

(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6月9日(09.06.2022)



(10) 国際公開番号

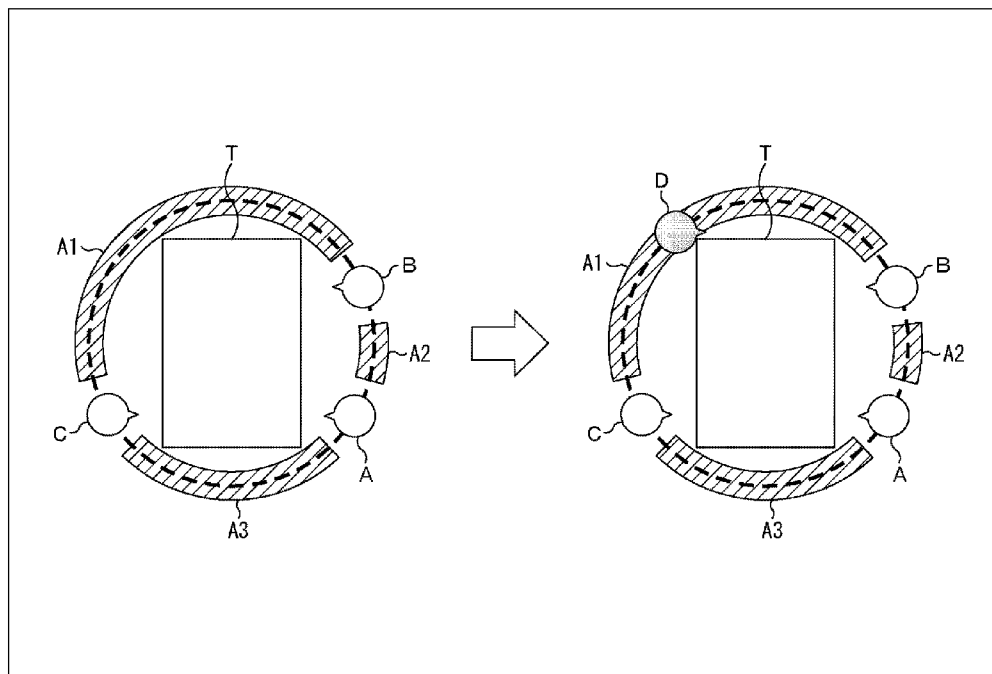
WO 2022/118671 A1

- (51) 国際特許分類:
H04M 3/56 (2006.01) *H04S 7/00* (2006.01)
- (21) 国際出願番号: PCT/JP2021/042528
- (22) 国際出願日: 2021年11月19日(19.11.2021)
- (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 2020-201905 2020年12月4日(04.12.2020) 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).
- (72) 発明者: 木村 健太郎 (**KIMURA Kentaro**); 〒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).
古賀 康之 (**KOGA Yasuyuki**); 〒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: 西川 孝, 外 (**NISHIKAWA Takashi et al.**); 〒1700013 東京都豊島区東池袋3丁目9番10号 池袋F Nビル4階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

(54) **Title:** INFORMATION PROCESSING APPARATUS, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法、およびプログラム

FIG. 4



(57) **Abstract:** The present technology relates to an information processing apparatus, an information processing method, and a program that make it possible to easily discern a difference between the voice of a real participant and the voice of a remote participant. An information processing apparatus according to an aspect of the present technology is provided with a sound image localization processing unit for localizing the sound image of the voice of a remote participant, who is remotely participating in a conversation being conducted in a predetermined space, at a position different than the position



WO 2022/118671 A1

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))

of a real participant, who is a participant present in the predetermined space. The present technology may be applied to a computer for conducting a remote meeting.

(57) 要約：本技術は、現実の参加者の音声と遠隔の参加者の音声とを容易に聞き分けることができるようにする情報処理装置、情報処理方法、およびプログラムに関する。本技術の一側面の情報処理装置は、所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる音像定位処理部を備える。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、情報処理方法、およびプログラム 技術分野

[0001] 本技術は、情報処理装置、情報処理方法、およびプログラムに関し、特に、現実の参加者の音声と遠隔の参加者の音声とを容易に聞き分けることができるようにした情報処理装置、情報処理方法、およびプログラムに関する。

背景技術

[0002] 遠隔にいるユーザがPCなどの装置を使用して会議に参加するいわゆるリモート会議が普及してきている。マイクロホ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。これにより、それぞれの参加者は、他の参加者と会話を行うことができる。

[0003] 例えば特許文献1には、仮想的な発言位置を間隔を空けて設定しておき、会議で重要な発言をされると思われる参加者の音声を、頭部伝達関数を用いて受話者の正面に定位させる技術が開示されている。

先行技術文献

特許文献

[0004] 特許文献1：特開2006-279492号公報
特許文献2：特開2006-254064号公報

発明の概要

発明が解決しようとする課題

[0005] 特許文献1に記載の技術においては、受話者の前に現実にいる人や受話者がいる空間にある物などの存在が考慮されていない。したがって、遠隔にいる参加者の音声が、実際にいる人の位置に定位している場合、現実にいる人の位置から、別人である遠隔にいる参加者の声が聞こえてしまう。

[0006] 本技術はこのような状況に鑑みてなされたものであり、現実の参加者の音声と遠隔の参加者の音声とを容易に聞き分ける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

[0007] 本技術の一側面の情報処理装置は、所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる音像定位処理部を備える。

[0008] 本技術の一側面においては、所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像が、所定の空間にいる参加者である現実参加者の位置と異なる位置に定位する。

図面の簡単な説明

[0009] [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]定位位置の設定の例を示す図である。

[図4]定位位置の設定の例を示す図である。

[図5]リモート会議の様子を示す図である。

[図6]現実参加者の位置を考慮していない場合の定位位置の設定の例を示す図である。

[図7]イヤホンの外観を示す図である。

[図8]出力デバイスの例を示す図である。

[図9]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10]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11]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12]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13]会議開始前の全体の処理について説明するフローチャートである。

[図14]図13のステップS23において行われる定位位置設定処理について説明するフローチャートである。

[図15]遠隔参加者の音声の定位位置の設定の流れについて説明する図である。

[図16]遠隔参加者の音声の定位位置の設定の流れについて説明する図15に続く図である。

[図17]遠隔参加者の音声の定位位置の設定の流れについて説明する図16に続く図である。

[図18]遠隔参加者の音声の定位位置の設定の流れについて説明する図17に続く図である。

[図19]遠隔参加者の音声の定位位置の設定の流れについて説明する図18に続く図である。

[図20]遠隔参加者の音声の定位位置の設定の流れについて説明する図である。

[図21]会議開始後の全体の処理について説明するフローチャートである。

[図22]会議開始後の全体の他の処理について説明するフローチャートである。

[図23]会議開始後の全体の他の処理について説明するフローチャートである。

[図24]定位可能領域の設定の例を示す図である。

[図25]音声のアニメーションの例を示す図である。

[図26]音源の移動経路の例を示す図である。

[図27]会議画面の例を示す図である。

[図28]コミュニケーション管理サーバの他の機能構成例を示すブロック図である。

発明を実施するための形態

[0010] 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システムについて
2. 各装置の構成

3. Tele-communicationシステムの動作

4. 変形例

[0011] <<Tele-communicationシステムについて>>

<システム構成>

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。

[0012] 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乃至2Dが、会議の参加者であるユーザA乃至Dが使用するクライアント端末として示されている。

[0013] 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乃至2Dを区別する必要がない場合、適宜、クライアント端末2という。

[0014] ユーザA乃至Dは、同じ会議に参加するユーザである。ユーザA乃至Dは、例えば、ステレオタイプのイヤホン（インナーイヤーヘッドホン）をそれぞれ装着し、会議に参加する。例えば、耳穴を密閉しないオープンイヤー型（開放型）のイヤホンが用いられる。ユーザA乃至Dが使用するイヤホンの外観については後述する。

[0015] 開放型のイヤホンを用いることにより、ユーザA乃至Dは、クライアント端末2が出力する音声とともに、周囲の音を聞くことができる。

[0016] 例えばイヤホンの筐体の所定の位置にはマイク（マイクロホン）が設けられる。イヤホンとクライアント端末2は、ケーブルを介して有線によって、または、無線LANやBluetooth（登録商標）などの所定の規格の通信を介して無線によって接続される。イヤホンとクライアント端末2の間では音声の送受信が行われる。

[0017] 各ユーザは、クライアント端末2を用意し、イヤホンを装着した状態で会議に参加することになる。例えば、図1の下方に示すように、ユーザA乃至

Cは、オフィスにある会議室などの同じ空間から会議に参加する。また、ユーザDは、会議に自宅からリモート参加する。

[0018] 同じ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。同じ空間から参加するユーザの数とリモート参加するユーザの数も任意に変更可能である。

[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] 例えばユーザAが発話を行った場合、図2の上段の矢印#1に示すように、コミュニケーション管理サーバ1は、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、ユーザAが使用する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] 図2の下段の矢印#2に示すように、コミュニケーション管理サーバ1は、ユーザAの音声データをクライアント端末2Dに送信し、ユーザAの音声を出力させる。

[0022] これにより、ユーザB乃至Dは、ユーザAの音声を聞くことができる。上述したように、ユーザBとユーザCが使用するイヤホンは開放型のイヤホンであるから、ユーザAと同じ空間にいるユーザBとユーザCは、ユーザAの音声を直接聞くことができる。

[0023] なお、後述するように、ユーザA乃至Cが開放型のイヤホンではなく密閉型のイヤホンやヘッドホンを装着している場合、ユーザAの音声は、コミュニケーション管理サーバ1を介してユーザBとユーザCにも送信される。この場合、ユーザB、ユーザC、ユーザDは、それぞれ、コミュニケーション管理サーバ1を介して送信されてきたユーザAの音声を聞くことになる。

[0024] ユーザBまたはユーザCが発話を行った場合も同様に、クライアント端末2Bまたはクライアント端末2C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、クライアント端末2Dに送信される。

[0025] また、コミュニケーション管理サーバ1は、ユーザDが発話することに応じてクライアント端末2Dから送信されてきたユーザDの音声データを受信する。コミュニケーション管理サーバ1は、ユーザDの音声データをクライアント端末2A乃至2Cにそれぞれ送信し、ユーザDの音声を出力させる。これにより、ユーザA乃至Cは、ユーザDの音声を聞くことができる。

[0026] 以下、適宜、他の参加者ととともに同じ空間に実際に集まって会議に参加する参加者を、同じ空間に現実にいる参加者という意味で現実参加者という。図1の例の場合、オフィスの会議室から会議に参加するユーザA乃至Cが、現実参加者となる。例えば、基準となる位置を中心として所定の範囲の空間が同じ空間となる。

[0027] 一方、現実参加者がいる空間とは異なる空間から会議に参加する参加者を、遠隔にいる参加者という意味で遠隔参加者という。図1の例の場合、自宅から1人で会議に参加するユーザDが、遠隔参加者となる。

[0028] <会議中の音声の定位>

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の出力時、クライアント端末2は音像定位処理を行う。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は、空間内の所定の位置に定位して出力される。

[0029] 例えば、ユーザA乃至Cが使用するクライアント端末2は、遠隔参加者であるユーザDの音声を会議室の所定の位置に定位させるための音像定位処理を行い、音像定位処理を行うことによって得られたユーザDの音声をそれぞれのユーザが使用するイヤホンから出力させる。ユーザA乃至Cは、定位位置として設定された位置からユーザDの音声が聞こえるように、ユーザDの音声の音像を感じるようになる。

[0030] コミュニケーション管理サーバ1においては、遠隔参加者であるユーザDの音声の定位位置が会議室内の所定の位置に設定される。コミュニケーション管理サーバ1により設定された定位位置の情報がクライアント端末2に提供され、音像定位処理に用いられる。

- [0031] 図3は、定位位置の設定の例を示す図である。
- [0032] 上方からの様子を図3に示す空間は、ユーザA乃至Cがいる会議室である。会議室に設けられたテーブルTの右側にユーザAとユーザBが並んで座り、テーブルTの左側の、ユーザAの正面の位置にユーザCが座っている。ユーザA乃至CはそれぞれテーブルTに向かって座っている。例えばユーザAが発話を行った場合、ユーザAの音声は、当然、ユーザBにとっては左側から聞こえ、ユーザCにとっては正面から聞こえる。
- [0033] 現実参加者であるユーザA乃至Cが図3の左側に示す状態で座っている場合、コミュニケーション管理サーバ1は、図3の右側にハッチを付して示すように、ユーザA乃至Cがいない領域を、遠隔参加者であるユーザDの音声を定位させることが可能な領域である定位可能領域として設定する。
- [0034] 図3の右側の例においては、ユーザA乃至Cの位置を通る円上であって、ユーザBとユーザCの間の領域が定位可能領域A1として設定されている。また、ユーザAとユーザBの間の領域が定位可能領域A2として設定され、ユーザAとユーザCの間の領域が定位可能領域A3として設定されている。定位可能領域A1乃至A3は、所定の幅の円弧状の領域である。
- [0035] この場合、コミュニケーション管理サーバ1は、図4の右側に示すように、定位可能領域内の所定の位置を、ユーザDの音声の定位位置として設定する。図4の右側の例においては、定位可能領域A1内の位置が、ユーザDの音声の定位位置として設定されている。遠隔参加者の音声の定位位置の設定については後に詳述する。
- [0036] ユーザDの音声の定位位置がこのようにしてコミュニケーション管理サーバ1において設定され、クライアント端末2において音像定位処理が行われることにより、図5に示すように、ユーザDの音声は、ユーザAにとっては右斜め前方の位置から聞こえ、ユーザBにとっては略正面の位置から聞こえる。また、ユーザDの音声は、ユーザCにとっては略左方の位置から聞こえる。
- [0037] 図5において、吹き出しに示す発話はユーザDの発話である。吹き出し元

の多重の円は、ユーザDの音声の音像を模式的に示す。ユーザDの音声の音像は、現実参加者がだれもいない位置に感じられることになる。ユーザA乃至Cはそれぞれイヤホン3を装着している。

[0038] このように、現実参加者の位置を考慮して、現実参加者がいない位置に遠隔参加者の音声の定位位置が設定されるため、それぞれの現実参加者は、他の現実参加者の音声と遠隔参加者の音声とを容易に聞き分けることが可能となる。

[0039] 仮に、現実参加者の位置を考慮することなく遠隔参加者の音声の定位位置が設定されるとした場合、図6に示すように、遠隔参加者であるユーザDの音声の定位位置が現実参加者であるユーザBの位置と同じ位置に設定されることがある。この場合、ユーザDの音声ユーザBの位置から聞こえ、ユーザAとユーザCは、誰と話しているのかが分からない状態になってしまうが、そのような状態になることを防ぐことが可能となる。

[0040] <イヤホンの構成>

図7は、イヤホンの外観を示す図である。

[0041] 各ユーザが装着するイヤホン3は、右側ユニット3Rと左側ユニット3L（図示せず）により構成される。図7の吹き出しに拡大して示すように、右側ユニット3Rは、ドライバユニット31とリング状の装着部33が、U字状の音導管32を介して接合されることによって構成される。右側ユニット3Rは、装着部33を外耳孔の周りに押し当て、装着部33とドライバユニット31とで右耳を挟むようにして装着される。

[0042] 左側ユニット3Lも右側ユニット3Rと同じ構成を有している。左側ユニット3Lと右側ユニット3Rは有線または無線で接続される。

[0043] 右側ユニット3Rのドライバユニット31は、クライアント端末2から送信されてきたオーディオ信号を受信し、オーディオ信号に応じた音を、矢印#1に示すように音導管32の先端から出力させる。音導管32と装着部33の接合部には、外耳孔に向けて音を出力する孔部が形成されている。

[0044] 装着部33はリング状の形状を有している。外耳孔には、音導管32の先

端から出力された音声とともに、矢印# 2に示すように周囲の音も到達することになる。

[0045] このように、イヤホン3は、耳穴を密閉しない開放型のイヤホンである。例えば、ドライバユニット31にマイクが設けられる。会議の参加者の音声を聞くことに用いられる出力デバイスとして、イヤホン3以外のデバイスが用いられるようにしてもよい。

[0046] 図8は、出力デバイスの例を示す図である。

[0047] 音声を聞くことに用いられる出力デバイスとして、図8のAに示すような密閉型のヘッドホン（オーバーイヤーヘッドホン）、または、図8のBに示すような肩載せ型のネックバンドスピーカが用いられる。ネックバンドスピーカを構成する左右のユニットにはスピーカが設けられており、ユーザの耳に向けて音出力される。ヘッドホンやネックバンドスピーカにもマイクが設けられ、ユーザの音声が集音される。

[0048] <<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9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。

[0049] 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9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。

[0050] 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。

[0051] バス104には、さらに、入出力インタフェース105が接続される。入出力インタフェース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。

- [0052] 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。
- [0053] 図10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10に示す機能部のうちの少なくとも一部は、図9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。
- [0054] 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、位置情報取得部131、定位位置設定部132、定位位置情報送信部133、音声受信部134、および音声送信部135により構成される。
- [0055] 位置情報取得部131は、現実参加者の位置を取得する。現実参加者が使用するクライアント端末2からは、現実参加者の位置を表す位置情報が送信されてくる。複数の現実参加者がいる場合、それぞれのクライアント端末2から送信されてきた位置情報に基づいて、それぞれの現実参加者の位置が取得される。位置情報取得部131により取得された位置情報は定位位置設定部132に供給される。
- [0056] 定位位置設定部132は、同じ空間にいる現実参加者の位置に基づいて、現実参加者がいない位置、すなわち、現実参加者の位置と異なる位置に定位可能領域を設定する。また、定位位置設定部132は、定位可能領域内の所定の位置を、遠隔参加者の音声の定位位置として設定する。定位位置設定部132により設定された遠隔参加者の音声の定位位置の情報である定位位置情報は、現実参加者の位置情報とともに定位位置情報送信部133に供給される。
- [0057] 定位位置情報送信部133は、通信部109を制御し、定位位置設定部132から供給された定位位置情報を、それぞれの現実参加者が使用するクライアント端末2に送信する。

[0058] また、定位位置情報送信部133は、定位位置情報と現実参加者の位置情報を、それぞれの遠隔参加者が使用するクライアント端末2に送信する。例えば、遠隔参加者が使用するクライアント端末2においては、定位位置情報により表される遠隔参加者自身の位置を基準として、それぞれの現実参加者の音声、現実参加者の位置に応じた方向から聞こえるように音像定位処理が行われる。

[0059] 音声受信部134は、通信部109を制御し、発話を行った参加者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134により受信された音声データは音声送信部135に出力される。

[0060] 音声送信部135は、通信部109を制御し、音声受信部134から供給された音声データを、聴取者となる参加者が使用するクライアント端末2に送信する。

[0061] <クライアント端末2の構成>

図11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。

[0062] クライアント端末2は、制御部201に対して、メモリ202、音声入力部203、音声出力部204、操作部205、通信部206、ディスプレイ207、およびカメラ208が接続されることによって構成される。

[0063] 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。

[0064] 例えばクライアント端末2は、汎用のPCに専用のアプリケーションプログラムであるクライアントプログラム201Aをインストールすることによって実現される。クライアント端末2が汎用のPCにDSPボードとA/D、D/A変換ボードを搭載することによって実現されるようにしてもよいし、専用の装置によって実現されるようにしてもよい。

- [0065] 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。
- [0066] 音声入力部203は、イヤホン3との間で通信を行い、イヤホン3から送信されてきた音声を受信する。イヤホン3からは、イヤホン3に設けられたマイクにより集音されたユーザの音声を送信されてくる。音声入力部203により受信された音声は、マイク音声として制御部201に出力される。
- [0067] 音声の入力が、クライアント端末2に設けられたマイクを用いて行われるようにしてもよいし、クライアント端末2に接続された外部のマイクを用いて行われるようにしてもよい。
- [0068] 音声出力部204は、イヤホン3との間で通信を行い、制御部201から供給されたオーディオ信号を送信することによって、会議の参加者の音声をイヤホン3から出力させる。
- [0069] 操作部205は、キーボードなどの入力部や、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。
- [0070] 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、IP通信網であるネットワーク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。
- [0071] 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、リモート会議の画面などの各種の画面が表示される。
- [0072] カメラ208は、例えばRGBカメラにより構成される。カメラ208は、会議の参加者であるユーザを撮影し、制御部201に出力する。音声だけでな

く、適宜、カメラ208により撮影された画像の送受信が、それぞれのクライアント端末2の間でコミュニケーション管理サーバ1を介して行われる。

[0073] 図12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12に示す機能部のうちの少なくとも一部は、図11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。

[0074] 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、再生処理部221、音声送信部222、およびユーザ位置検出部223により構成される。図12の情報処理部211が、現実参加者が使用するクライアント端末2の構成であるものとして主に説明する。

[0075] 再生処理部221は、音声受信部241、定位位置取得部242、音像定位処理部243、HRTFデータ記憶部244、および出力制御部245により構成される。

[0076] 音声受信部241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。コミュニケーション管理サーバ1からは、遠隔参加者などの他の参加者の音声データが送信されてくる。音声受信部241により受信された音声データは音像定位処理部243に供給される。

[0077] 定位位置取得部242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定位位置情報を受信する。コミュニケーション管理サーバ1からは、遠隔参加者の音声の定位位置を表す定位位置情報が送信されてくる。定位位置取得部242により受信された定位位置情報は音像定位処理部243に供給される。

[0078] 音像定位処理部243は、現実参加者であるクライアント端末2のユーザの位置と、定位位置情報により表される遠隔参加者の音声の定位位置との位置関係（方向と距離の関係）に応じたHRTF(Head-Related Transfer Function)データをHRTFデータ記憶部244から読み出して取得する。

[0079] HRTFデータ記憶部244には、現実参加者がいる空間上のそれぞれの位置を聴取位置としたときの、様々な位置から聴取位置までの音の伝達特性を表

すHRTF（頭部伝達関数）のデータであるHRTFデータが記憶されている。現実参加者がいる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データがクライアント端末2に用意されている。

[0080] 音像定位処理部243は、発話を行った遠隔参加者の音声、遠隔参加者の音声の定位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遠隔参加者の音声データに対して行う。音像定位処理部243が行う音像定位処理には、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。

[0081] すなわち、遠隔参加者の音声は、オブジェクトオーディオの音声データとしてクライアント端末2において処理される。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出力制御部245に供給される。

[0082] 出力制御部245は、音像定位処理により生成されたオーディオデータを音声出力部204に出力し、イヤホン3から出力させる。

[0083] 音声送信部222は、通信部206を制御し、音声入力部203から供給されたマイク音声のデータをコミュニケーション管理サーバ1に送信する。

[0084] ユーザ位置検出部223は、現実参加者であるクライアント端末2のユーザの位置を検出する。ユーザ位置検出部223は、会議参加者の位置センサとして機能する。

[0085] クライアント端末2のユーザの位置は、例えば、GPSなどの測位システムの情報に基づいて検出される。また、ユーザの位置は、モバイル基地局の情報、無線LANのアクセスポイントの情報に基づいて検出される。Bluetooth（登録商標）の通信を用いてユーザの位置が検出されるようにしてもよいし、カメラ208により撮影された画像に基づいてユーザの位置が検出されるようにしてもよい。クライアント端末2に搭載された加速度センサやジャイロセンサによる計測結果に基づいてユーザの位置が検出されるようにしてもよい。

[0086] ユーザ位置検出部223は、通信部206を制御し、クライアント端末2

のユーザの位置を表す位置情報をコミュニケーション管理サーバ1に送信する。

[0087] なお、図12の情報処理部211が、遠隔参加者が使用するクライアント端末2に設けられる構成である場合、音声受信部241においては、コミュニケーション管理サーバ1から送信されてきた、現実参加者その他の遠隔参加者の音声データが受信される。

[0088] 定位位置取得部242においては、コミュニケーション管理サーバ1から送信されてきた定位位置情報と現実参加者の位置情報が受信され、現実参加者その他の遠隔参加者のそれぞれの位置関係に応じたHRTFデータを用いて、音像定位処理部243において音像定位処理が行われる。音像定位処理が行われることによって得られた現実参加者の音声と他の遠隔参加者の音声は、出力制御部245によって、遠隔参加者であるクライアント端末2のユーザが使用するイヤホン3から出力される。

[0089] なお、音声データのレンダリングについては、携帯電話機、PHS、VOIP電話機、デジタル交換機、ゲートウェイ、ターミナルアダプタ等の外部の装置が有する演算機能を用いて行われるようにしてもよい。

[0090] <<Tele-communicationシステムの動作>>

<会議の開始前の処理>

図13のフローチャートを参照して、会議開始前の全体の処理について説明する。

[0091] 図13には、クライアント端末2の処理として、現実参加者U1、U2が使用するクライアント端末2の処理と、遠隔参加者U11、U12が使用するクライアント端末2の処理が示されている。後述する他のシーケンス図においても同様である。

[0092] 現実参加者U1が使用するクライアント端末2をクライアント端末2-1とし、現実参加者U2が使用するクライアント端末2をクライアント端末2-2として説明する。同様に、遠隔参加者U11が使用するクライアント端末2をクライアント端末2-11とし、遠隔参加者U12が使用するクライ

アント端末2をクライアント端末2-12として説明する。

- [0093] 同じ会議に参加する他の現実参加者が使用するクライアント端末2の処理も、現実参加者U1, U2が使用するクライアント端末2-1, 2-2の処理と同様の処理になる。また、他の遠隔参加者が使用するクライアント端末2の処理も、遠隔参加者U11, U12が使用するクライアント端末2-11, 2-12の処理と同様の処理になる。
- [0094] ステップS1において、クライアント端末2-1のユーザ位置検出部223は、現実参加者U1の位置を検出し、位置情報をコミュニケーション管理サーバ1に送信する。
- [0095] ステップS11において、クライアント端末2-2のユーザ位置検出部223は、現実参加者U2の位置を検出し、位置情報をコミュニケーション管理サーバ1に送信する。
- [0096] ステップS21において、コミュニケーション管理サーバ1の位置情報取得部131は、クライアント端末2-1から送信されてきた位置情報を受信し、現実参加者U1の位置を取得する。
- [0097] ステップS22において、位置情報取得部131は、クライアント端末2-2から送信されてきた位置情報を受信し、現実参加者U2の位置を取得する。
- [0098] ステップS23において、定位位置設定部132は定位位置設定処理を行う。定位位置設定処理により、遠隔参加者U11, U12のそれぞれの音声の定位位置が設定される。定位位置設定処理の詳細については図14のフローチャートを参照して後述する。
- [0099] ステップS24において、定位位置情報送信部133は、遠隔参加者U11, U12のそれぞれの音声の定位位置を表す定位位置情報をクライアント端末2-1とクライアント端末2-2に送信する。
- [0100] ステップS25において、定位位置情報送信部133は、現実参加者U1, U2のそれぞれの位置を表す位置情報とともに、遠隔参加者U11, U12のそれぞれの音声の定位位置を表す定位位置情報をクライアント端末2-

1 1 とクライアント端末 2 - 1 2 に送信する。

- [0101] ステップ S 2 において、クライアント端末 2 - 1 の定位位置取得部 2 4 2 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定位位置情報を受信する。
- [0102] ステップ S 1 2 において、クライアント端末 2 - 2 の定位位置取得部 2 4 2 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定位位置情報を受信する。
- [0103] ステップ S 3 1 において、クライアント端末 2 - 1 1 の定位位置取得部 2 4 2 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、現実参加者 U 1 , U 2 のそれぞれの位置情報および遠隔参加者 U 1 2 の定位位置情報を受信する。
- [0104] ステップ S 4 1 において、クライアント端末 2 - 1 2 の定位位置取得部 2 4 2 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、現実参加者 U 1 , U 2 のそれぞれの位置情報および遠隔参加者 U 1 1 の定位位置情報を受信する。
- [0105] 図 1 4 のフローチャートを参照して、図 1 3 のステップ S 2 3 において行われる定位位置設定処理について説明する。
- [0106] ステップ S 5 1 において、コミュニケーション管理サーバ 1 の定位位置設定部 1 3 2 は、位置情報取得部 1 3 1 により取得された現実参加者の位置に基づいて定位可能領域を算出する。
- [0107] ステップ S 5 2 において、定位位置設定部 1 3 2 は、定位可能領域内の所定の位置を、遠隔参加者の音声の定位位置として設定する。その後、図 1 3 のステップ S 2 3 に戻り、それ以降の処理が行われる。
- [0108] 図 1 5 乃至図 2 0 を参照して、遠隔参加者の音声の定位位置の設定の流れについて説明する。
- [0109] 図 1 5 に示すように、現実参加者 U 1 乃至 U 3 が、それぞれ、会議室などの同じ空間内の位置 P 1 乃至 P 3 にいるものとする。位置 P 1 乃至 P 3 は、現実参加者 U 1 乃至 U 3 がそれぞれ使用するクライアント端末 2 から送信さ

れてきた位置情報に基づいて特定される。

[0110] 定位位置設定部 132 は、現実参加者がいる空間に所定の半径 R [m] の円を作成し、作成した円の内側にいる現実参加者からなるグループを、同じ会議の参加者の 1 つのグループとして形成する。

[0111] 図 16 に示すようにグループの作り方が複数ある場合、定位位置設定部 132 は、「会議室の参加者は集まってください」などの音声をクライアント端末 2 から出力させることによって、互いに近づくことを参加者に通知する。破線の円で囲んで示すように、図 16 の例においては、現実参加者 U_1 乃至 U_3 により 1 つのグループが形成されるとともに、現実参加者 U_1 , U_3 , U_4 により 1 つのグループが形成される。

[0112] このような処理が、1 つのグループが形成されるまで続けられる。例えば、1 つのグループを形成する円として半径 5 m の円が設定される。空間の広さに応じて、1 つのグループを形成する円の大きさを例えば現実参加者自身に変更できるようにしてもよい。

[0113] 同じグループを形成する現実参加者 U_1 乃至 U_N の位置 P_1 乃至 P_N の x 座標、 y 座標を (x_1, y_1) 乃至 (x_n, y_n) とすると ($n = 1$ 乃至 N)、定位位置設定部 132 は、下式 (1) により表される円において、点 (x_n, y_n) までの距離の和が最小となる x_c, y_c, r_1 を求める。

[数1]

$$(x_n - x_c)^2 + (y_n - y_c)^2 = r_1^2 \quad \dots (1)$$

[0114] x_c, y_c, r_1 を求めることは、図 17 の破線の円で示すように、位置 P_1 乃至 P_N にある点群の近似円を求めることに相当する。図 17 の例においては、現実参加者 U_1 乃至 U_4 がいる位置 P_1 乃至 P_4 に応じて半径 r_1 の近似円 C が設定されている。なお、 $N = 2$ の場合、位置 P_1 と位置 P_2 を通る最小半径の円が近似円として設定される。

[0115] 図 18 に斜線を付して示すように、定位位置設定部 132 は、近似円 C 上であって、現実参加者が実際にいる位置から離れた位置に定位可能領域を設定する。上述したように、定位可能領域は所定の幅の円弧状の領域である。

例えば、定位可能領域は、現実参加者が実際にいる位置に最も近い近似円C上の位置から r_2 [m]以上離れた位置に設定される。

[0116] 図18の例においては、現実参加者U1と現実参加者U3の間に、それぞれの現実参加者から所定の距離だけ離して定位可能領域A1が設定され、現実参加者U1と現実参加者U2の間に、それぞれの現実参加者から所定の距離だけ離して定位可能領域A2が設定されている。また、現実参加者U3と現実参加者U4の間に、それぞれの現実参加者から所定の距離だけ離して定位可能領域A3が設定され、現実参加者U2と現実参加者U4の間に、それぞれの現実参加者から所定の距離だけ離して定位可能領域A4が設定されている。

[0117] 現実参加者U1乃至U3を囲む実線の小円は、現実参加者の近似円C上の位置を中心とする半径 r_2 の円である。現実参加者U4の近傍に示す実線の小円も、近似円C上の位置を中心とする半径 r_2 の円である。

[0118] 定位可能領域を設定した後、定位位置設定部132は、定位可能領域内に遠隔参加者の音声の定位位置を設定する。図19の例においては、定位可能領域A1の略中心に、1人目の遠隔参加者である遠隔参加者U11の音声の定位位置が設定され、定位可能領域A2の略中心に、2人目の遠隔参加者である遠隔参加者U12の音声の定位位置が設定されている。

[0119] 音声の位置が近いと聞き分けが難しくなる。例えば、近似円Cの中心Oからの角度が分散するように、遠隔参加者の音声の定位位置が設定される。

[0120] 言い換えると、現実参加者が実際にいる位置 P_1 乃至 P_N 、M人の遠隔参加者の音声の定位位置 Q_m ($m=1$ 乃至 M)としたときに、角 P_iOP_j ($1 \leq i < j \leq N$)、角 Q_iOQ_j ($1 \leq i < j \leq M$)、角 P_iOQ_j ($1 \leq i \leq N, 1 \leq j \leq M$)の角度の最小値が最も大きくなるように、定位位置 Q_m が設定される。定位位置 Q_m を現実参加者の位置からできるだけ離す(角度を分散させる)ことにより、音声をより聞き分けやすくすることが可能となる。

[0121] 1つの定位可能領域に2人の遠隔参加者の音声の定位位置を設定する場合、それぞれの位置を離すように遠隔参加者の音声の定位位置が調整される。

例えば、図20の例においては、遠隔参加者U11の音声の定位位置が定位可能領域A1内に設定された後に、遠隔参加者U15が会議に参加したことによって、それぞれの位置を離すように、遠隔参加者U11と遠隔参加者U15の音声の定位位置が調整されている。

[0122] 遠隔参加者の音声の定位位置がこのようにして設定された後、音声の送受信が開始される。

[0123] <会議の開始後の処理>

・遠隔参加者が発話を行うケース

図21のフローチャートを参照して、会議開始後の全体の処理について説明する。図21に示す処理は、遠隔参加者が発話を行うことに応じて行われる処理である。

[0124] 例えば遠隔参加者U11が発話を行った場合、ステップS131において、クライアント端末2-11の音声送信部222（図12）は、遠隔参加者U11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。

[0125] ステップS121において、コミュニケーション管理サーバ1の音声受信部134（図10）は、クライアント端末2-11から送信されてきた音声データを受信する。

[0126] ステップS122において、音声送信部135は、遠隔参加者U11の音声データをクライアント端末2-1, 2-2, 2-12のそれぞれに送信する。

[0127] ステップS101において、クライアント端末2-1の音声受信部241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を受信する。

[0128] ステップS102において、音像定位処理部243は、発話を行った遠隔参加者U11の音声の定位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遠隔参加者U11の音声データに対して行う。

[0129] ステップS103において、出力制御部245は、音像定位処理により生成された遠隔参加者U11の音声を現実参加者U1が装着するイヤホン3か

ら出力させる。

[0130] クライアント端末2-2は、ステップS101乃至S103の処理と同様の処理をステップS111乃至S113において行うことによって、遠隔参加者U11の音声を現実参加者U2が装着するイヤホン3から出力させる。

[0131] クライアント端末2-12も同様に、ステップS101乃至S103の処理と同様の処理をステップS141乃至S143において行うことによって、遠隔参加者U11の音声を遠隔参加者U12が装着するイヤホン3から出力させる。

[0132] これにより、現実参加者である現実参加者U1、U2と、他の遠隔参加者である遠隔参加者U12は、遠隔参加者U11の音声を聞くことができる。遠隔参加者U11の音声が現実参加者U1、U2から離れた位置に定位して感じられるから、現実参加者U1、U2と遠隔参加者U12は、遠隔参加者U11の音声と他の参加者の音声とを聞き分けることができる。

[0133] ・現実参加者が発話を行うケース1

図22のフローチャートを参照して、会議開始後の全体の他の処理について説明する。図22に示す処理は、現実参加者が発話を行うことに応じて行われる処理である。

[0134] 例えば現実参加者U1が発話を行った場合、ステップS151において、クライアント端末2-1の音声送信部222は、現実参加者U1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。

[0135] ステップS161において、コミュニケーション管理サーバ1の音声受信部134は、クライアント端末2-1から送信されてきた音声データを受信する。

[0136] ステップS162において、音声送信部135は、現実参加者U1の音声データをクライアント端末2-11、2-12のそれぞれに送信する。

[0137] ステップS171において、クライアント端末2-11の音声受信部241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を受信する。

- [0138] ステップS 1 7 2において、音像定位処理部 2 4 3は、発話を行った現実参加者U 1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現実参加者U 1の音声データに対して行う。
- [0139] ステップS 1 7 3において、出力制御部 2 4 5は、音像定位処理により生成された現実参加者U 1の音声を遠隔参加者U 1 1が装着するイヤホン3から出力させる。
- [0140] クライアント端末 2 - 1 2は、ステップS 1 7 1乃至S 1 7 3の処理と同様の処理をステップS 1 8 1乃至S 1 8 3において行うことによって、現実参加者U 1の音声を遠隔参加者U 1 2が装着するイヤホン3から出力させる。
- [0141] これにより、遠隔参加者である遠隔参加者U 1 1, U 1 2は、現実参加者U 1の音声を聞くことができる。
- [0142] ・現実参加者が発話を行うケース2
- 上述したように、密閉型のヘッドホンなどを例えば現実参加者が装着している場合、発話を行った他の現実参加者の音声は、直接ではなく、コミュニケーション管理サーバ1を介して届けられる。例えば密閉型のヘッドホンを装着している現実参加者が使用するクライアント端末2においては、他の現実参加者の音声に対して音像定位処理が行われる。
- [0143] なお、現実参加者が開放型のイヤホン3を装着している場合であっても、他の現実参加者の音声がコミュニケーション管理サーバ1を介して届けられ、音像定位処理が行われるようにしてもよい。
- [0144] 図 2 3のフローチャートを参照して、会議開始後の全体の他の処理について説明する。図 2 3に示す処理は、現実参加者U 1が発話を行うことに応じて行われる処理である。現実参加者U 2が密閉型のヘッドホンを装着しているものとする。
- [0145] 図 2 3に示す処理は、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対する音像定位処理がクライアント端末 2 - 2において行われる点で、図 2 2に示す処理と異なる。重複する説明については適宜省略す

る。

[0146] コミュニケーション管理サーバ1からステップS162において送信された現実参加者U1の音声データは、クライアント端末2-2に対しても送信される。

[0147] ステップS201において、クライアント端末2-2の音声受信部241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を受信する。

[0148] ステップS202において、音像定位処理部243は、発話を行った現実参加者U1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現実参加者U1の音声データに対して行う。

[0149] ステップS203において、出力制御部245は、音像定位処理により生成された現実参加者U1の音声を現実参加者U2が装着するイヤホン3から出力させる。

[0150] 以上のように、会議参加者の位置情報に基づいて、現実の参加者とずらした位置から聞こえるように遠隔参加者の音声を提示することにより、それぞれの参加者は、現実の参加者のいない位置から遠隔参加者の音声を聞くことができ、音声の聞き分けが容易になる。

[0151] また、遠隔参加者の音声を提示可能な領域を会議参加者の位置情報に基づいて算出し、定位可能領域内に音声の定位位置を分散して配置することにより、それぞれの参加者は、音声の聞き分けが容易になる。

[0152] <<変形例>>

<除外領域の設定>

遠隔参加者の音声の定位位置とするべきではない領域が会議室などの空間にある場合、そのような領域を除外するように定位可能領域が設定され、遠隔参加者の音声の定位位置が設定される。すなわち、遠隔参加者の音声の定位位置とするべきではない領域が除外領域として設定される。

[0153] 例えば、現実参加者がいる会議室の形状や大きさが分かっている場合、コミュニケーション管理サーバ1においては、会議室の外の領域や壁がある領

域を除外するように、定位可能領域が設定される。

[0154] 図24は、定位可能領域の設定の例を示す図である。

[0155] 図24に示すように、現実参加者の位置に応じて設定された近似円Cに重なるように会議室の壁Wがある場合、壁Wの表面より右側の領域を除外するようにして定位可能領域が設定される。壁Wの位置などの環境についても、現実参加者の位置とともにクライアント端末2のユーザ位置検出部223により検出され、コミュニケーション管理サーバ1に提供される。例えば、カメラ208により撮影された画像に基づいて、現実参加者がいる環境がユーザ位置検出部223により検出される。

[0156] 図24の例においては、壁Wの近くにいる現実参加者U1と現実参加者U2の間、現実参加者U2と現実参加者U3の間、および、現実参加者U3と壁Wの間にそれぞれ定位可能領域A11、A12、A13が設定されている。近似円C上の、壁Wの表面より右側の領域には定位可能領域が設定されていない。

[0157] 遠隔参加者の音声の定位位置がこのようにして設定された定位可能領域に基づいて設定されることにより、遠隔参加者の音声が、壁がある位置から聞こえてくる状況になることを防ぐことができる。

[0158] 定位可能領域の設定対象から除外される領域は、壁がある領域に限られるものではない。遠隔参加者の音声の定位位置とすべきではない様々な領域が、定位可能領域の設定対象から除外される。

[0159] 例えば、会議の参加者が歩道を歩いている環境において、車道の方が検出された場合、事故を避けたり、騒音による聞き取りにくさを避けたりするため、車道側の領域は、定位可能領域の設定対象から除外される。

[0160] また、会議の参加者が駅のホームにいる環境において、線路がある場所が検出された場合、線路側の領域は、定位可能領域の設定対象から除外される。

[0161] さらに、会議の参加者が、テーマパークなどの広い空間を使ったエンターテインメント施設にいる環境において、立ち入り禁止区域が検出された場合

、立ち入り禁止区域は、定位可能領域の設定対象から除外される。

[0162] このように、音声の定位位置とした場合に不自然となる領域が、遠隔参加者の音声の定位位置とするべきではない除外領域として設定される。除外領域を除くように定位可能領域が設定されることにより、会議の参加者がいる環境に適した音像定位が可能となる。

[0163] <参加者の増減時の定位位置の移動>

参加者が追加されたり、参加者が会議から退出したりして人数が増減した場合、コミュニケーション管理サーバ1においては、上述した計算が行われ、遠隔参加者の音声の定位位置が更新される。この場合、更新前の位置である位置 P_{old} から、更新後の位置である位置 P_{new} に対する定位位置の移動が生じる。

[0164] ここで、遠隔参加者の音声の定位位置が瞬時に移動してしまうと、例えば現実参加者にとっては、想像していた位置と違う位置から遠隔参加者の音声が聞こえることになり、不自然である。

[0165] 遠隔参加者の音声の定位位置の移動時、定位位置の移動がアニメーションによって提示される。音声のアニメーションは、例えば、音像定位処理に用いるHRTFデータを変更し、位置 P_{old} から位置 P_{new} までの経路に沿って、音声の定位位置（音源の位置）を順次移動させることによって行われる。

[0166] 図25は、音声のアニメーションの例を示す図である。

[0167] 音声のアニメーション時、音源を位置 P_{old} から位置 P_{new} まで直線的に移動させた場合、会話の輪の中心近くを横切り、不自然な会話となってしまうことがある。そのため、図25の太線矢印で示すように、例えば、音源の移動は、位置 P_{old} から位置 P_{new} まで、近似円Cの円弧上を移動するようにして行われる。

[0168] 音源の位置の移動経路として円弧状の経路が設定され、基準となる位置である近似円Cの中心位置との距離を維持しながら音源の位置が移動することにより、近似円Cに沿って形成される会話の輪を保つことが可能となる。

[0169] 円弧状の経路上に現実参加者がいる場合、アニメーションとして出力され

る音声の音源が現実参加者と重なってしまい、違和感を生じさせてしまう。そのため、図26に示すように、音源の移動経路は、現実参加者の位置を回避するように設定される。

[0170] 図26の例においては、位置 P_{old} を移動開始位置として、近似円Cの中心から離れた位置である位置P31に移動し、近似円Cの中心から位置P31までの距離を維持しながら円弧状に移動する経路が設定されている。また、近似円Cの中心と位置 P_{new} を通る直線上の位置P32に到達したときに、位置P32から位置 P_{new} に移動する経路が設定されている。

[0171] このような移動経路に沿って音源が移動することにより、違和感を生じさせることなく、自然な形で、遠隔参加者の音声の定位位置を移動させることが可能になる。

[0172] 以上のような音源の移動は、現実参加者の人数が増減した場合と遠隔参加者の人数が増減した場合のうちの少なくともいずれかの場合だけでなく、例えば、現実参加者が移動した場合などにも行われる。音源の移動が参加者の各種の状況の変化に応じて行われるようにすることが可能である。

[0173] <その他の例>

・画面表示

それぞれの参加者が使用するクライアント端末2のディスプレイ207に会議画面が表示され、それにより、それぞれの参加者の位置関係が提示されるようにしてもよい。

[0174] 図27は、会議画面の例を示す図である。

[0175] 図27に示すように、会議画面には、会議を行う場所を表す背景画像に重ねて、それぞれの参加者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、それぞれの参加者の位置に応じた位置となる。

[0176] 図27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、クライアント端末2を使用する参加者の位置から、それぞれの参加者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 乃至 4 は、それぞれ現実参加者または遠隔参加者を表す。

[0177] ・レンダリングの例

レンダリングとバイノーラル処理を含む音像定位処理がクライアント端末 2 により行われるものとしたが、コミュニケーション管理サーバ 1 により行われるようにしてもよい。すなわち、音像定位処理がクライアント端末 2 側で行われるようにしてもよいし、コミュニケーション管理サーバ 1 側で行われるようにしてもよい。

[0178] 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、図 12 の再生処理部 221 は、図 28 に示すように、コミュニケーション管理サーバ 1 の情報処理部 121 において実現される。

[0179] 情報処理部 121 の再生処理部 221 を構成する音声受信部 241 (図 12) は、音像定位処理の対象となる参加者の音声データを受信する。また、定位位置取得部 242 は、定位位置設定部 132 により設定された定位位置を取得する。

[0180] 音像定位処理部 243 は、音声受信部 241 により受信された音声データに対して、定位位置取得部 242 により取得された定位位置に応じた HRTF データを用いて音像定位処理を行う。出力制御部 245 は、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をクライアント端末 2 に対して送信し、イヤホン 3 から出力させる。

[0181] このように、音像定位処理がコミュニケーション管理サーバ 1 側で行われることにより、クライアント端末 2 の処理負担を軽減させることが可能となる。

[0182] ・会話の例

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。

[0183] ・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、プログラム記録媒体からインストールされる。

[0184] 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9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。

[0185] 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。

[0186] 本明細書において、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。

[0187] 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。

[0188] 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。

[0189] 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。

[0190] 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。

[0191] 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。

[0192] ・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。

[0193] (1)

所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる音像定位処理部を備える

情報処理装置。

(2)

前記遠隔参加者の音声の音像の定位位置を、前記現実参加者の位置に基づいて設定する定位位置設定部をさらに備える

前記(1)に記載の情報処理装置。

(3)

前記定位位置設定部は、それぞれの前記現実参加者の位置から離れた位置に前記定位位置を設定する

前記(2)に記載の情報処理装置。

(4)

前記定位位置設定部は、複数の前記遠隔参加者のそれぞれの音声の音像の前記定位位置を離れた位置に設定する

前記(2)または(3)に記載の情報処理装置。

(5)

前記定位位置設定部は、前記所定の空間の環境に応じて設定された除外領

域を除く領域内に前記定位位置を設定する

前記（２）乃至（４）のうちのいずれかに記載の情報処理装置。

（６）

前記定位位置設定部は、前記会話の参加者の状況の変化に応じて前記定位位置を移動させる

前記（２）乃至（５）のうちのいずれかに記載の情報処理装置。

（７）

前記定位位置設定部は、移動元の位置から移動先の位置まで、基準となる位置との距離を維持しながら前記定位位置を移動させる

前記（６）に記載の情報処理装置。

（８）

前記定位位置設定部は、移動元の位置から移動先の位置までの経路上に前記現実参加者がいる場合、前記現実参加者の位置を回避して前記定位位置を移動させる

前記（７）に記載の情報処理装置。

（９）

前記音像定位処理部は、前記現実参加者の位置に基づいて設定された定位位置に、前記遠隔参加者の音声の音像を定位させる

前記（１）に記載の情報処理装置。

（１０）

前記音像定位処理部は、それぞれの前記現実参加者の位置から離れた位置に設定された前記定位位置に前記遠隔参加者の音声の音像を定位させる

前記（９）に記載の情報処理装置。

（１１）

前記音像定位処理部は、複数の前記遠隔参加者のそれぞれの音声の音像を離れた位置に定位させる

前記（９）または（１０）に記載の情報処理装置。

（１２）

前記音像定位処理部は、前記所定の空間の環境に応じて設定された除外領域を除く領域内の位置に前記遠隔参加者の音声の音像を定位させる

前記（９）乃至（１１）のうちのいずれかに記載の情報処理装置。

（１３）

前記遠隔参加者の音声を、前記現実参加者が使用する出力デバイスから出力させる出力制御部をさらに備える

前記（１）乃至（１２）のうちのいずれかに記載の情報処理装置。

（１４）

情報処理装置が、

所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる

情報処理方法。

（１５）

コンピュータに、

所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる

処理を実行させるためのプログラム。

符号の説明

[0194] １ コミュニケーション管理サーバ、 ２ クライアント端末、 ３ イヤホン、 １２１ 情報処理部、 １３１ 位置情報取得部、 １３２ 定位位置設定部、 １３３ 定位位置情報送信部、 １３４ 音声受信部、 １３５ 音声送信部、 ２１１ 情報処理部、 ２２１ 再生処理部、 ２２２ 音声送信部、 ２２３ ユーザ位置検出部、 ２４１ 音声受信部、 ２４２ 定位位置取得部、 ２４３ 音像定位処理部、 ２４４ HRTFデータ記憶部、 ２４５ 出力制御部

請求の範囲

- [請求項1] 所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる音像定位処理部を備える
情報処理装置。
- [請求項2] 前記遠隔参加者の音声の音像の定位位置を、前記現実参加者の位置に基づいて設定する定位位置設定部をさらに備える
請求項1に記載の情報処理装置。
- [請求項3] 前記定位位置設定部は、それぞれの前記現実参加者の位置から離れた位置に前記定位位置を設定する
請求項2に記載の情報処理装置。
- [請求項4] 前記定位位置設定部は、複数の前記遠隔参加者のそれぞれの音声の音像の前記定位位置を離れた位置に設定する
請求項2に記載の情報処理装置。
- [請求項5] 前記定位位置設定部は、前記所定の空間の環境に応じて設定された除外領域を除く領域内に前記定位位置を設定する
請求項2に記載の情報処理装置。
- [請求項6] 前記定位位置設定部は、前記会話の参加者の状況の変化に応じて前記定位位置を移動させる
請求項2に記載の情報処理装置。
- [請求項7] 前記定位位置設定部は、移動元の位置から移動先の位置まで、基準となる位置との距離を維持しながら前記定位位置を移動させる
請求項6に記載の情報処理装置。
- [請求項8] 前記定位位置設定部は、移動元の位置から移動先の位置までの経路上に前記現実参加者がいる場合、前記現実参加者の位置を回避して前記定位位置を移動させる
請求項7に記載の情報処理装置。
- [請求項9] 前記音像定位処理部は、前記現実参加者の位置に基づいて設定され

た定位位置に、前記遠隔参加者の音声の音像を定位させる

請求項 1 に記載の情報処理装置。

[請求項10] 前記音像定位処理部は、それぞれの前記現実参加者の位置から離れた位置に設定された前記定位位置に前記遠隔参加者の音声の音像を定位させる

請求項 9 に記載の情報処理装置。

[請求項11] 前記音像定位処理部は、複数の前記遠隔参加者のそれぞれの音声の音像を離れた位置に定位させる

請求項 9 に記載の情報処理装置。

[請求項12] 前記音像定位処理部は、前記所定の空間の環境に応じて設定された除外領域を除く領域内の位置に前記遠隔参加者の音声の音像を定位させる

請求項 9 に記載の情報処理装置。

[請求項13] 前記遠隔参加者の音声を、前記現実参加者が使用する出力デバイスから出力させる出力制御部をさらに備える

請求項 1 に記載の情報処理装置。

[請求項14] 情報処理装置が、

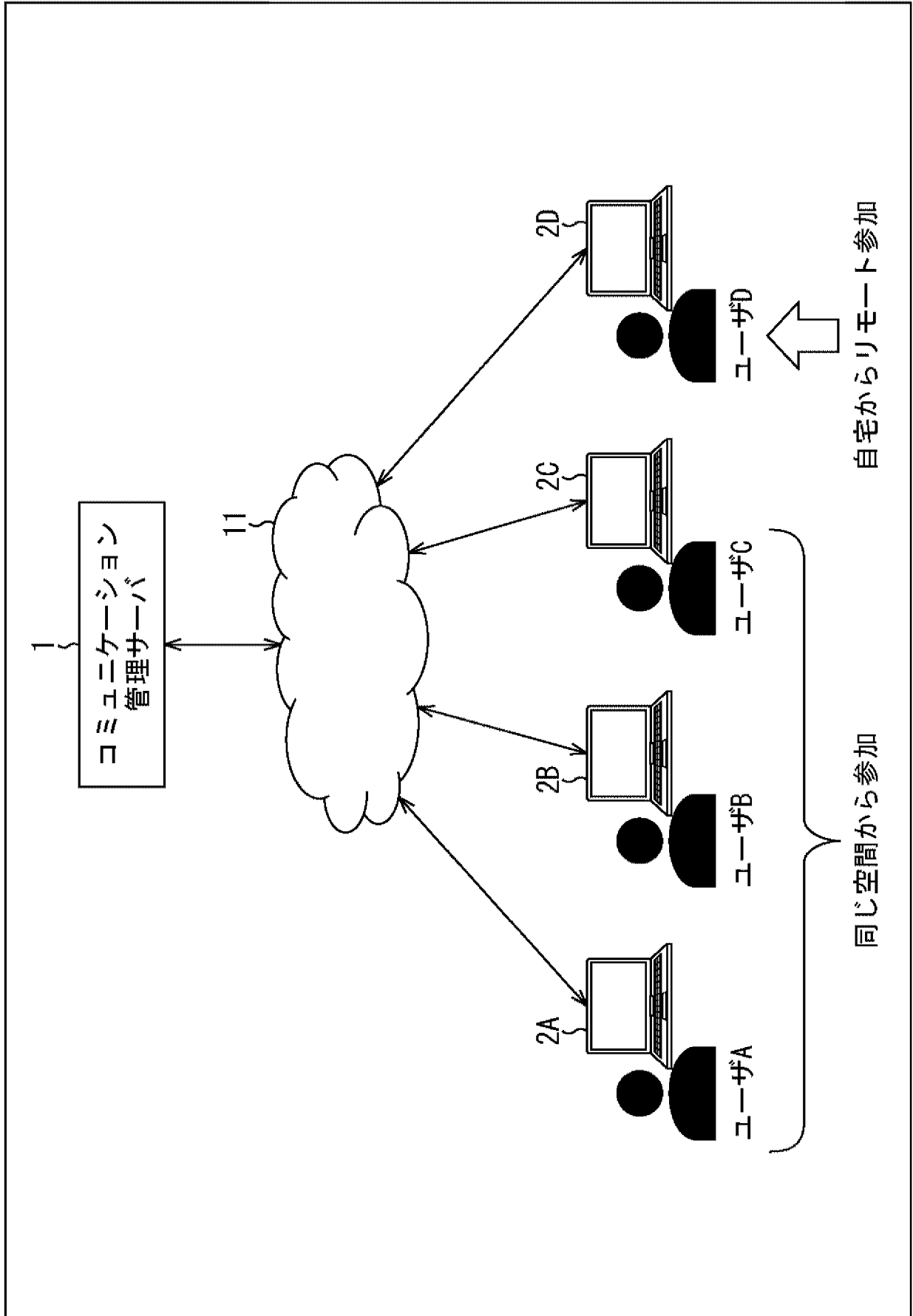
所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる

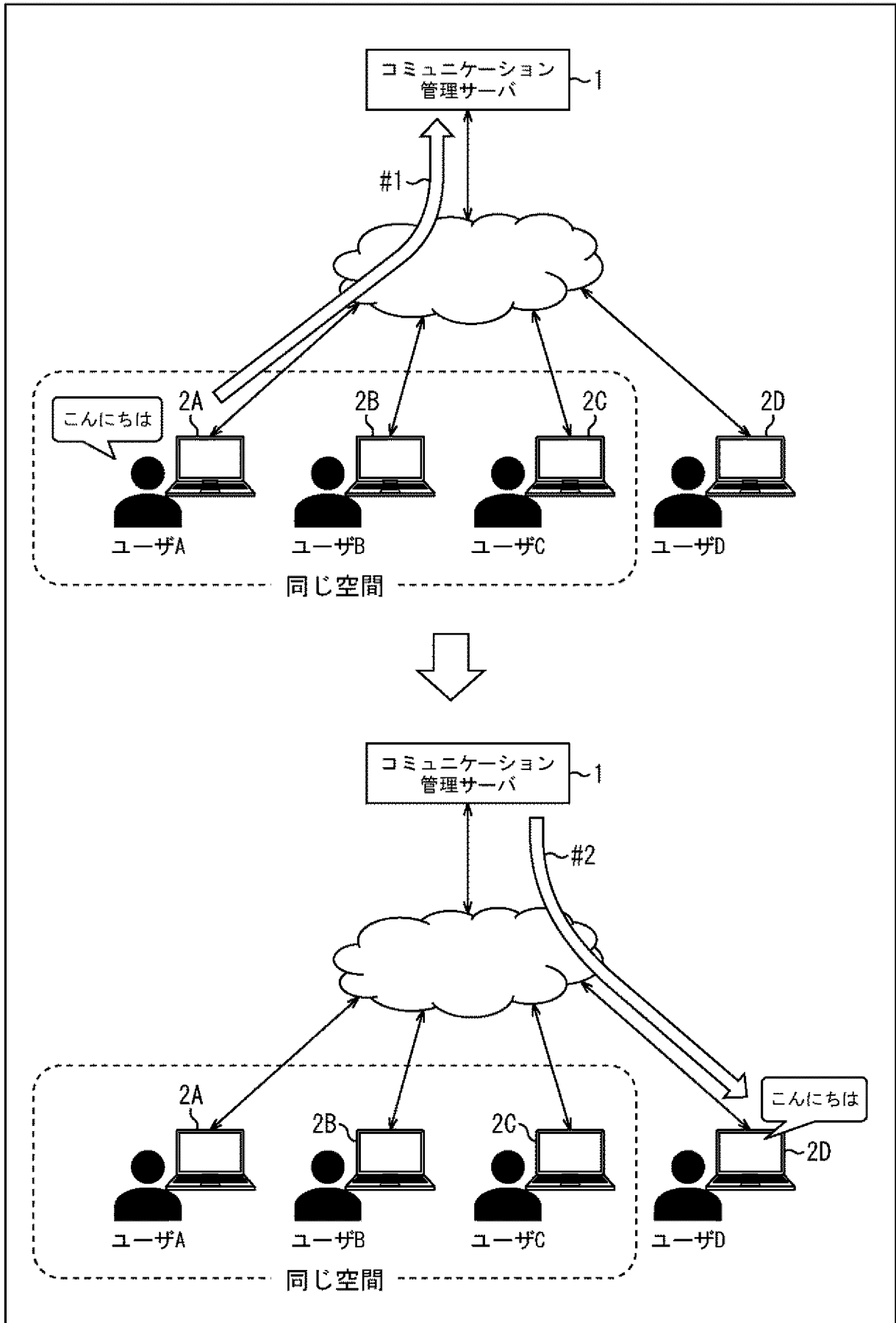
情報処理方法。

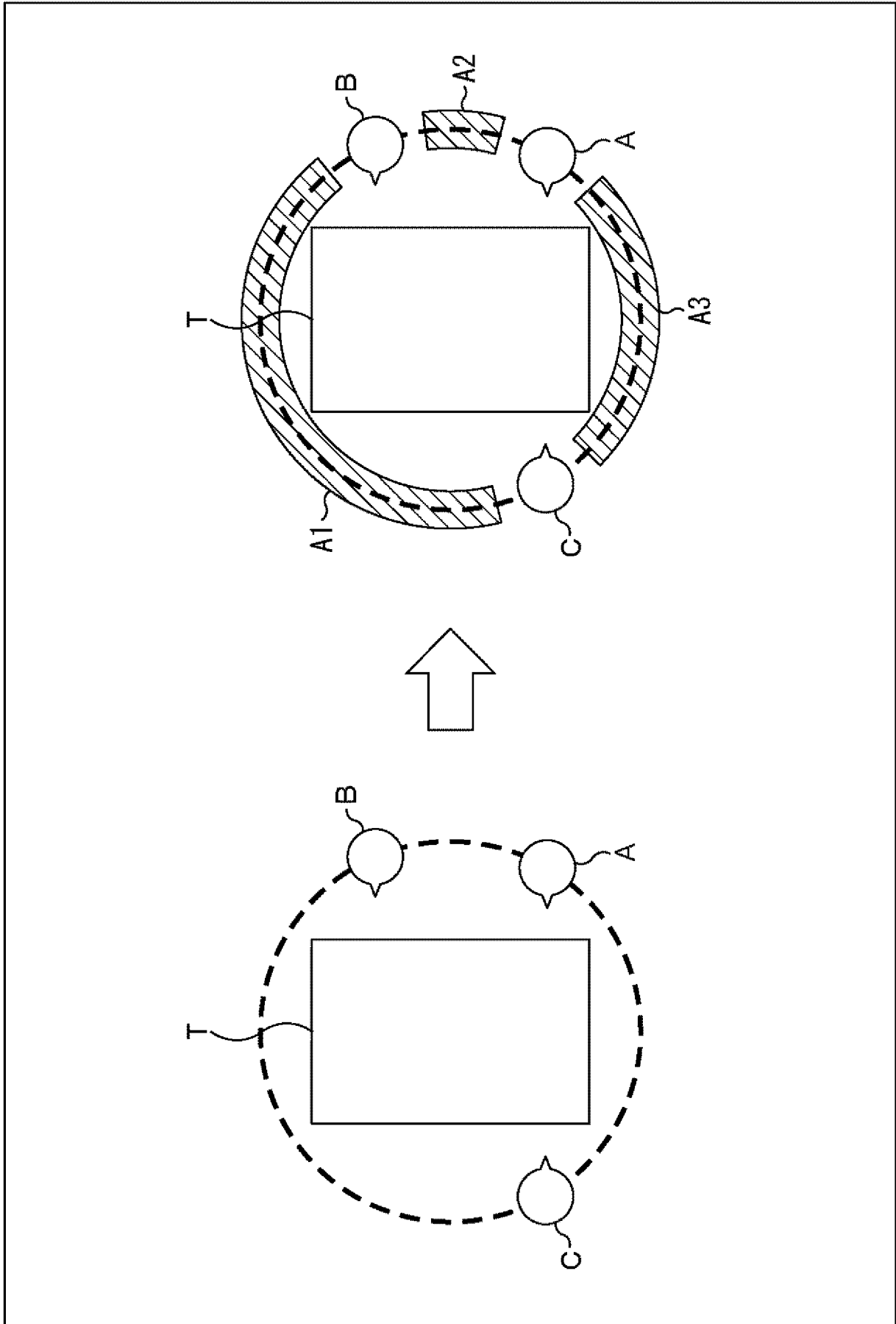
[請求項15] コンピュータに、

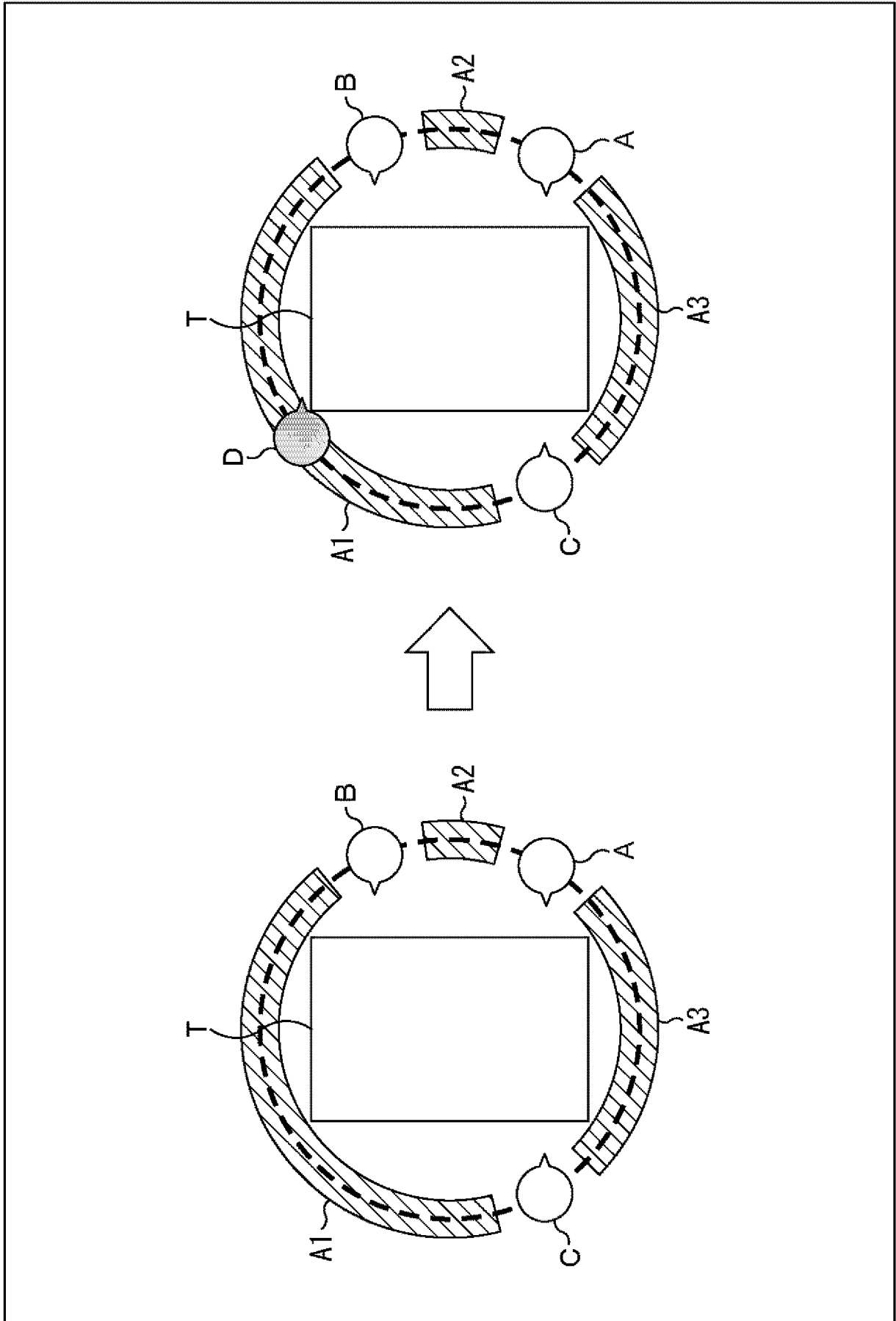
所定の空間で行われる会話に遠隔から参加する遠隔参加者の音声の音像を、前記所定の空間にいる参加者である現実参加者の位置と異なる位置に定位させる

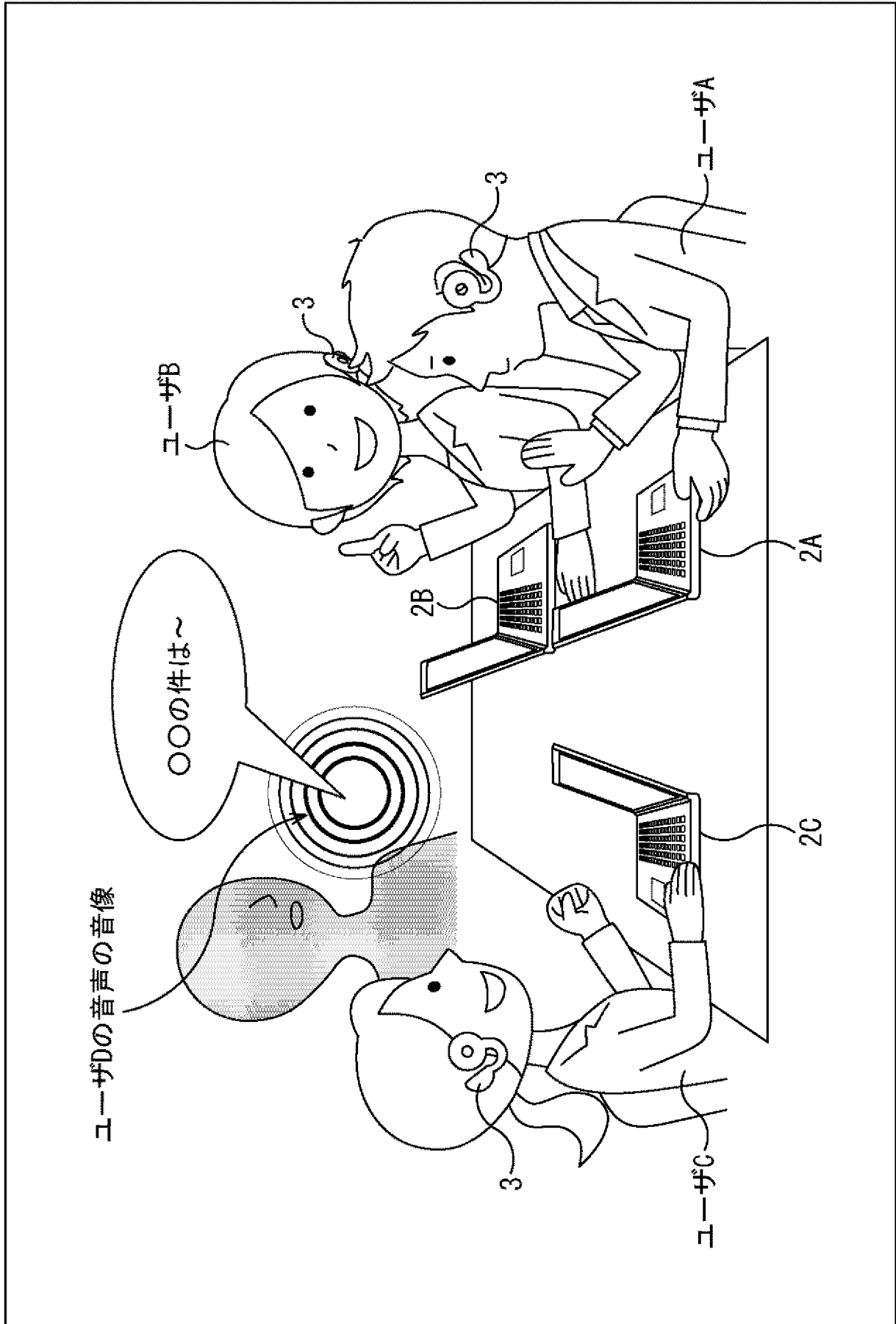
処理を実行させるためのプログラム。

[図1]
FIG. 1

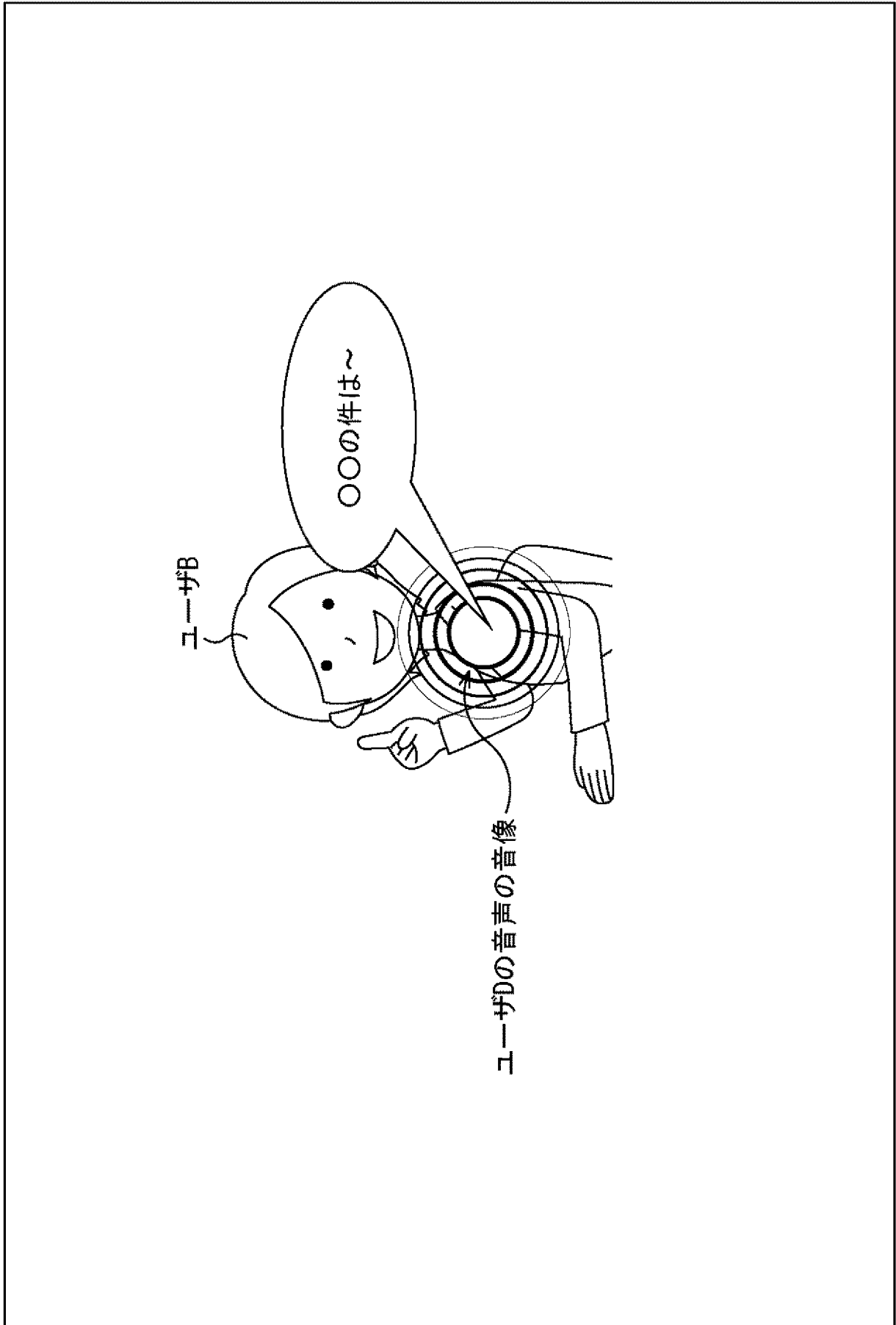
[図2]
FIG. 2

[図3]
FIG. 3

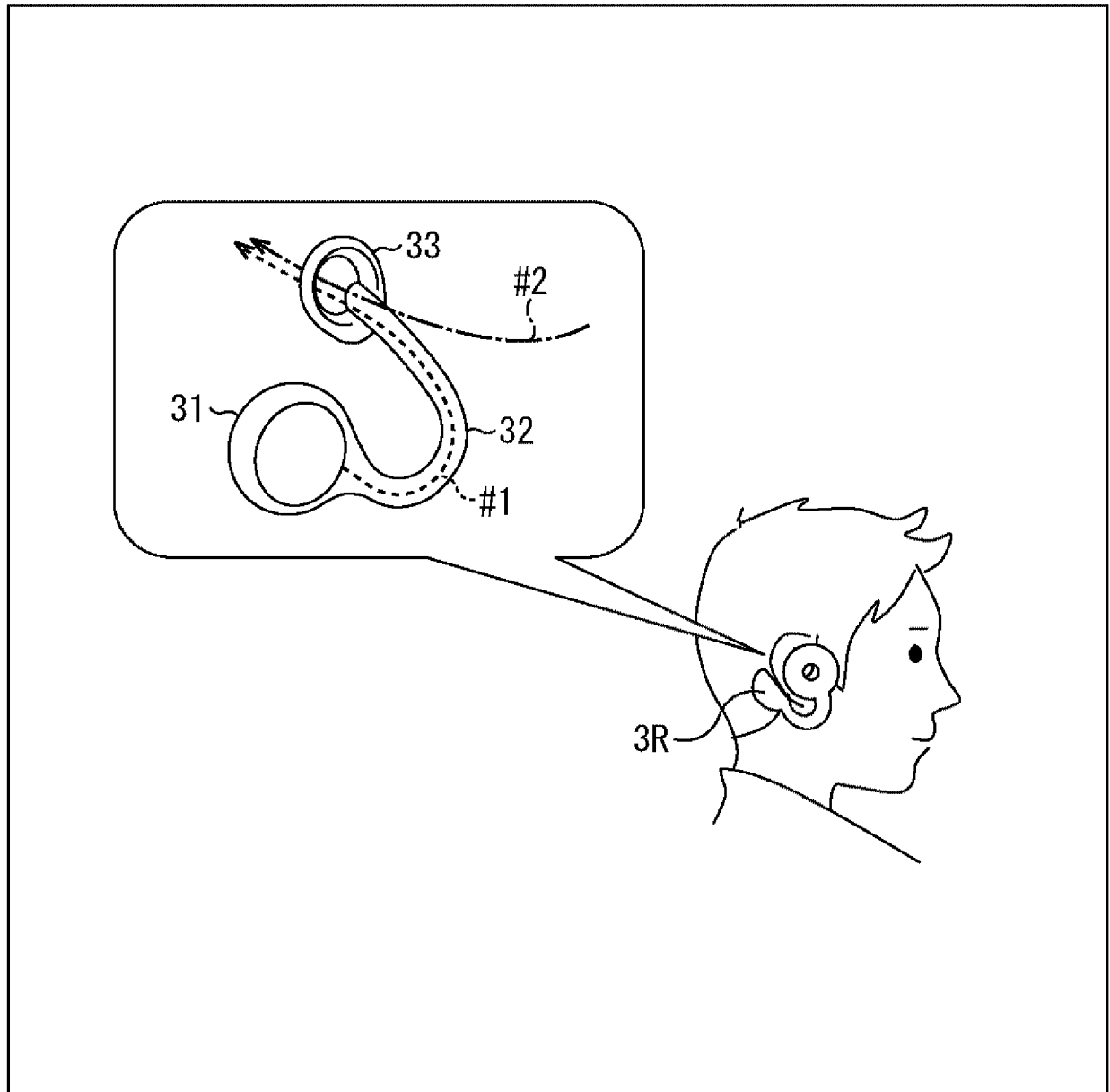
[図4]
FIG. 4

[図5]
FIG. 5

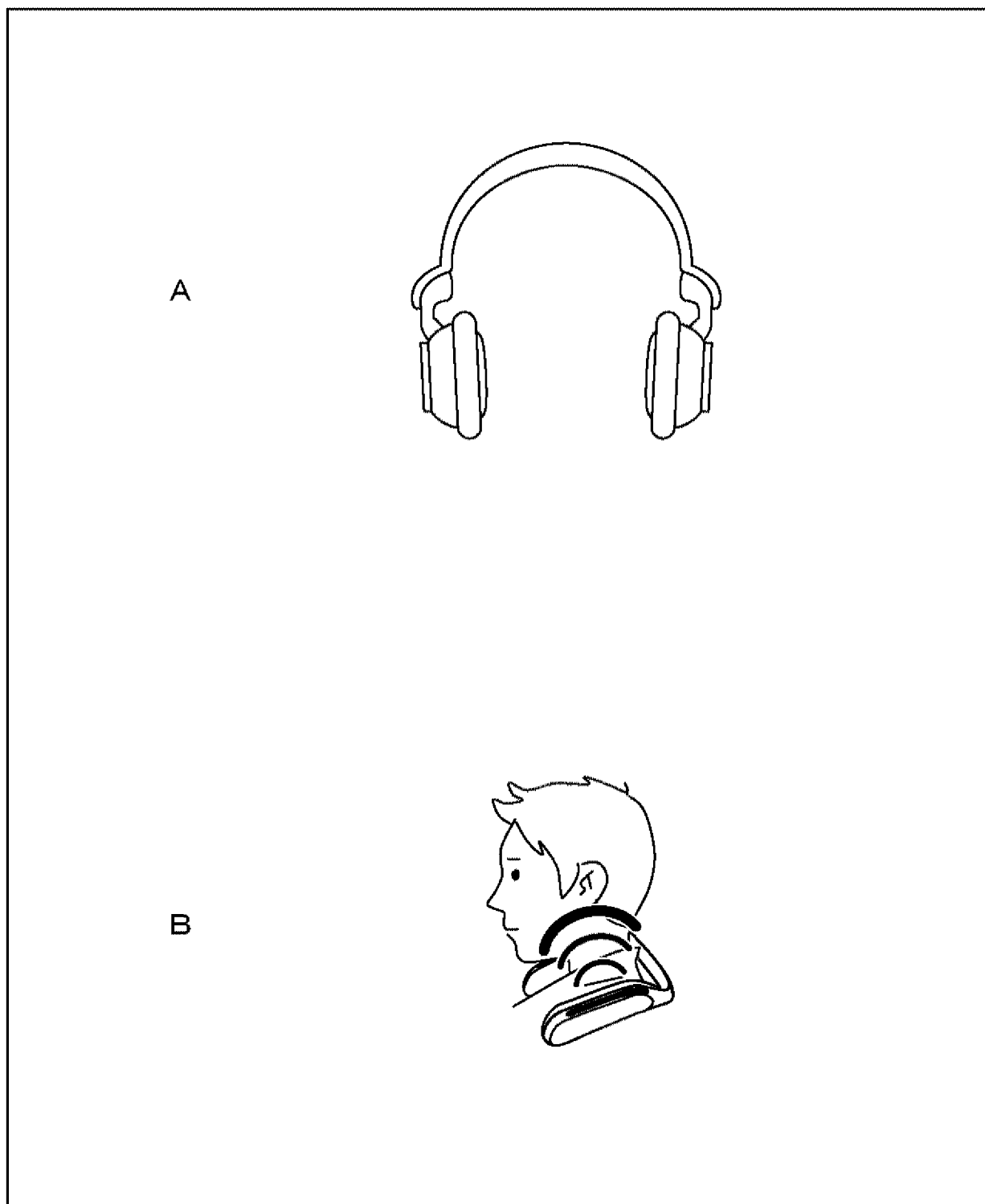
[図6]
FIG. 6

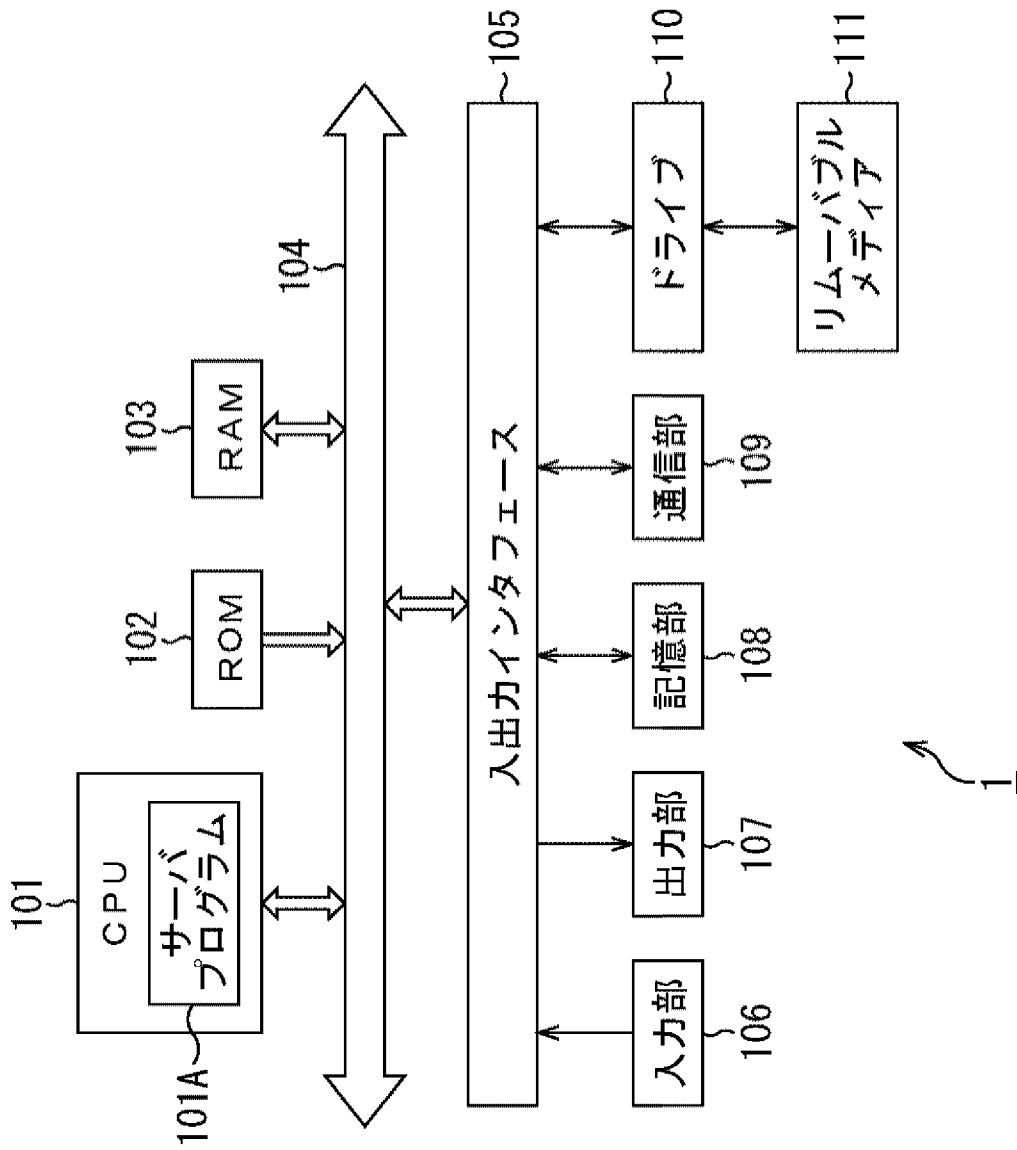


[図7]
FIG. 7

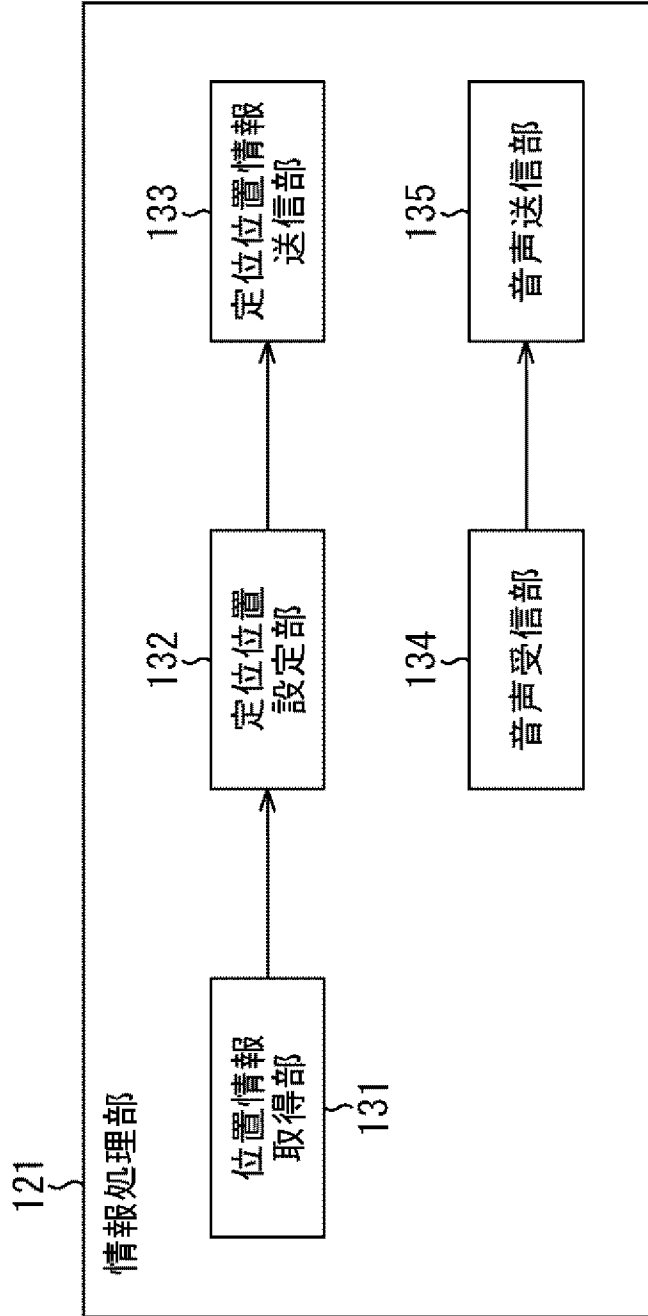


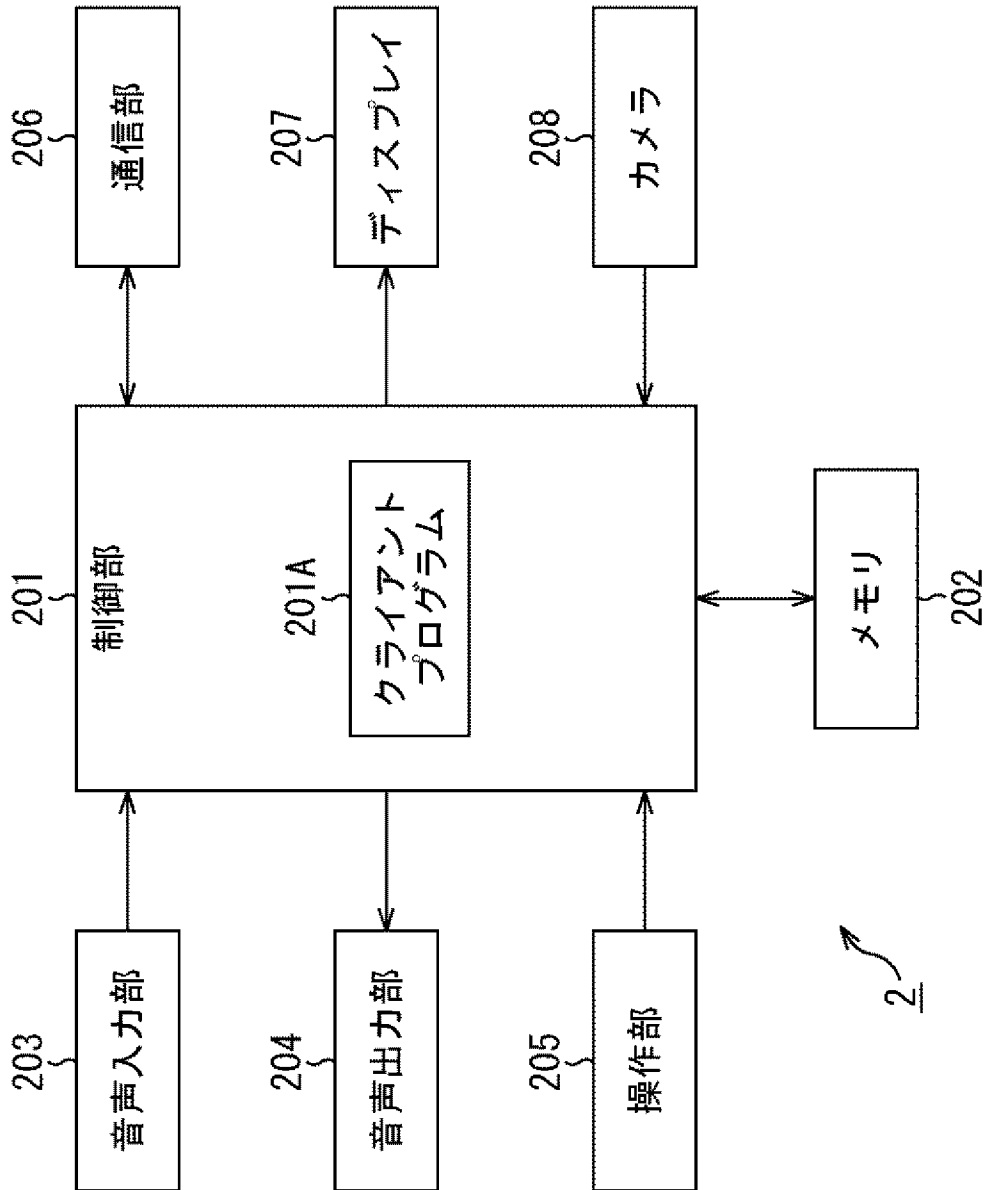
[図8]
FIG. 8

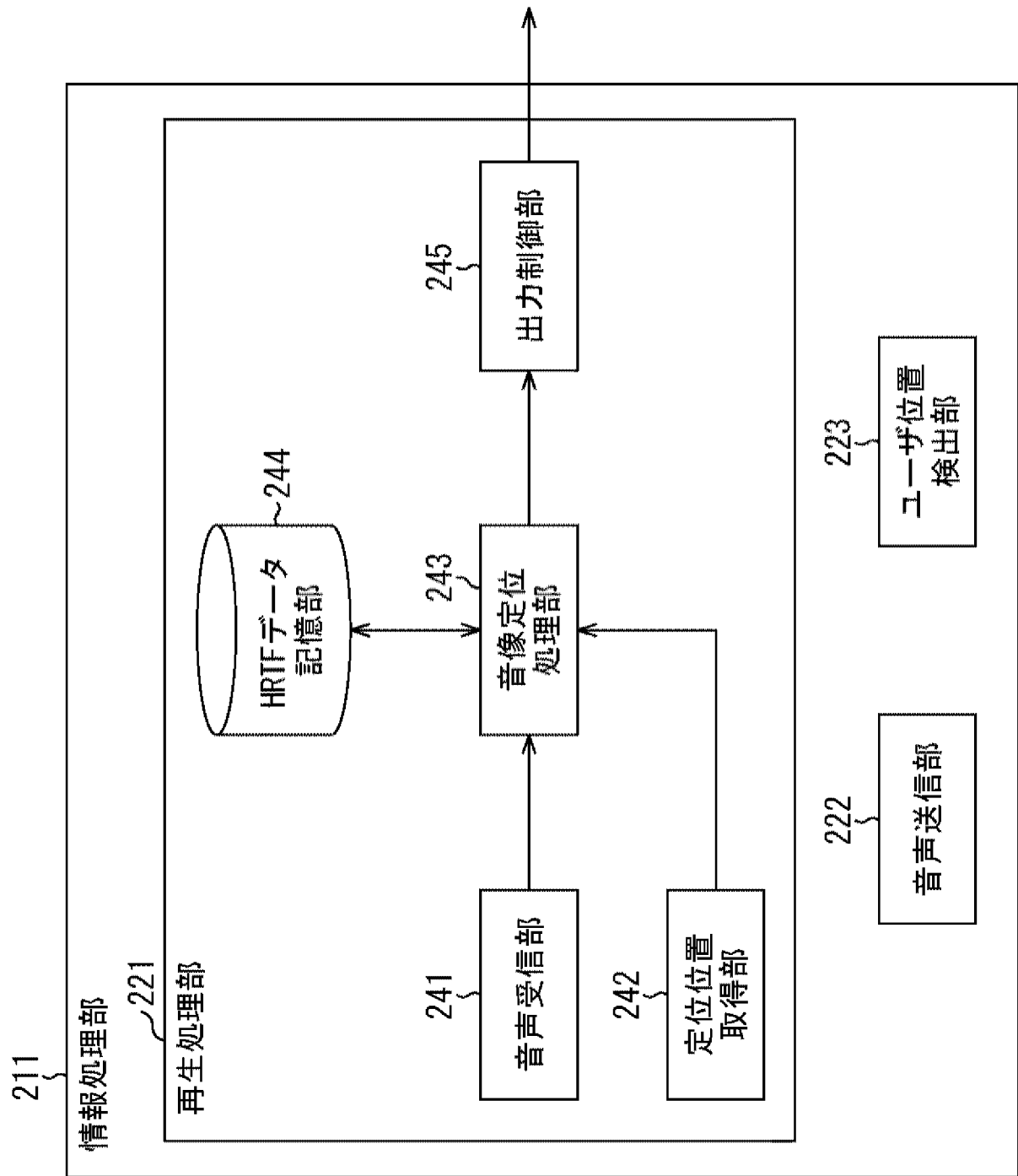


[図9]
FIG. 9

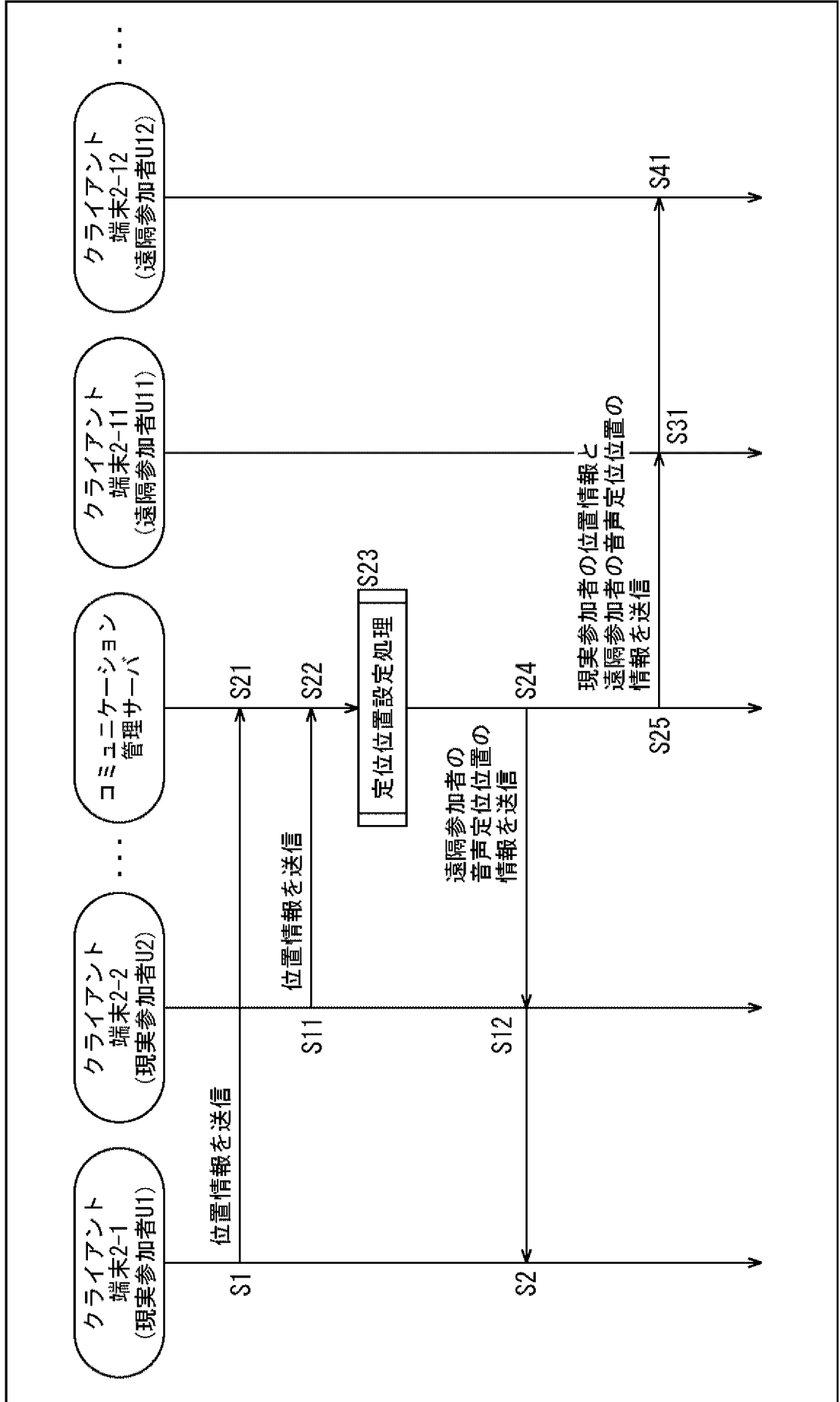
[図10]
FIG. 10



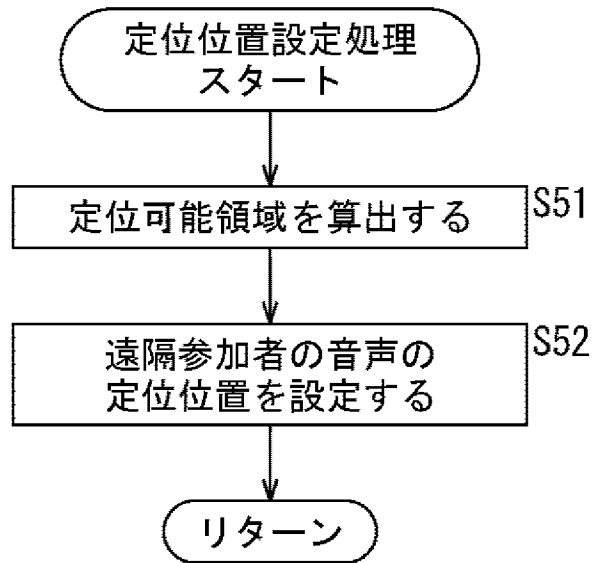
[図11]
FIG. 11

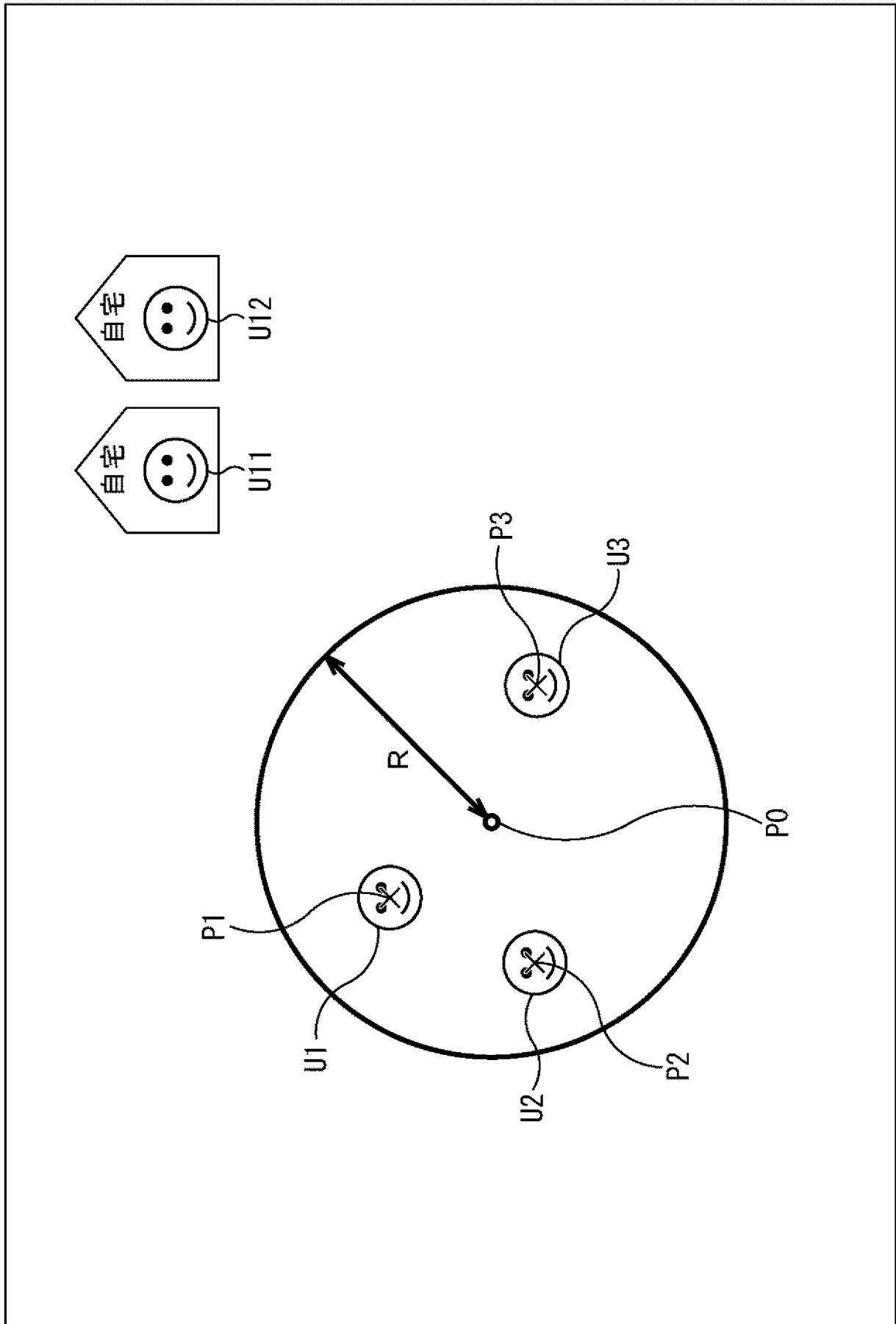
[図12]
FIG. 12

[図13]
FIG. 13

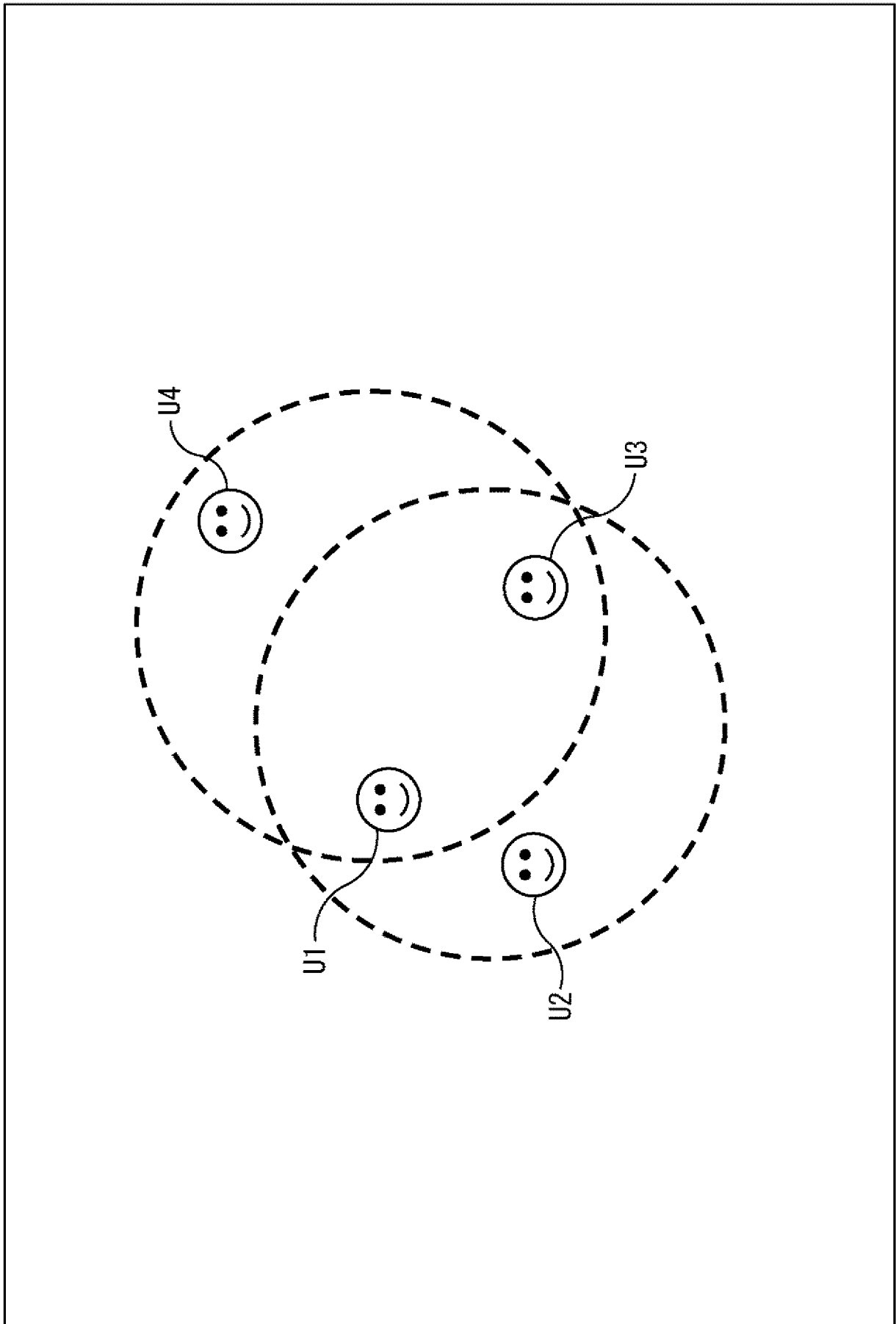


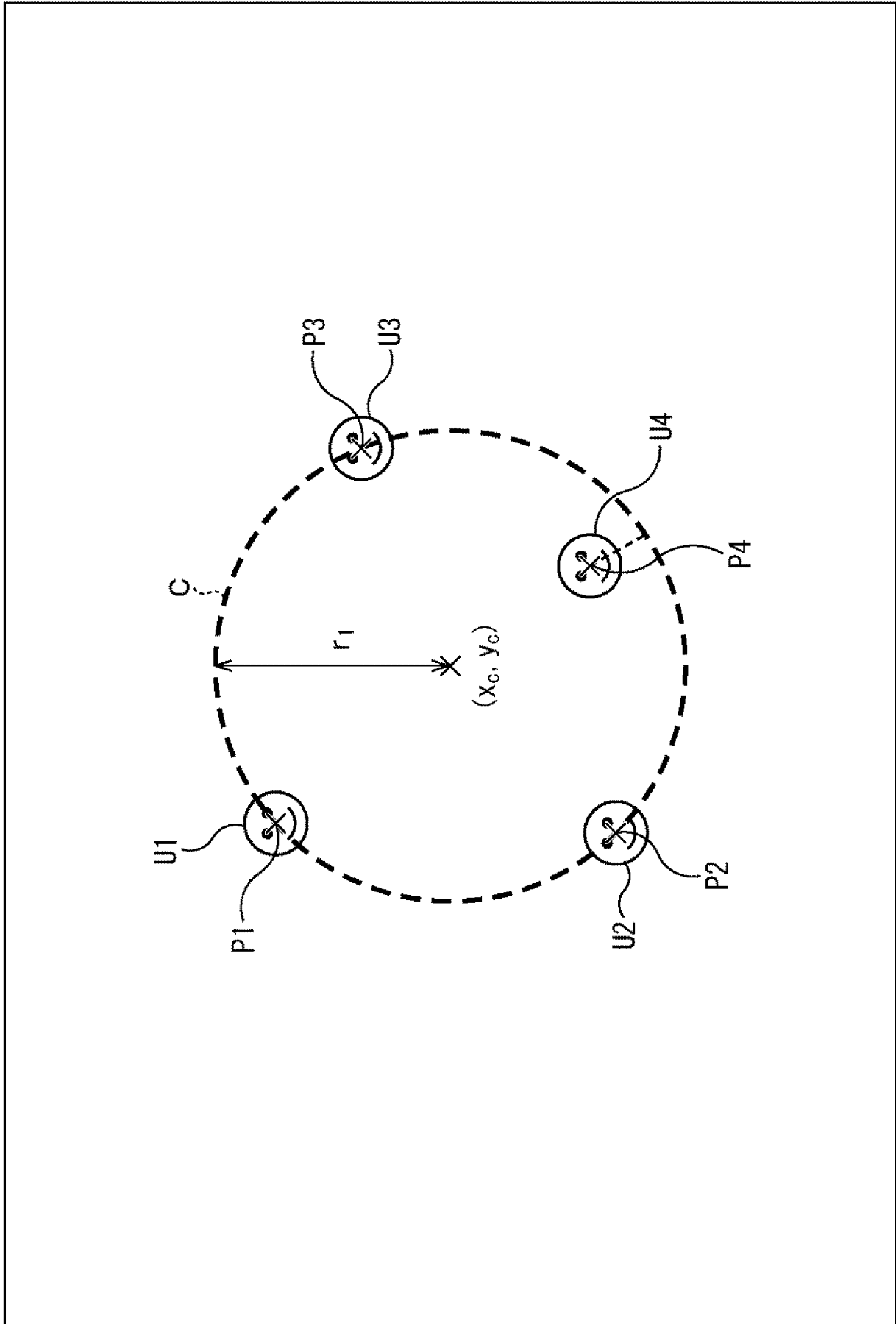
[図14]
FIG. 14

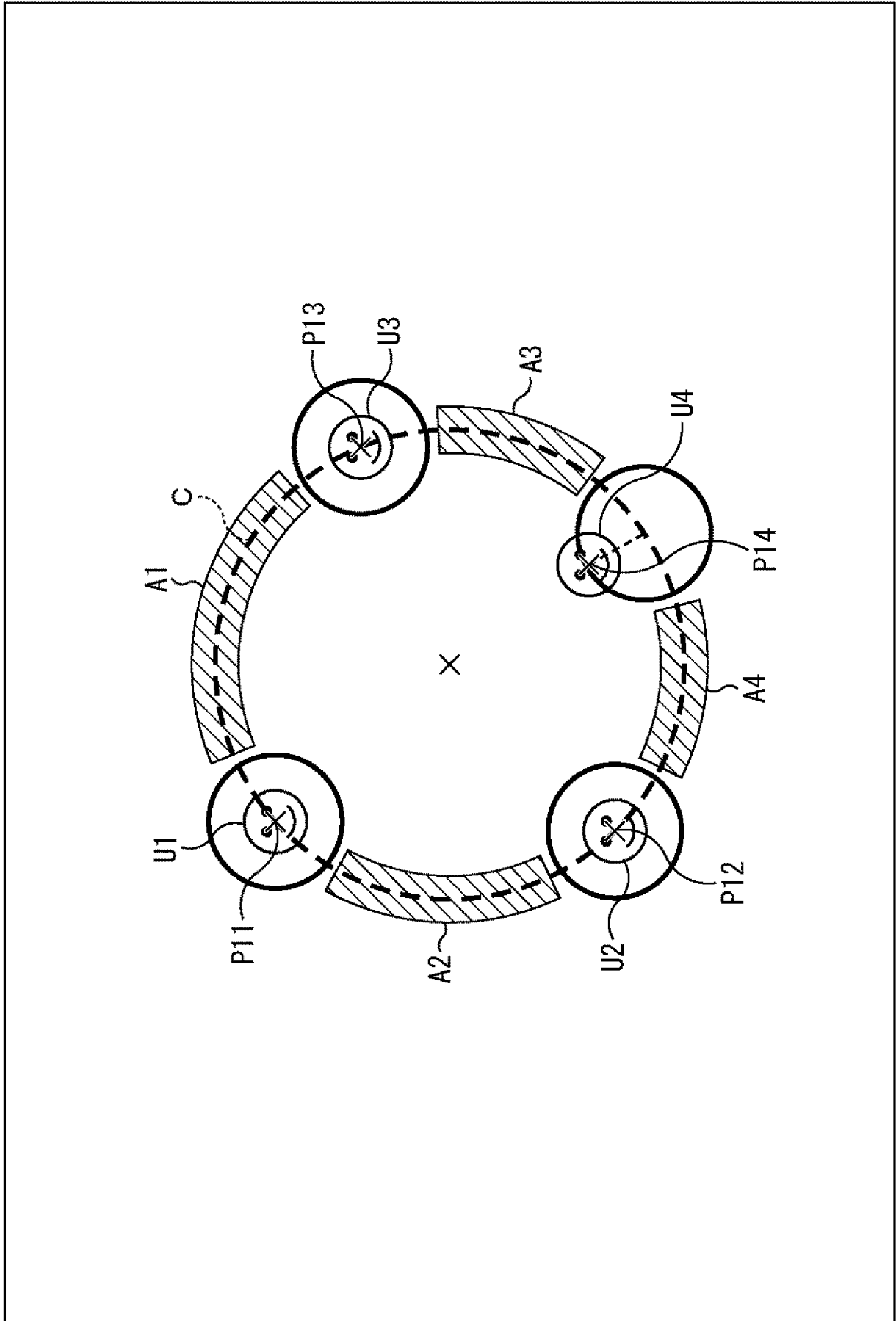



[図15]
FIG. 15

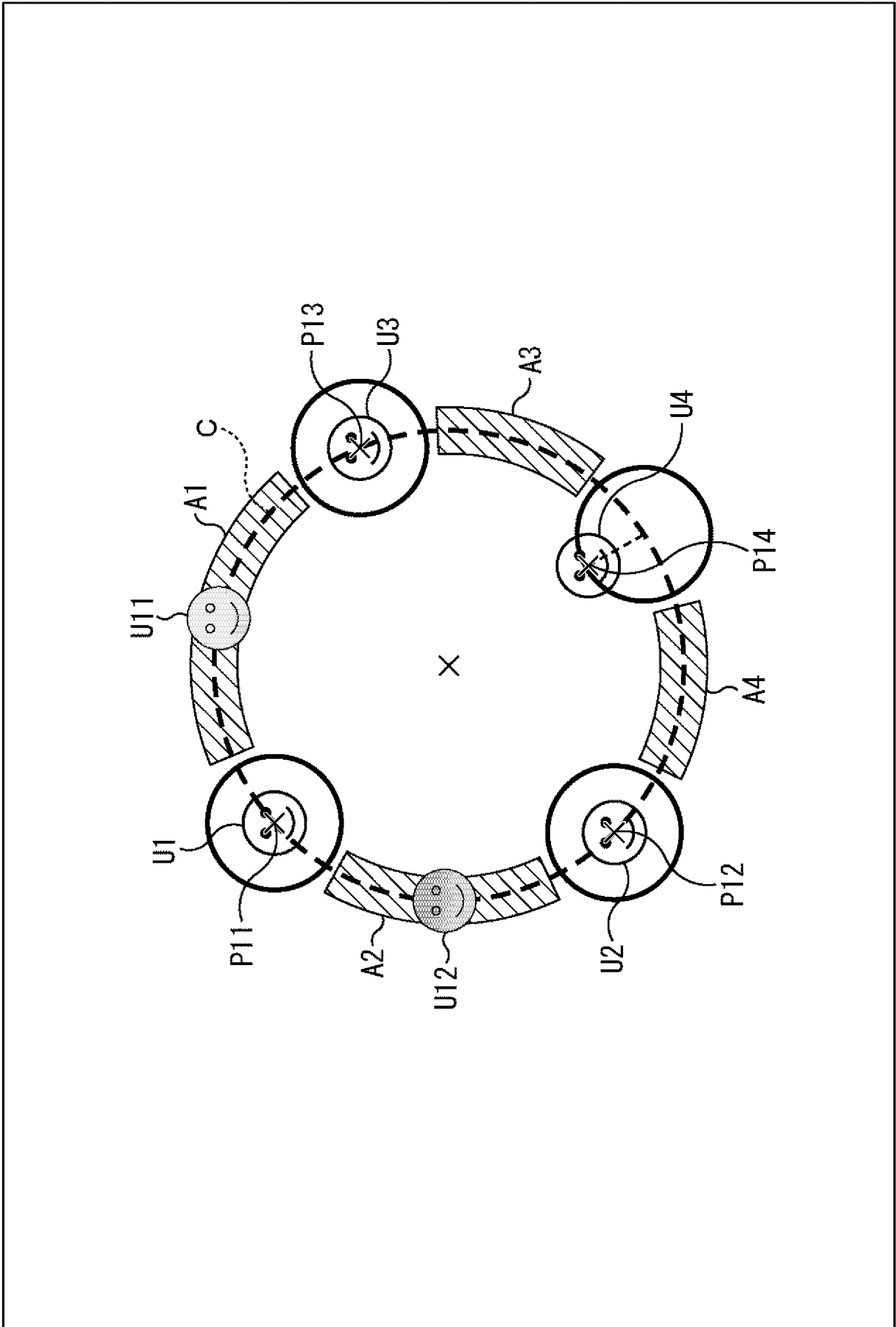
[図16]
FIG. 16




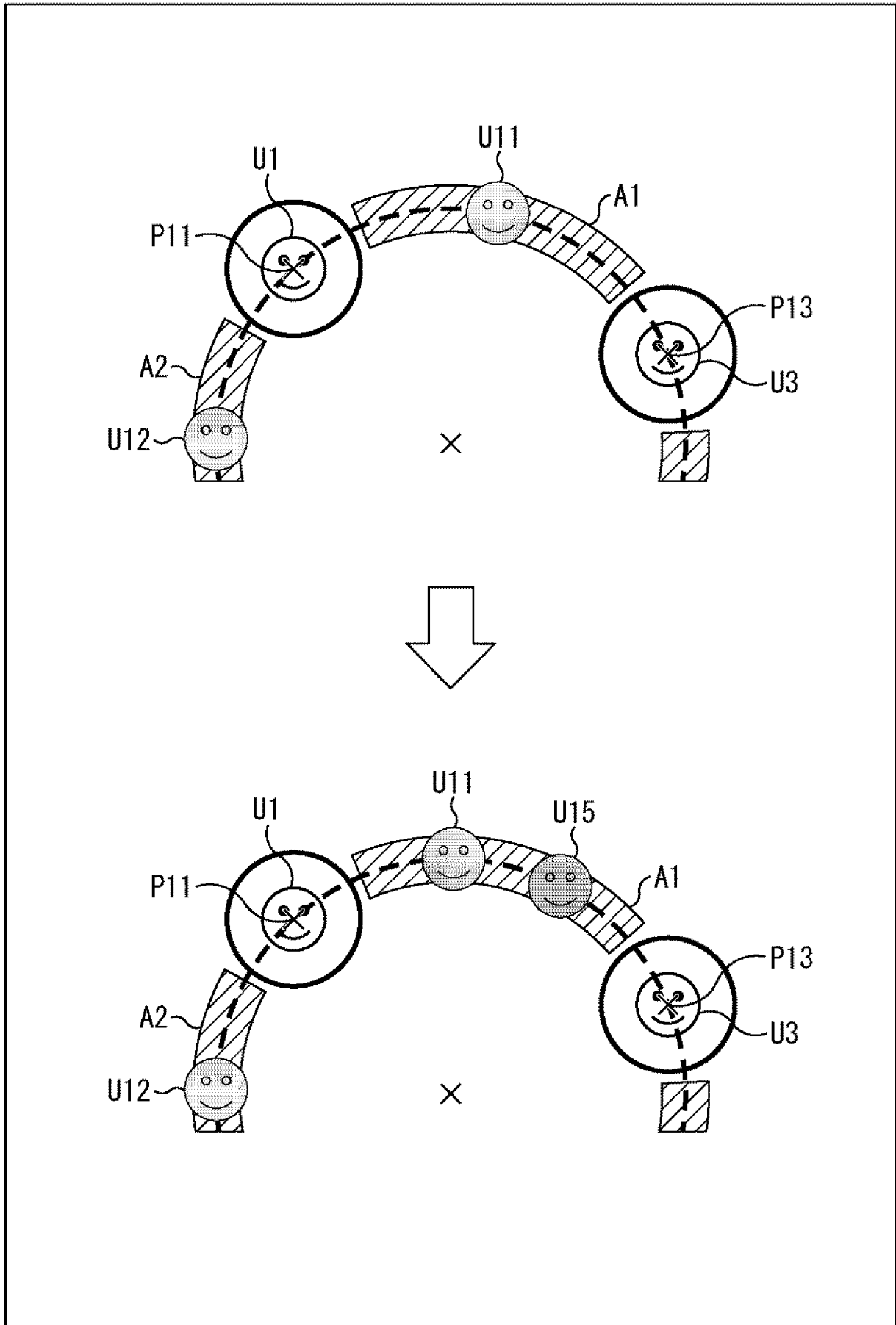
[図17]
FIG. 17

[図18]
FIG. 18

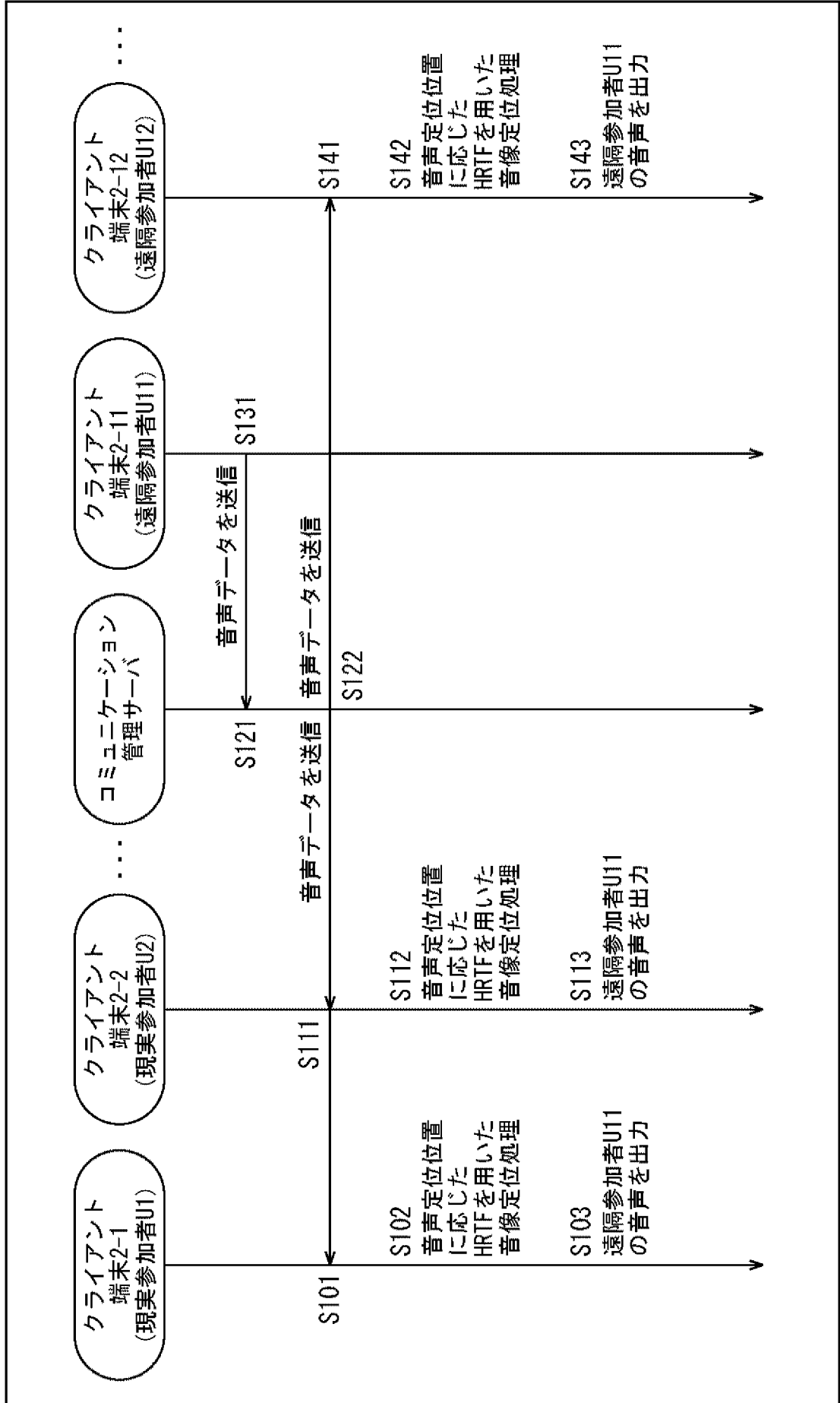
[19]
FIG. 19



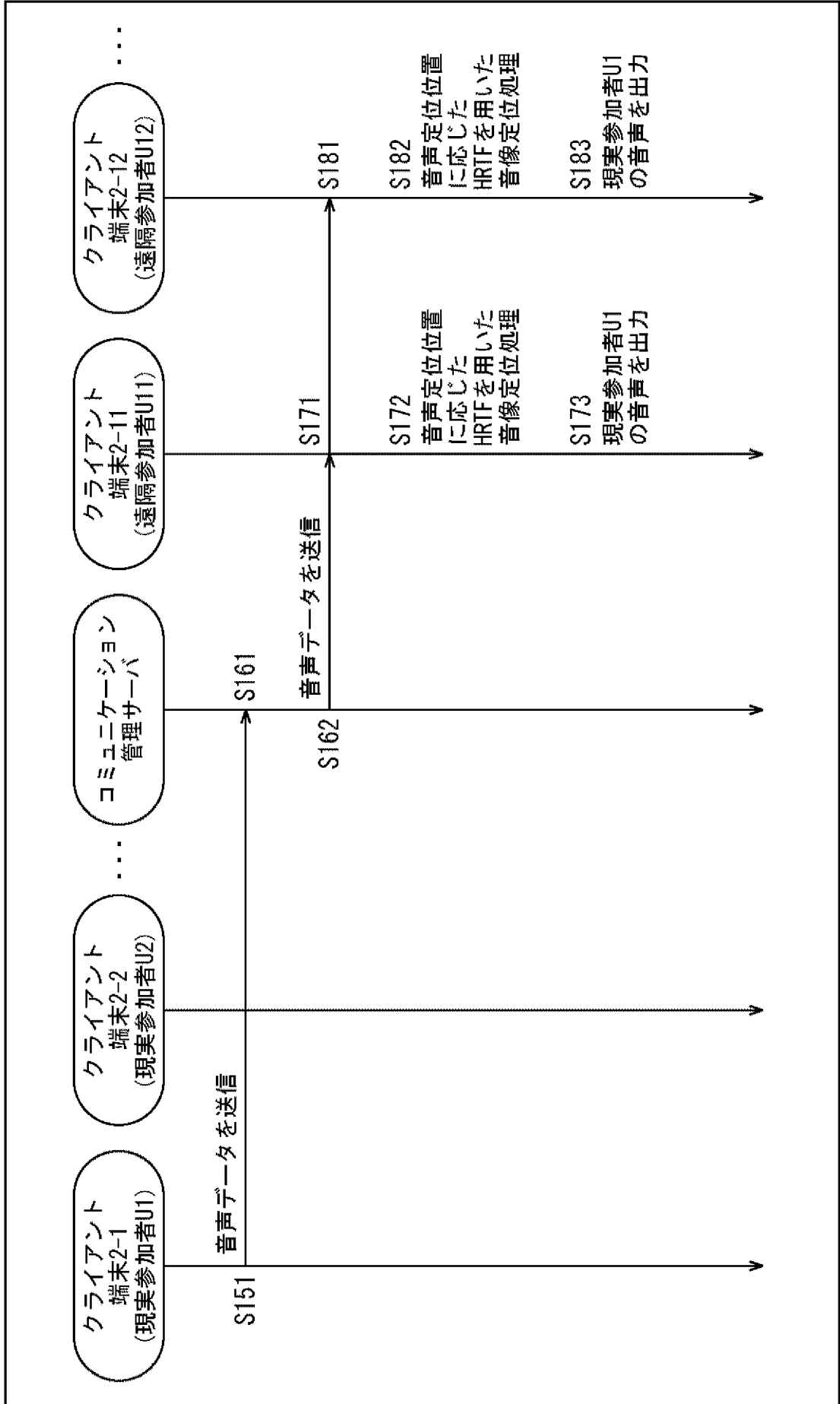
[20]
FIG. 20



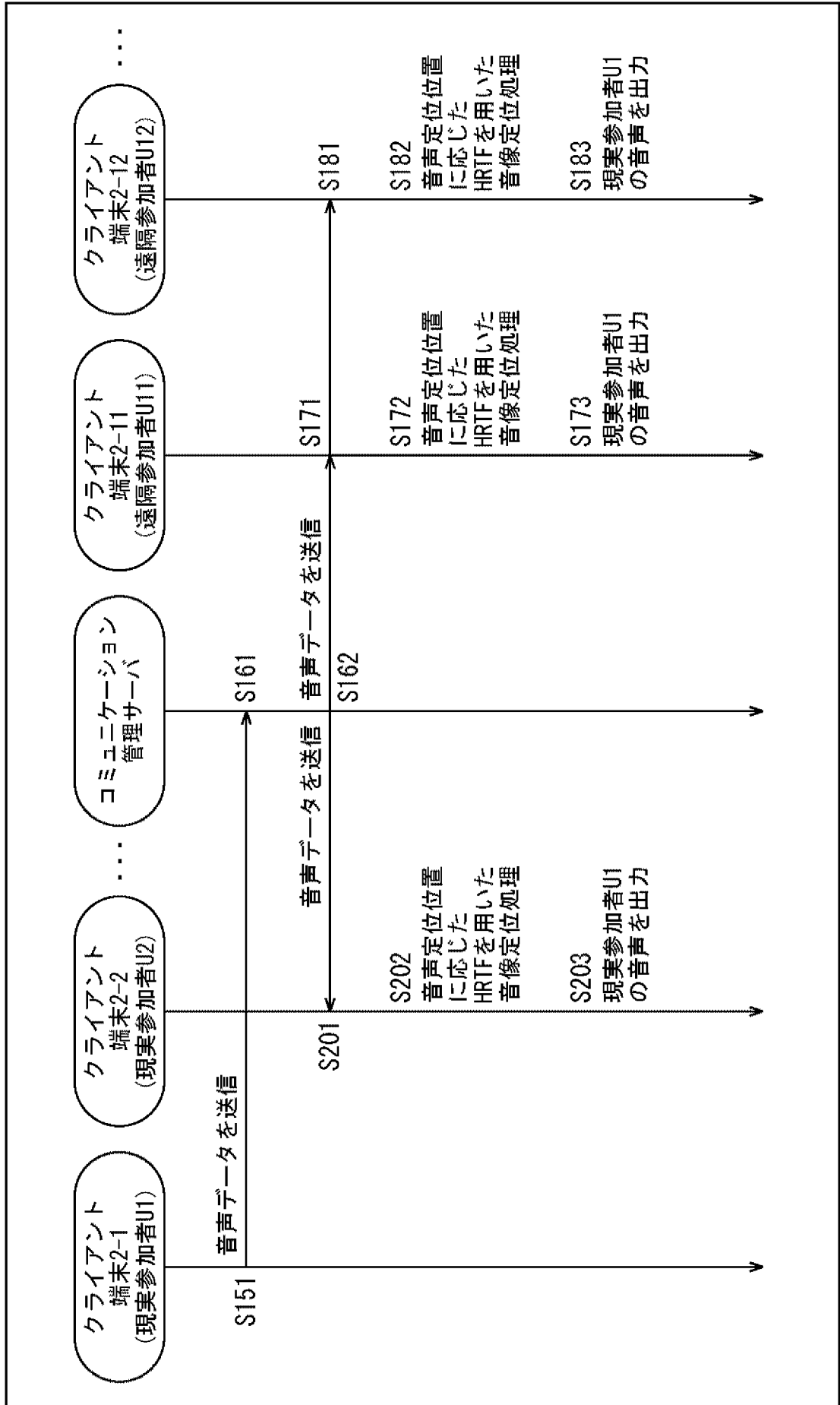
[図21]
FIG. 21

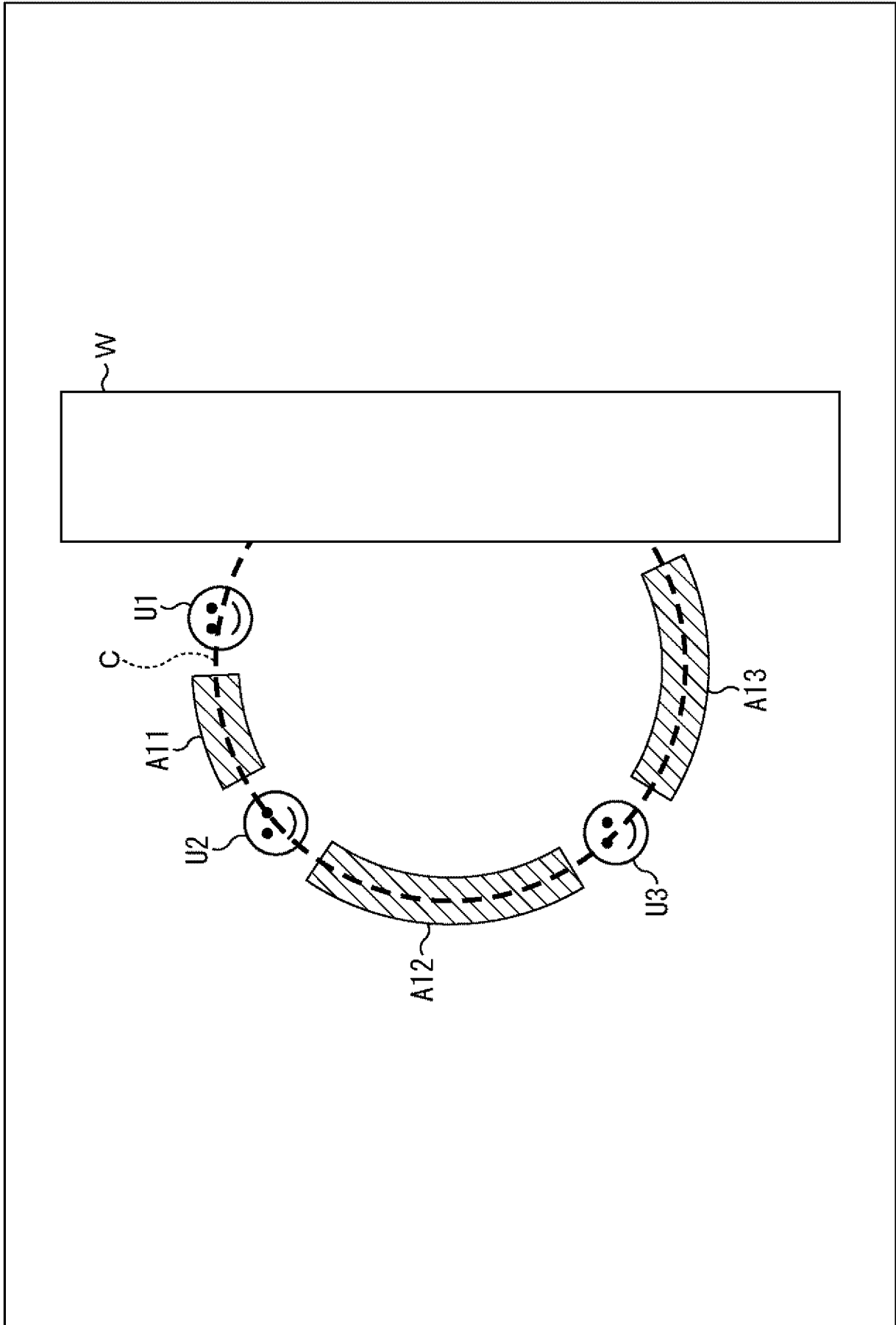



[図22]
FIG. 22

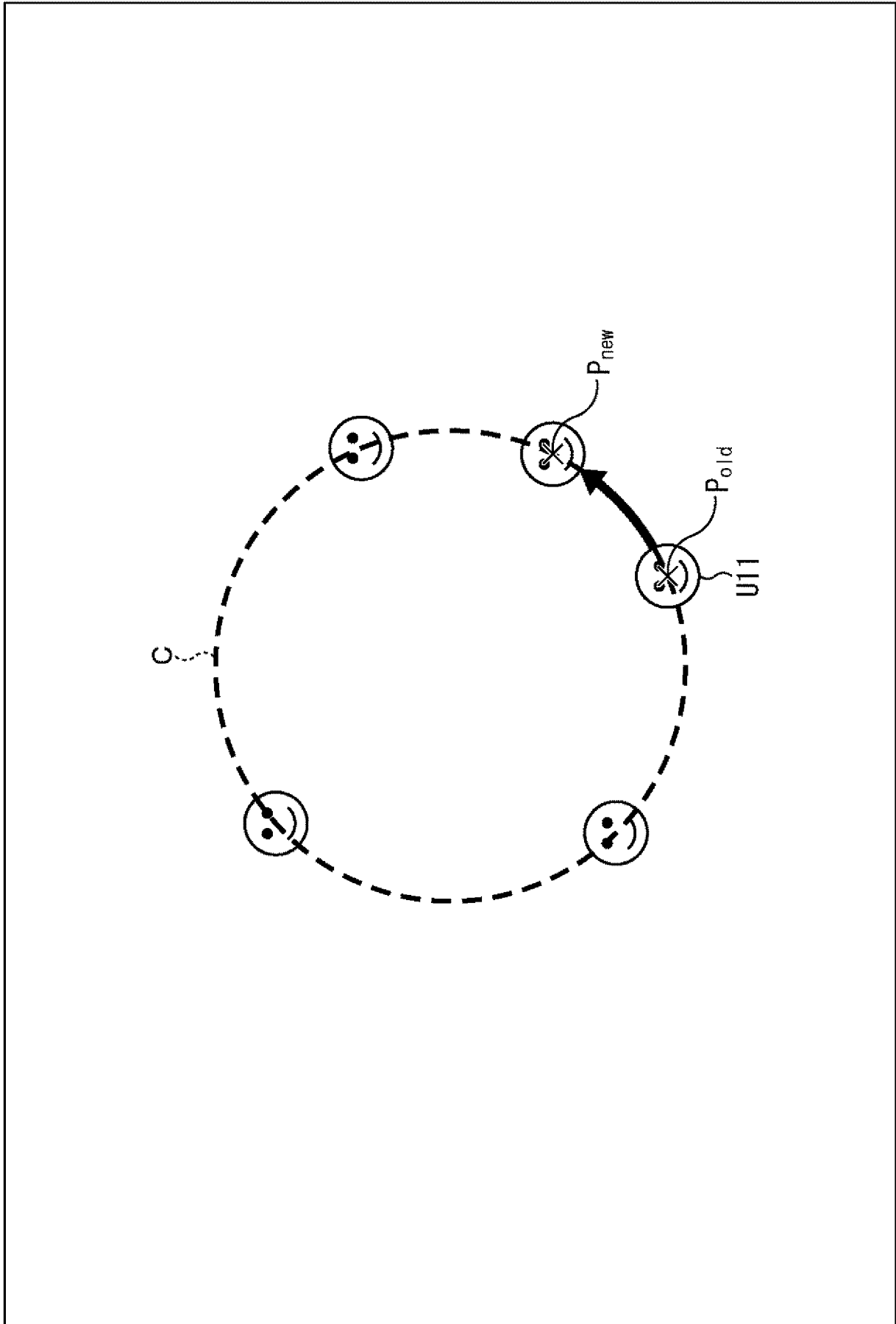



[図23]
FIG. 23

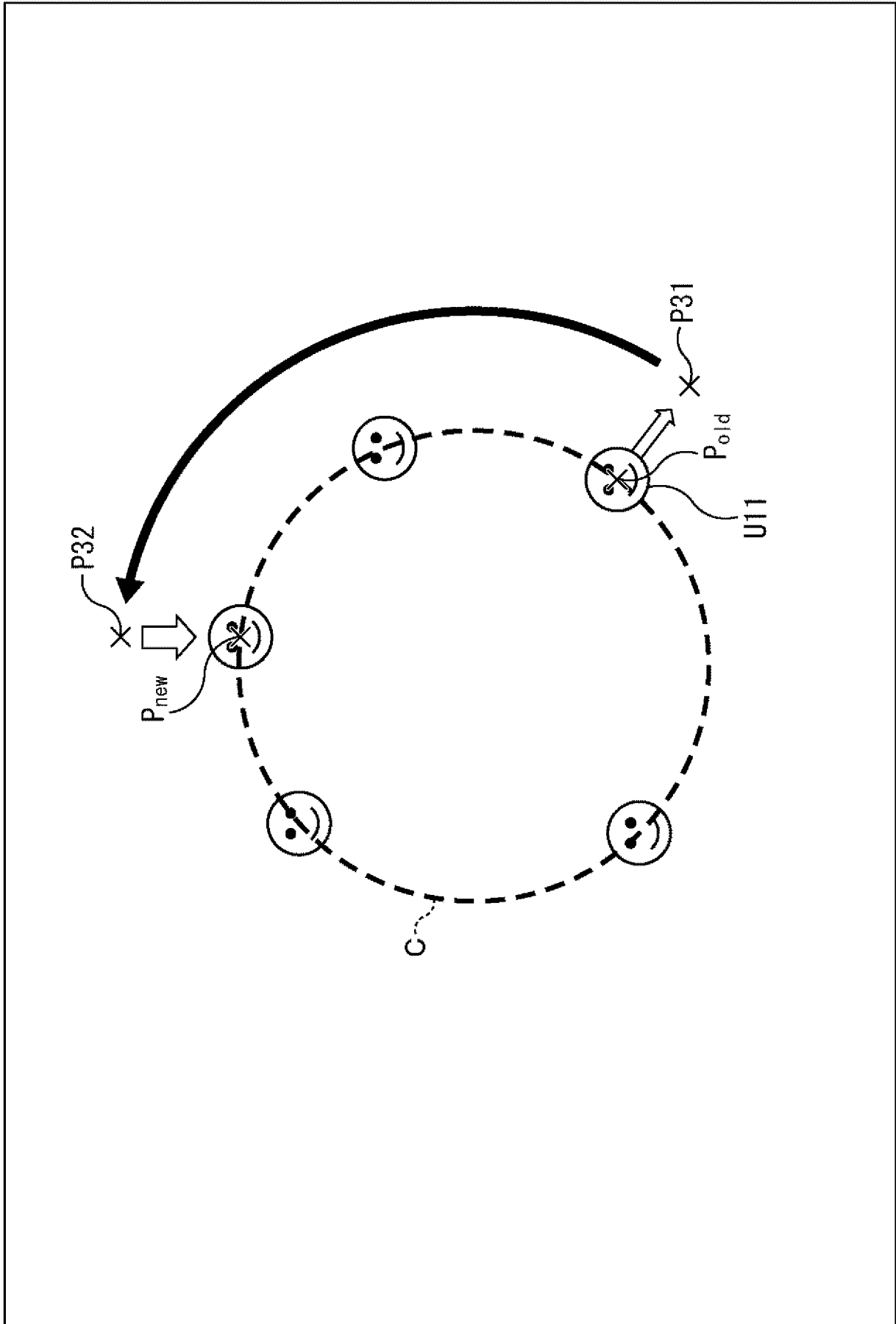


[図24]
FIG. 24

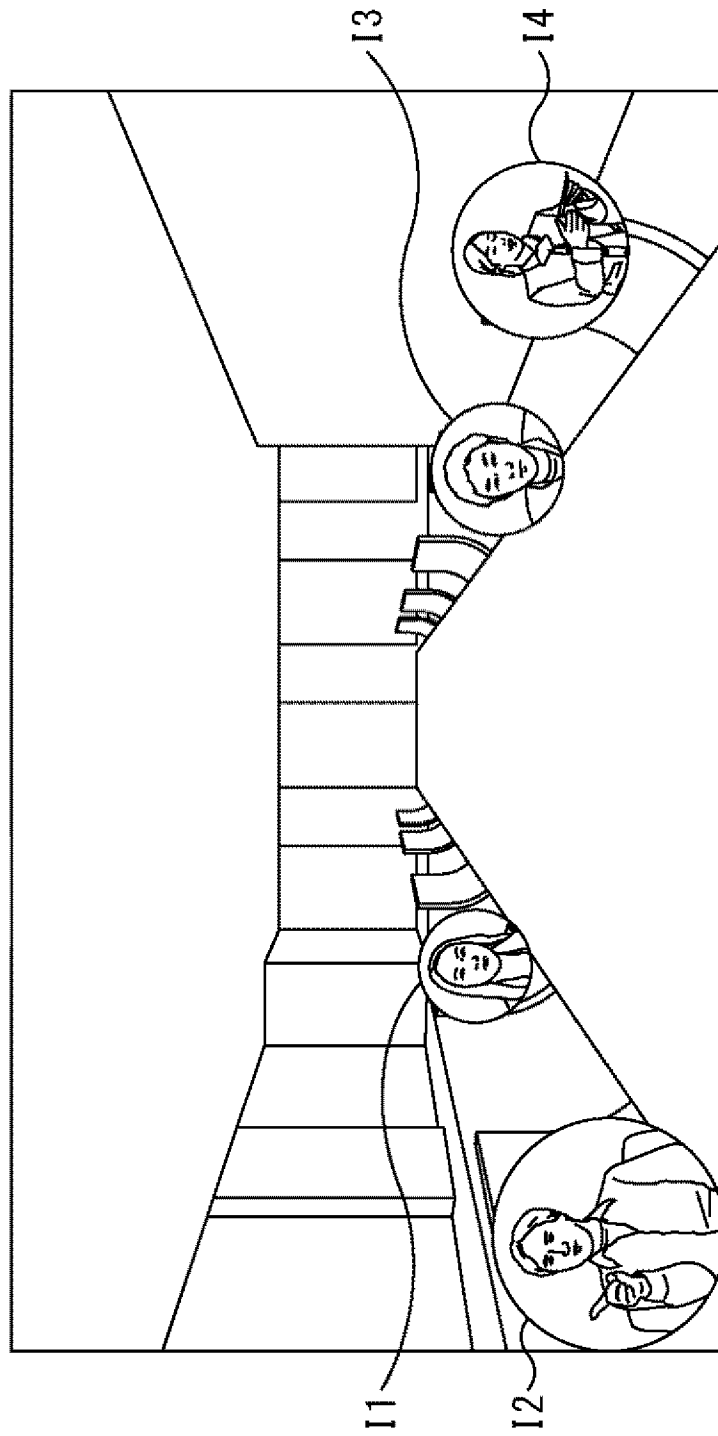
[25]
FIG. 25



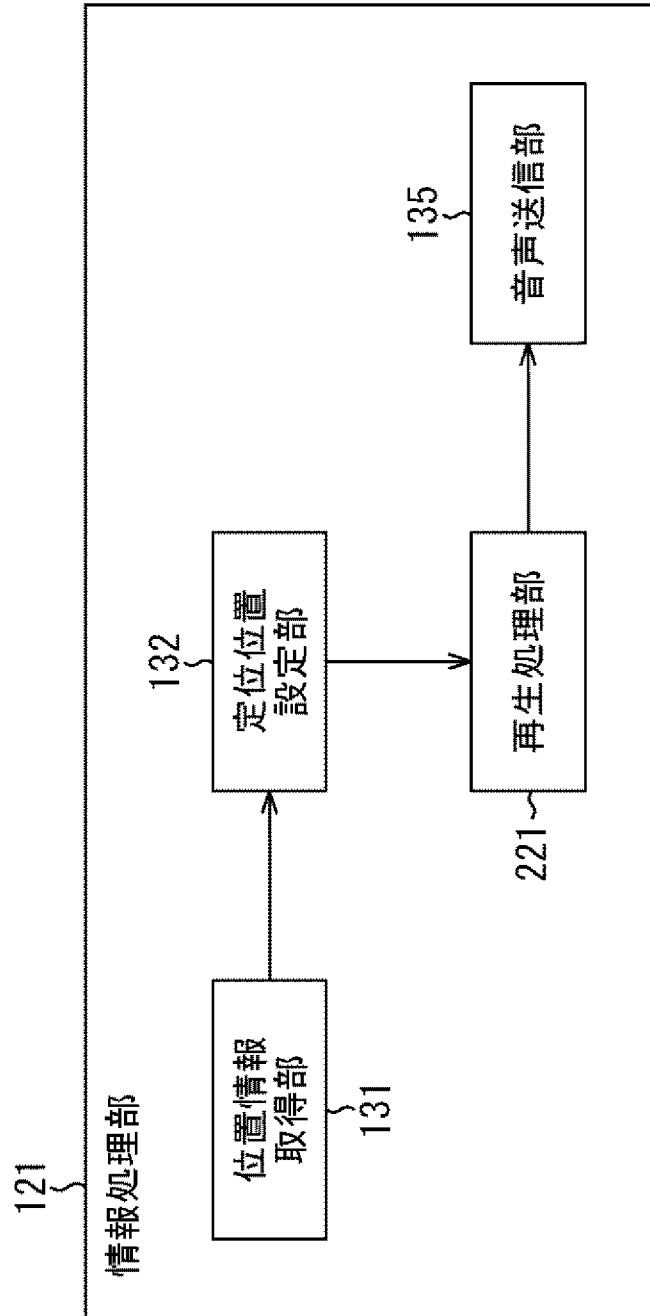
[26]
FIG. 26



[図27]
FIG. 27



[図28]
FIG. 28



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/042528

A. CLASSIFICATION OF SUBJECT MATTER		
<i>H04M 3/56</i> (2006.01)i; <i>H04S 7/00</i> (2006.01)i FI: 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		
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	JP 2020-88516 A (TAKENAKA KOMUTEN CO., LTD.) 04 June 2020 (2020-06-04) paragraphs [0031]-[0066], fig. 1	1-7, 9-15
A	paragraphs [0031]-[0066], fig. 1	8
A	WO 2020/022154 A1 (SHARP K. K.) 30 January 2020 (2020-01-30) paragraphs [0013]-[0106]	8
A	JP 2008-131193 A (YAMAHA CORP.) 05 June 2008 (2008-06-05) paragraphs [0010]-[0040]	8
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 22 December 2021		Date of mailing of the international search report 11 January 2022
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/JP2021/042528

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2020-88516 A	04 June 2020	(Family: none)	
WO 2020/022154 A1	30 January 2020	US 2021/0185174 A1 paragraphs [0024]-[0117]	
JP 2008-131193 A	05 June 2008	(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 FI: 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 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 1971-2021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		
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2020-88516 A（株式会社竹中工務店）04.06.2020（2020-06-04） 段落[0031]-[0066]、図1	1-7,9-15
A	段落[0031]-[0066]、図1	8
A	WO 2020/022154 A1（シャープ株式会社）30.01.2020（2020-01-30） 段落[0013]-[0106]	8
A	JP 2008-131193 A（ヤマハ株式会社）05.06.2008（2008-06-05） 段落[0010]-[0040]	8
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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	“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 “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	
国際調査を完了した日	国際調査報告の発送日	
22.12.2021	11.01.2022	
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） 大石 剛 5Z 4882 電話番号 03-3581-1101 内線 3591	

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2021/042528

引用文献	公表日	パテントファミリー文献	公表日
JP 2020-88516 A	04.06.2020	(ファミリーなし)	
WO 2020/022154 A1	30.01.2020	US 2021/0185174 A1 Paragraphs[0024]-[0117]	
JP 2008-131193 A	05.06.2008	(ファミリーなし)	