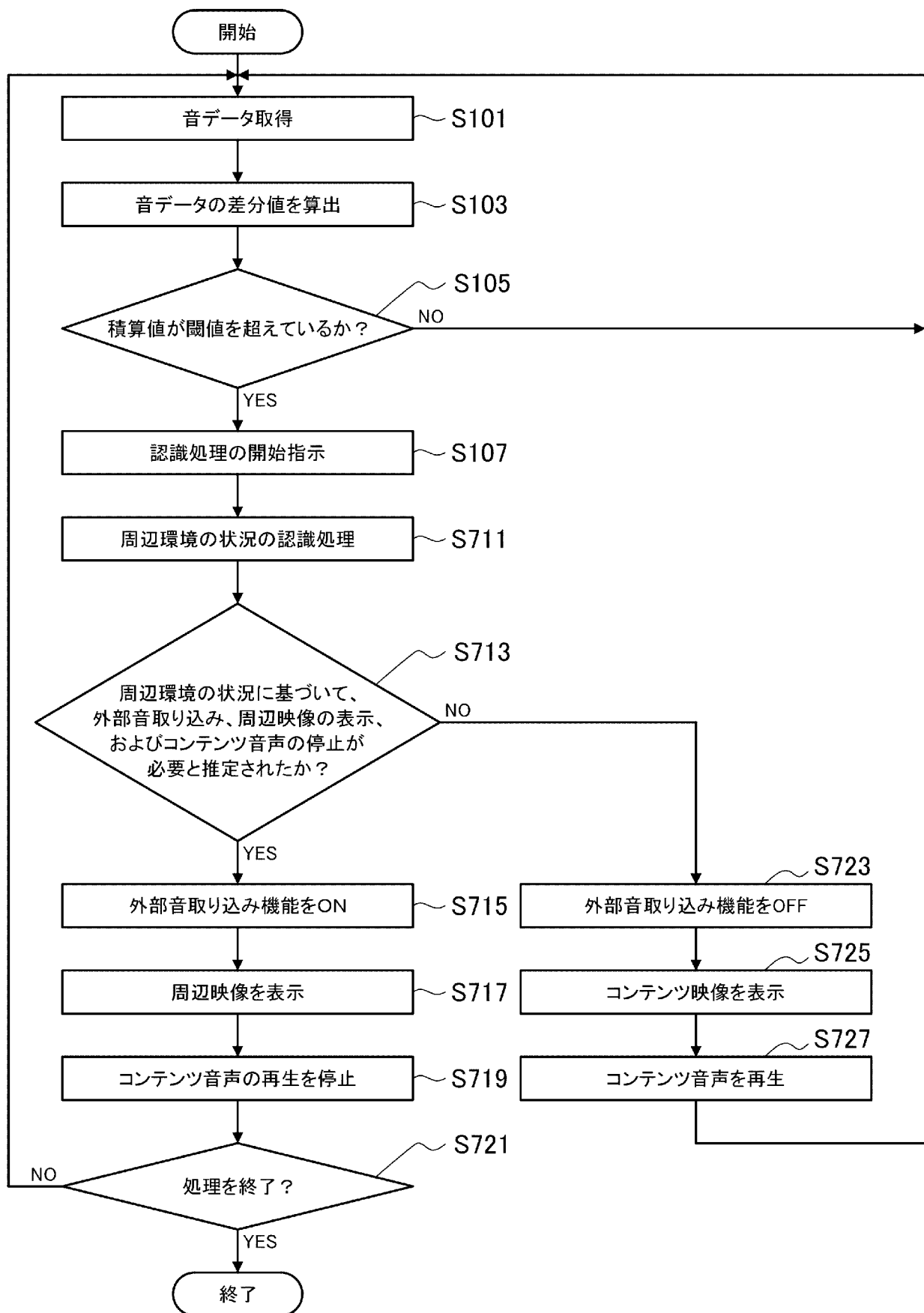
[図19]



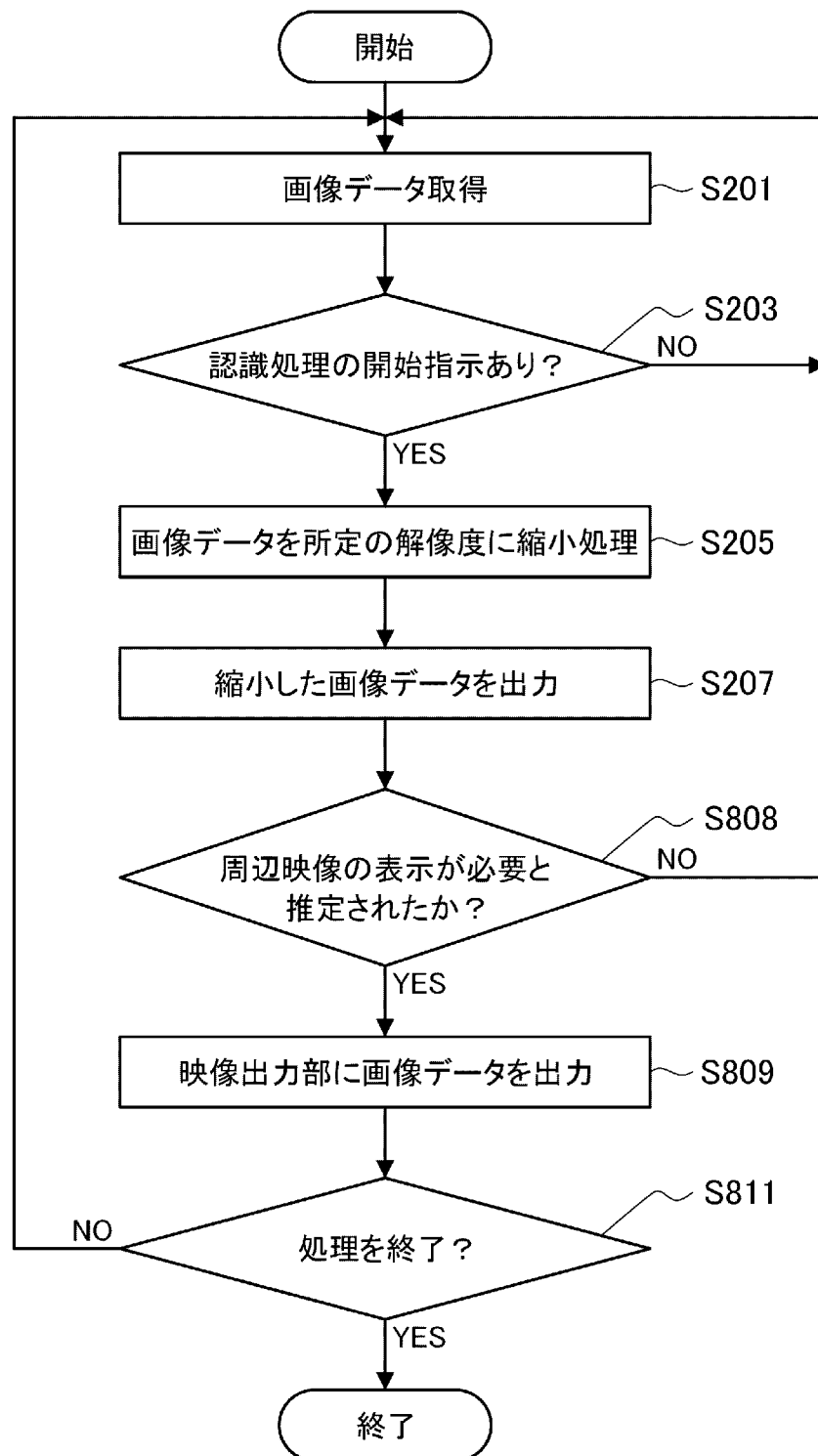
[図20]



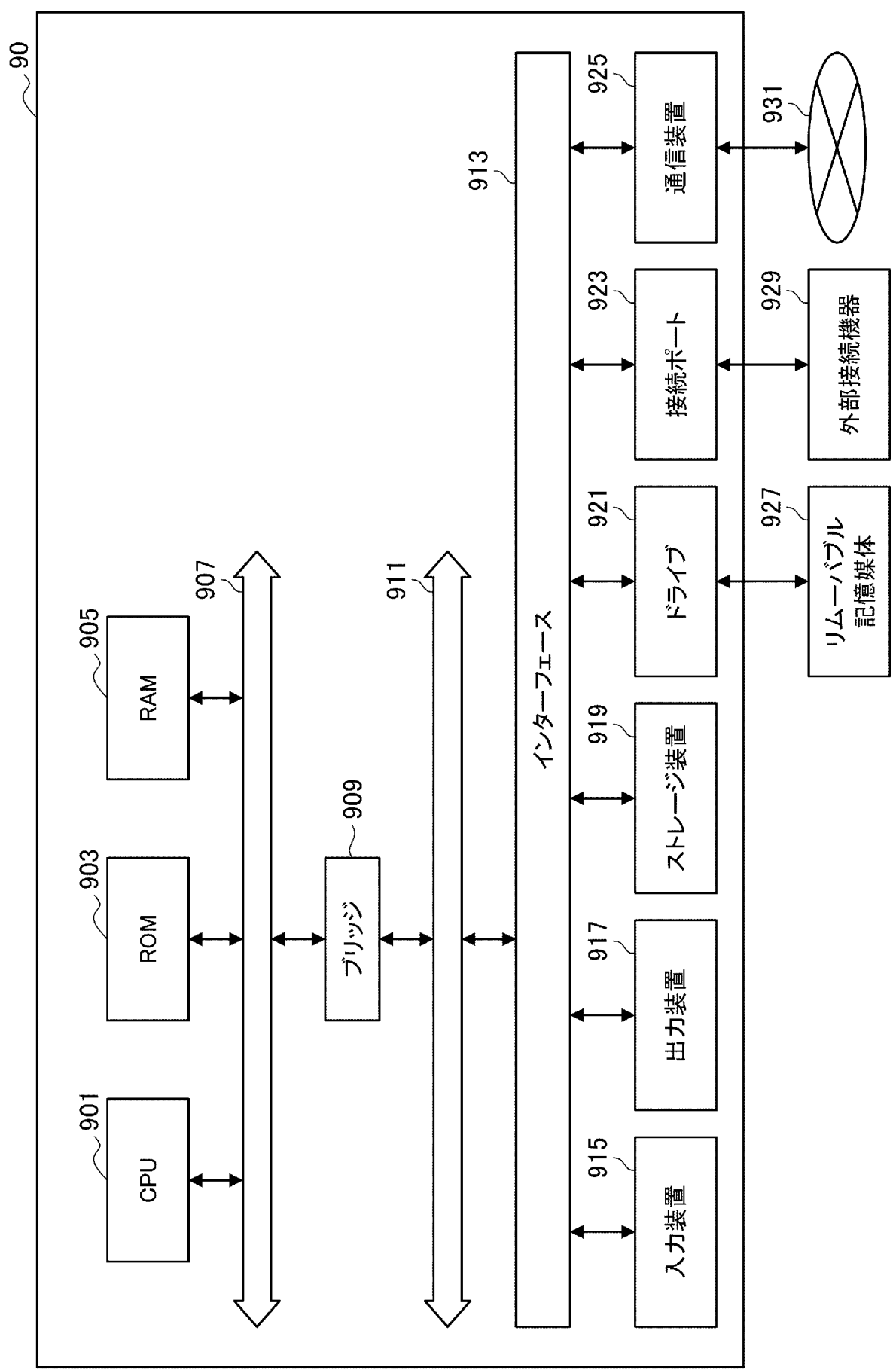
[図21]



[図22]



[図23]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/027195

A. CLASSIFICATION OF SUBJECT MATTER**G06F 3/01**(2006.01)i; **G06F 3/16**(2006.01)i

FI: G06F3/01 510; G06F3/16 650

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)

G06F3/01; G06F3/16

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 1971-2023
 Registered utility model specifications of Japan 1996-2023
 Published registered utility model applications of Japan 1994-2023

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 2019-197497 A (NEC PERSONAL COMPUTERS, LTD.) 14 November 2019 (2019-11-14) paragraphs [0063]-[0076]	1-11, 16-18
Y	JP 2017-026661 A (FUJI XEROX CO., LTD.) 02 February 2017 (2017-02-02) paragraph [0054]	1-11, 16-18
Y	JP 2016-014484 A (TOSHIBA CORP.) 28 January 2016 (2016-01-28) paragraphs [0146], [0147]	1-11, 16-18
Y	JP 2005-184700 A (ENERGIA COMMUNICATIONS, INC.) 07 July 2005 (2005-07-07) paragraph [0010]	2-11, 16
Y	JP 2022-105941 A (RYOYO ELECTRO CORP.) 15 July 2022 (2022-07-15) paragraphs [0017], [0018]	6-11
Y	JP 2015-191124 A (SOFTBANK BB CORP.) 02 November 2015 (2015-11-02) paragraphs [0027]-[0054]	1-6, 12-18



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

27 September 2023

Date of mailing of the international search report

10 October 2023

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
 Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/027195

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2013-131824 A (NIKON CORP.) 04 July 2013 (2013-07-04) paragraphs [0038]-[0040], [0056]	1-6, 12-18

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2023/027195

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2019-197497	A	14 November 2019	(Family: none)	
JP	2017-026661	A	02 February 2017	(Family: none)	
JP	2016-014484	A	28 January 2016	WO 2014/208585 A1 CN 105308403 A KR 10-2016-0003211 A	
JP	2005-184700	A	07 July 2005	(Family: none)	
JP	2022-105941	A	15 July 2022	WO 2022/149360 A1	
JP	2015-191124	A	02 November 2015	(Family: none)	
JP	2013-131824	A	04 July 2013	US 2014/0292512 A1 paragraphs [0045]-[0047], [0063] WO 2013/094091 A1 CN 103765860 A	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 3/01(2006.01)i; G06F 3/16(2006.01)i FI: G06F3/01 510; G06F3/16 650		
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）） G06F3/01; G06F3/16		
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 1971-2023年 日本国実用新案登録公報 1996-2023年 日本国登録実用新案公報 1994-2023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 2019-197497 A (NECパーソナルコンピュータ株式会社) 14.11.2019 (2019 - 11 - 14) 段落[0063]-[0076]	1-11, 16-18
Y	JP 2017-026661 A (富士ゼロックス株式会社) 02.02.2017 (2017 - 02 - 02) 段落[0054]	1-11, 16-18
Y	JP 2016-014484 A (株式会社東芝) 28.01.2016 (2016 - 01 - 28) 段落[0146]-[0147]	1-11, 16-18
Y	JP 2005-184700 A (株式会社エネルギア・コミュニケーションズ) 07.07.2005 (2005 - 07 - 07) 段落[0010]	2-11, 16
Y	JP 2022-105941 A (菱洋エレクトロ株式会社) 15.07.2022 (2022 - 07 - 15) 段落[0017]-[0018]	6-11
Y	JP 2015-191124 A (ソフトバンクBB株式会社) 02.11.2015 (2015 - 11 - 02) 段落[0027]-[0054]	1-6, 12-18
☒ 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、一般的な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 27.09.2023		国際調査報告の発送日 10.10.2023
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） 円子 英紀 5E 3979 電話番号 03-3581-1101 内線 3521

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 2013-131824 A (株式会社ニコン) 04.07.2013 (2013 - 07 - 04) 段落[0038]-[0040], [0056]	1-6, 12-18

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/JP2023/027195

引用文献			公表日	パテントファミリー文献	公表日
JP	2019-197497	A	14.11.2019	(ファミリーなし)	
JP	2017-026661	A	02.02.2017	(ファミリーなし)	
JP	2016-014484	A	28.01.2016	WO 2014/208585	A1
				CN 105308403	A
				KR 10-2016-0003211	A
JP	2005-184700	A	07.07.2005	(ファミリーなし)	
JP	2022-105941	A	15.07.2022	WO 2022/149360	A1
JP	2015-191124	A	02.11.2015	(ファミリーなし)	
JP	2013-131824	A	04.07.2013	US 2014/0292512	A1
				段落[0045]-[0047], [0063]	
				WO 2013/094091	A1
				CN 103765860	A