

(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17日(17.03.2022)



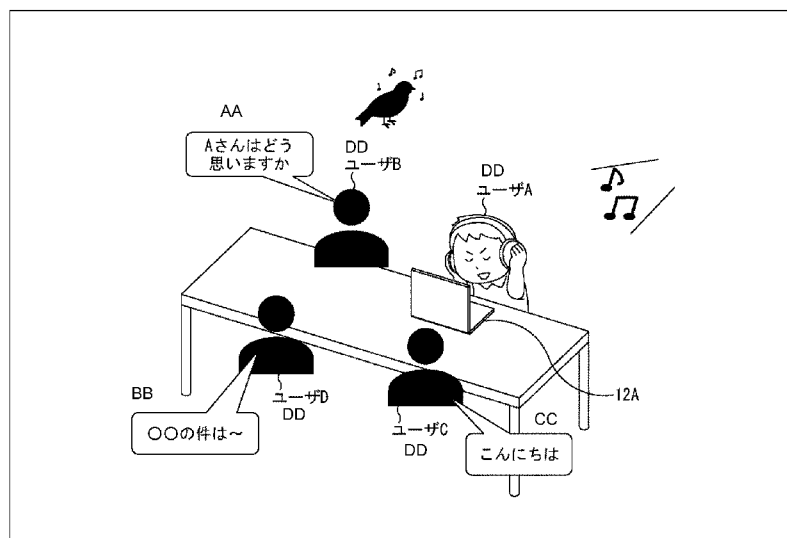
(10) 国際公開番号

WO 2022/054899 A1

- (51) 国際特許分類:
H04M 3/56 (2006.01) *H04N 7/15* (2006.01)
H04S 7/00 (2006.01) *H04N 21/233* (2011.01)
ト (SONY INTERACTIVE ENTERTAINMENT INC.) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 Tokyo (JP).
- (21) 国際出願番号: PCT/JP2021/033279 (72) 発明者: 大西 拓人(ONISHI, Takuto); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニーグループ株式会社内 Tokyo (JP). 北原 恵一(KITAHARA, Keiichi); 〒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). 寺坂 勇(TERASAKA, Isamu); 〒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). 藤原 真志(FUJIHARA, Masashi); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニーグループ株式会社内 Tokyo (JP). 中川 亨(NAKAGAWA, Jun); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニーグループ株式会社内 Tokyo (JP).
- (22) 国際出願日: 2021年9月10日(10.09.2021)
- (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 2020-152418 2020年9月10日(10.09.2020) JP
- (71) 出願人: ソニーグループ株式会社(SONY GROUP CORPORATION) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 Tokyo (JP). 株式会社ソニー・インタラクティブエンタテインメント

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

(54) 発明の名称: 情報処理装置、情報処理端末、情報処理方法、およびプログラム



AA What do you think, Mr. A?
BB Concerning the matter of ...
CC Hello
DD User

(57) Abstract: An information processing device according to one aspect of the present technology is provided with: a storage unit for storing HRTF data corresponding to a plurality of positions with reference to a listening position; and a sound image localizing process unit for performing a sound image localizing process on the basis of the HRTF data corresponding to the position in a virtual space of a participant to a conversation participating via a network, and voice data of the participant. The present technology may be applied to a computer for conducting a conference remotely.

Toru); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニーグループ株式会社内 Tokyo (JP).

(74) 代理人: 特許業務法人 酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関 3 丁目 8 番 1 号 虎の門三井ビルディング 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, 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, 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))

(57) 要約: 本技術の一側面の情報処理装置は、聴取位置を基準とした複数の位置に対応する HRTF データを記憶する記憶部と、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する HRTF データと、参加者の音声データとに基づいて音像定位処理を行う音像定位処理部とを備える。本技術は、リモート会議を行うコンピュータに適用することができる。

明 細 書

発明の名称：

情報処理装置、情報処理端末、情報処理方法、およびプログラム

技術分野

[0001] 本技術は、特に、臨場感のある会話を行うことができるようにした情報処理装置、情報処理端末、情報処理方法、およびプログラムに関する。

背景技術

[0002] 遠隔にいる複数の参加者がPCなどの装置を使用して会議を行ういわゆるリモート会議が行われている。PCにインストールされたWebブラウザや専用のアプリケーションを起動させ、会議毎に割り当てられたURLにより指定されるアクセス先にアクセスすることにより、URLを知っているユーザは、参加者として会議に参加することができる。

[0003] マイクロフォンにより集音された参加者の音声は、他の参加者が使用する装置にサーバを介して送信され、ヘッドホンやスピーカから出力される。また、カメラにより撮影された参加者が映る映像は、他の参加者が使用する装置にサーバを介して送信され、装置のディスプレイに表示される。

[0004] これにより、それぞれの参加者は、他の参加者の顔を見ながら会話を行うことができる。

先行技術文献

特許文献

[0005] 特許文献1：特開平11-331992号公報

発明の概要

発明が解決しようとする課題

[0006] 複数の参加者が同時に発話した場合の音声の聞き取りが難しい。

[0007] また、参加者の音声が平面的に出力されるだけであるため、音像などを感じることができず、参加者が実在している感覚を音声から得ることが難しい。

[0008] 本技術はこのような状況に鑑みてなされたものであり、臨場感のある会話を行うことができるようにするものである。

課題を解決するための手段

[0009] 本技術の一側面の情報処理装置は、聴取位置を基準とした複数の位置に対応するHRTFデータを記憶する記憶部と、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う音像定位処理部とを備える。

[0010] 本技術の他の側面の情報処理端末は、聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う情報処理装置から送信されてきた、前記音像定位処理を行うことによって得られた発話者となる前記参加者の音声データを受信し、前記発話者の音声を出力する音声受信部を備える。

[0011] 本技術の一側面においては、聴取位置を基準とした複数の位置に対応するHRTFデータが記憶され、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理が行われる。

[0012] 本技術の他の側面においては、聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う情報処理装置から送信されてきた、前記音像定位処理を行うことによって得られた発話者となる前記参加者の音声データが受信され、前記発話者の音声が出力される。

図面の簡単な説明

[0013] [図1]本技術の一実施形態に係るTele-communicationシステムの構成例を示す図である。

[図2]音声データの送受信の例を示す図である。

[図3]仮想空間上のユーザの位置の例を示す平面図である。

[図4]リモート会議画面の表示例を示す図である。

[図5]音声の聞こえ方の例を示す図である。

[図6]音声の聞こえ方の他の例を示す図である。

[図7]会議に参加しているユーザの様子を示す図である。

[図8]コミュニケーション管理サーバの基本処理について説明するフローチャートである。

[図9]クライアント端末の基本処理について説明するフローチャートである。

[図10]コミュニケーション管理サーバのハードウェア構成例を示すブロック図である。

[図11]コミュニケーション管理サーバの機能構成例を示すブロック図である。

[図12]参加者情報の例を示す図である。

[図13]クライアント端末のハードウェア構成例を示すブロック図である。

[図14]クライアント端末の機能構成例を示すブロック図である。

[図15]グループ設定画面の例を示す図である。

[図16]発話ユーザのグルーピングに関する処理の流れを示す図である。

[図17]コミュニケーション管理サーバの制御処理について説明するフローチャートである。

[図18]位置設定画面の例を示す図である。

[図19]位置情報の共有に関する処理の流れを示す図である。

[図20]コミュニケーション管理サーバの制御処理について説明するフローチャートである。

[図21]背景音の設定に用いられる画面の例を示す図である。

[図22]背景音の設定に関する処理の流れを示す図である。

[図23]コミュニケーション管理サーバの制御処理について説明するフローチャートである。

[図24]背景音の設定に関する処理の流れを示す図である。

[図25]コミュニケーション管理サーバの制御処理について説明するフローチャートである。

ャートである。

[図26] 音像定位処理の動的切り替えに関する処理の流れを示す図である。

[図27] コミュニケーション管理サーバの制御処理について説明するフローチャートである。

[図28] 音響設定の管理に関する処理の流れを示す図である。

発明を実施するための形態

[0014] 以下、本技術を実施するための形態について説明する。説明は以下の順序で行う。

1. Tele-communicationシステムの構成
2. 基本的な動作
3. 各装置の構成
4. 音像定位のユースケース
5. 変形例

[0015] <<Tele-communicationシステムの構成>>

図1は、本技術の一実施形態に係るTele-communicationシステムの構成例を示す図である。

[0016] 図1のTele-communicationシステムは、コミュニケーション管理サーバ1に対して、会議の参加者が使用する複数のクライアント端末がインターネットなどのネットワーク11を介して接続されることによって構成される。図1の例においては、PCであるクライアント端末2Aから2Dが、会議の参加者であるユーザAからDが使用するクライアント端末として示されている。

[0017] マイクフォン（マイク）などの音声入力デバイスと、ヘッドホンやスピーカなどの音声出力デバイスとを有する、スマートフォンやタブレット端末などの他のデバイスがクライアント端末として用いられるようにしてもよい。クライアント端末2Aから2Dを区別する必要がある場合、適宜、クライアント端末2という。

[0018] ユーザAからDは、同じ会議に参加するユーザである。なお、会議に参加するユーザの数は4人に限定されるものではない。

[0019] コミュニケーション管理サーバ1は、複数のユーザがオンライン上で会話を行うことによって進められる会議を管理する。コミュニケーション管理サーバ1は、クライアント端末2間の音声の送受信を制御し、いわゆるリモート会議を管理する情報処理装置である。

[0020] 例えば、コミュニケーション管理サーバ1は、図2の上段の矢印A1に示すように、ユーザAが発話することに応じてクライアント端末2Aから送信されてきたユーザAの音声データを受信する。クライアント端末2Aからは、クライアント端末2Aに設けられたマイクにより集音されたユーザAの音声データが送信されてくる。

[0021] コミュニケーション管理サーバ1は、ユーザAの音声データを、図2の下段の矢印A11からA13に示すようにクライアント端末2Bから2Dのそれぞれに送信し、ユーザAの音声を出力させる。ユーザAが発話者として発話した場合、ユーザBからDが聴取者となる。以下、適宜、発話者となるユーザを発話ユーザといい、聴取者となるユーザを聴取ユーザという。

[0022] 他のユーザが発話を行った場合も同様に、発話ユーザが使用するクライアント端末2から送信された音声データは、コミュニケーション管理サーバ1を経由して、聴取ユーザが使用するクライアント端末2に送信される。

[0023] コミュニケーション管理サーバ1は、それぞれのユーザの仮想空間上の位置を管理する。仮想空間は、会議を行う場所として仮想的に設定された例えば3次元の空間である。仮想空間上の位置は3次元の座標で表される。

[0024] 図3は、仮想空間上のユーザの位置の例を示す平面図である。

[0025] 図3の例においては、矩形の枠Fで示される仮想空間の略中央に縦長長方形のテーブルTが配置され、テーブルTの周りの位置である位置P1からP4が、それぞれ、ユーザAからDの位置として設定されている。それぞれのユーザの正面方向は、それぞれのユーザの位置からテーブルTの方向である。

[0026] 会議中、それぞれのユーザが使用するクライアント端末2の画面には、図4に示すように、会議を行う場所を表す背景画像に重ねて、ユーザを視覚的

に表す情報である参加者アイコンが表示される。参加者アイコンの画面上の位置は、仮想空間上のそれぞれのユーザの位置に応じた位置となる。

[0027] 図4の例においては、参加者アイコンは、ユーザの顔を含む円形状の画像として構成されている。参加者アイコンは、仮想空間に設定された基準の位置からそれぞれのユーザの位置までの距離に応じた大きさで表示される。参加者アイコン1から4は、それぞれユーザAからDを表す。

[0028] 例えば、それぞれのユーザの位置は、会議に参加したときにコミュニケーション管理サーバ1により自動的に設定される。図4の画面上で参加者アイコンを移動させるなどして、仮想空間上の位置がユーザ自身により設定されるようにしてもよい。

[0029] コミュニケーション管理サーバ1は、仮想空間上のそれぞれの位置を聴取位置としたときの、複数の位置から聴取位置までの音の伝達特性を表すHRTF(Head-Related Transfer Function)(頭部伝達関数)のデータであるHRTFデータを有している。仮想空間上のそれぞれの聴取位置を基準とした、複数の位置に対応するHRTFデータがコミュニケーション管理サーバ1に用意されている。

[0030] コミュニケーション管理サーバ1は、それぞれの聴取ユーザにとって、発話ユーザの音声 that 当該発話ユーザの仮想空間上の位置から聞こえるように、HRTFデータを用いた音像定位処理を音声データに対して行い、音像定位処理を行うことによって得られた音声データを送信する。

[0031] 上述したようにしてクライアント端末2に送信される音声データは、コミュニケーション管理サーバ1において音像定位処理が行われることによって得られた音声データとなる。音像定位処理には、位置情報に基づくVBAP(Vector Based Amplitude Panning)などのレンダリング、HRTFデータを用いたバイノーラル処理が含まれる。

[0032] すなわち、それぞれの発話ユーザの音声は、オブジェクトオーディオの音声データとしてコミュニケーション管理サーバ1において処理される。コミュニケーション管理サーバ1における音像定位処理により生成された、例え

ばL／Rの2チャンネルのチャンネルベースのオーディオデータがコミュニケーション管理サーバ1からそれぞれのクライアント端末2に送信され、クライアント端末2に設けられたヘッドホンなどから、発話ユーザの音声出力される。

[0033] 聴取ユーザ自身の位置と発話ユーザの位置との相対的な位置関係に応じたHRTFデータを用いた音像定位処理が行われることにより、それぞれの聴取ユーザは、発話ユーザの音声を、発話ユーザの位置から聞こえるように感じるようになる。

[0034] 図5は、音声の聞こえ方の例を示す図である。

[0035] 位置P1が仮想空間上の位置として設定されているユーザAを聴取ユーザとして注目すると、ユーザBの音声は、図5の矢印で示すように、位置P2を音源位置とする位置P2－位置P1間のHRTFデータに基づいて音像定位処理が行われることにより、右隣から聞こえる。クライアント端末2Aに顔を向けて会話を行っているユーザAの正面は、クライアント端末2Aの方向である。

[0036] また、ユーザCの音声は、位置P3を音源位置とする位置P3－位置P1間のHRTFデータに基づいて音像定位処理が行われることにより、正面から聞こえる。ユーザDの音声は、位置P4を音源位置とする位置P4－位置P1間のHRTFデータに基づいて音像定位処理が行われることにより、右奥から聞こえる。

[0037] 他のユーザが聴取ユーザである場合も同様である。例えば、図6に示すように、ユーザAの音声は、クライアント端末2Bに顔を向けて会話を行っているユーザBにとっては左隣から聞こえ、クライアント端末2Cに顔を向けて会話を行っているユーザCにとっては正面から聞こえる。また、ユーザAの音声は、クライアント端末2Dに顔を向けて会話を行っているユーザDにとっては右奥から聞こえる。

[0038] このように、コミュニケーション管理サーバ1においては、それぞれの聴取ユーザ用の音声データが、それぞれの聴取ユーザの位置と発話ユーザの位

置との位置関係に応じて生成され、発話ユーザの音声の出力に用いられる。それぞれの聴取ユーザに対して送信される音声データは、それぞれの聴取ユーザの位置と発話ユーザの位置との位置関係に応じて聞こえ方が異なる音声データとなる。

[0039] 図7は、会議に参加しているユーザの様子を示す図である。

[0040] 例えばヘッドホンを装着して会議に参加しているユーザAは、右隣、正面、右奥のそれぞれの位置に音像が定位しているユーザBからDの音声を聞き、会話を行うことになる。図5等を参照して説明したように、ユーザAの位置を基準とすると、ユーザBからDの位置は、それぞれ、右隣、正面、右奥の位置である。なお、図7においてユーザBからDに色を付して示していることは、ユーザBからDが、ユーザAが会議を行っている空間と同じ空間に実在していないことを表す。

[0041] なお、後述するように、鳥のさえずりやBGMなどの背景音についても、所定の位置に音像が定位するように、音像定位処理によって得られた音声データに基づいて出力される。

[0042] コミュニケーション管理サーバ1が処理対象とする音声には、発話音声だけでなく、環境音や背景音などの音も含まれる。以下、適宜、それぞれの音の種類を区別する必要がある場合、コミュニケーション管理サーバ1が処理対象とする音を単に音声として説明する。実際には、コミュニケーション管理サーバ1が処理対象とする音には、音声以外の種類の音も含まれる。

[0043] 発話ユーザの音声が入仮想空間における位置に応じた位置から聞こえることにより、聴取ユーザは、参加者が複数いる場合であっても、それぞれのユーザの音声を容易に聞き分けることができる。例えば複数のユーザが同時に発話を行った場合であっても、聴取ユーザは、それぞれの音声を聞き分けることが可能となる。

[0044] また、発話ユーザの音声が入立体的に感じられるため、聴取ユーザは、音像の位置に発話ユーザが実在している感覚を音声から得ることができる。聴取ユーザは、臨場感のある会話を他のユーザとの間で行うことができる。

[0045] <<基本的な動作>>

ここで、コミュニケーション管理サーバ1とクライアント端末2の基本的な動作の流れについて説明する。

[0046] <コミュニケーション管理サーバ1の動作>

図8のフローチャートを参照して、コミュニケーション管理サーバ1の基本処理について説明する。

[0047] ステップS1において、コミュニケーション管理サーバ1は、クライアント端末2から音声データが送信されてきたか否かを判定し、音声データが送信されてきたと判定するまで待機する。

[0048] クライアント端末2から音声データが送信されてきたとステップS1において判定した場合、ステップS2において、コミュニケーション管理サーバ1は、クライアント端末2から送信されてきた音声データを受信する。

[0049] ステップS3において、コミュニケーション管理サーバ1は、それぞれのユーザの位置情報に基づいて音像定位処理を行い、それぞれの聴取ユーザ用の音声データを生成する。

[0050] 例えば、ユーザA用の音声データは、発話ユーザの音声の音像が、ユーザAの位置を基準としたときに、その発話ユーザの位置に応じた位置に定位するようにして生成される。

[0051] また、ユーザB用の音声データは、発話ユーザの音声の音像が、ユーザBの位置を基準としたときに、その発話ユーザの位置に応じた位置に定位するようにして生成される。

[0052] 他の聴取ユーザ用の音声データについても同様に、聴取ユーザの位置を基準として、発話ユーザとの位置の相対的な位置関係に応じたHRTFデータを用いて生成される。それぞれの聴取ユーザ用の音声データは異なるデータとなる。

[0053] ステップS4において、コミュニケーション管理サーバ1は、それぞれの聴取ユーザに対して音声データを送信する。以上の処理が、発話ユーザが使用するクライアント端末2から音声データが送信されてくる毎に行われる。

[0054] <クライアント端末2の動作>

図9のフローチャートを参照して、クライアント端末2の基本処理について説明する。

[0055] ステップS11において、クライアント端末2は、マイク音声が入力されたか否かを判定する。マイク音声は、クライアント端末2に設けられたマイクにより集音された音声である。

[0056] マイク音声が入力されたステップS11において判定した場合、ステップS12において、クライアント端末2は、音声データをコミュニケーション管理サーバ1に送信する。マイク音声が入力されていないとステップS11において判定された場合、ステップS12の処理はスキップされる。

[0057] ステップS13において、クライアント端末2は、コミュニケーション管理サーバ1から音声データが送信されてきたか否かを判定する。

[0058] 音声データが送信されてきたステップS13において判定した場合、ステップS14において、コミュニケーション管理サーバ1は、音声データを受信し、発話ユーザの音声を出力する。

[0059] 発話ユーザの音声出力された後、または、音声データが送信されてきないとステップS13において判定された場合、ステップS11に戻り、上述した処理が繰り返し行われる。

[0060] <<各装置の構成>>

<コミュニケーション管理サーバ1の構成>

図10は、コミュニケーション管理サーバ1のハードウェア構成例を示すブロック図である。

[0061] コミュニケーション管理サーバ1はコンピュータにより構成される。コミュニケーション管理サーバ1が、図10に示す構成を有する1台のコンピュータにより構成されるようにしてもよいし、複数台のコンピュータにより構成されるようにしてもよい。

[0062] CPU101、ROM102、RAM103は、バス104により相互に接続される。CPU101は、サーバプログラム101Aを実行し、コミュニケーション管

理サーバ1の全体の動作を制御する。サーバプログラム101Aは、Tele-communicationシステムを実現するためのプログラムである。

[0063] バス104には、さらに、入出カインタフェース105が接続される。入出カインタフェース105には、キーボード、マウスなどよりなる入力部106、ディスプレイ、スピーカなどよりなる出力部107が接続される。

[0064] また、入出カインタフェース105には、ハードディスクや不揮発性のメモリなどよりなる記憶部108、ネットワークインタフェースなどよりなる通信部109、リムーバブルメディア111を駆動するドライブ110が接続される。例えば、通信部109は、それぞれのユーザが使用するクライアント端末2との間でネットワーク11を介して通信を行う。

[0065] 図11は、コミュニケーション管理サーバ1の機能構成例を示すブロック図である。図11に示す機能部のうちの少なくとも一部は、図10のCPU101によりサーバプログラム101Aが実行されることによって実現される。

[0066] コミュニケーション管理サーバ1においては情報処理部121が実現される。情報処理部121は、音声受信部131、信号処理部132、参加者情報管理部133、音像定位処理部134、HRTFデータ記憶部135、システム音声管理部136、2chミックス処理部137、および音声送信部138から構成される。

[0067] 音声受信部131は、通信部109を制御し、発話ユーザが使用するクライアント端末2から送信されてきた音声データを受信する。音声受信部131により受信された音声データは、信号処理部132に出力される。

[0068] 信号処理部132は、音声受信部131から供給された音声データに対して、所定の信号処理を適宜施し、信号処理を施すことによって得られた音声データを音像定位処理部134に出力する。例えば、発話ユーザの音声と環境音を分離する処理が信号処理部132により行われる。マイク音声には、発話ユーザの音声の他に、発話ユーザがいる空間の騒音やノイズなどの環境音が含まれる。

[0069] 参加者情報管理部133は、通信部109を制御し、クライアント端末2

と通信を行うなどして、会議の参加者に関する情報である参加者情報を管理する。

[0070] 図 1 2 は、参加者情報の例を示す図である。

[0071] 図 1 2 に示すように、参加者情報には、ユーザ情報、位置情報、設定情報、ボリューム情報が含まれる。

[0072] ユーザ情報は、あるユーザが設定した会議に参加するユーザの情報である。例えば、ユーザのIDなどがユーザ情報に含まれる。参加者情報に含まれる他の情報が例えばユーザ情報に紐付けて管理される。

[0073] 位置情報は、仮想空間上のそれぞれのユーザの位置を表す情報である。

[0074] 設定情報は、会議で使用する背景音の設定などの、会議に関する設定の内容を表す情報である。

[0075] ボリューム情報は、それぞれのユーザの音声を出力するときの音量を表す情報である。

[0076] 参加者情報管理部 1 3 3 が管理する参加者情報は、音像定位処理部 1 3 4 に供給される。参加者情報管理部 1 3 3 が管理する参加者情報は、適宜、システム音声管理部 1 3 6、2chミックス処理部 1 3 7、音声送信部 1 3 8 等に対しても供給される。このように、参加者情報管理部 1 3 3 は、それぞれのユーザの仮想空間上の位置を管理する位置管理部として機能するとともに、背景音の設定を管理する背景音管理部として機能する。

[0077] 音像定位処理部 1 3 4 は、参加者情報管理部 1 3 3 から供給された位置情報に基づいて、それぞれのユーザの位置関係に応じたHRTFデータをHRTFデータ記憶部 1 3 5 から読み出して取得する。音像定位処理部 1 3 4 は、信号処理部 1 3 2 から供給された音声データに対して、HRTFデータ記憶部 1 3 5 から読み出したHRTFデータを用いた音像定位処理を行い、それぞれの聴取ユーザ用の音声データを生成する。

[0078] また、音像定位処理部 1 3 4 は、システム音声管理部 1 3 6 から供給されたシステム音声のデータに対して、所定のHRTFデータを用いた音像定位処理を行う。システム音声は、コミュニケーション管理サーバ 1 側で発生させ、

発話ユーザの音声とともに聴取ユーザに聴かせる音声である。システム音声には、例えば、BGMなどの背景音や、効果音が含まれる。システム音声は、ユーザの音声とは異なる音声である。

[0079] すなわち、コミュニケーション管理サーバ1においては、背景音や効果音などの、発話ユーザの音声以外の音声についても、オブジェクトオーディオとして処理が行われる。システム音声の音声データに対しても、仮想空間の所定の位置に音像を定位させるための音像定位処理が行われる。例えば、参加者の位置よりも遠い位置に音像を定位させるための音像定位処理が、背景音の音声データに対して施される。

[0080] 音像定位処理部134は、音像定位処理を行うことによって得られた音声データを2chミックス処理部137に出力する。2chミックス処理部137に対しては、発話ユーザの音声データと、適宜、システム音声の音声データが出力される。

[0081] HRTFデータ記憶部135は、仮想空間上のそれぞれの聴取位置を基準とした、複数の位置に対応するHRTFデータを記憶する。

[0082] システム音声管理部136は、システム音声を管理する。システム音声管理部136は、システム音声の音声データを音像定位処理部134に出力する。

[0083] 2chミックス処理部137は、音像定位処理部134から供給された音声データに対して2chミックス処理を行う。2chミックス処理が施されることにより、発話ユーザの音声とシステム音声のそれぞれのオーディオ信号Lとオーディオ信号Rの成分を含む、チャンネルベースのオーディオデータが生成される。2chミックス処理が施されることによって得られた音声データは音声送信部138に出力される。

[0084] 音声送信部138は、通信部109を制御し、2chミックス処理部137から供給された音声データをそれぞれの聴取ユーザが使用するクライアント端末2に送信する。

[0085] <クライアント端末2の構成>

図13は、クライアント端末2のハードウェア構成例を示すブロック図である。

- [0086] クライアント端末2は、制御部201に対して、メモリ202、音声入力デバイス203、音声出力デバイス204、操作部205、通信部206、ディスプレイ207、およびセンサ部208が接続されることによって構成される。
- [0087] 制御部201は、CPU、ROM、RAMなどにより構成される。制御部201は、クライアントプログラム201Aを実行することによって、クライアント端末2の全体の動作を制御する。クライアントプログラム201Aは、コミュニケーション管理サーバ1が管理するTele-communicationシステムを利用するためのプログラムである。クライアントプログラム201Aには、送信側の処理を実行する送信側モジュール201A-1と、受信側の処理を実行する受信側モジュール201A-2が含まれる。
- [0088] メモリ202は、フラッシュメモリなどにより構成される。メモリ202は、制御部201が実行するクライアントプログラム201Aなどの各種の情報を記憶する。
- [0089] 音声入力デバイス203は、マイクにより構成される。音声入力デバイス203により集音された音声は、マイク音声として制御部201に出力される。
- [0090] 音声出力デバイス204は、ヘッドホンやスピーカなどの機器により構成される。音声出力デバイス204は、制御部201から供給されたオーディオ信号に基づいて、会議の参加者の音声などを出力させる。
- [0091] 以下、適宜、音声入力デバイス203がマイクであるとして説明する。また、音声出力デバイス204がヘッドホンであるとして説明する。
- [0092] 操作部205は、各種のボタンや、ディスプレイ207に重ねて設けられたタッチパネルにより構成される。操作部205は、ユーザの操作の内容を表す情報を制御部201に出力する。
- [0093] 通信部206は、5G通信などの移動通信システムの無線通信に対応した

通信モジュール、無線LANなどに対応した通信モジュールである。通信部206は、基地局が出力する電波を受信し、ネットワーク11を介して、コミュニケーション管理サーバ1などの各種の装置との間で通信を行う。通信部206は、コミュニケーション管理サーバ1から送信されてきた情報を受信し、制御部201に出力する。また、通信部206は、制御部201から供給された情報をコミュニケーション管理サーバ1に送信する。

[0094] ディスプレイ207は、有機ELディスプレイ、LCDなどにより構成される。ディスプレイ207には、リモート会議画面などの各種の画面が表示される。

[0095] センサ部208は、RGBカメラ、デプスカメラ、ジャイロセンサ、加速度センサなどの各種のセンサにより構成される。センサ部208は、計測を行うことによって得られたセンサデータを制御部201に出力する。センサ部208により計測されたセンサデータに基づいて、ユーザの状況の認識などが適宜行われる。

[0096] 図14は、クライアント端末2の機能構成例を示すブロック図である。図14に示す機能部のうちの少なくとも一部は、図13の制御部201によりクライアントプログラム201Aが実行されることによって実現される。

[0097] クライアント端末2においては情報処理部211が実現される。情報処理部211は、音声処理部221、設定情報送信部222、ユーザ状況認識部223、および表示制御部224により構成される。

[0098] 情報処理部211は、音声受信部231、出力制御部232、マイク音声取得部233、および音声送信部234により構成される。

[0099] 音声受信部231は、通信部206を制御し、コミュニケーション管理サーバ1から送信されてきた音声データを受信する。音声受信部231により受信された音声データは出力制御部232に供給される。

[0100] 出力制御部232は、コミュニケーション管理サーバ1から送信されてきた音声データに応じた音声を音声出力デバイス204から出力させる。

[0101] マイク音声取得部233は、音声入力デバイス203を構成するマイクに

より集音されたマイク音声の音声データを取得する。マイク音声取得部 2 3 3 により取得されたマイク音声の音声データは音声送信部 2 3 4 に供給される。

[0102] 音声送信部 2 3 4 は、通信部 2 0 6 を制御し、マイク音声取得部 2 3 3 から供給されたマイク音声の音声データをコミュニケーション管理サーバ 1 に送信する。

[0103] 設定情報送信部 2 2 2 は、ユーザの操作に応じて、各種の設定の内容を表す設定情報を生成する。設定情報送信部 2 2 2 は、通信部 2 0 6 を制御し、設定情報をコミュニケーション管理サーバ 1 に送信する。

[0104] ユーザ状況認識部 2 2 3 は、センサ部 2 0 8 により計測されたセンサデータに基づいてユーザの状況を認識する。ユーザ状況認識部 2 2 3 は、通信部 2 0 6 を制御し、ユーザの状況を表す情報をコミュニケーション管理サーバ 1 に送信する。

[0105] 表示制御部 2 2 4 は、通信部 2 0 6 を制御することによってコミュニケーション管理サーバ 1 との間で通信を行い、コミュニケーション管理サーバ 1 から送信されてきた情報に基づいて、リモート会議画面をディスプレイ 2 0 7 に表示させる。

[0106] <<音像定位のユースケース>>

会議の参加者による発話音声を含む各種の音声の音像定位のユースケースについて説明する。

[0107] <発話ユーザのグルーピング>

複数の話題を聞きやすくするために、それぞれのユーザは、発話ユーザをグルーピングすることができる。発話ユーザのグルーピングは、クライアント端末 2 のディスプレイ 2 0 7 に GUI として表示される設定画面を用いて、会議が始まる前などの所定のタイミングで行われる。

[0108] 図 1 5 は、グループ設定画面の例を示す図である。

[0109] グループ設定画面上でのグループの設定は、例えば、参加者アイコンをドラッグ&ドロップで移動させることによって行われる。

- [0110] 図 1 5 の例においては、Group1を表す矩形領域 3 0 1 とGroup2を表す矩形領域 3 0 2 がグループ設定画面に表示されている。参加者アイコン 1 1 1 と参加者アイコン 1 1 2 が矩形領域 3 0 1 に移動され、参加者アイコン 1 1 3 が、カーソルで矩形領域 3 0 1 に移動中とされている。また、参加者アイコン 1 1 4 から 1 1 7 が矩形領域 3 0 2 に移動されている。
- [0111] 参加者アイコンが矩形領域 3 0 1 に移動された発話ユーザはGroup1に属するユーザとなり、矩形領域 3 0 2 に移動された発話ユーザはGroup2に属するユーザとなる。このような画面を用いて、それぞれの発話ユーザのグループが設定される。グループが割り当てられた領域に参加者アイコンを移動させるのではなく、複数の参加者アイコンを重ねることによりグループが形成されるようにしてもよい。
- [0112] 図 1 6 は、発話ユーザのグルーピングに関する処理の流れを示す図である。
- [0113] 図 1 5 のグループ設定画面を用いて設定されたグループを表す設定情報であるグループ設定情報は、矢印 A 1 に示すように、クライアント端末 2 からコミュニケーション管理サーバ 1 に送信される。
- [0114] 矢印 A 2, A 3 に示すようにクライアント端末 2 からマイク音声を送信されてきた場合、コミュニケーション管理サーバ 1 においては、グループ毎に異なるHRTFを用いて音像定位処理が行われる。例えば、グループ毎に異なる位置から聞こえるように、同じグループに属する発話ユーザの音声データに対して、同じHRTFデータを用いた音像定位処理が行われる。
- [0115] 音像定位処理により生成された音声データは、矢印 A 4 に示すようにそれぞれの聴取ユーザが使用するクライアント端末 2 に送信され、出力される。
- [0116] なお、図 1 6 において、複数のブロックを用いて最上段に示すマイク音声 # 1 から # N は、それぞれ、異なるクライアント端末 2 において検出された発話ユーザの音声である。また、1つのブロックを用いて最下段に示す音声出力は、1人の聴取ユーザが使用するクライアント端末 2 での出力を表す。
- [0117] 図 1 6 の左側に示すように、例えば、グループの設定とグループ設定情報

の送信に関する矢印A1で示される機能は、受信側モジュール201A-2により実現される。また、マイク音声の送信に関する矢印A2, A3で示される機能は、送信側モジュール201A-1により実現される。HRTFデータを用いた音像定位処理は、サーバプログラム101Aにより実現される。

[0118] 図17のフローチャートを参照して、発話ユーザのグルーピングに関するコミュニケーション管理サーバ1の制御処理について説明する。

[0119] コミュニケーション管理サーバ1の制御処理のうち、図8を参照して説明した内容と重複する内容については適宜説明を省略する。後述する図20等においても同様である。

[0120] ステップS101において、参加者情報管理部133（図11）は、それぞれのユーザにより設定された発話グループを表すグループ設定情報を受信する。クライアント端末2からは、発話ユーザのグループの設定が行われることに応じて、グループ設定情報が送信されてくる。参加者情報管理部133においては、クライアント端末2から送信されてきたグループ設定情報が、グループを設定したユーザの情報と紐付けて管理される。

[0121] ステップS102において、音声受信部131は、発話ユーザが使用するクライアント端末2から送信されてきた音声データを受信する。音声受信部131により受信された音声データは、信号処理部132を介して音像定位処理部134に供給される。

[0122] ステップS103において、音像定位処理部134は、同じグループに属する発話ユーザの音声データに対して、同じHRTFデータを用いた音像定位処理を行う。

[0123] ステップS104において、音声送信部138は、音像定位処理によって得られた音声データを聴取ユーザが使用するクライアント端末2に送信する。

[0124] 図15の例の場合、Group1に属する発話ユーザの音声データとGroup2に属する発話ユーザの音声データに対してそれぞれ異なるHRTFデータを用いた音像定位処理が行われる。また、そのグループ設定を行ったユーザ（聴取ユー

ザ) が使用するクライアント端末 2 においては、Group1 と Group2 のそれぞれのグループに属する発話ユーザの音声の音像が、異なる位置に定位して感じられることになる。

[0125] ユーザは、例えば同じ話題の会話をしているユーザ毎にグループを設定することにより、それぞれの話題を聞き取りやすくすることが可能となる。

[0126] 例えば、デフォルトの状態ではグループが作成されておらず、すべてのユーザを表す参加者アイコンが等間隔にレイアウトされる。この場合、グループ設定画面上における参加者アイコンのレイアウトに応じて、等間隔に離れた位置に音像が定位するように音像定位処理が行われる。

[0127] <位置情報の共有>

仮想空間上の位置情報がユーザ全員の間で共有されるようにしてもよい。図 15 等を参照して説明した例においては、それぞれのユーザが、他のユーザの音声の定位をカスタマイズすることができるのに対して、この例においては、それぞれのユーザが設定した自分の位置がユーザ全員の間で共通に用いられる。

[0128] この場合、それぞれのユーザは、クライアント端末 2 のディスプレイ 207 に GUI として表示される設定画面を用いて、会議が始まる前などの所定のタイミングで自分の位置を設定する。

[0129] 図 18 は、位置設定画面の例を示す図である。

[0130] 図 18 の位置設定画面に表示される三次元空間は仮想空間を表す。それぞれのユーザは、人の形をした参加者アイコンを移動させ、好みの位置を選択する。図 18 に示す参加者アイコン 131 から 134 はそれぞれユーザを表す。

[0131] 例えば、デフォルトの状態では、それぞれのユーザの位置として、仮想空間上の空いている位置が自動的に設定される。複数の聴取位置が設定されており、その中からユーザの位置が選択されるようにしてもよいし、仮想空間上の任意の位置を選択することができるようにしてもよい。

[0132] 図 19 は、位置情報の共有に関する処理の流れを示す図である。

- [0133] 図18の位置設定画面を用いて設定された仮想空間上の位置を表す位置情報は、矢印A11, A12に示すように、それぞれのユーザが使用するクライアント端末2からコミュニケーション管理サーバ1に送信される。コミュニケーション管理サーバ1においては、それぞれのユーザが自身の位置を設定することに同期して、それぞれのユーザの位置情報が共有の情報として管理される。
- [0134] 矢印A13, A14に示すようにクライアント端末2からマイク音声を送信されてきた場合、コミュニケーション管理サーバ1においては、共有の位置情報に基づいて、聴取ユーザとそれぞれの発話ユーザとの位置関係に応じたHRTFデータを用いて音像定位処理が行われる。
- [0135] 音像定位処理により生成された音声データは、矢印A15に示すように聴取ユーザが使用するクライアント端末2に送信され、出力される。
- [0136] クライアント端末2に設けられたカメラにより撮影された画像に基づいて、矢印A16に示すように聴取ユーザの頭部の位置の推定が行われる場合、位置情報のヘッドトラッキングが行われるようにしてもよい。聴取ユーザの頭部の位置の推定が、センサ部208を構成するジャイロセンサや加速度センサなどの他のセンサにより検出されたセンサデータに基づいて行われるようにしてもよい。
- [0137] 例えば聴取ユーザの頭が右方向に30度回転した場合、全てのユーザの位置を左方向に30度回転させることによって、それぞれのユーザの位置が補正され、補正後の位置に応じたHRTFデータを用いて音像定位処理が行われる。
- [0138] 図20のフローチャートを参照して、位置情報の共有に関するコミュニケーション管理サーバ1の制御処理について説明する。
- [0139] ステップS111において、参加者情報管理部133は、それぞれのユーザにより設定された位置を表す位置情報を受信する。それぞれのユーザが使用するクライアント端末2からは、仮想空間上の位置の設定が行われることに応じて、位置情報が送信されてくる。参加者情報管理部133においては

、クライアント端末2から送信されてきた位置情報が、それぞれのユーザの情報と紐付けて管理される。

[0140] ステップS 1 1 2において、参加者情報管理部 1 3 3は、それぞれのユーザの位置情報を共有の情報として管理する。

[0141] ステップS 1 1 3において、音声受信部 1 3 1は、発話ユーザが使用するクライアント端末2から送信されてきた音声データを受信する。

[0142] ステップS 1 1 4において、音像定位処理部 1 3 4は、共有の位置情報に基づいて、聴取ユーザとそれぞれの発話ユーザとの位置関係に応じたHRTFデータをHRTFデータ記憶部 1 3 5から読み出して取得する。音像定位処理部 1 3 4は、発話ユーザの音声データに対してHRTFデータを用いた音像定位処理を行う。

[0143] ステップS 1 1 5において、音声送信部 1 3 8は、音像定位処理によって得られた音声データを聴取ユーザが使用するクライアント端末2に送信する。

[0144] 以上の処理により、聴取ユーザが使用するクライアント端末2においては、発話ユーザの音声の音像が、それぞれの発話ユーザが設定した位置に定位して感じられることになる。

[0145] <背景音の設定>

発話ユーザの音声を聞きやすくするために、それぞれのユーザは、マイク音声に含まれる環境音を別の音声である背景音に変えることができる。背景音の設定は、クライアント端末2のディスプレイ 2 0 7にGUIとして表示される画面を用いて、会議が始まる前などの所定のタイミングで行われる。

[0146] 図 2 1 は、背景音の設定に用いられる画面の例を示す図である。

[0147] 背景音の設定は、例えば、リモート会議画面に表示されるメニューを用いて行われる。

[0148] 図 2 1 の例においては、背景音設定メニュー 3 2 1 がリモート会議画面の右上に表示されている。背景音設定メニュー 3 2 1 には、BGMなどの背景音のタイトルが複数表示される。ユーザは、背景音設定メニュー 3 2 1 に表示さ

れる音声の中から所定の音声を背景音として設定することができる。

[0149] なお、デフォルトの状態では背景音のオフが設定される。この場合、発話ユーザがいる空間の環境音がそのまま聞こえることになる。

[0150] 図 2 2 は、背景音の設定に関する処理の流れを示す図である。

[0151] 図 2 2 の画面を用いて設定された背景音を表す設定情報である背景音設定情報は、矢印 A 2 1 に示すように、クライアント端末 2 からコミュニケーション管理サーバ 1 に送信される。

[0152] 矢印 A 2 2, A 2 3 に示すようにクライアント端末 2 からマイク音声を送信されてきた場合、コミュニケーション管理サーバ 1 においては、それぞれのマイク音声から環境音が分離される。

[0153] 環境音が分離されることによって得られた発話ユーザの音声データに対しては、矢印 A 2 4 に示すように背景音が追加（合成）され、発話ユーザの音声データと、背景音の音声データのそれぞれに対して、位置関係に応じた HRTF データを用いた音像定位処理が行われる。例えば、発話ユーザの位置よりも遠い位置に音像を定位させるための音像定位処理が、背景音の音声データに対して施される。

[0154] 背景音の種類毎（タイトル毎）に異なる HRTF データが用いられるようにしてもよい。例えば、鳥のさえずりの背景音が選択された場合、音像を高い位置に定位させるための HRTF データが用いられ、波の音の背景音が選択された場合、音像を低い位置に定位させるための HRTF データが用いられる。このように、背景音の種類毎に HRTF データが用意される。

[0155] 音像定位処理により生成された音声データは、矢印 A 2 5 に示すように背景音の設定を行った聴取ユーザが使用するクライアント端末 2 に送信され、出力される。

[0156] 図 2 3 のフローチャートを参照して、背景音の設定に関するコミュニケーション管理サーバ 1 の制御処理について説明する。

[0157] ステップ S 1 2 1 において、参加者情報管理部 1 3 3 は、それぞれのユーザにより設定された背景音の設定内容を表す背景音設定情報を受信する。ク

クライアント端末２からは、背景音の設定が行われることに応じて背景音設定情報が送信されてくる。参加者情報管理部１３３においては、クライアント端末２から送信されてきた背景音設定情報が、背景音を設定したユーザの情報と紐付けて管理される。

[0158] ステップＳ１２２において、音声受信部１３１は、発話ユーザが使用するクライアント端末２から送信されてきた音声データを受信する。音声受信部１３１により受信された音声データは、信号処理部１３２に供給される。

[0159] ステップＳ１２３において、信号処理部１３２は、音声受信部１３１から供給された音声データから環境音の音声データを分離する。環境音の音声データを分離することによって得られた、発話ユーザの音声データは音像定位処理部１３４に供給される。

[0160] ステップＳ１２４において、システム音声管理部１３６は、聴取ユーザにより設定された背景音の音声データを音像定位処理部１３４に出力し、音像定位処理の対象の音声データとして追加する。

[0161] ステップＳ１２５において、音像定位処理部１３４は、聴取ユーザの位置と発話ユーザの位置との位置関係に応じたHRTFデータと、聴取ユーザの位置と背景音の位置（音像を定位させる位置）との位置関係に応じたHRTFデータをHRTFデータ記憶部１３５から読み出して取得する。音像定位処理部１３４は、発話ユーザの音声データに対して発話音声用のHRTFデータを用いた音像定位処理を行い、背景音の音声データに対して背景音用のHRTFデータを用いた音像定位処理を行う。

[0162] ステップＳ１２６において、音声送信部１３８は、音像定位処理によって得られた音声データを聴取ユーザが使用するクライアント端末２に送信する。以上の処理が、それぞれの聴取ユーザを対象として行われる。

[0163] 以上の処理により、聴取ユーザが使用するクライアント端末２においては、発話ユーザの音声の音像と、聴取ユーザ自身が選択した背景音の音像とが異なる位置に定位して感じられることになる。

[0164] 聴取ユーザは、発話ユーザの音声と、発話ユーザがいる環境の騒音などの

環境音が同じ位置から聞こえる場合に比べて、発話ユーザの音声を聞き取りやすくすることが可能となる。また、聴取ユーザは、好みの背景音を用いて会話を行うことができる。

[0165] 背景音の追加がコミュニケーション管理サーバ1側で行われるのではなく、クライアント端末2側において、受信側モジュール201A-2により行われるようにしてもよい。

[0166] <背景音の共有>

BGMなどの背景音の設定がユーザ全員の間で共有されるようにしてもよい。図21等を参照して説明した例においては、それぞれのユーザが、他のユーザの音声に合成させる背景音を個別に設定してカスタマイズすることができるのに対して、この例においては、任意のユーザが設定した背景音が、他のユーザが聴取ユーザとなる場合の背景音として共通に用いられる。

[0167] この場合、任意のユーザは、クライアント端末2のディスプレイ207にGUIとして表示される設定画面を用いて、会議が始まる前などの所定のタイミングで背景音を設定する。背景音の設定は、図21に示す画面と同様の画面を用いて行われる。例えば、背景音設定メニューには、背景音の共有のオン／オフを設定するための表示も設けられる。

[0168] なお、デフォルトの状態では背景音の共有はオフとなる。この場合、背景音の合成なしに、発話ユーザの音声そのまま聞こえることになる。

[0169] 図24は、背景音の設定に関する処理の流れを示す図である。

[0170] 背景音の共有のオン／オフと、共有のオンが設定された場合に選択された背景音を表す設定情報である背景音設定情報は、矢印A31に示すように、クライアント端末2からコミュニケーション管理サーバ1に送信される。

[0171] 矢印A32, A33に示すようにクライアント端末2からマイク音声を送信されてきた場合、コミュニケーション管理サーバ1においては、それぞれのマイク音声から環境音が分離される。環境音の分離が行われなくともよい。

[0172] 環境音が分離されることによって得られた発話ユーザの音声データに対し

ては、背景音が追加され、発話ユーザの音声データと、背景音の音声データのそれぞれに対して、位置関係に応じたHRTFデータを用いた音像定位処理が行われる。例えば、発話ユーザの位置よりも遠い位置に音像を定位させるための音像定位処理が、背景音の音声データに対して施される。

[0173] 音像定位処理により生成された音声データは、矢印A34, A35に示すようにそれぞれの聴取ユーザが使用するクライアント端末2に送信され、出力される。それぞれの聴取ユーザが使用するクライアント端末2においては、共通の背景音が、発話ユーザの音声とともに出力される。

[0174] 図25のフローチャートを参照して、背景音の共有に関するコミュニケーション管理サーバ1の制御処理について説明する。

[0175] 図25に示す制御処理は、背景音の設定がそれぞれのユーザにより個別に行われるのではなく、1人のユーザにより行われる点を除いて、図23を参照して説明した処理と同様の処理である。重複する説明については省略する。

[0176] すなわち、ステップS131において、参加者情報管理部133は、任意のユーザにより設定された背景音の設定内容を表す背景音設定情報を受信する。参加者情報管理部133においては、クライアント端末2から送信されてきた背景音設定情報が、ユーザ全員のユーザの情報と紐付けて管理される。

[0177] ステップS132において、音声受信部131は、発話ユーザが使用するクライアント端末2から送信されてきた音声データを受信する。音声受信部131により受信された音声データは、信号処理部132に供給される。

[0178] ステップS133において、信号処理部132は、音声受信部131から供給された音声データから環境音の音声データを分離する。環境音の音声データを分離することによって得られた、発話ユーザの音声データは音像定位処理部134に供給される。

[0179] ステップS134において、システム音声管理部136は、共通の背景音の音声データを音像定位処理部134に出力し、音像定位処理の対象の音声

データとして追加する。

[0180] ステップS 1 3 5において、音像定位処理部 1 3 4 は、聴取ユーザの位置と発話ユーザの位置との位置関係に応じたHRTFデータと、聴取ユーザの位置と背景音の位置との位置関係に応じたHRTFデータをHRTFデータ記憶部 1 3 5から読み出して取得する。音像定位処理部 1 3 4 は、発話ユーザの音声データに対して発話音声用のHRTFデータを用いた音像定位処理を行い、背景音の音声データに対して背景音用のHRTFデータを用いた音像定位処理を行う。

[0181] ステップS 1 3 6において、音声送信部 1 3 8 は、音像定位処理によって得られた音声データを聴取ユーザが使用するクライアント端末 2 に送信する。

[0182] 以上の処理により、聴取ユーザが使用するクライアント端末 2 においては、発話ユーザの音声の音像と、会議において共通に使用されている背景音の音像が異なる位置に定位して感じられることになる。

[0183] 背景音の共有が以下のようにして行われるようにしてもよい。

[0184] (A) 複数人が同時に同じ講演を仮想の講演会場で聴く場合、講演者の音声を共通の背景音として遠くに定位させ、ユーザの音声を近くに定位させるように音像定位処理が行われる。発話ユーザの音声に対しては、それぞれのユーザの位置の関係と空間音響を考慮したレンダリングなどの音像定位処理が行われる。

[0185] (B) 映画コンテンツを複数人が同時に仮想の映画館で鑑賞する場合、共通の背景音となる映画コンテンツの音声をスクリーンの近くに定位させるように音像定位処理が行われる。映画コンテンツの音声に対しては、それぞれのユーザが自分の席として選択した映画館内の席の位置とスクリーンの位置との関係と、映画館の音響を考慮したレンダリングなどの音像定位処理が行われる。

[0186] (C) あるユーザがいる空間の環境音がマイク音声から分離され、共通の背景音として用いられる。この場合、それぞれのユーザは、発話ユーザの音声とともに、他のユーザがいる空間の環境音と同じ音を聴くことになる。これ

により、任意の空間の環境音をユーザ全員に共有させることが可能となる。

[0187] <音像定位処理の動的切り替え>

レンダリングなどを含むオブジェクトオーディオの処理である音像定位処理をコミュニケーション管理サーバ1側で行うのか、クライアント端末2側で行うのかが動的に切り替えられる。

[0188] この場合、コミュニケーション管理サーバ1の図11に示す構成のうちの、少なくとも、音像定位処理部134、HRTFデータ記憶部135、2chミックス処理部137と同様の構成が、クライアント端末2にも設けられる。音像定位処理部134、HRTFデータ記憶部135、2chミックス処理部137と同様の構成は、例えば、受信側モジュール201A-2によって実現される。

[0189] 聴取ユーザの位置情報などの、音像定位処理に用いるパラメータの設定が会議中に変更され、その変更をリアルタイムで音像定位処理に反映させる場合、音像定位処理はクライアント端末2側で行われる。音像定位処理がローカルで行われることにより、パラメータの変更に対するレスポンスを早くすることが可能となる。

[0190] 一方、パラメータの設定変更が一定時間以上ない場合、音像定位処理はコミュニケーション管理サーバ1側で行われる。音像定位処理がサーバ上で行われることにより、コミュニケーション管理サーバ1-クライアント端末2間のデータ通信量を抑えることが可能となる。

[0191] 図26は、音像定位処理の動的切り替えに関する処理の流れを示す図である。

[0192] 音像定位処理がクライアント端末2側で行われる場合、矢印A101、A102に示すようにクライアント端末2から送信されてきたマイク音声は、矢印A103に示すように、そのまま、クライアント端末2に送信される。マイク音声の送信元となるクライアント端末2は発話ユーザが使用するクライアント端末2であり、マイク音声の送信先となるクライアント端末2は聴取ユーザが使用するクライアント端末2である。

[0193] 聴取ユーザの位置などの、音像の定位に関するパラメータの設定が矢印A

104に示すように聴取ユーザにより変更された場合、設定の変更をリアルタイムで反映して、コミュニケーション管理サーバ1から送信されてきたマイク音声に対して音像定位処理が行われる。

[0194] クライアント端末2側での音像定位処理により生成された音声データに応じた音声、矢印A105に示すように出力される。

[0195] クライアント端末2においては、パラメータの設定の変更内容が保存され、変更内容を表す情報が矢印A106に示すようにコミュニケーション管理サーバ1に送信される。

[0196] 音像定位処理がコミュニケーション管理サーバ1側で行われる場合、矢印A107、A108に示すようにクライアント端末2から送信されてきたマイク音声に対しては、変更後のパラメータを反映して、音像定位処理が行われる。

[0197] 音像定位処理により生成された音声データは、矢印A109に示すように聴取ユーザが使用するクライアント端末2に送信され、出力される。

[0198] 図27のフローチャートを参照して、音像定位処理の動的切り替えに関するコミュニケーション管理サーバ1の制御処理について説明する。

[0199] ステップS201において、パラメータの設定変更が一定時間以上ないかが判定される。この判定は、例えば、聴取ユーザが使用するクライアント端末2から送信されてくる情報に基づいて、参加者情報管理部133により行われる。

[0200] パラメータの設定変更があるとステップS201において判定された場合、ステップS202において、音声送信部138は、参加者情報管理部133により受信された発話ユーザの音声データを、そのまま、聴取ユーザが使用するクライアント端末2に送信する。送信される音声データは、オブジェクトオーディオのデータとなる。

[0201] クライアント端末2においては、変更後の設定を用いて音像定位処理が行われ、音声の出力が行われる。また、変更後の設定の内容を表す情報がコミュニケーション管理サーバ1に対して送信される。

[0202] ステップS203において、参加者情報管理部133は、クライアント端末2から送信されてきた、設定変更の内容を表す情報を受信する。クライアント端末2から送信されてきた情報に基づいて、聴取ユーザの位置情報の更新などが行われた後、ステップS201に戻り、それ以降の処理が行われる。コミュニケーション管理サーバ1側で行われる音像定位処理は、更新後の位置情報に基づいて行われる。

[0203] 一方、パラメータの設定変更がないとステップS201において判定された場合、ステップS204において、コミュニケーション管理サーバ1側の音像定位処理が行われる。ステップS204において行われる処理は、基本的には、図8を参照して説明した処理と同様の処理である。

[0204] 以上の処理が、位置の変更だけでなく、背景音の設定の変更などの、他のパラメータが変更された場合にも行われる。

[0205] <音響設定の管理>

背景音に適した音響設定がデータベース化され、コミュニケーション管理サーバ1において管理されるようにしてもよい。例えば、背景音の種類毎に、音像を定位させる位置として適した位置が設定され、設定された位置に応じたHRTFデータが保存される。リバーブなどの、他の音響設定に関するパラメータが保存されるようにしてもよい。

[0206] 図28は、音響設定の管理に関する処理の流れを示す図である。

[0207] 発話ユーザの音声に背景音を合成させる場合、コミュニケーション管理サーバ1においては、背景音が再生され、矢印A121に示すように、背景音に適したHRTFデータなどの音響設定を用いて音像定位処理が行われる。

[0208] 音像定位処理により生成された音声データは、矢印A122に示すように聴取ユーザが使用するクライアント端末2に送信され、出力される。

[0209] <<変形例>>

複数のユーザにより行われる会話がリモート会議での会話であるものとしたが、食事の場面での会話、講演会での会話などの、複数人がオンライン経由で参加する会話であれば、様々な種類の会話に上述した技術は適用可能で

ある。

[0210] ・プログラムについて

上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、汎用のパーソナルコンピュータなどにインストールされる。

[0211] インストールされるプログラムは、光ディスク（CD-ROM(Compact Disc-Read Only Memory)、DVD(Digital Versatile Disc)等）や半導体メモリなどよりなる図10に示されるリムーバブルメディア111に記録して提供される。また、ローカルエリアネットワーク、インターネット、デジタル放送といった、有線または無線の伝送媒体を介して提供されるようにしてもよい。プログラムは、ROM102や記憶部108に、あらかじめインストールしておくことができる。

[0212] なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。

[0213] 明細書において、システムとは、複数の構成要素（装置、モジュール（部品）等）の集合を意味し、すべての構成要素が同一筐体中にあるか否かは問わない。したがって、別個の筐体に収納され、ネットワークを介して接続されている複数の装置、及び、1つの筐体の中に複数のモジュールが収納されている1つの装置は、いずれも、システムである。

[0214] 本明細書に記載された効果はあくまで例示であって限定されるものではなく、また他の効果があってもよい。

[0215] 本技術の実施の形態は、上述した実施の形態に限定されるものではなく、本技術の要旨を逸脱しない範囲において種々の変更が可能である。音声出力デバイスとしてヘッドホンまたはスピーカが用いられるものとしたが、他の

デバイスが用いられるようにしてもよい。例えば、通常のイヤホン（インナーイヤーヘッドホン）や、環境音の取り込みが可能な開放型のイヤホンが音声出力デバイスとして用いられるようにすることが可能である。

[0216] 例えば、本技術は、1つの機能をネットワークを介して複数の装置で分担、共同して処理するクラウドコンピューティングの構成をとることができる。

[0217] また、上述のフローチャートで説明した各ステップは、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0218] さらに、1つのステップに複数の処理が含まれる場合には、その1つのステップに含まれる複数の処理は、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0219] ・構成の組み合わせ例

本技術は、以下のような構成をとることもできる。

[0220] (1)

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶する記憶部と、

ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う音像定位処理部と

を備える情報処理装置。

(2)

前記音像定位処理部は、聴取者となる前記参加者の位置と発話者となる前記参加者の位置との関係に応じた前記HRTFデータを用いて、前記発話者の音声データに対して前記音像定位処理を行う

前記(1)に記載の情報処理装置。

(3)

前記音像定位処理を行うことによって得られた前記発話者の音声データを、それぞれの前記聴取者が使用する端末に送信する送信処理部をさらに備え

る

前記（２）に記載の情報処理装置。

（４）

前記参加者が使用する端末に表示された画面上における、前記参加者を視覚的に表す視覚情報の位置に基づいて、それぞれの前記参加者の仮想空間上の位置を管理する位置管理部をさらに備える

前記（１）から（３）のいずれかに記載の情報処理装置。

（５）

前記位置管理部は、前記参加者による設定に従って前記参加者のグループを形成し、

前記音像定位処理部は、同じ前記グループに属する前記参加者の音声データに対して、同じ前記HRTFデータを用いた前記音像定位処理を行う

前記（４）に記載の情報処理装置。

（６）

前記音像定位処理部は、前記参加者の音声とは異なる音である背景音のデータに対して、仮想空間上の所定の位置に対応する前記HRTFデータを用いた前記音像定位処理を行い、

前記送信処理部は、前記音像定位処理によって得られた前記背景音のデータを、前記発話者の音声データとともに、前記聴取者が使用する端末に送信する

前記（３）に記載の情報処理装置。

（７）

前記参加者による設定に従って、前記背景音を選択する背景音管理部をさらに備える

前記（６）に記載の情報処理装置。

（８）

前記送信処理部は、前記背景音のデータを、前記背景音を選択した前記聴取者が使用する端末に送信する

前記（７）に記載の情報処理装置。

（９）

前記送信処理部は、前記背景音のデータを、前記背景音を選択した前記参加者を含む全ての前記参加者が使用する端末に送信する

前記（７）に記載の情報処理装置。

（１０）

それぞれの前記参加者の仮想空間上の位置を、全ての前記参加者の間で共通に用いられる位置として管理する位置管理部をさらに備える

前記（１）に記載の情報処理装置。

（１１）

情報処理装置が、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、

ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う

情報処理方法。

（１２）

コンピュータに、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、

ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う

処理を実行させるプログラム。

（１３）

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う情報処理装置から送信されてきた、前記音像定位処理を行うことによって得られた

発話者となる前記参加者の音声データを受信し、前記発話者の音声を入力する音声受信部を備える

情報処理端末。

(14)

前記情報処理端末のユーザの音声データを、前記発話者の音声データとして前記情報処理装置に送信する音声送信部をさらに備える

前記(13)に記載の情報処理端末。

(15)

前記参加者を視覚的に表す視覚情報を、それぞれの前記参加者の仮想空間上の位置に対応する位置に表示させる表示制御部をさらに備える

前記(13)または(14)に記載の情報処理端末。

(16)

前記情報処理端末のユーザにより設定された前記参加者のグループを表す設定情報を前記情報処理装置に送信する設定情報生成部をさらに備え、

前記音声受信部は、同じ前記グループに属する前記参加者の音声データに対して、同じ前記HRTFデータを用いた前記音像定位処理を行うことによって前記情報処理装置において得られた前記発話者の音声データを受信する

前記(13)から(15)のいずれかに記載の情報処理端末。

(17)

前記情報処理端末のユーザにより選択された、前記参加者の音声とは異なる音である背景音の種類を表す設定情報を前記情報処理装置に送信する設定情報生成部をさらに備え、

前記音声受信部は、仮想空間上の所定の位置に対応する前記HRTFデータを用いた前記音像定位処理を前記背景音のデータに対して行うことによって前記情報処理装置において得られた前記背景音のデータを、前記発話者の音声データとともに受信する

前記(13)から(15)のいずれかに記載の情報処理端末。

(18)

情報処理端末が、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う情報処理装置から送信されてきた、前記音像定位処理を行うことによって得られた発話者となる前記参加者の音声データを受信し、

前記発話者の音声を出力する

情報処理方法。

(19)

コンピュータに、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う情報処理装置から送信されてきた、前記音像定位処理を行うことによって得られた発話者となる前記参加者の音声データを受信し、

前記発話者の音声を出力する

処理を実行させるプログラム。

符号の説明

[0221] 1 コミュニケーション管理サーバ, 2 A~2 D クライアント端末,
1 2 1 情報処理部, 1 3 1 音声受信部, 1 3 2 信号処理部,
1 3 3 参加者情報管理部, 1 3 4 音像定位処理部, 1 3 5 HRTFデータ記憶部,
1 3 6 システム音声管理部, 1 3 7 2chミックス処理部,
, 1 3 8 音声送信部, 2 0 1 制御部, 2 1 1 情報処理部, 2
2 1 音声処理部, 2 2 2 設定情報送信部, 2 2 3 ユーザ状況認識部,
2 3 1 音声受信部, 2 3 3 マイク音声取得部

請求の範囲

- [請求項1] 聴取位置を基準とした複数の位置に対応するHRTFデータを記憶する記憶部と、
- ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う音像定位処理部と
- を備える情報処理装置。
- [請求項2] 前記音像定位処理部は、聴取者となる前記参加者の位置と発話者となる前記参加者の位置との関係に応じた前記HRTFデータを用いて、前記発話者の音声データに対して前記音像定位処理を行う
- 請求項1に記載の情報処理装置。
- [請求項3] 前記音像定位処理を行うことによって得られた前記発話者の音声データを、それぞれの前記聴取者が使用する端末に送信する送信処理部をさらに備える
- 請求項2に記載の情報処理装置。
- [請求項4] 前記参加者が使用する端末に表示された画面上における、前記参加者を視覚的に表す視覚情報の位置に基づいて、それぞれの前記参加者の仮想空間上の位置を管理する位置管理部をさらに備える
- 請求項1に記載の情報処理装置。
- [請求項5] 前記位置管理部は、前記参加者による設定に従って前記参加者のグループを形成し、
- 前記音像定位処理部は、同じ前記グループに属する前記参加者の音声データに対して、同じ前記HRTFデータを用いた前記音像定位処理を行う
- 請求項4に記載の情報処理装置。
- [請求項6] 前記音像定位処理部は、前記参加者の音声とは異なる音である背景音のデータに対して、仮想空間上の所定の位置に対応する前記HRTFデータを用いた前記音像定位処理を行い、

前記送信処理部は、前記音像定位処理によって得られた前記背景音のデータを、前記発話者の音声データとともに、前記聴取者が使用する端末に送信する

請求項 3 に記載の情報処理装置。

[請求項7] 前記参加者による設定に従って、前記背景音を選択する背景音管理部をさらに備える

請求項 6 に記載の情報処理装置。

[請求項8] 前記送信処理部は、前記背景音のデータを、前記背景音を選択した前記聴取者が使用する端末に送信する

請求項 7 に記載の情報処理装置。

[請求項9] 前記送信処理部は、前記背景音のデータを、前記背景音を選択した前記参加者を含む全ての前記参加者が使用する端末に送信する

請求項 7 に記載の情報処理装置。

[請求項10] それぞれの前記参加者の仮想空間上の位置を、全ての前記参加者の間で共通に用いられる位置として管理する位置管理部をさらに備える

請求項 1 に記載の情報処理装置。

[請求項11] 情報処理装置が、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、
ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う

情報処理方法。

[請求項12] コンピュータに、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、
ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う

処理を実行させるプログラム。

- [請求項13] 聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像定位処理を行う情報処理装置から送信されてきた、前記音像定位処理を行うことによって得られた発話者となる前記参加者の音声データを受信し、前記発話者の音声を出力する音声受信部を備える
情報処理端末。
- [請求項14] 前記情報処理端末のユーザの音声データを、前記発話者の音声データとして前記情報処理装置に送信する音声送信部をさらに備える
請求項13に記載の情報処理端末。
- [請求項15] 前記参加者を視覚的に表す視覚情報を、それぞれの前記参加者の仮想空間上の位置に対応する位置に表示させる表示制御部をさらに備える
請求項13に記載の情報処理端末。
- [請求項16] 前記情報処理端末のユーザにより設定された前記参加者のグループを表す設定情報を前記情報処理装置に送信する設定情報生成部をさらに備え、
前記音声受信部は、同じ前記グループに属する前記参加者の音声データに対して、同じ前記HRTFデータを用いた前記音像定位処理を行うことによって前記情報処理装置において得られた前記発話者の音声データを受信する
請求項13に記載の情報処理端末。
- [請求項17] 前記情報処理端末のユーザにより選択された、前記参加者の音声とは異なる音である背景音の種類を表す設定情報を前記情報処理装置に送信する設定情報生成部をさらに備え、
前記音声受信部は、仮想空間上の所定の位置に対応する前記HRTFデータを用いた前記音像定位処理を前記背景音のデータに対して行うことによって前記情報処理装置において得られた前記背景音のデータを

、前記発話者の音声データとともに受信する

請求項 13 に記載の情報処理端末。

[請求項18]

情報処理端末が、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、
ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像
定位処理を行う情報処理装置から送信されてきた、前記音像定位処理
を行うことによって得られた発話者となる前記参加者の音声データを
受信し、

前記発話者の音声を出力する

情報処理方法。

[請求項19]

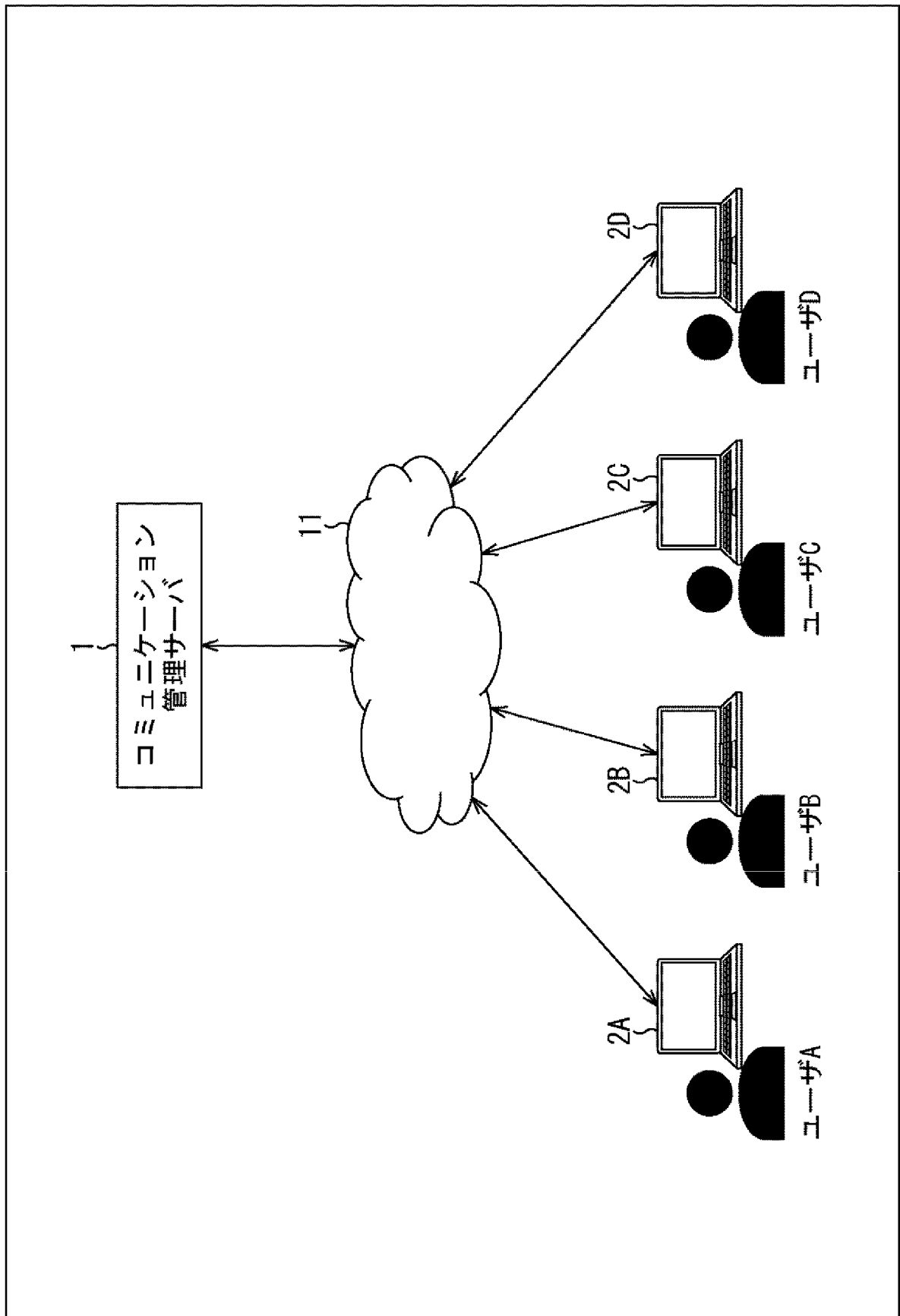
コンピュータに、

聴取位置を基準とした複数の位置に対応するHRTFデータを記憶し、
ネットワークを介して参加する会話の参加者の仮想空間上の位置に対応する前記HRTFデータと、前記参加者の音声データとに基づいて音像
定位処理を行う情報処理装置から送信されてきた、前記音像定位処理
を行うことによって得られた発話者となる前記参加者の音声データを
受信し、

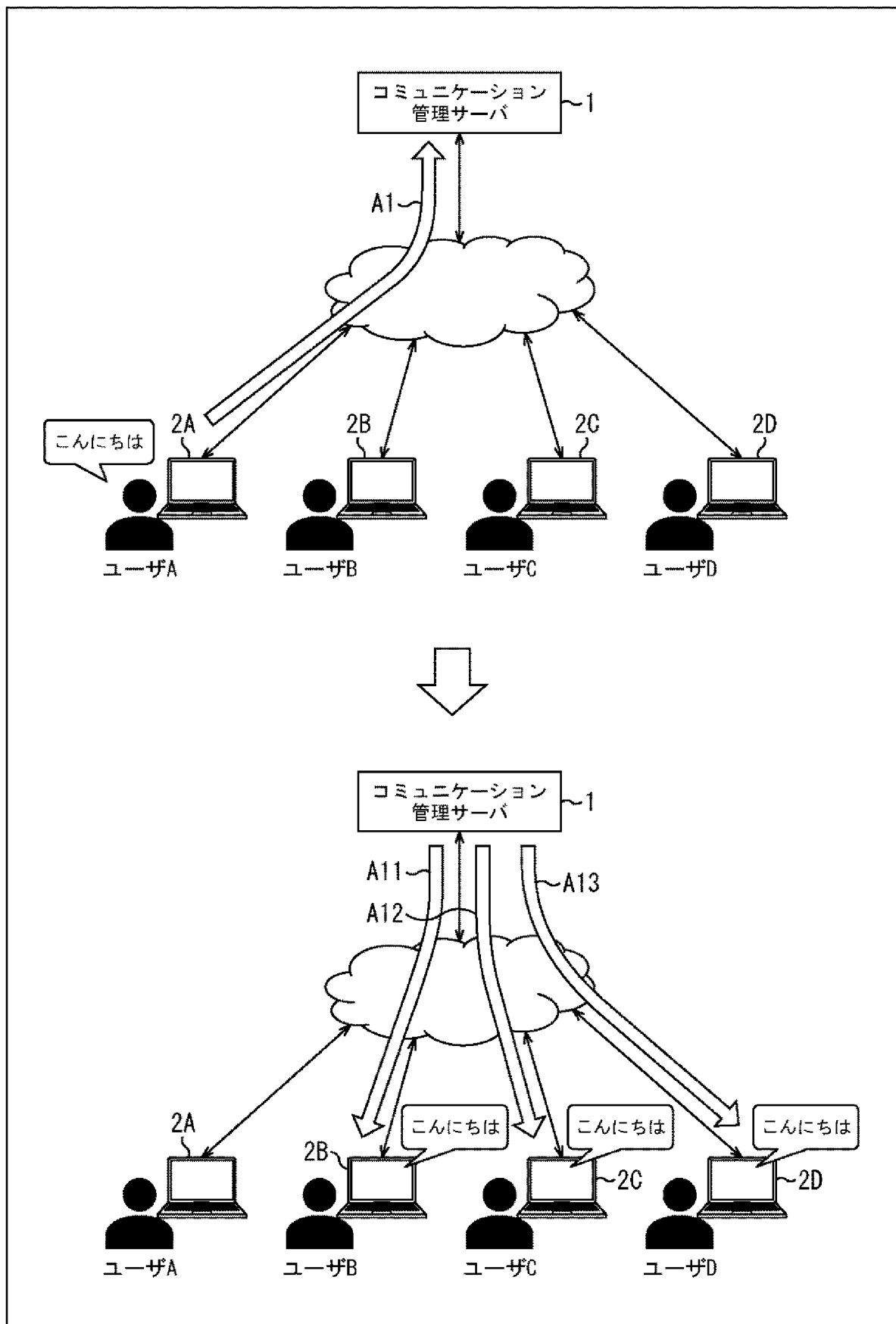
前記発話者の音声を出力する

処理を実行させるプログラム。

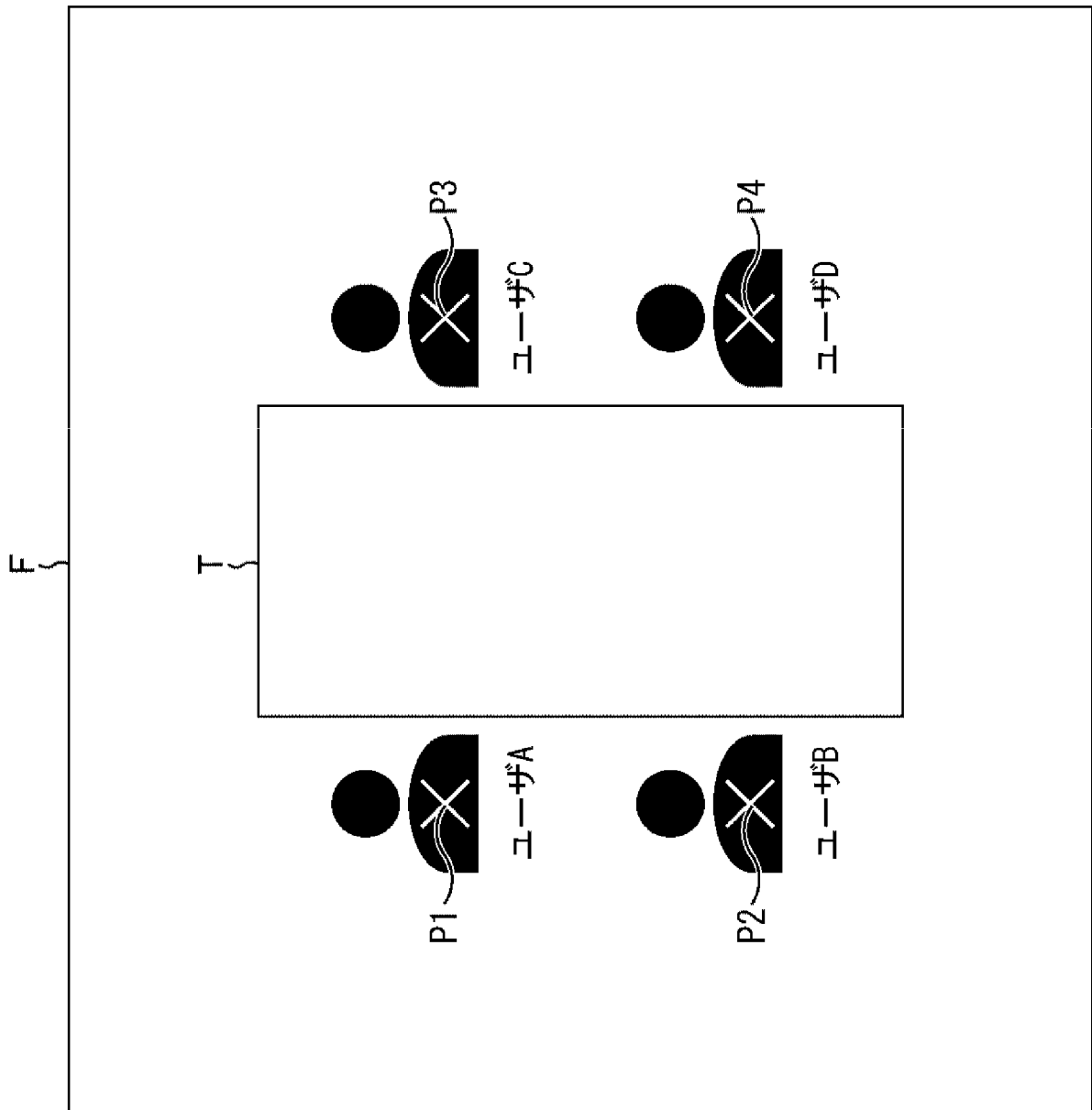
[図1]



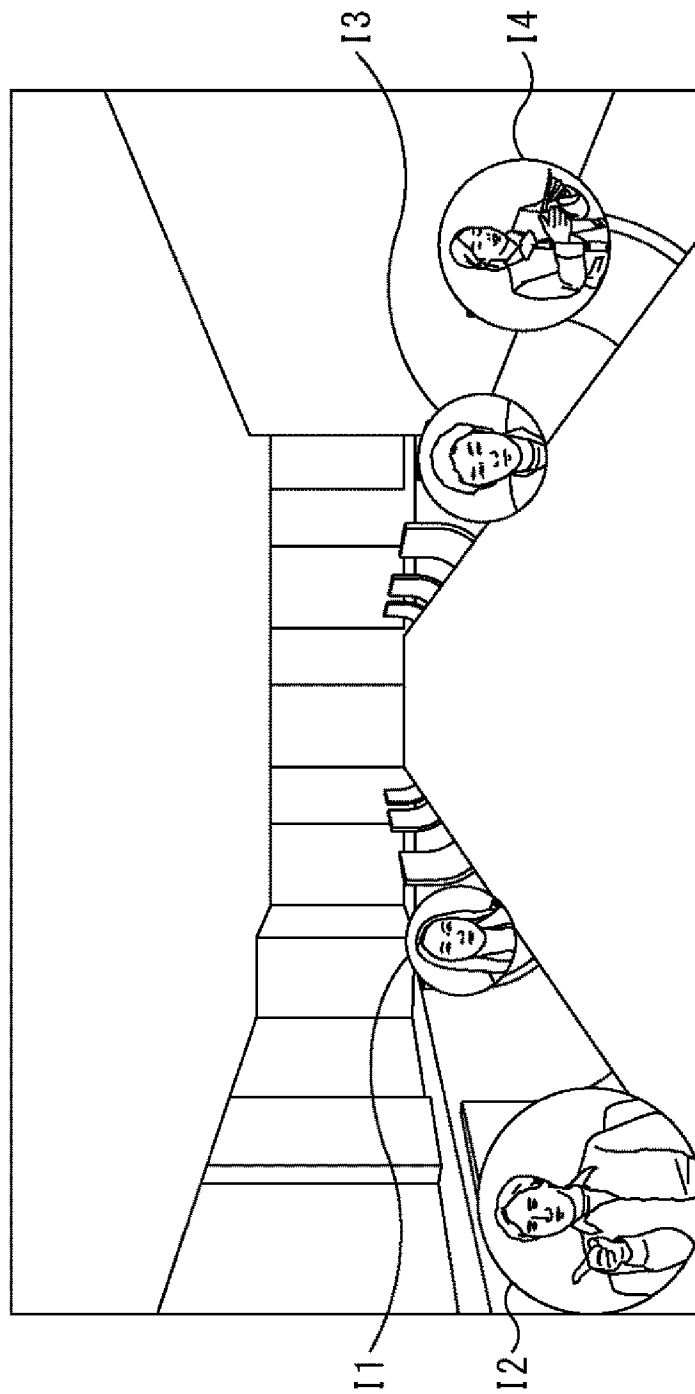
[図2]



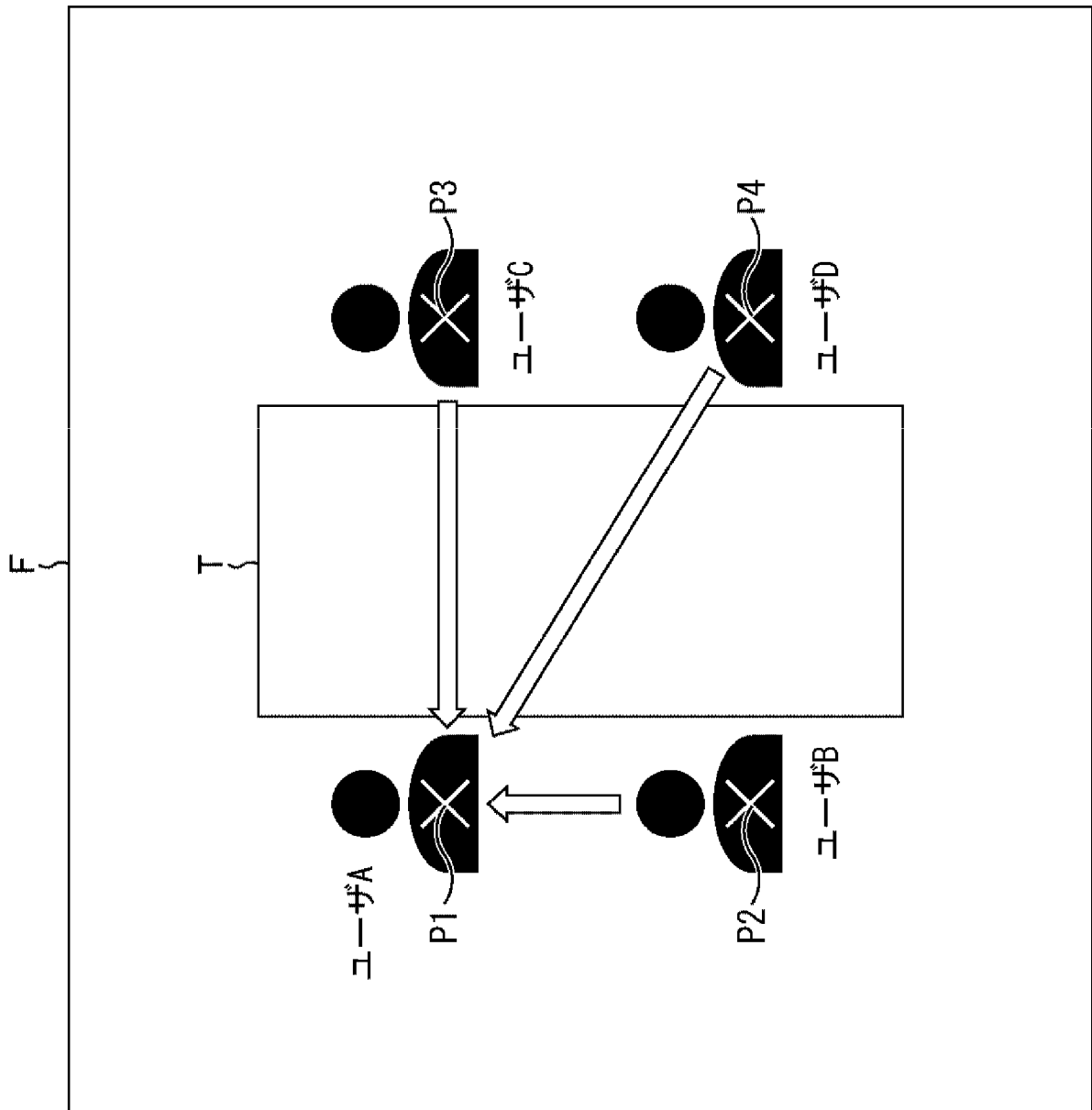
[図3]



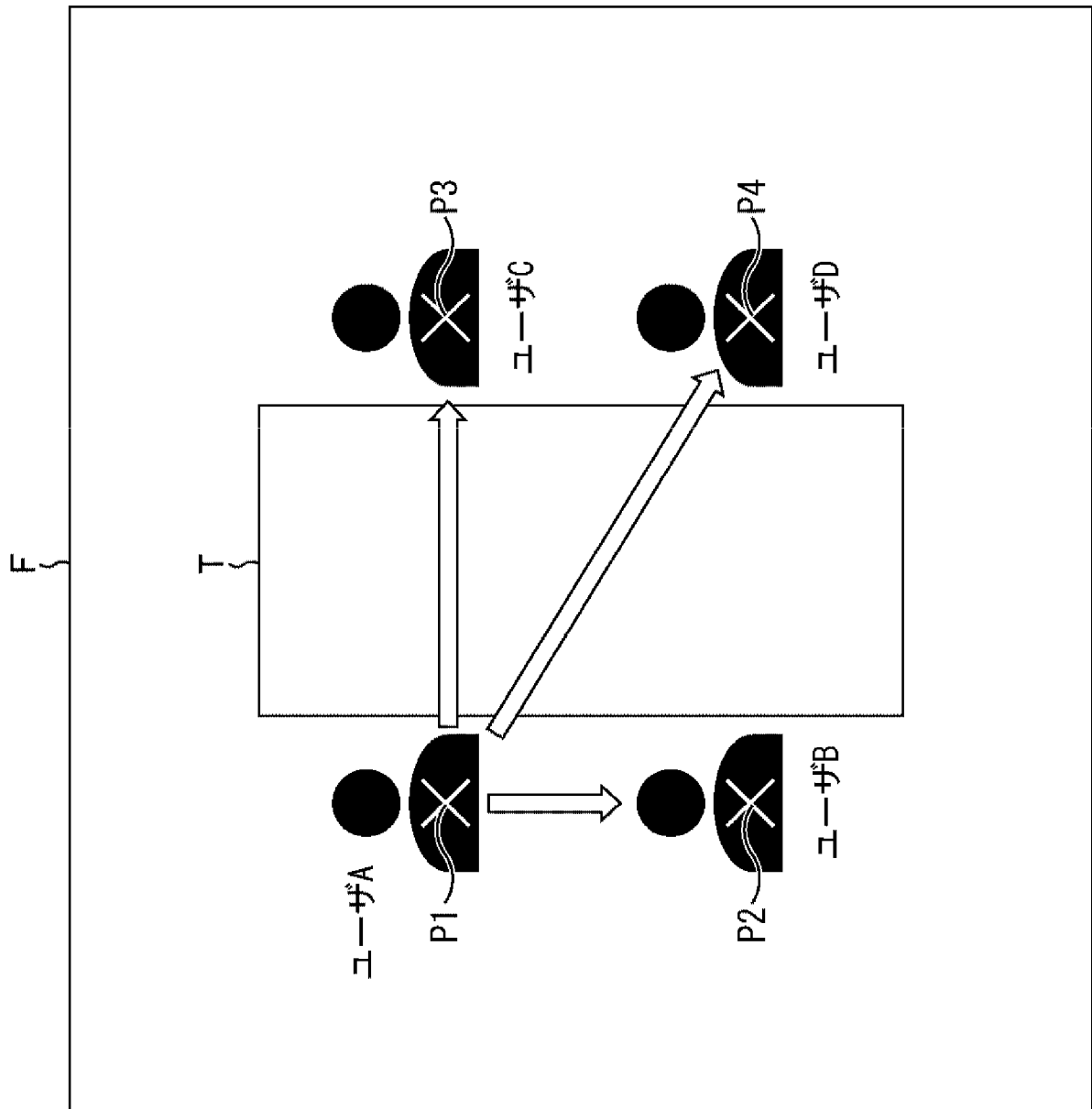
[図4]



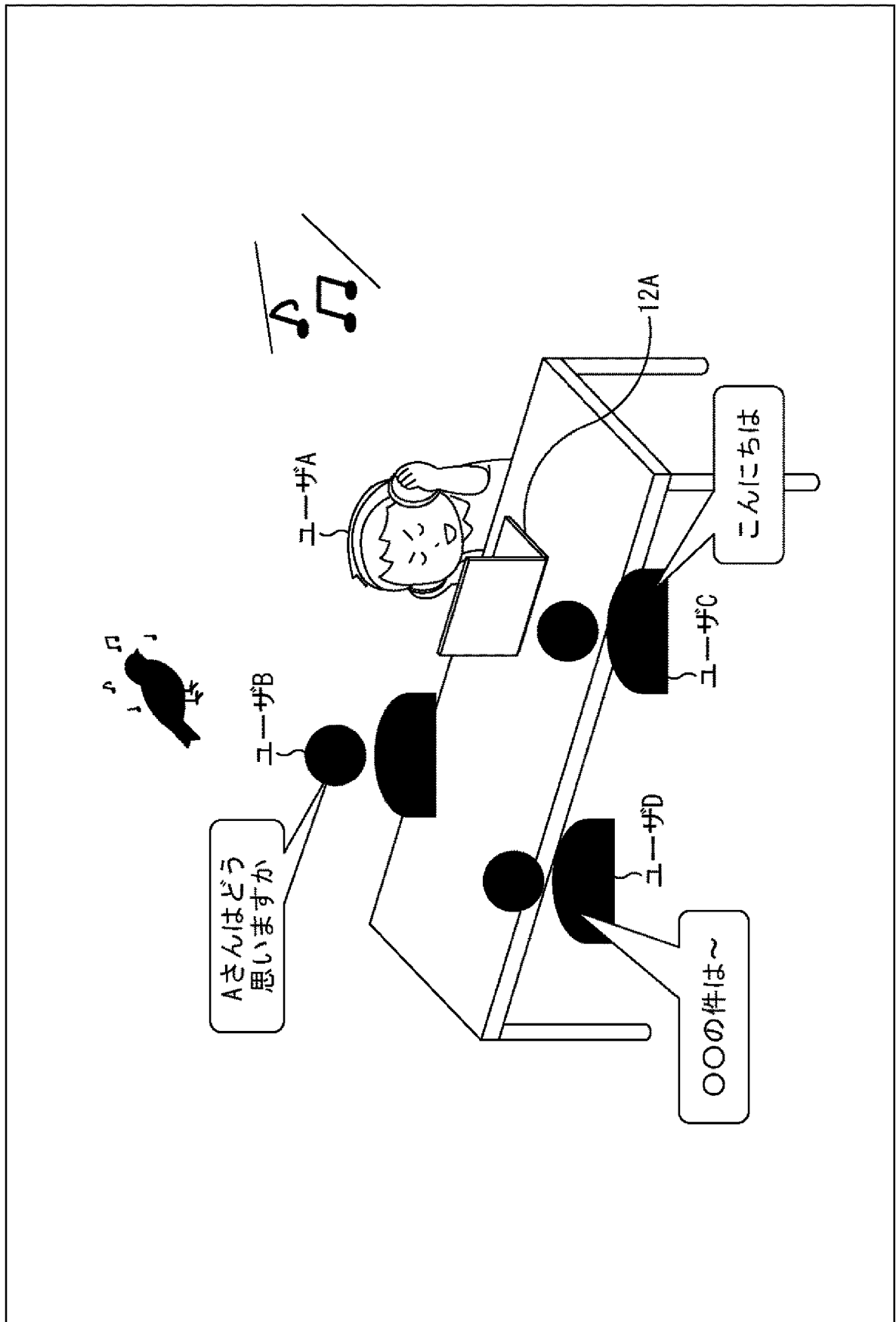
[図5]



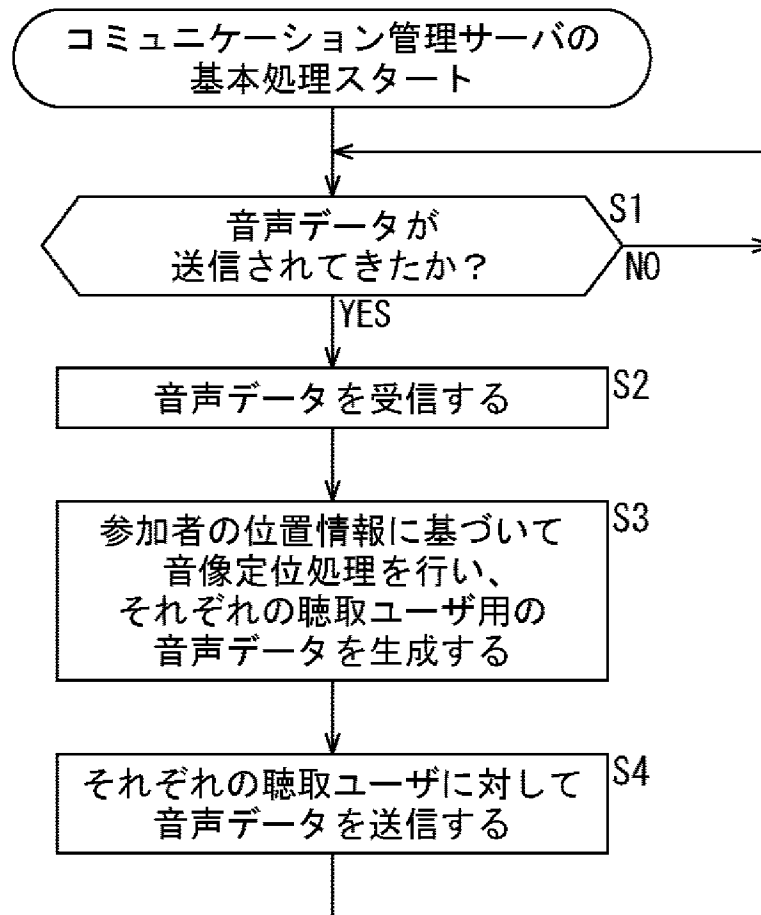
[図6]



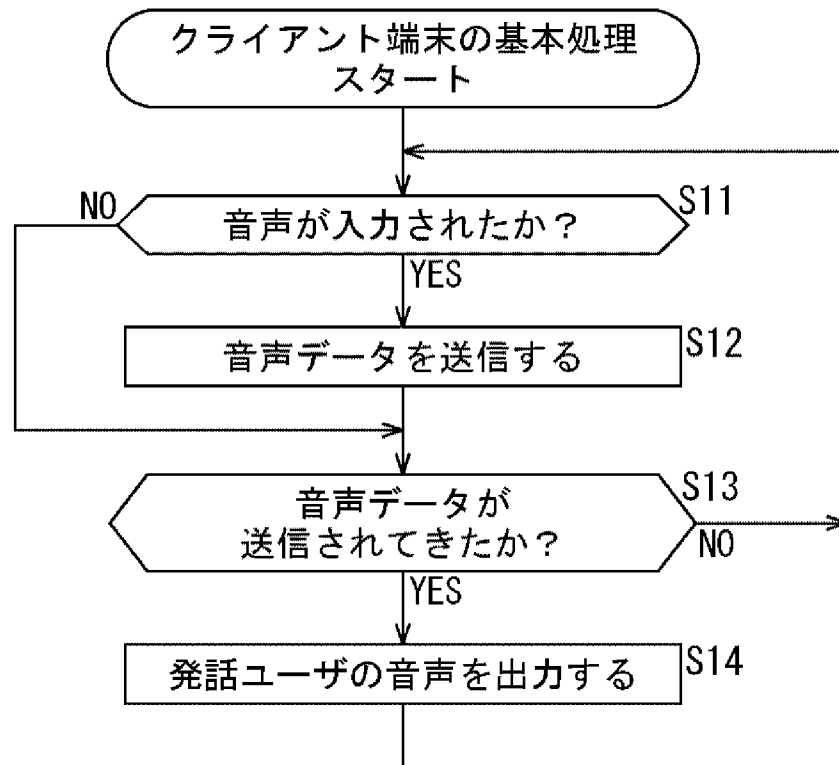
[図7]



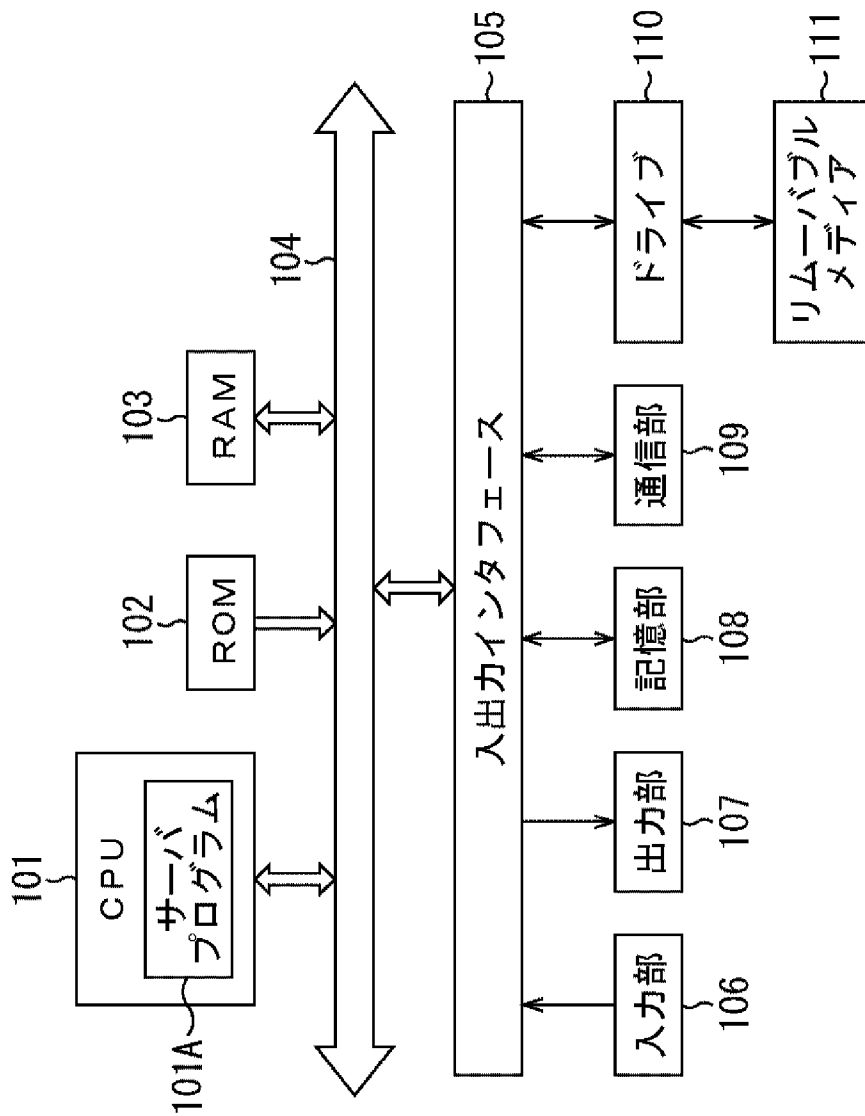
[図8]



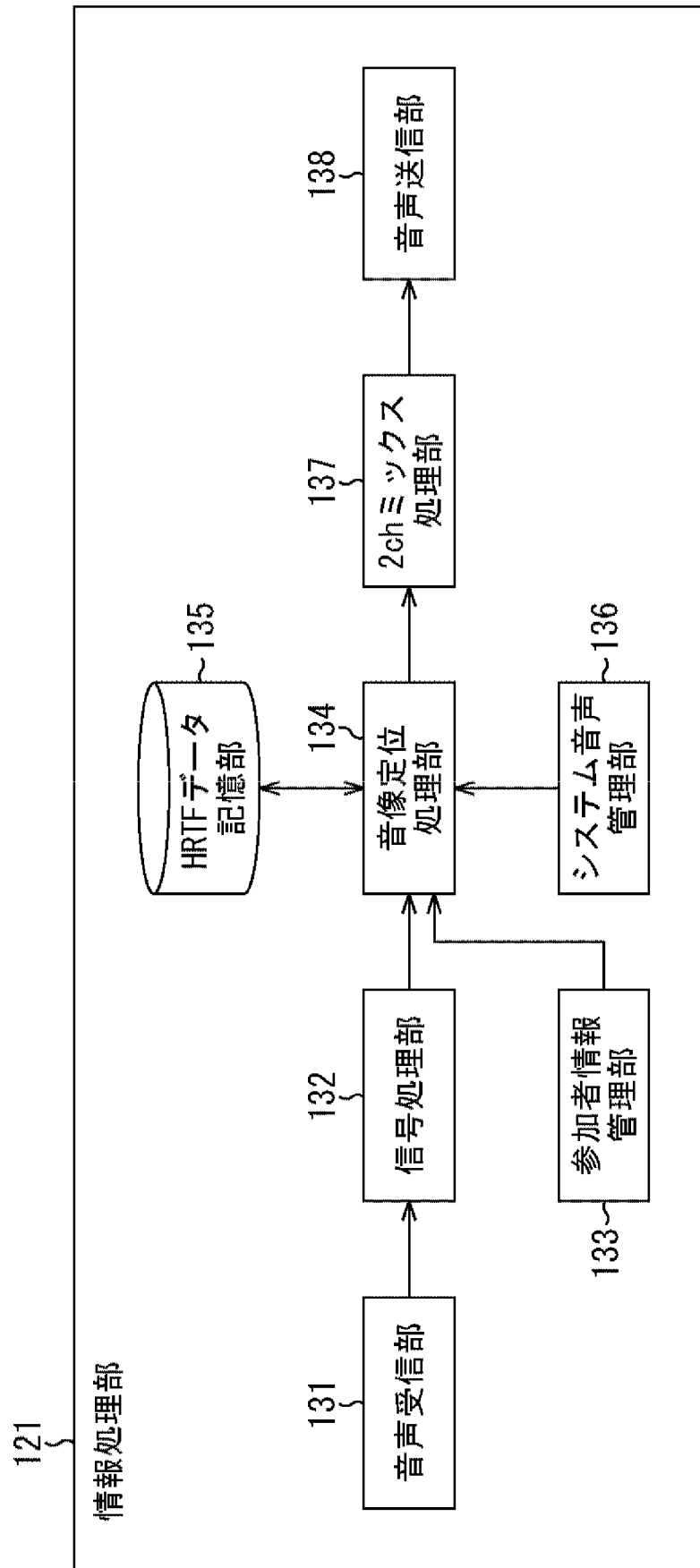
[図9]



[図10]

↑
I

[図11]

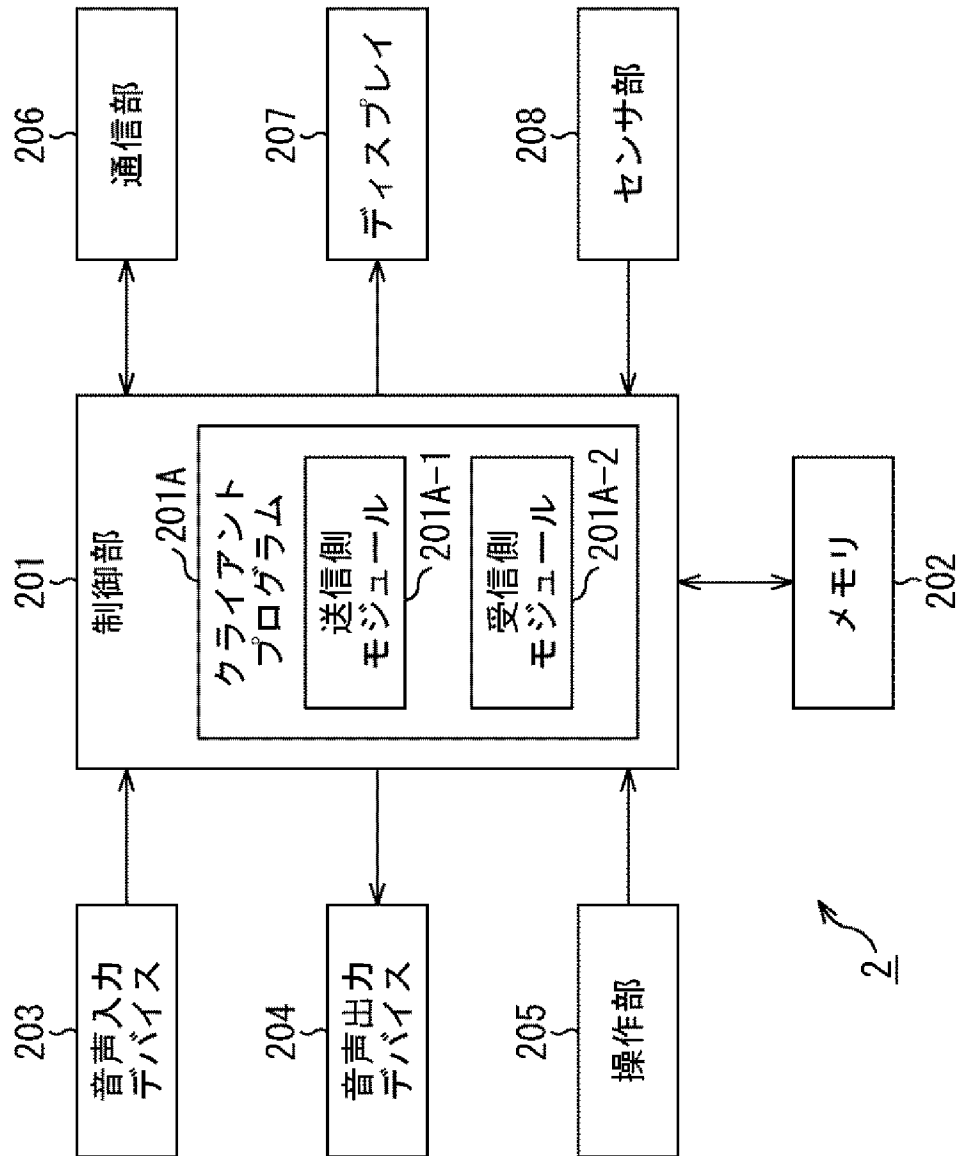


[図12]

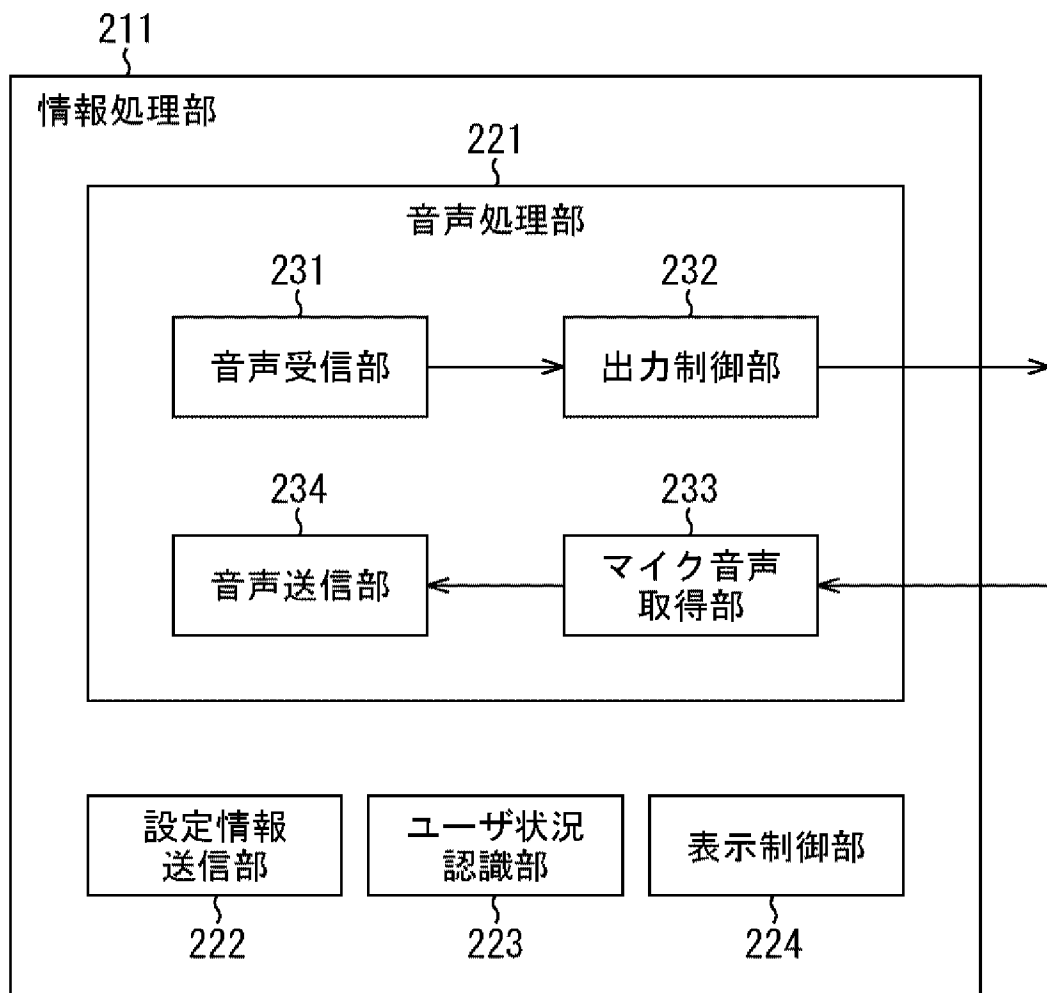
参加者情報

- ・ ユーザ情報
- ・ 位置情報
- ・ 設定情報（背景音など）
- ・ ボリューム情報

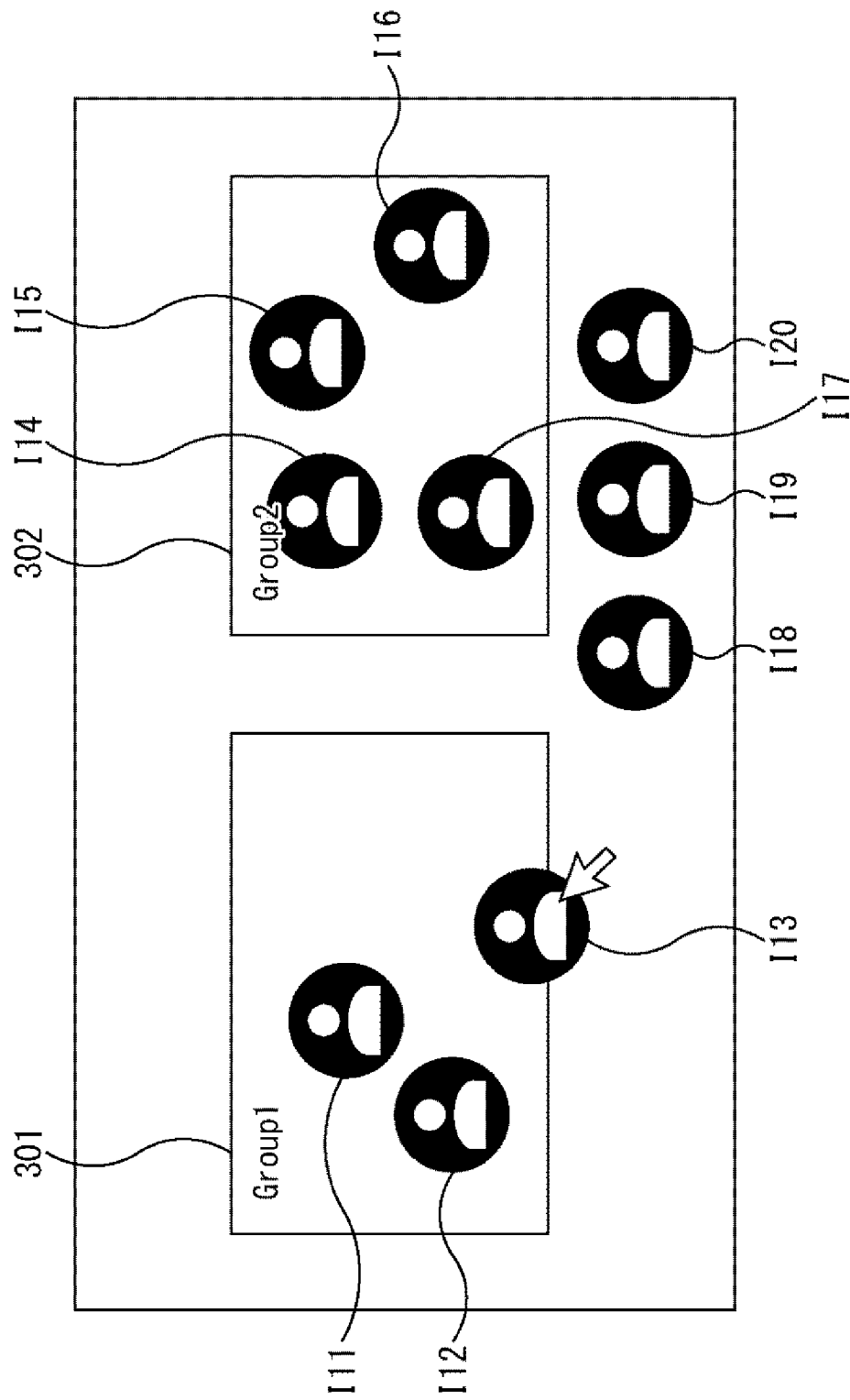
[図13]



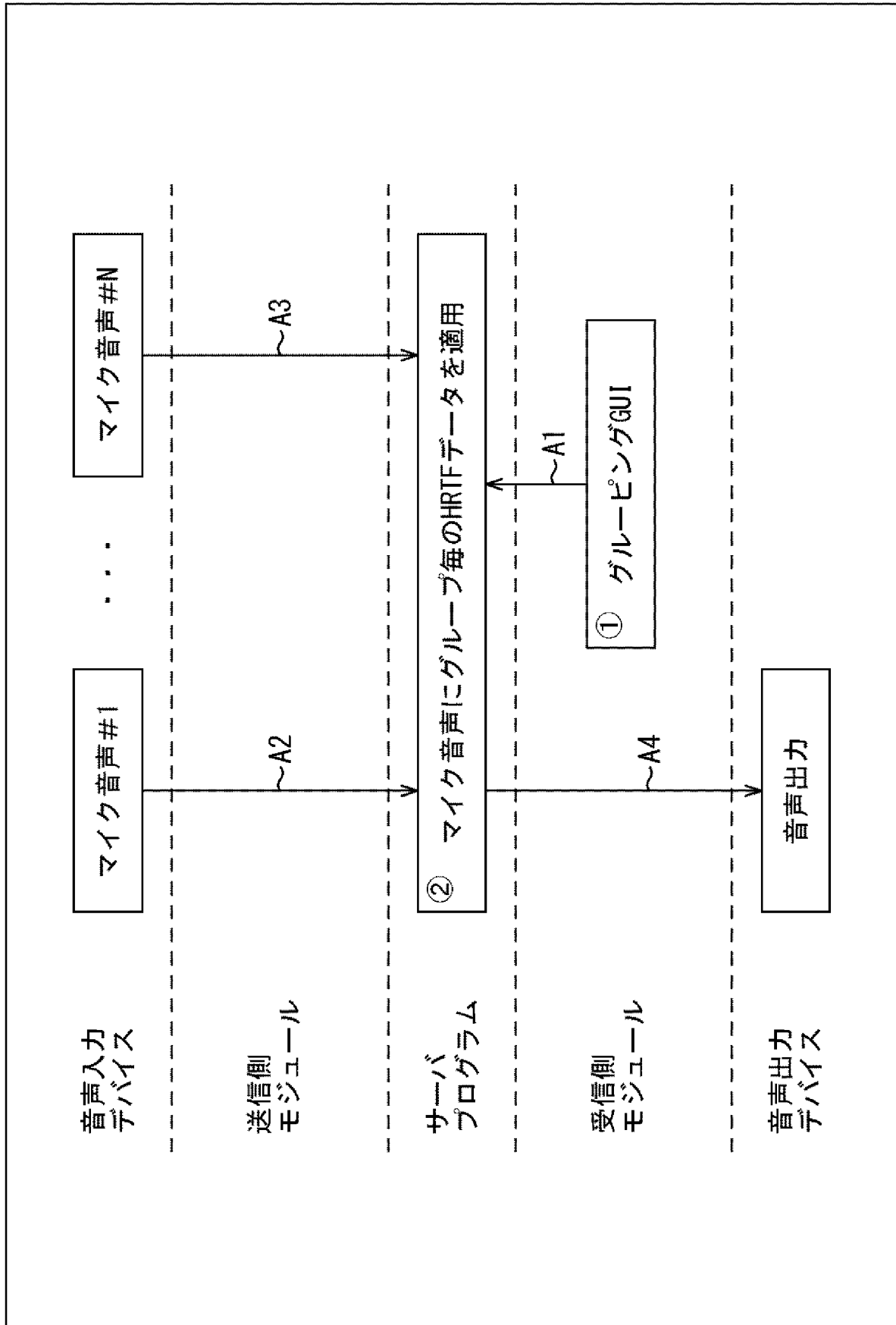
[図14]



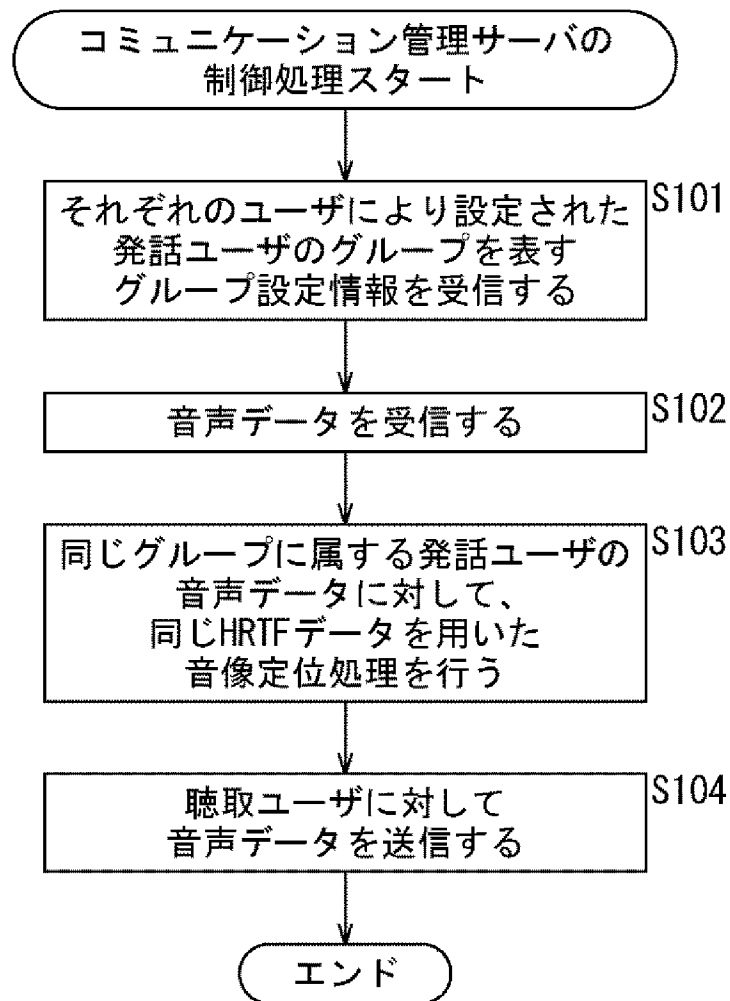
[図15]



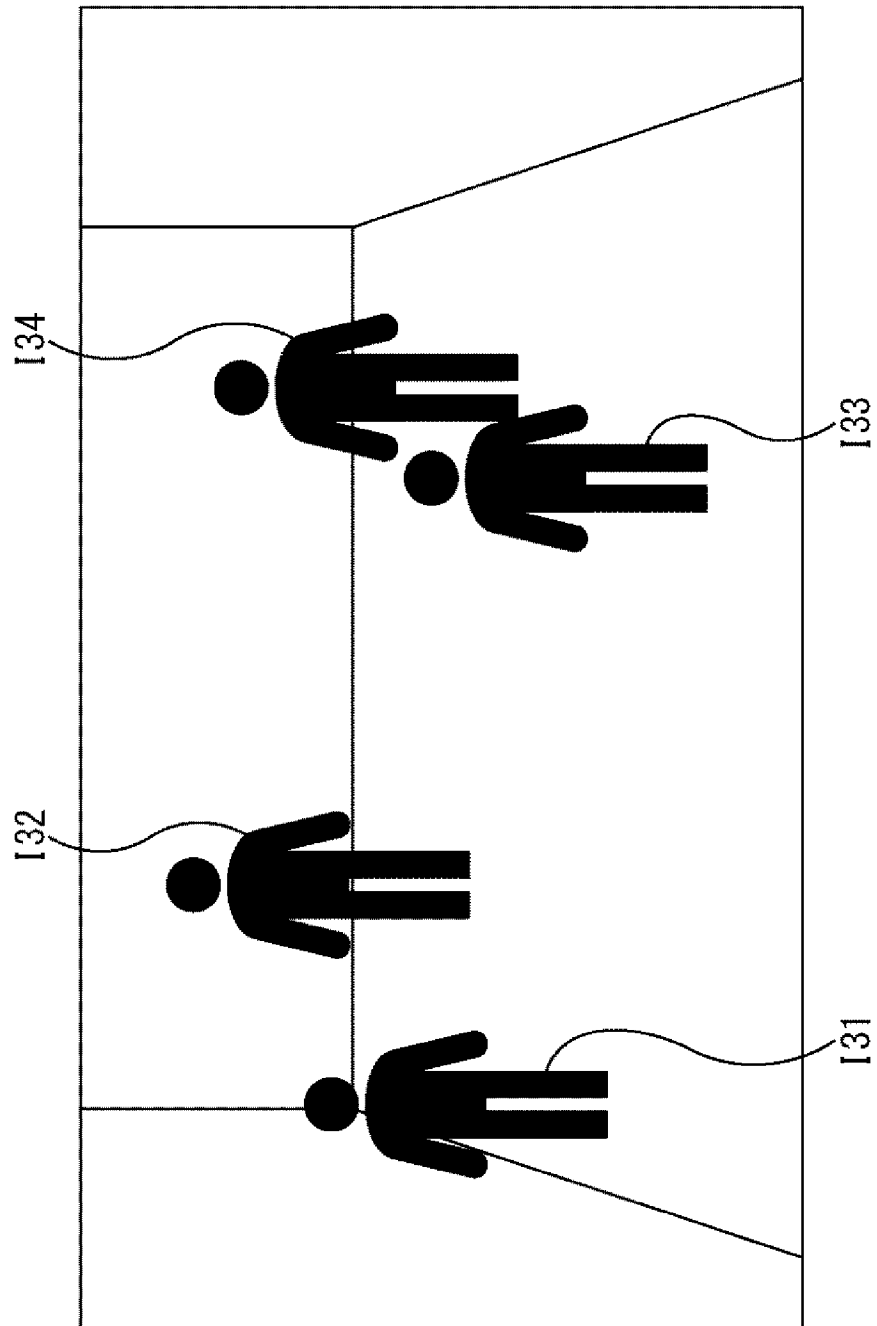
[図16]



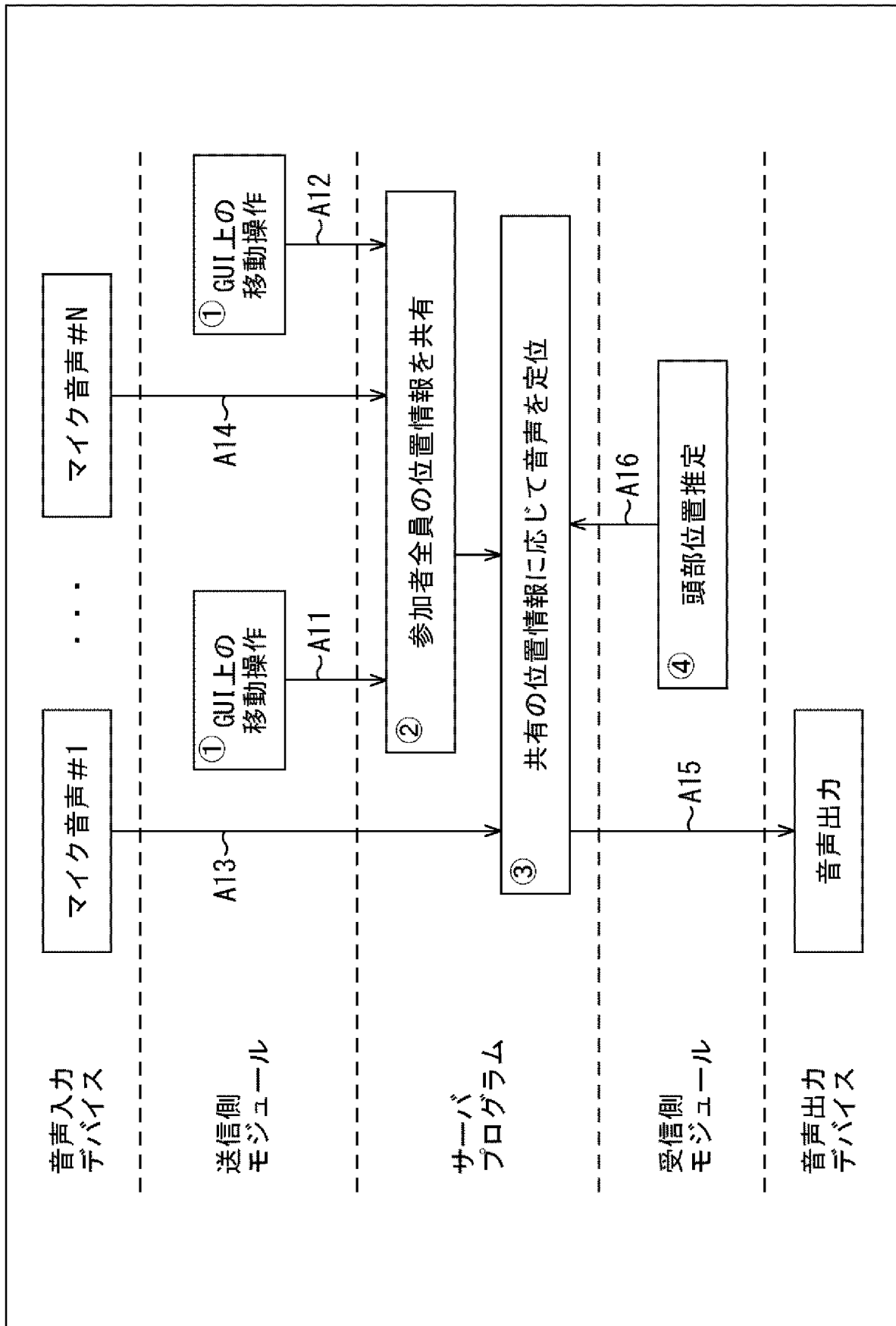
[図17]



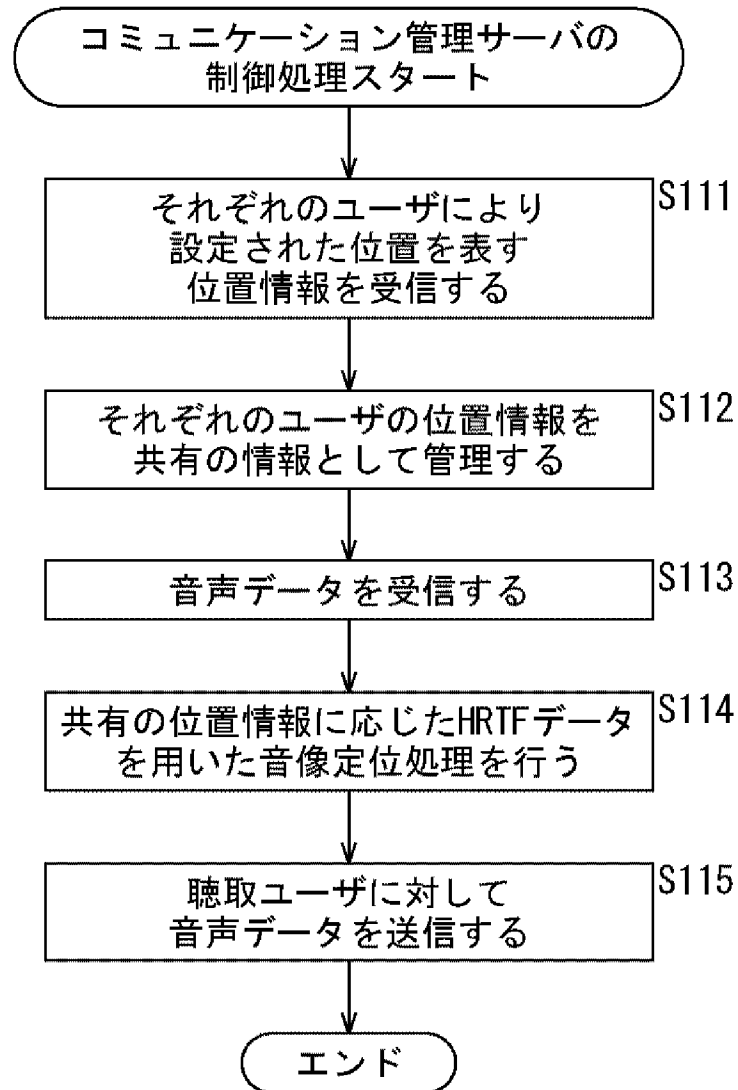
[図18]



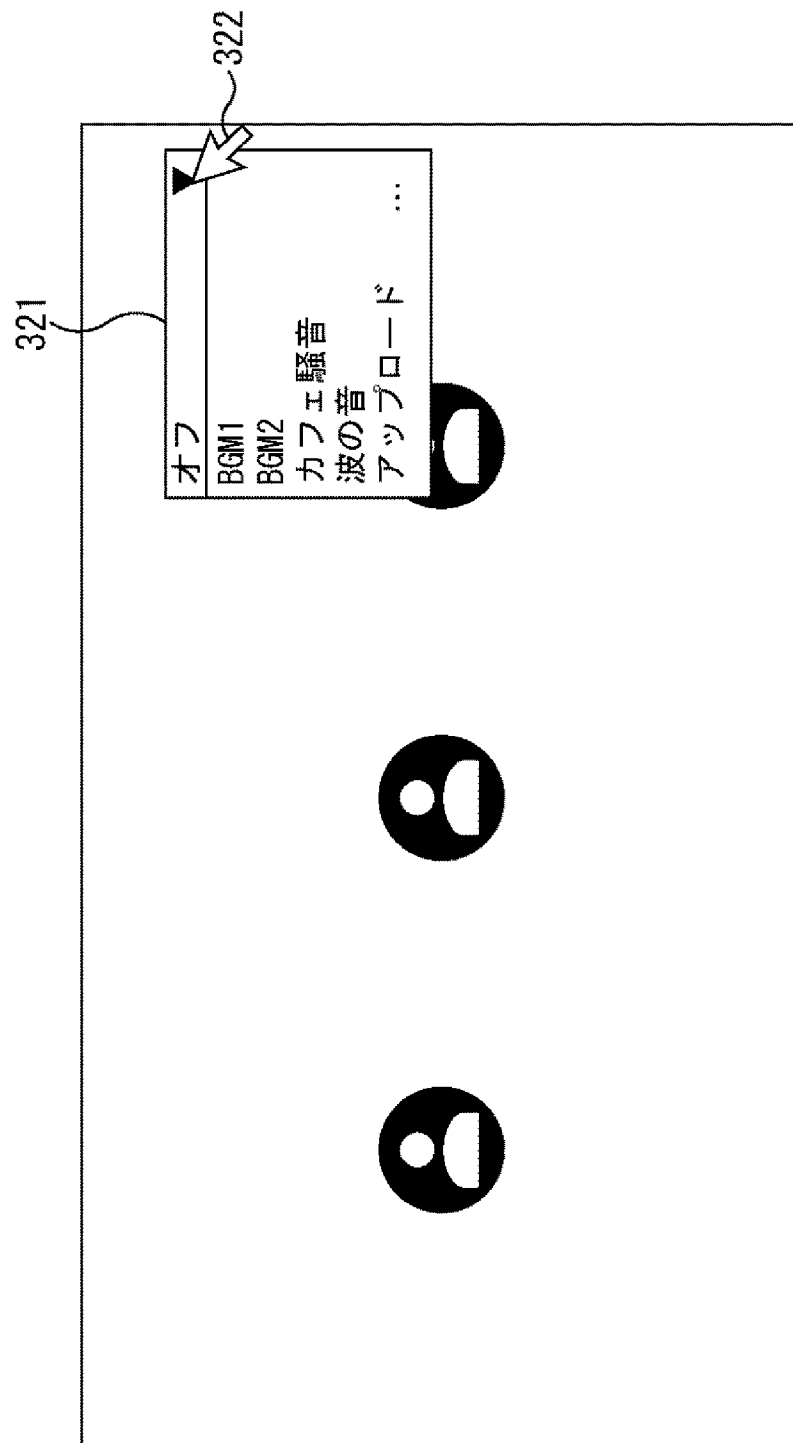
[図19]



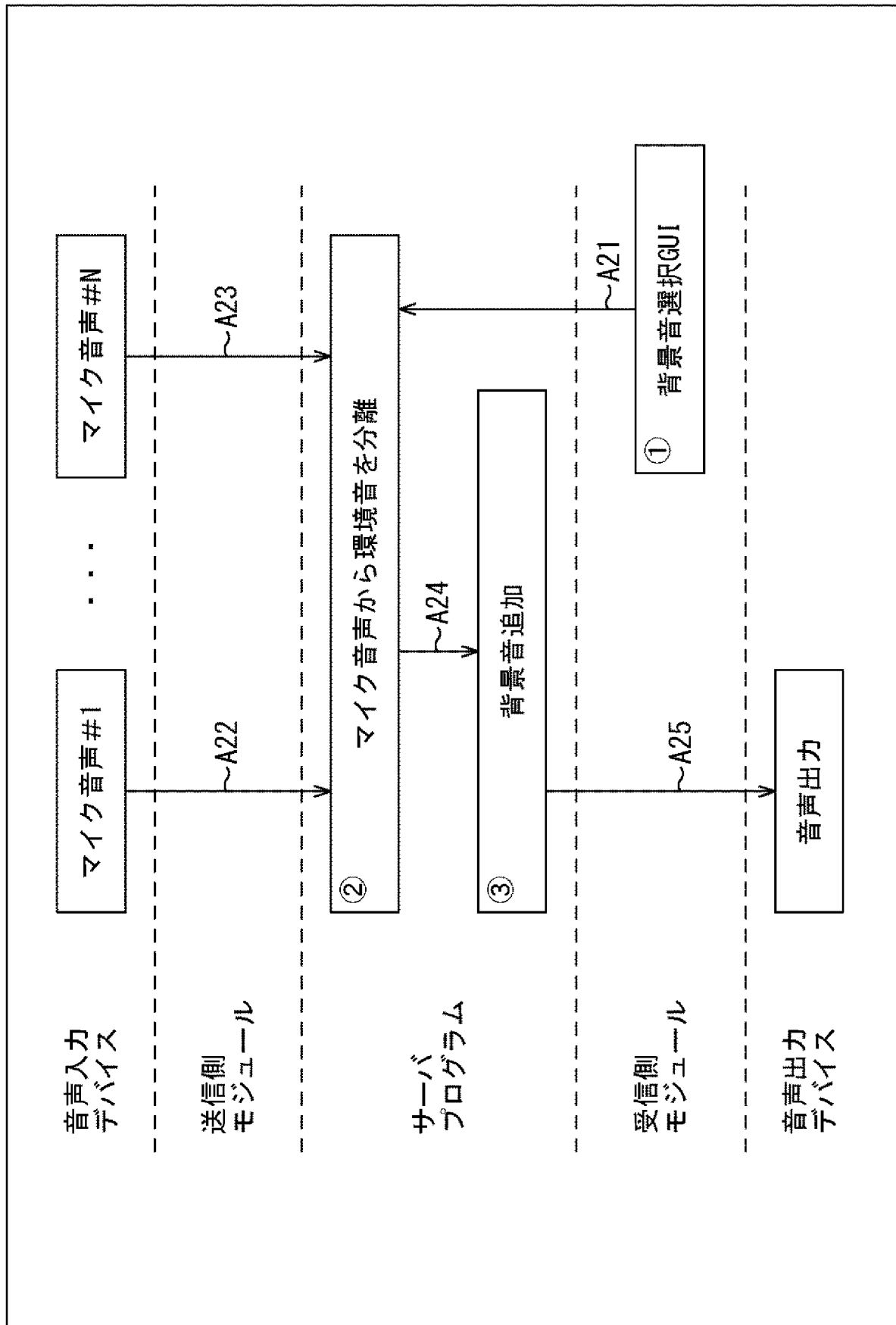
[図20]



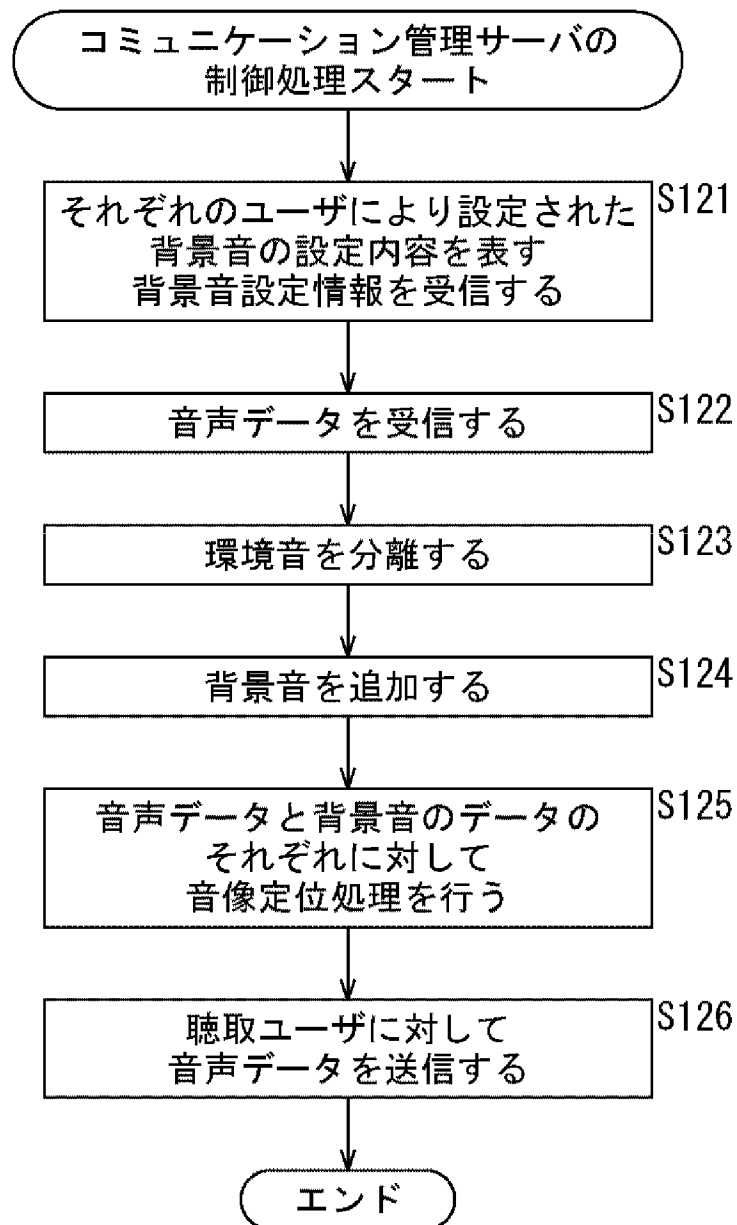
[図21]



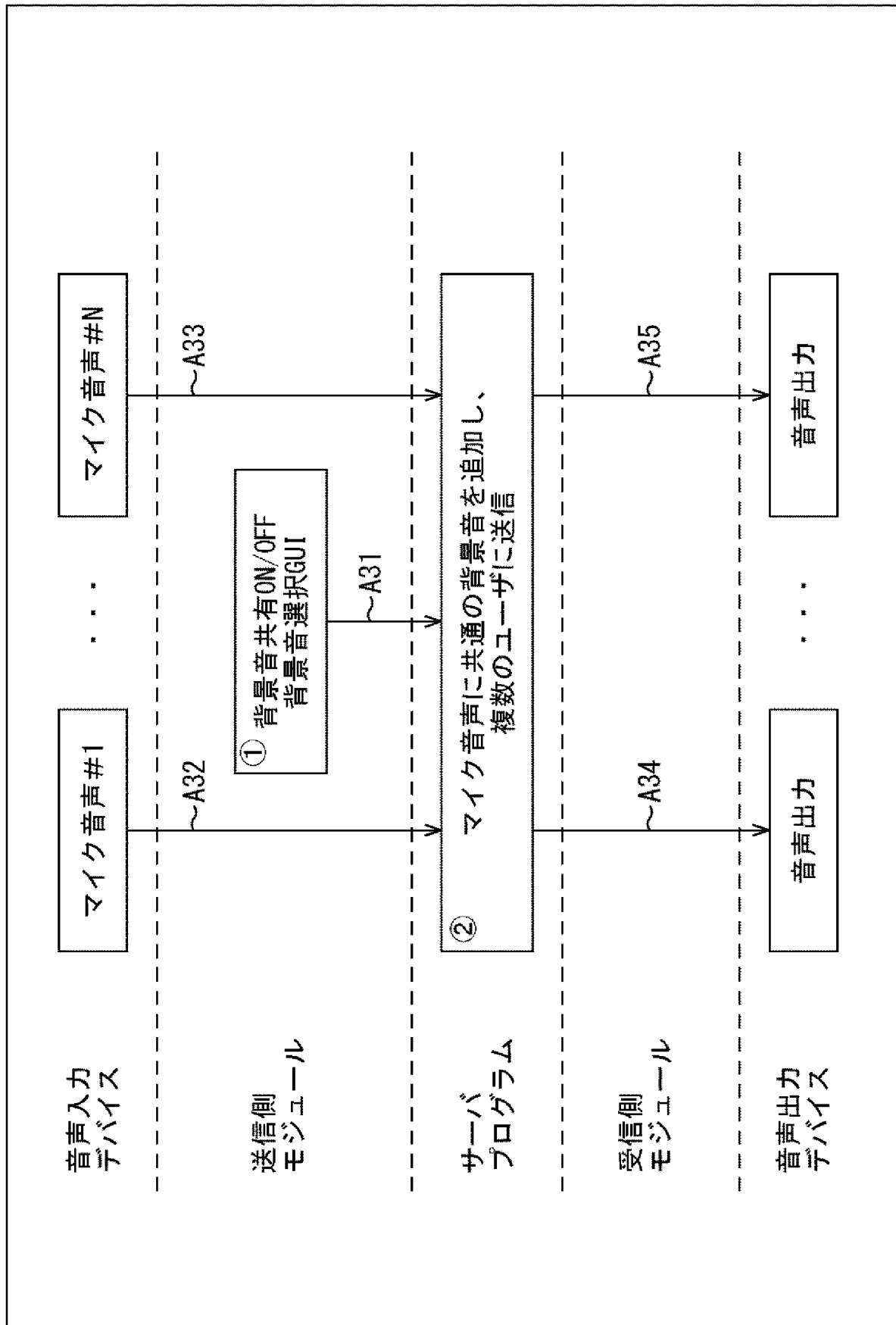
[図22]



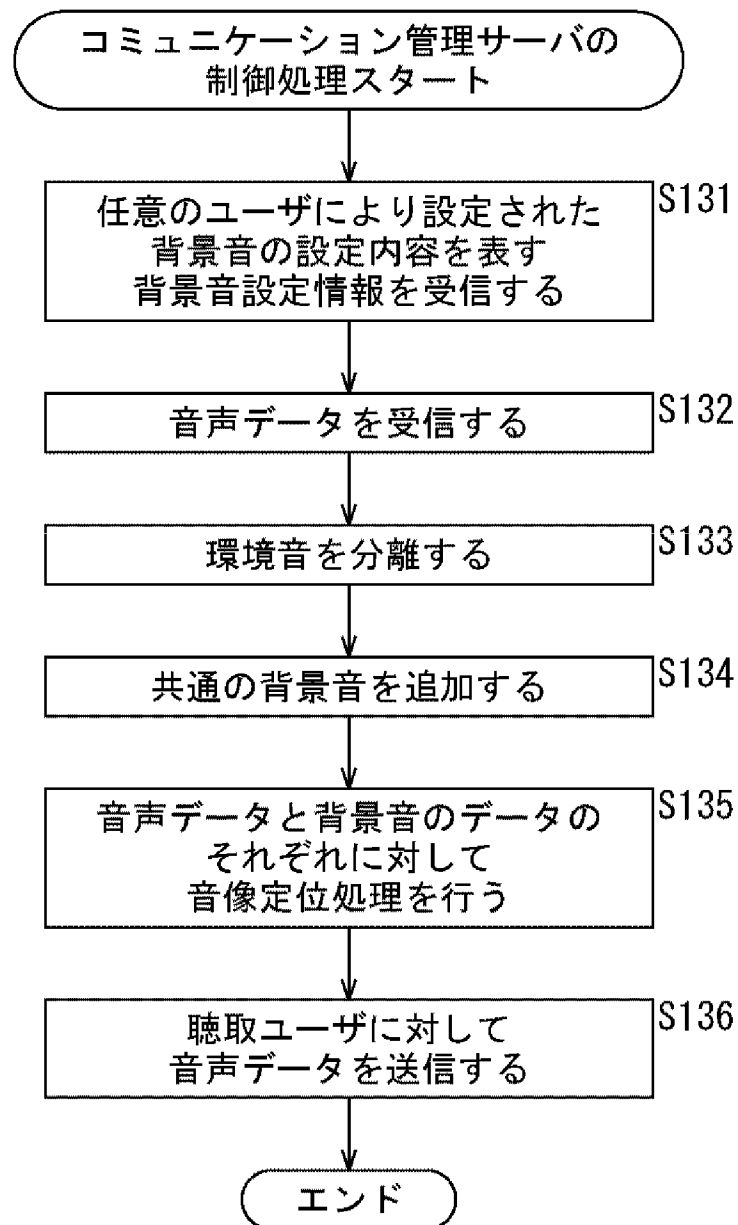
[図23]



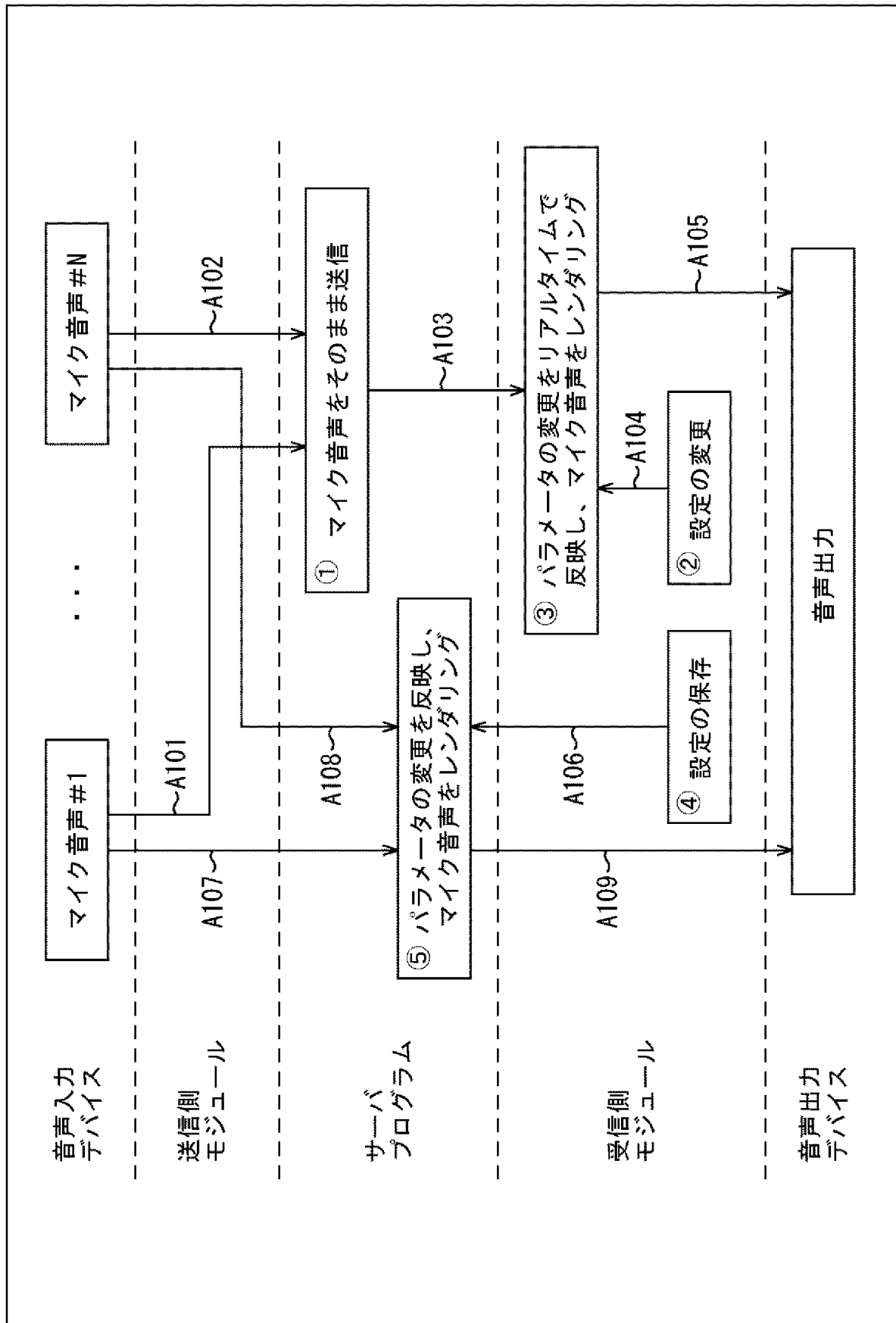
[図24]



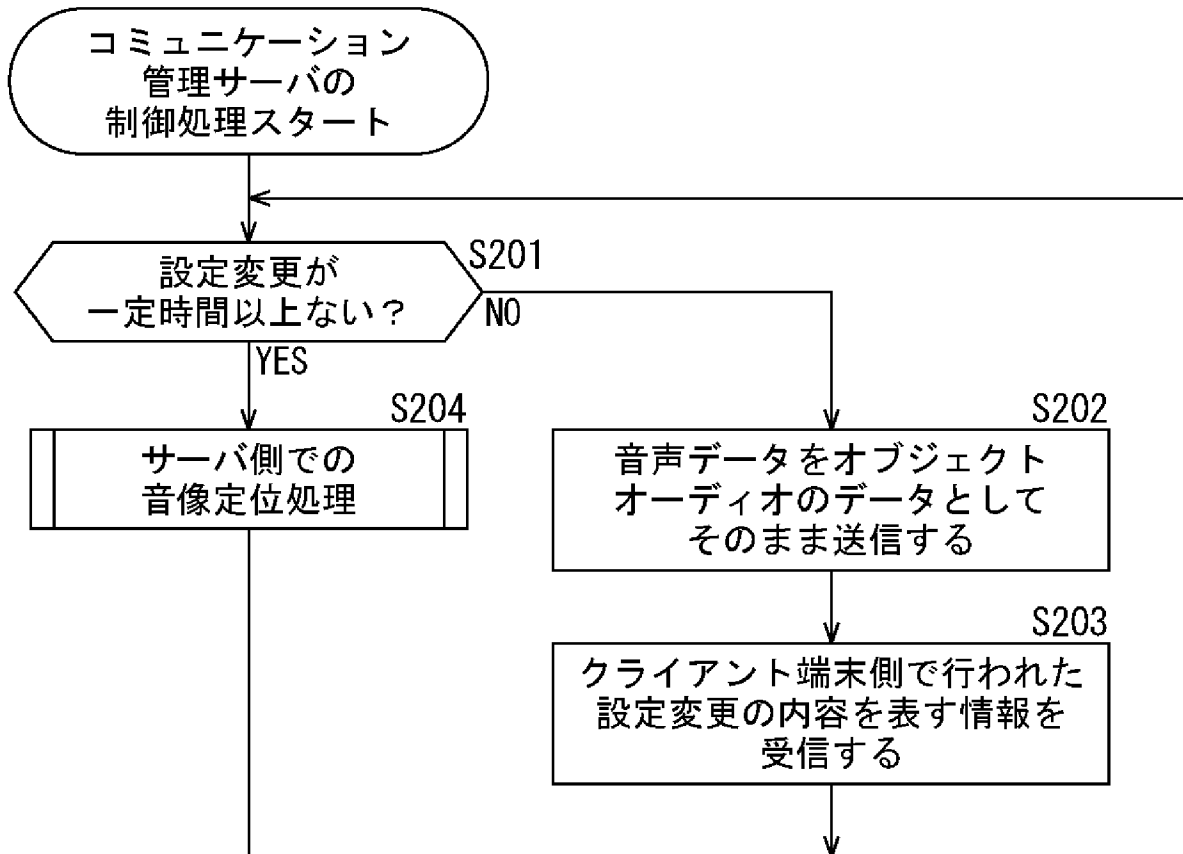
[図25]



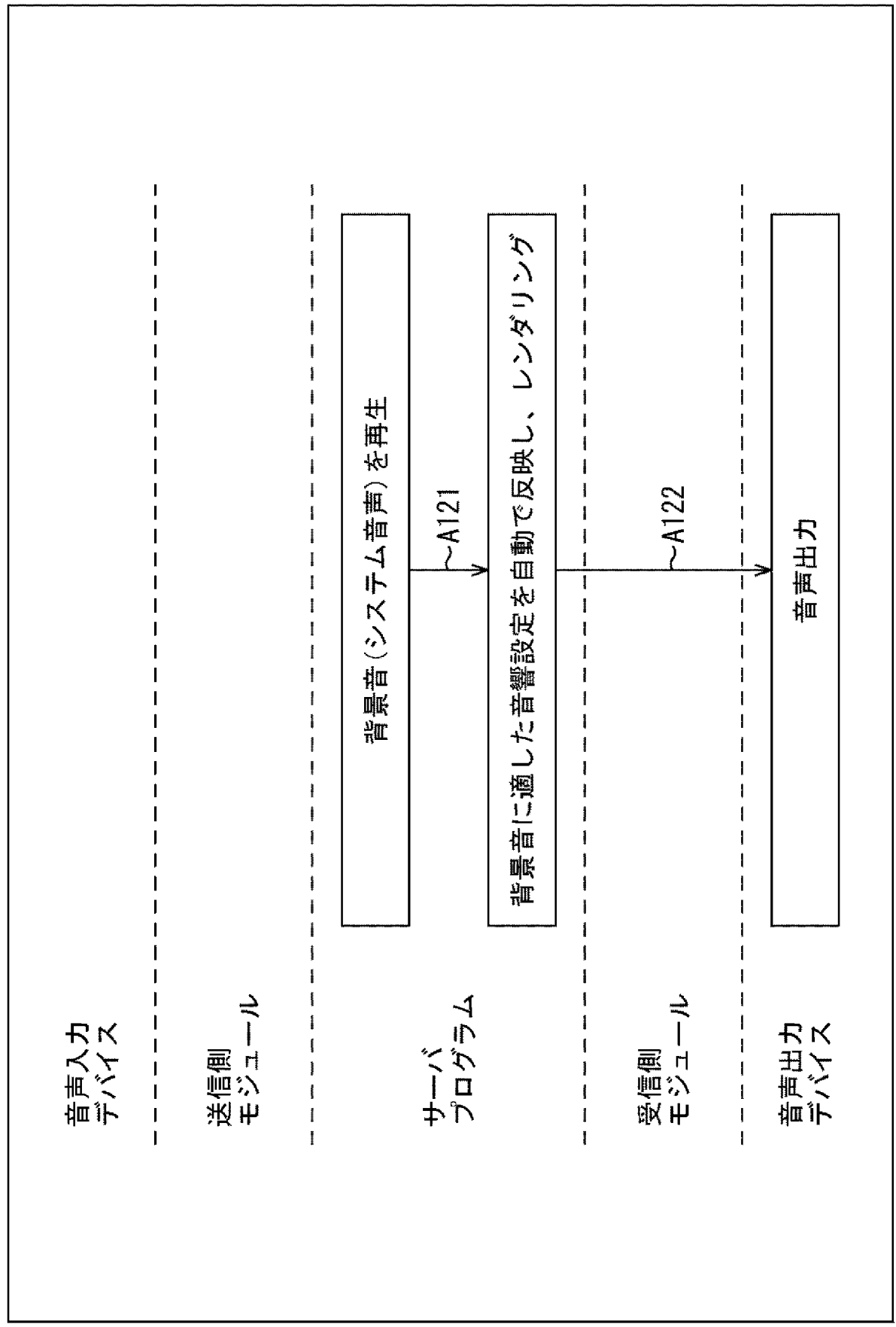
[図26]



[図27]



[図28]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/033279

A. CLASSIFICATION OF SUBJECT MATTER**H04M 3/56**(2006.01)i; **H04S 7/00**(2006.01)i; **H04N 7/15**(2006.01)i; **H04N 21/233**(2011.01)i

FI: H04N7/15 170; H04N21/233; H04S7/00 300; H04M3/56 C

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)

H04M3/56; H04S7/00; H04N7/15; H04N21/233

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 1971-2021
 Registered utility model specifications of Japan 1996-2021
 Published registered utility model applications of Japan 1994-2021

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	US 2020/0014792 A1 (LYREN, Philip Scott, et al.) 09 January 2020 (2020-01-09) paragraphs [0085]-[0102], [0218], [0219], [0262]	1-4, 6-9, 11-15, 17-19
Y		5, 10, 16
Y	US 2010/0215164 A1 (SANDGREN, Patrik et al.) 26 August 2010 (2010-08-26) paragraphs [0089], [0090]	5, 16
Y	JP 2001-274912 A (SEIKO EPSON CORP.) 05 October 2001 (2001-10-05) paragraphs [0035]-[0038]	10

☐ 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

12 November 2021

Date of mailing of the international search report

22 November 2021

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
Information on patent family members

International application No.

PCT/JP2021/033279

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2020/0014792	A1	09 January 2020	US	2017/0295278	A1	
				US	2019/0182377	A1	
				US	2019/0387095	A1	
US	2010/0215164	A1	26 August 2010	WO	2008/143561	A1	
				EP	2158752	A1	
				CN	101690149	A	
JP	2001-274912	A	05 October 2001	(Family: none)			

A. 発明の属する分野の分類（国際特許分類（IPC）） H04M 3/56(2006.01)i; H04S 7/00(2006.01)i; H04N 7/15(2006.01)i; H04N 21/233(2011.01)i FI: H04N7/15 170; H04N21/233; H04S7/00 300; H04M3/56 C																	
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）） H04M3/56; H04S7/00; H04N7/15; H04N21/233 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 <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2021年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2021年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2021年</td> </tr> </table> 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）			日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2021年	日本国実用新案登録公報	1996-2021年	日本国登録実用新案公報	1994-2021年							
日本国実用新案公報	1922-1996年																
日本国公開実用新案公報	1971-2021年																
日本国実用新案登録公報	1996-2021年																
日本国登録実用新案公報	1994-2021年																
C. 関連すると認められる文献 <table border="1"> <thead> <tr> <th>引用文献の カテゴリー*</th> <th>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</th> <th>関連する 請求項の番号</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>US 2020/0014792 A1 (LYREN PHILIP SCOTT, et al.) 09.01.2020 (2020-01-09) 段落[0085]-[0102], [0218]-[0219], [0262]</td> <td>1-4, 6-9, 11-15, 17-19</td> </tr> <tr> <td>Y</td> <td></td> <td>5, 10, 16</td> </tr> <tr> <td>Y</td> <td>US 2010/0215164 A1 (SANDGREN PATRIK, et al.) 26.08.2010 (2010-08-26) 段落[0089]-[0090]</td> <td>5, 16</td> </tr> <tr> <td>Y</td> <td>JP 2001-274912 A (セイコーエプソン株式会社) 05.10.2001 (2001-10-05) 段落[0035]-[0038]</td> <td>10</td> </tr> </tbody> </table>			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	US 2020/0014792 A1 (LYREN PHILIP SCOTT, et al.) 09.01.2020 (2020-01-09) 段落[0085]-[0102], [0218]-[0219], [0262]	1-4, 6-9, 11-15, 17-19	Y		5, 10, 16	Y	US 2010/0215164 A1 (SANDGREN PATRIK, et al.) 26.08.2010 (2010-08-26) 段落[0089]-[0090]	5, 16	Y	JP 2001-274912 A (セイコーエプソン株式会社) 05.10.2001 (2001-10-05) 段落[0035]-[0038]	10
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	US 2020/0014792 A1 (LYREN PHILIP SCOTT, et al.) 09.01.2020 (2020-01-09) 段落[0085]-[0102], [0218]-[0219], [0262]	1-4, 6-9, 11-15, 17-19															
Y		5, 10, 16															
Y	US 2010/0215164 A1 (SANDGREN PATRIK, et al.) 26.08.2010 (2010-08-26) 段落[0089]-[0090]	5, 16															
Y	JP 2001-274912 A (セイコーエプソン株式会社) 05.10.2001 (2001-10-05) 段落[0035]-[0038]	10															
☐ 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 12. 11. 2021		国際調査報告の発送日 22. 11. 2021															
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） 富樫 明 5C 5890 電話番号 03-3581-1101 内線 3541															

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/JP2021/033279

引用文献			公表日	パテントファミリー文献			公表日
US	2020/0014792	A1	09.01.2020	US	2017/0295278	A1	
				US	2019/0182377	A1	
				US	2019/0387095	A1	
US	2010/0215164	A1	26.08.2010	WO	2008/143561	A1	
				EP	2158752	A1	
				CN	101690149	A	
JP	2001-274912	A	05.10.2001	(ファミリーなし)			