(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9月12日(12.09.2024)

WIPO | PCT

(10) 国際公開番号
WO 2024/185334 A1

(51) 国際特許分類:
H04R 3/00 (2006.01)　*H04N 7/14* (2006.01)
G10K 15/04 (2006.01)　*H04N 7/15* (2006.01)
H04M 1/72 (2021.01)　*H04S 7/00* (2006.01)
H04M 3/56 (2006.01)

(21) 国際出願番号：　PCT/JP2024/002446

(22) 国際出願日：　2024年1月26日(26.01.2024)

(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 2023-035310　2023年3月8日(08.03.2023)　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).

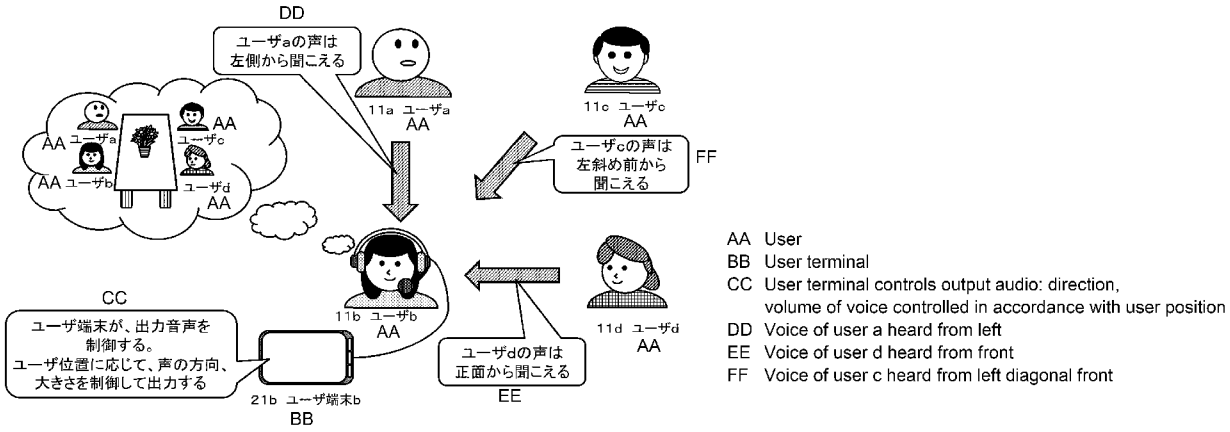
(72) 発明者: 深谷　裕美 (**FUKAYA, Hiromi**); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 宮崎　秀平 (**MIYAZAKI, Shuhei**); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP).

(74) 代理人:宮田 正昭，外(**MIYATA, Masaaki** et al.); 〒1040032 東京都中央区八丁堀三丁目２５番９号 Ｄａｉｗａ八丁堀駅前ビル西館８階 弁理士法人 大同特許事務所 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, 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, JP, 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.

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

(54) 発明の名称： 情報処理装置、および情報処理方法、並びにプログラム



(57) **Abstract:** The present invention performs audio direction control such that user speech by a conversation partner over a network is heard as speech from a user position for the conversation partner relative to a predefined self position. The present invention has a communication unit that receives user speech by a conversation partner over a network and an output audio control unit that performs output control of the user speech. The output audio control unit performs audio direction control and volume control such that the user speech is heard as speech from a user position for the conversation partner relative to a predefined self position. The user position for the conversation partner relative to the self position is set at a predetermined fixed position or is determined in accordance with intimacy with the conversation partner and set closer to the self position the greater the intimacy.

WO 2024/185334 A1

［続葉有］

(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))

(57) 要約: ネットワークを介した会話相手のユーザ発話を予め規定した自己位置に対する会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する。ネットワークを介した会話相手のユーザ発話を受信する通信部と、ユーザ発話の出力制御を実行する出力音声制御部を有する。出力音声制御部がユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御や音量制御を実行する。自己位置に対する会話相手のユーザ位置は、予め決定された固定位置、または会話相手に対する親密度に応じて決定し、親密度が高いほど自己位置に近い位置に設定する。

明　細　書

発明の名称 ：

情報処理装置、および情報処理方法、並びにプログラム

技術分野

[0001] 本開示は、情報処理装置、および情報処理方法、並びにプログラムに関する。さらに詳細には、複数のユーザ端末をネットワーク接続して会話やミーティングを行う際、各ユーザ端末に例えばカフェ等、ある実環境の背景音や背景画像を出力することで各ユーザがカフェ等の実環境で会話をしているような感覚を得ることを可能とした情報処理装置、および情報処理方法、並びにプログラムに関する。

背景技術

[0002] 昨今、通信端末を利用して音声、画像データを送受信して行うリモートミーティング等、ネットワークを介した会話やミーティングが盛んに行われている。

[0003] ネットワークを介した会話では、会話参加ユーザ各々が所有するＰＣやスマホ（スマートフォン）などのユーザ端末をインターネット等の通信ネットワークに接続し、通信ネットワークを介して各端末間で画像や音声を送受信する。

[0004] しかし、会話の参加ユーザが実際に存在する場所は、それぞれ各ユーザの自宅当、異なる場所である場合が多く、それぞれの居場所の環境は異なる。

一方、現実世界での実際の会話やミーティングは例えば会議室やカフェ等、１つの場所に集まって行われる。

[0005] このようにネットワークを介した会話では、各ユーザの環境が異なるため、１つの場所で会話を行っている感覚がなく一体感を得にくいという問題がある。

[0006] なお、ユーザ端末をネットワーク接続して会話を行うシステムを開示した従来技術として、例えば特許文献１（国際公開ＷＯ２０１９／１５５７３５

号公報）がある。

この特許文献１は、ユーザ端末に会話相手の仮想画像を表示し、表示した仮想画像の向きや表情を実際の会話相手と同様に変化させる構成を開示している。

[0007] しかし、この開示構成は会話ユーザの表示制御のみであり、背景音や背景画像を制御するものではなく、会話を行っているユーザが同一空間にいるような感覚を得られるものではない。

先行技術文献

特許文献

[0008] 特許文献1：国際公開ＷＯ２０１９／１５５７３５号公報

発明の概要

発明が解決しようとする課題

[0009] 本開示は、例えば、上記問題点に鑑みてなされたものであり、通信ネットワークを介して会話を行うユーザのユーザ端末に例えばカフェ等の実環境の背景音や背景画像を出力することで、ユーザがカフェ等の実環境で会話をしているような感覚を得ることを可能とした情報処理装置、および情報処理方法、並びにプログラムを提供することを目的とする。

課題を解決するための手段

[0010] 本開示の第１の側面は、

ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記出力音声制御部は、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する情報処理装置にある。

[0011] さらに、本開示の第２の側面は、

情報処理装置において実行する情報処理方法であり、

前記情報処理装置は、

ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記出力音声制御部が、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する情報処理方法にある。

[0012] さらに、本開示の第３の側面は、

情報処理装置が情報処理を実行させるプログラムであり、

前記情報処理装置は、

ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記プログラムは、前記出力音声制御部に、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行させるプログラムにある。

[0013] なお、本開示のプログラムは、例えば、様々なプログラム・コードを実行可能な情報処理装置やコンピュータ・システムに対して、コンピュータ可読な形式で提供する記憶媒体、通信媒体によって提供可能なプログラムである。このようなプログラムをコンピュータ可読な形式で提供することにより、情報処理装置やコンピュータ・システム上でプログラムに応じた処理が実現される。

[0014] 本開示のさらに他の目的、特徴や利点は、後述する本開示の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

[0015] 本開示の一実施例の構成によれば、ネットワークを介した会話相手のユーザ発話を予め規定した自己位置に対する会話相手のユーザ位置からの発話と

して聞こえるように音声方向制御を実行する構成が実現される。

具体的には、例えば、ネットワークを介した会話相手のユーザ発話を受信する通信部と、ユーザ発話の出力制御を実行する出力音声制御部を有する。出力音声制御部がユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御や音量制御を実行する。自己位置に対する会話相手のユーザ位置は、予め決定された固定位置、または会話相手に対する親密度に応じて決定し、親密度が高いほど前記自己位置に近い位置に設定する。

本構成により、ネットワークを介した会話相手のユーザ発話を予め規定した自己位置に対する会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する構成が実現される。

なお、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、また付加的な効果があってもよい。

図面の簡単な説明

[0016] [図1]本開示の情報処理システムの構成と実行する処理の概要について説明する図である。

[図2]背景データ提供サーバがユーザ端末に提供する背景音声データや背景画像データの具体例について説明する図である。

[図3]背景データ（背景音声データ、背景画像データ）をユーザ端末に格納する場合の処理例を説明する図である。

[図4]実施例１において利用されるユーザ端末の一構成例について説明する図である。

[図5]ユーザ端末の出力音声制御部が実行する音声制御処理の具体例について説明する図である。

[図6]背景の設定や各ユーザの位置を決定する処理シーケンスの例について説明する図である。

[図7]他のユーザの発話を各ユーザ設定位置から発せられたように聞くことを可能とした設定について説明する図である。

[図8]出力音声制御部が実行する出力音声制御処理に適用する音声出力制御パラメータの算出処理例について説明する図である。

[図9]出力音声制御部が実行する出力音声制御処理に適用する音声出力制御パラメータの算出処理例について説明する図である。

[図10]ユーザａのユーザ端末ａの出力音声制御部が実行する他ユーザｂ～ｄの発話音声の出力制御処理例について説明する図である。

[図11]ユーザｂのユーザ端末ｂの出力音声制御部が実行する処理の一例について説明する図である。

[図12]ユーザｂのユーザ端末ｂの出力音声制御部が実行する処理の一例について説明する図である。

[図13]４人のユーザａ～ｄが、様々な背景音が存在するカフェの背景データを利用して会話する場合の例について説明する図である。

[図14]背景がカフェの設定において、ユーザ端末の出力音声制御部が実行する出力音声制御処理の具体例について説明する図である。

[図15]４人のユーザａ～ｄが周囲にたくさんの人がいるカフェの背景データを利用して会話する場合の例について説明する図である。

[図16]背景が周囲にたくさんの人がいるカフェの設定において、ユーザ端末の出力音声制御部が実行する出力音声制御処理の具体例について説明する図である。

[図17]実施例２の具体的処理例について説明する図である。

[図18]ユーザ間の親密度に応じたユーザ（ユーザ対応のアバターや実画像）の表示制御処理例について説明する図である。

[図19]親密度算出部が算出した親密度情報に従って出力画像制御部が決定した各ユーザの表示位置に従って、画像出力部が表示部に表示する画像の具体例について説明する図である。

[図20]ユーザ端末ａの出力音声制御部が実行する親密度に応じたユーザ発話出力音量の制御処理例を示すグラフについて説明する図である。

[図21]ユーザａとユーザｂ～ｄ間の具体的な親密度に対応する出力音声制御

処理例について説明する図である。

[図22]ユーザａとユーザｂ～ｄ間の具体的な親密度に対応する出力音声制御処理例について説明する図である。

[図23]実施例２において利用されるユーザ端末の一構成例について説明する図である。

[図24]親密度算出部の詳細な構成例と具体的な親密度算出処理例について説明する図である。

[図25]親密度計算部が算出する「ユーザ好みベース親密度」の算出処理例を示す図である。

[図26]親密度計算部が算出する「会話密度ベース親密度」の算出処理例を示す図である。

[図27]複数ユーザａ～ｄ間の会話実行中に出力画像制御部が実行する表示態様の変更処理例について説明する図である。

[図28]出力画像制御部が実行する親密度の変更に伴うユーザ画像（アバター画像、または実画像）の変更処理例について説明する図である。

[図29]ネットワークを介して会話を行うユーザのユーザ端末各々が異なる背景データを各ユーザ端末に出力している例を示す図である。

[図30]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処理例について説明する図である。

[図31]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処理例について説明する図である。

[図32]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処理例について説明する図である。

[図33]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処理例について説明する図である。

[図34]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処理例について説明する図である。

[図35]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処

理例について説明する図である。

[図36]ユーザ間の会話に応じて自端末に出力する背景データを切り替える処理例について説明する図である。

[図37]ユーザａがユーザ端末ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する図である。

[図38]ユーザａがユーザ端末ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する図である。

[図39]ユーザａがユーザ端末ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する図である。

[図40]ユーザａがユーザ端末ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する図である。

[図41]ユーザａがユーザ端末ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する図である。

[図42]ユーザａがユーザ端末ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する図である。

[図43]ユーザ端末やサーバのハードウェア構成例について説明する図である。

発明を実施するための形態

[0017] 以下、図面を参照しながら本開示の情報処理装置、および情報処理方法、並びにプログラムの詳細について説明する。なお、説明は以下の項目に従って行なう。

１．本開示の情報処理システムの構成と実行する処理の概要について

２．（実施例１）ユーザの位置に基づく音声制御を実行する実施例について

２－１．実施例１のユーザ端末の構成例について

２－２．ユーザ位置に基づく音声制御処理の具体例について

３．（実施例２）ユーザ間の親密度に応じた音声出力制御、表示制御を実行する実施例について

３－１．ユーザ間の親密度に応じた音声出力制御、表示制御処理の具体的処理例について

３－２．実施例２のユーザ端末の構成例について

４．（実施例３）各ユーザ端末で異なる背景データを利用する実施例について

４－１．（処理例１）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例について

４－２．（処理例２）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみを会話相手のユーザ端末に設定された背景音声に切り替える設定とした場合の処理例について

４－３．（処理例３）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみならず、背景画像データも会話相手のユーザ端末に設定された背景データに切り替える設定とした場合の処理例について

４－４．（処理例４）他ユーザから話しかけられて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例について

４－５．（処理例５）複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について

４－６．（処理例６）複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データと背景画像データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について

４－７．（処理例７）複数ユーザの会話中に別の新たなユーザから話しかけられた場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について

５．ユーザ端末に背景データを出力してユーザ間で会話を行う具体的な処理シーケンスの例について

６．ユーザ端末、およびサーバのハードウェア構成例について

７．本開示の構成のまとめ

[0018] ［１．本開示の情報処理システムの構成と実行する処理の概要について］

まず、図１以下を参照して本開示の情報処理システムの構成と実行する処理の概要について説明する。

[0019] 図１は、例えばリモート会議や、リモートミーティング、あるいはオンラインゲームなどを実行するシステムの一例であり、通信ネットワークを介してユーザ間で会話を行うことが可能な情報処理システムの構成例を示している。

[0020] 図１には、通信ネットワークを介した会話の参加ユーザであるユーザａ，１１ａ～ユーザｄ，１１ｄと各ユーザが利用するユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄと、通信実行環境を提供するサーバである通信管理サーバ５０と、様々な背景音声データや、背景画像データを提供する背景データ提供サーバ７０を示している。

[0021] 通信管理サーバ５０は、例えばリモート会議実行環境を提供するリモート会議管理サーバや、ゲーム実行環境を提供するゲームサーバ等である。

背景データ提供サーバ７０は、ユーザａ，１１ａ～ユーザｄ，１１ｄが会議等の会話を行う場所、例えばカフェや会議室などの様々な場所の背景音声データや背景画像データを各ユーザ端末に提供するサーバである。

[0022] なお、通信管理サーバ５０と背景データ提供サーバ７０は、図では別サーバとして示しているが、これらは１つのサーバとしてもよい。

[0023] 通信ネットワークを介した会話を実行する場合、ユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄと通信管理サーバ５０は通信ネットワーク３０を介して接続され、ユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄから出力される音声や画像は、通信管理サーバ５０を介してユーザ端末ａ，２１ａ～ユーザ

端末ｄ，２１ｄで送受信される。

[0024] なお、背景データ提供サーバ７０が提供する様々な背景音声データや、背景画像データは、通信ネットワークを介した会話を実行する前にユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄに格納しておくことが可能である。

[0025] ユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄが、背景音声データや背景画像データを事前取得している場合は、通信ネットワークを介した会話実行時には、背景データ提供サーバ７０とユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄとの接続は不要である。

[0026] さらに、ユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄが、背景データ提供サーバ７０が提供する様々な背景音声データや、背景画像データを取得しながら通信ネットワークを介した会話を実行することも可能である。この場合は、通信ネットワークを介した会話実行時に背景データ提供サーバ７０とユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄとの接続を維持する。

[0027] ユーザ端末２１ａ～２１ｄは、例えばＰＣ、スマホ（スマートフォン）、タブレット端末等の通信可能な情報処理装置によって構成される。

これらユーザ端末２１ａ～２１ｄの各々は、マイクやカメラを有しており、ユーザ端末２１において取得されたユーザ発話等の音声データやユーザの顔画像等の画像データは通信管理サーバ５０を介して他のユーザ端末２１に送信される。

[0028] なお、背景データ提供サーバ７０が提供する背景音声データや背景画像データを利用した処理を行う場合、各ユーザ端末２１ａ～２１ｄには共通の環境、例えばカフェ等、ある１つの環境の背景音や背景画像を出力しながら会話することが可能となる。

[0029] ネットワークを介した通信処理を実行するユーザ端末２１ａ～２１ｄの各々は、各ユーザの会話音の音声出力制御に加えて、例えばある１つの環境の背景音や背景画像を出力するための音声出力制御や画像出力制御を実行する。

[0030] 図２を参照して、背景データ提供サーバ７０がユーザ端末２１に提供する

背景音声データや背景画像データの具体例について説明する。

[0031] 図２には、背景データ提供サーバ７０の記憶部に格納された背景音声データや背景画像データの具体例を示している。

図２に示すように、背景データ提供サーバ７０の記憶部には、例えば以下の様々な背景対応の背景データ（背景音声データ、背景画像データ）が格納されている。

（１）会議室

（２）カフェ

（３）公園

（４）ライブハウス

（５）電車の中

（６）駅

（７）空港

[0032] 背景音声データは、背景を構成する空間内で発生する様々な音声データである。背景音声データは、背景空間に空間を構成する壁、天井、床等がある場合はこれら壁等からの反響音も考慮して生成された音声データである。

[0033] 例えば背景音声データがカフェの背景音声データの場合、音声データにはカフェ空間内にあるコーヒーサイフォンの音や、カフェ内の人の話し声等が含まれ、さらにカフェの壁からの反響音も含む音声データとなる。

[0034] 背景データ提供サーバ７０は、様々な環境における音声データファイルを生成して記憶部に格納する。例えばリアルな空間のインパルス応答の解析などを実行して、実際の様々な空間に対応した音声データを格納した音声データファイルを生成する。

[0035] また、画像データファイルに格納される画像データは３次元画像データ、または２次元画像データであり、様々な視点方向からの画像を表示可能な画像データである。各ユーザはユーザ端末の表示部を操作（指でスライド等）することで、様々な方向からの画像を表示することができる。

[0036] 上述したように、背景データ提供サーバ７０が保持する様々な背景音声デ

ータや、背景画像データは、通信ネットワークを介した会話を実行する前にユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄに格納しておくことが可能である。

[0037] 図３は、背景データ（背景音声データ、背景画像データ）をユーザ端末２１に格納する場合の処理例を説明する図である。

[0038] 図３には、背景データ提供サーバ７０と、ユーザ端末ａ，２１ａを示している。ユーザ端末ａ，２１ａは、背景データ提供サーバ７０にアクセスし、背景データ提供サーバ７０から様々な背景データ（背景音声データ、背景画像データ）を取得（ダウンロード）してユーザ端末ａ，２１ａの記憶部に格納することができる。

[0039] 図３に示す例では、ユーザ端末ａ，２１ａの記憶部である音声データ格納部に、背景データ提供サーバ７０から取得した様々な背景音声データを格納し、画像データ格納部に、背景データ提供サーバ７０から取得した様々な背景画像データを格納した例を示している。

[0040] なお、図３にはユーザ端末ａ，２１ａによる背景データ取得処理例を示しているが、その他のユーザ端末ｂ～ｄも同様の処理を実行して、背景データ提供サーバ７０から様々な背景データ（背景音声データ、背景画像データ）を取得（ダウンロード）して各ユーザ端末の記憶部に格納することができる。

[0041] ［２．（実施例１）ユーザの位置に基づく音声制御を実行する実施例］

次に実施例１として、ユーザの位置に基づく音声制御を実行する実施例について説明する。

[0042] ［２－１．実施例１のユーザ端末の構成例について］

まず、図４を参照して、本実施例１において利用されるユーザ端末の一構成例について説明する。

[0043] 図４には図１を参照して説明した通信ネットワークを介した会話の参加ユーザであるユーザａ，１１ａが利用するユーザ端末ａ，２１ａの構成例を示している。

なお、ユーザａ，１１ａ～ユーザｄ，１１ｄが利用するユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄは、全て図４に示す構成例とほぼ同様の構成を有する。

[0044] なお、先に説明したように、ユーザ端末２１ａ～２１ｄは、例えばＰＣ、スマホ（スマートフォン）、タブレット端末等の通信可能な情報処理装置によって構成される。

[0045] 図４に示すように、ユーザ端末ａ，２１ａは、通信部１０１、ユーザ位置決定部（ＵＩ）１０２、ユーザ位置情報格納部１０３、背景データ取得部１０４、音声データ格納部１０５、画像データ格納部１０６、音声データ受信部１０７、出力音声制御部１０８、音声出力部１０９、出力画像制御部１１０、画像出力部１１１、表示部１１２、音声入力部１１３、カメラ１１４、画像入力部１１５、データ送信部１１６を有する。

[0046] 通信部１０１は、通信ネットワークを介した通信処理を実行する。他のユーザ端末や通信管理サーバ５０、背景データ提供サーバ７０とのデータ送受信処理を実行する。

[0047] ユーザ位置決定部（ＵＩ）１０２は、複数のユーザでネットワークを介した会話を行う場合に、各ユーザの位置を決定する処理を実行する。例えばＵＩ（ユーザインタフェース）を利用してユーザ位置を決定することができる。

例えば図１を参照して説明したように４人のユーザａ，１１ａ～ユーザｄ，１１ｄがネットワークを介した会話を行う場合には、これら４人のユーザの位置を決定する処理を実行する。

[0048] このユーザ位置決定処理は、各ユーザの声が聞こえる方向を制御するために実行する処理である。このユーザ位置の決定処理や、決定したユーザ位置に応じた音声制御処理の具体例については後段で説明する。

[0049] ユーザ位置情報格納部１０３は、ユーザ位置決定部（ＵＩ）１０２において決定したユーザ位置情報を格納するための記憶部である。例えば、図１を参照して説明したように４人のユーザａ，１１ａ～ユーザｄ，１１ｄがネッ

トワークを介した会話を行う場合には、これら４人のユーザの位置情報が格納される。

[0050] ユーザ位置情報格納部１０３に格納されたユーザ位置情報は、出力音声制御部１０８や、出力画像制御部１１０に出力される。

出力音声制御部１０８や、出力画像制御部１１０は、ユーザ位置情報格納部１０３に格納されたユーザ位置情報に従って、各ユーザの発話音声の出力制御（各ユーザ発話の聞こえる方向の制御等）や、ユーザ画像（アバター画像、または実画像）の出力制御等を実行する。

[0051] 背景データ取得部１０４は、通信部１０１を介して背景データ提供サーバ７０から様々な背景データ（背景音声データ、背景画像データ）を取得（ダウンロード）する。

背景データ取得部１０４が背景データ提供サーバ７０から取得した背景データを構成する背景音声データは音声データ格納部１０５に格納される。また背景画像データは画像データ格納部１０６に格納される。

[0052] 音声データ格納部１０５に格納された背景音声データは、出力音声制御部１０８の制御の下、選択取得されて音声出力部１０９を介して出力される。

また、画像データ格納部１０６に格納された背景画像データは、出力画像制御部１１０の制御の下、選択取得され画像出力部１１１を介して表示部１１２に出力される。

[0053] 音声データ受信部１０７は、他の会話参加ユーザの声などの音声データを、通信部１０１を介して受信し、出力音声制御部１０８に出力する。

[0054] 出力音声制御部１０８は、音声データ受信部１０７から入力する他の会話参加ユーザの声などの音声データと、音声データ格納部１０５から選択取得した背景音声データとを併せて音声出力部１１０９を介してユーザの装着したヘッドホン等のスピーカーに出力する。

[0055] なお、出力音声制御部１０８は、ユーザ位置情報格納部１０３に格納された各ユーザの位置に応じて、各ユーザの発話音声の出力制御、具体的には各ユーザ発話の聞こえる方向や声の大きさの制御も実行する。

[0056] 出力画像制御部１１０は、画像データ格納部１０６から選択取得した背景画像データを、画像出力部１１１を介して表示部１１２に出力する制御を実行する。

なお、出力画像制御部１１０は、ユーザ位置情報格納部１０３に格納された各ユーザの位置に応じて、各ユーザを示すアバター画像等、各ユーザを示す仮想画像（キャラクタ画像）、あるいは各ユーザの実画像を背景画像に重畳して表示する制御も行う。

[0057] なお、各ユーザを示すアバター画像等、各ユーザを示す仮想画像（キャラクタ画像）は、予め画像データ格納部１０６に格納されたものを利用する。あるいは通信部１０１を介して各ユーザ端末から受信した画像を利用してもよい。

各ユーザの実画像を表示する場合は、通信部１０１を介して各ユーザ端末から受信した画像を利用する。

[0058] 音声入力部１１３は、ユーザの発話音声等の音声データを、マイクを介して入力し、データ送信部１１６、通信部１０１を介してネットワーク接続された各ユーザ端末等の各装置に送信する。

[0059] 画像入力部１１５は、カメラ１１４によって撮影されたユーザの顔画像等の画像データを入力し、データ送信部１１６、通信部１０１を介してネットワーク接続された各ユーザ端末等の各装置に送信する。

[0060] ［２－２．ユーザ位置に基づく音声制御処理の具体例について］

次に本実施例１のユーザ端末２１が実行する処理、すなわちユーザ位置に基づく音声制御処理の具体例について説明する。

[0061] 先に図４を参照して説明したように、ユーザ端末の出力音声制御部１０８は、ユーザ位置情報格納部１０３に格納された各ユーザの位置に応じて、各ユーザの発話音声の出力制御、具体的には各ユーザ発話の聞こえる方向の制御や声の大きさの制御を実行する。

[0062] このユーザ端末の出力音声制御部１０８が実行する音声制御処理の具体例について図５以下を参照して説明する。

[0063] 図１に示す４人のユーザａ，１１ａ～ユーザｄ，１１ｄがオンライン会議を行う場合の処理例について説明する。

４人のユーザａ，１１ａ～ユーザｄ，１１ｄは、図５に示すように１つの会議室で会議を行う設定としたオンライン会議を行う。

この場合、背景画像データとして会議室の画像データが利用され、背景音声データとして会議室の音声データが利用される。なお、会議室の音声データは、例えばほぼ無音の音声データ、あるいは空調音の実の音声データ等である。

[0064] まず、オンライン会議を始める前に、各ユーザの位置を決定する。具体的には４人のユーザａ，１１ａ～ユーザｄ，１１ｄの位置を例えば図５に示すように決定する処理を行う。

[0065] 例えばユーザａ，１１ａ～ユーザｄ，１１ｄが、各々のユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄを、通信ネットワークを介して接続した後、背景の設定や各ユーザの位置を相談して決定する。

図６を参照して、背景の設定や各ユーザの位置を決定する処理シーケンスの例について説明する。

[0066] 背景の設定や各ユーザの位置を決定する処理シーケンスは、例えば図６のステップ（Ｓ０１）～（Ｓ０４）の手順で実行することが可能である。

[0067] まず、ステップＳ０１において、ユーザａ，１１ａが、その他のユーザｂ～ｄに対してオンライン会議の設定の提案として、

「会議室で会議する設定にしよう」

とする提案を行う。

[0068] ステップＳ０２で全てのユーザａ～ｄがこの提案に賛成すると、各ユーザは、それぞれのユーザ端末を操作して背景データ（背景画像データ、背景音声データ）として会議室の背景データを選択するための操作を実行する。例えば、各ユーザ端末には、背景データ選択用ＵＩが表示され、ユーザはこのＵＩを利用して会議室の背景データを選択する。

[0069] この背景データ選択処理によって、各ユーザ端末の出力音声制御部１０８

は、音声データ格納部１０５から会議室の音声データを入力して、音声出力部１０９を介して会議室の音声データを出力するように設定される。

同様に、各ユーザ端末の出力画像制御部１１０は、画像データ格納部１０６から会議室の画像データを入力して、画像出力部１１１を介して会議室の背景画像を表示部１１２に出力するように設定される。

[0070] 次に、図６に示すようにステップＳ０３において、ユーザａ，１１ａが、その他のユーザｂ～ｄに対してオンライン会議におけるユーザ位置の設定について、以下のように提案する。

「ユーザ位置は、自分（ａ）の前がｃさん、自分（ａ）の隣がｂさん、ｂさんの前がｄさんでいいかな？」

このようなユーザ位置の提案を行う。

[0071] この処理は、図４を参照して説明したユーザ端末のユーザ位置決定部（ＵＩ）１０２を利用して実行することができる。

例えば図６の（Ｓ０３）の右側に示すようなユーザ配置データが各ユーザ端末に表示され、各ユーザは、ユーザａ，１１ａの提案したユーザ位置を確認することができる。

[0072] 最後に、ステップＳ０４で、ユーザａ，１１ａが提案した会議室内の各ユーザａ～ｄの位置について全てのユーザａ～ｄが賛成すると、その後、オンライン会議が開始される。

なお、ユーザａ，１１ａがユーザ端末ａ，２１ａのユーザ位置決定部（ＵＩ）１０２を利用して設定したユーザａ～ｄの位置情報は、各ユーザ端末ｂ～ｄにも送信され、このユーザ位置情報が各ユーザ端末内のユーザ位置情報格納部１０３に格納される。

[0073] その後、各ユーザ端末の出力音声制御部１０８は、ユーザ位置情報格納部１０３に格納されたユーザ位置に従って、各ユーザの発話音声の音声方向制御や音量制御を実行する。

[0074] また、各ユーザ端末の出力画像制御部１１０は、ユーザ位置情報格納部１０３に格納されたユーザ位置に従って、各ユーザのアバター画像、または実

画像を会議室の背景画像データ上に重畳して表示する。

すなわち、各ユーザａ～ｄのユーザ端末の表示部には、図５に示すような画像データが表示されることになる。

[0075] これらの設定が完了した後、オンライン会議が開始される。

オンライン会議開始後、各ユーザａ～ｄは、他のユーザの発話を、図６を参照して説明したシーケンスに従って設定された各ユーザの位置から発せられた音声として聞くことができる。

[0076] すなわち、各ユーザａ～ｄのユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄの出力音声制御部１０８は、それぞれユーザ位置情報格納部１０３に格納されたユーザ位置に従って、各ユーザの発話音声の音声方向制御や音量制御を実行する。この出力音声制御処理によって、各ユーザは他のユーザの発話を各ユーザ設定位置から発せられたように聞くことができる。

[0077] 図７を参照して具体例について説明する。

図７は、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの制御によってユーザａ，１１ａが他のユーザｂ～ｄの発話がどの方向からなされたかを認識する具体例を示している。

[0078] 図７に示すように、ユーザａ，１１ａによる他のユーザｂ～ｄの発話の認識方向は以下の通りとなる。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右側から聞こえる。

ユーザｃ，１１ｃの声はユーザａ，１１ａの正面から聞こえる。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右斜め前から聞こえる。

[0079] これは、ユーザ端末ａ，２１ａの出力音声制御部１０８の出力音声制御処理による効果である。

ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザ位置情報格納部１０３に格納されたユーザ位置に従って、各ユーザの発話音声の音声方向制御や音量制御を実行する。この出力音声制御処理によって、ユーザａ，１１ａは他のユーザｂ～ｄの発話を、図５に示すユーザ位置から発せられたように聞くことができる。

[0080] ユーザ端末ａ，２１ａの出力音声制御部１０８が実行する出力音声制御処理は、ユーザ位置情報格納部１０３に格納されたユーザ位置に従ったユーザ発話音声の音声方向制御処理、および音量制御処理である。

[0081] この処理を実行するために、出力音声制御部１０８は、音声を聞くユーザの位置を中心位置として、その周囲の様々な音源位置対応の音声出力制御パラメータを予め算出し保持し、この制御パラメータを利用して、各位置からの音声の出力制御を行う。

[0082] 図８を参照して、出力音声制御部１０８が実行する出力音声制御処理に適用する音声出力制御パラメータの算出処理例について説明する。

図８は音声を聞くユーザの位置＝聴取位置をｘｙｚ３次元空間の中心位置（ｘ０，ｙ０，ｚ０）とし、その周囲に様々な仮想音源位置（ｘ１，ｙ１，ｚ１）～（ｘｎ，ｙｎ，ｚｎ）を設定した例を示している。

[0083] 仮想音源位置（ｘ１，ｙ１，ｚ１）～（ｘｎ，ｙｎ，ｚｎ）各々からの音声出力チャネルをｃｈ１～ｃｈｎとする。

[0084] 音声出力チャネルｃｈ１～ｃｈｎ各々について、チャネル対応の音声制御パラメータを算出する。すなわち、中心位置（ｘ０，ｙ０，ｚ０）に位置するユーザのユーザ端末の音声出力部（スピーカ）、またはユーザが装着したヘッドホンやイヤホンのＬＲ各音声出力部（スピーカ）に出力する音声の出力制御パラメータを算出する。

[0085] チャネル対応の音声制御パラメータは、チャネルｃｈ１～ｃｈｎ各々の音声を中心位置（ｘ０，ｙ０，ｚ０）に位置するユーザが、各チャネルの音源位置（ｘ１，ｙ１，ｚ１）～（ｘｎ，ｙｎ，ｚｎ）からの出力音として認識させるためのパラメータである。

[0086] なお、チャネル対応の音声制御パラメータは、音声の聞こえる方向の制御パラメータのみならず、音の大きさを調整する音量制御パラメータも含むパラメータである。

すなわち、中心位置（ｘ０，ｙ０，ｚ０）と各チャネル位置との距離に応じた音量制御パラメータも含んでおり、中心位置（ｘ０，ｙ０，ｚ０）と近

いチャネル位置では音量が大きく、遠い位置では音量を小さくする制御がなされる。

[0087] このように、出力音声制御部１０８は、音声を聞くユーザの位置を中心位置として、その周囲の様々な音源位置対応の音声出力制御パラメータを予め算出して保持する。

出力音声制御部１０８は、予め算出したこの制御パラメータを利用して、発話ユーザの設定位置に対応するパラメータを利用して各ユーザの発話に対する音声出力制御を実行する。

[0088] なお、発話ユーザの設定位置が図８に示す複数のチャネル位置の間にある場合は、発話ユーザの設定位置周囲の複数のチャネル位置の制御パラメータを合成して、発話ユーザの設定位置に対応するパラメータを算出して、算出パラメータを利用してユーザ発話に対する音声出力制御を実行する。

[0089] この制御処理によって、図７を参照して説明したように、ユーザａ，１１ａによる他のユーザｂ～ｄの発話の認識方向は以下の通りとなる。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右側から聞こえる。

ユーザｃ，１１ｃの声はユーザａ，１１ａの正面から聞こえる。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右斜め前から聞こえる。

[0090] なお、図８に示す制御パラメータの算出例は、聴取者となるユーザの位置を中心とした３次元空間の周囲に複数のチャネルを設定しているが、例えば図９に示すようにｘｙ２次元平面上に仮想音源位置を設定して、各位置対応の制御パラメータを算出して利用する構成としてもよい。

このような簡易的な処理を行っても、少なくとも２次元方向の発話方向認識は可能となる。

[0091] 図９は音声を聞くユーザの位置＝聴取位置をｘｙ２次元空間の中心位置（ｘ０，ｙ０）とし、その周囲に様々な仮想音源位置（ｘ１，ｙ１）～（ｘｎ，ｙｎ）を設定している。

仮想音源位置（ｘ１，ｙ１）～（ｘｎ，ｙｎ）各々からの音声出力チャネルをｃｈ１～ｃｈｎとする。

[0092] 音声出力チャネルｃｈ１～ｃｈｎ各々について、チャネル対応の音声制御パラメータを算出する。すなわち、中心位置（ｘ０，ｙ０）に位置するユーザの装着したヘッドホンやイヤホンのＬＲ各音声出力部（スピーカ）に出力する音声の出力制御パラメータを算出する。

[0093] このような簡易的な処理によって次元平面上のチャネル対応制御パラメータを算出して利用する場合でも、少なくとも２次元方向の発話方向認識は可能となる。

[0094] 図９を参照して説明した処理によって算出される２次元平面上のチャネル対応制御パラメータを利用して、先に説明した図５に示すユーザ配置とした場合の制御処理例について図１０を参照して説明する。

[0095] 図１０に示す例は、ユーザａ，１１ａのユーザ端末ａ，２１ａの出力音声制御部１０８が実行する他ユーザｂ～ｄの発話音声の出力制御処理例を説明する図である。

図１０に示すｘｙ２次元平面の中心位置（ｘ０，ｙ０）にユーザａ，１１ａの位置を設定する。

[0096] その他のユーザｂ，１１ｂ～ユーザｄ，１１ｄは、先に説明した図６に示すシーケンスに従って決定された位置、すなわち先に説明した図５に示す配置となる。

すなわち、

ユーザａ，１１ａの右側にユーザｂ，１１ｂ、

ユーザａ，１１ａの前にユーザｃ，１１ｃ、

ユーザａ，１１ａの右前にユーザｄ，１１ｄ、

この配置となる。

[0097] このユーザ配置を図１０に示すようにチャネル対応のパラメータを算出したｘｙ２次元平面の仮想音源位置に対応付ける。

この結果、図１０に示すように、

ユーザｂ，１１ｂは、ｃｈｑ（ｘｑ，ｙｑ）の位置に設定される。

ユーザｃ，１１ｃは、ｃｈｐ（ｘｐ，ｙｐ）の位置に設定される。

ユーザｄ，１１ｄは、ｃｈｎ（ｘｎ，ｙｎ）の位置に設定される。

[0098] このように、ユーザ位置を各チャネル位置に対応付けることが可能となる。

ユーザａ，１１ａのユーザ端末ａ，２１ａの出力音声制御部１０８は、図１０に示すようなユーザｂ～ｄに対応するチャネル位置に従って、各ユーザｂ～ｄの発話音声の出力制御を、各ユーザ対応のチャネル位置に対応する音声出力制御パラメータを利用して各ユーザ発話の出力制御を実行する。

[0099] すなわち、

ネットワークを介して入力するユーザｂ，１１ｂの発話音声に対しては、ｃｈｑ対応の制御パラメータを利用した制御音声を、ユーザａ，１１ａの音声出力部（ヘッドホン）に出力する。

ユーザｃ，１１ｃの発話音声に対しては、ｃｈｐ対応の制御パラメータを利用した制御音声を、ユーザａ，１１ａの音声出力部（ヘッドホン）に出力する。

ユーザｄ，１１ｄの発話音声に対しては、ｃｈｎ対応の制御パラメータを利用した制御音声を、ユーザａ，１１ａの音声出力部（ヘッドホン）に出力する。

[0100] なお、発話ユーザの設定位置が図１０に示す複数のチャネル位置の間にある場合は、前述したように発話ユーザの設定位置周囲の複数のチャネル位置の制御パラメータを合成して、発話ユーザの設定位置に対応するパラメータを算出して、算出パラメータを利用してユーザ発話に対する音声出力制御を実行する。

[0101] このような音声出力制御を実行することで、先に図７を参照して説明したように、ユーザａ，１１ａによる他のユーザｂ～ｄの発話の認識方向が以下のように設定されることになる。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右側から聞こえる。

ユーザｃ，１１ｃの声はユーザａ，１１ａの正面から聞こえる。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右斜め前から聞こえる。

[0102] このように、ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザ位置情報格納部１０３に格納されたユーザ位置に従って、各ユーザの発話音声の音声方向制御や音量制御を実行する。この出力音声制御処理によって、ユーザａ，１１ａは他のユーザｂ～ｄの発話を、図７に示すように、それぞれのユーザ位置から発せられたように聞くことができる。

[0103] 先に説明した図７は、ユーザ端末ａ，２１ａの出力音声制御部１０８による処理例を示しているが、その他のユーザｂ～ｄも各々のユーザ端末ｂ，２１ｂ～ユーザ端末ｄ，２１ｄにおいて同様の処理を実行する。

[0104] ユーザｂ～ｄ各々のユーザ端末ｂ，２１ｂ～ユーザ端末ｄ，２１ｄは、それぞれのユーザ自身を中心位置に設定し、その他の各ユーザの位置を解析して、他ユーザの発話を各ユーザ位置から発せられたように各ユーザの音声制御を行う。

[0105] 図１１、図１２を参照してユーザｂ，１１ｂのユーザ端末ｂ，２１ｂの出力音声制御部１０８の処理例について説明する。

[0106] 図１１は、ユーザｂ，１１ｂが利用するユーザ端末ｂ，２１ｂの制御によってユーザｂ，１１ｂが他のユーザａ，ｃ，ｄの発話がどの方向からなされたかを認識する具体例を示している。

[0107] 図１１に示すように、ユーザｂ，１１ｂによる他のユーザａ，ｃ，ｄの発話の認識方向は以下の通りとなる。

ユーザａ，１１ａの声はユーザｂ，１１ｂの左側から聞こえる。

ユーザｃ，１１ｃの声はユーザｂ，１１ｂの左斜め前から聞こえる。

ユーザｂ，１１ｂの声はユーザｂ，１１ｂの正面から聞こえる。

[0108] これは、ユーザ端末ｂ，２１ｂの出力音声制御部１０８の出力音声制御処理による効果である。

ユーザ端末ｂ，２１ｂの出力音声制御部１０８は、ユーザ位置情報格納部１０３に格納されたユーザ位置に従って、各ユーザの発話音声の音声方向制御や音量制御を実行する。この出力音声制御処理によって、ユーザｂ，１１ｂは他のユーザａ，ｃ，ｄの発話を、図１１に示すユーザ位置から発せられ

たように聞くことができる。

[0109] 図１２は、ユーザ端末ｂ，２１ｂの出力音声制御部１０８の出力音声制御処理の具体例を示す図である。

ユーザ端末ｂ，２１ｂの出力音声制御部１０８は、図１２に示すｘｙ２次元平面の中心位置（ｘ０，ｙ０）にユーザｂ，１１ｂの位置を設定する。

[0110] その他のユーザａ，ｃ，ｄは、先に説明した図６に示すシーケンスに従って決定された位置、すなわち先に説明した図５に示す配置となる。

すなわち、

ユーザｂ，１１ｂの左側にユーザａ，１１ａ、

ユーザｂ，１１ｂの左前にユーザｃ，１１ｃ、

ユーザｂ，１１ｂの前にユーザｄ，１１ｄ、

この配置となる。

[0111] このユーザ配置を図１２に示すようにチャネル対応のパラメータを算出したｘｙ２次元平面の仮想音源位置に対応付ける。

この結果、図１０に示すように、

ユーザａ，１１ａは、ｃｈｒ（ｘｒ，ｙｒ）の位置に設定される。

ユーザｃ，１１ｃは、ｃｈｓ（ｘｓ，ｙｓ）の位置に設定される。

ユーザｄ，１１ｄは、ｃｈｔ（ｘｔ，ｙｔ）の位置に設定される。

[0112] このように、ユーザ位置を各チャネル位置に対応付けることが可能となる。

ユーザｂ，１１ｂのユーザ端末ｂ，２１ｂの出力音声制御部１０８は、図１２に示すようなユーザａ，ｃ，ｄに対応するチャネル位置に従って、各ユーザａ，ｃ，ｄの発話音声の出力制御を、各ユーザ対応のチャネル位置に対応する音声出力制御パラメータを利用して各ユーザ発話の出力制御を実行する。

[0113] この処理によって、ユーザｂ，１１ｂは、その他のユーザａ，ｃ，ｄの発話を、図１１に示すように、それぞれのユーザ位置から発せられたように聞くことができる。

なお、出力音声制御部１０８は、ユーザーに応じた発話の方向制御のみならずユーザ位置に応じた音量制御も実行する。すなわち遠いユーザの音量は、近いユーザの音量より小さくする制御を行う。

[0114] このように、本実施例では、ユーザ位置情報格納部１０３に格納された各ユーザの位置に応じて、各ユーザの発話が、各ユーザの設定位置から聞こえるように制御され、各ユーザが１つの同一空間に実際に居るのと同様の感覚を味わうことができる。

[0115] なお、図５～図１２を参照して説明した処理例は、４人のユーザａ～ｄが会議室で会議を行う設定の例である。この場合、各ユーザ端末ａ，２１ａ～ｄ，２１ｄの出力音声制御部１０８は、背景音を格納した音声データ格納部１０５から会議室の背景音を取得して、音声出力部１０９から出力する。

[0116] すなわち、各ユーザは、会議室の背景音に併せて各ユーザの発話音を聞くことになる。

ただし、会議室の背景音は、前述したようにほぼ無音の音声データ、あるいは空調音の実の音声データ等である。

[0117] 図１３以下を参照して、４人のユーザａ～ｄが、様々な背景音が存在する環境で会話する場合の例について説明する。

図１３は、４人のユーザａ～ｄが、様々な背景音が存在するカフェの背景データを利用して会話する場合の例を示している。

ユーザ配置は、先に図５を参照して説明した会議室の場合と同様の配置である。

[0118] 各ユーザ端末ａ，２１ａ～ｄ，２１ｄの出力音声制御部１０８は、背景音を格納した音声データ格納部１０５からカフェの背景音を取得して、音声出力部１０９から出力する。

図１３に示すように背景がカフェである場合、背景音には例えばコーヒーサイフォンの音や、コーヒーカップの音、ＢＧＭなどが含まれる。

[0119] このように背景がカフェの設定において、ユーザ端末の出力音声制御部１０８が実行する出力音声制御処理の具体例について、図１４を参照して説明

する。

図１４には、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの出力音声制御部１０８が実行する出力音声制御処理例を示している。

[0120] 図１４に示すように、ユーザ端末ａ，２１ａの出力音声制御部１０８は、音声データ格納部１０５からカフェの背景音、例えばコーヒーサイフォンの音や、コーヒーカップの音、ＢＧＭなどから構成される背景音を取得して、音声出力部１０９を介して出力する。さらに、以下のように他のユーザｂ～ｄの発話の方向制御も併せて実行する。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右側から聞こえるように制御する。

ユーザｃ，１１ｃの声はユーザａ，１１ａの正面から聞こえるように制御する。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右斜め前から聞こえるように制御する。

[0121] このような制御を実行することで、ユーザａ，１１ａは４人のユーザａ～ｄが１つのカフェに集まって会話を行っているような感覚を得ることが可能となる。

[0122] 次に、図１５を参照して、周囲にたくさんの人がいるカフェを背景データとした場合の処理例について説明する。

図１５は、４人のユーザａ～ｄが周囲にたくさんの人がいるカフェの背景データを利用して会話する場合の例を示している。

ユーザ配置は、先に図５を参照して説明した会議室の場合と同様の配置である。

[0123] 各ユーザ端末ａ，２１ａ～ｄ，２１ｄの出力音声制御部１０８は、背景音を格納した音声データ格納部１０５から周囲にたくさんの人がいるカフェの背景音を取得して、音声出力部１０９から出力する。

図１５に示すように背景が周囲にたくさんの人がいるカフェである場合、背景音には、多くの人の話し声、コーヒーカップの音、ＢＧＭなどが含まれ

る。

[0124] このように背景が周囲にたくさんの人がいるカフェの設定において、ユーザ端末の出力音声制御部１０８が実行する出力音声制御処理の具体例について、図１６を参照して説明する。

図１６には、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの出力音声制御部１０８が実行する出力音声制御処理例を示している。

[0125] 図１６に示すように、ユーザ端末ａ，２１ａの出力音声制御部１０８は、音声データ格納部１０５から周囲にたくさんの人がいるカフェの背景音、例えば多くの人の話し声、コーヒーカップの音、ＢＧＭなどから構成される背景音を取得して、音声出力部１０９を介して出力する。さらに、以下のように他のユーザｂ～ｄの発話の方向制御も併せて実行する。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右側から聞こえるように制御する。

ユーザｃ，１１ｃの声はユーザａ，１１ａの正面から聞こえるように制御する。

ユーザｂ，１１ｂの声はユーザａ，１１ａの右斜め前から聞こえるように制御する。

[0126] このような制御を実行することで、ユーザａ，１１ａは４人のユーザａ～ｄが多くの人がいるカフェに集まって会話を行っているような感覚を得ることが可能となる。

[0127] ［３．（実施例２）ユーザ間の親密度に応じた音声出力制御、表示制御を実行する実施例について］

次に、実施例２として、ユーザ間の親密度に応じた音声出力制御、表示制御を実行する実施例について説明する。

[0128] ［３－１．ユーザ間の親密度に応じた音声出力制御、表示制御処理の具体的処理例について］

まず、ユーザ間の親密度に応じた音声出力制御、表示制御処理の具体的処理例について説明する。

[0129] 先に説明した実施例１では、例えば会話に参加する複数のユーザ各々の位置を予め、決定し、決定した各ユーザの位置に応じてユーザ発話の音声出力制御や、ユーザ画像（アバターや実画像）の表示制御を実行する実施例であった。

[0130] 以下に説明する実施例は、例えば会話に参加する複数のユーザ各々の位置を予め、決定せず、各ユーザ間の親密度に応じてユーザ発話の音声出力制御や、ユーザ画像（アバターや実画像）の表示制御を実行する実施例である。

[0131] 例えばユーザａ，１１ａのユーザ端末ａ，２１ａは、ユーザａ，１１ａと仲がよい（親密度が高い）ユーザの発話を大きな音量とする音声出力制御を実行し、さらに表示位置もユーザａの表示位置に近い位置や正面位置にするといった表示制御を実行する。

一方、ユーザａ，１１ａと仲が悪い（親密度が低い）ユーザの発話を小さな音量とする音声出力制御を実行し、さらに表示位置もユーザａの表示位置に遠い位置や後ろの位置にするといった表示制御を実行する。

[0132] 図１７以下を参照して本実施例２の具体的処理例について説明する。

図１７は、先の実施例１において説明した事前処理と同様の処理である。すなわちユーザ間の会話開始前に実行する背景決定処理やユーザ配置の決定処理のシーケンスを説明する図である。

背景の設定や各ユーザの位置を決定する処理シーケンスは、例えば図１７のステップ（Ｓ１１）～（Ｓ１４）の手順で実行することが可能である。

[0133] まず、ステップＳ１１において、ユーザａ，１１ａが、その他のユーザｂ～ｄに対して会話設定の提案として、

「公園で集まる設定にしよう」

とする提案を行う。

[0134] ステップＳ１２で全てのユーザａ～ｄがこの提案に賛成すると、各ユーザは、それぞれのユーザ端末を操作して背景データ（背景画像データ、背景音声データ）として公園の背景データを選択するための操作を実行する。例えば、各ユーザ端末には、背景データ選択用ＵＩが表示され、ユーザはこのＵ

Ｉを利用して公園の背景データを選択する。

[0135] この背景データ選択処理によって、各ユーザ端末の出力音声制御部１０８は、音声データ格納部１０５から公園の音声データを入力して、音声出力部１０９を介して公園の音声データを出力するように設定される。例えば鳥のさえずりや小川のせせらぎの音などを含む公園の音声データを出力するように設定される。

[0136] 同様に、各ユーザ端末の出力画像制御部１１０は、画像データ格納部１０６から公園の画像データを入力して、画像出力部１１１を介して公園の背景画像を表示部１１２に出力するように設定される。

[0137] 次に、図１７に示すようにステップＳ１３において、ユーザａ，１１ａが、その他のユーザｂ～ｄに対してユーザ位置の設定について、以下のように提案する。

「仲良し同士が自由に会話する設定でいいかな？」

このような提案を行う。

[0138] この処理は、図４を参照して説明したユーザ端末のユーザ位置決定部（ＵＩ）１０２を利用して実行することができる。

例えば図１７の（Ｓ１３）の右側に示すようなユーザ配置データが各ユーザ端末に表示され、各ユーザは、ユーザａ，１１ａの提案したユーザ位置を確認することができる。

[0139] 最後に、ステップＳ１４で、ユーザａ，１１ａが提案した設定、すなわち、

「公園で仲良し同士が自由に会話する設定」

この設定について全てのユーザａ～ｄが賛成すると、その後、会話が開始される。

[0140] ユーザ端末２１には、ユーザ間の親密度を算出する親密度算出部が設けられている。

ユーザ端末２１の親密度算出部は、各ユーザ間の会話状況を解析し、さらにユーザが入力した他のユーザの好み情報等を解析して、ユーザ間の親密度

を、遂次、算出し更新する。

なお、親密度算出処理の具体例については後述する。

[0141] 図１８を参照して、ユーザ間の親密度に応じたユーザ位置設定例、すなわち、ユーザ端末２１の表示部に表示する各ユーザ（ユーザ対応のアバターや実画像）の表示制御処理例について説明する。

[0142] 図１８には、ユーザａ，１１ａのユーザ端末ａ，２１ａの親密度算出部１２１が実行する親密度算出例と、出力画像制御部１１０が実行する算出親密度に従った各ユーザの表示位置の制御例を示している。

[0143] 親密度算出部１２１は、ユーザａ，１１ａがユーザ端末ａ，２１ａに入力した各ユーザの好み情報や、過去や現在のユーザ間の会話量等を解析して、ユーザａ，１１ａと他のユーザｂ～ｄ間のユーザ間親密度を算出する。

[0144] 図１８（ａ）に示すグラフは、親密度算出部１２１が算出した親密度の例を示すグラフである。

この例では、ユーザ間親密度は０～１０の数値として表現される。ユーザ間親密度＝０は親密度の最低値であり、ユーザ同士の仲が最も良くない状態であることを示している。ユーザ間親密度＝１０は親密度の最高値であり、ユーザ同士の仲が最も良い状態であることを示している。

[0145] 図１８（ａ）に示すグラフは以下のようなユーザ間親密度状態を示している。

ユーザａとユーザｂ間の親密度は１０であり、ユーザａとユーザｂの仲が最も良い状態であることを示している。

ユーザａとユーザｃ間の親密度は２であり、ユーザａとユーザｃの仲はあまり良くない状態であることを示している。

ユーザａとユーザｄ間の親密度は５であり、ユーザａとユーザｄの仲は良くも悪くもない普通の状態であることを示している。

[0146] ユーザ端末ａ，２１ａの出力画像制御部１１０は、この図１８（ａ）に示す親密度、すなわち親密度算出部１２１が算出した親密度情報を入力して、算出された親密度情報に従って各ユーザの表示位置を制御する。

出力画像制御部１１０が決定した親密度情報に従ったユーザ表示位置は図１８（ｂ）に示すような設定となる。

[0147] すなわち、出力画像制御部１１０は、各ユーザの表示位置を、親密度情報に従って、以下のように決定する。

ユーザａと最も仲が良い状態にある親密度１０のユーザｂの表示位置は、ユーザａに極めて近い（距離Ｌ１）位置とし、かつユーザａのほぼ正面位置とする。

ユーザａと仲があまり良い状態にない親密度２のユーザｃの表示位置は、ユーザａから遠い（距離Ｌ３）位置とし、かつユーザａのほぼ背面位置とする。

ユーザａと仲が良くも悪くもない状態にある親密度５のユーザｄの表示位置は、ユーザａから近くも遠くもない位置（距離Ｌ２）とし、かつユーザａの斜め前位置を表示位置とする。

なお、距離Ｌ１，Ｌ２，Ｌ３の大小関係は、Ｌ１＜Ｌ２＜Ｌ３である。

[0148] このように、出力画像制御部１１０は、親密度が高いユーザほど、自身のユーザ表示位置からの距離を短く設定し、かつ自身のユーザ表示位置の正面に近い位置に表示する制御を行う。

一方、親密度が低いユーザほど、自身のユーザ表示位置からの距離を長く設定し、かつ自身のユーザ表示位置の正面から離れた位置に表示する制御を行う。

[0149] 図１９は、親密度算出部１２１が算出した親密度情報に従って出力画像制御部１１０が決定した各ユーザの表示位置に従って、画像出力部１１１が表示部１１２に表示する画像の具体例を示す図である。

[0150] 図１９（ｂ）は、先に図１８（ｂ）を参照して説明した出力画像制御部１１０が親密度算出部１２１の算出親密度情報に従って決定した各ユーザの表示位置を示している。

[0151] ユーザ端末ａ，２１ａの画像出力部１１１は、この出力画像制御部１１０が決定した各ユーザの表示位置情報を入力して、図１９（ｃ）に示すように

、背景画像（公園の背景画像）上に各ユーザ、すなわちユーザａ，１１ａ～ユーザｄ，１１ｄの画像（アバター、または実画像）を表示する。

[0152] ユーザａ，１１ａは、ユーザ端末ａ，２１ａに表示されたこの画像を見ながら、各ユーザとの会話を進めることになる。

ユーザａ，１１ａと最も仲の良いユーザｂがユーザａの近くの正面位置に表示され、この結果、ユーザａ，１１ａは、より積極的にユーザｂ，１１ｂに話しかけやすくなる。

一方、ユーザａ，１１ａと最も仲の悪いユーザｃはユーザａから離れた背面位置に表示される。この結果、ユーザａ，１１ａは、ユーザｃ，１１ｃとはあまり話さない状態となる。

[0153] なお、各ユーザの表示位置は、各ユーザの利用するユーザ端末の親密度算出部１２１の算出した親密度に従って決定される。従って、ユーザａ～ｄの利用するユーザ端末ａ～ｄ各々において、各ユーザの表示位置は異なる設定となる場合がある。

[0154] 次に、ユーザが利用するユーザ端末２１が実行する親密度に従った音声制御処理について説明する。

前述したように、例えばユーザａ，１１ａのユーザ端末ａ，２１ａは、ユーザａ，１１ａと仲がよい（親密度が高い）ユーザの発話を大きな音量とする音声出力制御を実行する。

一方、ユーザａ，１１ａと仲が悪い（親密度が低い）ユーザの発話を小さな音量とする音声出力制御を実行する。

[0155] 図２０には、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの出力音声制御部１０８が実行する親密度に応じたユーザ発話出力音量の制御処理例を示すグラフを示している。

[0156] 図２０に示すグラフは横軸にユーザ間親密度、縦軸にユーザ発話出力音量を示すグラフである。

グラフから理解されるように、ユーザ間親密度が高いほど、ユーザ発話の出力音量は大きくなる。

[0157] すなわち、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザａ，１１ａとの親密度の高いユーザの発話は、出力音量を大きくする出力音量制御処理を実行する。

一方、ユーザａ，１１ａとの親密度の低いユーザの発話は、出力音量を小さくする出力音量制御処理を実行する。

[0158] 図２１を参照して、ユーザａ，１１ａとユーザｂ～ｄ間の具体的な親密度に対応する出力音声制御処理例について説明する。

ユーザ間の親密度は、親密度算出部１２１が算出する親密度である。

ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの親密度算出部１２１が算出したユーザａ，１１ａとユーザｂ～ｄ間の親密度を先に図１８を参照して説明したと同様の密度とする。

[0159] すなわち、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの親密度算出部１２１が以下のようなユーザ間親密度を算出したとする。

ユーザａとユーザｂ間の親密度は１０、すなわち高い親密度である。

ユーザａとユーザｃ間の親密度は２、すなわち低い親密度である。

ユーザａとユーザｄ間の親密度は５、すなわち、中程度の親密度である。

[0160] この場合、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの出力音声制御部１０８は、図２１に示すように、ユーザｂ～ｄの各発話の出力音量を以下のように制御する。

[0161] 親密度＝１０、すなわち高い親密度のユーザｂ，１１ｂの発話音量については大きな音量（Ｖｏｌ．３）に設定する音量制御を実行して、音声出力部１０９を介してユーザのヘッドホン等のスピーカーに出力する。

親密度＝５、すなわち中程度の親密度のユーザｄ，１１ｄの発話音量については中程度の音量（Ｖｏｌ．２）に設定する音量制御を実行して、音声出力部１０９を介してユーザのヘッドホン等のスピーカーに出力する。

親密度＝２、すなわち低い親密度のユーザｃ，１１ｃの発話音量については小さな音量（Ｖｏｌ．１）に設定する音量制御を実行して、音声出力部１０９を介してユーザのヘッドホン等のスピーカーに出力する。

[0162] このように、ユーザ端末２１の出力音声制御部１０８は、そのユーザ端末２１を利用するユーザとの親密度の高いユーザの発話は、出力音量を大きくし、親密度の低いユーザの発話は、出力音量を小さくする出力音量制御処理を実行する。

[0163] なお、ユーザ端末２１の出力音声制御部１０８は、このユーザ発話の出力音量制御処理に併せて、背景音の出力音量制御処理も実行する。

背景音の出力音量制御処理は図２２に示すように、ユーザ端末２１を利用するユーザとの親密度の高いユーザの発話が実行される場合の背景音の出力音量を小さくし、親密度の低いユーザの発話が実行される場合は、背景音の出力音量を大きくする出力音量制御処理を実行する。

[0164] すなわち、ユーザａ，１１ａが利用するユーザ端末ａ，２１ａの出力音声制御部１０８は、図２２に示すように以下のような制御を実行する。

親密度＝１０、すなわち高い親密度のユーザｂ，１１ｂが発話を行っている間は、背景音の出力音量を小さな音量（Ｖｏｌ．ｂ１）に設定する音量制御を実行して、音声出力部１０９を介してユーザのヘッドホン等のスピーカーに出力する。

[0165] 親密度＝５、すなわち中程度の親密度のユーザｄ，１１ｄが発話を行っている間は、背景音の出力音量を中程度の音量（Ｖｏｌ．ｂ２）に設定する音量制御を実行して、音声出力部１０９を介してユーザのヘッドホン等のスピーカーに出力する。

[0166] 親密度＝２、すなわち低い親密度のユーザｃ，１１ｃが発話を行っている間は、背景音の出力音量を大きな音量（Ｖｏｌ．ｂ３）に設定する音量制御を実行して、音声出力部１０９を介してユーザのヘッドホン等のスピーカーに出力する。

[0167] このように、ユーザ端末２１の出力音声制御部１０８は、ユーザ発話の出力音量制御処理に併せて、背景音の出力音量制御処理も実行する。

[0168] すなわち、本実施例２のユーザ端末２１の出力音声制御部１０８は、親密度の高いユーザについては、そのユーザ発話の音量を高めて、背景音の音量

を低下させる出力音量制御処理を実行し、一方、親密度の低いユーザについては、そのユーザ発話の音量を低下させて、背景音の音量を上昇させる出力音量制御処理を実行する。

[0169] ［３－２．実施例２のユーザ端末の構成例について］

次に、図２３を参照して、本実施例２において利用されるユーザ端末の一構成例について説明する。

[0170] 図２３には図１を参照して説明した通信ネットワークを介した会話の参加ユーザであるユーザａ，１１ａが利用するユーザ端末ａ，２１ａの構成例を示している。

なお、ユーザａ，１１ａ～ユーザｄ，１１ｄが利用するユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄは、全て図２３に示す構成例とほぼ同様の構成を有する。

[0171] なお、先に説明したように、ユーザ端末２１ａ～２１ｄは、例えばＰＣ、スマホ（スマートフォン）、タブレット端末等の通信可能な情報処理装置によって構成される。

[0172] 図２３に示すように、ユーザ端末ａ，２１ａは、通信部１０１、ユーザ位置決定部（ＵＩ）１０２、ユーザ位置情報格納部１０３、背景データ取得部１０４、音声データ格納部１０５、画像データ格納部１０６、音声データ受信部１０７、出力音声制御部１０８、音声出力部１０９、出力画像制御部１１０、画像出力部１１１、表示部１１２、音声入力部１１３、カメラ１１４、画像入力部１１５、データ送信部１１６と、親密度算出部１２１を有する。

[0173] この図２３に示すユーザ端末ａ，２１ａの構成は、先の実施例１において図４を参照して説明したユーザ端末ａ，２１ａの構成に親密度算出部１２１を追加した構成である。

親密度算出部１２１以外の構成は、実施例１において図４を参照して説明した構成と同様であるので、説明を省略する。

[0174] 本実施例２のユーザ端末ａ，２１ａは、例えば会話に参加する複数のユー

ザ各々の親密度に応じてユーザ発話の音声出力制御や、ユーザ画像（アバターや実画像）の表示制御を実行する。

親密度算出部１２１は、この音声出力制御処理や画像表示制御処理のベースとなる親密度を算出する。

[0175] 親密度算出部１２１において算出された親密度は、図２３に示すように出力音声制御部１０８と、出力画像制御部１１０に出力される。

出力音声制御部１０８と、出力画像制御部１１０は、親密度算出部１２１が算出した親密度に応じた音声出力制御、画像出力制御を実行する。

すなわち、例えば先に図１８～図２２を参照して説明した制御である。

[0176] 前述したように、親密度算出部１２１は、ユーザａ，１１ａがユーザ端末ａ，２１ａに入力した各ユーザの好み情報や、過去や現在のユーザ間の会話量等を解析して、ユーザａ，１１ａと他のユーザｂ～ｄ間のユーザ間親密度を算出する。

[0177] 図２４以下を参照して親密度算出部１２１の詳細な構成例と具体的な親密度算出処理例について説明する。

[0178] 図２４は、親密度算出部１２１の詳細構成例を示す図である。

図２４に示すように、親密度算出部１２１は、ユーザ好み入力部（ＵＩ）１４１、ユーザ好み解析部１４２、ユーザ好み情報格納部１４３、会話密度解析部１４４、親密度計算部１４５を有する。

[0179] ユーザ好み入力部（ＵＩ）１４１は、ユーザ自身が、他のユーザについて好みのレベルを直接、入力することを可能とした入力部（ＵＩ）である。

例えばユーザａ，１１ａが、他のユーザｂ～ｄ各々についての好みのレベル（例えばＬｅｖ．０～Ｌｅｖ．５）を入力する。

[0180] ユーザ好み入力部（ＵＩ）１４１を介してユーザが入力した他ユーザ各々についての好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）は、ユーザ好み解析部１４２に入力される。

[0181] ユーザ好み解析部１４２は、ユーザ好み入力部（ＵＩ）１４１に入力された他ユーザの好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）と、通信部を介して入

力するユーザ間の会話の解析結果に基づいて、ユーザ端末利用ユーザから見た他ユーザの最終的な好みレベル（例えばＬｅｖ．０～Ｌｅｖ．５）を算出する。

[0182] なお、ユーザ好み解析部１４２は、通信部を介して入力するユーザ間の会話の解析処理として、例えば以下のような解析処理を実行する。

ユーザａ，１１ａとよく意見が一致するユーザ、ユーザａ，１１ａとあまり意見が一致しないユーザ等を識別し、意見が一致するユーザほど好みレベルが高いと判定する処理。

お互いに笑いあっているユーザや、喧嘩をよくするユーザ等を識別し、お互い笑いあっているユーザほど好みレベルが高いと判定する処理。

[0183] ユーザ好み解析部１４２は、例えばこれらの会話解析処理を行って各ユーザ間の親密度を推定する。

なお、これら会話解析結果に基づく好みレベルの推定には、例えば様々な会話データの学習処理結果を利用する構成としてもよい。

[0184] このように、ユーザ好み解析部１４２は、ユーザ好み入力部（ＵＩ）１４１から入力した他ユーザ各々についての好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）と、通信部を介して入力するユーザ間の会話の解析によって取得した好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）を合成して、ユーザ端末利用ユーザから見た他ユーザ各々に対する最終的な好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）を算出する。

[0185] ユーザ好み解析部１４２によって算出された他ユーザ各々に対する最終的な好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）データは、ユーザ好み情報格納部１４３に格納される。

[0186] 会話密度解析部１４４は、通信部を介して入力するユーザ間の会話の解析処理を実行し、各ユーザ間の会話密度を算出する。

会話密度解析部１４４は、ユーザ間の会話の解析処理を実行し、各ユーザ間の会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）を算出する。

[0187] 会話密度解析部１４４は、例えば直接的なユーザ間の会話量や、ボイスチ

ャット量、さらにユーザの名前を呼ぶ回数等も考慮してユーザ間の会話の解析処理を実行し、各ユーザ間の会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）を算出する。

[0188] 頻繁に会話が行われ、会話密度が高いほど会話密度レベルは高い値（Ｌｅｖ．５に近い値）として算出される。一方、あまり会話がなされず会話密度が低いほど会話密度レベルは低い値（Ｌｅｖ．０に近い値）として算出される。

[0189] 会話密度解析部１４４が算出した会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）は親密度計算部１４５に入力される。

[0190] 親密度計算部１４５は、ユーザ好み情報格納部１４３に格納されたユーザ好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）と、会話密度解析部１４４が算出した会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）を利用して、ユーザ間の親密度レベルを算出する。

[0191] 例えばユーザ端末ａ，２１ａの親密度計算部１４５は、ユーザ端末ａ，２１ａの利用ユーザであるユーザａ，１１に対する他ユーザｂ～ｄ各々に対する親密度を算出する。

[0192] 例えば、親密度計算部１４５は、まず、ユーザ好み情報格納部１４３に格納されたユーザ好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）に応じた親密度である「ユーザ好みベース親密度」を算出する。

[0193] 図２５に示すグラフは、親密度計算部１４５が算出する「ユーザ好みベース親密度」の算出処理例を示す。図２５には、横軸にユーザ好み情報格納部１４３に格納されたユーザ好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）、縦軸に「ユーザ好みベース親密度」を示すグラフである。

[0194] 親密度計算部１４５は、例えばこのようなグラフ（＝関係式）を利用して、ユーザ好み情報格納部１４３に格納されたユーザ好みレベル（Ｌｅｖ．０～Ｌｅｖ．５）に基づいて「ユーザ好みベース親密度」を算出する。

[0195] 図２５に示す例は、ユーザａ，１１ａの好みが最も高いユーザが、ユーザｂ，１１ｂであり、このユーザｂ，１１ｂに対する「ユーザ好みベース親密

度」が最も高い値（約９．０）として算出される。

次に、ユーザａ，１１ａの好みが高いユーザが、ユーザｄ，１１ｄであり、このユーザｄ，１１ｄに対する「ユーザ好みベース親密度」が次に高い値（約５．８）として算出される。

さらに、ユーザａ，１１ａの好みが最も低いユーザが、ユーザｃ，１１ｃであり、このユーザｃ，１１ｃに対する「ユーザ好みベース親密度」が最も低い値（約２．１）として算出される。

[0196] 親密度計算部１４５は、次に、会話密度解析部１４４が算出した会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）に応じた親密度である「会話密度ベース親密度」を算出する。

[0197] 図２６に示すグラフは、親密度計算部１４５が算出する「会話密度ベース親密度」の算出処理例を示す。図２５には、横軸に会話密度解析部１４４が算出した会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）、縦軸に「会話密度ベース親密度」を示すグラフである。

[0198] 親密度計算部１４５は、例えばこのようなグラフ（＝関係式）を利用して、会話密度解析部１４４が算出した会話密度レベル（Ｌｅｖ．０～Ｌｅｖ．５）に基づいて、「会話密度ベース親密度」を算出する。

[0199] 図２６に示す例は、ユーザａ，１１ａとの会話密度が最も高いユーザが、ユーザｂ，１１ｂであり、このユーザｂ，１１ｂに対する「会話密度ベース親密度」が最も高い値（約７．２）として算出される。

次に、ユーザａ，１１ａとの会話密度が高いユーザが、ユーザｄ，１１ｄであり、このユーザｄ，１１ｄに対する「会話密度ベース親密度」が次に高い値（約４．５）として算出される。

さらに、ユーザａ，１１ａの好みが最も低いユーザが、ユーザｃ，１１ｃであり、このユーザｃ，１１ｃに対する「会話密度ベース親密度」が最も低い値（約１．８）として算出される。

[0200] 最後に、親密度計算部１４５は、図２５に示すグラフに従って算出した「ユーザ好みベース親密度」と、図２６に示すグラフに従って算出した「会話

密度ベース親密度」を用いた演算処理を実行して最終的なユーザ対応の「親密度」を算出する。

[0201] 例えば、「ユーザ好みベース親密度」の値をｐ（ｐ＝０～１０）、「会話密度ベース親密度」の値をｑ（ｑ＝０～１０）とし、最終的な親密度をｒ（ｒ＝０～１０）とした場合、以下のような演算式（重み加算演算式）を用いて、最終的な親密度ｒ（ｒ＝０～１０）を算出する。

最終的な親密度ｒ＝αｐ＋βｑ

なお、　ｐ：ユーザ好みベース親密度

ｑ：会話密度ベース親密度

であり、

α、βは乗算重み係数（ただしα＋β＝１）である。これらの重み係数は事前に設定する。

[0202] ユーザａ，１１ａの利用するユーザ端末ａ，２１ａでは、

ユーザａ，１１ａとユーザｂ，１１ｂの親密度、

ユーザａ，１１ａとユーザｃ，１１ｃの親密度、

ユーザａ，１１ａとユーザｄ，１１ｄの親密度、

これら３人のユーザｂ～ｄに対する最終的な親密度ｒ（０～１０）が算出される。

[0203] 親密度計算部１４５が算出したこれら３人のユーザｂ～ｄに対する最終的な親密度ｒ（０～１０）の値は、図２３や図２４に示すように、出力音声制御部１０７と、出力画像制御部１１０に入力される。

[0204] 出力音声制御部１０７と、出力画像制御部１１０は、親密度算出部１２１から入力した３人のユーザｂ～ｄに対する最終的な親密度ｒ（０～１０）に応じた音声出力制御、画像出力制御を実行する。

すなわち、例えば先に図１８～図２２を参照して説明したような音声出力制御、画像出力制御を実行する。

[0205] なお、ユーザ好み解析部１４２や、会話密度解析部１４４は、ユーザ間で会話が行われている間も通信部１０１を介して、これらの会話情報を入力し

て解析処理を継続し、遂次「ユーザ好みベース親密度」や、「会話密度ベース親密度」を更新し、更新データを親密度計算部１４５に入力する。

[0206] 親密度計算部１４５は、ユーザ好み解析部１４２や、会話密度解析部１４４から入力する最新の「ユーザ好みベース親密度」や、「会話密度ベース親密度」を利用して、最終的な親密度の値も、遂次更新する処理を行い、更新された値を継続的に出力音声制御部１０７と、出力画像制御部１１０に入力している。

[0207] 従って、出力音声制御部１０７と、出力画像制御部１１０は、会話の実行中に更新される最新の親密度の値に応じて、制御態様を逐次、変更する処理を行うことになる。

[0208] 図２７を参照して、複数ユーザａ～ｄ間の会話実行中に出力画像制御部１１０が実行する表示態様の変更処理例について説明する。

[0209] 図２７は、複数ユーザａ～ｄ間の会話実行中の出力画像制御部１１０が実行する表示制御処理例を示す図であり、以下の２つの時間における表示データの例を示している。

（ａ）時間＝ｔ０

（ｂ）時間＝ｔ１

なお、時間ｔ１は時間ｔ０から一定時間、経過した後の時間である。

[0210] まず、時間ｔ０では、ユーザａとユーザｂ～ｄの親密度が以下の設定である。

ユーザａとユーザｂの親密度＝１０

ユーザａとユーザｃの親密度＝２

ユーザａとユーザｄの親密度＝５

[0211] この親密度設定では、出力画像制御部１１０は図２７上段の（ａ）に示すような表示制御処理を実行する。

ユーザａと最も仲が良い状態にある親密度１０のユーザｂの表示位置は、ユーザａに極めて近い（距離Ｌ１）位置とし、かつユーザａのほぼ正面位置とする。

ユーザａと仲があまり良い状態にない親密度２のユーザｃの表示位置は、ユーザａから遠い（距離Ｌ３）位置とし、かつユーザａのほぼ背面位置とする。

ユーザａと仲が良くも悪くもない状態にある親密度５のユーザｄの表示位置は、ユーザａから近くも遠くもない位置（距離Ｌ２）とし、かつユーザａの斜め前位置を表示位置とする。

なお、距離Ｌ１，Ｌ２，Ｌ３の大小関係は、Ｌ１＜Ｌ２＜Ｌ３である。

[0212] さらに、出力音声制御部１０７は、各ユーザの発話に対する出力音声制御処理として、各ユーザの発話を各ユーザの表示位置から聞こえるように音声方向制御を実行し、さらにユーザａとの表示位置の距離（＝親密度）に応じた音量制御も実行する。

ユーザａとの距離が近い（＝親密度が高い）ほど、発話の出力音を大きな音量とする。一方、背景音はユーザａとの距離が近い（＝親密度が高い）ほど小さくする音量制御処理を実行する。

[0213] 図２７の下段に示す（ｂ）は、時間ｔ０から一定時間、後の時間ｔ１における表示制御処理例を示している。

時間ｔ０～ｔ１の間にもユーザａ～ｄ間の会話が行われ、ユーザ好み解析部１４２や、会話密度解析部１４４は、通信部１０１を介して会話情報を入力して解析処理を継続する。

[0214] ユーザ好み解析部１４２や、会話密度解析部１４４は、この解析処理の結果、遂次「ユーザ好みベース親密度」や、「会話密度ベース親密度」を更新し、更新データを親密度計算部１４５に入力する。

親密度計算部１４５は、ユーザ好み解析部１４２や、会話密度解析部１４４から入力する最新の「ユーザ好みベース親密度」や、「会話密度ベース親密度」を利用して、最終的な親密度の値を更新する。この更新値が出力音声制御部１０７と、出力画像制御部１１０に入力される。

[0215] 図２７（ｂ）に示す例は、時間＝ｔ１においてユーザａとユーザｂ～ｄの親密度が以下のように更新された例を示している。

ユーザａとユーザｂの親密度＝８

ユーザａとユーザｃの親密度＝４

ユーザａとユーザｄの親密度＝７

[0216] この親密度設定では、出力画像制御部１１０は図２７下段の（ｂ）に示すような表示制御処理を実行する。

ユーザａと最も仲が良い状態にある親密度８のユーザｂの表示位置は、ユーザａに極めて近い（距離Ｌ１’）位置とし、かつユーザａのほぼ正面位置とする。

[0217] 時間ｔ０では、ユーザａとの親密度＝２で、仲があまり良い状態になかったユーザｃは、時間ｔ１で親密度＝４に改善している。ユーザｃの表示位置は、時間ｔ０よりユーザａから近い位置（距離Ｌ３’）に変更され、さらにユーザａの背面位置から斜め前に移動して表示される。

[0218] 時間ｔ０で親密度＝５のユーザｄは、時間ｔ１で親密度＝７に親密度が低下している。これに伴い、ユーザｄの表示位置は、時間ｔ０よりユーザａから遠い位置（距離Ｌ２’）に変更される。

[0219] このように、時間経過に伴う親密度の変化により、各ユーザの表示位置が変更される。

[0220] さらに、出力音声制御部１０７も、各ユーザ対応の親密度の変更や、表示位置の変更にともない、各ユーザ発話に対する出力音声制御処理の態様も変更する。

すなわち、各ユーザの発話を各ユーザの新たな表示位置から聞こえるように音声方向制御を実行し、さらにユーザａとの新たな表示位置の距離（＝親密度）に応じた出力音量の変更処理を実行する。

[0221] なお、図２７では、出力画像制御部１１０が実行する親密度の変更に伴う表示制御処理として、ユーザの表示位置を変更する例を説明したが、出力画像制御部１１０は、さらに、表示するユーザの画像（アバター画像、または実画像）を変更する処理を行ってもよい。

[0222] 次に、図２８を参照して、出力画像制御部１１０が実行する親密度の変更

に伴うユーザ画像（アバター画像、または実画像）の変更処理例について説明する。

[0223] 例えば図２８（ａ）に示すように、親密度が高い状態では、正面向きのユーザ画像（アバター画像、または実画像）を表示し、親密度が低い状態では、横向きのユーザ画像（アバター画像、または実画像）を表示する。

出力画像制御部１１０は、このような親密度に従ったユーザ画像（アバター画像、または実画像）の表示制御を実行してもよい。

[0224] さらに、図２８（ｂ）に示すように、親密度が高い状態では、正面向きのユーザ画像（アバター画像、または実画像）を表示し、親密度が中程度の状態では、横向きのユーザ画像（アバター画像、または実画像）を表示し、親密度が低い状態では、後ろ向きのユーザ画像（アバター画像、または実画像）を表示する。

出力画像制御部１１０は、このような親密度に従ったユーザ画像（アバター画像、または実画像）の表示制御を実行してもよい。

[0225] ［４．（実施例３）各ユーザ端末で異なる背景データを利用する実施例について］

次に、実施例３として、各ユーザ端末で異なる背景データを利用する実施例について説明する。

[0226] 上述した実施例１，２では、基本的にネットワークを介して会話を行っている各ユーザのユーザ端末は、共通の背景データを表示して会話を行う例として説明した。

[0227] しかし、背景データの設定は、各ユーザ端末が個別に行うことが可能であり、ネットワークを介して会話を行っている各ユーザのユーザ端末各々が個別の異なる背景データをそれぞれのユーザ端末に出力しながら、ネットワークを介した会話を行うことも可能である。

以下に説明する実施例３は、このような処理を行う実施例である。

[0228] 図２９は、ネットワークを介して会話を行うユーザのユーザ端末各々が異なる背景データを各ユーザ端末に出力している例を示す図である。

[0229] 図２９に示す４人のユーザ、すなわちユーザａ，１１ａ～ユーザｄ，１１ｄは、それぞれのユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄを用いて通信を実行して会話を行っている。

[0230] ユーザａ，１１ａはユーザ端末ａ，２１ａにカフェの背景データを出力する設定をしており、ユーザ端末ａ，２１ａはカフェの背景画像とカフェの背景音を出力している。

一方、ユーザｂ，１１ｂはユーザ端末ｂ，２１ｂに公園の背景データを出力する設定をしており、ユーザ端末ｂ，２１ｂは公園の背景画像と公園の背景音を出力している。

[0231] また、ユーザｃ，１１ｃはユーザ端末ｃ，２１ｃにライブハウスの背景データを出力する設定をしており、ユーザ端末ｃ，２１ｃはライブハウスの背景画像とライブハウスの背景音を出力している。

さらに、ユーザｄ，１１ｄはユーザ端末ｄ，２１ｄに、ユーザ端末ｂ，２１ｂと同様、公園の背景データを出力する設定をしており、ユーザ端末ｄ，２１ｄは公園の背景画像と公園の背景音を出力している。

[0232] このようにユーザａ，１１ａ～ユーザｄ，１１ｄがそれぞれのユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄに異なる背景データを出力しながら相互に会話が行われる。

[0233] なお、各ユーザ端末に表示されるユーザの位置についても、各ユーザ端末において決定した端末ごとに異なるユーザ位置とすることが可能である。

例えば実施例１において説明した固定的なユーザ位置を設定することも可能であり、実施例２において説明した親密度に応じて決定されるユーザ位置としてもよい。

[0234] なお、このようにネットワークを介して接続されたユーザ端末にそれぞれ異なる背景データが出力されている場合、自端末に設定した背景データを継続して出力する設定も可能であるが、他のユーザとの会話時にのみ、会話相手のユーザ端末に設定された背景データを会話相手のユーザ発話データに併せて受信して自端末に出力する設定とすることも可能である。

[0235] 他ユーザとの会話の有無に関わらず自端末に設定した背景データを継続して出力する設定とするか、他ユーザとの会話時に会話相手のユーザ端末に設定された背景データを自端末に出力する設定とするかについては、各ユーザ端末において設定することが可能である。

これらの設定は各ユーザ端末において個別に行うことが可能であり、例えばＵＩ等を利用して設定することができる。

[0236] ネットワークを介して接続されたユーザ端末にそれぞれ異なる背景データが出力されている構成において、ユーザ間の会話に応じて自端末に出力する背景データを切り替える複数の処理例について、図３０以下を参照して説明する。

[0237] 図３０以下を参照して、以下の各処理例１～４について、順次、説明する。

（処理例１）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例

（処理例２）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみを会話相手のユーザ端末に設定された背景音声に切り替える設定とした場合の処理例

（処理例３）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみならず、背景画像データも会話相手のユーザ端末に設定された背景データに切り替える設定とした場合の処理例

（処理例４）他ユーザから話しかけられて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例

[0238] ［４－１．（処理例１）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例について］

[0239] まず、図３０を参照して、（処理例１）、すなわち、他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データを継続して出力す

る設定とした場合の処理例について説明する。

[0240] 図３０には、通信による会話を行うユーザａ，１１ａとユーザｂ，１１ｂを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。

ユーザｂ，１１ｂは、公園の背景データの出力設定をしたユーザ端末ｂ，２１ｂを利用している。

[0241] この設定で、ユーザａ，１１ａは、ユーザｂ，１１ｂに話かける。

ユーザｂ，１１ｂは、ユーザａ，１１ａからの話しかけに対する応答発話を実行する。

なお、ユーザｂ，１１ｂは、ユーザａ，１１ａからの話しかけを、ユーザ端末ｂ，２１ｂに設定した背景データである公園内にいるユーザａ，１１ａからの話しかけであると認識する。

[0242] ユーザｂ，１１ｂによる応答発話は、ユーザ端末ｂ，２１ｂから通信ネットワークを介してユーザ端末ａ，２１ａに送信される。

ユーザ端末ｂ，２１ｂからユーザ端末ａ，２１ａに送信されるデータはユーザｂ，１１ｂの発話音声データのみであり、ユーザ端末ｂ，２１ｂに設定されている背景データ、すなわち公園の背景画像データや背景音声データは送信されない。

[0243] ユーザ端末ａ，２１ａは、ユーザ端末ｂ，２１ｂからユーザｂ，１１ｂの発話音声データのみを受信する。ユーザ端末ａ，２１ａの出力音声制御部１０８は、受信したユーザｂ，１１ｂの発話音声データに対する出力制御を実行して制御された音声が出力される。例えばユーザ端末ａ，２１ａに接続されたユーザａ，１１ａが装着したヘッドホンから出力される。

[0244] なお、このユーザｂ，１１ｂの応答発話に対して実行される制御処理は、例えば先に説明した実施例１に従った制御処理である。すなわち図３０に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｂ，１１ｂの相対位置から、ユーザｂ，１１ｂの発話が聞こえるように音声出力

制御がなされて出力される。

[0245] なお、ユーザａ，１１ａが装着したヘッドホンからはユーザ端末ａ，１１ａに設定しているカフェの背景音も継続して出力され、ユーザｂ，１１ｂの応答発話は、カフェの背景音に併せてヘッドホンから出力される。

[0246] この出力音声制御処理によって、図３０に示す例では、ユーザａ，１１ａは、ユーザｂ，１１ｂもユーザａ，１１ａと同じカフェにいると認識しながら会話を行うことができる。

[0247] ［４－２．（処理例２）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみを会話相手のユーザ端末に設定された背景音声に切り替える設定とした場合の処理例について］

[0248] 次に、図３１を参照して、（処理例２）、すなわち、他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみを会話相手のユーザ端末に設定された背景音声に切り替える設定とした場合の処理例について説明する。

[0249] 図３１にも、図３０と同様、通信による会話を行うユーザａ，１１ａとユーザｂ，１１ｂを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。

ユーザｂ，１１ｂは、公園の背景データの出力設定をしたユーザ端末ｂ，２１ｂを利用している。

[0250] この設定で、ユーザａ，１１ａは、ユーザｂ，１１ｂに話かける。

ユーザｂ，１１ｂは、ユーザａ，１１ａからの話しかけに対する応答発話を実行する。

なお、ユーザｂ，１１ｂは、ユーザａ，１１ａからの話しかけを、ユーザ端末ｂ，２１ｂに設定した背景データである公園内にいるユーザａ，１１ａからの話しかけであると認識する。

[0251] ユーザｂ，１１ｂによる応答発話は、ユーザ端末ｂ，２１ｂから通信ネットワークを介してユーザ端末ａ，２１ａに送信される。

この（処理例２）では、ユーザ端末ｂ，２１ｂからユーザ端末ａ，２１ａに対して、ユーザｂ，１１ｂの発話音声データのみならず、ユーザ端末ｂ，２１ｂに設定されている公園の背景音声データも送信される。ただし、公園の背景画像データは送信されない。

[0252] ユーザ端末ａ，２１ａは、ユーザ端末ｂ，２１ｂからユーザｂ，１１ｂの発話音声データと、公園の背景音声データを受信する。ユーザ端末ａ，２１ａの出力音声制御部１０８は、受信した公園の背景音声データと、ユーザｂ，１１ｂの発話音声データに対する出力制御を実行して制御された音声を出力する。例えばユーザ端末ａ，２１ａに接続されたユーザａ，１１ａが装着したヘッドホンから出力する。

[0253] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザｂ，１１ｂの応答発話に対して、図３０を参照して説明した（処理例１）と同様の制御処理を行う。すなわち、先に説明した実施例１に従った制御処理であり、図３１に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｂ，１１ｂの相対位置から、ユーザｂ，１１ｂの発話が聞こえるように音声出力制御がなされて出力される。

[0254] なお、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｂ，１１ｂの発話の出力タイミングにおいてのみ、ユーザ端末ｂ，１１ｂから受信した公園の背景音、例えば鳥のさえずりの音などからなる公園の背景音が出力される。

[0255] すなわち、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｂ，１１ｂの発話の出力タイミング以外は、ユーザ端末ａ，２１ａに設定した背景音であるカフェの背景音声データが出力されるが、ユーザｂ，１１ｂの発話の出力タイミングにおいてのみ、ユーザ端末ｂ，１１ｂから受信した公園の背景音、例えば鳥のさえずりの音などからなる公園の背景音に切り替えられる。

[0256] この出力音声制御処理によって、図３１に示す例では、ユーザａ，１１ａは、ユーザｂ，１１ｂが公園にいる設定で会話をしていると認識することが

できる。

[0257] ［４－３．（処理例３）他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみならず、背景画像データも会話相手のユーザ端末に設定された背景データに切り替える設定とした場合の処理例について］

[0258] 次に、図３２を参照して、（処理例３）、すなわち、他ユーザに話しかけて他ユーザと会話する場合、自端末に設定した背景データ中、背景音声データのみならず、背景画像データも会話相手のユーザ端末に設定された背景データに切り替える設定とした場合の処理例について説明する。

[0259] 図３２にも、図３０、図３１と同様、通信による会話を行うユーザａ，１１ａとユーザｂ，１１ｂを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。なお、図３２では、ユーザ端末ａ，２１ａの背景画像が公園の背景データに切り替えられた後の状態を示している。

ユーザｂ，１１ｂは、公園の背景データの出力設定をしたユーザ端末ｂ，２１ｂを利用している。

[0260] この設定で、ユーザａ，１１ａは、ユーザｂ，１１ｂに話かける。

ユーザｂ，１１ｂは、ユーザａ，１１ａからの話しかけに対する応答発話を実行する。

なお、ユーザｂ，１１ｂは、ユーザａ，１１ａからの話しかけを、ユーザ端末ｂ，２１ｂに設定した背景データである公園内にいるユーザａ，１１ａからの話しかけであると認識する。

[0261] ユーザｂ，１１ｂによる応答発話は、ユーザ端末ｂ，２１ｂから通信ネットワークを介してユーザ端末ａ，２１ａに送信される。

この（処理例３）では、ユーザ端末ｂ，２１ｂからユーザ端末ａ，２１ａに対して、ユーザｂ，１１ｂの発話音声データのみならず、ユーザ端末ｂ，２１ｂに設定されている公園の背景音声データと公園の背景画像データが送信される。

[0262] ユーザ端末ａ，２１ａは、ユーザ端末ｂ，２１ｂからユーザｂ，１１ｂの発話音声データと、公園の背景音声データ、および公園の背景画像データを受信する。

[0263] ユーザ端末ａ，２１ａの出力音声制御部１０８は、受信した公園の背景音声データと、ユーザｂ，１１ｂの発話音声データに対する出力制御を実行して制御された音声を出力する。例えばユーザ端末ａ，２１ａに接続されたユーザａ，１１ａが装着したヘッドホンから出力する。

[0264] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザｂ，１１ｂの応答発話に対して、図３０を参照して説明した（処理例１）と同様の制御処理を行う。すなわち、先に説明した実施例１に従った制御処理であり、図３１に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｂ，１１ｂの相対位置から、ユーザｂ，１１ｂの発話が聞こえるように音声出力制御がなされて出力される。

[0265] また、ユーザ端末ａ，２１ａの出力画像制御部１０８は、ユーザｂ，１１ｂの応答発話の出力タイミングに合わせて、ユーザ端末ｂ，２１ｂから受信した公園の背景画像データをユーザ端末ａ，２１ａの表示部に出力する。

図３２に示すユーザ端末ａ，２１ａの表示部の状態である。

[0266] ただし、公園の背景画像上に配置する各ユーザのユーザ画像（アバター画像、または実画像）は、ユーザ端末ａ，２１ａのユーザ位置情報格納部１０３に格納されたユーザ位置に従って配置される。すなわち、ユーザ端末ａ，２１ａに設定した背景データであるカフェの画像上に配置したユーザ位置と同様の位置に表示されることになる。

[0267] これらの処理によって、ユーザｂ，１１ｂの発話の出力タイミングでは、ユーザ端末ａ，２１ａの表示部に公園の背景データが表示され、ユーザａ，１１ａが装着したヘッドホンからは、公園の背景音、例えば鳥のさえずりの音などからなる公園の背景音が出力される。

[0268] これらの出力画像制御処理と出力音声制御処理によって、図３２に示す例では、ユーザａ，１１ａは、ユーザｂ，１１ｂの発話の出力タイミングにお

いて、ユーザａ，１１ａがユーザｂ，１１ｂとともに公園にいる設定で会話をしているように認識することができる。

[0269] なお、ユーザｂ，１１ｂの発話が終了すると、ユーザ端末ａ，２１ａの背景データは、元のカフェの背景データに切り替えられる。すなわち表示部にはカフェの背景画像が表示され、ヘッドホンからはカフェの背景音が出力される。

[0270] すなわち、ユーザａ，１１ａは、ユーザｂ，１１ｂの発話の出力タイミングにおいてのみ、瞬間的にカフェから公園に移動して会話をするような認識を味わい、ユーザｂ，１１ｂの発話が終了すると元のカフェに戻るような感覚を味わうことができる。

[0271] ［４－４．（処理例４）他ユーザから話しかけられて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例について］

[0272] 次に、図３３を参照して、（処理例４）、すなわち、他ユーザから話しかけられて他ユーザと会話する場合、自端末に設定した背景データを継続して出力する設定とした場合の処理例について説明する。

[0273] 図３３には、図３０～図３２と同様、通信による会話を行うユーザａ，１１ａとユーザｂ，１１ｂを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。

ユーザｂ，１１ｂは、公園の背景データの出力設定をしたユーザ端末ｂ，２１ｂを利用している。

[0274] この設定で、ユーザａ，１１ａは、ユーザｂ，１１ｂから話かけられる。

なお、ユーザｂ，１１ｂは、ユーザ端末ｂ，２１ｂに設定した背景データである公園内にいるユーザａ，１１ａに対する話しかけを行っていると認識する。

[0275] ユーザｂ，１１ｂによる発話は、ユーザ端末ｂ，２１ｂから通信ネットワークを介してユーザ端末ａ，２１ａに送信される。

ユーザ端末ｂ，２１ｂからユーザ端末ａ，２１ａに送信されるデータはユーザｂ，１１ｂの発話音声データのみであり、ユーザ端末ｂ，２１ｂに設定されている背景データ、すなわち公園の背景画像データや背景音声データは送信されない。

[0276] ユーザ端末ａ，２１ａは、ユーザ端末ｂ，２１ｂからユーザｂ，１１ｂの発話音声データのみを受信する。ユーザ端末ａ，２１ａの出力音声制御部１０８は、受信したユーザｂ，１１ｂの発話音声データに対する出力制御を実行して制御された音声が出力される。例えばユーザ端末ａ，２１ａに接続されたユーザａ，１１ａが装着したヘッドホンから出力される。

[0277] なお、このユーザｂ，１１ｂの応答発話に対して実行される制御処理は、例えば先に説明した実施例１に従った制御処理である。すなわち図３０に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｂ，１１ｂの相対位置から、ユーザｂ，１１ｂの発話が聞こえるように音声出力制御がなされて出力される。

[0278] なお、ユーザａ，１１ａが装着したヘッドホンからはユーザ端末ａ，１１ａに設定しているカフェの背景音も継続して出力され、ユーザｂ，１１ｂの応答発話は、カフェの背景音に併せてヘッドホンから出力される。

[0279] この出力音声制御処理によって、図３３に示す例では、ユーザａ，１１ａは、ユーザｂ，１１ｂもユーザａ，１１ａと同じカフェにいると認識しながら会話を行うことができる。

[0280] 一方、ユーザｂ，１１ｂは、ユーザａ，１１ａとともに、ユーザ端末ｂ，２１ｂに設定した背景データである公園にいると認識しながら会話を行うことができる。

[0281] 次に、図３４以下を参照して、複数ユーザの会話中に、別の新たなユーザに話しかける場合、あるいは新たなユーザから話しかけられた場合の処理例について説明する。

[0282] 以下の３種類の処理例について、順次、説明する。

（処理例５）複数ユーザの会話中に別の新たなユーザに話しかける場合、

新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例

（処理例６）複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データと背景画像データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例

（処理例７）複数ユーザの会話中に別の新たなユーザから話しかけられた場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例

[0283] ［４－５．（処理例５）複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について］

[0284] まず、図３４を参照して、（処理例５）、すなわち、複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について説明する。

[0285] 図３４には、会話中のユーザａ，１１ａとユーザｂ，１１ｂ、さらに、別の新たなユーザｃ，１１ｃを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。

ユーザｃ，１１ｃは、ライブハウスの背景データの出力設定をしたユーザ端末ｃ，２１ｃを利用している。

[0286] ユーザｂ，１１ｂのユーザ端末ｂ，２１ｂについては示していないが、ユーザ端末ｂ，２１ｂの背景データの設定は様々な設定が可能である。ユーザ端末ｂ，２１ｂの背景データの切り替えは、ユーザ端末ａ，２１ａと同様に行われる。

[0287] この設定で、ユーザａ，１１ａ、またはユーザｂ，１１ｂが、ユーザｃ，１１ｃに話かける。

ユーザｃ，１１ｃは、例えばユーザａ，１１ａからの話しかけに対する応

答発話を実行する。

なお、ユーザｃ，１１ｃは、ユーザａ，１１ａからの話しかけを、ユーザ端末ｃ，２１ｃに設定した背景データであるライブハウス内にいるユーザａ，１１ａからの話しかけであると認識する。

[0288] ユーザｃ，１１ｃによる応答発話は、ユーザ端末ｃ，２１ｃから通信ネットワークを介してユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂに送信される。

この（処理例５）では、ユーザ端末ｃ，２１ｃからユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂに対して、ユーザｃ，１１ｃの発話音声データのみならず、ユーザ端末ｃ，２１ｃに設定されているライブハウスの背景音声データも送信される。ただし、ライブハウスの背景画像データは送信されない。

[0289] ユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂは、ユーザ端末ｃ，２１ｃからユーザｃ，１１ｃの発話音声データと、ライブハウスの背景音声データを受信する。ユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂの出力音声制御部１０８は、受信したライブハウスの背景音声データと、ユーザｃ，１１ｃの発話音声データに対する出力制御を実行して制御された音声を出力する。例えばユーザ端末ａ，２１ａやユーザ端末ｂ，２１ｂに接続されたヘッドホンから出力する。

[0290] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザｃ，１１ｃの応答発話に対して、図３０を参照して説明した（処理例１）と同様の制御処理を行う。すなわち、先に説明した実施例１に従った制御処理であり、図３４に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｃ，１１ｃの相対位置から、ユーザｃ，１１ｃの発話が聞こえるように音声出力制御がなされて出力される。

[0291] なお、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、ユーザ端末ｃ，１１ｃから受信したライブハウスの背景音が出力される。

[0292] すなわち、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｃ，１１ｃの発話の出力タイミング以外は、ユーザ端末ａ，２１ａに設定した背景音であるカフェの背景音声データが出力されるが、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、ユーザ端末ｃ，１１ｃから受信したライブハウスの背景音に切り替えられる。

[0293] この出力音声制御処理によって、ユーザａ，１１ａは、ユーザｃ，１１ｃがライブハウスにいる設定で会話をしていると認識することができる。

[0294] なお、図には示していないが、ユーザ端末ｂ，２１ｂにおいても同様の処理が実行される。ユーザｃ，１１ｃの発話音声と、ライブハウスの背景音がユーザ端末ｂ，２１ｂにも送信され、これらの音声データがユーザ端末ｂ，２１ｂを介して出力される。

例えばユーザ端末ｂ，２１ｂに設定された背景データが公園の背景データである場合、ユーザｃ，１１ｃの発話音声の出力タイミングにおいてのみ、公園の背景音がライブハウスの背景音に切り替えられて出力される。

[0295] この出力音声制御処理によって、ユーザｂ，１１ｂは、ユーザｃ，１１ｃがライブハウスにいる設定で会話をしていると認識することができる。

[0296] ［４－６．（処理例６）複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データと背景画像データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について］

[0297] 次に、図３５を参照して、（処理例６）、すなわち、複数ユーザの会話中に別の新たなユーザに話しかける場合、新たなユーザのユーザ端末の背景音声データと背景画像データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について説明する。

[0298] 図３５には、図３４と同様、会話中のユーザａ，１１ａとユーザｂ，１１ｂ、さらに、別の新たなユーザｃ，１１ｃを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。なお、図３５では、ユーザ端末ａ，２１ａの背景

画像がライブハウスの背景データに切り替えられた後の状態を示している。

ユーザｃ，１１ｃは、ライブハウスの背景データの出力設定をしたユーザ端末ｃ，２１ｃを利用している。

[0299] ユーザｂ，１１ｂのユーザ端末ｂ，２１ｂについては示していないが、ユーザ端末ｂ，２１ｂの背景データの設定は様々な設定が可能である。ユーザ端末ｂ，２１ｂの背景データの切り替えは、ユーザ端末ａ，２１ａと同様に行われる。

[0300] この設定で、ユーザａ，１１ａ、またはユーザｂ，１１ｂが、ユーザｃ，１１ｃに話かける。

ユーザｃ，１１ｃは、例えばユーザａ，１１ａからの話しかけに対する応答発話を実行する。

なお、ユーザｃ，１１ｃは、ユーザａ，１１ａからの話しかけを、ユーザ端末ｃ，２１ｃに設定した背景データであるライブハウス内にいるユーザａ，１１ａからの話しかけであると認識する。

[0301] ユーザｃ，１１ｃによる応答発話は、ユーザ端末ｃ，２１ｃから通信ネットワークを介してユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂに送信される。

この（処理例６）では、ユーザ端末ｃ，２１ｃからユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂに対して、ユーザｃ，１１ｃの発話音声データのみならず、ユーザ端末ｃ，２１ｃに設定されているライブハウスの背景音声データと背景画像データも送信される。

[0302] ユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂは、ユーザ端末ｃ，２１ｃからユーザｃ，１１ｃの発話音声データと、ライブハウスの背景音声データと背景画像データを受信する。

[0303] ユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂの出力音声制御部１０８は、受信したライブハウスの背景音声データと、ユーザｃ，１１ｃの発話音声データに対する出力制御を実行して制御された音声を出力する。例えばユーザ端末ａ，２１ａやユーザ端末ｂ，２１ｂに接続されたヘッドホンから出力

する。

[0304] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザｃ，１１ｃの応答発話に対して、図３０を参照して説明した（処理例１）と同様の制御処理を行う。すなわち、先に説明した実施例１に従った制御処理であり、図３４に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｃ，１１ｃの相対位置から、ユーザｃ，１１ｃの発話が聞こえるように音声出力制御がなされて出力される。

[0305] なお、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、ユーザ端末ｃ，１１ｃから受信したライブハウスの背景音が出力される。

[0306] すなわち、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｃ，１１ｃの発話の出力タイミング以外は、ユーザ端末ａ，２１ａに設定した背景音であるカフェの背景音声データが出力されるが、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、ユーザ端末ｃ，１１ｃから受信したライブハウスの背景音に切り替えられる。

[0307] また、ユーザ端末ａ，２１ａの出力画像制御部１０８は、ユーザｃ，１１ｃの応答発話の出力タイミングに合わせて、ユーザ端末ｃ，２１ｃから受信したライブハウスの背景画像データをユーザ端末ａ，２１ａの表示部に出力する。

図３５に示すユーザ端末ａ，２１ａの表示部の状態である。

[0308] ただし、ユーザ端末ａ，２１ａの表示部に表示されるライブハウスの背景画像上に配置する各ユーザのユーザ画像（アバター画像、または実画像）は、ユーザ端末ａ，２１ａのユーザ位置情報格納部１０３に格納されたユーザ位置に従って配置される。すなわち、ユーザ端末ａ，２１ａに設定した背景データであるカフェの画像上に配置したユーザ位置と同様の位置に表示されることになる。

[0309] これらの処理によって、ユーザｃ，１１ｃの発話の出力タイミングでは、ユーザ端末ａ，２１ａの表示部にライブハウスの背景データが表示され、ユ

ーザａ，１１ａが装着したヘッドホンからは、ライブハウスの背景音が出力される。

[0310] なお、図には示していないが、ユーザ端末ｂ，２１ｂにおいても同様の処理が実行される。ユーザｃ，１１ｃの発話音声と、ライブハウスの背景画像データと背景音声データがユーザ端末ｂ，２１ｂにも送信され、これらの画像データ、音声データがユーザ端末ｂ，２１ｂを介して出力される。

例えばユーザ端末ｂ，２１ｂに設定された背景データが公園の背景データである場合、ユーザｃ，１１ｃの発話音声の出力タイミングにおいてのみ、公園の背景画像がライブハウスの背景画像に切り替えられて表示され、背景音も公園の背景音からライブハウスの背景音に切り替えられて出力される。

[0311] これらの画像出力制御と出力音声制御処理によって、ユーザｂ，１１ｂも、ユーザｃ，１１ｃとともにライブハウスにいる設定で会話をしていると認識することができる。

[0312] これらの出力画像制御処理と出力音声制御処理によって、図３５に示す例では、ユーザａ，１１ａとユーザｂ，１１ｂは、ユーザｃ，１１ｃの発話の出力タイミングにおいて、ユーザａ，１１ａとユーザｂ，１１ｂが、ユーザｃ，１１ｃとともにライブハウス会場にいる設定で会話をしているように認識することができる。

[0313] なお、ユーザｃ，１１ｃの発話が終了すると、ユーザ端末ａ，２１ａの背景データは、元のカフェの背景データに切り替えられ、ユーザ端末ｂ，２１ｂの背景データも元の背景データ、例えば公園の背景データに切り替えられる。

[0314] すなわちユーザ端末ａ，２１ａの表示部にはカフェの背景画像が表示され、ヘッドホンからはカフェの背景音が出力される。また、ユーザ端末ｂ，２１ｂの表示部には公園の背景画像が表示され、ヘッドホンからは公園の背景音が出力される。

[0315] すなわち、ユーザａ，１１ａとユーザｂ，１１ｂは、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、瞬間的にカフェや公園からライブハウ

ス会場に移動して会話をするような認識を味わい、ユーザｃ，１１ｃの発話が終了すると元のカフェや公園に戻るような感覚を味わうことができる。

[0316] ［４－７．（処理例７）複数ユーザの会話中に別の新たなユーザから話しかけられた場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について］

[0317] 次に、図３６を参照して、（処理例７）、すなわち、複数ユーザの会話中に別の新たなユーザから話しかけられた場合、新たなユーザのユーザ端末の背景音声データを、会話中の複数ユーザのユーザ端末に送信して出力する設定とした処理例について説明する。

[0318] 図３６には、図３４、図３５と同様、会話中のユーザａ，１１ａとユーザｂ，１１ｂ、さらに、別の新たなユーザｃ，１１ｃを示している。

ユーザａ，１１ａは、カフェの背景データの出力設定をしたユーザ端末ａ，２１ａを利用している。

ユーザｃ，１１ｃは、ライブハウスの背景データの出力設定をしたユーザ端末ｃ，２１ｃを利用している。

[0319] ユーザｂ，１１ｂのユーザ端末ｂ，２１ｂについては示していないが、ユーザ端末ｂ，２１ｂの背景データの設定は様々な設定が可能である。ユーザ端末ｂ，２１ｂの背景データの切り替えは、ユーザ端末ａ，２１ａと同様に行われる。

[0320] この設定で、ユーザａ，１１ａ、またはユーザｂ，１１ｂが、ユーザｃ，１１ｃから話かけられる。

なお、ユーザｃ，１１ｃは、例えばユーザａ，１１ａに対する話しかけを、ユーザ端末ｃ，２１ｃに設定した背景データであるライブハウス内にいるユーザａ，１１ａに対する話しかけであると認識する。

[0321] ユーザｃ，１１ｃの発話は、ユーザ端末ｃ，２１ｃから通信ネットワークを介してユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂに送信される。

この（処理例７）では、ユーザ端末ｃ，２１ｃからユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂに対して、ユーザｃ，１１ｃの発話音声データのみ

ならず、ユーザ端末ｃ，２１ｃに設定されているライブハウスの背景音声データも送信される。ただし、ライブハウスの背景画像データは送信されない。

[0322] ユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂは、ユーザ端末ｃ，２１ｃからユーザｃ，１１ｃの発話音声データと、ライブハウスの背景音声データを受信する。ユーザ端末ａ，２１ａとユーザ端末ｂ，２１ｂの出力音声制御部１０８は、受信したライブハウスの背景音声データと、ユーザｃ，１１ｃの発話音声データに対する出力制御を実行して制御された音声を出力する。例えばユーザ端末ａ，２１ａやユーザ端末ｂ，２１ｂに接続されたヘッドホンから出力する。

[0323] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、ユーザｃ，１１ｃの発話に対して、図３０を参照して説明した（処理例１）と同様の制御処理を行う。すなわち、先に説明した実施例１に従った制御処理であり、図３４に示すユーザ端末ａ，２１ａに表示されるユーザａ，１１ａに対するユーザｃ，１１ｃの相対位置から、ユーザｃ，１１ｃの発話が聞こえるように音声出力制御がなされて出力される。

[0324] なお、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、ユーザ端末ｃ，１１ｃから受信したライブハウスの背景音が出力される。

[0325] すなわち、ユーザａ，１１ａが装着したヘッドホンからは、ユーザｃ，１１ｃの発話の出力タイミング以外は、ユーザ端末ａ，２１ａに設定した背景音であるカフェの背景音声データが出力されるが、ユーザｃ，１１ｃの発話の出力タイミングにおいてのみ、ユーザ端末ｃ，１１ｃから受信したライブハウスの背景音に切り替えられる。

[0326] この出力音声制御処理によって、ユーザａ，１１ａは、ユーザｃ，１１ｃがライブハウスにいる設定で会話をしていると認識することができる。

[0327] なお、図には示していないが、ユーザ端末ｂ，２１ｂにおいても同様の処理が実行される。ユーザｃ，１１ｃの発話音声と、ライブハウスの背景音が

ユーザ端末ｂ，２１ｂにも送信され、これらの音声データがユーザ端末ｂ，２１ｂを介して出力される。

例えばユーザ端末ｂ，２１ｂに設定された背景データが公園の背景データである場合、ユーザｃ，１１ｃの発話音声の出力タイミングにおいてのみ、公園の背景音がライブハウスの背景音に切り替えられて出力される。

[0328] この出力音声制御処理によって、ユーザｂ，１１ｂは、ユーザｃ，１１ｃがライブハウスにいる設定で会話をしていると認識することができる。

[0329] 以上、図２９～図３６を参照して、各ユーザ端末で異なる背景データを利用している場合の複数の処理例について説明した。

このようにユーザａ，１１ａ～ユーザｄ，１１ｄがそれぞれのユーザ端末ａ，２１ａ～ユーザ端末ｄ，２１ｄに異なる背景データを出力しながら相互に会話を行うことが可能であり、各ユーザ端末に表示するユーザ位置も、各ユーザ端末において異なる位置とすることが可能である。すなわち、実施例１において説明した固定的なユーザ位置や、実施例２において説明した親密度に応じた位置など、自由に設定できる。

[0330] 図３０～図３６を参照して説明したように、ユーザ端末に異なる背景データが出力されている場合、自端末に設定した背景データを継続して出力する処理や、会話相手のユーザ端末に設定された背景データを受信して自端末に出力する処理、いずれかの処理を行うことが可能であり、これらの設定は各ユーザ端末において、例えばＵＩ等を利用して個別に設定することができる。

[0331] ［５．ユーザ端末に背景データを出力してユーザ間で会話を行う具体的な処理シーケンスの例について］

次に、ユーザ端末に背景データを出力してユーザ間で会話を行う具体的な処理シーケンスの例について説明する。

[0332] 図３７以下を参照して、ユーザａ，１１ａがユーザ端末ａ，２１ａを利用して他のユーザとネットワークを介して会話を行う場合の処理例について説明する。

図３７以下に示す各ステップＳ２１～の処理について、順次、説明する。

[0333] （ステップＳ２１）

まず、ユーザａ，１１ａは、ユーザ端末ａ，２１ａの表示部にネットワークを介した会話を開始するための背景データの選択処理を行う。

[0334] 図３７（Ｓ２１）に示す例は、ネットワークを介した会話の開催場所を選択するための背景データ選択ＵＩの一例である。

図に示す例は、トークルームの設定場所となる背景データの候補として、「会議室」、「カフェ」の２候補が選択可能な背景データとして表示された例である。

ユーザａ，１１ａは、背景データとして「カフェ」を選択して「入店」ボタンをタッチするユーザ操作を行う。

[0335] （ステップＳ２２）

ステップＳ２１におけるユーザ操作によって、背景データとして「カフェ」の背景データが選択されると、ステップＳ２２において、ユーザ端末ａ，２１ａの表示部にカフェの背景画像が表示される。さらにユーザａ，１１ａのユーザ画像（アバター画像）がカフェの背景画像上に表示される。

なお、ユーザａ，１１ａのユーザ画像（アバター画像）の表示位置は、ユーザａ，１１ａの好みに応じて背景画像内の任意の位置に設定することが可能である。あるいは、例えば、図３７（Ｓ２２）に示す例のように、背景画像の左下等、予め決められた位置に表示するようにしてもよい。

[0336] また、ユーザ端末ａ，２１ａのスピーカー、またはユーザ端末ａ，２１ａに接続されたヘッドホンからカフェの背景音が出力される。

[0337] （ステップＳ２３）

次のステップＳ２３では、ユーザｂ，１１ｂがカフェに入店する。

ユーザｂ，１１ｂも自身のユーザ端末２１ｂに対して上述したステップＳ２１～Ｓ２２の処理と同様の操作を実行することで、ユーザａ，１１ａが入店している同じカフェに入店することができる。

[0338] なお、図３７（Ｓ２３）に示すユーザ端末ａ，２１ａの表示データ中のユ

ーザｂ，１１ｂの表示位置は、先に説明した実施例２に従った処理、すなわち親密度に応じて決定された位置である。

この状態で、ユーザａ，１１ａとユーザｂ，１１ｂは会話を開始する。

[0339] （ステップＳ２４）

図３８に示す次のステップＳ２４では、ユーザｃ，１１ｃがカフェに入店する。

ユーザｃ，１１ｃも自身のユーザ端末２１ｃに対して上述したステップＳ２１～Ｓ２２の処理と同様の操作を実行することで、ユーザａ，１１ａとユーザｂ，１１ｂが入店している同じカフェに入店することができる。

[0340] なお、図３８（Ｓ２４）に示すユーザ端末ａ，２１ａの表示データ中のユーザｃ，１１ｃの表示位置も、先に説明した実施例２に従った処理、すなわち親密度に応じて決定された位置である。

この状態で、ユーザａ，１１ａとユーザｂ，１１ｂ、ユーザｃ，１１ｃは互いに会話することができる。

[0341] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、各ユーザの表示位置に応じて各ユーザの発話の音声方向や音量を調整する制御を実行する。

[0342] 出力音声制御部１０８は、ユーザａ，１１ａの利用するスピーカーやヘッドホンからの出力音が、ユーザ端末ａ，２１ａの表示部に表示されたユーザａ，１１ａの位置において、ユーザｂ，ｃの各声をそれぞれユーザｂ，１１ｂとユーザｃ，１１ｃの表示位置から聞いていると同様の音声となるように制御する。

[0343] （ステップＳ２５）

次のステップＳ２５は、ユーザａ，１１ａとユーザｂ，１１ｂ、ユーザｃ，１１ｃが互いに会話を行った結果、ユーザａ，１１ａのユーザｃ，１１ｃに対する親密度が上昇し、この親密度上昇に伴うユーザの表示位置更新処理を行った例を示している。

[0344] この親密度上昇に伴い、ユーザ端末ａ，２１ａの出力画像制御部１１０は、ユーザｃ，１１ｃの表示位置をユーザａ，１１ａの表示位置に近い位置に

移動させるユーザ表示位置の更新処理を実行する。

[0345] （ステップＳ２６）

図３９に示す次のステップＳ２６では、ユーザｃ，１１ｃがカフェから退店し、ユーザｄ，１１ｄがカフェに入店する。

ユーザｄ，１１ｄも自身のユーザ端末２１ｄに対して上述したステップＳ２１～Ｓ２２の処理と同様の操作を実行することで、ユーザａ，１１ａとユーザｂ，１１ｂが入店している同じカフェに入店することができる。

[0346] なお、図３９（Ｓ２６）に示すユーザ端末ａ，２１ａの表示データ中のユーザｄ，１１ｄの表示位置も、先に説明した実施例２に従った処理、すなわち親密度に応じて決定された位置である。

この状態で、ユーザａ，１１ａとユーザｂ，１１ｂ、ユーザｄ，１１ｄは互いに会話することができる。

[0347] なお、ユーザ端末ａ，２１ａの出力音声制御部１０８は、各ユーザの表示位置に応じて各ユーザの発話の音声方向や音量を調整する制御を実行する。

[0348] 出力音声制御部１０８は、ユーザａ，１１ａの利用するスピーカーやヘッドホンからの出力音が、ユーザ端末ａ，２１ａの表示部に表示されたユーザａ，１１ａの位置において、ユーザｂ，ｄの各声をそれぞれユーザｂ，１１ｂとユーザｄ，１１ｄの表示位置から聞いていると同様の音声となるように制御する。

[0349] （ステップＳ２７）

次のステップＳ２７は、ユーザａ，１１ａとユーザｂ，１１ｂ、ユーザｄ，１１ｄが互いに会話を行った結果、ユーザａ，１１ａのユーザｂ，１１ｂと、ユーザｃ，１１ｃに対する親密度が変化し、この親密度変化に伴う各ユーザの表示位置更新処理を行った例を示している。

[0350] この親密度変化に伴い、ユーザ端末ａ，２１ａの出力画像制御部１１０は、ユーザｃ，１１ｃと、ユーザｄ，１１ｄの表示位置を変更するユーザ表示位置更新処理を実行する。

[0351] 次に、図４０と図４１を参照して、背景データ（画像，音声）の切り替え

処理例について説明する。

まず、図４０を参照して、同じカフェの背景画像であるが、異なる方向から見たカフェの画像に切り替える処理例について説明する。

[0352] （ステップＳ３１）

ステップＳ３１は、ユーザ端末ａ，２１ａにユーザａ，１１ａが設定した背景データであるカフェの背景画像、および背景音声を出力しながらユーザａ，１１ａと、ユーザｂ，１１ｂと、ユーザｄ，１１ｄで会話を行っている状態である。

[0353] （ステップＳ３２）

ステップＳ３２は、ユーザｄ，１１ｄの発話時に、ユーザｄ，１１ｄの利用端末であるユーザ端末ｄ，２１ｄからユーザ発話に併せて、ユーザ端末ｄ，２１ｄに設定された背景データ、すなわち異なる方向から見たカフェの画像データを含む背景データ（画像データ、音声データ）を入力して、ユーザ端末ａ，２１ａに出力した例である。

[0354] このように、同じカフェの背景データでも、各ユーザによって異なる位置から見た背景画像が利用されている場合、これらの画像を、各ユーザ端末から受信して切り替えて出力することが可能である。

[0355] 次に、図４１を参照して、カフェの背景データからライブハウスの背景データに切り替える処理例について説明する。

[0356] （ステップＳ４１）

ステップＳ４１は、ユーザ端末ａ，２１ａにユーザａ，１１ａが設定した背景データであるカフェの背景画像、および背景音声を出力しながらユーザａ，１１ａと、ユーザｂ，１１ｂと、ユーザｄ，１１ｄで会話を行っている状態である。

[0357] （ステップＳ４２）

ステップＳ４２は、ユーザａ，１１ａが、カフェには入店していないユーザｃ，１１ｃからの呼びかけに応じた場合の処理例を示している。

[0358] ユーザ端末ａ，１１ａは、カフェに入店しないユーザｃ，１１ｃの発話の

受信時に、ユーザｃ，１１ｃの利用端末であるユーザ端末ｃ，２１ｃからユーザ発話の受信に併せて、ユーザ端末ｃ，２１ｃに設定された背景データ、すなわちライブハウスの画像データを含む背景データ（画像データ、音声データ）を入力する。

[0359] ユーザ端末ａ，２１ａは、ユーザｃ，１１ｃからの呼びかけ発話をスピーカーやヘッドホンを介して出力する際、ユーザ端末ｃ，２１ｃから受信した背景データ、すなわちライブハウスの画像データを表示部に表示し、ライブハウスの音声データをスピーカーやヘッドホンを介して出力する。

[0360] この処理によって、ユーザａ，１１ａし、ユーザｃ，１１ｃがライブハウスから呼びかけていることを認識することが可能となり、興味があれば、ライブハウスに移動してユーザｃ，１１ｃと会話を行うことが可能となる。

[0361] 次に、図４２を参照して、ユーザ端末に表示するユーザ画像をアバター画像からも、ユーザの実画像（カメラ撮影画像）に切り替える処理例について説明する。

[0362] （ステップＳ５１）

ステップＳ５１は、ユーザ端末ａ，２１ａにユーザａ，１１ａが設定した背景データであるカフェの背景画像、および背景音声を出力しながらユーザａ，１１ａと、ユーザｂ，１１ｂと、ユーザｄ，１１ｄで会話を行っている状態である。

[0363] ここでユーザ端末ａ，２１ａの表示部に表示されているユーザａ，１１ａの画像は仮想的なキャラクタ画像、すなわちユーザａを示すアバター画像である。

[0364] （ステップＳ５２）

ステップＳ５２は、ユーザａ，１１ａが、ユーザ端末ａ，２１ａのカメラを起動して、ユーザａ，１１ａの顔画像を撮影して、ユーザ端末ａ，２１ａの表示部に表示されたユーザａ，１１ａのアバター画像を、ユーザａ，１１ａの実画像、すなわちカメラ撮影画像に切り替えた例を示している。

[0365] なお、ユーザ端末ａ，２１ａのカメラによって撮影されたユーザａ，１１

ａの顔画像は、ネットワークを介した会話を実行中の他のユーザであるユーザｂ，１１ｂのユーザ端末ｂ，２１ｂと、ユーザｄ，１１ｄのユーザ端末ｄ，２１ｄにも送信され、これらのユーザ端末の表示部にもユーザａ，１１ａの実画像（カメラ撮影画像）が表示される。

[0366] ［６．ユーザ端末、およびサーバのハードウェア構成例について］

次に、ユーザ端末、およびサーバのハードウェア構成例について説明する。

[0367] 図４３は、本開示のユーザ端末２１や、サーバのハードウェア構成の一例を示す図である。

以下、図４３に示すハードウェア構成について説明する。

[0368] ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）３０１は、ＲＯＭ（Ｒｅａｄ Ｏｎｌｙ Ｍｅｍｏｒｙ）３０２、または記憶部３０８に記憶されているプログラムに従って各種の処理を実行する制御部やデータ処理部として機能する。例えば、上述した実施例において説明したシーケンスに従った処理を実行する。ＲＡＭ（Ｒａｎｄｏｍ Ａｃｃｅｓｓ Ｍｅｍｏｒｙ）３０３には、ＣＰＵ３０１が実行するプログラムやデータなどが記憶される。これらのＣＰＵ３０１、ＲＯＭ３０２、およびＲＡＭ３０３は、バス３０４により相互に接続されている。

[0369] ＣＰＵ３０１はバス３０４を介して入出力インタフェース３０５に接続され、入出力インタフェース３０５には、各種スイッチ、キーボード、マウス、マイクロホン、センサーなどよりなる入力部３０６、ディスプレイ、スピーカーなどよりなる出力部３０７が接続されている。ＣＰＵ３０１は、入力部３０６から入力される指令に対応して各種の処理を実行し、処理結果を例えば出力部３０７に出力する。

[0370] 入出力インタフェース３０５に接続されている記憶部３０８は、例えばハードディスク等からなり、ＣＰＵ３０１が実行するプログラムや各種のデータを記憶する。通信部３０９は、Ｗｉ－Ｆｉ通信、ブルートゥース（登録商標）（ＢＴ）通信、その他インターネットやローカルエリアネットワークな

どのネットワークを介したデータ通信の送受信部として機能し、外部の装置と通信する。

[0371] 入出力インタフェース３０５に接続されているドライブ３１０は、磁気ディスク、光ディスク、光磁気ディスク、あるいはメモリカード等の半導体メモリなどのリムーバブルメディア３１１を駆動し、データの記録あるいは読み取りを実行する。

[0372] ［７．本開示の構成のまとめ］

以上、特定の実施例を参照しながら、本開示の実施例について詳解してきた。しかしながら、本開示の要旨を逸脱しない範囲で当業者が実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本開示の要旨を判断するためには、特許請求の範囲の欄を参酌すべきである。

[0373] なお、本明細書において開示した技術は、以下のような構成をとることができる。

（１） ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記出力音声制御部は、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する情報処理装置。

[0374] （２）前記出力音声制御部は、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音量制御を実行する（１）に記載の情報処理装置。

[0375] （３）前記自己位置に対する前記会話相手のユーザ位置は、予め決定された固定位置である（１）または（２）に記載の情報処理装置。

[0376] （４）前記自己位置に対する前記会話相手のユーザ位置は、前記会話相手

に対する親密度に応じて決定する位置である（１）～（３）いずれかに記載の情報処理装置。

[0377] （５）前記自己位置に対する前記会話相手のユーザ位置は、

前記会話相手に対する親密度が高いほど前記自己位置に近く、

前記会話相手に対する親密度が低いほど前記自己位置から遠い位置に決定される（４）に記載の情報処理装置。

[0378] （６）前記情報処理装置は、

会話相手ユーザに対する親密度を算出する親密度算出部を有し、

前記自己位置に対する前記会話相手のユーザ位置は、

前記親密度算出部が算出した親密度に応じて決定される位置である（１）～（５）いずれかに記載の情報処理装置。

[0379] （７）前記親密度算出部は、

前記情報処理装置の利用ユーザの前記会話相手ユーザに対する好みに応じて親密度を算出する（６）に記載の情報処理装置。

[0380] （８）前記親密度算出部は、

前記情報処理装置の利用ユーザの前記会話相手ユーザに対する好みを、過去の履歴に基づいて解析する（７）に記載の情報処理装置。

[0381] （９）前記親密度算出部は、

前記情報処理装置の利用ユーザと、前記会話相手ユーザとの会話密度に応じて親密度を算出する（６）～（８）いずれかに記載の情報処理装置。

[0382] （１０）前記情報処理装置は、

表示部に対する画像出力制御を行う出力画像制御部を有し、

前記出力画像制御部は、

自己を示す自己ユーザ画像と、前記会話相手のユーザ画像を表示部に表示する処理を実行する（１）～（９）いずれかに記載の情報処理装置。

[0383] （１１）前記出力画像制御部は、

前記情報処理装置を利用するユーザが決定した背景画像を表示部に表示し、

背景画像上に会話を行うユーザのユーザ画像を表示する（１０）に記載の情報処理装置。

[0384] （１２）前記出力画像制御部は、

自己を示す自己ユーザ画像と、前記会話相手のユーザ画像を予め決定された固定位置に表示する処理を実行する（１０）または（１１）に記載の情報処理装置。

[0385] （１３）前記情報処理装置は、

表示部に対する画像出力制御を行う出力画像制御部を有し、

前記出力画像制御部は、

自己を示す自己ユーザ画像に対する前記会話相手のユーザ画像の表示位置を、前記親密度算出部が算出した親密度に応じて決定する（６）～（１２）いずれかに記載の情報処理装置。

[0386] （１４）前記出力画像制御部は、

前記会話相手のユーザ発話を出力するタイミングにおいて、

表示部に表示する背景画像を、自端末で設定した背景画像から、前記会話相手のユーザ端末が設定した背景画像に切り替える処理を実行する（１０）～（１３）いずれかに記載の情報処理装置。

[0387] （１５）前記出力音声制御部は、

前記情報処理装置を利用するユーザが決定した背景音声を、音声出力部を介して出力する処理を実行する（１）～（１４）いずれかに記載の情報処理装置。

[0388] （１６）前記出力音声制御部は、

前記会話相手のユーザ発話を出力するタイミングにおいて、

音声出力部に出力する背景音声を、自端末で設定した背景音声から、前記会話相手のユーザ端末が設定した背景音声に切り替える処理を実行する（１）～（１５）いずれかに記載の情報処理装置。

[0389] （１７）情報処理装置において実行する情報処理方法であり、

前記情報処理装置は、

ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記出力音声制御部が、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する情報処理方法。

[0390] （１８）情報処理装置が情報処理を実行させるプログラムであり、

前記情報処理装置は、

ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記プログラムは、前記出力音声制御部に、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行させるプログラム。

[0391] また、明細書中において説明した一連の処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。ソフトウェアによる処理を実行する場合は、処理シーケンスを記録したプログラムを、専用のハードウェアに組み込まれたコンピュータ内のメモリにインストールして実行させるか、あるいは、各種処理が実行可能な汎用コンピュータにプログラムをインストールして実行させることが可能である。例えば、プログラムは記録媒体に予め記録しておくことができる。記録媒体からコンピュータにインストールする他、ＬＡＮ（Ｌｏｃａｌ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）、インターネットといったネットワークを介してプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

[0392] なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。また、本明細書においてシステムとは

、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

産業上の利用可能性

[0393] 以上、説明したように、本開示の一実施例の構成によれば、ネットワークを介した会話相手のユーザ発話を予め規定した自己位置に対する会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する構成が実現される。

具体的には、例えば、ネットワークを介した会話相手のユーザ発話を受信する通信部と、ユーザ発話の出力制御を実行する出力音声制御部を有する。出力音声制御部がユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御や音量制御を実行する。自己位置に対する会話相手のユーザ位置は、予め決定された固定位置、または会話相手に対する親密度に応じて決定し、親密度が高いほど自己位置に近い位置に設定する。

本構成により、ネットワークを介した会話相手のユーザ発話を予め規定した自己位置に対する会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する構成が実現される。

符号の説明

[0394]
１１　ユーザ
２１　ユーザ端末
５０　通信管理サーバ
７０　背景データ提供サーバ
１０１　通信部
１０２　ユーザ位置決定部（ＵＩ）
１０３　ユーザ位置情報格納部
１０４　背景データ取得部
１０５　音声データ格納部
１０６　画像データ格納部

１０７　音声データ受信部

１０８　出力音声制御部

１０９　音声出力部

１１０　出力画像制御部

１１１　画像出力部

１１２　表示部

１１３　音声入力部

１１４　カメラ

１１５　画像入力部

１１６　データ送信部

１２１　親密度算出部

１４１　ユーザ好み入力部（ＵＩ）

１４２　ユーザ好み解析部

１４３　ユーザ好み情報格納部

１４４　会話密度解析部

１４５　親密度計算部

３０１　ＣＰＵ

３０２　ＲＯＭ

３０３　ＲＡＭ

３０４　バス

３０５　入出力インタフェース

３０６　入力部

３０７　出力部

３０８　記憶部

３０９　通信部

３１０　ドライブ

３１１　リムーバブルメディア

請求の範囲

[請求項1] ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記出力音声制御部は、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する情報処理装置。

[請求項2] 前記出力音声制御部は、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音量制御を実行する請求項１に記載の情報処理装置。

[請求項3] 前記自己位置に対する前記会話相手のユーザ位置は、予め決定された固定位置である請求項１に記載の情報処理装置。

[請求項4] 前記自己位置に対する前記会話相手のユーザ位置は、前記会話相手に対する親密度に応じて決定する位置である請求項１に記載の情報処理装置。

[請求項5] 前記自己位置に対する前記会話相手のユーザ位置は、

前記会話相手に対する親密度が高いほど前記自己位置に近く、

前記会話相手に対する親密度が低いほど前記自己位置から遠い位置に決定される請求項４に記載の情報処理装置。

[請求項6] 前記情報処理装置は、

会話相手ユーザに対する親密度を算出する親密度算出部を有し、

前記自己位置に対する前記会話相手のユーザ位置は、

前記親密度算出部が算出した親密度に応じて決定される位置である請求項１に記載の情報処理装置。

[請求項7] 前記親密度算出部は、

前記情報処理装置の利用ユーザの前記会話相手ユーザに対する好みに応じて親密度を算出する請求項６に記載の情報処理装置。

[請求項8] 前記親密度算出部は、

前記情報処理装置の利用ユーザの前記会話相手ユーザに対する好みを、過去の履歴に基づいて解析する請求項７に記載の情報処理装置。

[請求項9] 前記親密度算出部は、

前記情報処理装置の利用ユーザと、前記会話相手ユーザとの会話密度に応じて親密度を算出する請求項６に記載の情報処理装置。

[請求項10] 前記情報処理装置は、

表示部に対する画像出力制御を行う出力画像制御部を有し、

前記出力画像制御部は、

自己を示す自己ユーザ画像と、前記会話相手のユーザ画像を表示部に表示する処理を実行する請求項１に記載の情報処理装置。

[請求項11] 前記出力画像制御部は、

前記情報処理装置を利用するユーザが決定した背景画像を表示部に表示し、

背景画像上に会話を行うユーザのユーザ画像を表示する請求項１０に記載の情報処理装置。

[請求項12] 前記出力画像制御部は、

自己を示す自己ユーザ画像と、前記会話相手のユーザ画像を予め決定された固定位置に表示する処理を実行する請求項１０に記載の情報処理装置。

[請求項13] 前記情報処理装置は、

表示部に対する画像出力制御を行う出力画像制御部を有し、

前記出力画像制御部は、

自己を示す自己ユーザ画像に対する前記会話相手のユーザ画像の表示位置を、前記親密度算出部が算出した親密度に応じて決定する請求項６に記載の情報処理装置。

[請求項14] 前記出力画像制御部は、

前記会話相手のユーザ発話を出力するタイミングにおいて、

表示部に表示する背景画像を、自端末で設定した背景画像から、前記会話相手のユーザ端末が設定した背景画像に切り替える処理を実行する請求項１０に記載の情報処理装置。

[請求項15] 前記出力音声制御部は、

前記情報処理装置を利用するユーザが決定した背景音声を、音声出力部を介して出力する処理を実行する請求項１に記載の情報処理装置。

[請求項16] 前記出力音声制御部は、

前記会話相手のユーザ発話を出力するタイミングにおいて、

音声出力部に出力する背景音声を、自端末で設定した背景音声から、前記会話相手のユーザ端末が設定した背景音声に切り替える処理を実行する請求項１に記載の情報処理装置。

[請求項17] 情報処理装置において実行する情報処理方法であり、

前記情報処理装置は、

ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

前記出力音声制御部が、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行する情報処理方法。

[請求項18] 情報処理装置が情報処理を実行させるプログラムであり、

前記情報処理装置は、

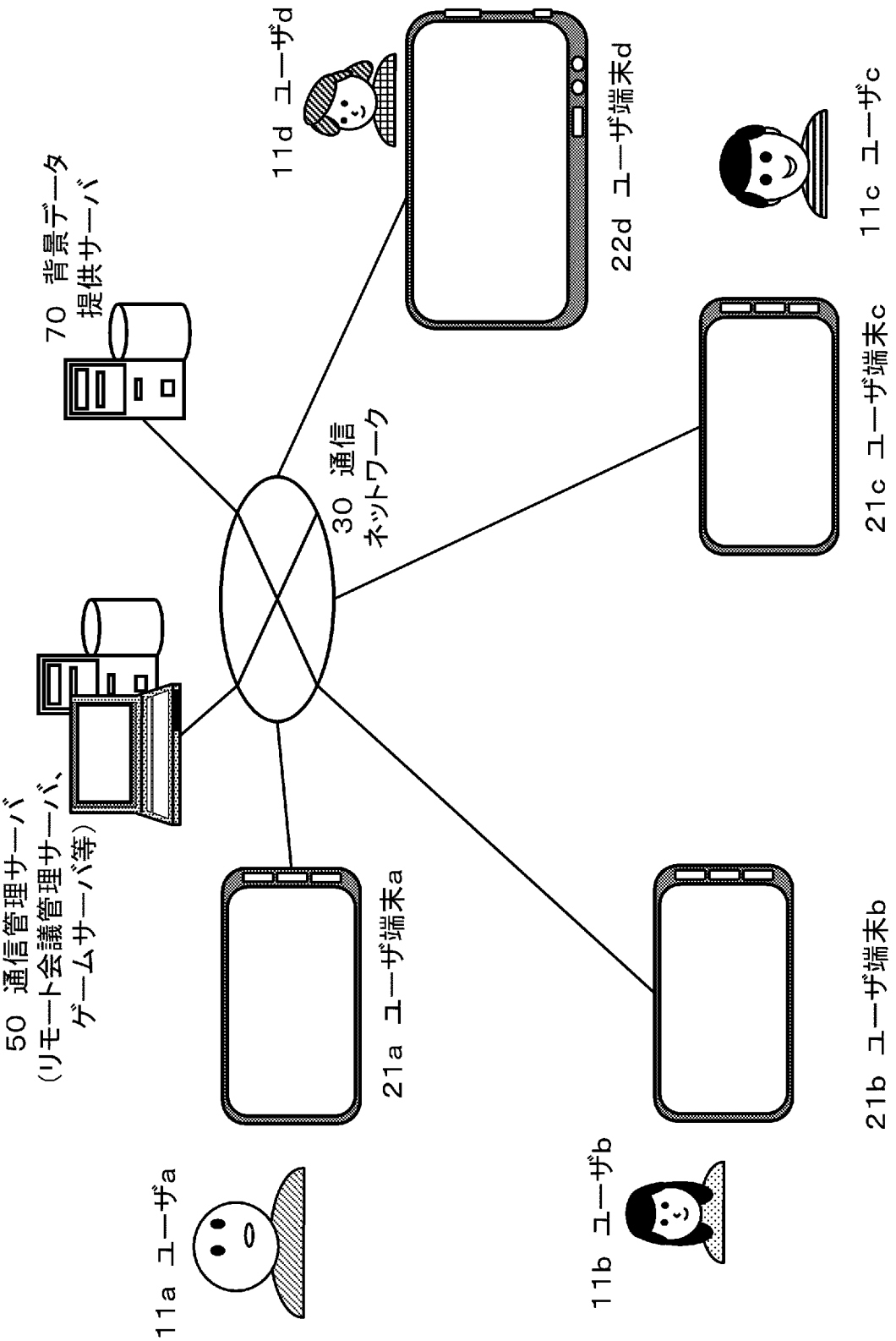
ネットワークを介した会話相手のユーザ発話を受信する通信部と、

前記ユーザ発話の出力制御を実行する出力音声制御部を有し、

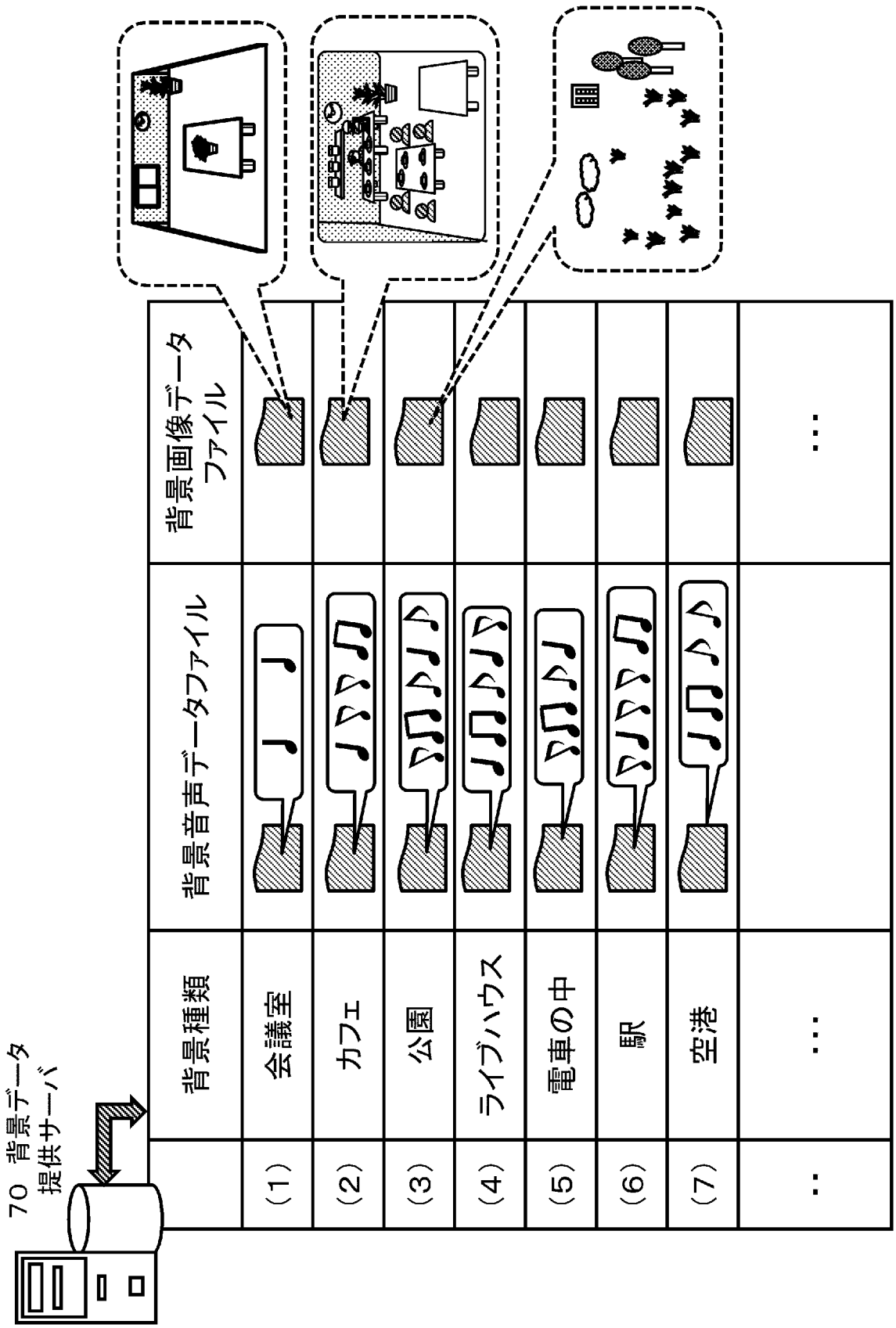
前記プログラムは、前記出力音声制御部に、

前記ユーザ発話を、予め規定した自己位置に対する前記会話相手のユーザ位置からの発話として聞こえるように音声方向制御を実行させるプログラム。

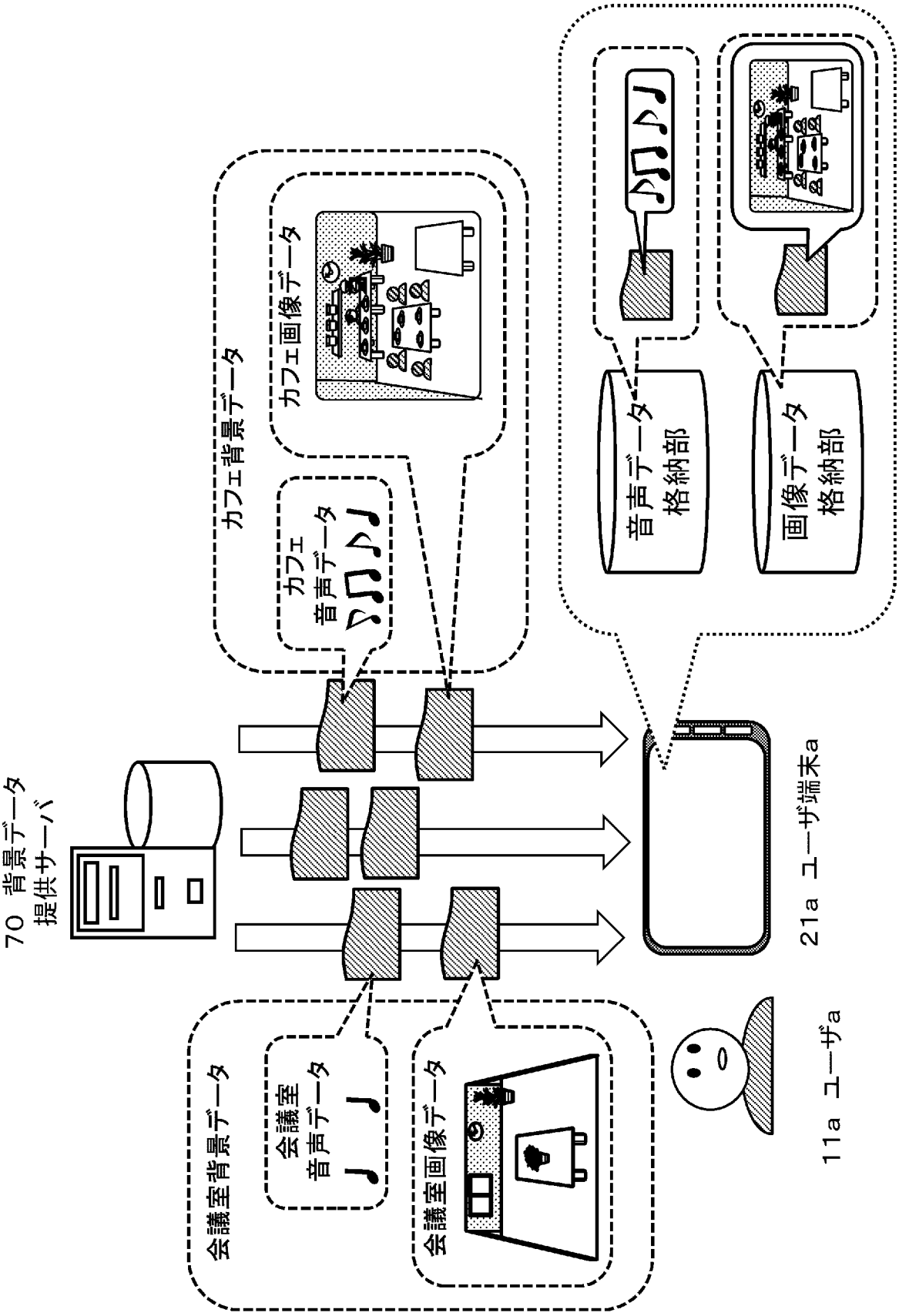
[図1]


11d ユーザd
22d ユーザ端末d
70 背景データ
提供サーバ
11c ユーザc
30 通信
ネットワーク
21c ユーザ端末c
50 通信管理サーバ
(リモート会議管理サーバ、
ゲームサーバ等)
21a ユーザ端末a
21b ユーザ端末b
11a ユーザa
11b ユーザb

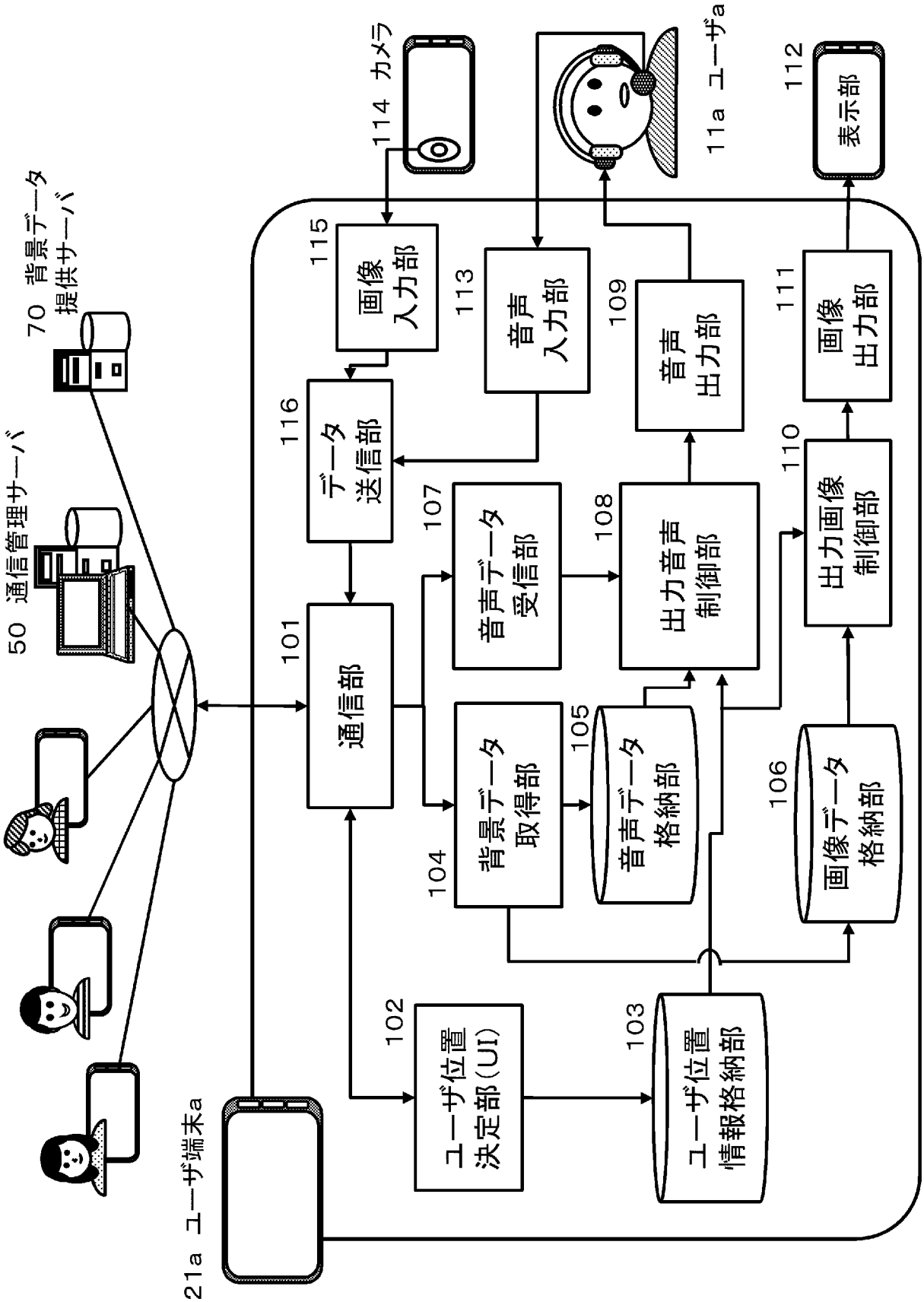
[図2]


70 背景データ提供サーバ
背景種類
背景音声データファイル
背景画像データファイル
(1)
会議室
(2)
カフェ
(3)
公園
(4)
ライブハウス
(5)
電車の中
(6)
駅
(7)
空港

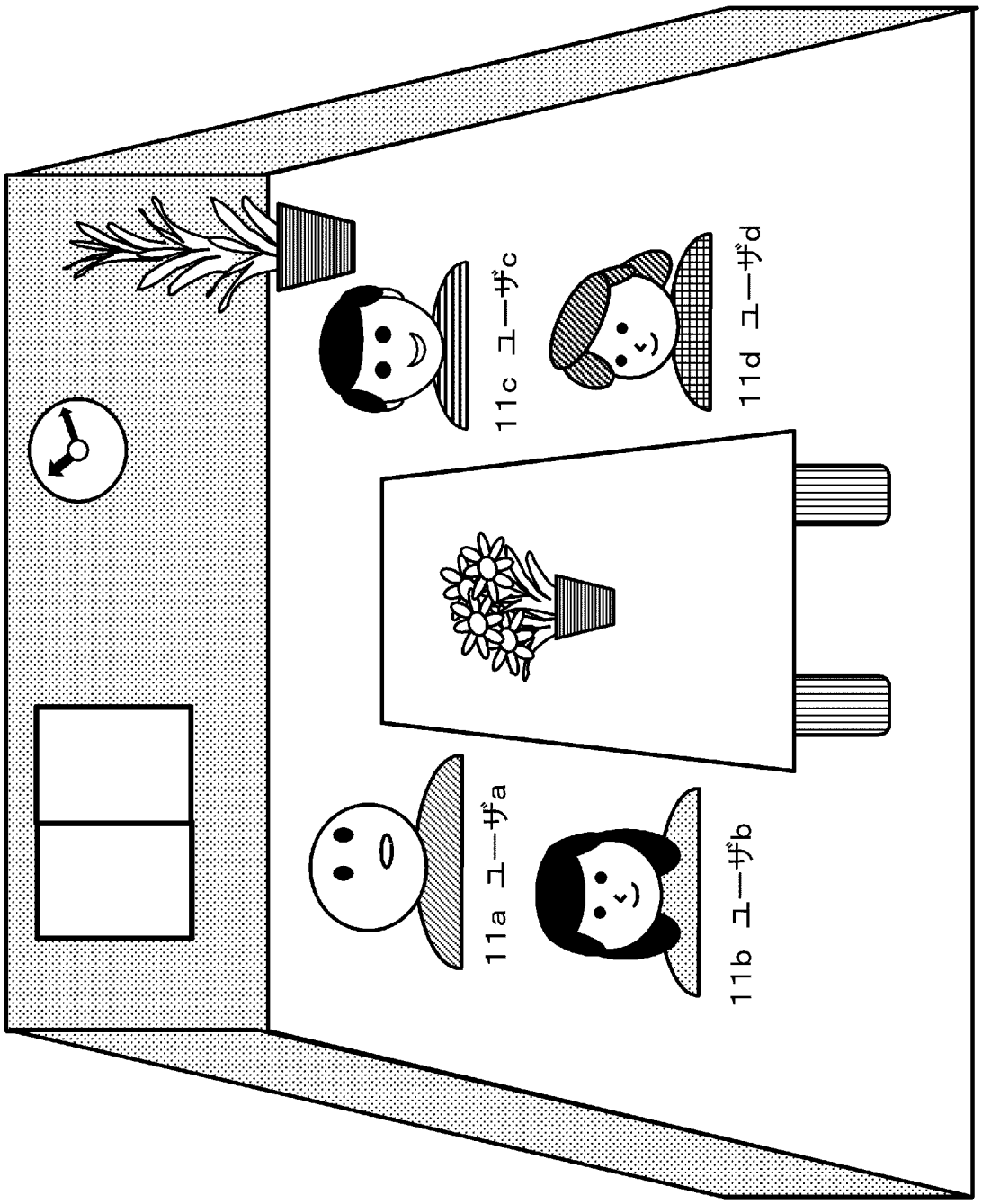
[図3]


70 背景データ
提供サーバ
カフェ背景データ
カフェ画像データ
カフェ
音声データ
音声データ
格納部
画像データ
格納部
21a ユーザ端末a
会議室背景データ
会議室
音声データ
会議室画像データ
11a ユーザa

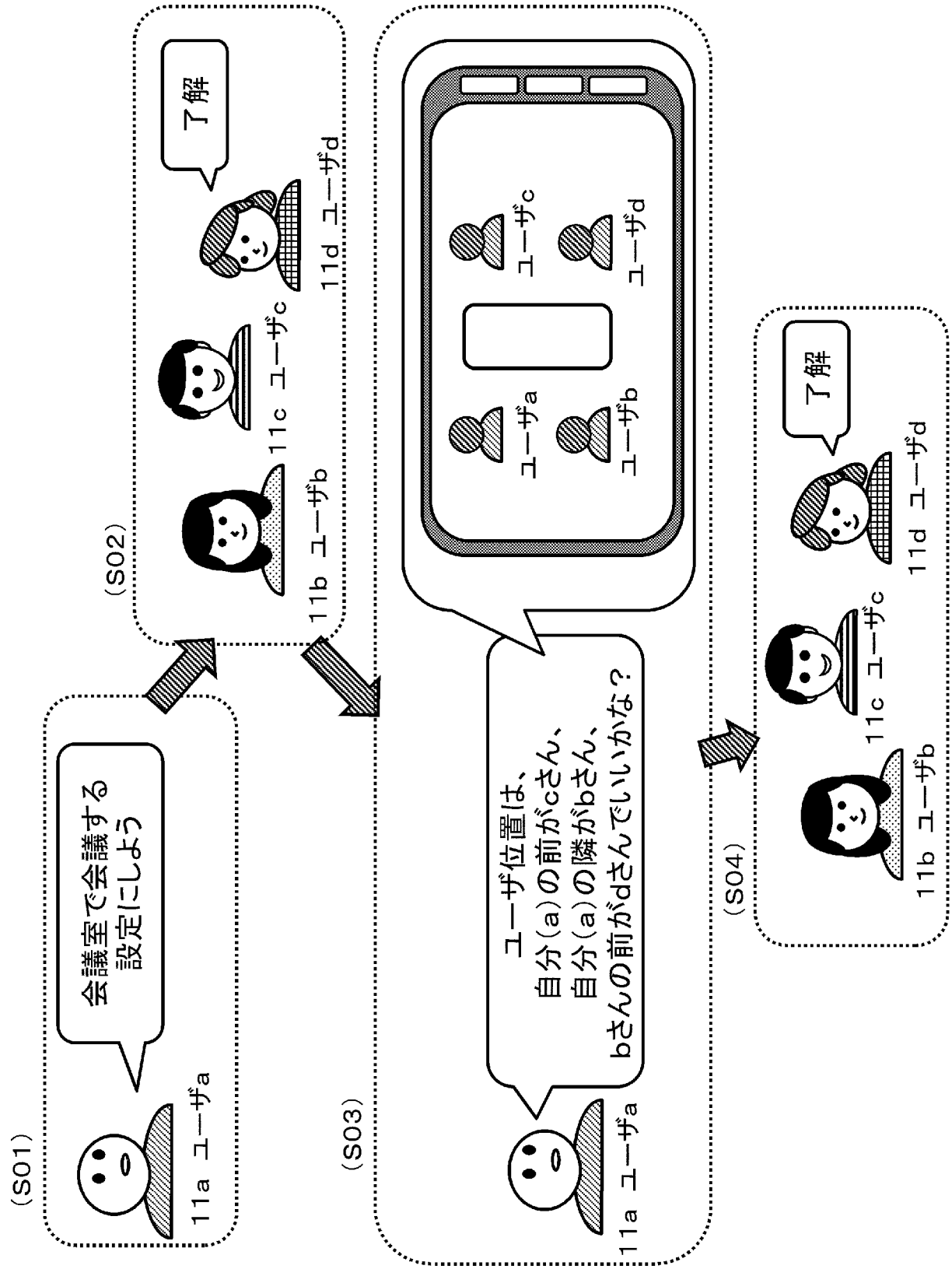
[図4]


70 背景データ提供サーバ
50 通信管理サーバ
21a ユーザ端末a
101 通信部
102 ユーザ位置決定部(UI)
103 ユーザ位置情報格納部
104 背景データ取得部
105 音声データ格納部
106 画像データ格納部
107 音声データ受信部
108 出力音声制御部
109 音声出力部
110 出力画像制御部
111 画像出力部
112 表示部
113 音声入力部
114 カメラ
115 画像入力部
116 データ送信部
11a ユーザa

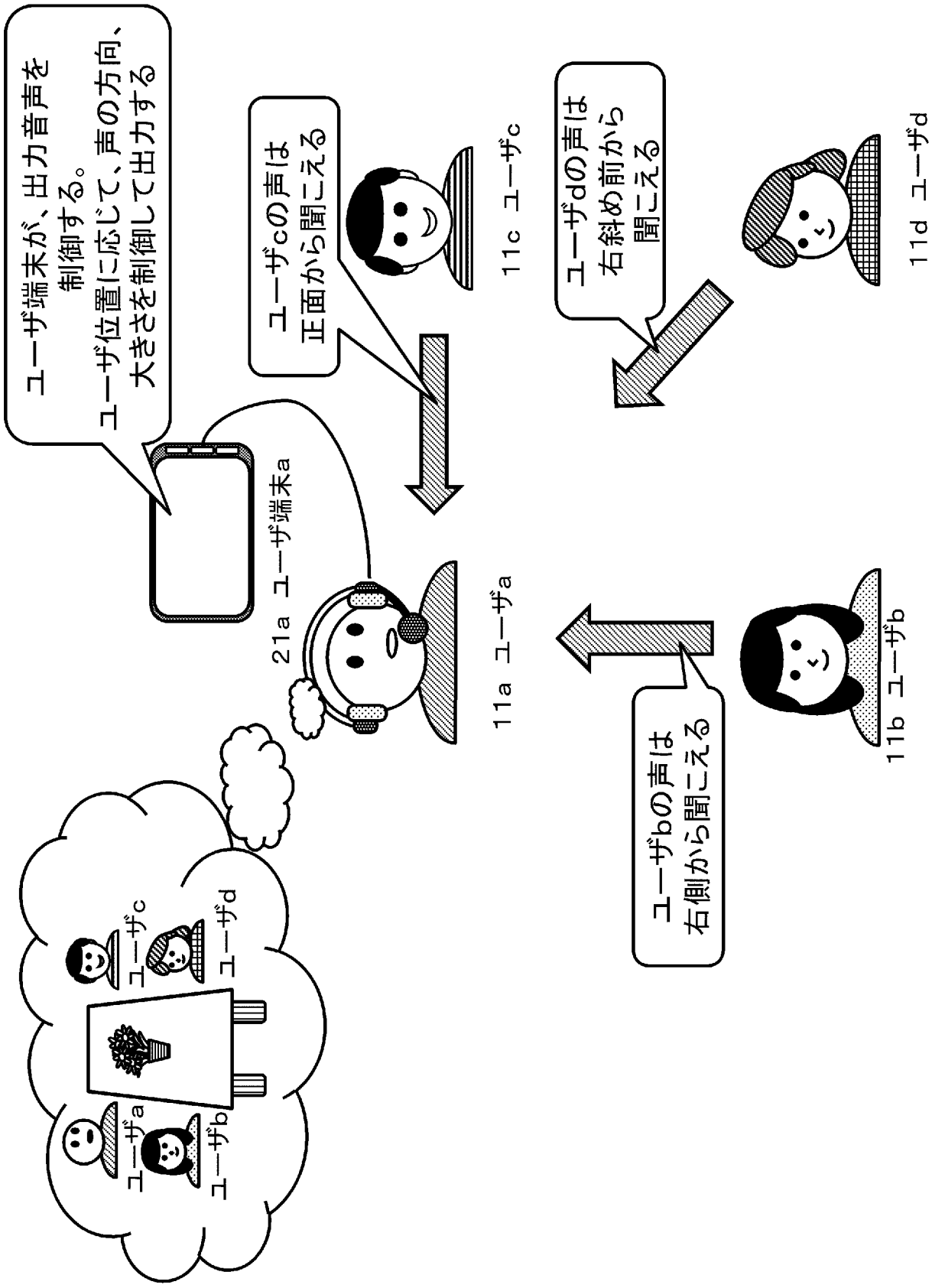
[図5]



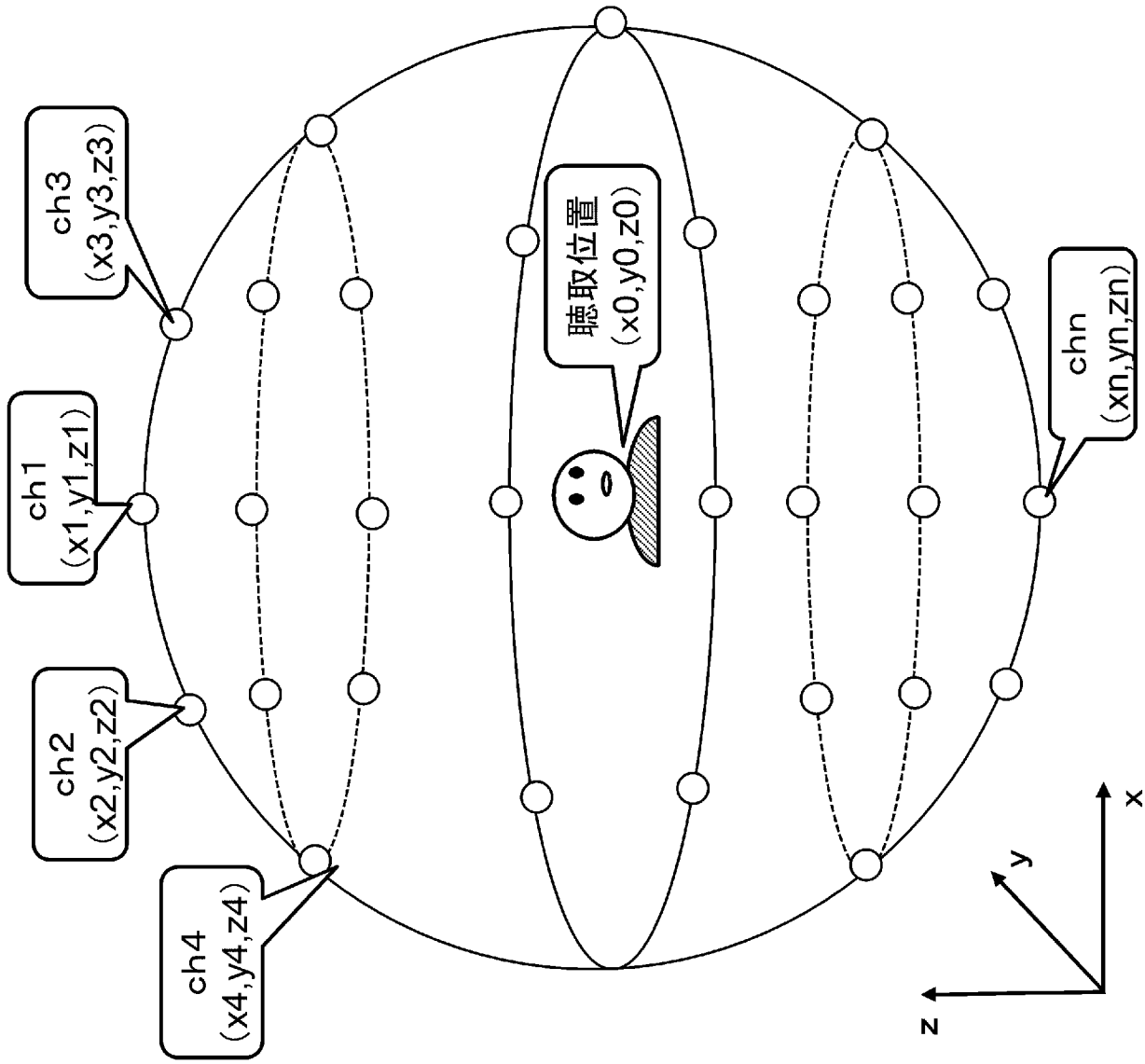
[図6]


(S01)
会議室で会議する
設定にしよう
11a ユーザa
(S02)
了解
11b ユーザb
11c ユーザc
11d ユーザd
(S03)
ユーザ位置は、
自分(a)の前がcさん、
自分(a)の隣がbさん、
bさんの前がdさんでいいかな？
11a ユーザa
ユーザa
ユーザb
ユーザc
ユーザd
(S04)
了解
11b ユーザb
11c ユーザc
11d ユーザd

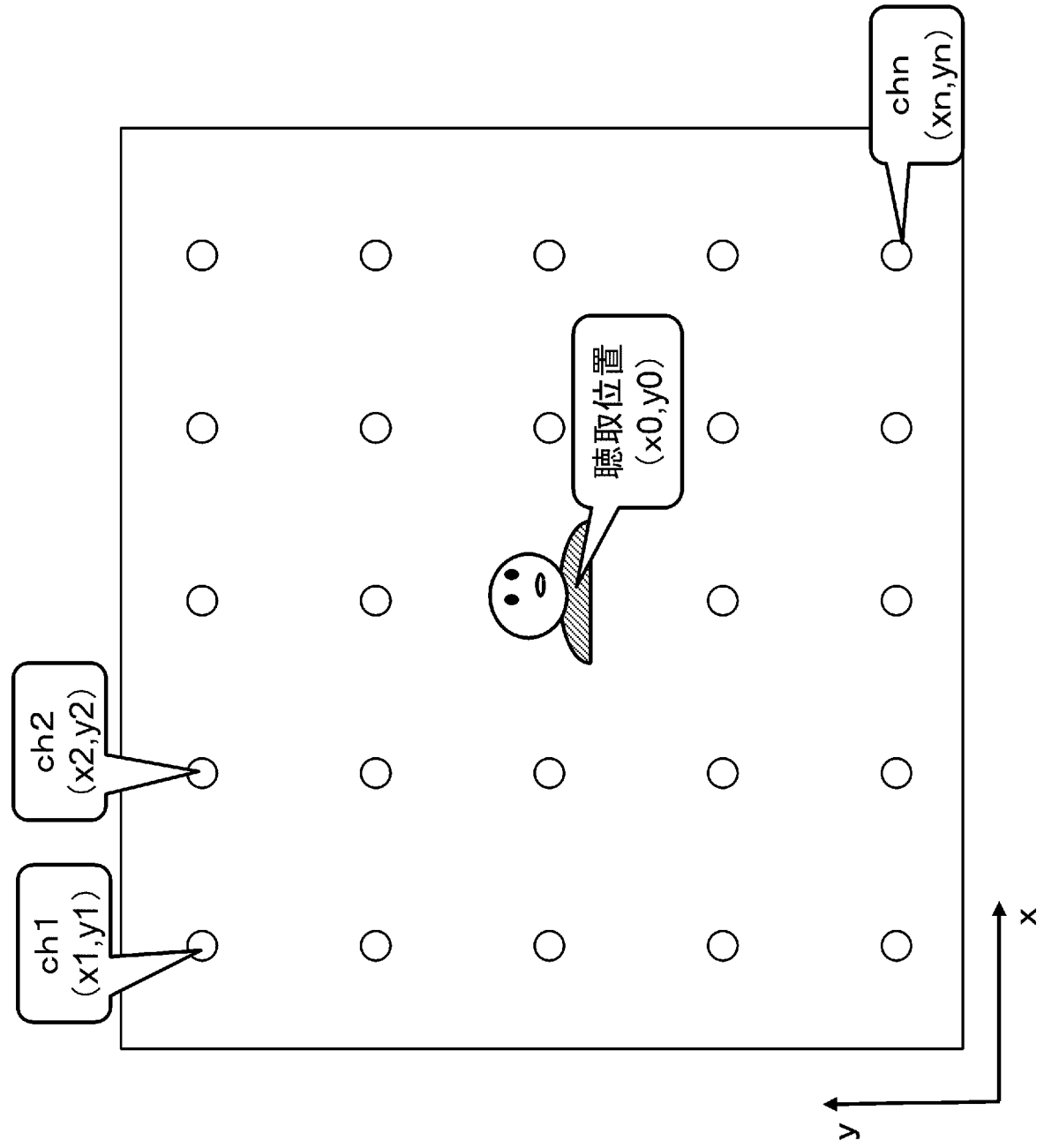
[図7]


ユーザ端末が、出力音声を制御する。
ユーザ位置に応じて、声の方向、大きさを制御して出力する
ユーザcの声は正面から聞こえる
ユーザdの声は右斜め前から聞こえる
ユーザbの声は右側から聞こえる
21a ユーザ端末a
11a ユーザa
11b ユーザb
11c ユーザc
11d ユーザd
ユーザa
ユーザb
ユーザc
ユーザd

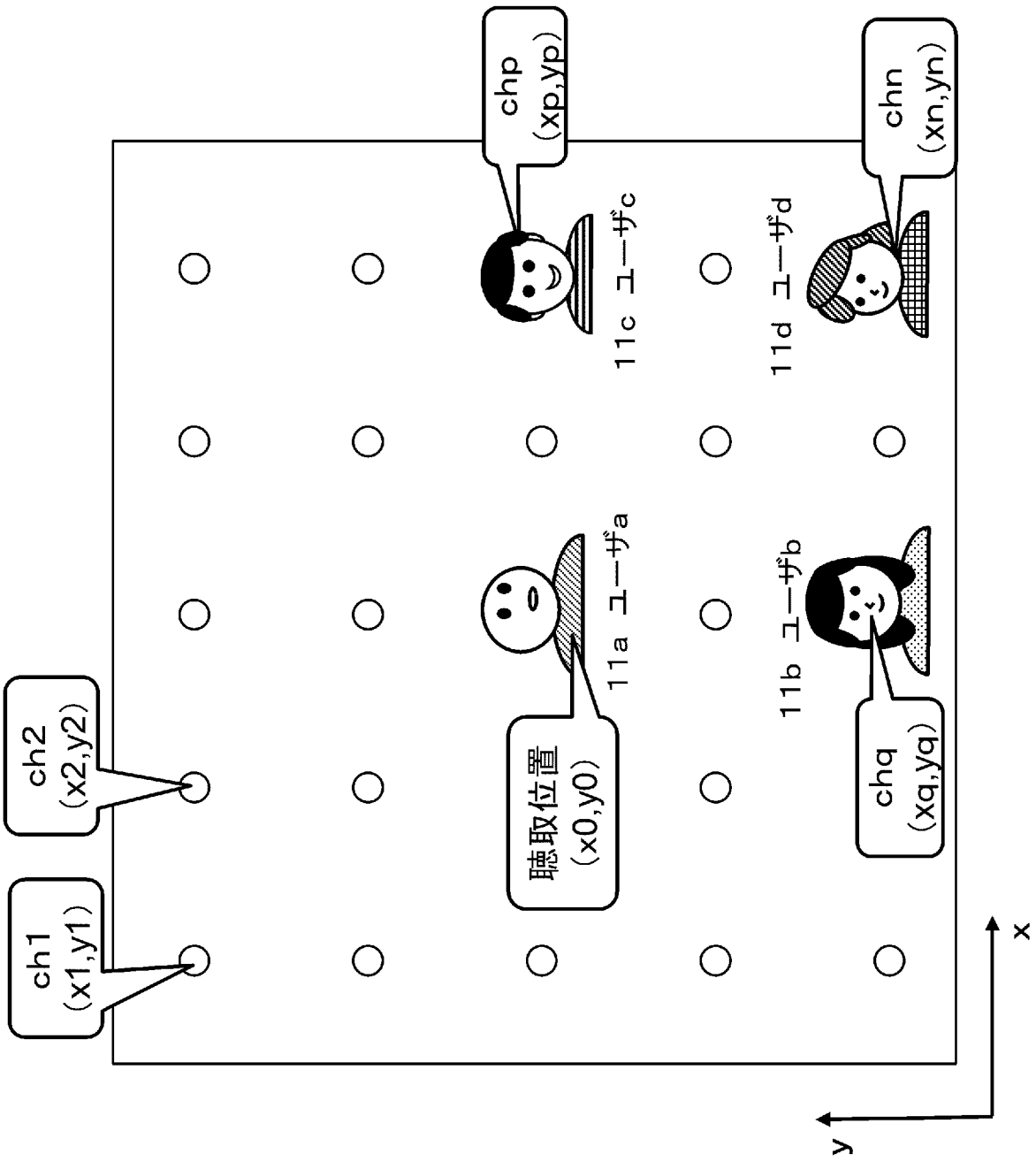
[図8]


ch3
(x3,y3,z3)
ch1
(x1,y1,z1)
ch2
(x2,y2,z2)
ch4
(x4,y4,z4)
聴取位置
(x0,y0,z0)
chn
(xn,yn,zn)
x
y
z

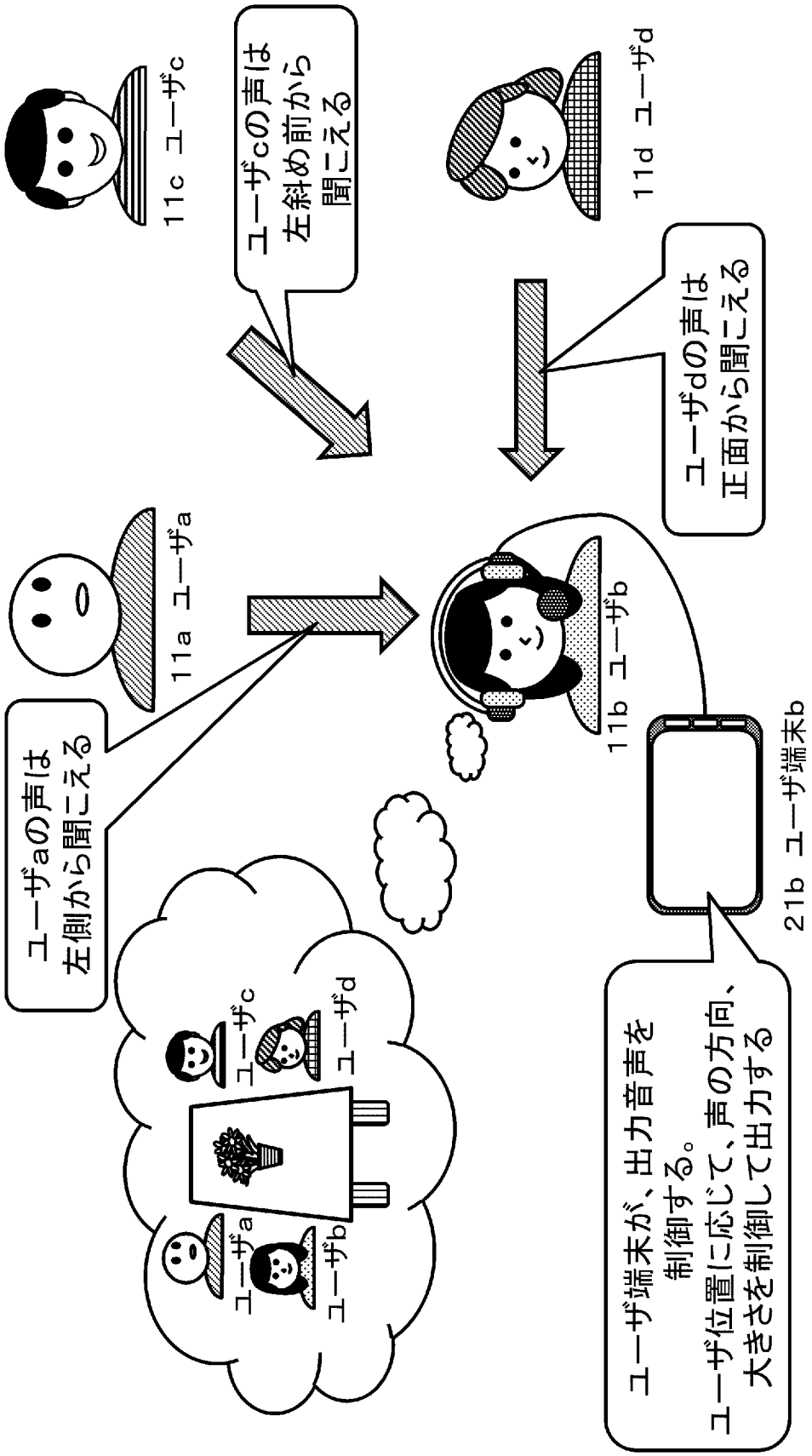
[図9]



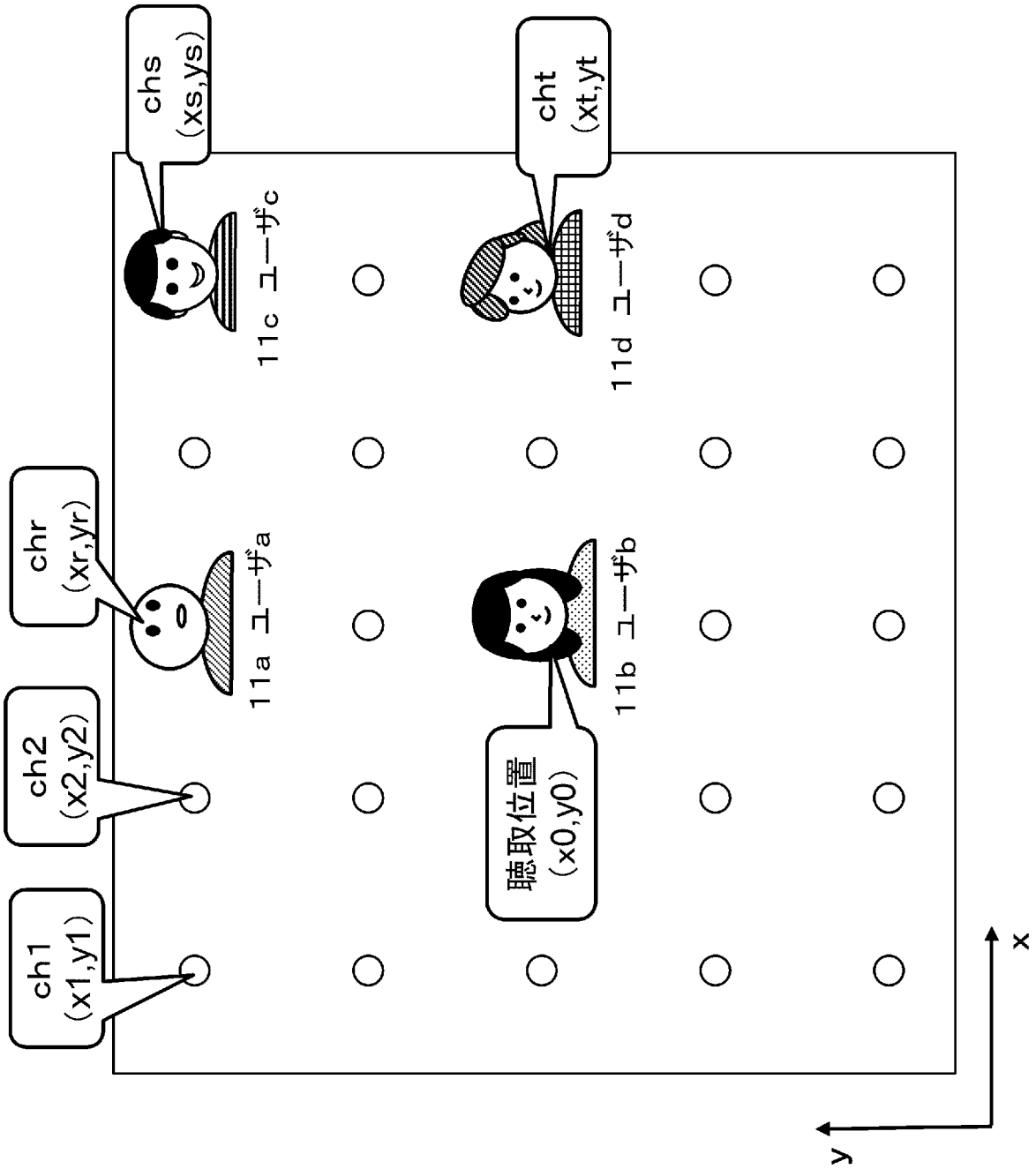
[図10]


ch1
(x1,y1)
ch2
(x2,y2)
聴取位置
(x0,y0)
11a ユーザa
11b ユーザb
chq
(xq,yq)
chp
(xp,yp)
11c ユーザc
11d ユーザd
chn
(xn,yn)
x
y

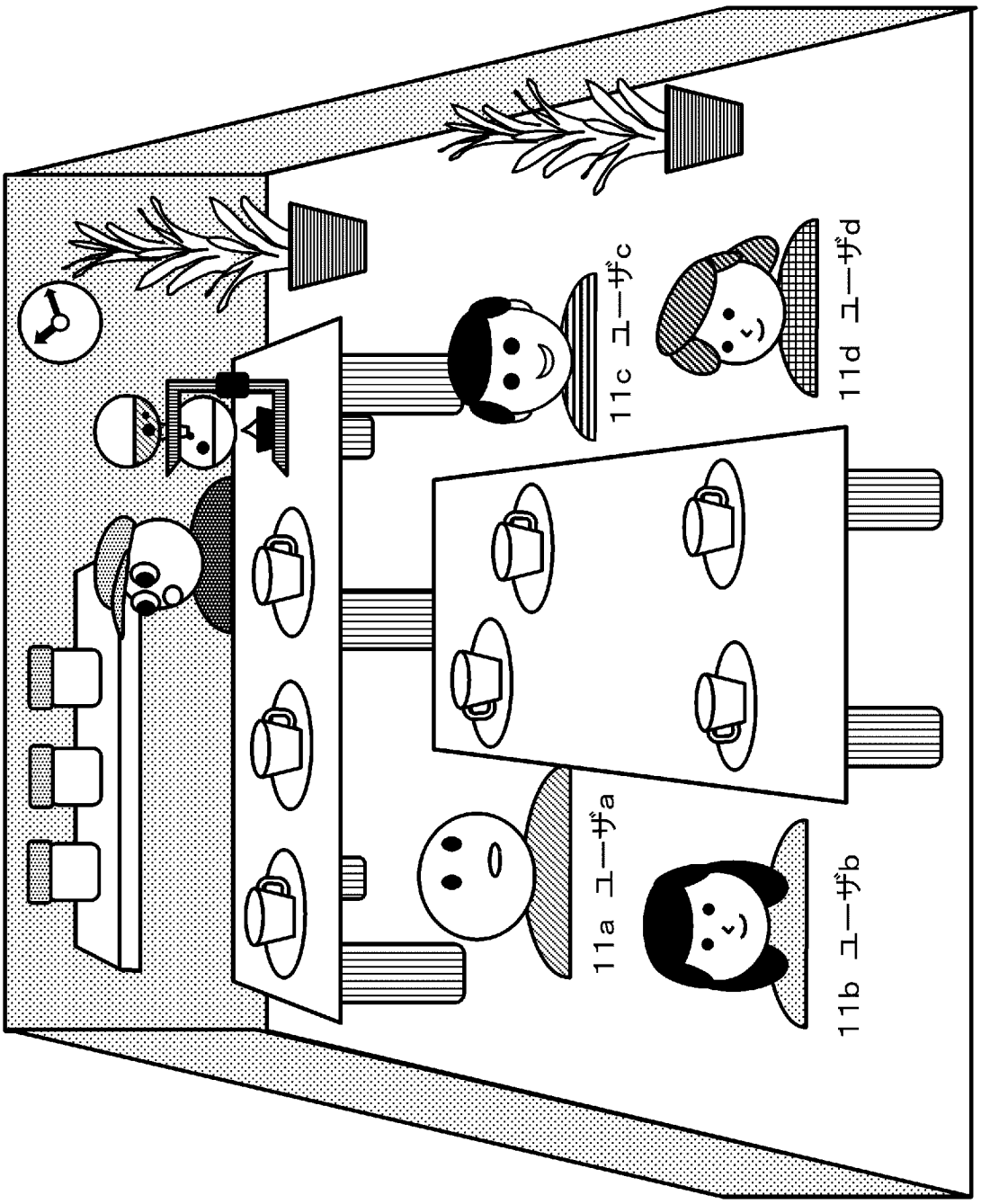
[図11]


11c ユーザc
ユーザcの声は
左斜め前から
聞こえる
11d ユーザd
ユーザdの声は
正面から聞こえる
11a ユーザa
ユーザaの声は
左側から聞こえる
11b ユーザb
ユーザa
ユーザb
ユーザc
ユーザd
21b ユーザ端末b
ユーザ端末が、出力音声を
制御する。
ユーザ位置に応じて、声の方向、
大きさを制御して出力する

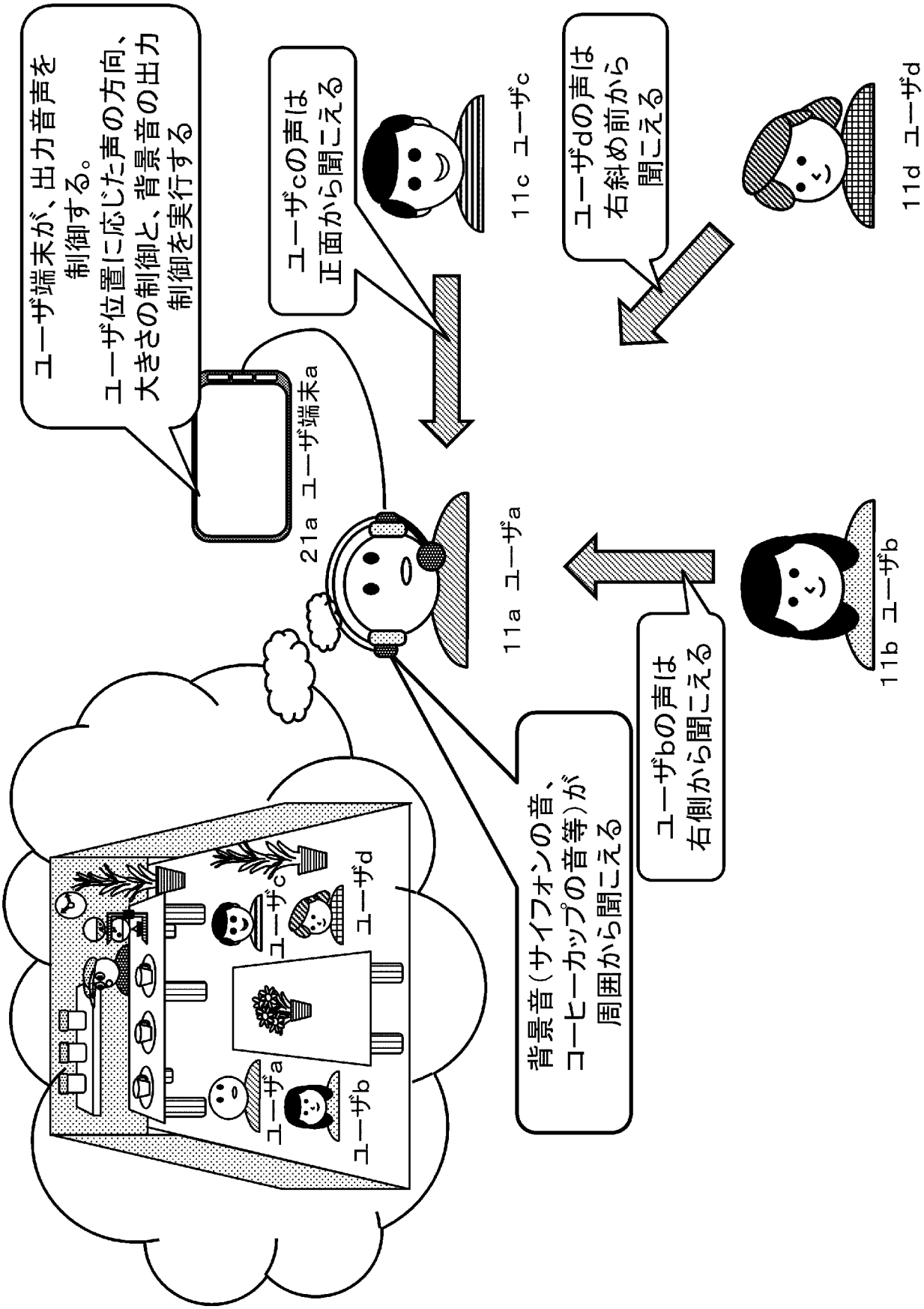
[図12]


chs
(xs,ys)
cht
(xt,yt
chr
(xr,yr)
ch2
(x2,y2)
ch1
(x1,y1)
聴取位置
(x0,y0)
11a ユーザa
11b ユーザb
11c ユーザc
11d ユーザd
x
y

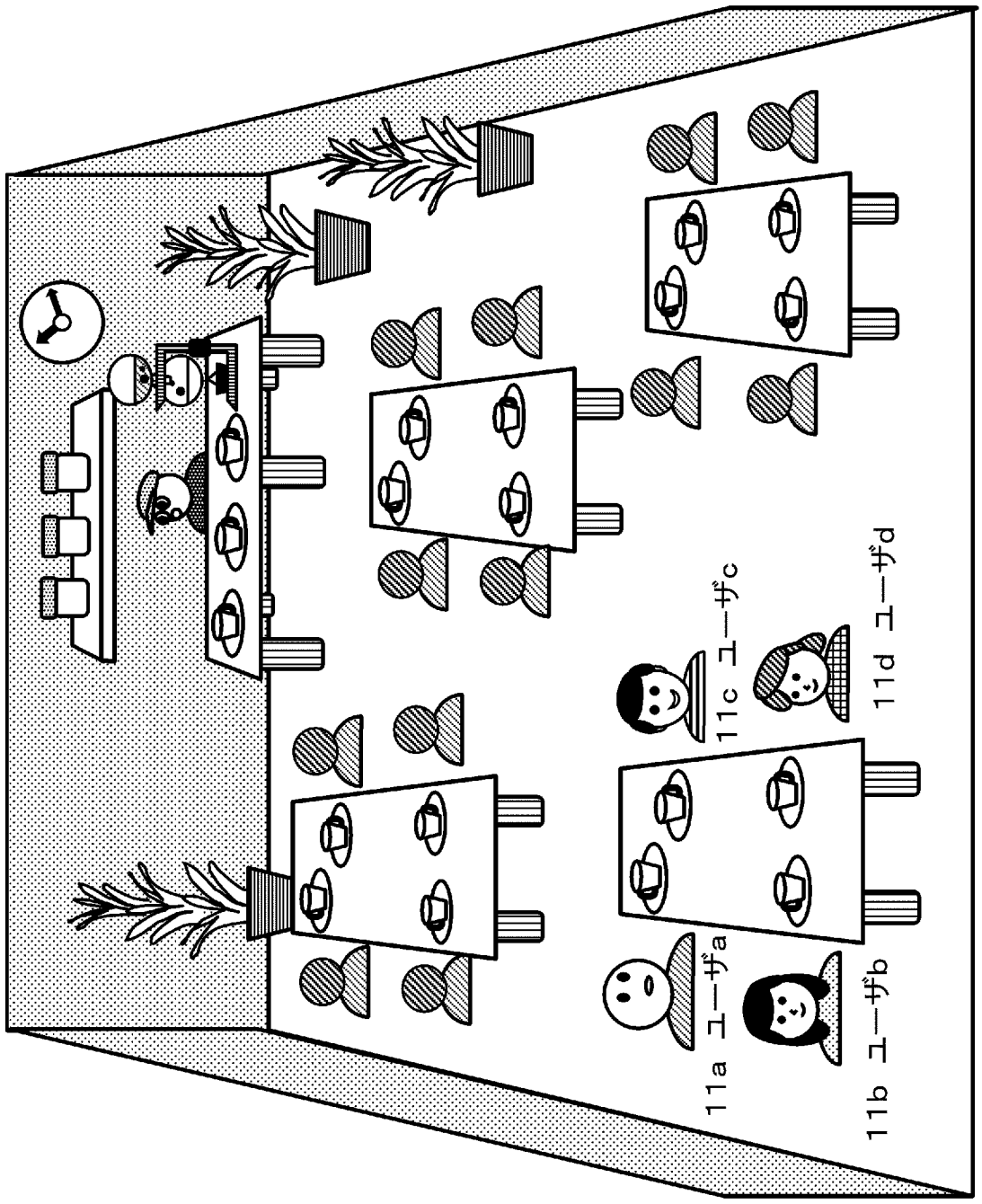
[図13]



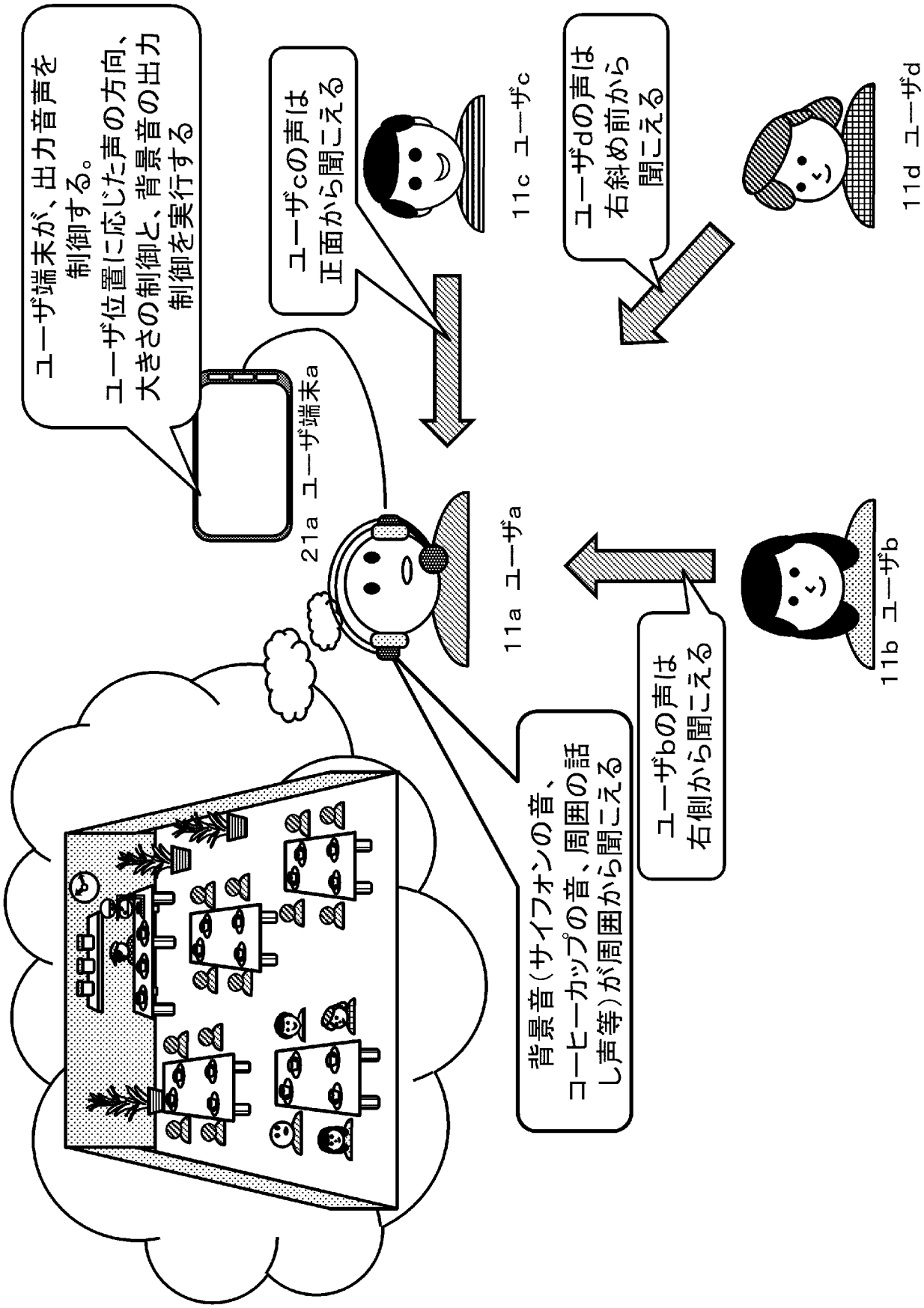
[図14]


ユーザ端末が、出力音声を
制御する。
ユーザ位置に応じた声の方向、
大きさの制御と、背景音の出力
制御を実行する
21a ユーザ端末a
ユーザcの声は
正面から聞こえる
11c ユーザc
ユーザdの声は
右斜め前から
聞こえる
11d ユーザd
11a ユーザa
背景音（サイフォンの音、
コーヒーカップの音等）が
周囲から聞こえる
ユーザbの声は
右側から聞こえる
11b ユーザb
ユーザa
ユーザb
ユーザc
ユーザd

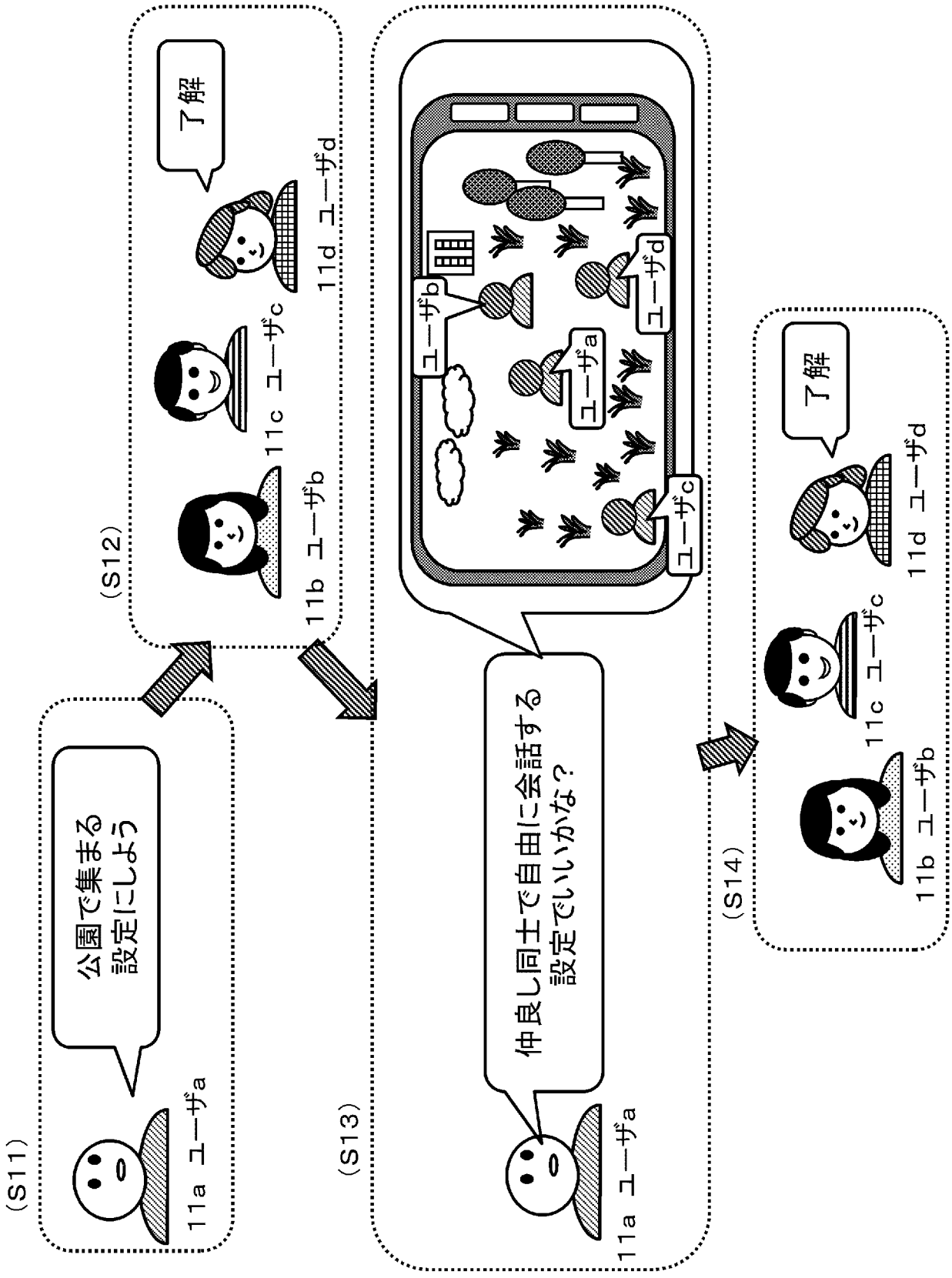
[図15]



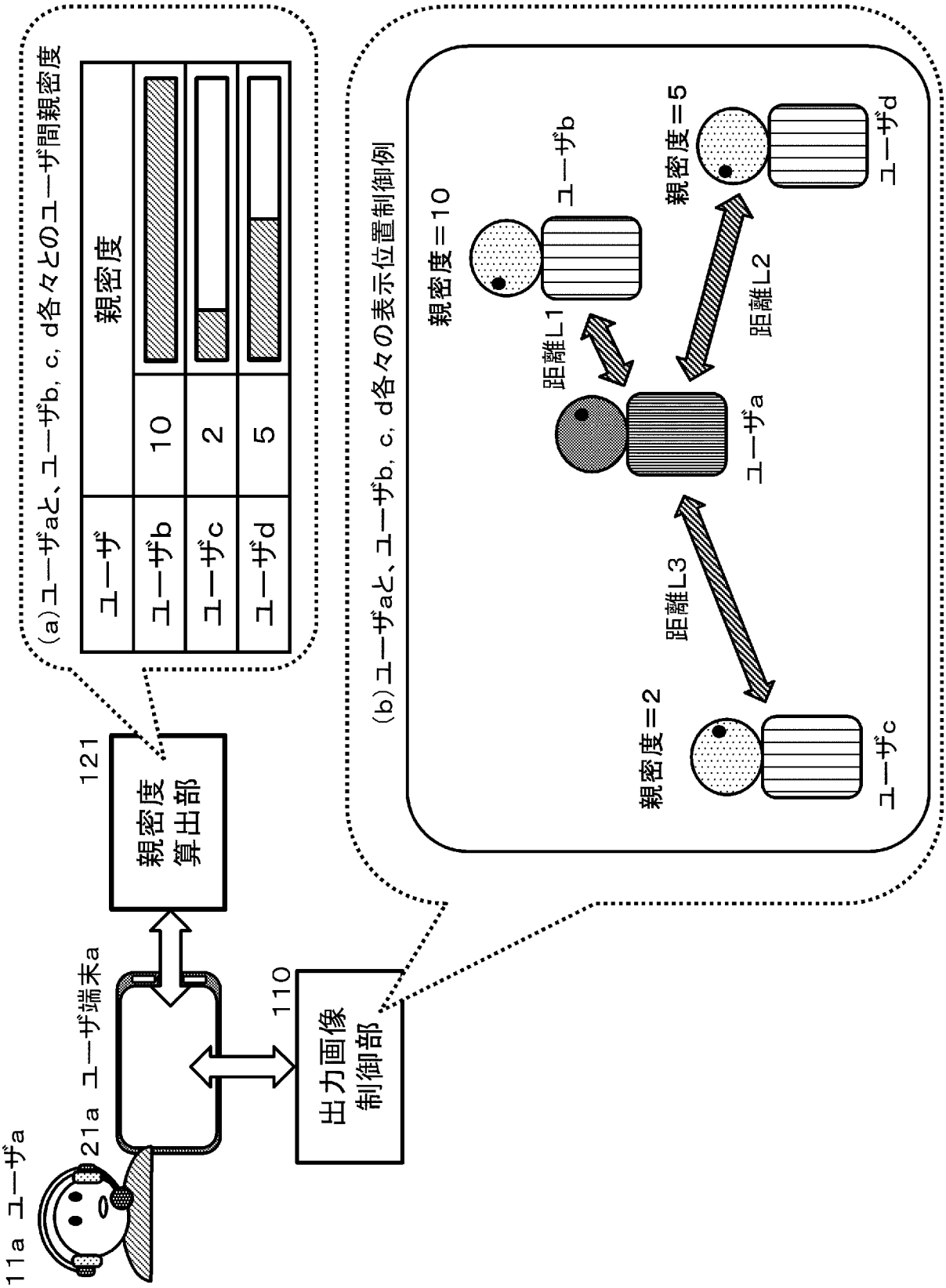
[図16]


ユーザ端末が、出力音声を制御する。
ユーザ位置に応じた声の方向、大きさの制御と、背景音の出力制御を実行する
21a ユーザ端末a
ユーザcの声は正面から聞こえる
11c ユーザc
ユーザdの声は右斜め前から聞こえる
11d ユーザd
11a ユーザa
背景音（サイフォンの音、コーヒーカップの音、周囲の話し声等）が周囲から聞こえる
ユーザbの声は右側から聞こえる
11b ユーザb

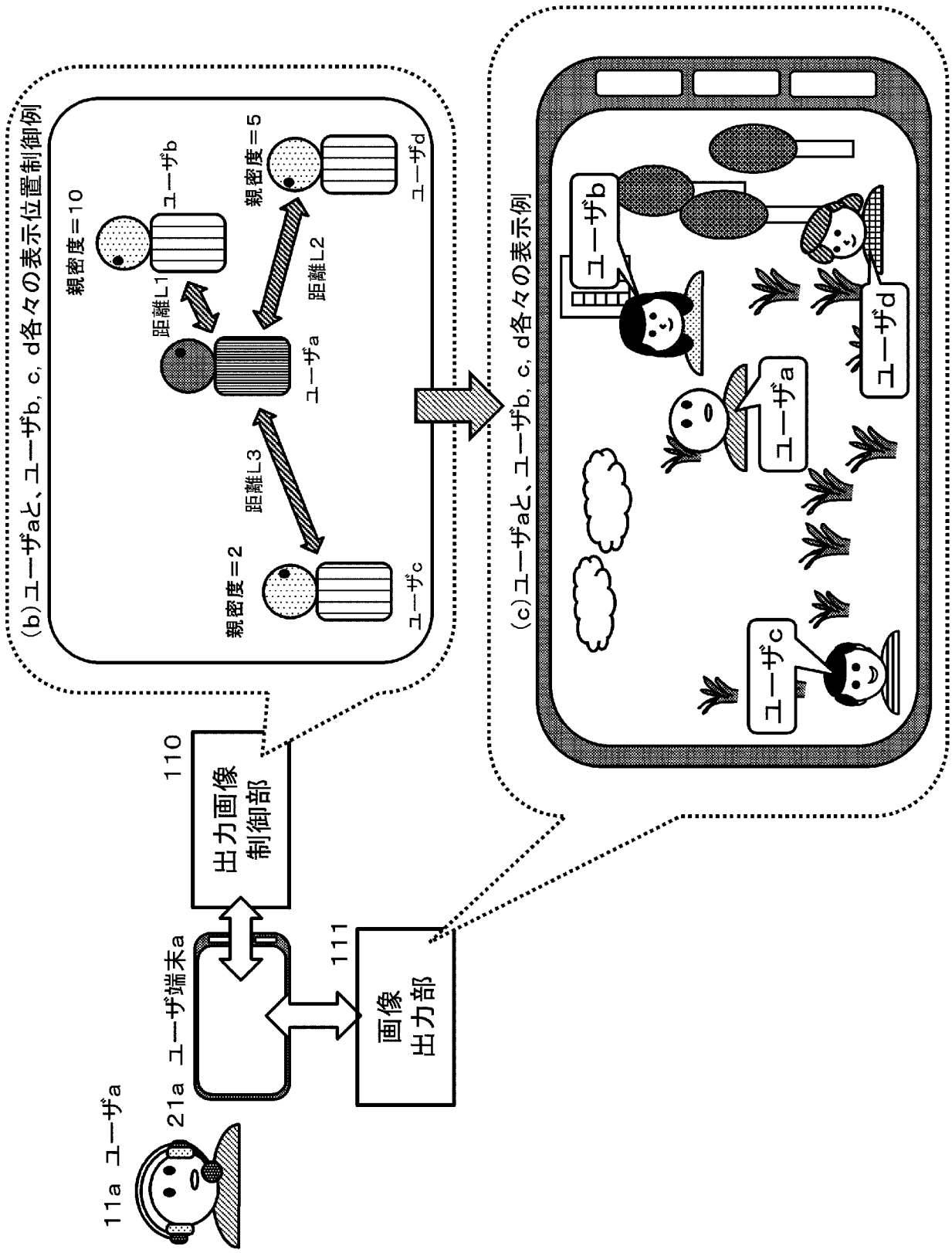
[図17]


(S11)
公園で集まる
設定にしよう
11a ユーザa
(S12)
了解
11b ユーザb
11c ユーザc
11d ユーザd
(S13)
ユーザb
ユーザa
ユーザd
ユーザc
仲良し同士で自由に会話する
設定でいいかな？
11a ユーザa
(S14)
了解
11c ユーザc
11d ユーザd
11b ユーザb

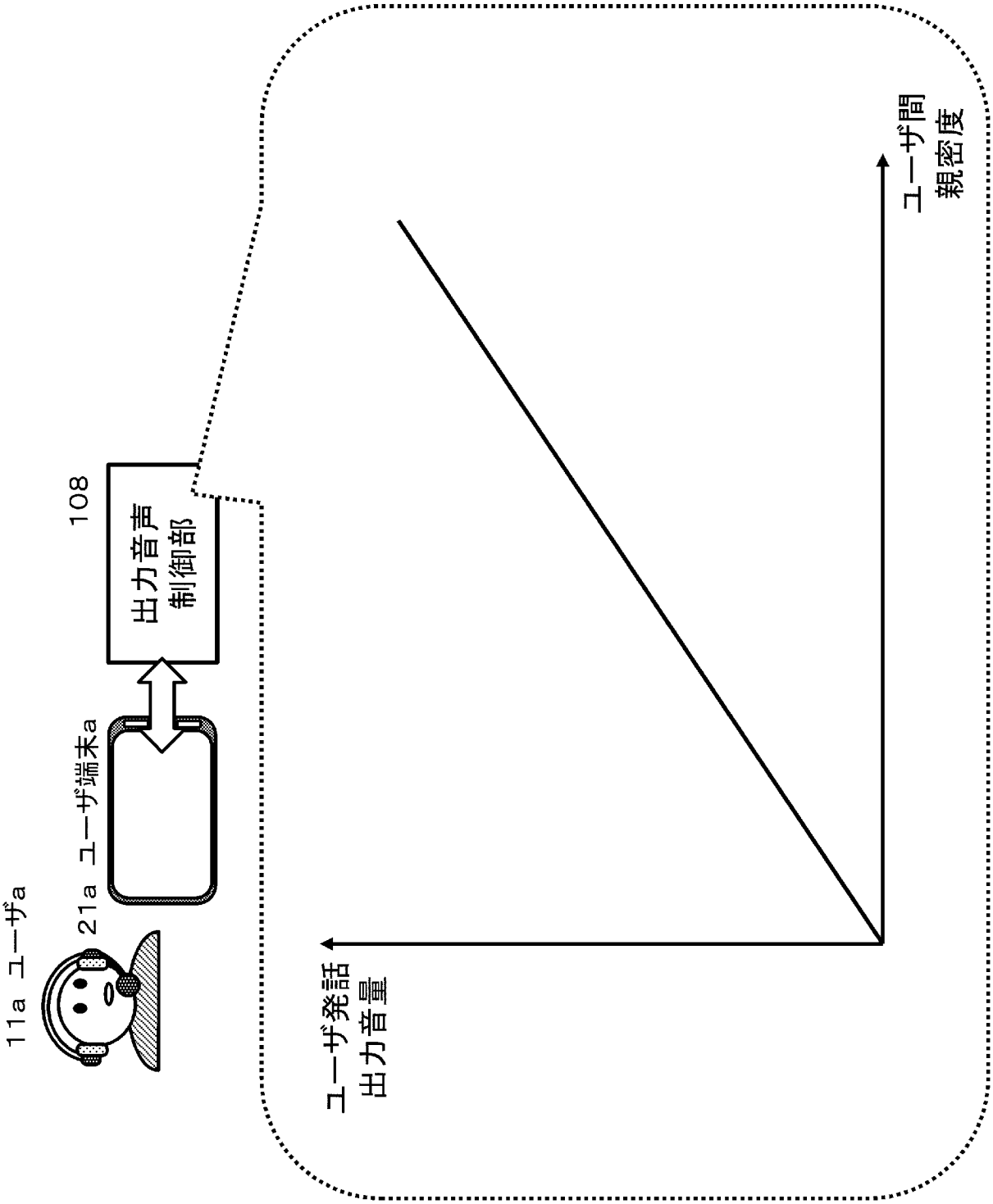
[図18]


11a ユーザa
21a ユーザ端末a
121
親密度算出部
110
出力画像制御部
(a)ユーザaと、ユーザb, c, d各々とのユーザ間親密度
ユーザ
親密度
ユーザb
10
ユーザc
2
ユーザd
5
(b)ユーザaと、ユーザb, c, d各々の表示位置制御例
親密度＝10
ユーザb
距離L1
親密度＝5
ユーザd
距離L2
ユーザa
距離L3
親密度＝2
ユーザc

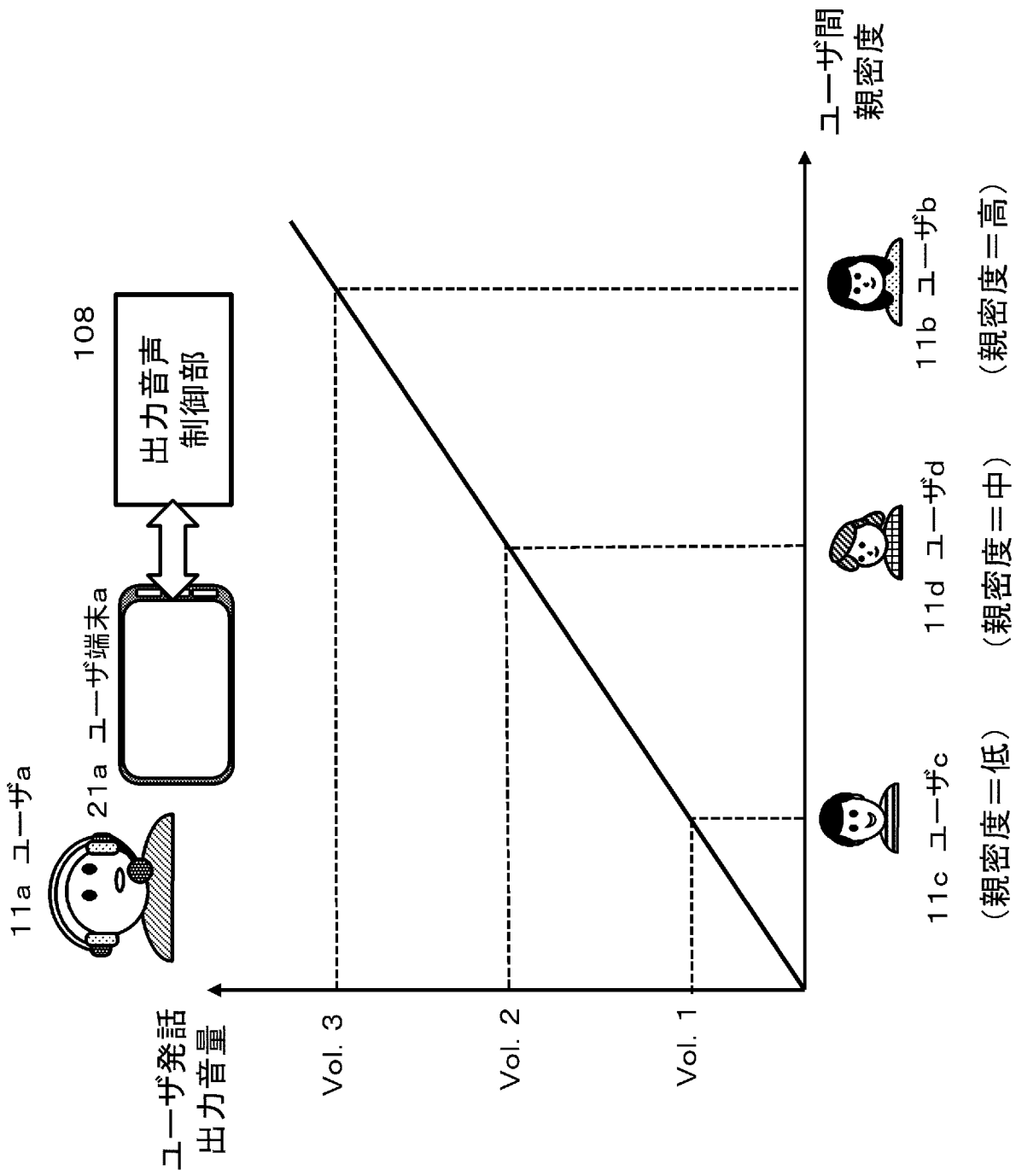
[図19]


11a ユーザa
21a ユーザ端末a
110
出力画像
制御部
111
画像
出力部
(b)ユーザaと、ユーザb, c, d各々の表示位置制御例
親密度＝10
ユーザb
距離L1
ユーザa
距離L2
親密度＝5
ユーザd
距離L3
親密度＝2
ユーザc
(c)ユーザaと、ユーザb, c, d各々の表示例
ユーザb
ユーザa
ユーザd
ユーザc

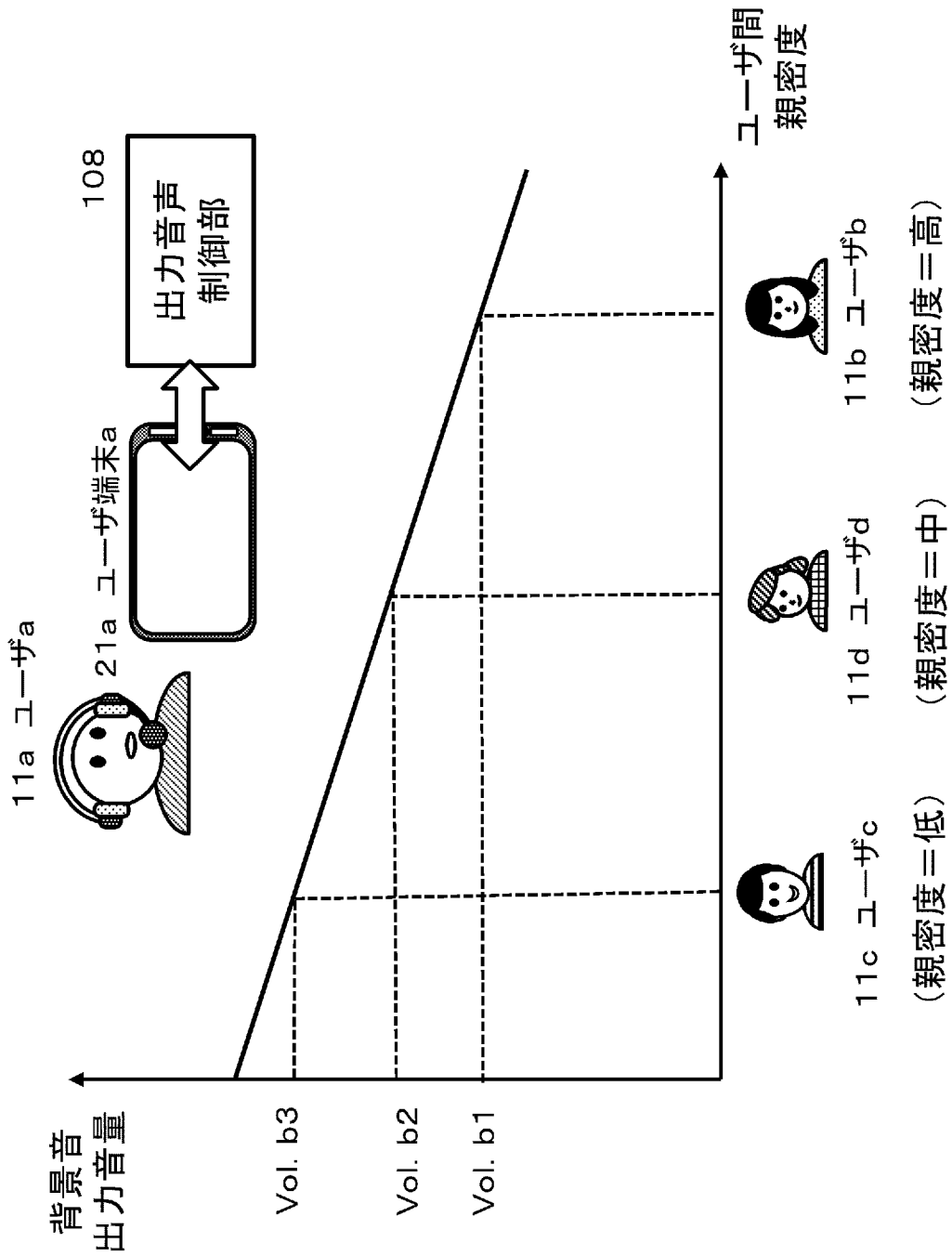
[図20]



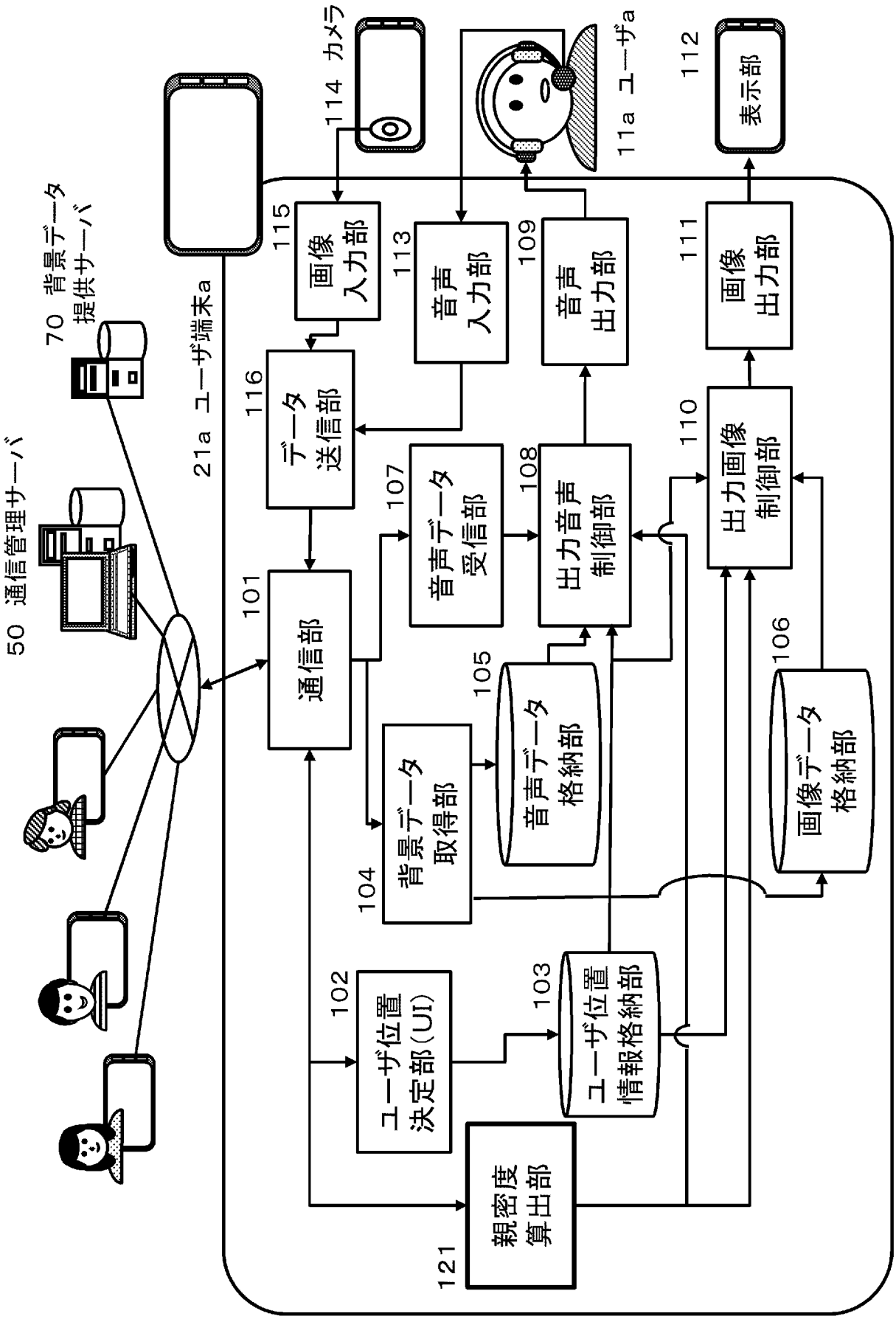
[図21]


11a ユーザa
21a ユーザ端末a
108
出力音声
制御部
ユーザ発話
出力音量
Vol. 3
Vol. 2
Vol. 1
ユーザ間
親密度
11c ユーザc
（親密度＝低）
11d ユーザd
（親密度＝中）
11b ユーザb
（親密度＝高）

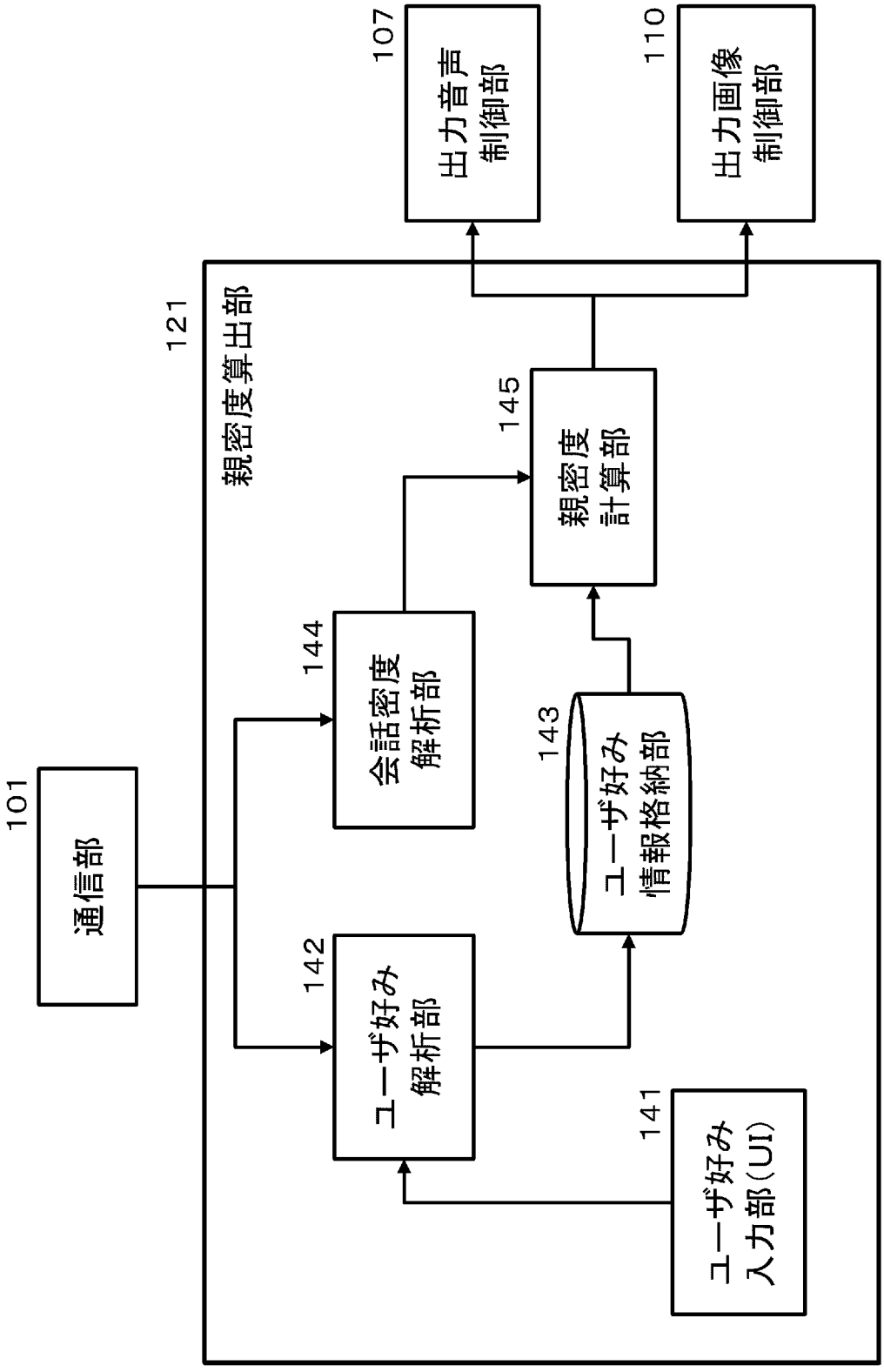
[図22]


108
出力音声
制御部
21a ユーザ端末a
11a ユーザa
背景音
出力音量
Vol. b3
Vol. b2
Vol. b1
ユーザ間
親密度
11c ユーザc
（親密度＝低）
11d ユーザd
（親密度＝中）
11b ユーザb
（親密度＝高）

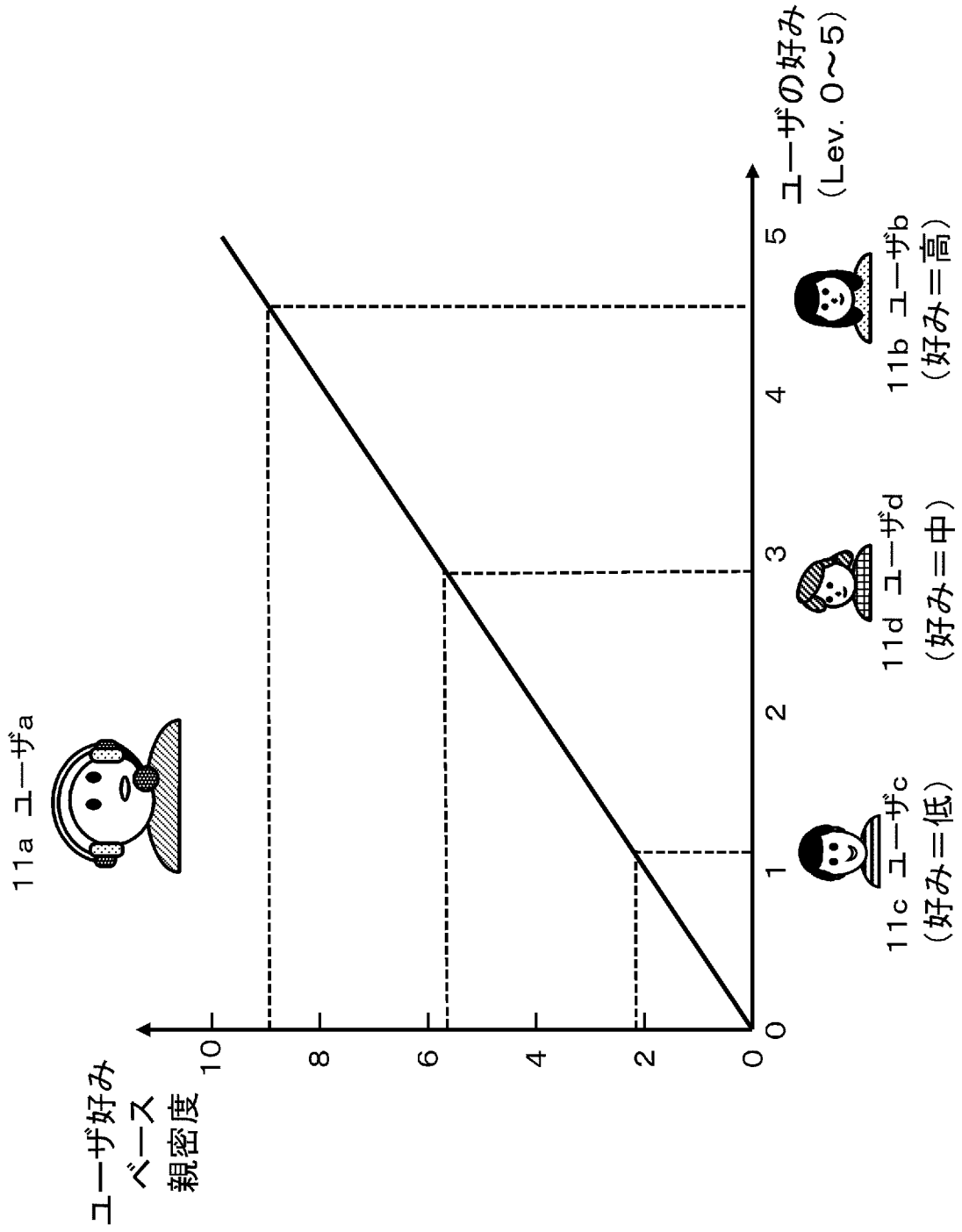
[図23]


50 通信管理サーバ
70 背景データ提供サーバ
21a ユーザ端末a
11a ユーザa
114 カメラ
112 表示部
115 画像入力部
113 音声入力部
109 音声出力部
111 画像出力部
116 データ送信部
107 音声データ受信部
108 出力音声制御部
110 出力画像制御部
101 通信部
105 音声データ格納部
106 画像データ格納部
104 背景データ取得部
102 ユーザ位置決定部(UI)
103 ユーザ位置情報格納部
121 親密度算出部

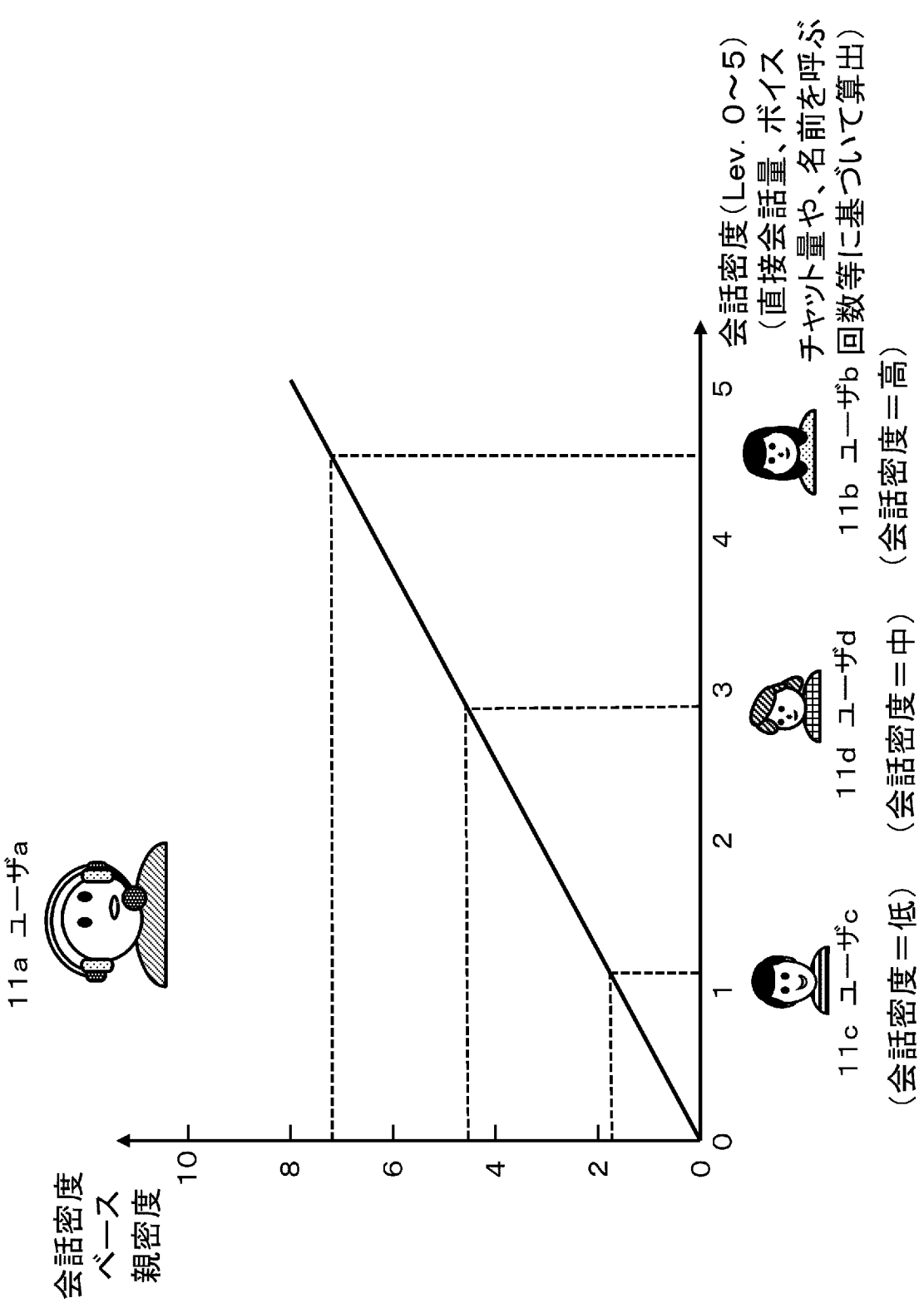
[図24]


107
出力音声
制御部
110
出力画像
制御部
121
親密度算出部
145
親密度
計算部
144
会話密度
解析部
143
ユーザ好み
情報格納部
101
通信部
142
ユーザ好み
解析部
141
ユーザ好み
入力部(UI)

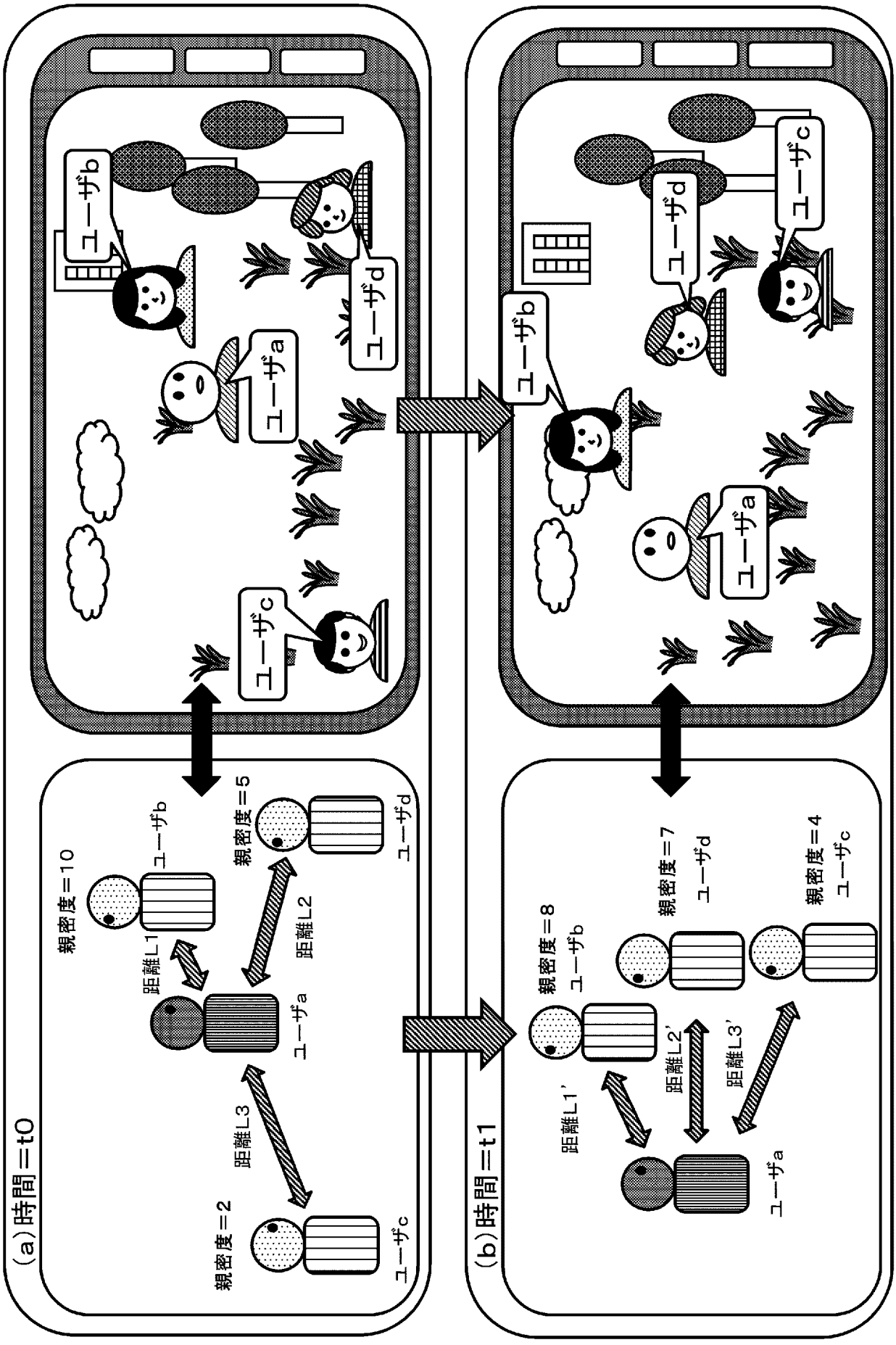
[図25]


ユーザ好み
ベース
親密度
10
8
6
4
2
0
0
1
2
3
4
5
ユーザの好み
(Lev. 0～5)
11a ユーザa
11b ユーザb
(好み＝高)
11c ユーザc
(好み＝低)
11d ユーザd
(好み＝中)

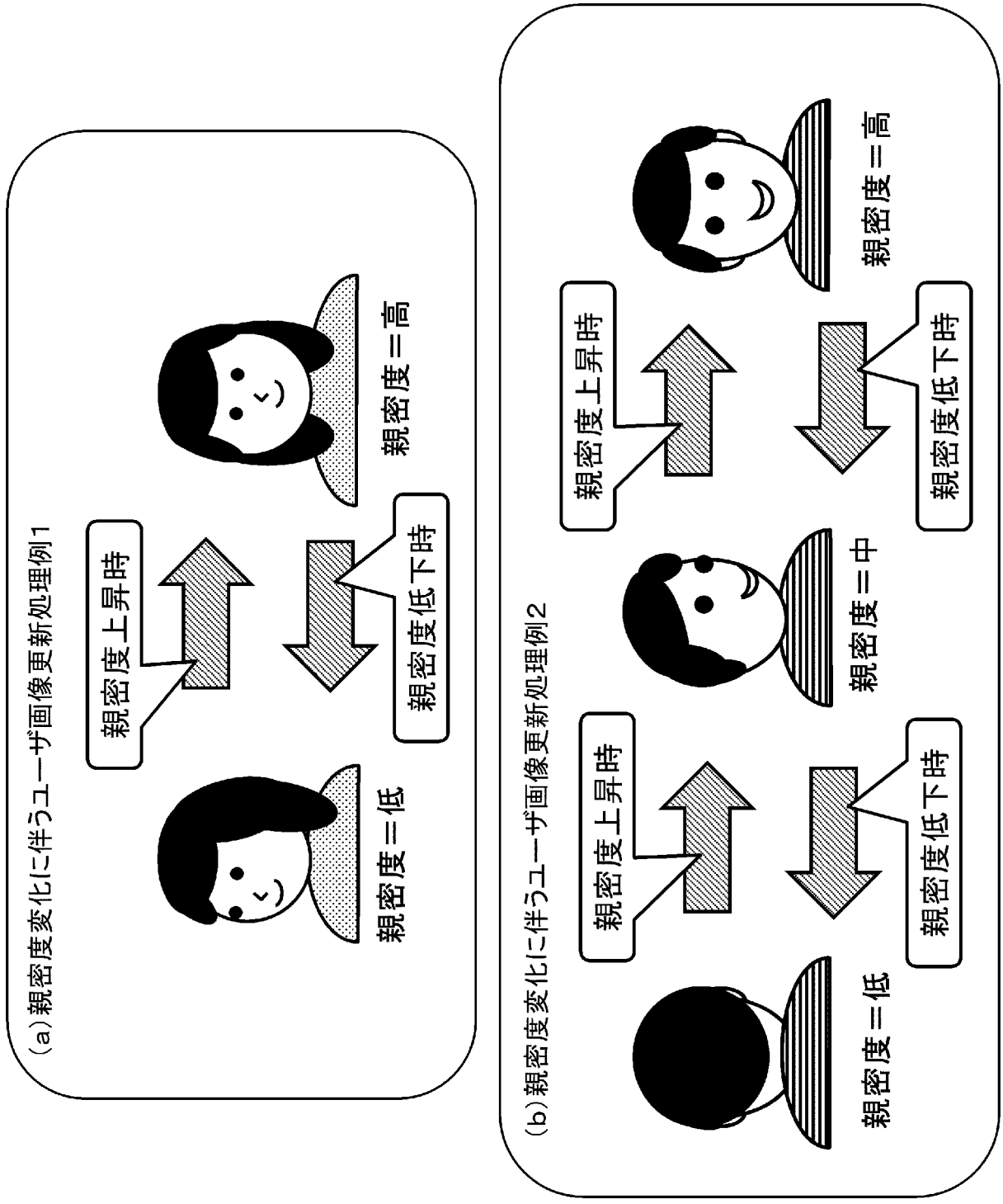
[図26]


会話密度
ベース
親密度
10
8
6
4
2
0
0
1
2
3
4
5
会話密度(Lev. 0～5)
(直接会話量、ボイス
チャット量や、名前を呼ぶ
回数等に基づいて算出)
11a ユーザa
11b ユーザb
11c ユーザc
11d ユーザd
(会話密度＝低)
(会話密度＝中)
(会話密度＝高)

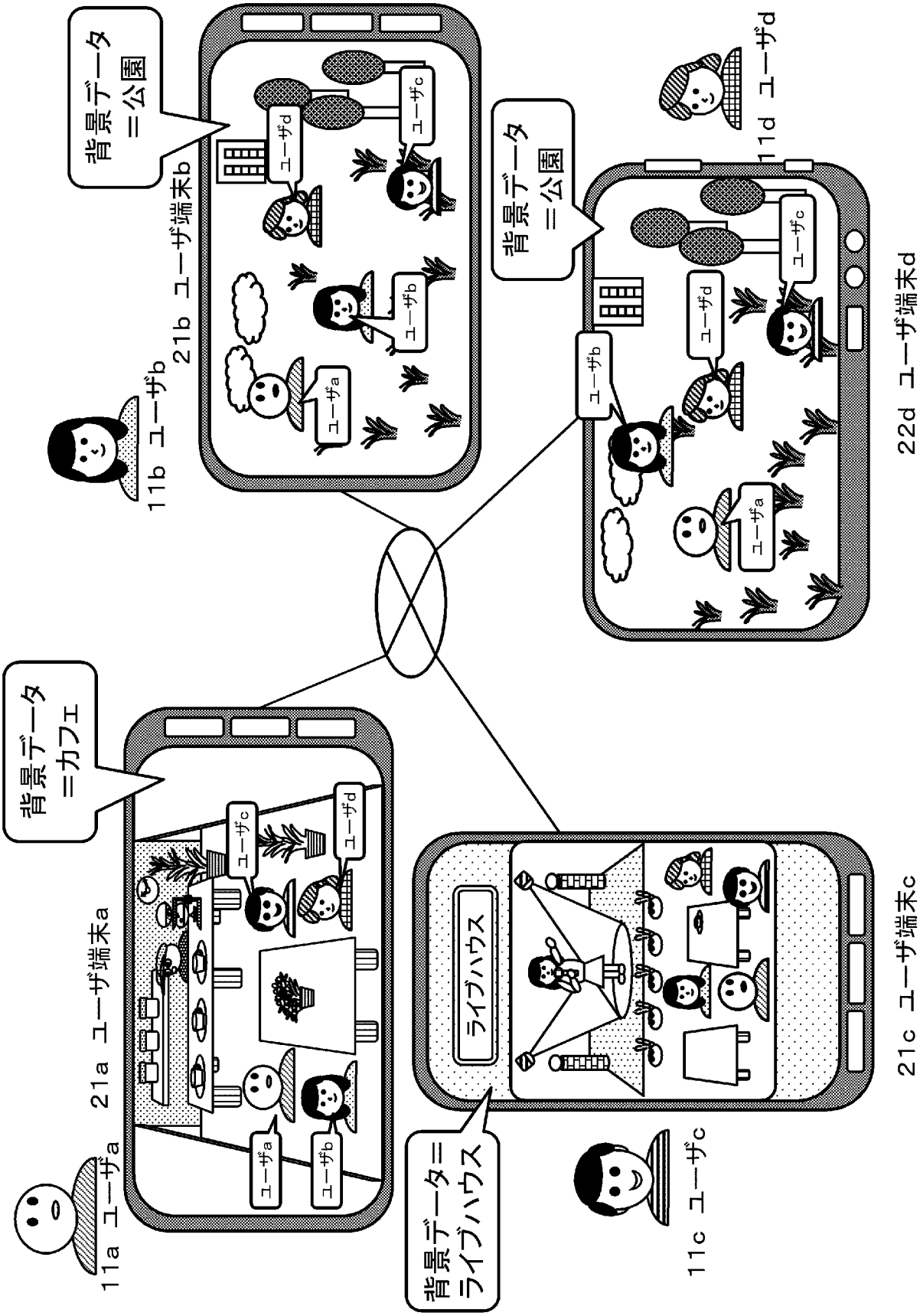
[図27]


(a)時間＝t0
ユーザb
ユーザa
ユーザd
ユーザc
親密度＝10
ユーザb
距離L1
ユーザa
親密度＝5
距離L2
ユーザd
距離L3
親密度＝2
ユーザc
(b)時間＝t1
ユーザb
ユーザd
ユーザc
ユーザa
親密度＝8
ユーザb
親密度＝7
ユーザd
親密度＝4
ユーザc
距離L1'
距離L2'
距離L3'
ユーザa

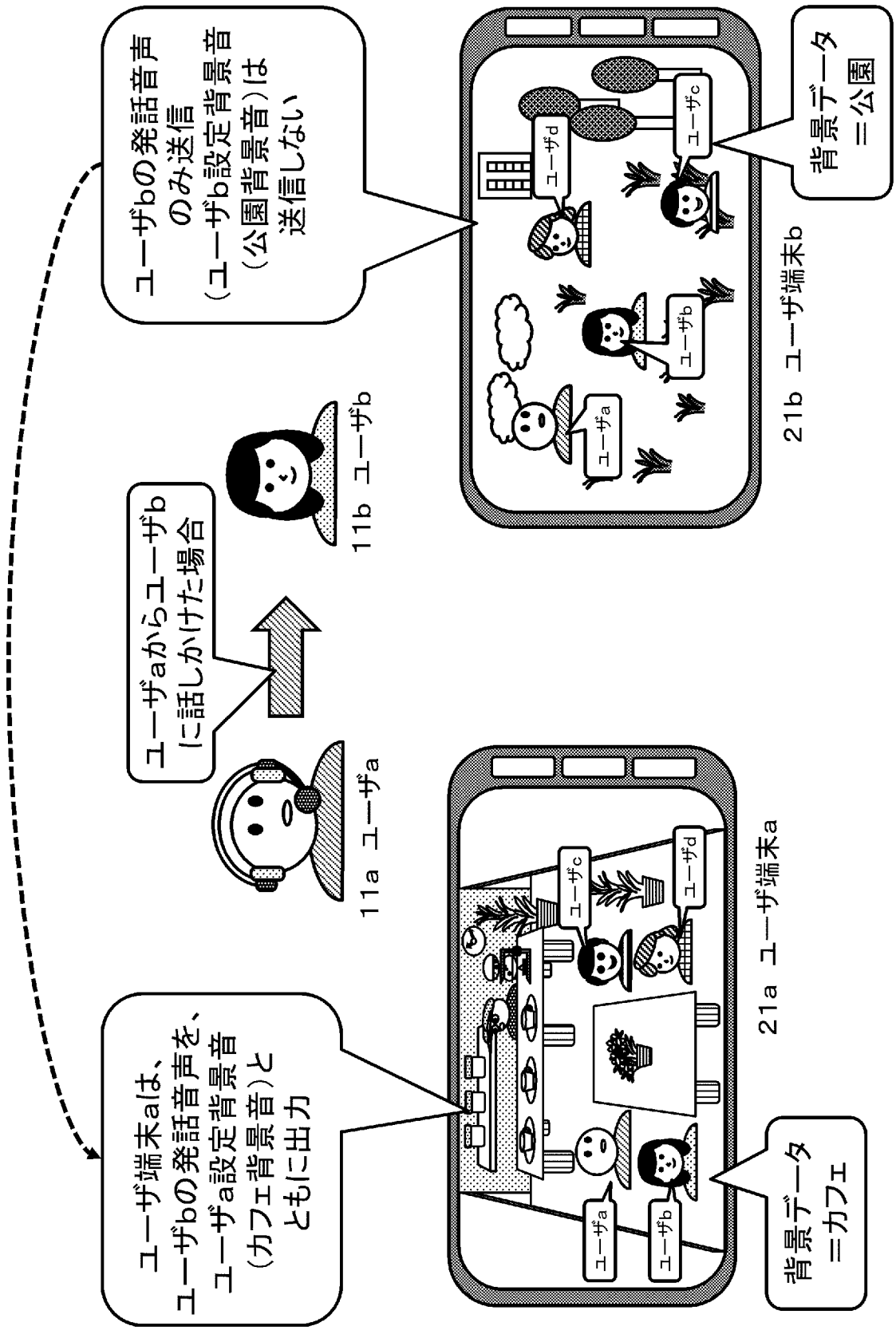
[図28]


(a)親密度変化に伴うユーザ画像更新処理例1
親密度上昇時
親密度低下時
親密度＝低
親密度＝高
(b)親密度変化に伴うユーザ画像更新処理例2
親密度上昇時
親密度低下時
親密度上昇時
親密度低下時
親密度＝低
親密度＝中
親密度＝高

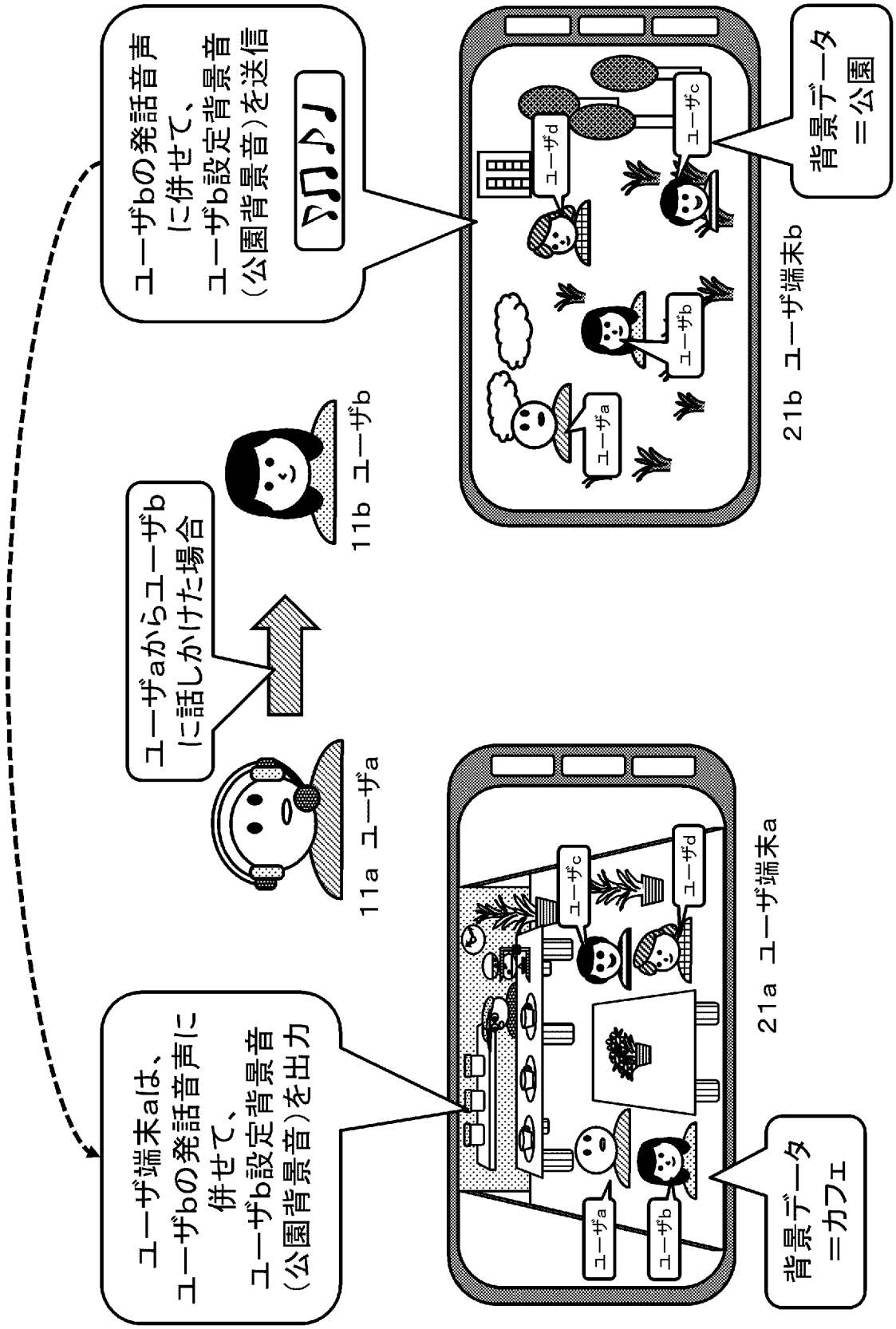
[図29]


背景データ
＝公園
21b ユーザ端末b
11b ユーザb
ユーザa
ユーザb
ユーザc
ユーザd
背景データ
＝公園
11d ユーザd
22d ユーザ端末d
ユーザa
ユーザb
ユーザc
ユーザd
背景データ
＝カフェ
21a ユーザ端末a
11a ユーザa
ユーザa
ユーザb
ユーザc
ユーザd
背景データ＝
ライブハウス
ライブハウス
11c ユーザc
21c ユーザ端末c

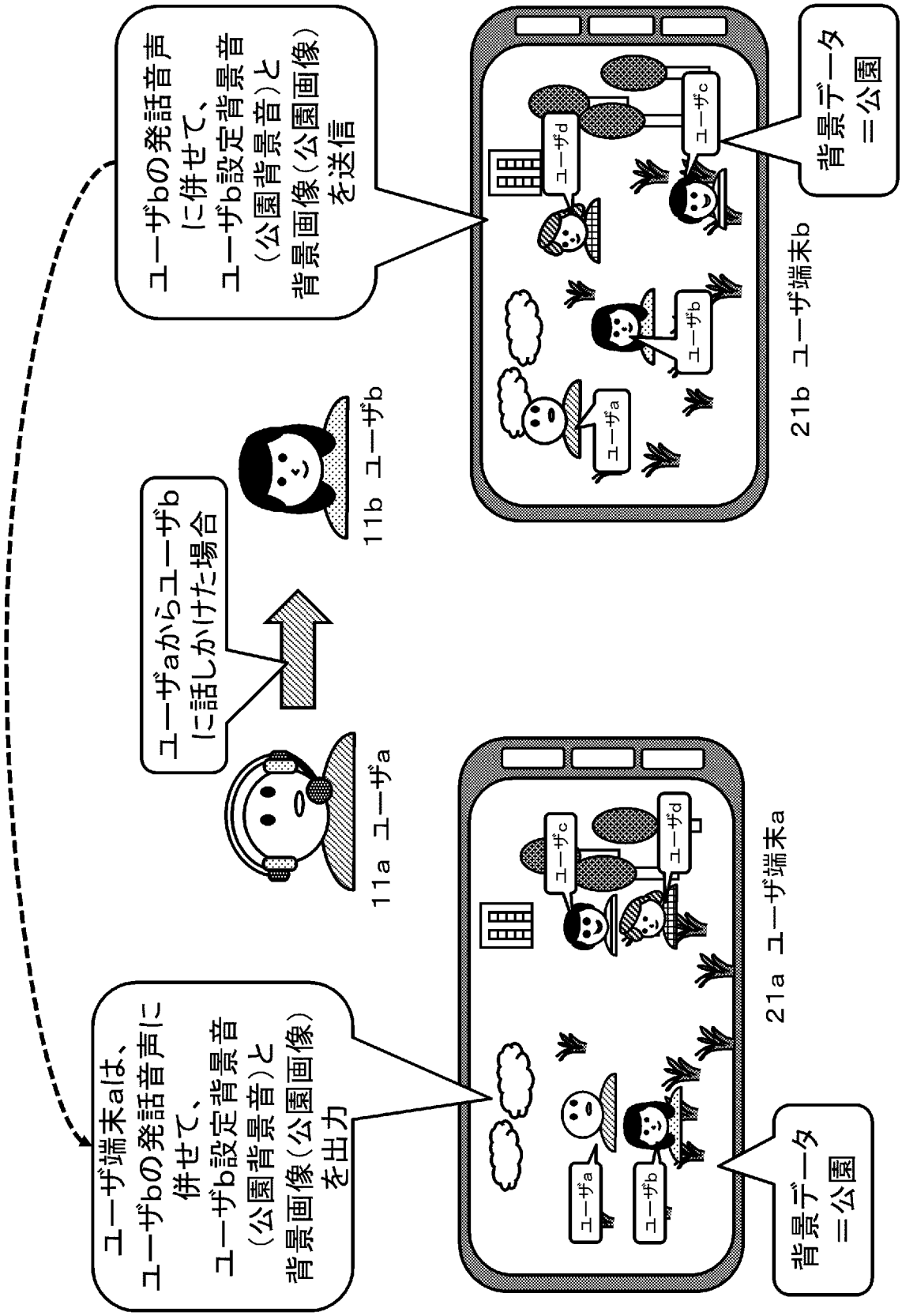
[図30]


ユーザbの発話音声
のみ送信
（ユーザb設定背景音）は
送信しない
ユーザaからユーザb
に話しかけた場合
ユーザ端末aは、
ユーザbの発話音声を、
ユーザa設定背景音
（カフェ背景音）と
ともに出力
11a ユーザa
11b ユーザb
21a ユーザ端末a
21b ユーザ端末b
ユーザa
ユーザb
ユーザc
ユーザd
背景データ
＝カフェ
背景データ
＝公園

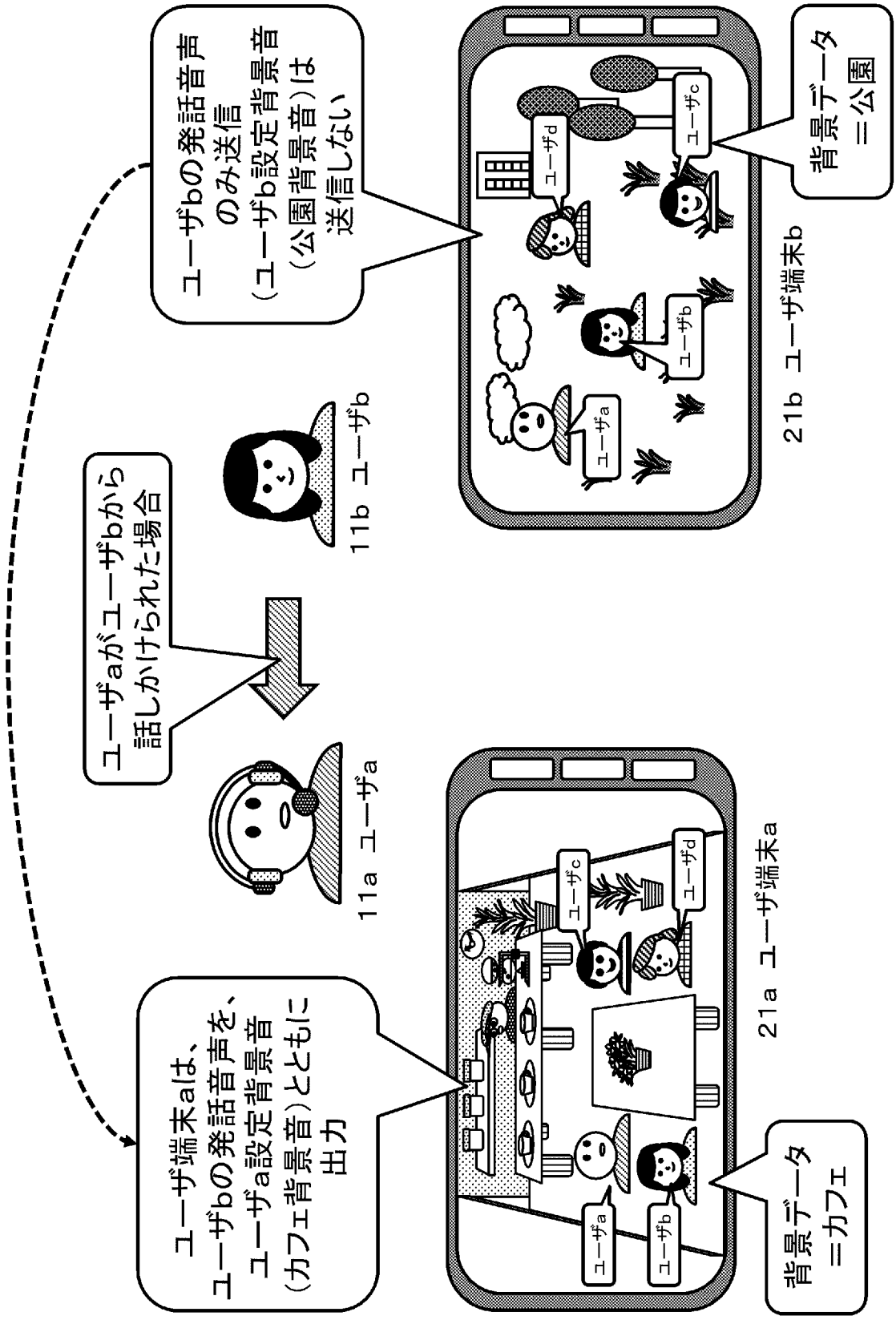
[図31]


ユーザbの発話音声
に併せて、
ユーザb設定背景音
(公園背景音)を送信
背景データ
=公園
21b ユーザ端末b
ユーザd
ユーザc
ユーザa
ユーザb
ユーザaからユーザb
に話しかけた場合
11b ユーザb
11a ユーザa
ユーザ端末aは、
ユーザbの発話音声に
併せて、
ユーザb設定背景音
(公園背景音)を出力
21a ユーザ端末a
ユーザc
ユーザd
ユーザa
ユーザb
背景データ
=カフェ

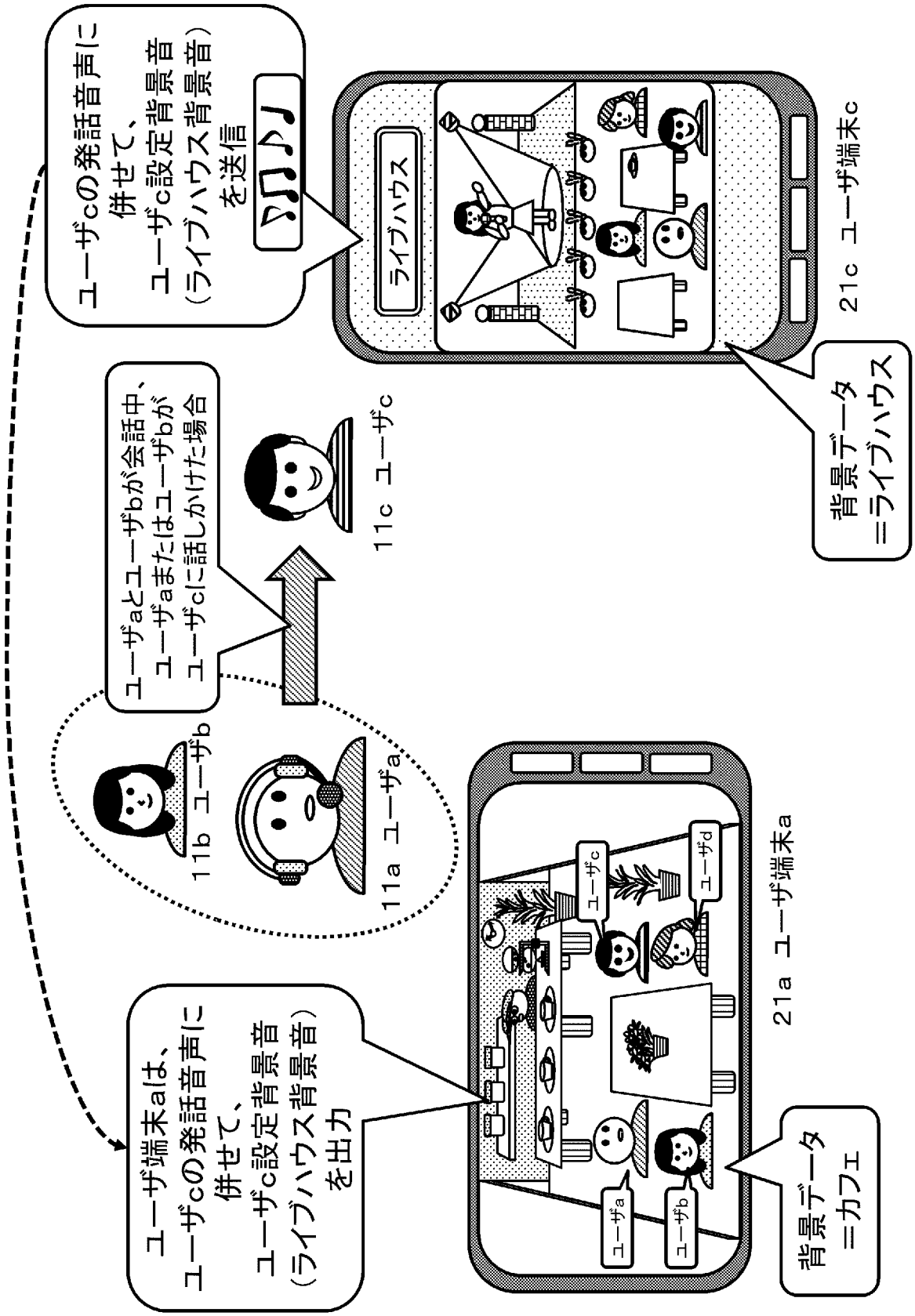
[図32]


ユーザbの発話音声
に併せて、
ユーザb設定背景音
（公園背景音）と
背景画像（公園画像）
を送信
背景データ
＝公園
ユーザd
ユーザc
ユーザa
ユーザb
21b ユーザ端末b
11b ユーザb
ユーザaからユーザb
に話しかけた場合
11a ユーザa
ユーザ端末aは、
ユーザbの発話音声に
併せて、
ユーザb設定背景音
（公園背景音）と
背景画像（公園画像）
を出力
ユーザc
ユーザd
ユーザa
ユーザb
21a ユーザ端末a
背景データ
＝公園

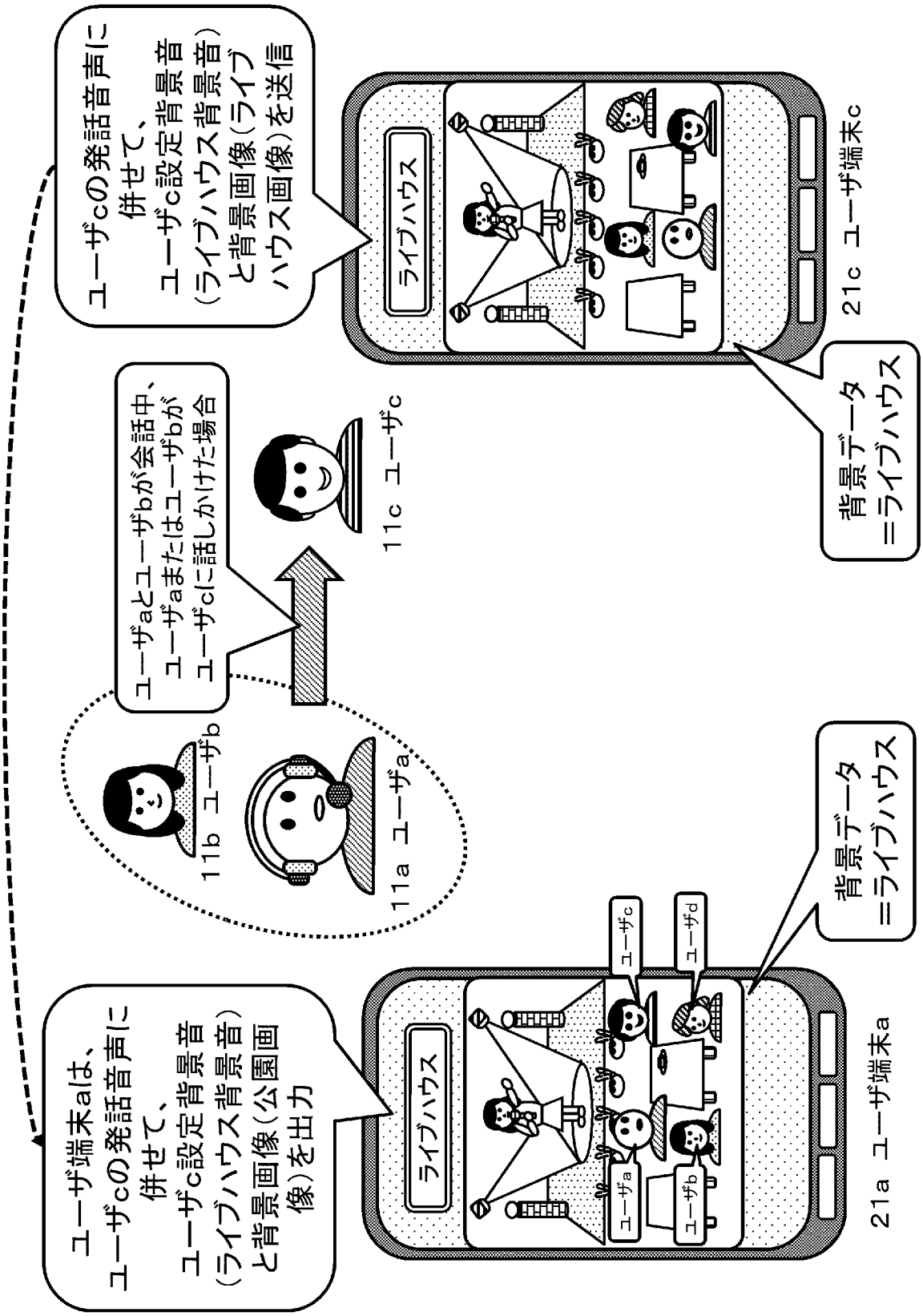
[図33]


ユーザbの発話音声
のみ送信
（ユーザb設定背景音
（公園背景音）は
送信しない
ユーザaがユーザbから
話しかけられた場合
ユーザ端末aは、
ユーザbの発話音声を、
ユーザa設定背景音
（カフェ背景音）とともに
出力
11b ユーザb
11a ユーザa
21b ユーザ端末b
21a ユーザ端末a
背景データ
＝公園
背景データ
＝カフェ
ユーザa
ユーザb
ユーザc
ユーザd

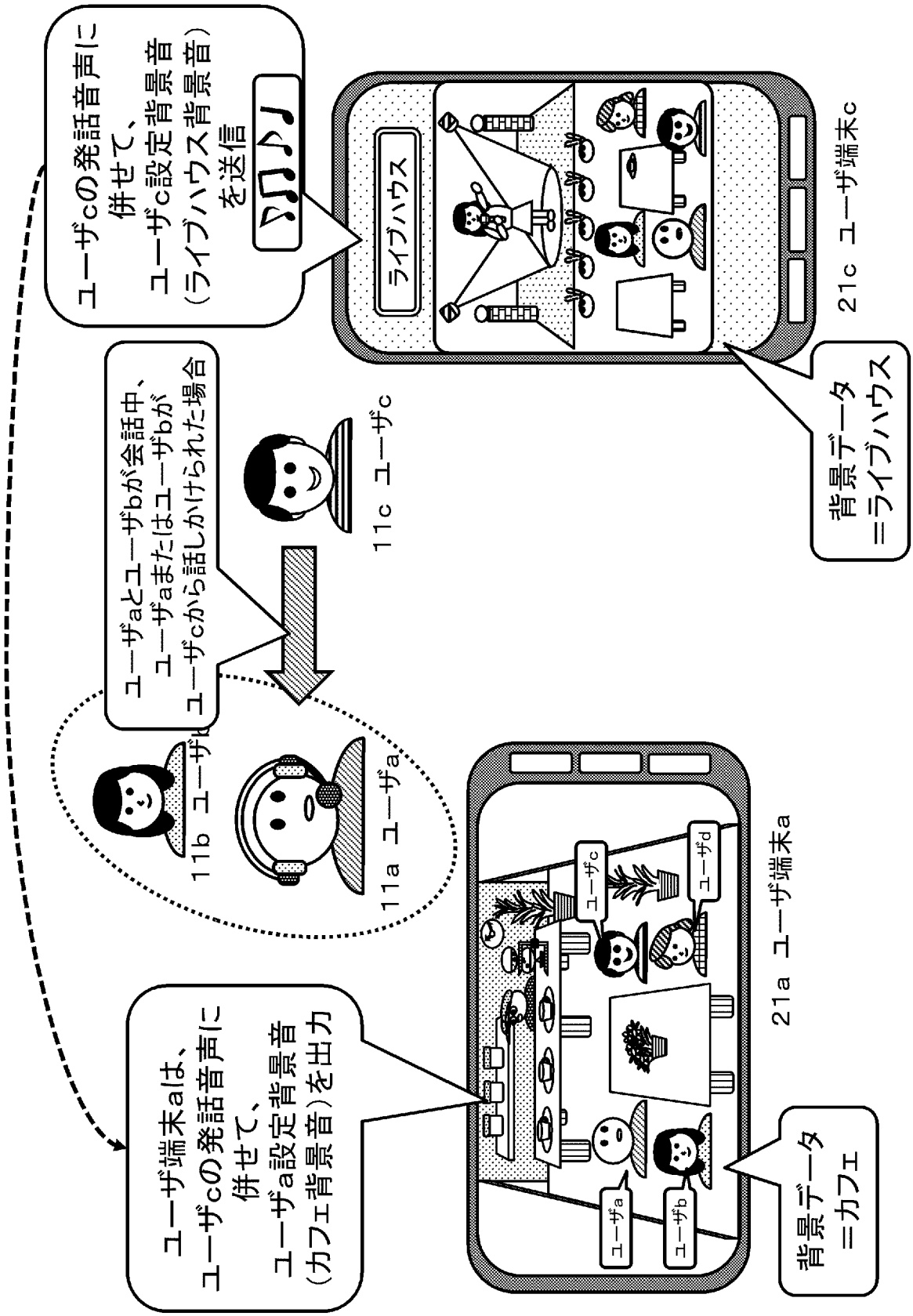
[図34]


ユーザcの発話音声に
併せて、
ユーザc設定背景音
（ライブハウス背景音）
を送信
ライブハウス
21c ユーザ端末c
背景データ
＝ライブハウス
ユーザaとユーザbが会話中、
ユーザaまたはユーザbが
ユーザcに話しかけた場合
11b ユーザb
11a ユーザa
11c ユーザc
ユーザ端末aは、
ユーザcの発話音声に
併せて、
ユーザc設定背景音
（ライブハウス背景音）
を出力
ユーザa
ユーザb
ユーザc
ユーザd
21a ユーザ端末a
背景データ
＝カフェ

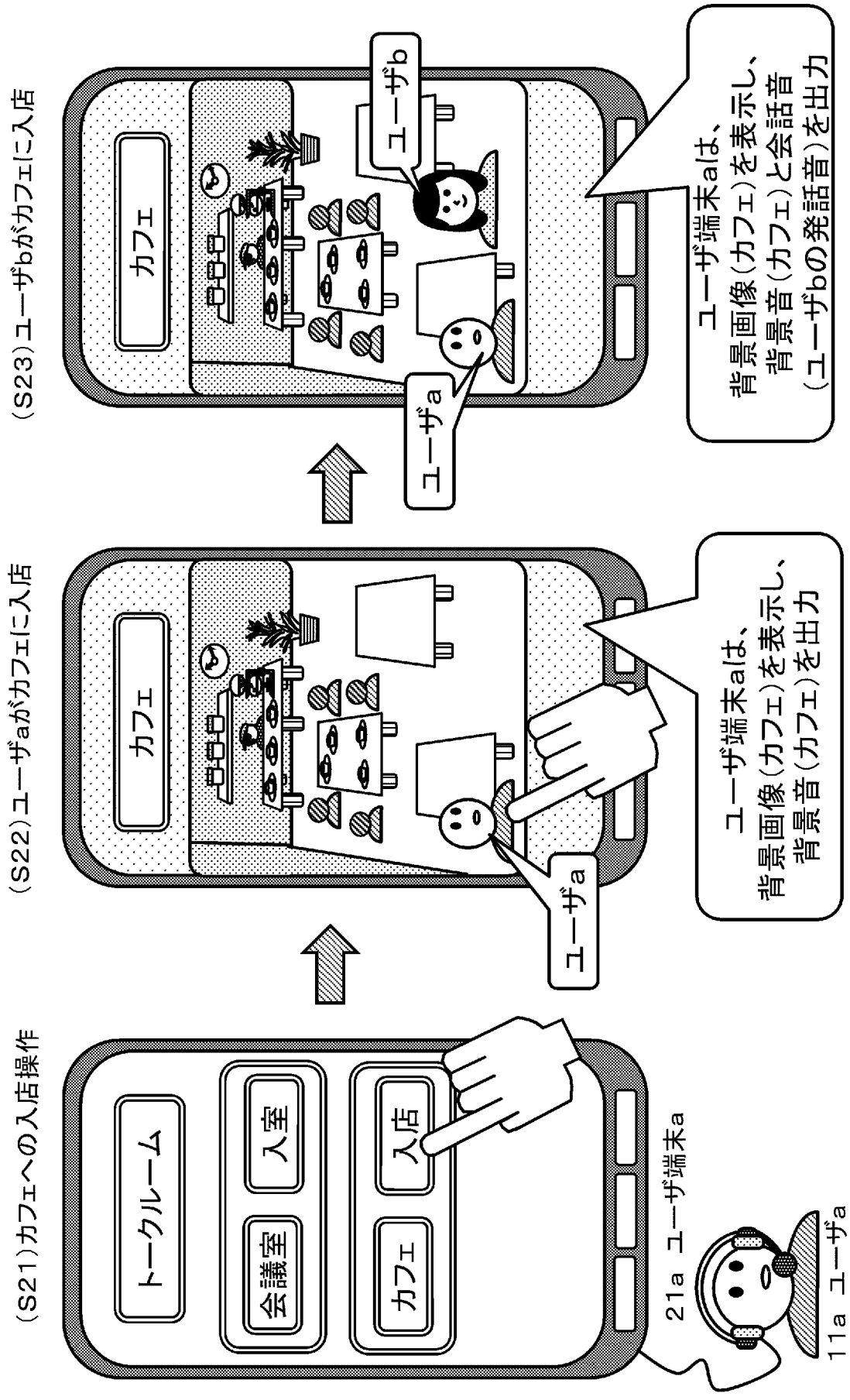
[図35]


ユーザ端末aは、
ユーザcの発話音声に
併せて、
ユーザc設定背景音
(ライブハウス背景音)
と背景画像(公園画
像)を出力
ユーザaとユーザbが会話中、
ユーザaまたはユーザbが
ユーザcに話しかけた場合
ユーザcの発話音声に
併せて、
ユーザc設定背景音
(ライブハウス背景音)
と背景画像(ライブ
ハウス画像)を送信
11b ユーザb
11a ユーザa
11c ユーザc
ライブハウス
ユーザa
ユーザb
ユーザc
ユーザd
21a ユーザ端末a
背景データ
=ライブハウス
ライブハウス
21c ユーザ端末c
背景データ
=ライブハウス

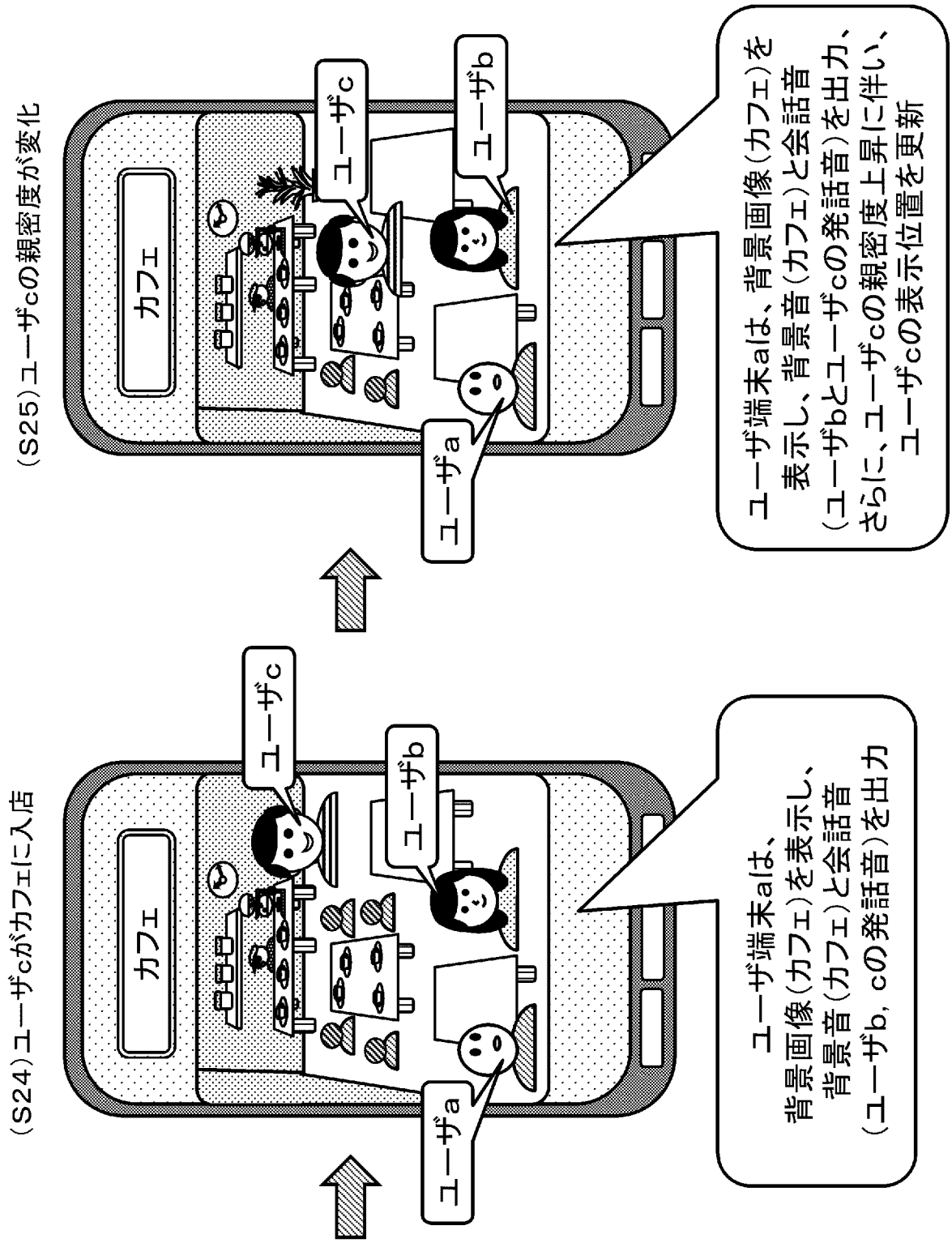
[図36]


ユーザcの発話音声に
併せて、
ユーザc設定背景音
（ライブハウス背景音）
を送信
ライブハウス
21c ユーザ端末c
背景データ
=ライブハウス
ユーザaとユーザbが会話中、
ユーザaまたはユーザbが
ユーザcから話しかけられた場合
11c ユーザc
11b ユーザb
11a ユーザa
ユーザ端末aは、
ユーザcの発話音声に
併せて、
ユーザa設定背景音
（カフェ背景音）を出力
ユーザc
ユーザd
ユーザa
ユーザb
21a ユーザ端末a
背景データ
=カフェ

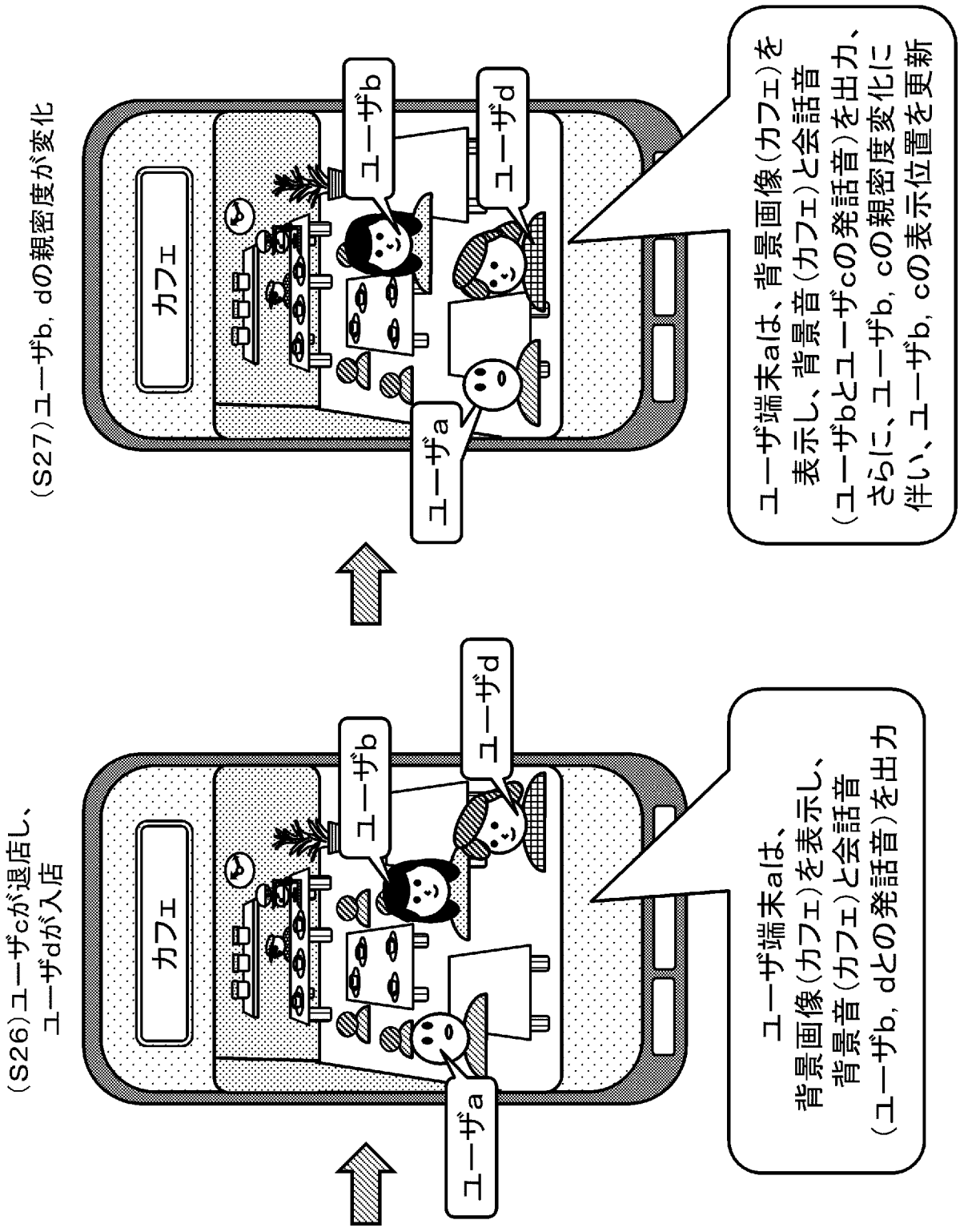
[図37]


(S21)カフェへの入店操作
トークルーム
会議室
入室
カフェ
入店
21a ユーザ端末a
11a ユーザa
(S22)ユーザaがカフェに入店
カフェ
ユーザa
ユーザ端末aは、
背景画像(カフェ)を表示し、
背景音(カフェ)を出力
(S23)ユーザbがカフェに入店
カフェ
ユーザb
ユーザa
ユーザ端末aは、
背景画像(カフェ)を表示し、
背景音(カフェ)と会話音
(ユーザbの発話音)を出力

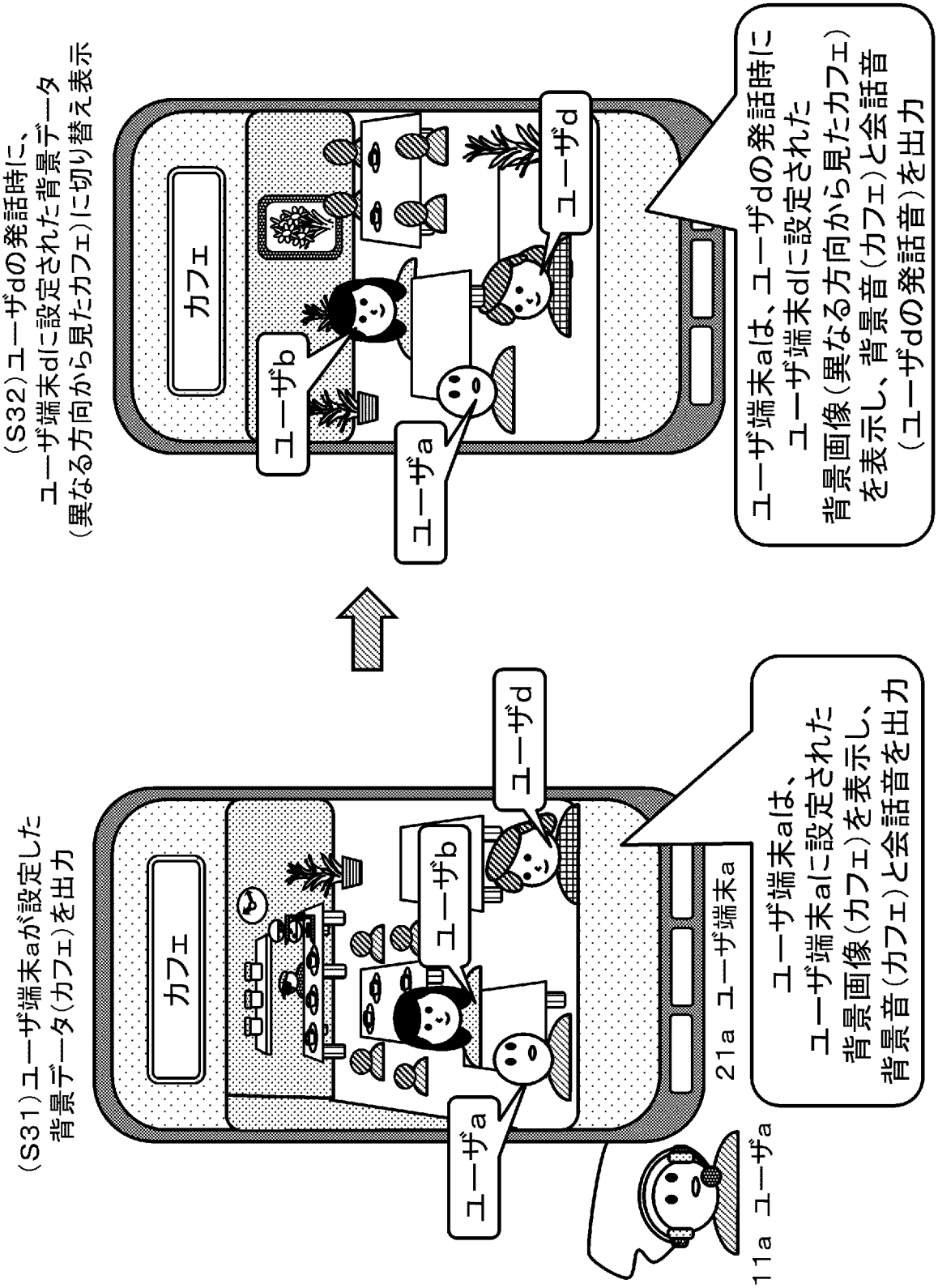
[図38]


(S24)ユーザcがカフェに入店
カフェ
ユーザa
ユーザb
ユーザc
ユーザ端末aは、
背景画像(カフェ)を表示し、
背景音(カフェ)と会話音
(ユーザb, cの発話音)を出力
(S25)ユーザcの親密度が変化
カフェ
ユーザa
ユーザb
ユーザc
ユーザ端末aは、背景画像(カフェ)を
表示し、背景音(カフェ)と会話音
(ユーザbとユーザcの発話音)を出力、
さらに、ユーザcの親密度上昇に伴い、
ユーザcの表示位置を更新

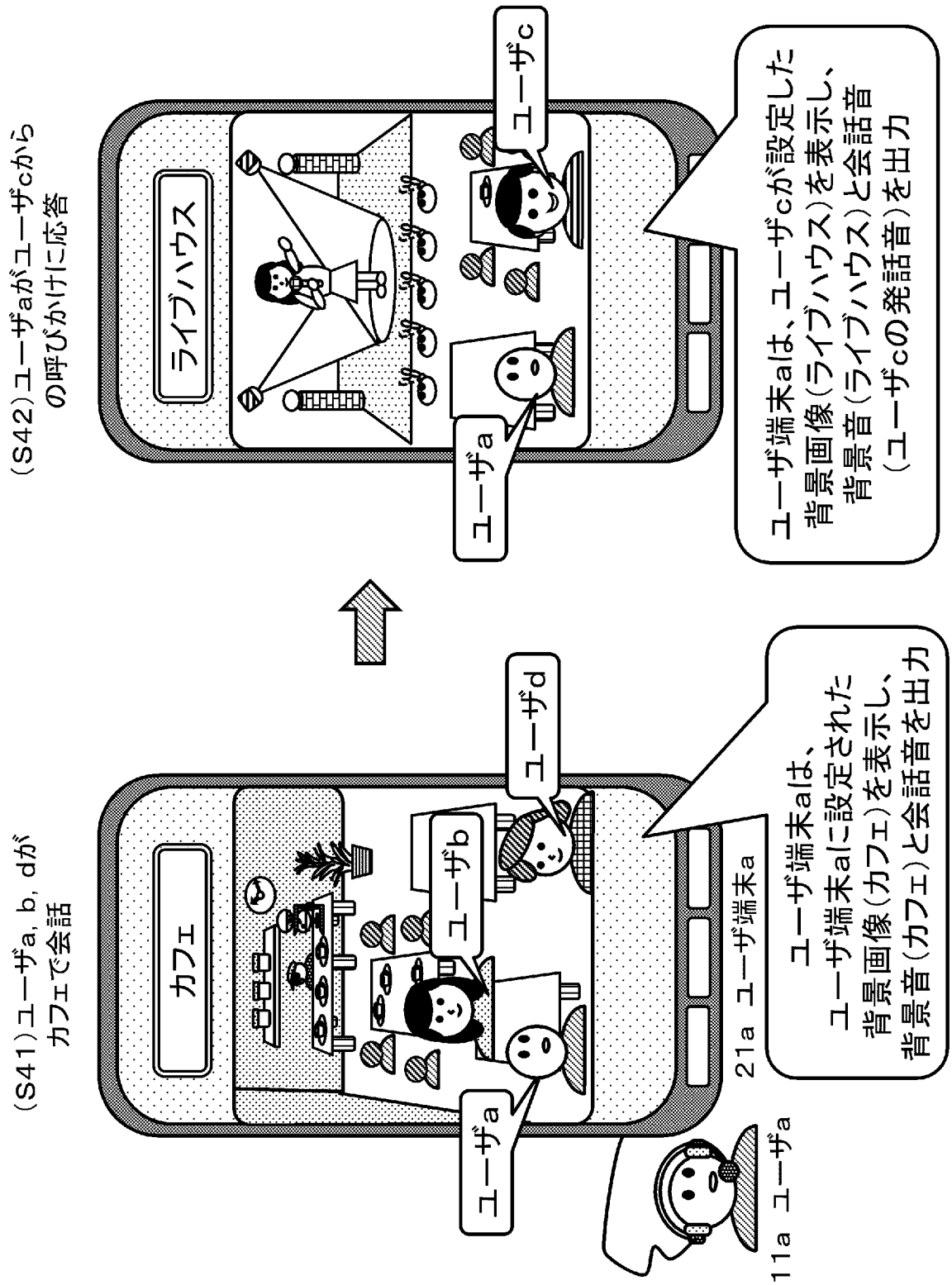
[図39]


(S26)ユーザcが退店し、
ユーザdが入店
カフェ
ユーザa
ユーザb
ユーザd
ユーザ端末aは、
背景画像(カフェ)を表示し、
背景音(カフェ)と会話音
(ユーザb，dとの発話音)を出力
(S27)ユーザb，dの親密度が変化
カフェ
ユーザa
ユーザb
ユーザd
ユーザ端末aは、背景画像(カフェ)を
表示し、背景音(カフェ)と会話音
(ユーザbとユーザcの発話音)を出力、
さらに、ユーザb，cの親密度変化に
伴い、ユーザb，cの表示位置を更新

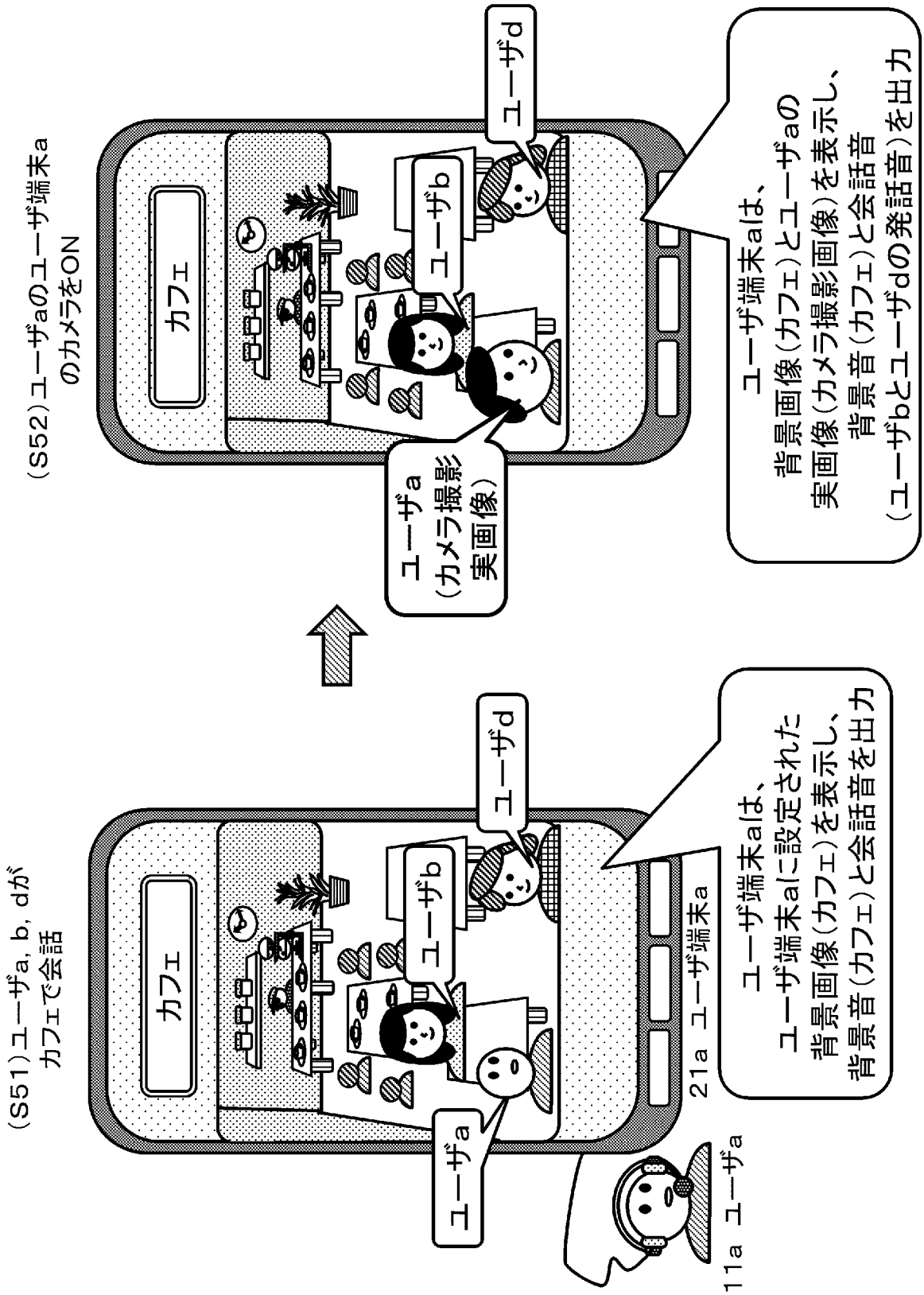
[図40]


(S31)ユーザ端末aが設定した
背景データ(カフェ)を出力
カフェ
ユーザa
ユーザb
ユーザd
11a ユーザa
21a ユーザ端末a
ユーザ端末aは、
ユーザ端末aに設定された
背景画像(カフェ)を表示し、
背景音(カフェ)と会話音を出力
(S32)ユーザdの発話時に、
ユーザ端末dに設定された背景データ
(異なる方向から見たカフェ)に切り替え表示
カフェ
ユーザb
ユーザa
ユーザd
ユーザ端末aは、ユーザdの発話時に
ユーザ端末dに設定された
背景画像(異なる方向から見たカフェ)
を表示し、背景音(カフェ)と会話音
(ユーザdの発話音)を出力

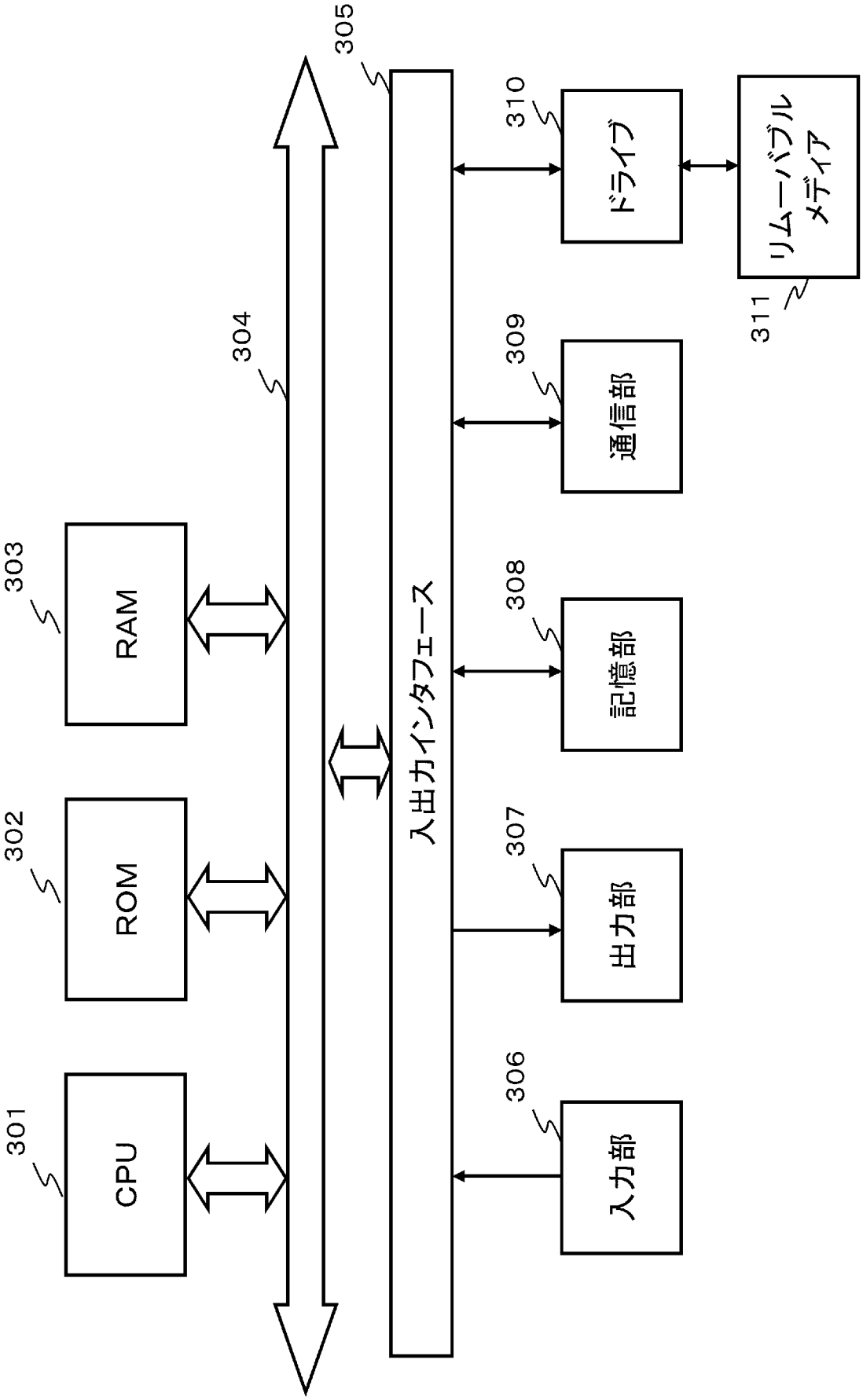
[図41]


(S41)ユーザa, b, dが
カフェで会話
カフェ
ユーザa
ユーザb
ユーザd
11a ユーザa
21a ユーザ端末a
ユーザ端末aは、
ユーザ端末aに設定された
背景画像(カフェ)を表示し、
背景音(カフェ)と会話音を出力
(S42)ユーザaがユーザcから
の呼びかけに応答
ライブハウス
ユーザa
ユーザc
ユーザ端末aは、ユーザcが設定した
背景画像(ライブハウス)を表示し、
背景音(ライブハウス)と会話音
(ユーザcの発話音)を出力

[図42]


(S51)ユーザa, b, dが
カフェで会話
カフェ
ユーザa
ユーザb
ユーザd
11a ユーザa
21a ユーザ端末a
ユーザ端末aは、
ユーザ端末aに設定された
背景画像(カフェ)を表示し、
背景音(カフェ)と会話音を出力
(S52)ユーザaのユーザ端末a
のカメラをON
ユーザa
(カメラ撮影
実画像)
ユーザ端末aは、
背景画像(カフェ)とユーザaの
実画像(カメラ撮影画像)を表示し、
背景音(カフェ)と会話音
(ユーザbとユーザdの発話音)を出力

[図43]


CPU
301
ROM
302
RAM
303
304
305
入出力インタフェース
入力部
306
出力部
307
記憶部
308
通信部
309
ドライブ
310
リムーバブル
メディア
311

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2024/002446

A. CLASSIFICATION OF SUBJECT MATTER

H04R 3/00(2006.01)i; ***G10K 15/04***(2006.01)i; ***H04M 1/72***(2021.01)i; ***H04M 3/56***(2006.01)i; *H04N 7/14*(2006.01)n; *H04N 7/15*(2006.01)n; *H04S 7/00*(2006.01)n
FI: H04R3/00 310; G10K15/04 302F; H04M1/72; H04M3/56 C; H04N7/14 110; H04N7/15 170; H04S7/00 320

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)

H04R3/00; G10K15/04; H04M1/72; H04M3/56; H04N7/14; H04N7/15; 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 1971-2024
Registered utility model specifications of Japan 1996-2024
Published registered utility model applications of Japan 1994-2024

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 2013-017027 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 24 January 2013 (2013-01-24) paragraphs [0015], [0021], [0027]-[0046], fig. 1-6	1-2, 4, 6-9, 17-18
Y		4-5, 10-11, 13, 15
A		14, 16
X	JP 2001-036881 A (CANON KABUSHIKI KAISHA) 09 February 2001 (2001-02-09) paragraphs [0019], [0025]-[0027], [0030]-[0038], fig. 2-3	1-3, 10, 12, 17-18
Y		4-5, 10-11, 13, 15
A		6-9, 14, 16
Y	JP 2006-140595 A (SONY CORPORATION) 01 June 2006 (2006-06-01) paragraphs [0022]-[0030], [0043]-[0051], fig. 1-3	4-5

☑ 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
"D" document cited by the applicant in the international application
"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	Date of mailing of the international search report
03 April 2024	**16 April 2024**
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/JP2024/002446

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2022-116906 A (UNIVERISTY OF WASEDA) 10 August 2022 (2022-08-10) paragraphs [0012]-[0015], [0022]-[0023], [0030]-[0031], [0035], [0050]-[0051], fig. 1-3, 6, 8	10, 13
Y	JP 6-165173 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 10 June 1994 (1994-06-10) paragraphs [0008]-[0009], [0013]-[0014], fig. 1-3	10-11, 15
A		14, 16
A	JP 2014-011509 A (SHARP KABUSHIKI KAISHA) 20 January 2014 (2014-01-20) paragraph [132]	5
A	WO 2021/125081 A1 (NEC CORPORATION) 24 June 2021 (2021-06-24) paragraph [134]	5

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2024/002446

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2013-017027 A	24 January 2013	(Family: none)	
JP 2001-036881 A	09 February 2001	(Family: none)	
JP 2006-140595 A	01 June 2006	(Family: none)	
JP 2022-116906 A	10 August 2022	(Family: none)	
JP 6-165173 A	10 June 1994	(Family: none)	
JP 2014-011509 A	20 January 2014	(Family: none)	
WO 2021/125081 A1	24 June 2021	US 2023/0007428 A1 paragraph [168]	

国際調査報告

国際出願番号
PCT/JP2024/002446

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

H04R 3/00(2006.01)i; G10K 15/04(2006.01)i; H04M 1/72(2021.01)i; H04M 3/56(2006.01)i;
H04N 7/14(2006.01)n; H04N 7/15(2006.01)n; H04S 7/00(2006.01)n
FI: H04R3/00 310; G10K15/04 302F; H04M1/72; H04M3/56 C; H04N7/14 110; H04N7/15 170; H04S7/00 320

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））

H04R3/00; G10K15/04; H04M1/72; H04M3/56; H04N7/14; H04N7/15; 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	１９７１－２０２４年
日本国実用新案登録公報	１９９６－２０２４年
日本国登録実用新案公報	１９９４－２０２４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 2013-017027 A (日本電信電話株式会社) 24.01.2013 (2013-01-24) 段落15,21,27-46, 図1-6	1-2,4,6-9,17-18
Y		4-5,10-11,13,15
A		14,16
X	JP 2001-036881 A (キヤノン株式会社) 09.02.2001 (2001-02-09) 段落19,25-27,30-38, 図2-3	1-3,10,12,17-18
Y		4-5,10-11,13,15
A		6-9,14,16
Y	JP 2006-140595 A (ソニー株式会社) 01.06.2006 (2006-06-01) 段落22-30,43-51, 図1-3	4-5
Y	JP 2022-116906 A (学校法人早稲田大学) 10.08.2022 (2022-08-10) 段落12-15,22-23,30-31,35,50-51, 図1-3,6,8	10,13

☑ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
"A" 特に関連のある文献ではなく、一般的技術水準を示すもの
"D" 国際出願で出願人が先行技術文献として記載した文献
"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	国際調査報告の発送日
03.04.2024	16.04.2024

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）

中嶋　樹理 5Z 1161

電話番号 03-3581-1101 内線 3549

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 6-165173 A (日本電信電話株式会社) 10.06.1994 (1994-06-10) 段落8-9,13-14,図1-3	10-11,15
A		14,16
A	JP 2014-011509 A (シャープ株式会社) 20.01.2014 (2014-01-20) 段落132	5
A	WO 2021/125081 A1 (日本電気株式会社) 24.06.2021 (2021-06-24) 段落134	5

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/JP2024/002446

引用文献	公表日	パテントファミリー文献	公表日
JP 2013-017027 A	24.01.2013	（ファミリーなし）	
JP 2001-036881 A	09.02.2001	（ファミリーなし）	
JP 2006-140595 A	01.06.2006	（ファミリーなし）	
JP 2022-116906 A	10.08.2022	（ファミリーなし）	
JP 6-165173 A	10.06.1994	（ファミリーなし）	
JP 2014-011509 A	20.01.2014	（ファミリーなし）	
WO 2021/125081 A1	24.06.2021	US 2023/0007428 A1 段落168	