

(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
2014年1月30日(30.01.2014)



(10) 国際公開番号
WO 2014/017134 A1

- (51) 国際特許分類:
H04R 3/00 (2006.01) H04R 1/40 (2006.01)
G10L 25/51 (2013.01) H04S 5/02 (2006.01)
- (21) 国際出願番号: PCT/JP2013/061894
- (22) 国際出願日: 2013年4月23日(23.04.2013)
- (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 2012-166977 2012年7月27日(27.07.2012) JP
- (71) 出願人: ソニー株式会社 (SONY CORPORATION)
[JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 佐古 曜一郎 (SAKO, Yoichiro); 〒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). 荒谷 勝久 (ARATANI, Katsuhisa); 〒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). 迫田 和之 (SAKODA, Kazuyuki); 〒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). 浅田 宏平 (ASADA, Kohei); 〒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). 竹原 充 (TAKEHARA, Mitsuru); 〒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). 中村 隆俊 (NAKAMURA, Takatoshi); 〒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). 渡邊 一弘 (WATANABE, Kazuhiro); 〒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). 丹下 明 (TANGE, Akira); 〒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). 花谷 博幸 (HANAYA, Hiroyuki); 〒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). 甲賀 有希 (KOGA, Yuki); 〒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). 大沼 智也 (ONUMA, Tomoya); 〒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: 亀谷 美明, 外 (KAMEYA, Yoshiaki et al.); 〒1600004 東京都新宿区四谷3-1-3 第一富澤ビル はづき国際特許事務所 四谷オフィス Tokyo (JP).

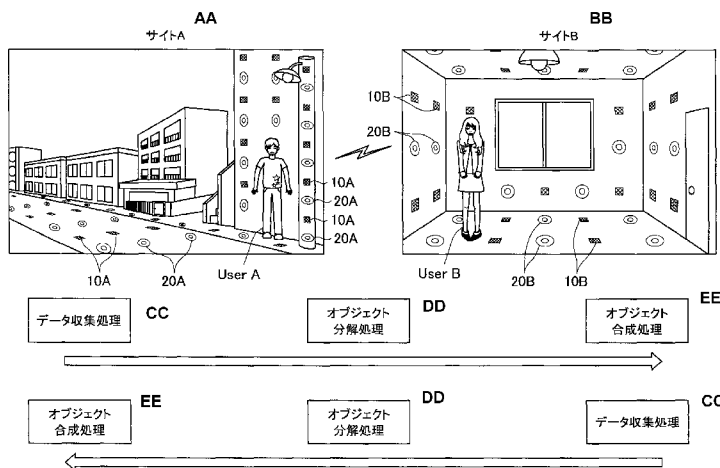
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH,

[続葉有]

(54) Title: INFORMATION PROCESSING SYSTEM AND STORAGE MEDIUM

(54) 発明の名称: 情報処理システムおよび記憶媒体

[図1]

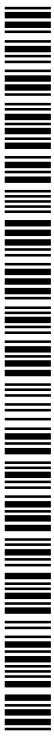


AA Site A
 BB Site B
 CC Data collection processing
 DD Object decomposition processing
 EE Object synthesis processing

(57) Abstract: [Problem] To provide an information processing system and storage medium capable of providing a sense of immersion into a third space when linking a space in the vicinity of a user and another space together. [Solution] The information processing system is equipped with: a recognizing unit that recognizes a first object and a second object on the basis of a signal detected by a plurality of sensors placed in the vicinity of a specific user; an identification unit that identifies the first object and second object recognized by the recognizing unit; an estimating unit that estimates the position of the specific user according to a signal detected by any of the plurality of sensors; and a signal processing unit that processes each signal acquired from sensors in the vicinity of the first and second objects identified by the identifying unit so that the sound when output from a plurality of actuators placed in the vicinity of the specific user is located in the vicinity of the position of the specific user estimated by the estimating unit.

(57) 要約:

[続葉有]



WO 2014/017134 A1



PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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, ML, MR, NE, SN, TD, TG).

- (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, 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,

添付公開書類:

- 国際調査報告 (条約第 21 条(3))

【課題】 ユーザ周辺の空間を他の空間と相互連携させる際に、第 3 の空間への没入感を提供することが可能な情報処理システムおよび記憶媒体を提供する。 【解決手段】 特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第 1 の対象および第 2 の対象を認識する認識部と、前記認識部により認識された前記第 1 および第 2 の対象を同定する同定部と、前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、前記特定ユーザの周辺に配される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第 1 および第 2 の対象の周辺のセンサから取得した各信号を処理する信号処理部と、を備える、情報処理システム。

明 細 書

発明の名称： 情報処理システムおよび記憶媒体

技術分野

[0001] 本開示は、情報処理システムおよび記憶媒体に関する。

背景技術

[0002] 近年、データ通信の分野において様々な技術が提案されている。例えば、下記特許文献1では、M2M (Machine-to-Machine) ソリューションに関する技術が提案されている。具体的には、特許文献1に記載の遠隔管理システムは、インターネットプロトコル (IP) のマルチメディアサブシステム (IMS) プラットフォーム (IS) を利用し、装置によるプレゼンス情報の公開や、ユーザと装置の間のインスタントメッセージングを介して、権限のあるユーザのクライアント (UC) と機械のクライアント (DC) の相互作用が実現される。

[0003] 一方、音響技術の分野において、音響ビームを形成することができるアレイスピーカーが種々開発されている。例えば、下記特許文献2には、複数のスピーカーをその波面を共通にして一つのキャビネットに取り付け、各スピーカーから発する音の遅延量とレベルを制御するアレイスピーカーについて記載されている。また、下記特許文献2には、同様の原理によるアレイマイクも開発されている旨が記載され、当該アレイマイクは、各マイクの出力信号のレベルと遅延量とを調整することにより、その集音点を任意に設定でき、これにより効率のよい集音が可能となる。

先行技術文献

特許文献

[0004] 特許文献1：特開2006-279565号公報

特許文献2：特表2008-543137号公報

発明の概要

発明が解決しようとする課題

[0005] しかしながら、上述した特許文献 1、2 では、大量のイメージセンサ、マイク、スピーカー等を広範囲に配し、ユーザの身体拡張を実現する手段として捉える技術やコミュニケーション方法については、何ら言及されていない。

[0006] そこで、本開示では、ユーザ周辺の空間を他の空間と相互連携させる際に、第 3 の空間への没入感を提供することが可能な、新規かつ改良された情報処理システムおよび記憶媒体を提案する。

課題を解決するための手段

[0007] 本開示によれば、特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第 1 の対象および第 2 の対象を認識する認識部と、前記認識部により認識された前記第 1 および第 2 の対象を同定する同定部と、前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、前記特定ユーザの周辺に配される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第 1 および第 2 の対象の周辺のセンサから取得した各信号を処理する信号処理部と、を備える情報処理システムを提案する。

[0008] 本開示によれば、特定ユーザの周辺のセンサにより検知された信号に基づいて、第 1 の対象および第 2 の対象を認識する認識部と、前記認識部により認識された前記第 1 および第 2 の対象を同定する同定部と、前記同定部により同定された前記第 1 および第 2 の対象の周辺に配される複数のセンサから取得された信号に基づき、前記特定ユーザの周辺のアクチュエータから出力する信号を生成する信号処理部と、を備える情報処理システムを提案する。

[0009] 本開示によれば、コンピュータを、特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第 1 の対象および第 2 の対象を認識する認識部と、前記認識部により認識された前記第 1 および第 2 の対象を同定する同定部と、前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、前記特定ユーザの周辺に配

される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第1および第2の対象の周辺のセンサから取得した各信号を処理する信号処理部、として機能させるためのプログラムが記憶された記憶媒体を提案する。

[0010] 本開示によれば、コンピュータを、特定ユーザの周辺のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、前記認識部により認識された前記第1および第2の対象を同定する同定部と、前記同定部により同定された前記第1および第2の対象の周辺に配される複数のセンサから取得された信号に基づき、前記特定ユーザの周辺のアクチュエータから出力する信号を生成する信号処理部、として機能させるためのプログラムが記憶された記憶媒体を提案する。

発明の効果

[0011] 以上説明したように本開示によれば、ユーザ周辺の空間を他の空間と相互連携させる際に、第3の空間への没入感を提供することが可能となる。

図面の簡単な説明

[0012] [図1]本開示の一実施形態による音響システムの概要を説明するための図である。

[図2]本開示の一実施形態による音響システムのシステム構成を示す図である。

[図3]本実施形態による信号処理装置の構成を示すブロック図である。

[図4]本実施形態による音響閉曲面の形状について説明するための図である。

[図5]本実施形態による管理サーバの構成を示すブロック図である。

[図6]本実施形態による音響システムの基本処理を示すフローチャートである。

[図7]本実施形態によるコマンド認識処理を示すフローチャートである。

[図8]本実施形態による收音処理を示すフローチャートである。

[図9]本実施形態による第3の空間の音場構築について説明するための図であ

る。

[図10] サイトCの音場構築手法を説明するための図である。

[図11] 本実施形態による管理サーバの他の構成を示すブロック図である。

[図12] 音響パラメータの測定について説明するための図である。

[図13] 測定環境における複数のマイクの配置と、聴取環境における複数のスピーカーの配置を比較して示す図である。

[図14] 本実施形態による測定環境における閉曲面の形状について説明するための図である。

[図15] サイトCへの没入感を提供するよう音場構築を行う音場再生信号処理部の構成を示すブロック図である。

[図16A] サイトCにおけるインパルス応答の測定について説明するための図である。

[図16B] 本実施形態によるMatrix Convolution部によるインパルス応答群を用いた演算について説明するための図である。

[図17] 本実施形態による音場再生処理を示すフローチャートである。

[図18A] サイトBで構築される音場が固定されている場合について説明するための図である。

[図18B] サイトBで構築される音場が流動する場合について説明するための図である。

[図19A] 測定対象空間における測定について説明するための図である。

[図19B] 無響室における測定について説明するための図である。

[図19C] 再生対象空間における再現について説明するための図である。

[図20] 本実施形態による音響システムの他のシステム構成を示す図である。

[図21] 本実施形態による自律型音響システムのシステム構成の一例を示す図である。

[図22] 本実施形態による自律型音響システムのデバイスの構成を示すブロック図である。

[図23] 本実施形態による自律型音響システムの動作処理を示すフローチャー

トである。

[図24]本実施形態による自律型音響システムにおいて、ユーザの移動に応じた動作デバイスの変更について説明するための図である。

[図25]本実施形態による自律型音響システムにおいて、複数ユーザにサービス提供を行う場合について説明するための図である。

発明を実施するための形態

[0013] 以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

[0014] また、説明は以下の順序で行うものとする。

1. 本開示の一実施形態による音響システムの概要
2. 基本構成
 - 2-1. システム構成
 - 2-2. 信号処理装置
 - 2-3. 管理サーバ
3. 動作処理
 - 3-1. 基本処理
 - 3-2. コマンド認識処理
 - 3-3. 收音処理
4. 第3空間の音場構築
 - 4-1. 管理サーバの構成
 - 4-2. 音場再生信号処理部の構成
 - 4-3. 音場再生処理
5. 補足
6. まとめ

[0015] <1. 本開示の一実施形態による音響システムの概要>

まず、本開示の一実施形態による音響システム（情報処理システム）の概

要について、図1を参照して説明する。図1は、本開示の一実施形態による音響システムの概要を説明するための図である。図1に示すように、本実施形態による音響システムでは、部屋、家、ビル、屋外、地域、国等の世界の至る所に大量のマイクロフォン10、イメージセンサ（不図示）、およびスピーカー20等の各種センサおよびアクチュエータが配置されている状況を想定する。

[0016] 図1に示す例では、ユーザAが現在居る屋外の一のエリア「サイトA」の道路等に、複数のセンサの一例として、複数のマイクロフォン（以下、マイクと称す）10A、および複数のアクチュエータの一例として、複数のスピーカー20Aが配されている。また、ユーザBが現在居る屋内の一のエリア「サイトB」では、壁、床、天井等に、複数のマイク10Bおよび複数のスピーカー20Bが配されている。なお、サイトA、Bには、センサの一例として、図示しない人感知センサやイメージセンサがさらに配されていてもよい。

[0017] ここで、サイトAとサイトBはネットワークを介して接続可能であって、サイトAの各マイクおよびスピーカーで入出力される信号と、サイトBの各マイクおよびスピーカーで入出力される信号は、互いに送受信される。

[0018] これにより、本実施形態による音響システムは、所定対象（人物、場所、建物等）に対応する音声や画像をユーザの周囲に配された複数のスピーカーやディスプレイでリアルタイムに再生する。また、本実施形態による音響システムは、ユーザの音声をユーザの周囲に配された複数のマイクにより收音して所定対象の周囲でリアルタイムに再生することができる。このように、本実施形態による音響システムでは、ユーザ周辺の空間を他の空間と相互連携させることが可能となる。

[0019] また、屋内や屋外の至る所に配されるマイクロフォン10、スピーカー20、イメージセンサ等を用いて、実質的にユーザの口、目、耳等の身体を広範囲に拡張させることが可能となり、新たなコミュニケーション方法を実現することができる。

[0020] さらに、本実施形態による音響システムでは、至る所にマイクロフォンやイメージセンサ等が配されているので、ユーザはスマートフォンや携帯電話端末を所有する必要がなく、声やジェスチャーで所定対象を指示し、所定対象周辺の空間と接続させることができる。以下、サイトAに居るユーザAがサイトBに居るユーザBと会話したい場合における本実施形態による音響システムの適用について簡潔に説明する。

[0021] (データ収集処理)

サイトAでは、複数のマイク10A、イメージセンサ（不図示）、および人感センサ（不図示）等により継続的にデータ収集処理が行われている。具体的には、本実施形態による音響システムは、複数のマイク10Aで収録した音声、イメージセンサで撮像した撮像画像、または人感センサの検知結果を収集し、これによりユーザの位置を推定する。

[0022] また、本実施形態による音響システムは、予め登録された複数のマイク10Aの位置情報、および推定されたユーザの位置に基づいて、ユーザの声が十分収録可能な位置に配されているマイク群を選出してもよい。そして、本実施形態による音響システムは、選出した各マイクにより収録されたオーディオ信号のストリーム群に対してマイクアレイ処理を行う。特に、本実施形態による音響システムは、ユーザAの口元に収録点が合うような遅延和アレイを行ってもよく、これによりアレイマイクの超指向性を形成できる。よって、ユーザAのつぶやき程度の小さな声も収録され得る。

[0023] また、本実施形態による音響システムは、収録したユーザAの音声に基づいてコマンドを認識し、コマンドに従った動作処理を実行する。例えば、サイトAに居るユーザAが「Bさんと話したい」とつぶやくと、「ユーザBへの発呼要求」がコマンドとして認識される。この場合、本実施形態による音響システムは、ユーザBの現在位置を同定し、ユーザBが現在居るサイトBとユーザAが現在居るサイトAを接続させる。これにより、ユーザAは、ユーザBと通話を行うことができる。

[0024] (オブジェクト分解処理)

通話中においては、サイトAの複数のマイクで収録されたオーディオ信号（ストリームデータ）に対して、音源分離（ユーザAの周囲のノイズ成分や、ユーザAの周囲の人物の会話などを分離）、残響抑制、ノイズ／エコー処理等のオブジェクト分解処理が行われる。これにより、S／N比のよい、残響感も抑制されたストリームデータがサイトBに送られる。

[0025] なお、ユーザAが移動しながら話している場合も想定されるが、本実施形態による音響システムは、上記データ収集を継続的に行うことで対応することができる。具体的には、本実施形態による音響システムは、複数のマイク、イメージセンサ、および人感センサ等に基づいて継続的にデータ収集を行い、ユーザAの移動経路や向いている方向を把握する。そして、本実施形態による音響システムは、移動しているユーザAの周囲に配される適切なマイク群の選出を継続的に更新し、また、移動しているユーザAの口元に常に収録点が合うようアレイマイク処理を継続的に行う。これにより、本実施形態による音響システムは、ユーザAが移動しながら話す場合にも対応することができる。

[0026] また、音声のストリームデータとは別に、ユーザAの移動方向や向き等がメタデータ化され、ストリームデータと共にサイトBに送られる。

[0027] （オブジェクト合成）

そして、サイトBに送られたストリームデータは、サイトBに居るユーザBの周囲に配されたスピーカーから再生される。この際、本実施形態による音響システムは、サイトBにおいて、複数のマイク、イメージセンサ、および人感センサによりデータ収集を行い、収集したデータに基づいてユーザBの位置を推定し、さらにユーザBの周囲を音響閉曲面で囲う適切なスピーカー群を選出する。サイトBに送られたストリームデータは、このように選出したスピーカー群から再生され、音響閉曲面内側のエリアが適切な音場として制御される。なお、本明細書において、ある対象物（例えばユーザ）を取り囲むような形で、近接する複数のスピーカーまたは複数のマイクの位置を繋いだ場合に形成される面を、概念的に「音響閉曲面」と称す。また、「音

響閉曲面」は、必ずしも完全な閉曲面を構成するものではなく、おおよそ対象物（例えばユーザ）を取り囲むような形であればよい。

[0028] また、ここでの音場は、ユーザBが自ら任意で選択できるようにしてもよい。例えば、本実施形態による音響システムは、ユーザBが、サイトAを音場に指定した場合、サイトAの環境がサイトBで再現される。具体的には、例えばリアルタイムに収録されるアンビエントとしての音情報や、予め取得されたサイトAに関するメタ情報に基づいて、サイトAの環境がサイトBで再現される。

[0029] また、本実施形態による音響システムは、サイトBにおいてユーザBの周辺に配された複数のスピーカー20Bを用いて、ユーザAの音像を制御することも可能である。すなわち、本実施形態による音響システムは、アレイスピーカー（ビームフォーミング）を形成することで、ユーザBの耳元や、音響閉曲面の外側にユーザAの声（音像）を再現することも可能である。また、本実施形態による音響システムは、ユーザAの移動経路や向きのメタデータを利用して、サイトBにおいて、ユーザAの実際の移動に合わせてユーザAの音像をユーザBの周囲で移動させてもよい。

[0030] 以上、データ収集処理、オブジェクト分解処理、およびオブジェクト合成処理の各ステップに分けてサイトAからサイトBへの音声通信について概要を説明したが、サイトBからサイトAの音声通信においても当然に同様の処理が行われる。これにより、サイトAおよびサイトBで双方向の音声通信が可能となる。

[0031] 以上、本開示の一実施形態における音響システム（情報処理システム）の概要について説明した。続いて、本実施形態による音響システムの構成について図2～図5を参照して詳細に説明する。

[0032] <2. 基本構成>

[2-1. システム構成]

図2は、本実施形態による音響システムの全体構成を示す図である。図2に示すように、音響システムは、信号処理装置1A、信号処理装置1B、お

よび管理サーバ3を有する。

[0033] 信号処理装置1Aおよび信号処理装置1Bは、有線／無線によりネットワーク5に接続し、ネットワーク5を介して互いにデータの送受信が可能である。また、ネットワーク5には管理サーバ3が接続され、信号処理装置1Aおよび信号処理装置1Bは、管理サーバ3とデータの送受信を行うことも可能である。

[0034] 信号処理装置1Aは、サイトAに配される複数のマイク10Aおよび複数のスピーカー20Aにより入出力される信号を処理する。また、信号処理装置1Bは、サイトBに配される複数のマイク10Bおよび複数のスピーカー20Bにより入出力される信号を処理する。なお、信号処理装置1A、1Bを区別して説明する必要がある場合は、信号処理装置1と称する。

[0035] 管理サーバ3は、ユーザの認証処理や、ユーザの絶対位置（現在位置）を管理する機能を有する。さらに、管理サーバ3は、場所や建物の位置を示す情報（IPアドレス等）を管理してもよい。

[0036] これにより、信号処理装置1は、ユーザにより指定された所定の対象（人物、場所、建物等）の接続先情報（IPアドレス等）を管理サーバ3に問い合わせ取得することができる。

[0037] [2-2. 信号処理装置]

次に、本実施形態による信号処理装置1の構成について詳細に説明する。図3は、本実施形態による信号処理装置1の構成を示すブロック図である。図3に示すように、本実施形態による信号処理装置1は、複数のマイク10（アレイマイク）、アンプ・ADC（アナログデジタルコンバータ）部11、信号処理部13、マイク位置情報DB（データベース）15、ユーザ位置推定部16、認識部17、同定部18、通信I/F（インターフェース）19、スピーカー位置情報DB21、アンプ・DAC（デジタルアナログコンバータ）部23、および複数のスピーカー20（アレイスピーカー）を有する。以下、各構成について説明する。

[0038] （アレイマイク）

複数のマイク10は、上述したように、あるエリア（サイト）の至る所に配置されている。例えば、屋外であれば、道路、電柱、街灯、家やビルの外壁等、屋内であれば、床、壁、天井等に配置される。また、複数のマイク10は、周囲の音を收音し、アンプ・ADC部11に各々出力する。

[0039] (アンプ・ADC部)

アンプ・ADC部11は、複数のマイク10から各々出力された音波の増幅機能 (amplifier)、および音波 (アナログデータ) をオーディオ信号 (デジタルデータ) に変換する機能 (Analog-to-Digital Converter) を有する。アンプ・ADC部11は、変換した各オーディオ信号を信号処理部13に出力する。

[0040] (信号処理部)

信号処理部13は、マイク10により收音され、アンプ・ADC部11を介して送られた各オーディオ信号や、DAC・アンプ部23を介してスピーカー20から再生する各オーディオ信号を処理する機能を有する。また、本実施形態による信号処理部13は、マイクアレイ処理部131、高S/N化処理部133、および音場再生信号処理部135として機能する。

[0041] ・マイクアレイ処理部

マイクアレイ処理部131は、アンプ・ADC部11から出力された複数のオーディオ信号に対するマイクアレイ処理として、ユーザの音声にフォーカスするよう (收音位置がユーザの口元になるよう) 指向性制御を行う。

[0042] この際、マイクアレイ処理部131は、ユーザ位置推定部16により推定されたユーザの位置や、マイク位置情報DB15に登録されている各マイク10の位置に基づいて、ユーザの音声收音に最適な、ユーザを内包する音響閉曲面を形成するマイク群を選択してもよい。そして、マイクアレイ処理部131は、選択したマイク群により取得されたオーディオ信号に対して指向性制御を行う。また、マイクアレイ処理部131は、遅延和アレイ処理、Null生成処理によりアレイマイクの超指向性を形成してもよい。

[0043] ・高S/N化処理部

高S/N化処理部133は、アンプ・ADC部11から出力された複数のオーディオ信号に対して、明瞭度が高くS/N比がよいモノラル信号となるよう処理する機能を有する。具体的には、高S/N化処理部133は、音源を分離し、残響・ノイズ抑制を行う。

[0044] なお、高S/N化処理部133は、マイクアレイ処理部131の後段に設けられてもよい。また、高S/N化処理部133により処理されたオーディオ信号（ストリームデータ）は、認識部17による音声認識に用いられ、通信部I/F19を介して外部に送信されたりする。

[0045] ・音場再生信号処理部

音場再生信号処理部135は、複数のスピーカ20から再生するオーディオ信号に関する信号処理を行い、ユーザの位置付近に音場が定位するよう制御する。具体的には、例えば音場再生信号処理部135は、ユーザ位置推定部16により推定されたユーザの位置やスピーカ位置情報DB21に登録されている各スピーカ20の位置に基づいて、ユーザを内包する音響閉曲面を形成する最適なスピーカ群を選択する。そして、音場再生信号処理部135は、選択したスピーカ群に応じた複数のチャンネルの出力バッファに、信号処理したオーディオ信号を書き込む。

[0046] また、音場再生信号処理部135は、音響閉曲面の内側のエリアを適切な音場として制御する。音場の制御方法は、例えばキルヒホッフ・ヘルムホルツの積分則、またはレイリー積分則として知られるものであり、これを応用した波面合成法(WFS: Wave Field Synthesis)等が一般的に知られている。また、音場再生信号処理部135は、特許第4674505号、および特許第4735108号等に記載の信号処理技術を応用してもよい。

[0047] なお、上述したマイクまたはスピーカにより形成される音響閉曲面の形状は、ユーザを取り囲む立体的な形状であれば特に限定されず、例えば、図4に示すような楕円型の音響閉曲面40-1、円柱型の音響閉曲面40-2、または多角形型の音響閉曲面40-3であってもよい。図4に示す例では

、一例としてサイトBにおいてユーザBの周辺に配される複数のスピーカ—20B-1~20B-12による音響閉曲面の形状を示すが、複数のマイク10による音響閉曲面の形状についても同様である。

[0048] (マイク位置情報DB)

マイク位置情報DB15は、サイトに配される複数のマイク10の位置情報を記憶する記憶部である。複数のマイク10の位置情報は、予め登録されていてもよい。

[0049] (ユーザ位置推定部)

ユーザ位置推定部16は、ユーザの位置を推定する機能を有する。具体的には、ユーザ位置推定部16は、複数のマイク10から收音した音声の解析結果、イメージセンサにより撮像した撮像画像の解析結果、または人感センサによる検知結果に基づいて、複数のマイク10または複数のスピーカ—20に対するユーザの相対位置を推定する。また、ユーザ位置推定部16は、GPS (Global Positioning System) 情報を取得し、ユーザの絶対位置 (現在位置情報) を推定してもよい。

[0050] (認識部)

認識部17は、複数のマイク10により收音され、信号処理部13により処理されたオーディオ信号に基づいてユーザの音声を解析し、コマンドを認識する。例えば、認識部17は、「Bさんと話したい」というユーザの音声を形態素解析し、ユーザに指定された所定の対象「B」および要求「話す」に基づき、発呼要求コマンドを認識する。

[0051] (同定部)

同定部18は、認識部17により認識された所定の対象を同定する機能を有する。具体的には、例えば同定部18は、所定の対象に対応する音声や画像を取得するための接続先情報を決定してもよい。同定部18は、例えば所定の対象を示す情報を通信部I/F19から管理サーバ3に送信し、管理サーバ3から所定の対象に対応する接続先情報 (IPアドレス等) を取得してもよい。

[0052] (通信 I / F)

通信 I / F 19 は、ネットワーク 5 を通じて他の信号処理装置や管理サーバ 3 との間でデータの送受信を行うための通信モジュールである。例えば、本実施形態による通信 I / F 19 は、管理サーバ 3 に対して所定の対象に対応する接続先情報の問い合わせを行ったり、接続先である他の信号処理装置に、マイク 10 で收音して信号処理部 13 で処理したオーディオ信号を送信したりする。

[0053] (スピーカー位置情報 DB)

スピーカー位置情報 DB 21 は、サイトに配される複数のスピーカー 20 の位置情報を記憶する記憶部である。複数のスピーカー 20 の位置情報は、予め登録されていてもよい。

[0054] (DAC・アンプ部)

DAC・アンプ部 23 は、複数のスピーカー 20 から各々再生するための各チャンネルの出力バッファに書き込まれたオーディオ信号（デジタルデータ）を音波（アナログデータ）に変換する機能（Digital・to・Analog Converter）を有する。

[0055] さらに、DAC・アンプ部 23 は、変換した音波を増幅（amplifier）し、複数のスピーカー 20 から各々再生（出力）する。

[0056] (アレイスピーカー)

複数のスピーカー 20 は、上述したように、あるエリア（サイト）の至る所に配置されている。例えば、屋外であれば、道路、電柱、街灯、家やビルの外壁等、屋内であれば、床、壁、天井等に配置される。また、複数のスピーカー 20 は、DAC・アンプ部 23 から出力された音波（音声）を再生する。

[0057] 以上、本実施形態による信号処理装置 1 の構成について詳細に説明した。続いて、本実施形態による管理サーバ 3 の構成について図 5 を参照して説明する。

[0058] [2-3. 管理サーバ]

図5は、本実施形態による管理サーバ3の構成を示すブロック図である。図5に示すように、管理サーバ3は、管理部32、検索部33、ユーザ位置情報DB35、および通信I/F39を有する。以下、各構成について説明する。

[0059] (管理部)

管理部32は、信号処理装置1から送信されたユーザID (Identification) 等に基づいて、ユーザが現在居る場所 (サイト) に関する情報を管理する。例えば管理部32は、ユーザIDに基づいてユーザを識別し、識別したユーザの氏名等に、送信元の信号処理装置1のIPアドレス等を接続先情報として対応付けてユーザ位置情報DB35に記憶させる。なお、ユーザIDは、氏名、暗証番号、または生体情報等を含んでもよい。また、管理部32は、送信されたユーザIDに基づいてユーザの認証処理を行ってもよい。

[0060] (ユーザ位置情報DB)

ユーザ位置情報DB35は、管理部32による管理に応じて、ユーザが現在居る場所に関する情報を記憶する記憶部である。具体的には、ユーザ位置情報DB35は、ユーザのID、および接続先情報 (ユーザが居るサイトに対応する信号処理装置のIPアドレス等) を対応付けて記憶する。また、各ユーザの現在位置情報は時々刻々と更新されてもよい。

[0061] (検索部)

検索部33は、信号処理装置1からの接続先 (発呼先) 問い合わせに応じて、ユーザ位置情報DB35を参照し、接続先情報を検索する。具体的には、検索部33は、接続先問い合わせに含まれる対象ユーザの氏名等に基づいて、対応付けられた接続先情報をユーザ位置情報DB35から検索して抽出する。

[0062] (通信I/F)

通信I/F39は、ネットワーク5を通じて信号処理装置1との間でデータの送受信を行うための通信モジュールである。例えば、本実施形態による

通信 I / F 3 9 は、信号処理装置 1 からユーザの ID を受信したり、接続先問い合わせを受信したりする。また、通信 I / F 3 9 は、接続先問い合わせに応じて、対象ユーザの接続先情報を送信する。

[0063] 以上、本開示の一実施形態による音響システムの各構成について詳細に説明した。次に、本実施形態による音響システムの動作処理について図 6 ~ 図 9 を参照して詳細に説明する。

[0064] < 3. 動作処理 >

[3 - 1. 基本処理]

図 6 は、本実施形態による音響システムの基本処理を示すフローチャートである。図 6 に示すように、まず、ステップ S 1 0 3 において、信号処理装置 1 A はサイト A に居るユーザ A の ID を管理サーバ 3 に送信する。信号処理装置 1 A は、ユーザ A の ID を、ユーザ A が所有している R F I D (R a d i o F r e q u e n c y I D e n t i f i c a t i o n) 等のタグから取得してもよいし、ユーザ A の音声から認識してもよい。また、信号処理装置 1 A は、ユーザ A の身体 (顔、目、手等) から生体情報を読み取り、 I D として取得してもよい。

[0065] 一方、ステップ S 1 0 6 において、信号処理装置 1 B も同様にサイト B に居るユーザ B の ID を管理サーバ 3 に送信する。

[0066] 次に、ステップ S 1 0 9 において、管理サーバ 3 は、各信号処理装置 1 から送信されたユーザ ID に基づいてユーザを識別し、識別したユーザの氏名等に、送信元の信号処理装置 1 の I P アドレス等を接続先情報として対応付けて登録する。

[0067] 次いで、ステップ S 1 1 2 において、信号処理装置 1 B は、サイト B に居るユーザ B の位置を推定する。具体的には、信号処理装置 1 B は、サイト B に配された複数のマイクに対するユーザ B の相対位置を推定する。

[0068] 次に、ステップ S 1 1 5 において、信号処理装置 1 B は、推定したユーザ B の相対位置に基づき、サイト B に配された複数のマイクにより收音されたオーディオ信号に対して、ユーザ B の口元に收音位置がフォーカスするよう

マイクアレイ処理を行う。このように、信号処理装置 1 B は、ユーザ B が何らかの発言を行う場合に備える。

[0069] 一方、ステップ S 1 1 8 において、信号処理装置 1 A も同様に、ユーザ A の口元に收音位置がフォーカスするようサイト A に配された複数のマイクにより收音されたオーディオ信号に対してマイクアレイ処理を行い、ユーザ A が何らかの発言を行う場合に備える。そして、信号処理装置 1 A は、ユーザ A の音声（発言）に基づいてコマンドを認識する。ここでは、一例としてユーザ A が「Bさんと話したい」とつぶやいて、信号処理装置 1 A が「ユーザ B に対する発呼要求」コマンドとして認識した場合について説明を続ける。なお、本実施形態によるコマンド認識処理については、後述の [3-2. コマンド認識処理] において詳細に説明する。

[0070] 次に、ステップ S 1 2 1 において、信号処理装置 1 A は、接続先問い合わせを管理サーバ 3 に対して行う。上述したように、コマンドが「ユーザ B に対する発呼要求」であった場合、信号処理装置 1 A は、ユーザ B の接続先情報を問い合わせる。

[0071] 次いで、ステップ S 1 2 5 において、管理サーバ 3 は、信号処理装置 1 A からの接続先問い合わせに応じて、ユーザ B の接続先情報を検索し、続くステップ S 1 2 6 において、検索結果を信号処理装置 1 A に送信する。

[0072] 次に、ステップ S 1 2 7 において、信号処理装置 1 A は、管理サーバ 3 から受信したユーザ B の接続先情報により接続先を同定（決定）する。

[0073] 次いで、ステップ S 1 2 8 において、信号処理装置 1 A は、同定したユーザ B の接続先情報、例えばユーザ B が現在居るサイト B に対応する信号処理装置 1 B の IP アドレスに基づいて、信号処理装置 1 B に対して発呼処理を行う。

[0074] 次に、ステップ S 1 3 1 において、信号処理装置 1 B は、ユーザ A からの呼び出しに応答するか否かをユーザ B に問うメッセージを出力する（呼び出し通知）。具体的には、例えば信号処理装置 1 B は、ユーザ B の周辺に配されるスピーカーから当該メッセージを再生してもよい。また、信号処理装置 1

Bは、ユーザBの周辺に配された複数のマイクから收音したユーザBの音声に基づいて、呼び出し通知に対するユーザBの回答を認識する。

[0075] 次いで、ステップS 1 3 4において、信号処理装置1 Bは、ユーザBの回答を信号処理装置1 Aに送信する。ここでは、ユーザBがOK回答を行い、ユーザA（信号処理装置1 A側）とユーザB（信号処理装置1 B側）の双方向通信が開始される。

[0076] 具体的には、ステップS 1 3 7において、信号処理装置1 Aは、信号処理装置1 Bとの通信を開始すべく、サイトAにおいてユーザAの音声を收音し、音声ストリーム（オーディオ信号）をサイトB（信号処理装置1 B側）に送信する收音処理を行う。なお、本実施形態による收音処理については、後述の「3-3. 收音処理」において詳細に説明する。

[0077] そして、ステップS 1 4 0において、信号処理装置1 Bは、ユーザBの周辺に配された複数のスピーカによりユーザBを内包する音響閉曲面を形成し、信号処理装置1 Aから送信された音声ストリームに基づいて音場再生処理を行う。なお、本実施形態による音場再生処理では、さらに第3の空間（サイトC）の音場を構築し、他の空間に居る他ユーザと通話するユーザに第3の空間への没入感を提供することも可能である。このような音場再生処理については、後述の「4. 第3空間の音場構築」において詳細に説明する。

[0078] なお、上記ステップS 1 3 7～S 1 4 0では、一例として一方向の通信を示したが、本実施形態は双方向通信が可能であるので、上記ステップS 1 3 7～S 1 4 0とは逆に、信号処理装置1 Bで收音処理、信号処理装置1 Aで音場再生処理を行ってもよい。

[0079] 以上、本実施形態による音響システムの基本処理について説明した。これにより、ユーザAは、携帯電話端末やスマートフォン等を所持する必要なく、「Bさんと話したい」とつぶやくだけで、周辺に配された複数のマイクおよび複数のスピーカを利用して他の場所に居るユーザBと通話を行うことができる。続いて、上記ステップS 1 1 8に示したコマンド認識処理について図7を参照して詳細に説明する。

[0080] [3-2. コマンド認識処理]

図7は、本実施形態によるコマンド認識処理を示すフローチャートである。図7に示すように、まず、ステップS203において、信号処理装置1のユーザ位置推定部16は、ユーザの位置を推定する。例えばユーザ位置推定部16は、複数のマイク10から收音した音、イメージセンサにより撮像した撮像画像、およびマイク位置情報DB15に記憶されている各マイクの配置等に基づき、各マイクに対するユーザの相対的な位置、向き、および口の位置を推定してもよい。

[0081] 次に、ステップS206において、信号処理部13は、推定したユーザの相対的な位置、向き、および口の位置に応じて、ユーザを内包する音響閉曲面を形成するマイク群を選出する。

[0082] 次に、ステップS209において、信号処理部13のマイクアレイ処理部131は、選出したマイク群から收音したオーディオ信号に対してマイクアレイ処理を行い、ユーザの口元にフォーカスするようマイクの指向性を制御する。これにより、信号処理装置1は、ユーザが何らかの発言を行う場合に備えることができる。

[0083] 次に、ステップS212において、高S/N化処理部133は、マイクアレイ処理部131により処理したオーディオ信号に対して、残響・ノイズ抑制等の処理を行い、S/N比を向上させる。

[0084] 次に、ステップS215において、認識部17は、高S/N化処理部133から出力されたオーディオ信号に基づいて、音声認識（音声解析）を行う。

[0085] そして、ステップS218において、認識部17は、認識した音声（オーディオ信号）に基づいて、コマンド認識処理を行う。コマンド認識処理の具体的な内容については特に限定しないが、例えば認識部17は、予め登録された（学習した）要求パターンと認識した音声を比較し、コマンドを認識してもよい。

[0086] 上記ステップS218において、コマンドを認識できなかった場合（S2

18/N_o)、信号処理装置1は、ステップS203～S215に示す処理を繰り返す。この際、S203およびS206も繰り返されるので、信号処理部13は、ユーザの移動に応じてユーザを内包する音響閉曲面を形成するマイク群を更新することが可能である。

[0087] [3-3. 收音処理]

次に、図6のステップS137に示す收音処理について、図8を参照して詳細に説明する。図8は、本実施形態による收音処理を示すフローチャートである。図8に示すように、まず、ステップS308において、信号処理部13のマイクアレイ処理部131は、選出/更新した各マイクから收音したオーディオ信号に対してマイクアレイ処理を行い、ユーザの口元にフォーカスするようマイクの指向性を制御する。

[0088] 次に、ステップS312において、高S/N化処理部133は、マイクアレイ処理部131により処理したオーディオ信号に対して、残響・ノイズ抑制等の処理を行い、S/N比を向上させる。

[0089] そして、ステップS315において、通信I/F19は、高S/N化処理部133から出力されたオーディオ信号を、上記ステップS126（図6参照）で同定した対象ユーザの接続先情報で示される接続先（例えば、信号処理装置1B）に送信する。これにより、ユーザAがサイトAで発した音声は、ユーザAの周辺に配された複数のマイクにより收音され、サイトB側に送信される。

[0090] 以上、本実施形態によるコマンド認識処理および收音処理について説明した。続いて、本実施形態による音場再生処理について詳細に説明する。

[0091] <4. 第3空間の音場構築>

上述したように、本実施形態による音場再生処理（図6のステップS140）では、第3の空間（サイトC）の音場を構築し、他の空間に居る他ユーザと通話するユーザに第3の空間への没入感を提供することも可能である。以下、このような第3の空間への没入感を提供するための音場構築の概要について図9を参照して説明する。

[0092] 図9は、本実施形態による第3の空間の音場構築について説明するための図である。図9に示すように、サイトAに居るユーザAおよびサイトBに居るユーザBが通話する際、本実施形態による音響システムは、第3の空間であるサイトCの音場42を各サイトA、Bで構築する。ここで、一例として、サイトA、サイトBおよびサイトCは、各々離れた場所（遠隔地）とする。これにより、例えば東京（サイトB）に居るユーザBが、ユーザAと一緒に旅行で行く予定のイタリア（サイトC）の空間に没入しながら、アメリカ（サイトA）に居るユーザAと通話することができる。

[0093] 具体的には、本実施形態による音響システムは、サイトCにおいて予め測定された音響情報パラメータ（インパルス応答等のパラメータの特性）や、サイトCで収録した音響コンテンツ（環境音）を用いてサイトCの音場42を構築してもよい。なお、このような第3の空間の音響情報パラメータや音響コンテンツは、予め第3の空間において取得され、管理サーバに蓄積されていてもよい。

[0094] （サイトCの音場構築の手法）

ここで、サイトAに居るユーザAおよびサイトBに居るユーザBが通話している際に、各サイトでサイトCの音場を構築する場合の各手法について図10を参照して説明する。図10は、サイトCの音場構築手法を説明するための図である。図10に示す例では、一例として、ユーザAと通話するユーザBが居るサイトBにおいて、サイトCの音場を構築（サイトCへの没入感を提供）する場合について説明する。

[0095] 図10に示すように、手法1としては、ユーザBを内包するよう複数のスピーカー20Bで形成される音響閉曲面40Bの外側にユーザAの声が存在するよう音像を定位し、さらに、ユーザAの音声はサイトCで反響したように感じられるよう音響情報パラメータを用いて加工する。

[0096] ここで、音響閉曲面40の外側にユーザAの音像を定位させる場合、サイトBにおいて、図10に示すように、音響閉曲面40Bの外側に居るユーザAが発する音声は音響閉曲面40Bに交差する時の波面を想定する。そして

、音響閉曲面40Bの内側にその想定波面を創造するよう複数のスピーカー20から再生することで音像を定位させる。

[0097] また、サイトCでユーザAが音声を発したと想定した場合、ユーザAの音声はサイトCの構造物や障害物により、反射音（各材質・構造ごとに異なる反射音）を伴って音響閉曲面40Bに到達する場合もある。よって、本実施形態による音響システムは、予めサイトCで測定された音響情報パラメータ（インパルス応答）を用いてユーザAの音声を加工することで、ユーザAの音声がサイトCで反響したように感じられる音場42をサイトBに構築する。これにより、ユーザBは、サイトCへの没入感をさらに得ることができる。

[0098] 手法2としては、音響閉曲面40の内側に居るユーザBの音声を收音し、当該音声をサイトCの音響情報パラメータを用いて加工し、音響閉曲面40を形成する複数のスピーカー20Bから再生する。すなわち、音響閉曲面40の内側に居るユーザBが、サイトCの音場を、臨場感を持って体感し、サイトCへの没入感にさらに浸り、サイトCの空間の大きさを感じるためには、通話相手の音声の加工（手法1）の他、自らが発した音声の変化も重要である（エコーロケーション）。よって、手法2では、ユーザBが発した音声が、サイトCで反響したように感じられる音場42をサイトBに構築する。これにより、ユーザBは、サイトCの臨場感や、サイトCへの没入感をさらに得ることができる。なお、手法2の具体的な実現方法については、図16A、図16Bを参照して後述する。

[0099] 手法3としては、ユーザBを内包する音響閉曲面40を形成する複数のスピーカー20Bから、サイトCのざわめき声、環境音等の音響コンテンツを再生することで、サイトCの臨場感や、サイトCへの没入感を増加させる。サイトCの音響コンテンツは、予め録音されたものであってもよいし、リアルタイムで收音されてもよい。

[0100] 以上、サイトCへの没入感を提供するための音場構築における3つの手法について図10を参照して説明した。本実施形態による音響システムでは、

上記3つの手法のうち1の手法により音場構築してもよいし、2以上の手法を組み合わせて音場構築してもよい。

[0101] (サイトCの指定)

また、本実施形態において、第3の空間(サイトC)は、ユーザが任意に指定してもよいし、予め設定された場所としてもよい。例えば、サイトAに居るユーザAが、「ユーザBと(第1の対象)と、サイトC(第2の対象)で話したい」と発言すると、周辺に配された複数のマイク10A(図1参照)により当該発言が收音され、信号処理装置1Aによりコマンドとして認識される。

[0102] 次に、信号処理装置1Aは、管理サーバに対して「ユーザB」と通話するための接続先情報、および指定した場所の音場構築用データを要求する。そして、管理サーバは、信号処理装置1Aに対して、接続先情報(ここでは、ユーザBが居るサイトBの信号処理装置1BのIPアドレス等)および音場構築用データ(ここでは、サイトCの音響情報パラメータおよび音響コンテンツ)を送信する。

[0103] また、信号処理装置1Aと信号処理装置1Bとの間で通信が開始された場合(ユーザBがユーザAからの発呼に対してOK回答した場合)、音場構築用データは、信号処理装置1Bにも送信される。これにより、サイトAおよびサイトBでサイトCの音場が構築され、異なるサイトに居るユーザAおよびユーザBは、同じ場所への没入感を共有することができる。

[0104] 以上、第3の空間への没入感を提供するための音場構築の概要について説明した。続いて、第3の空間の音響情報パラメータや音響コンテンツを蓄積する管理サーバの構成について図11を参照して説明する。

[0105] [4-1. 管理サーバの構成]

図11は、本実施形態による管理サーバの他の構成を示すブロック図である。図11に示すように、管理サーバ3'は、管理部32、検索部34、ユーザ位置情報DB35、通信I/F39、音響情報パラメータDB36、および音響コンテンツDB37を有する。管理部32、ユーザ位置情報DB3

5、および通信 I / F 3 9 は、図 5 を参照して上述した通りであるので、ここでの説明は省略する。

[0106] (検索部)

検索部 3 4 は、まず、上述した検索部 3 3 と同様に、信号処理装置 1 からの接続先（発呼先）問い合わせに応じて、ユーザ位置情報 DB 3 5 を参照し、接続先情報を検索する。具体的には、検索部 3 4 は、接続先問い合わせに含まれる対象ユーザの氏名等に基づいて、対応付けられた接続先情報をユーザ位置情報 DB 3 5 から検索して抽出する。

[0107] さらに、検索部 3 4 は、信号処理装置 1 からの音場構築用データの要求に応じて、音響情報パラメータ DB 3 6 から、指定されたサイトの音響情報パラメータを検索して抽出する。また、検索部 3 4 は、信号処理装置 1 からの音場構築用データの要求に応じて、音響コンテンツ DB 3 7 から、指定されたサイトの音響コンテンツを検索して抽出する。

[0108] (音響情報パラメータ)

音響情報パラメータ DB 3 6 は、各サイトで予め測定された音響情報パラメータを記憶する記憶部である。音響パラメータとは、各サイトにおいて、任意の 1 点または複数点（音像を定位させたい位置）からのインパルス応答を測定したものであってもよい。また、インパルス応答を測定する際に、TSP (Time Stretched Pulse) 応答、Swept-Sine 法、M 系列応答等を利用することで、S / N 比が向上する。

[0109] ここで、音響情報パラメータの測定について、図 1 2 を参照して説明する。図 1 2 に示す測定 1 では、図 1 0 を参照して説明した手法 1 において音響閉曲面 4 0 の外側の任意の位置に定位される通話相手ユーザの音声を加工する際に用いられる音響情報パラメータ（第 1 の音響情報パラメータ）の測定を説明する。図 1 2 に示すように、サイト C に配置された、複数の外向き有指向性マイク 1 0 C により、複数のマイク 1 0 C で形成される閉曲面 4 3 の外側の任意の位置に設置された音源（スピーカー 2 0 C）から各マイク 1 0 C まで、どのように伝達されるか（インパルス応答）を測定する。

[0110] 図12に示す例では、測定1において、スピーカー20Cが一つ配置されているが、これに限定されず、複数のスピーカー20Cを閉曲面43の外側に配置し、各スピーカー20からの各マイク10Cへの伝達を各々測定してもよい。これにより、上記手法1において、ユーザAの音源を定位できる個所を増やすことができる。

[0111] また、図12に示す測定2では、図10を参照して説明した手法2において音響閉曲面40の内側に居るユーザ自身の音声を加工する際に用いられる音響情報パラメータ（第2の音響情報パラメータ）の測定を説明する。図12に示すように、サイトCに配置された、複数の外向き有指向性マイク10Cにより、複数のマイク10Cで形成される閉曲面43の内側に設置された音源（スピーカー20C）から出力された音（測定用信号）が、サイトCにおける反射・反響の影響を受けて各マイク10Cまでどのように伝達されるか（インパルス応答）を測定する。図12に示す例では、測定2において、一例としてスピーカー20Cが一つ配置されているが、本実施形態はこれに限定されず、複数のスピーカー20Cが閉曲面43の内側に配置され、各スピーカー20からの各マイク10Cへの伝達が各々測定される。

[0112] （音響コンテンツ）

音響コンテンツDB37は、各サイトで収録された音響コンテンツを記憶する記憶部である。音響コンテンツとは、例えば各サイトにおいて録音（測定）された周囲の音（環境音、ざわめき声等）である。

[0113] 音響コンテンツの測定は、例えば図12の測定3に示すように、サイトCに配置された、複数の外向き有指向性マイク10Cにより、周囲の音を測定（録音）する。また、周囲の音の測定は、時間別、平日／休日別に行ってもよい。これにより、本実施形態による音響システムは、サイトCの時間別、平日／休日別の音場を構築することができる。また、再生環境であるサイトBにおいて、現在の時刻に近い音響コンテンツを再生することもできる。

[0114] なお、図12に示す複数のマイク10Cにより形成される閉曲面43は、聴取環境（再生環境）の音響閉曲面より大きく形成されてもよい。以下、図

13を参照して説明する。図13は、測定環境（ここでは、サイトC）における複数のマイク10Cの配置と、聴取環境（ここでは、サイトB）における複数のスピーカー20Bの配置を比較して示す図である。

[0115] 図13に示すように、ユーザBを内包するよう複数のスピーカー20Bにより形成される音響閉曲面40に対して、サイトCにおいて測定時に用いられる複数のマイク10Cは、音響閉曲面40より大きい閉曲面43を形成するよう配置される。

[0116] また、図4を参照して上述したように、聴取環境（再生環境）のサイトBでは、複数のスピーカー20B-1～20B-12により立体的な音響閉曲面40-1、40-2、40-3が形成される。よって、測定環境のサイトCにおいても、図14に示すように、複数の外側に向けられた有指向性マイク10C-1～10C-12により立体的な閉曲面43-1、43-2、40-3が形成され得る。

[0117] 以上、本実施形態による管理サーバ3'の各構成について詳細に説明した。続いて、上記手法1～手法3（図12参照）を用いてサイトCの音場を構築する聴取環境（再生環境）のサイトB側の制御について説明する。サイトB側では、信号処理装置1Bの音場再生信号処理部135（図3参照）により、最適な音場が形成される。以下、図15を参照し、上記手法1～手法3を実現して音場を構築する音場再生信号処理部135の構成について具体的に説明する。

[0118] [4-2. 音場再生信号処理部の構成]

図15は、サイトCへの没入感を提供するよう音場構築を行う音場再生信号処理部135の構成について説明するためのブロック図である。また、図15では、信号処理装置1Bのうち、ここでの説明に関係する主要な構成を示し、他の構成は省略している。

[0119] 図15に示すように、音場再生信号処理部135は、Convolution（畳み込み積分）部136、ハウリング抑制部137、139、Matrix Convolution（畳み込み行列）部138として機能する

。

[0120] (Convolution部)

Convolution部136は、上記手法1（ユーザAの音像定位、ユーザAの反響音声加工）を実現する機能を有する。具体的には、Convolution部136は、通信I/F19を介してサイトAの信号処理装置1Aから取得（受信）したオーディオ信号b（ユーザAの音声）を、サイトCの音響情報パラメータc（第1の音響情報パラメータ）を用いて、各出力スピーカー別にレンダリングする。また、この際、Convolution部136は、ユーザAの音像を定位させる位置のパラメータaを考慮し、定位位置に応じたサイトCの音響情報パラメータc（インパルス応答）を用いてもよい。なお、ユーザAの音像を定位させる位置のパラメータaは、信号処理装置1Aまたは管理サーバ3'から通信I/F19を介して送信されたものであってもよいし、ユーザBの指示に基づいて信号処理装置1Bで算出されたものであってもよい。また、Convolution部136は、サイトCの音響情報パラメータc（インパルス応答）を、通信I/F19を介して管理サーバ3'から取得してもよい。

[0121] そして、Convolution部136は、図15に示すように、各出力スピーカー（ユーザBを内包する音響閉曲面40Bを形成する複数のスピーカー20B）の出力バッファに対して、上記信号処理を行ったオーディオ信号を書き込む。

[0122] (ハウリング抑制部)

ハウリング抑制部137、139は、フィードバックによるハウリングを避けるため、必要に応じて、図15に示すようにマイクのアンプ・ADC部11の後段およびスピーカーのDAC・アンプ部23の前段に、両者が連携処理できるよう設けられる。上述したように手法2では、ユーザBの周辺に配された複数のマイク10Bから收音した音を、音響情報パラメータ（インパルス応答）を用いてレンダリングし、ユーザBの周辺に配された複数のスピーカー20Bから再生する。この時、マイクとスピーカーの位置が近いの

で、両者の作用で過剰な発振動作が生じてしまう恐れがある。このため、図 15 に示す例では、ハウリング抑制部 137、139 を設け、ハウリング抑制処理を実施している。なお、音場再生信号処理部 135 は、上記過剰な発振動作を防止するために、ハウリング抑制部 137、139 の他、エコーキャンセルを有していてもよい。

[0123] (Matrix Convolution 部)

Matrix Convolution 部 138 は、上記手法 2 (ユーザ B の反響音声加工) を実現する機能を有する。具体的には、Matrix Convolution 部 138 は、サイト B に配された複数のマイク 10 B により收音されたオーディオ信号 (音響閉曲面 40 B 内で発生した音) を、サイト C の音響情報パラメータ c (第 2 の音響情報パラメータ; インパルス応答群) を用いて、各出力スピーカー別にレンダリングする。これにより、サイト B の音響閉曲面 40 B 内で発生した音、例えばユーザ B 自身の音声 がサイト C で反響したように感じられる音場をサイト B で構築するためのオーディオ信号が出力される。

[0124] ここで、本実施形態による手法 2 の実現方法について、図 16 A および図 16 B を参照して具体的に説明する。図 16 A は、サイト C におけるインパルス応答の測定について説明するための図である。図 16 A に示すように、まず、サイト C に配置され、閉曲面 43 の外側を向いた各スピーカー 20 C から、同じくサイト C に配置され、閉曲面 43 の外側を向いた各マイク 10 C までのインパルス応答が測定される。

[0125] 具体的には、閉曲面 43 上の単一のスピーカーから、同じく閉曲面 43 上の複数マイク群へのインパルス応答が測定される。このインパルス応答は、周波数軸で考えると、サイト C の構造物/障害物などの空間音響の影響を受けた伝達関数としても考えることができる。

[0126] ここで、図 16 A に示す例では、閉曲面 43 上の各マイク・スピーカの位置を R_1 、 R_2 、 \dots 、 R_N と表現する。そして、図 16 A に示すように、 R_1 に配置されたスピーカー (SP) から、 R_1 に配置されたマイク、 R_2

に配置されたマイク・・・ R_N に配置されたマイクまでの各伝達関数が測定される。次いで、 R_2 に配置されたスピーカーから、 R_1 に配置されたマイク、 R_2 に配置されたマイク・・・ R_N に配置されたマイクまでの各伝達関数が測定される。

[0127] 次に、 R_1 に位置するスピーカーから R_1 に位置するマイクまでの伝達関数を R_{11} 、 R_1 に位置するスピーカーから R_2 に位置するマイクまでの伝達関数を R_{12} 、と記述すると、図16Aに示す式(1)に示すように、伝達関数 R を用いた行列表現が可能となる。

[0128] 当該行列データは、音響情報パラメータとして、管理サーバ3'等に蓄積され、サイトBで、サイトCの音場を構築する際に使用される。続いて、サイトBで行列データを用いてサイトCの音場を構築する場合について、図16Bを参照して説明する。

[0129] 図16Bは、Matrix Convolution部138によるインパルス応答群を用いた演算について説明するための図である。図16Bに示す例では、サイトB(再生環境)側において、サイトCにおける測定時と略同じ大きさ・形状の閉曲面を想定する。また、サイトBに配置される複数のマイク10Bおよび複数のスピーカー20Bの数も、サイトCにおける測定時と同じであって、配置位置もサイトCにおける測定時と同じ R_1 、 R_2 、・・・ R_N である場合を想定する。ただし、図16Bに示すように、複数のマイク10Bおよび複数のスピーカー20Bは、音響閉曲面40Bの内側を向いている。

[0130] また、図16Bに示すように、サイトBでの R_1 、 R_2 、・・・ R_N の位置にある各マイクで收音される周波数軸表現を、 V_1 、 V_2 、・・・ V_N とする。また、サイトBでの R_1 、 R_2 、・・・ R_N の位置にある各スピーカーから出力(再生)される出力信号(オーディオ信号)を、 W_1 、 W_2 、・・・ W_N とする。

[0131] この場合、サイトBの音響閉曲面40Bの内側で発生した音(ユーザAの音声や物音)の波面が、音響閉曲面40Bに到達し、 R_1 、 R_2 、・・・ R

Nの位置にある内側向きマイク10Bに收音され、それぞれのマイク10BにてV1、V2、・・・VNの收音信号が取得される。

[0132] そして、Matrix Convolution部138は、V1、V2、・・・VNの信号群（マイク入力）、および図16Aを参照して説明した伝達関数群の行列（式1）を用いて、図16Bに示す式2を実行し、各スピーカー20Bからの出力W1、W2、・・・WNを算出する。

[0133] 以上説明したように、Matrix Convolution部138は、複数のマイク10Bにより收音したオーディオ信号（V1、V2、・・・VN）に対して、サイトCの音響情報パラメータ（伝達関数群）を用いて信号処理を行う。また、Matrix Convolution部138は、図15に示すように、各出力スピーカーの出力バッファに対して、上記信号処理を行ったオーディオ信号（W1、W2、・・・WN）を加算する。

[0134] （音響コンテンツの加算）

図15に示すように、音場再生信号処理部135は、通信I/F19を介して管理サーバ3'から受信したサイトCの音響コンテンツdを、各出力スピーカーの出力バッファに加算するよう処理することで、上記手法3を実現する。

[0135] 以上、本実施形態による信号処理装置1Bの音場再生信号処理部135の構成について詳細に説明した。次に、サイトBにおいてサイトCの音場を構築する際の音場再生処理について図17を参照して具体的に説明する。

[0136] [4-3. 音場再生処理]

図17は、本実施形態による音場再生処理を示すフローチャートである。図17に示すように、まず、ステップS403において、信号処理装置1Bのユーザ位置推定部16（図3参照）は、ユーザBの位置を推定する。例えばユーザ位置推定部16は、複数のマイク10Bから收音した音、イメージセンサにより撮像した撮像画像、およびスピーカー位置情報DB21に記憶されている各スピーカーの配置等に基づき、各スピーカー20Bに対するユーザBの相対的な位置、向き、口の位置および耳の位置を推定してもよい。

- [0137] 次いで、ステップS406において、信号処理部13は、推定したユーザBの相対的な位置、向き、口の位置および耳の位置に応じて、ユーザを内包する音響閉曲面を形成するマイク群・スピーカー群を選出する。
- [0138] 次に、ステップS407において、信号処理部13の音場再生信号処理部135は、Convolution部136により、受信したオーディオ信号b（サイトAで収録されたユーザAの音声）に対して図10に示す手法1の処理を行う。具体的には、Convolution部136は、図15に示すように、サイトAの信号処理装置1Aから受信したオーディオ信号bを、サイトCの音響情報パラメータc（第1の音響情報パラメータ）を用いて、選出された各出力スピーカー別にレンダリングする。そして、Convolution部136は、選出された各出力スピーカーの出力バッファに対して、手法1の処理を行ったオーディオ信号を書き込む。
- [0139] 次いで、ステップS409において、音場再生信号処理部135は、Matrix Convolution部138により、選出したマイク群によりサイトBで収録したユーザBの音声に対して図10に示す手法2の処理を行う。具体的には、Matrix Convolution部138は、ユーザBを内包する音響閉曲面を形成するマイク群（複数のマイク10B）により収録されたオーディオ信号をサイトCの音響情報パラメータc（第2の音響情報パラメータ）を用いて各出力スピーカー別にレンダリングする。そして、Matrix Convolution部138は、選出された各出力スピーカーの出力バッファに対して、手法2の処理を行ったオーディオ信号を加算する。
- [0140] 次に、ステップS411において、音場再生信号処理部135は、図10に示す手法3の処理として、サイトCの音響コンテンツdを、選出された各出力スピーカーの出力バッファに加算する。
- [0141] そして、ステップS415において、信号処理装置1Bは、各出力バッファの内容を、DAC・アンプ部23を介して、上記ステップS406で選出されたスピーカー群から出力する。

- [0142] 以上説明したように、本実施形態による音響システムでは、サイトAで収録されたユーザAの音声、サイトCで測定された第1の音響情報パラメータを用いてレンダリングされ、サイトCでの反響を伴って、サイトBの複数のスピーカー20Bから再生される。また、サイトBで収録されたユーザB自身の音声、サイトCで測定された第2の音響情報パラメータを用いてレンダリングされ、サイトCでの反響を伴って、サイトBの複数のスピーカー20Bから再生される。さらに、サイトCで収録された音響コンテンツが、サイトBの複数のスピーカー20Bから再生される。
- [0143] これにより、本実施形態による音響システムは、一のサイト（ここではサイトB）において、他のサイト（ここではサイトA）と相互連携（例えば通話）させる際に、第3の空間（ここではサイトC）への没入感を提供することができる。ユーザBは、ユーザAと共にサイトCに居るような音場感を得ることができ、より豊かな臨場感に浸ることができる。
- [0144] また、音場再生信号処理部135は、ユーザBの周辺に配されたスピーカー群を用いて、受信したオーディオ信号（ユーザAの音声）の音像を制御することも可能である。例えば、複数のスピーカーによりアレイスピーカー（ビームフォーミング）を形成することで、音場再生信号処理部135は、ユーザBの耳元でユーザAの音声を再現したり、ユーザBを内包する音響閉曲面の外側にユーザAの音像を再現したりすることが可能である。
- [0145] なお、上記S403およびS406を継続的に行うことで、信号処理部13は、ユーザBの移動に応じてユーザBを内包する音響閉曲面を形成するスピーカー群を更新することが可能である。以下、図18Aおよび図18Bを参照して具体的に説明する。
- [0146] 図18Aは、サイトBで構築される音場42が固定されている場合について説明するための図である。図18Aに示すように、まず、ユーザBを内包する音響閉曲面40を形成するために複数のスピーカー20Bが選出され（ステップS403、S406）、サイトCへの没入感を提供する音場42が構築されている場合を想定する。この場合に、ユーザBが部屋の中を移動し

たり部屋から出たり等して、音響閉曲面40から出てしまうと、音場42から外れてしまい、サイトCへの没入感が得られない。

[0147] そこで、上述したように、上記S403およびS406を継続的に行い、ユーザBの移動に応じてユーザBを内包する音響閉曲面を形成するスピーカー群を更新する。図18Bは、サイトBで構築される音場42が流動する場合について説明するための図である。

[0148] 図18Bに示すように、ユーザBの移動に応じて、新たにユーザBを内包する音響閉曲面40'を形成するスピーカー群（スピーカー20B'）が選出（更新）され、更新された複数のスピーカー20B'により、新たに音場42'が構築される。

[0149] 以上、本実施形態による音響システムの各動作処理について詳細に説明した。続いて、本実施形態の補足について説明する。

[0150] <5. 補足>

[5-1. コマンド入力の変形例]

上記実施形態では、音声にてコマンドを入力していたが、本開示による音響システムのコマンド入力方法は音声入力に限定されず、他の入力方法であってもよい。例えば、本実施形態による信号処理装置1は、ユーザの周辺に配される各スイッチ（操作入力部の一例）に対するユーザ操作を検出し、発呼要求等のコマンドを認識してもよい。また、この場合、信号処理装置1は、発呼先の指定（対象ユーザの氏名等）や、没入対象の場所の指定（地名等）も、ユーザの周辺に配されるタッチパネル等（操作入力部の一例）により受け付けることが可能である。

[0151] また、信号処理装置1の認識部17は、ユーザの周辺に配される撮像部により撮像された画像や、赤外線／熱センサにより取得した検知結果に基づいて、ユーザのジェスチャーを解析し、コマンドとして認識してもよい。例えば、認識部17は、ユーザが電話をかけるジェスチャーを行った場合、発呼要求コマンドとして認識する。また、この場合、信号処理装置1は、発呼先の指定（対象ユーザの氏名等）や、没入対象の場所の指定（地名等）を、ユ

ーザの周辺に配されるタッチパネル等から受け付けてもよいし、音声解析に基づいて判断してもよい。

[0152] また、他のサイトに居る他のユーザと通話中に（ユーザ周辺の複数のマイク10から他のユーザの音声再生されている場合に）、音が聴こえにくいと感じた場合、ユーザはジェスチャーにより再生音の制御を要求してもよい。具体的には、例えば認識部17は、広げた手を耳に近付けるジェスチャーや、両手を頭の上に近付けてウサギの耳を真似するジェスチャーを、音量アップのコマンドとして認識してもよい。

[0153] 以上説明したように、本開示による音響システムのコマンド入力方法は、音声入力に限定されず、スイッチ操作やジェスチャー入力等であってもよい。

[0154] [5-2. 他のコマンド例]

上記実施形態では、所定の対象として人物が指定され、発呼要求（通話要求）をコマンドとして認識する場合について説明したが、本開示による音響システムのコマンドは発呼要求（通話要求）に限定されず、他のコマンドであってもよい。例えば、信号処理装置1の認識部17は、所定の対象として指定された場所、建物、番組、曲等をユーザが居る空間で再現するコマンドを認識してもよい。

[0155] また、本実施形態による音響システムは、ユーザが居る空間に、他の空間をリアルタイムで再現してもよいし、指定された場所、建物等の過去の空間（例えば、有名な歌劇場で過去に行われた名コンサート等）を再現してもよい。

[0156] [5-3. 大空間から小空間への変換]

ここで、上述した実施形態では、サイトB（再生環境）側の閉曲面と、サイトC（測定環境）側の閉曲面が、略同じ大きさ・形状であることを想定しているが、本実施形態はこれに限定されない。例えば、本実施形態は、測定環境側の閉曲面より、再生環境側の閉曲面が小さい場合であっても、再生環境において測定環境の音場（空間の広がり）を再現することが可能である。

[0157] このような大空間から小空間への変換処理は、図10を参照して上述した手法1、3を実施する前に、信号処理装置1において、受信したオーディオ信号（ユーザAのオーディオ信号や音響コンテンツ等）に対して行ってもよい。また、本実施形態による音響システムは、かかる変換処理をリアルタイムで行うことで、測定環境側と再生環境側の各スピーカーおよびマイクの位置の対応関係の不整合問題を解決することができる。

[0158] 具体的には、例えば音場再生信号処理部135は、特許第4775487号で開示されている伝達関数を用いた信号処理を用いてもよい。特許第4775487号では、測定環境の音場において伝達関数（インパルス応答の測定データ）を求め、さらに再現環境において伝達関数に基づく演算処理を施された音声信号を再生し、再現環境で測定環境の音場（残響、音像定位等）を再現している。以下、図19A～図19Cを参照して伝達関数（インパルス応答の測定データ）を用いた信号処理について説明する。

[0159] 図19Aは、測定対象空間における測定について説明するための図である。まず、図19Aに示すように、測定対象空間（大空間）において、大きな閉曲面Pを形成するM個のマイクが配置され、Mch用（M個のスピーカー出力チャンネル用）の測定が行われる。また、各M個のマイクの位置をP1、P2、・・・PMとする。そして、閉曲面Pの外部に配置されたスピーカー（SP）から測定信号を出力し、スピーカーからP1、P2、・・・PMに配置された各マイクまでのインパルス応答が測定される。このように測定したインパルス応答（伝達関数）を図19Aの式（3）に示す。

[0160] 次に、無響室での測定について図19Bを参照して説明する。図19Bに示すように、無響室において、大きな閉曲面Pを形成するM個のスピーカーが配置され、閉曲面Pの内側には小さな閉曲面Qを形成するN個のマイクが配置され、Nch用（N個のスピーカー出力チャンネル用）の測定が行われる。ここで、各M個のスピーカーの位置を、図19Aと同様の位置P1、P2、・・・PMとする。また、各N個のマイクの位置をQ1、Q2、・・・QNとする。

- [0161] そして、P 1 に配置されたスピーカーから、図 19 A の P 1 に配置されたマイクにより收音された音（測定信号）が出力され、Q 1、Q 2、・・・Q N に配置された各マイクまでのインパルス応答が測定される。次に、P 2 に配置されたスピーカーから、図 19 A の P 2 に配置されたマイクにより收音された音（測定信号）が出力され、Q 1、Q 2、・・・Q N に配置された各マイクまでのインパルス応答が測定される。このように、M 個のスピーカーの各々から、Q 1、Q 2、・・・Q N に配置された各マイクまでのインパルス応答が全て測定される。
- [0162] このように測定した M 個の縦ベクトルを、 $M \times N$ 行列演算することで、N 個の出力に変換することができる。すなわち、このように測定したインパルス応答（伝達関数）を、図 19 B の式（4）に示すように行列化（伝達関数群の行列生成）することで、大空間（M c h 用係数）から小空間（N c h 用係数）への変換を実現する。
- [0163] 次に、再生対象空間（小空間）での再現について図 19 C を参照して説明する。図 19 C に示すように、再生対象空間において、ユーザ B を内包する小さな閉曲面 Q を形成する N 個のスピーカーが配置されている。ここで、各 N 個のスピーカーの位置を、図 19 B と同様の位置 Q 1、Q 2、・・・Q N とする。
- [0164] この場合、Q 1、Q 2、・・・Q N に配置される各スピーカーから、受信したオーディオ信号（例えばユーザ A の音声；音声信号 S）を出力する場合、各スピーカーの出力は、図 19 C に示す式（5）により求められる。式（5）は、上記式（3）および式（4）に示すインパルス応答（伝達関数）を用いた演算である。
- [0165] このように、例えば閉曲面 Q の外側にユーザ A の音像を定位させる場合、図 19 C に示すように、閉曲面 Q の外側に居るユーザ A が発する音声閉曲面 Q に交差する時の波面を想定し、閉曲面 Q の内側にその想定波面を創造する。この際、測定対象空間におけるマイクの数と、再生対象空間におけるスピーカーの数との不整合を、上記式（5）により変換することで、本実施形

態による音響システムは、小さな閉曲面Qにおいて、大きな閉曲面Pの音場を再現することができる。

[0166] [5-4. 映像構築]

さらに、上記実施形態では、第3の空間への没入感の提供を音場構築（音場再生処理）により実現しているが、本開示による音響システムはこれに限定されず、併せて映像構築を用いてもよい。

[0167] 例えば信号処理装置1は、第3の空間（サイトC）に配された複数のイメージセンサで撮像した映像を所定のサーバから受信し、サイトAに配された複数のマイクで収録した音声をサイトBで再生する際、サイトCの映像を再生してサイトCの空間を再現してもよい。

[0168] 映像の再生は、例えばホログラム再生による空間投影であってもよいし、部屋にあるテレビジョン、ディスプレイ、ユーザが装着するヘッドマウントディスプレイで再生してもよい。このように、音場構築と併せて映像構築も行うことで、ユーザは、第3の空間への没入感を得ることができ、より臨場感に浸ることができる。

[0169] [5-5. 他のシステム構成例]

図1～図2を参照して説明した上記実施形態による音響システムのシステム構成は、発呼側（サイトA）および着呼側（サイトB）の両者とも、ユーザの周辺に複数のマイクやスピーカーが配され、信号処理装置1A、1Bにより信号処理されている。しかし、本実施形態による音響システムのシステム構成は図1～図2に示す構成に限定されず、例えば図20に示すような構成であってもよい。

[0170] 図20は、本実施形態による音響システムの他のシステム構成を示す図である。図13に示すように、本実施形態による音響システムは、信号処理装置1、通信端末7、および管理サーバ3が、ネットワーク5を介して接続している。

[0171] 通信端末7は、携帯電話端末やスマートフォンといった通常の単数のマイクおよび単数のスピーカーを有し、本実施形態による複数のマイクおよび複

数のスピーカーが配される高機能なインターフェース空間に対して、レガシーなインターフェースである。

[0172] 本実施形態による信号処理装置 1 は、通常の通信端末 7 と接続し、通信端末 7 から受信した音声をユーザの周辺に配される複数のスピーカーから再生することができる。また、本実施形態による信号処理装置 1 は、ユーザの周辺に配される複数のマイクから收音したユーザの音声を、通信端末 7 に送信することができる。

[0173] 以上説明したように、本実施形態による音響システムによれば、周辺に複数のマイクおよび複数のスピーカーが配された空間に居る第 1 のユーザと、通常の通信端末 7 を所持する第 2 のユーザとの通話を実現することができる。すなわち、本実施形態による音響システムの構成は、発呼側および着呼側の一方が、本実施形態による複数のマイクおよび複数のスピーカーが配される高機能なインターフェース空間であってもよい。

[0174] [5-6. 自律型音響システム]

上記実施形態では、図 1～図 3 を参照して説明したように、信号処理装置 1 により、ユーザの周辺に配された複数のマイク 10 および複数のスピーカー 20 の入出力が制御されているが、本開示による音響システムの構成はこれに限定されない。例えば、ユーザの周辺に、自律型のマイクやスピーカーのデバイスが複数配され、各デバイス同士が通信し、各々の判断により、ユーザを内包する音響閉曲面を形成し、上述した音場構築を実現するようにしてもよい。以下、このような自律型音響システムについて図 21～図 24 を参照して具体的に説明する。なお、ここで説明する自律型音響システムでは、一例として、1 のマイク 10 および 1 のスピーカー 20 を有するデバイス 100 が、ユーザの周辺に複数配されている場合について説明する。

[0175] (システム構成)

図 21 は、本開示による自律型音響システムのシステム構成の一例を示す図である。図 21 に示すように、本開示による自律型音響システムは、複数のデバイス 100 (100-1～100-4)、管理サーバ 3、ユーザ ID

DB 6、サービスログDB 8、およびユーザ個人DB 9を有する。また、図 21 に示すように、管理サーバ3、ユーザIDDB 6、サービスログDB 8、およびユーザ個人DB 9は、ネットワーク5を介して接続されている。

[0176] ・デバイス

複数のデバイス100 (100-1~100-4) は、部屋、家、ビル、屋外、地域、国等の世界の至る所に配されている。図21に示す例では、デパートや駅等の公共の場の壁および床に複数のデバイス100が配されている場合について示す。また、複数のデバイス100 (100-1~100-4) は、無線/有線により互いに通信可能であって、互いに自デバイスのCapabilityを報知し合っている。また、複数のデバイス100 (100-1~100-4) のうち、少なくとも1のデバイス100 (例えばデバイス100-1) は、ネットワーク5とアクセス可能である。また、デバイス100は、各々マイク10およびスピーカー20を有する。なお、本実施形態によるデバイスの構成については、図22を参照して後述する。

[0177] ・情報の報知

上述したように、複数のデバイス100 (100-1~100-4) は、互いに自デバイスのCapability (特性情報) を報知し合っている。報知する特性情報には、デバイスID、自デバイスが提供可能なサービス、デバイスの持ち主ID、デバイスの貸出Attribute (属性) などが含まれる。ここで、持ち主IDとは、デバイス100の持ち主 (設置者) IDであり、図21に示す各デバイス100 (100-1~100-4) は、個人や法人により各々設置されたことが想定される。また、デバイスの貸出属性とは、デバイス設置者である個人や法人により予め貸出 (利用) が許可されたサービスの属性を示す情報である。

[0178] また、情報の報知は、定期的またはオンデマンドで、自律分散的に行われる。また、本実施形態による情報報知方法は、一般的にメッシュネットワーク構成の方法として知られる手順を利用してもよい (IEEE 802.11sのビーコニング等)。

[0179] また、デバイス100は、複数種類の通信I/F（インターフェース）を備えていてもよい。この場合、各デバイス100は、どの通信I/Fを使用した場合に、どのデバイスと通信可能であるかを定期的にチェックし、最も多くのデバイスと直接通信できる通信I/Fを高プライオリティで起動する。

[0180] また、各デバイス100は、自デバイス周辺のデバイスからの報知情報を、無線I/Fを使って数ホップ先のデバイスまで転送してもよいし、ネットワーク5を介して他装置に送信してもよい。

[0181] ・管理サーバ

管理サーバ3は、図2および図5を参照して説明した通り、各ユーザの絶対位置（現在位置）を管理する。または、管理サーバ3は、図11を参照して説明した第3の空間の音響情報パラメータ等を蓄積する管理サーバ3'であってもよい。

[0182] ・サービスログDB

サービスログDB8は、サービス内容、当該サービスの提供に寄与したデバイス100、およびサービスを提供したユーザを関連付けて記憶する記憶部である。これにより、サービスログDB8は、どのデバイスがどのようなサービス提供に用いられたか、また、どのユーザにどのようなサービスを提供したか等を把握することができる。

[0183] また、サービスログDB8に記憶されるサービスログは、後にサービスを利用したユーザ（利用者）への課金情報として用いられてもよいし、サービス提供に寄与したデバイス100の設置者（個人／法人）に対するキックバック情報として用いられてもよい。ここで、キックバック情報とは、デバイス100のサービス提供への寄与率（頻度）等に応じて、デバイス100の持ち主（設置者）に、利用料金の一部を提供する際に用いられる情報である。また、サービスログDB8に記憶されるサービスログは、ユーザの行動のメタデータとしてユーザ個人DB9に送信されてもよい。

[0184] ・ユーザ個人DB

ユーザ個人DB 9は、ユーザ所有のデータとして、サービスログDB 8から送信されたユーザの行動のメタデータを記憶する。ユーザ個人DB 9に記憶されたデータは、様々な個人化サービス等で利用され得る。

[0185] ・ユーザIDDB

ユーザIDDB 6は、登録されたユーザのID（氏名、暗証番号、生体情報等）と、当該ユーザに提供が許可されているサービスとを関連付けて記憶する記憶部である。ユーザIDDB 6は、デバイス100がユーザ認証を行う際に利用される。

[0186] 以上、本実施形態による自律型音響システムのシステム構成について図21を参照して説明した。続いて、本実施形態によるデバイス100（信号処理装置）の構成について図22を参照して説明する。

[0187] （デバイスの構成）

図22は、本実施形態によるデバイス100の構成を示すブロック図である。図22に示すように、デバイス100は、マイク10、アンプ・ADC部11、信号処理部200、認識部17、同定部18、通信I/F19、ユーザ認証部25、ユーザ位置推定部16、DAC・アンプ部23、およびスピーカー20を有する。マイク10、アンプ・ADC部11、認識部17、同定部18、通信I/F19、ユーザ位置推定部16、DAC・アンプ部23、およびスピーカー20は、図3を参照して説明したので、ここでの説明は省略する。

[0188] ・信号処理部

信号処理部200は、高S/N化処理部210および音場再生信号処理部220を有する。高S/N化処理部210は、図3に示す高S/N化処理部133と同様に、アンプ・ADC部11から出力されたオーディオ信号に対して、明瞭度が高くS/N比がよいモノラル信号となるよう処理する機能を有する。具体的には、高S/N化処理部210は、音源を分離し、残響・ノイズ抑制を行う。高S/N化処理部210により処理されたオーディオ信号は、認識部17に出力されてコマンド認識のために音声解析されたり、通信

I/F 19を介して外部装置に送信されたりする。

[0189] 音場再生信号処理部220は、スピーカー20から再生するオーディオ信号に関する信号処理を行い、ユーザの位置付近に音場が定位するよう制御する。また、音場再生信号処理部220は、隣接する他のデバイス100と連携し、ユーザを内包する音響閉曲面を形成するよう、スピーカー20からの出力内容（オーディオ信号）を制御する。

[0190] ・ユーザ認証部

ユーザ認証部25は、ユーザが所持するRFID等のタグから取得したユーザIDに基づいて、通信I/F19を介してネットワーク5上のユーザIDDB6に問い合わせを行い、ユーザ認証を行う。例えば、ユーザ認証部25は、取得したユーザIDが、ユーザIDDB6に予め登録されたIDと一致する場合、サービス提供が許可されたユーザとして認証する。

[0191] 以上、本実施形態によるデバイス100の構成について詳細に説明した。続いて、本実施形態による自律型音響システムの動作処理について図23を参照して説明する。

[0192] （動作処理）

図23は、本実施形態による自律型音響システムの動作処理を示すフローチャートである。図23に示すように、まず、ステップS503において、デバイス100は、事前準備処理を行う。具体的には、デバイス100は、他のデバイス100と、上述した特性情報の報知をし合って、連携可能な（信頼できる）デバイスを確認する。

[0193] 例えば、デバイス100-1は、隣接するデバイス100-2から受信した特性情報に含まれる持ち主IDや、貸出属性等に基づいて、隣接デバイス100-2が信頼に値するか否かを確認してもよい。信頼に値すると確認し合ったデバイス同士は、自デバイスのアクチュエータを動作させ、隣接デバイスのセンサでその出力結果をキャプチャする等して、デバイス同士の特性を組み合わせでどのような連携が可能であるかを把握することができる。このような確認手順は、定期的に行われてもよい。また、かかる確認手順を通

じて、隣接する複数のデバイス100（100-1～100-4）が配された空間で、どのようなサービスが提供可能であるかを、各デバイス100はゆるやかに把握することが可能である。

[0194] 次いで、ステップS506において、複数のデバイス100（100-1～100-4）が配された空間にユーザが登場した場合、デバイス100はユーザ認証を行う。例えば、図21に示すように、ユーザがRFID等のタグ60を所持する場合、周辺に配されたデバイス100-1～100-4は、タグ60から報知される特性情報を受信し、ユーザの登場を検出してもよい。そして、ユーザの登場を検出すると、各デバイス100は、タグ60から報知される特性情報に含まれるユーザIDに基づいて、ネットワーク5上のユーザIDDB6に問い合わせ、サービスを提供してもよいユーザか否かを認証する。

[0195] なお、ユーザがタグ60を所持しない場合、デバイス100-1～100-4は、各種センサ（マイク、カメラ、人感センサ、熱センサ等）によりユーザの登場を検出してもよい。また、デバイス100-1～100-4は、各種センサによる検知結果を解析することで、当該ユーザのID（生体情報等）を抽出してもよい。

[0196] また、図21に示す例では、デバイス100-1～100-4のうち、デバイス100-1が、ユーザIDDB6へのアクセス経路を保有している。この場合、ユーザIDを取得したデバイス100-2、100-3、または100-4は、当該ユーザIDをデバイス100-1に送信し、デバイス100-1がユーザIDDB6へ問い合わせを行い、ユーザ認証を行ってもよい。このように、複数のデバイス100-1～100-4の全てがユーザIDDB6へのアクセスを保持している必要はない。

[0197] また、デバイス100-1～100-4のうち1のデバイス100で行われたユーザ認証の結果は、周辺の他のデバイス100と共有され、デバイス100-1～100-4は、当該ユーザにサービスを提供できることを把握する。

- [0198] 次に、ステップS509において、デバイス100は、ユーザからのコマンド（サービス要求）を認識する。ここで、デバイス100は、認証されたユーザへ提供可能なサービスに関する情報を、タグ60に通知してもよい。タグ60は、スピーカまたは表示部等の各出力手段（不図示）により、この場所でどのようなサービスが受けられるかをユーザに通知することができる。また、タグ60は、ユーザからのコマンド入力（マイク、ジャイロ、キータッチ等）により、現在ユーザが所望しているサービスを特定し、周辺のデバイス100-1~100-4に通知する。
- [0199] なお、ユーザがタグ60を所持しない場合、デバイス100-1~100-4は、各種センサ（マイク、カメラ、人感センサ、熱センサ等）によりユーザの音声やジェスチャーを解析し、ユーザが所望しているサービスを認識してもよい。
- [0200] ここでデバイス100により認識されるコマンドは、上述したような発呼要求（通話要求）の他、所定の対象として指定した場所、建物、番組、曲等の再現を要求するコマンド等であってもよい。
- [0201] 次いで、ステップS512において、デバイス100-1~100-4は、要求されたサービスが当該ユーザに許可されているサービスである場合、当該サービスの提供を開始する。具体的には、例えばデバイス100-1~100-4は、センサ（例えばマイク10）やアクチュエータ（例えばスピーカ20）の動作を開始し、さらにデバイス同士の通信路を動作状態とする。また、デバイス100-1~100-4は、提供するサービスの種類・利用可能な通信リソース量等に基づいて、デバイス同士で連携し、自デバイスの動作を決定してもよい。
- [0202] また、デバイス100は、複数種類の通信I/Fを有している場合、提供する情報量等に基づいて、必要に応じてトラヒックの伝送に利用する通信I/Fを動作させてもよい。また、デバイス100は、必要な分だけ動作Duty Cycleを上げて、省電力モードを部分的に解除してもよい。さらに、デバイス100は、通信に利用する送受信時間帯を複数のデバイス間で

設定し合う等して、帯域の安定供給が可能な状態へと遷移させてもよい（IEEE 802.11sの予約アクセスの立ち上げ等）。

[0203] 次に、ステップS515において、デバイス100は、ユーザによりサービス終了が指示された場合、サービスの提供を終了する。具体的には、例えばデバイス100は、センサ（例えばマイク10）やアクチュエータ（例えばスピーカー20）の動作を終了させ、また、デバイス同士の通信路を停止状態とする。

[0204] 次に、ステップS518において、デバイス100は、自デバイスが今回のサービス提供に寄与した内容をサービスログDB8に通達する。また、デバイス100は、サービスを提供したユーザ（認証したユーザ）の情報も併せてサービスログDB8に通達してもよい。

[0205] 以上、本実施形態による自律型音響システムの動作処理について図23を参照して具体的に説明した。以下、本実施形態による自律型音響システムの追加説明を行う。

[0206] （サービスの継続）

本実施形態による自律型音響システムは、認証したユーザが歩く等して、場所を移動した場合でも、サービスを提供する（動作する）デバイス100を変更することで、継続してユーザにサービスを提供することが可能である。かかる変更処理は、例えばユーザが所有するタグ60からの電波強度や、各デバイスが有するセンサ（マイク、カメラ、人感センサ等）からの入力信号等に基づいて行われる。以下、図24を参照して具体的に説明する。

[0207] 図24は、本実施形態による自律型音響システムにおいて、ユーザの移動に応じた動作デバイスの変更について説明するための図である。図24に示すように、ここでは、ユーザがサービス提供を行うために動作しているデバイス100-1およびデバイス100-2から離れ、サービス提供のための動作を行っていないデバイス100-5、100-6に近づいた場合を想定する。

[0208] この場合、デバイス100-5、100-6は、ユーザが所持するタグ6

0からの電波強度や、デバイス100-5、100-6が有するセンサからの入力信号等に基づいて、ユーザが移動してきたことを検知する。そして、デバイス100-5、100-6は、隣接するデバイス100-2等から、当該ユーザのIDや提供してもよいサービス等の情報を受信する。

[0209] そして、デバイス100-5、100-6は、受信した情報に基づいて、当該ユーザへのサービス提供を開始する。一方、サービス提供を行っていたデバイス100-1、100-2は、自デバイスのセンサおよびアクチュエータによりサービス提供可能な範囲から当該ユーザが外れたと判断すると、サービス提供を終了し、デバイスの動作や通信路をダウンさせる。

[0210] このように、ユーザがサービス提供を受けている際に移動しても、移動先周辺に配されるデバイス100が、ユーザIDやサービスの内容を引き継いで継続的に提供することができる。

[0211] (ネットワーク5へのアクセス経路)

図21を参照して説明した本実施形態による自律型音響システムでは、デバイス100-1~100-4のうち、少なくとも1つのデバイス100(ここでは、デバイス100-1)がネットワーク5へのアクセス経路を有していた。しかしながら、本開示による自律型音響システムの構成は図21に示す例に限定されず、デバイス100-1~100-4が閉じたネットワークであって、外部(ネットワーク5)へのアクセス経路を保有しない場合も考えられる。

[0212] このような場合、デバイス100-1~100-4は、例えばユーザが保持タグ60を、外部へのアクセスゲートウェイとして利用してもよい。すなわち、デバイス100-1~100-4は、互いに特定情報の報知をし合う状態において、タグ60が登場した時点で、タグ60を介してネットワーク5上のユーザIDDB6に問い合わせ、ユーザの認証を行う。

[0213] (複数ユーザへのサービスの提供)

次に、デバイス100-1~100-4が配されている空間に、複数のユーザが登場した場合のサービス提供について図25を参照して説明する。

- [0214] 図25は、本実施形態による自律型音響システムにおいて、複数ユーザにサービス提供を行う場合について説明するための図である。図25に示すように、デバイス100-1~100-4が配されている空間に、複数のユーザが登場し、各々がサービス要求を行った場合、各デバイス100-1~100-4は、複数のサービス提供を実施する。
- [0215] この場合、各ユーザに対するデバイス100-1~100-4の動作は、図21~図24を参照して説明した通りであるが、ユーザ1が所持するタグ60にとっては、ユーザ2が所持するタグ65は、周辺に配されたデバイスの一つとしてみなされる。また、ユーザ2が所持するタグ65にとっても、ユーザ1が所持するタグ60は、周辺に配されたデバイスの一つとしてみなされる。
- [0216] よって、デバイス100-1~100-4が、タグ60またはタグ65とも特性情報の報知を行い、信頼できるデバイスであるか否かを確認することで、タグ60またはタグ65の特性をサービス提供に利用してもよい。
- [0217] 例えば、デバイス100-1~100-4が閉じたネットワークである場合、デバイス100-1~100-4は、タグ65との特性情報の報知により、タグ65が外部のネットワーク5へのアクセス経路を保持していることを把握する。そして、デバイス100-1~100-4は、タグ60を所持するユーザ1へのサービス提供において、ユーザ2が所持するタグ65を1つのデバイスとして利用し、外部のネットワーク5に接続することができる。
- [0218] このように、ユーザ1の周辺に配されたデバイス100-1~100-4に限定されず、付近に居るユーザ2が所持するタグ65が、タグ60を所持するユーザ1に対して外部ネットワークへのアクセスを提供するといったケースが想定できる。
- [0219] なお、このような場合、タグ65が提供したサービス内容はサービスログDB8に書き込まれ、後にタグ65を所持するユーザ2に対して、ユーザ1へのサービス提供に寄与したことに基づくキックバックを行う際に利用され

てもよい。

[0220] <6. まとめ>

上述したように、本実施形態による音響システムでは、ユーザ周辺の空間を他の空間と相互連携させる際に、第3の空間への没入感を提供することが可能となる。具体的には、本実施形態による音響システムは、第1の所定対象（人物、場所、建物等）に対応する音声や画像を、ユーザの周囲に配された複数のスピーカーやディスプレイから再生することができる。また、この際、本実施形態による音響システムは、第2の所定対象（場所等）の空間を再現し、第2の所定対象への没入感や臨場感を提供することができる。このように、屋内や屋外の至る所に配されるマイクロフォン10、スピーカー20、イメージセンサ等を用いて、実質的にユーザの口、目、耳等の身体を広範囲に拡張させることが可能となり、新たなコミュニケーション方法を実現することができる。

[0221] さらに、本実施形態による音響システムでは、至る所にマイクロフォンやイメージセンサ等が配されているので、ユーザはスマートフォンや携帯電話端末を所有する必要がなく、声やジェスチャーで所定対象を指示し、所定対象周辺の空間と接続させることができる。

[0222] また、このような新たなコミュニケーション方法を実現する音響システムの構成は、複数のマイクおよび複数のスピーカー等を制御する信号処理装置により実現されてもよい。また、本実施形態による音響システムは、自律した各マイクおよび各スピーカー等のデバイスが隣接する他のデバイスと連携することで実現されてもよい。

[0223] 以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本技術はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0224] 例えば、信号処理装置 1 の構成は図 3 に示す構成に限定されず、例えば図 3 に示す認識部 17 および同定部 18 が、信号処理装置 1 ではなくネットワークを介して接続するサーバ側に設けられる構成であってもよい。この場合、信号処理装置 1 は、信号処理部 13 から出力されるオーディオ信号を通信 I/F 19 を介してサーバに送信する。また、サーバは、受信したオーディオ信号に基づいて、コマンド認識や、所定の対象（人物、場所、建物、番組、曲等）を同定する処理を行い、認識結果および同定された所定の対象に対応する接続先情報を信号処理装置 1 に送信する。

[0225] なお、本技術は以下のような構成も取ることができる。

(1)

特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第 1 の対象および第 2 の対象を認識する認識部と、

前記認識部により認識された前記第 1 および第 2 の対象を同定する同定部と、

前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、

前記特定ユーザの周辺に配される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第 1 および第 2 の対象の周辺のセンサから取得した各信号を処理する信号処理部と、

を備える、情報処理システム。

(2)

前記第 1 の対象は所定の人物、前記第 2 の対象は所定の場所であって、

前記信号処理部は、前記所定の人物周辺のセンサにより取得された信号、および前記所定の場所周辺のセンサにより取得された信号を処理する、前記

(1) に記載の情報処理システム。

(3)

前記第 1 の対象は所定の人物、前記第 2 の対象は所定の場所であって、

前記信号処理部は、前記所定の人物周辺のセンサによりリアルタイムで取得された信号、および前記所定の場所周辺のセンサで既に取得され蓄積された信号を処理する、前記（１）に記載の情報処理システム。

（４）

前記第１の対象の周辺のセンサおよび前記第２の対象の周辺のセンサは、それぞれ遠隔地に分散して配置される、前記（１）～（３）のいずれか１項に記載の情報処理システム。

（５）

前記特定ユーザの周辺に配される複数のセンサは、マイクロフォンであって、

前記認識部は、前記マイクロフォンにより検知されたオーディオ信号に基づいて、前記第１および第２の対象を認識する、前記（１）～（４）のいずれか１項に記載の情報処理システム。

（６）

前記特定ユーザの周辺に配される複数のセンサは、イメージセンサであって、

前記認識部は、前記イメージセンサにより取得された撮像画像に基づいて、前記第１および第２の対象を認識する、前記（１）～（４）のいずれか１項に記載の情報処理システム。

（７）

前記第１の対象の周辺のセンサおよび前記第２の対象の周辺のセンサは、互いに異なる種類のセンサである、前記（１）～（６）のいずれか１項に記載の情報処理システム。

（８）

前記信号処理部は、前記第１の対象の周辺のセンサにより取得された信号を、前記第２の対象に対応するパラメータの特性に基づいて加工し、前記第２の対象の周辺のセンサにより取得された信号に加算する処理を行う、前記（１）～（７）のいずれか１項に記載の情報処理装置。

(9)

前記信号処理部は、前記特定ユーザの感覚器官付近に定位するよう、前記第1および第2の対象の周辺のセンサから取得した各信号を処理する、前記(1)～(8)のいずれか1項に記載の情報処理装置。

(10)

前記第1および第2の対象の周辺の各センサは、マイクロフォンであって、

前記特定ユーザの周辺に配される複数のアクチュエータは、複数のスピーカーであって、

前記信号処理部は、前記複数のスピーカーから出力された際に前記特定ユーザの位置付近に音場を形成するよう、前記複数のスピーカーの各位置および推定された前記ユーザの位置に基づいて、前記第1および第2の対象の周辺の前記マイクロフォンにより收音された各オーディオ信号を処理する、前記(1)～(9)のいずれか1項に記載の情報処理システム。

(11)

前記推定部は、前記特定ユーザの位置を継続的に推定し、

前記信号処理部は、前記特定ユーザの位置の変化に応じて、前記特定ユーザの位置付近に音場を形成するよう前記各オーディオ信号を処理する、前記(10)に記載の情報処理システム。

(12)

特定ユーザの周辺のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、

前記認識部により認識された前記第1および第2の対象を同定する同定部と、

前記同定部により同定された前記第1および第2の対象の周辺に配される複数のセンサから取得された信号に基づき、前記特定ユーザの周辺のアクチュエータから出力する信号を生成する信号処理部と、
を備える、情報処理システム。

(13)

前記第1の対象は所定の人物、前記第2の対象は所定の場所であって、
前記信号処理部は、前記所定の人物周辺に配された複数のセンサにより取得された信号、および前記所定の場所周辺に配された複数のセンサより取得された信号を処理する、前記(12)に記載の情報処理システム。

(14)

前記第1の対象は所定の人物、前記第2の対象は所定の場所であって、
前記信号処理部は、前記所定の人物周辺に配された複数のセンサによりリアルタイムで取得された信号、および前記所定の場所周辺のセンサで既に取得され蓄積された信号を処理する、前記(12)に記載の情報処理システム。

(15)

コンピュータを、
特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、
前記認識部により認識された前記第1および第2の対象を同定する同定部と、
前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、
前記特定ユーザの周辺に配される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第1および第2の対象の周辺のセンサから取得した各信号を処理する信号処理部、
として機能させるための、プログラムが記憶された記憶媒体。

(16)

コンピュータを、
特定ユーザの周辺のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、

前記認識部により認識された前記第 1 および第 2 の対象を同定する同定部と、

前記同定部により同定された前記第 1 および第 2 の対象の周辺に配される複数のセンサから取得された信号に基づき、前記特定ユーザの周辺のアクチュエータから出力する信号を生成する信号処理部、
として機能させるための、プログラムが記憶された記憶媒体。

符号の説明

- [0226] 1、1 A、1 B 信号処理装置
3、3' 管理サーバ
5 ネットワーク
6 ユーザ I D D B
7 通信端末
8 サービスログ D B
9 ユーザ個人 D B
1 0、1 0 A、1 0 B、1 0 C マイクロフォン (マイク)
1 1 アンプ・A D C (アナログデジタルコンバータ) 部
1 3、2 0 0 信号処理部
1 5 マイク位置情報 D B (データベース)
1 6 ユーザ位置推定部
1 7 認識部
1 8 同定部
1 9 通信 I / F (インターフェース)
2 0、2 0 A、2 0 B、2 0 C スピーカー
2 3 D A C (デジタルアナログコンバータ) ・アンプ部
2 5 ユーザ認証部
3 2 管理部
3 3、3 4 検索部
3 5 ユーザ位置情報 D B

- 36 音響情報パラメータDB
- 37 音響コンテンツDB
- 40、40-1、40-2、40-3 音響閉曲面
- 42 音場
- 43、43-1、43-2、43-3 閉曲面
- 60、65 タグ
- 100、100-1~100-4 デバイス
- 131 マイクアレイ処理部
- 133、210 高S/N化処理部
- 135、220 音場再生信号処理部
- 136 Convolution部
- 137、139 ハウリング抑制部
- 138 Matrix Convolution部

請求の範囲

- [請求項1] 特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、
前記認識部により認識された前記第1および第2の対象を同定する同定部と、
前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、
前記特定ユーザの周辺に配される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第1および第2の対象の周辺のセンサから取得した各信号を処理する信号処理部と、
を備える、情報処理システム。
- [請求項2] 前記第1の対象は所定の人物、前記第2の対象は所定の場所であって、
前記信号処理部は、前記所定の人物周辺のセンサにより取得された信号、および前記所定の場所周辺のセンサにより取得された信号を処理する、請求項1に記載の情報処理システム。
- [請求項3] 前記第1の対象は所定の人物、前記第2の対象は所定の場所であって、
前記信号処理部は、前記所定の人物周辺のセンサによりリアルタイムで取得された信号、および前記所定の場所周辺のセンサで既に取得され蓄積された信号を処理する、請求項1に記載の情報処理システム。
- [請求項4] 前記第1の対象の周辺のセンサおよび前記第2の対象の周辺のセンサは、それぞれ遠隔地に分散して配置される、請求項1に記載の情報処理システム。
- [請求項5] 前記特定ユーザの周辺に配される複数のセンサは、マイクロフォンであって、

前記認識部は、前記マイクロフォンにより検知されたオーディオ信号に基づいて、前記第1および第2の対象を認識する、請求項1に記載の情報処理システム。

[請求項6] 前記特定ユーザの周辺に配される複数のセンサは、イメージセンサであって、

前記認識部は、前記イメージセンサにより取得された撮像画像に基づいて、前記第1および第2の対象を認識する、請求項1に記載の情報処理システム。

[請求項7] 前記第1の対象の周辺のセンサおよび前記第2の対象の周辺のセンサは、互いに異なる種類のセンサである、請求項1に記載の情報処理システム。

[請求項8] 前記信号処理部は、前記第1の対象の周辺のセンサにより取得された信号を、前記第2の対象に対応するパラメータの特性に基づいて加工し、前記第2の対象の周辺のセンサにより取得された信号に加算する処理を行う、請求項1に記載の情報処理装置。

[請求項9] 前記信号処理部は、前記特定ユーザの感覚器官付近に定位するよう、前記第1および第2の対象の周辺のセンサから取得した各信号を処理する、請求項1に記載の情報処理装置。

[請求項10] 前記第1および第2の対象の周辺の各センサは、マイクロフォンであって、

前記特定ユーザの周辺に配される複数のアクチュエータは、複数のスピーカーであって、

前記信号処理部は、前記複数のスピーカーから出力された際に前記特定ユーザの位置付近に音場を形成するよう、前記複数のスピーカーの各位置および推定された前記ユーザの位置に基づいて、前記第1および第2の対象の周辺の前記マイクロフォンにより收音された各オーディオ信号を処理する、請求項1に記載の情報処理システム。

[請求項11] 前記推定部は、前記特定ユーザの位置を継続的に推定し、

前記信号処理部は、前記特定ユーザの位置の変化に応じて、前記特定ユーザの位置付近に音場を形成するよう前記各オーディオ信号を処理する、請求項10に記載の情報処理システム。

[請求項12] 特定ユーザの周辺のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、

前記認識部により認識された前記第1および第2の対象を同定する同定部と、

前記同定部により同定された前記第1および第2の対象の周辺に配される複数のセンサから取得された信号に基づき、前記特定ユーザの周辺のアクチュエータから出力する信号を生成する信号処理部と、を備える、情報処理システム。

[請求項13] 前記第1の対象は所定の人物、前記第2の対象は所定の場所であって、

前記信号処理部は、前記所定の人物周辺に配された複数のセンサにより取得された信号、および前記所定の場所周辺に配された複数のセンサより取得された信号を処理する、請求項12に記載の情報処理システム。

[請求項14] 前記第1の対象は所定の人物、前記第2の対象は所定の場所であって、

前記信号処理部は、前記所定の人物周辺に配された複数のセンサによりリアルタイムで取得された信号、および前記所定の場所周辺のセンサで既に取得され蓄積された信号を処理する、請求項12に記載の情報処理システム。

[請求項15] コンピュータを、

特定ユーザの周辺に配される複数のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、

前記認識部により認識された前記第1および第2の対象を同定する同定部と、

前記複数のセンサのいずれかにより検知された信号に応じて、前記特定ユーザの位置を推定する推定部と、

前記特定ユーザの周辺に配される複数のアクチュエータから出力される際に、前記推定部により推定された前記特定ユーザの位置付近に定位するよう、前記同定部により同定された前記第1および第2の対象の周辺のセンサから取得した各信号を処理する信号処理部、
として機能させるための、プログラムが記憶された記憶媒体。

[請求項16]

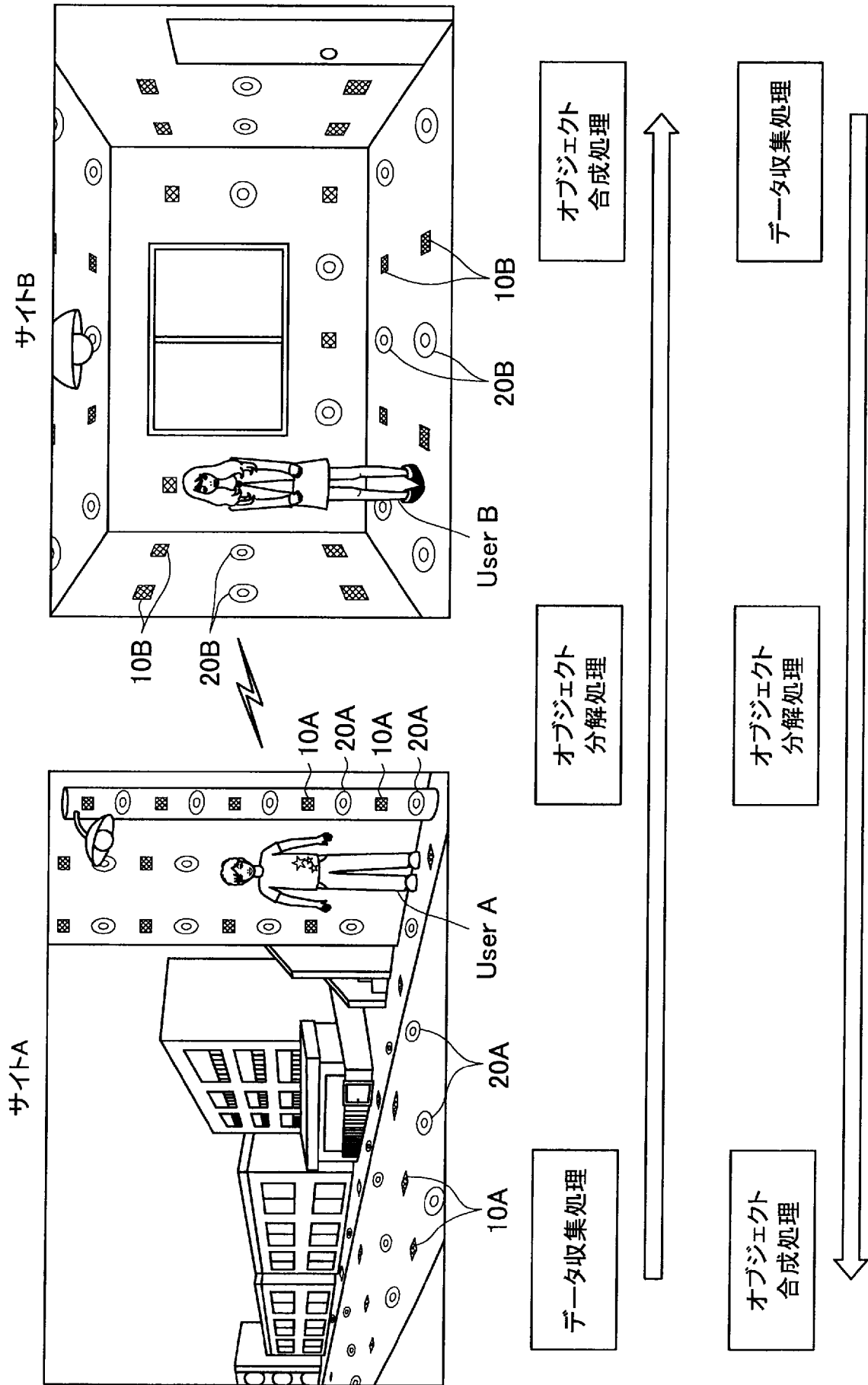
コンピュータを、

特定ユーザの周辺のセンサにより検知された信号に基づいて、第1の対象および第2の対象を認識する認識部と、

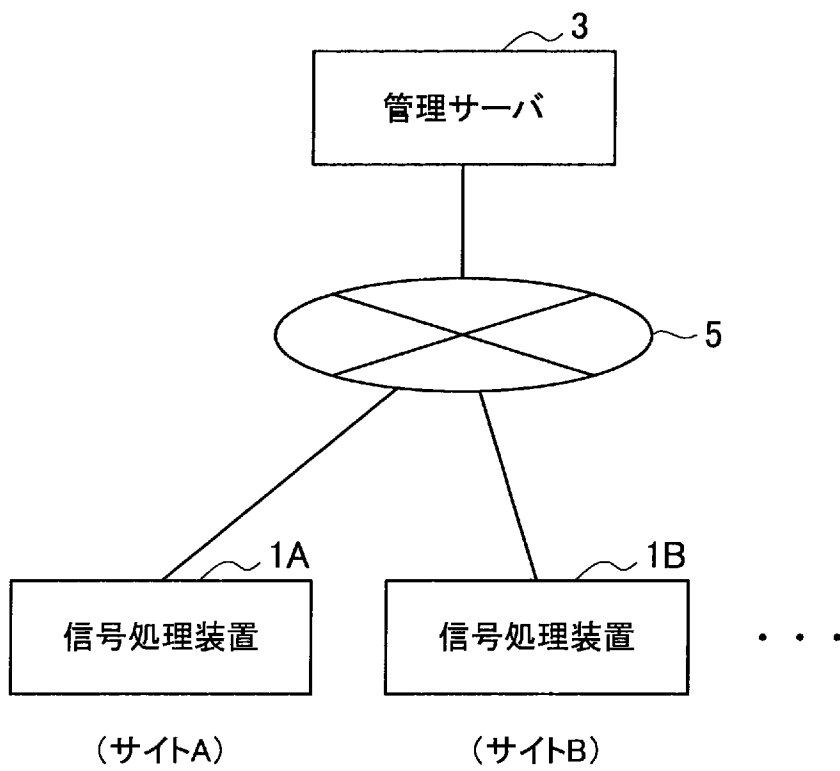
前記認識部により認識された前記第1および第2の対象を同定する同定部と、

前記同定部により同定された前記第1および第2の対象の周辺に配される複数のセンサから取得された信号に基づき、前記特定ユーザの周辺のアクチュエータから出力する信号を生成する信号処理部、
として機能させるための、プログラムが記憶された記憶媒体。

[図1]

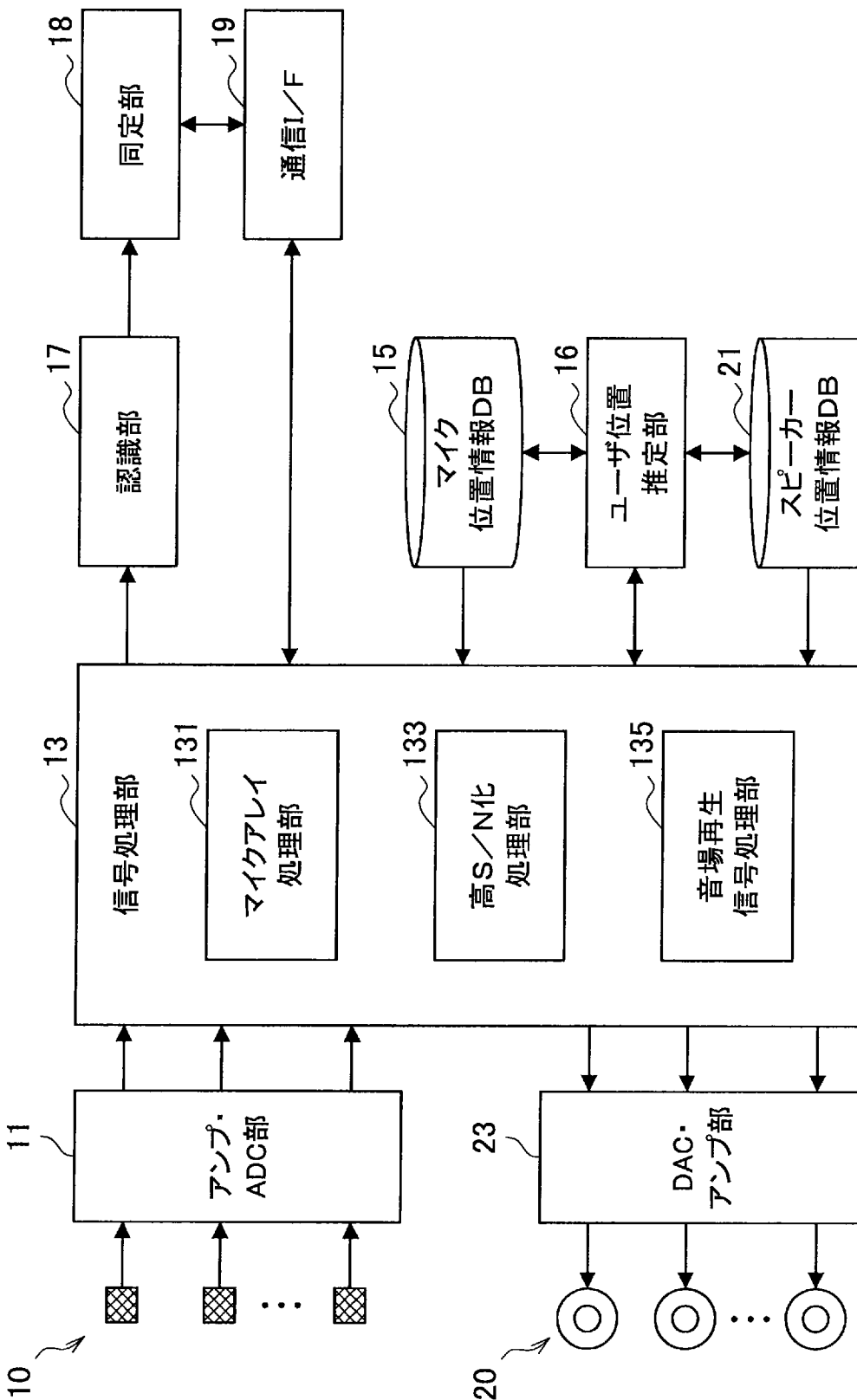


[図2]

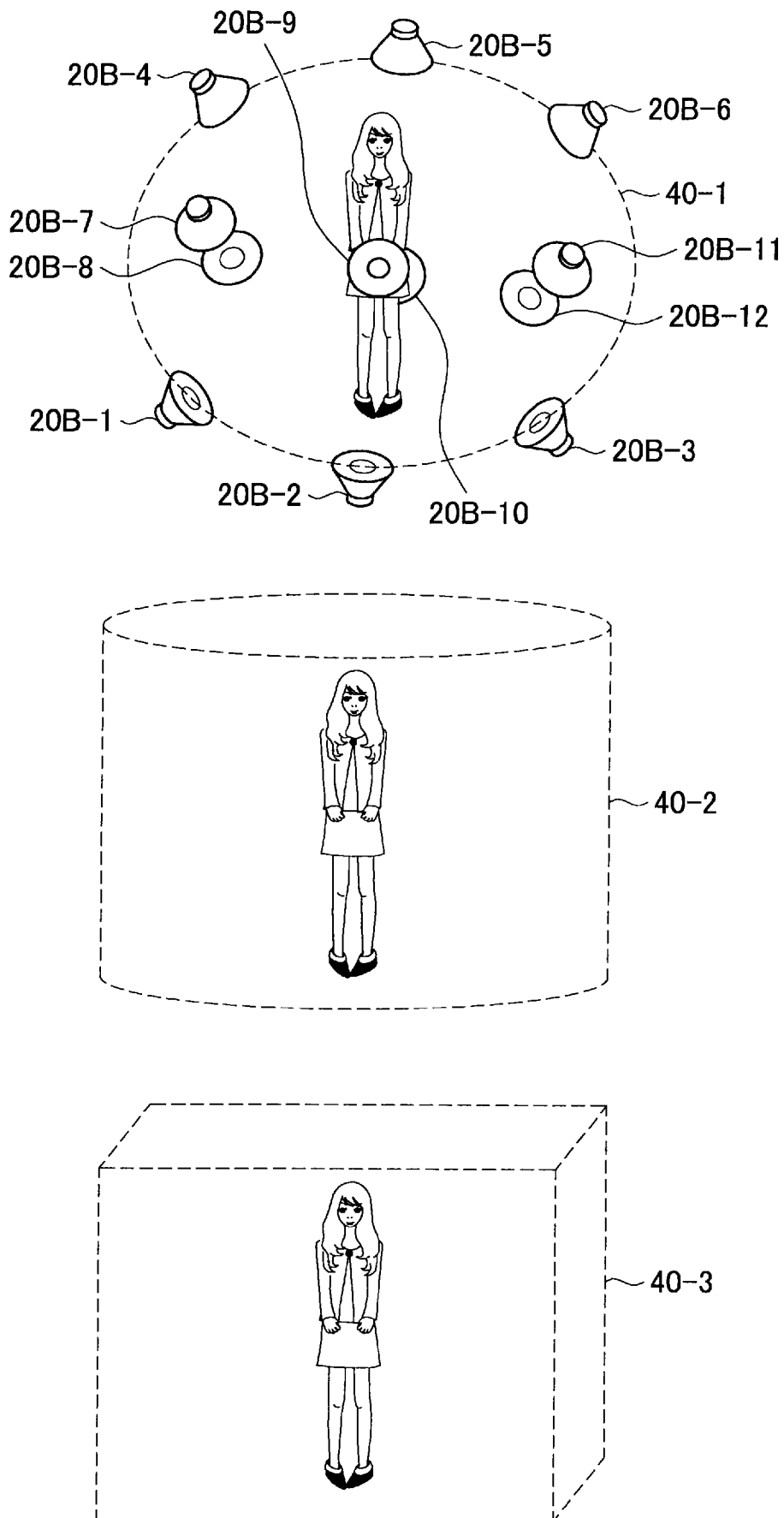


[図3]

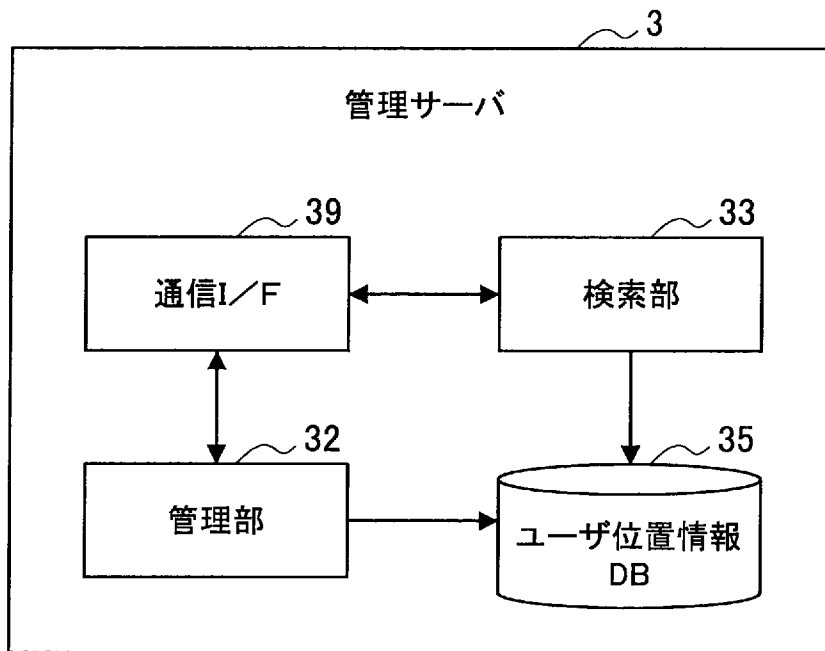
1:信号処理装置



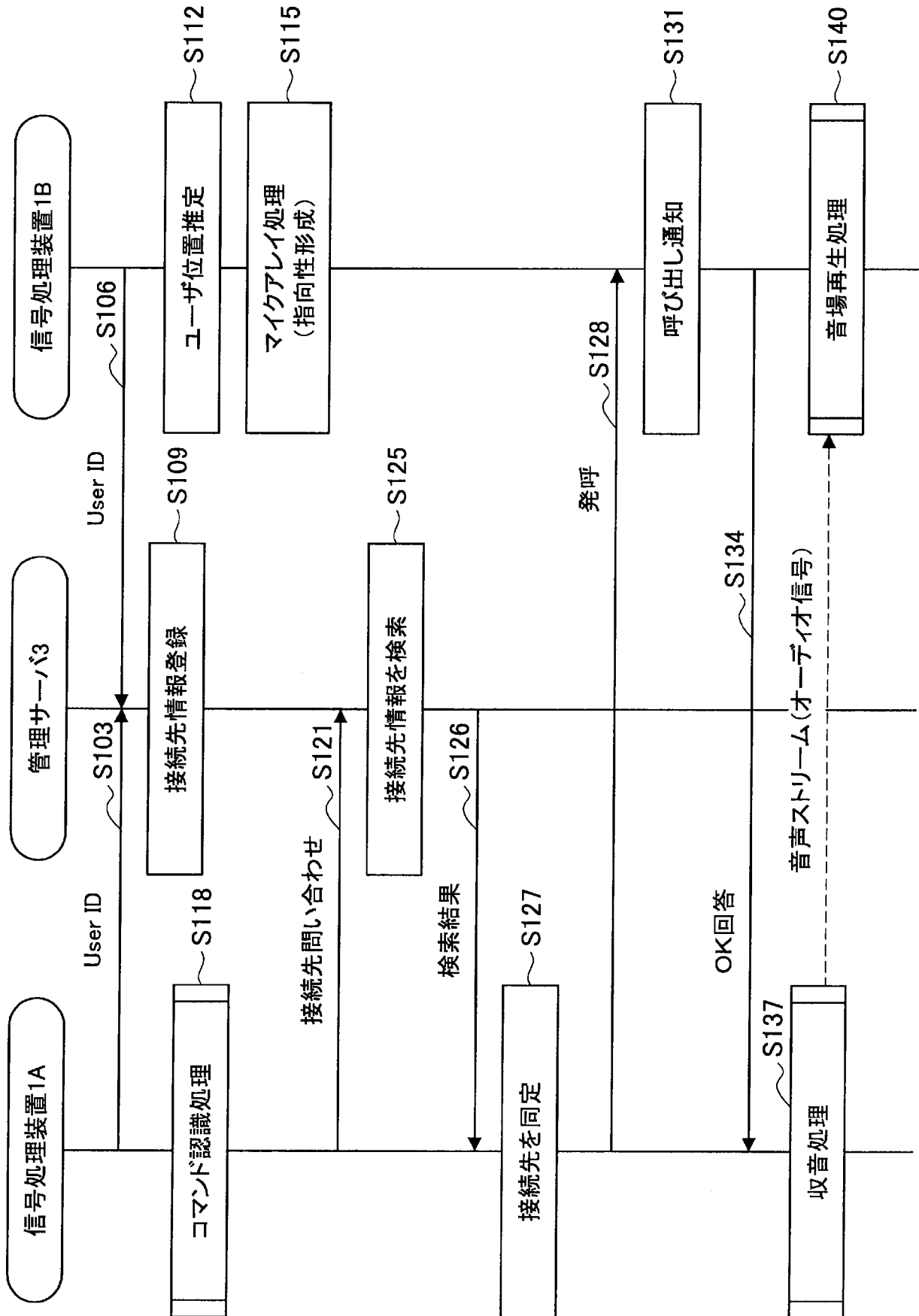
[図4]



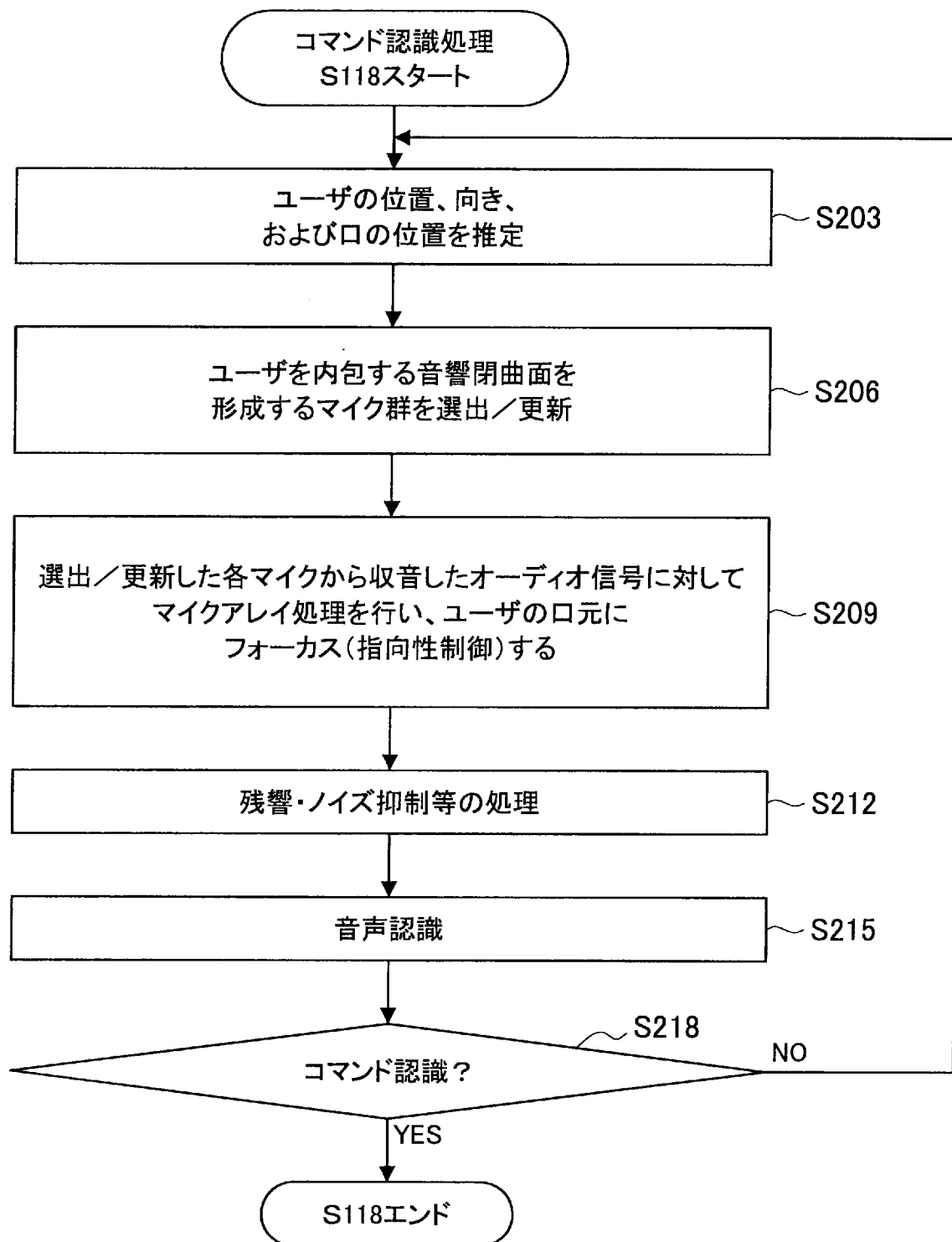
[図5]



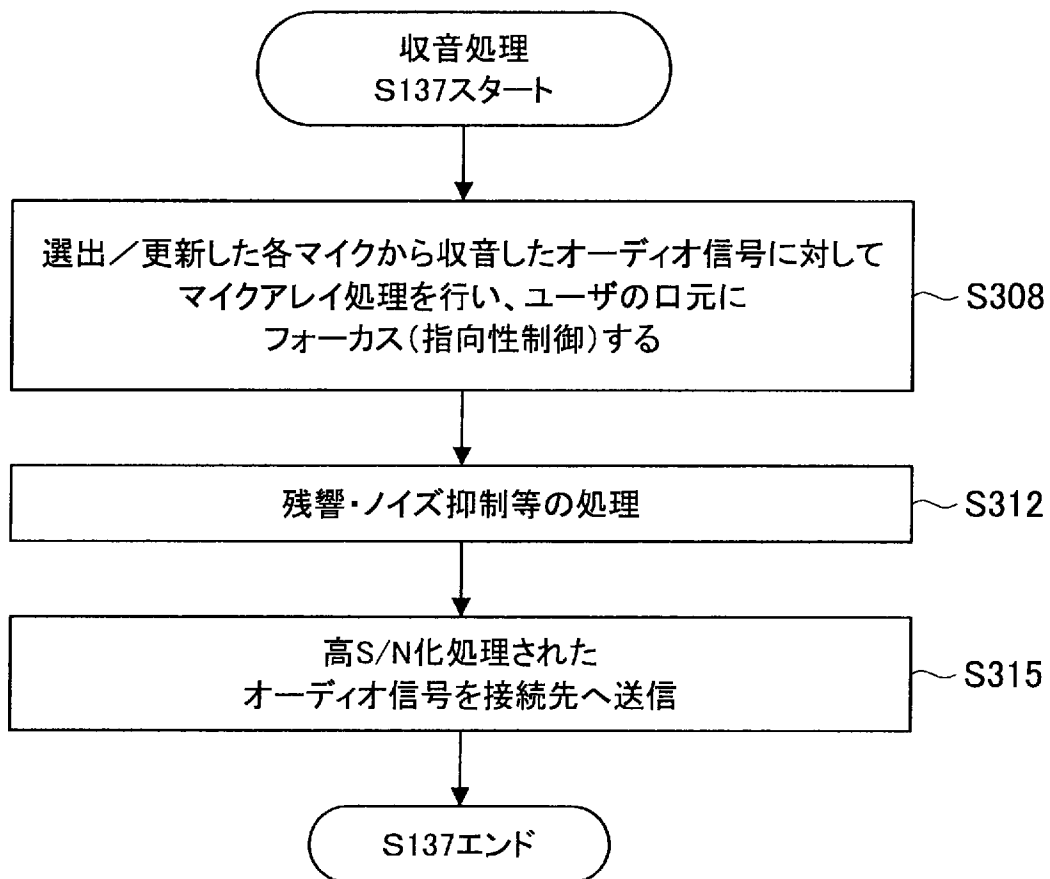
[図6]



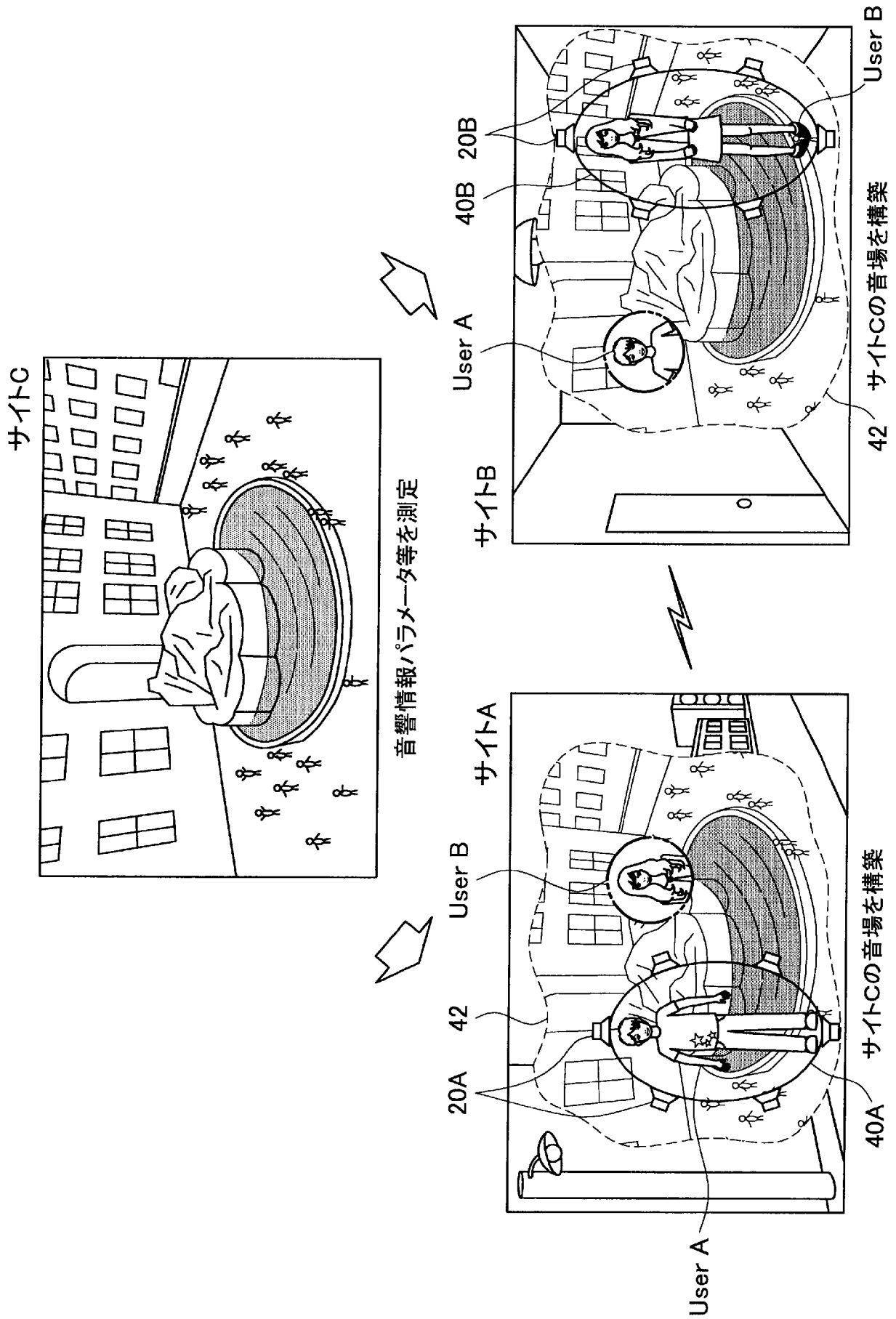
[図7]



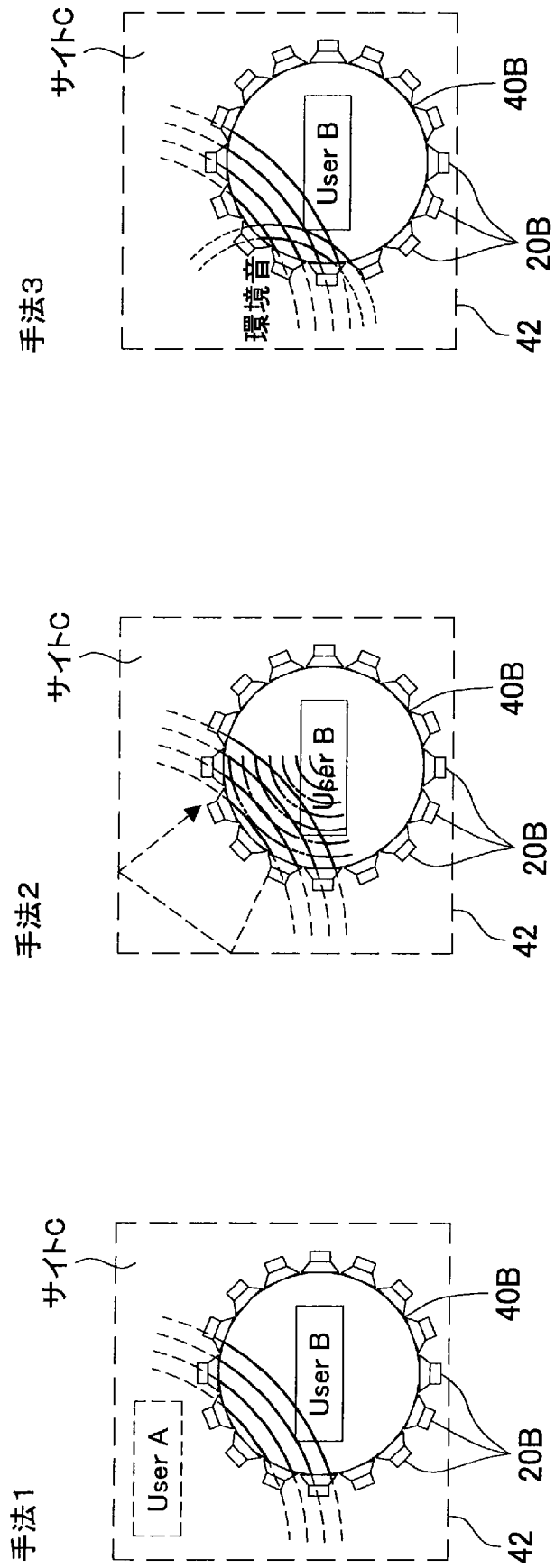
[図8]



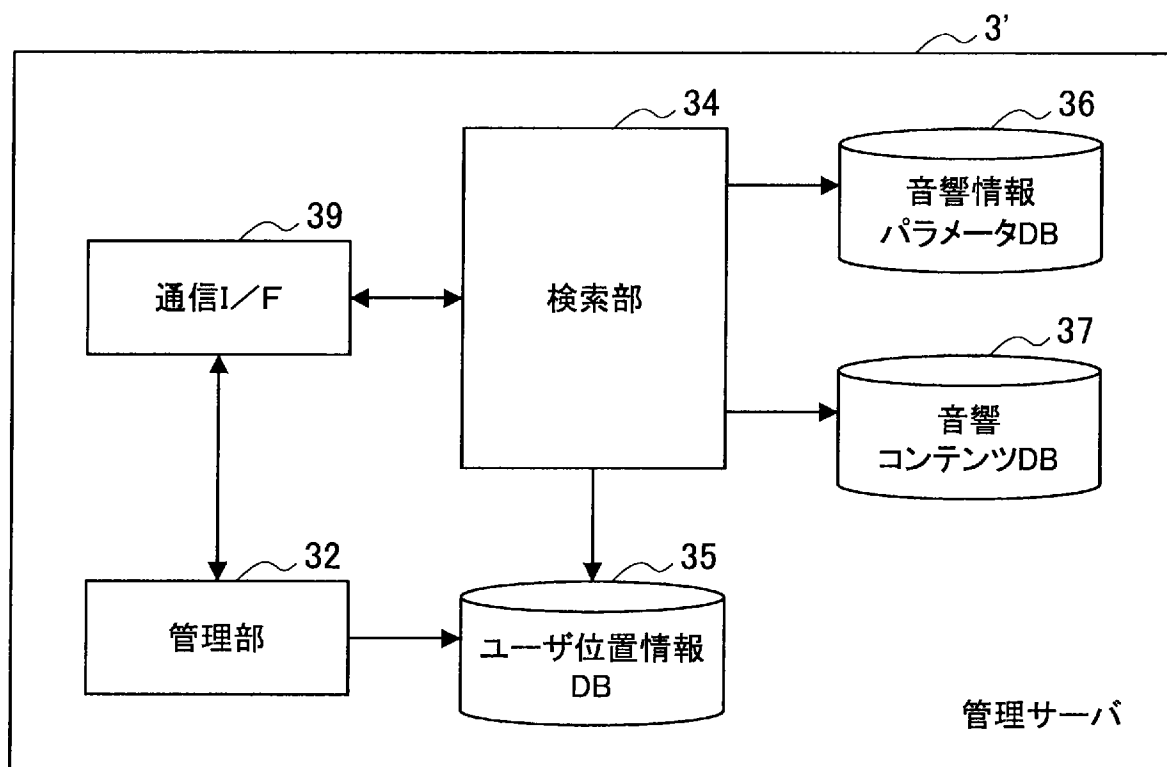
[図9]



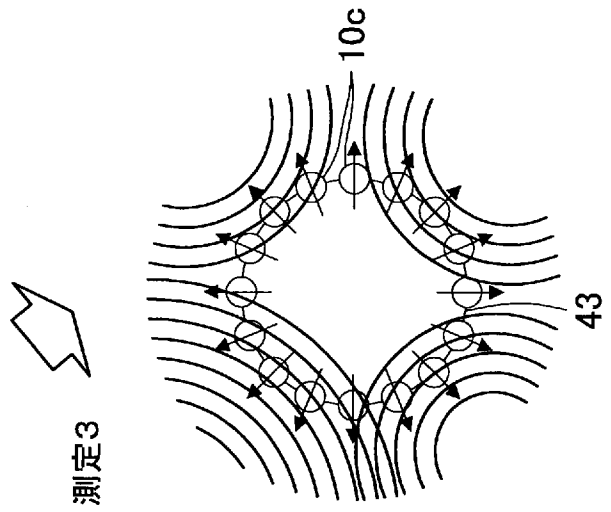
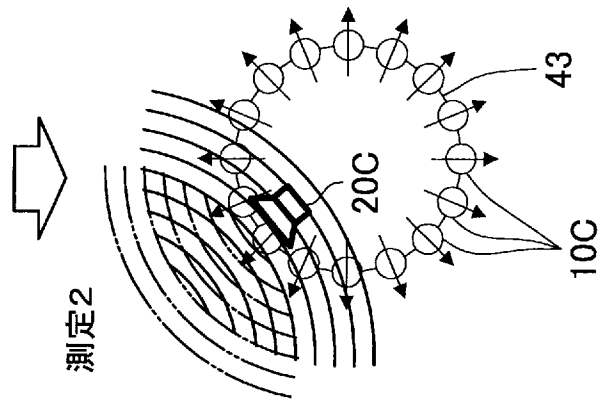
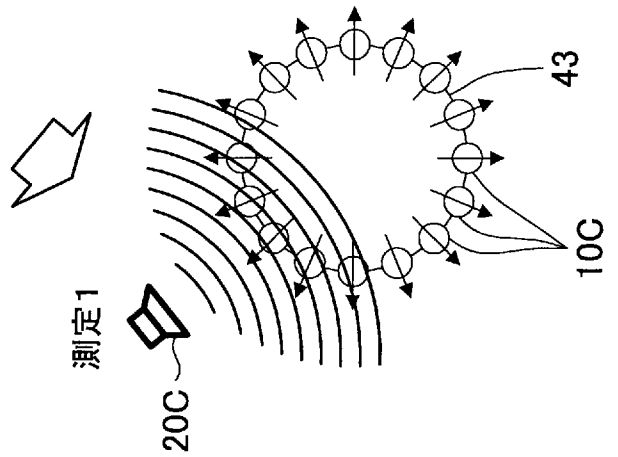
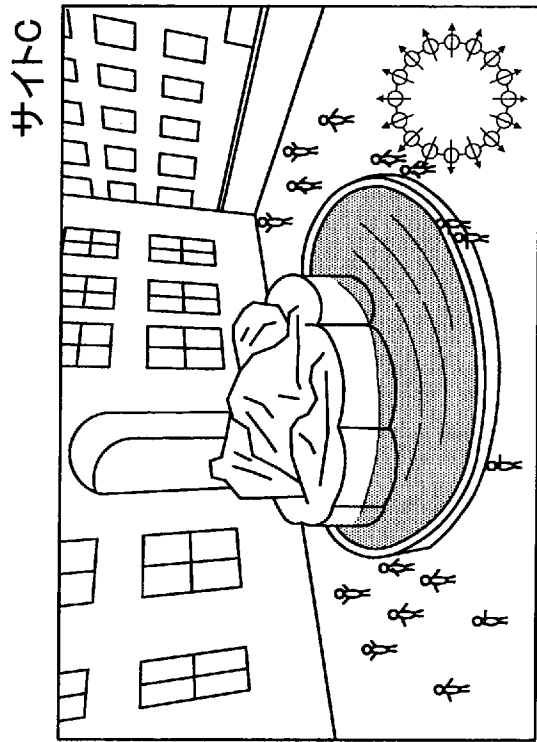
[図10]



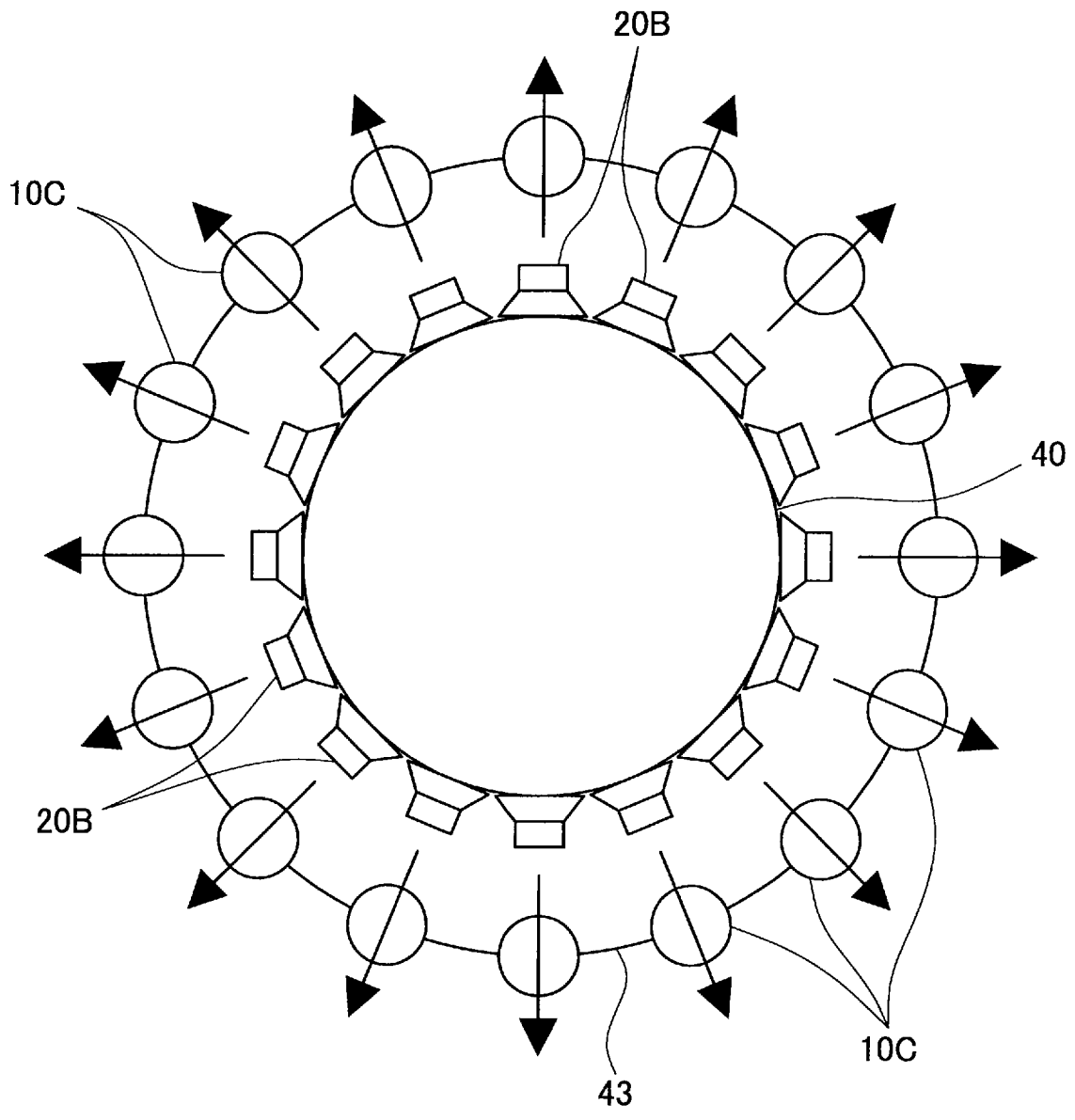
[図11]



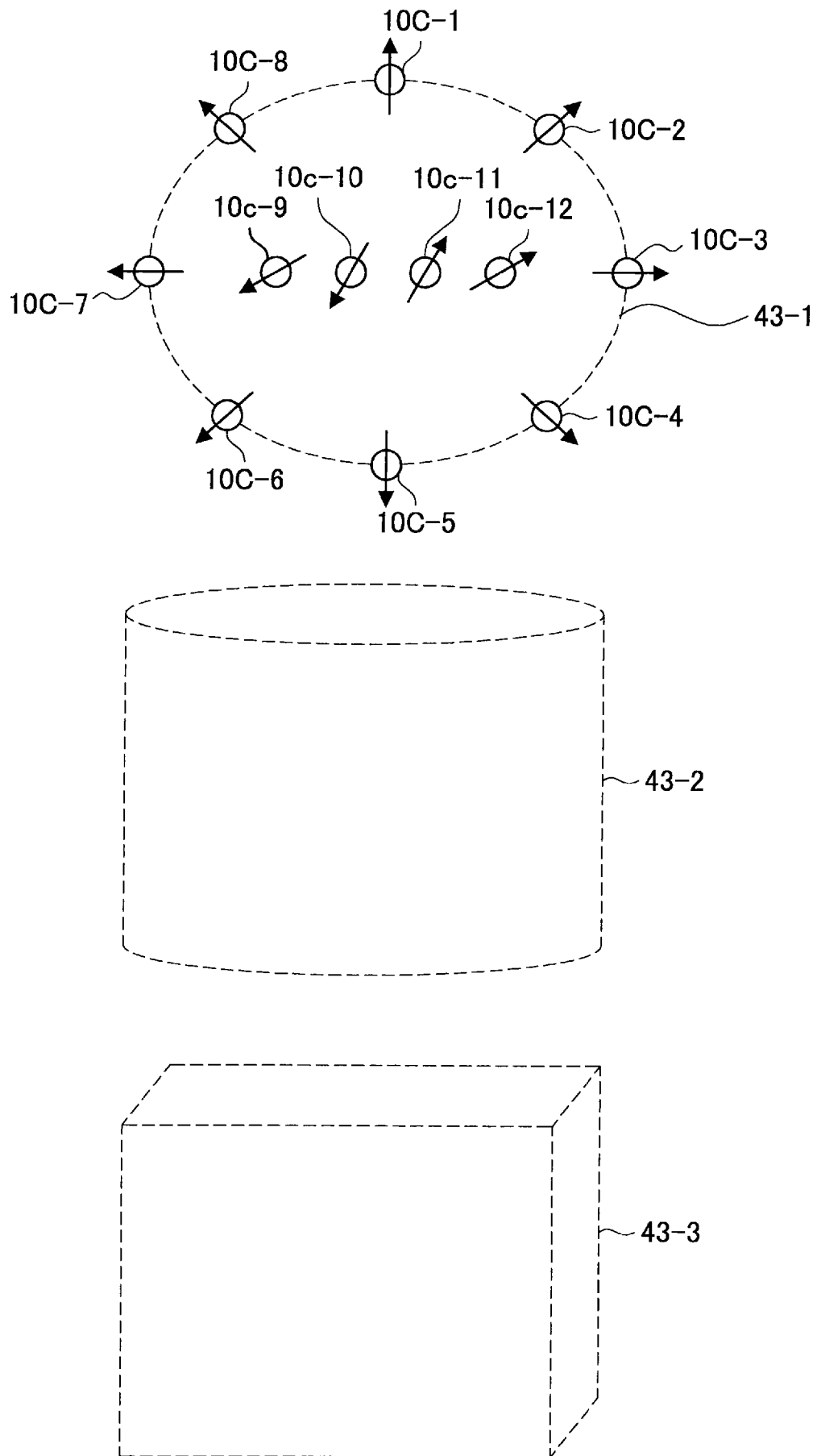
[図12]



[図13]

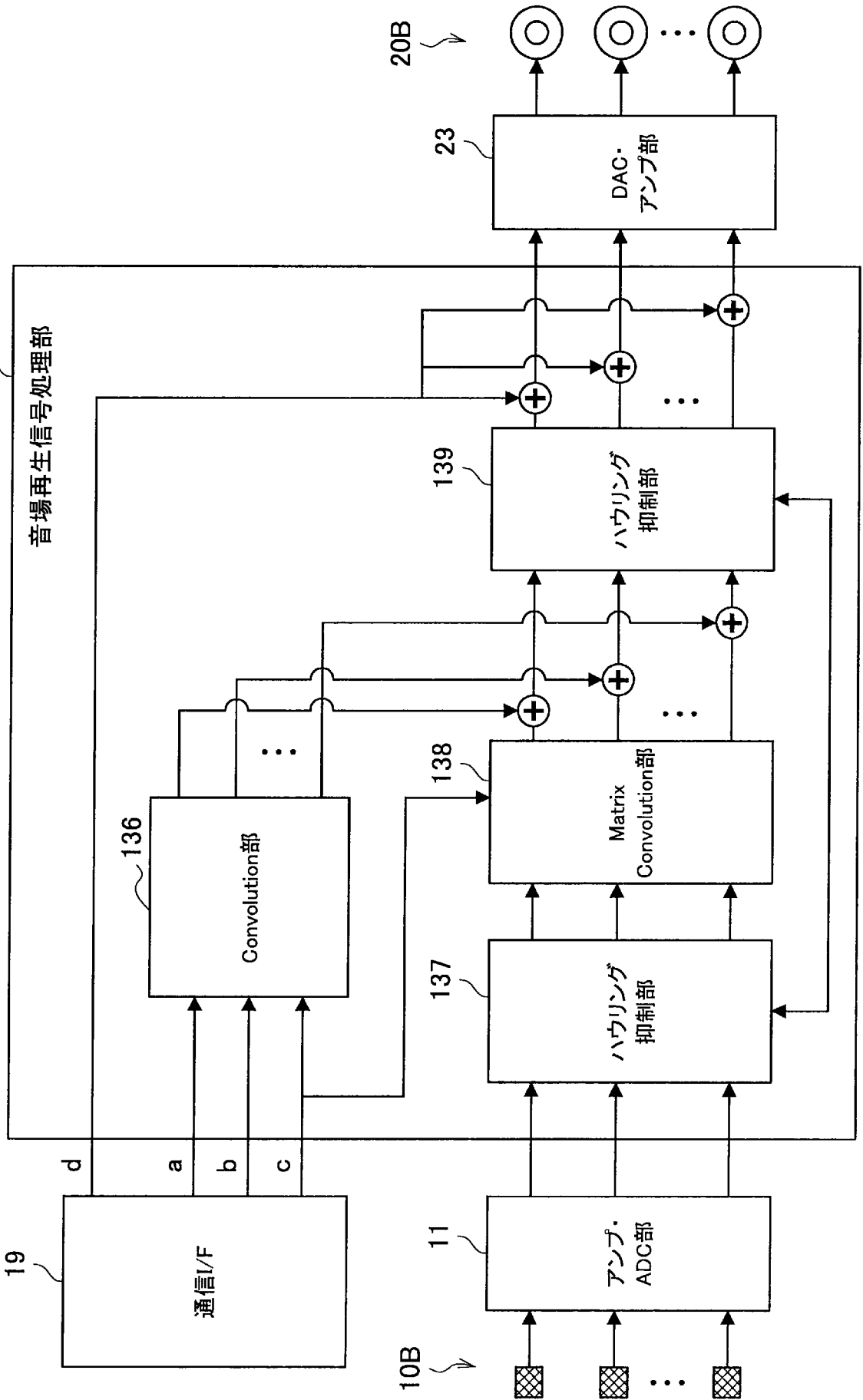


[図14]

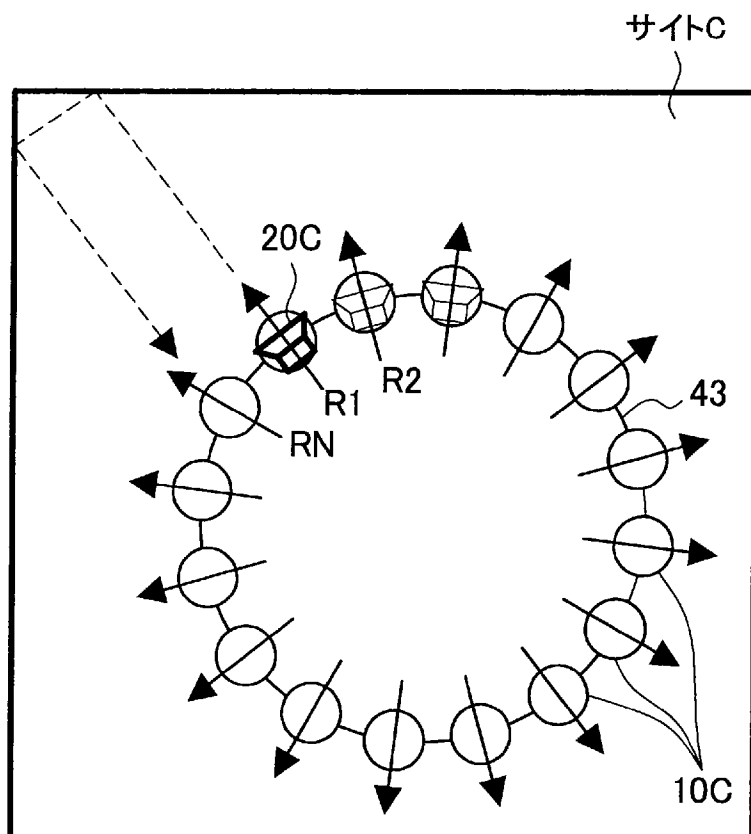


[図15]

1B:信号処理装置



[図16A]



インパルス応答(伝達関数)を測定

R1のSPから →(サイト環境での反射)→R1,R2...RNのマイクまで

R2のSPから →(サイト環境での反射)→R1,R2...RNのマイクまで

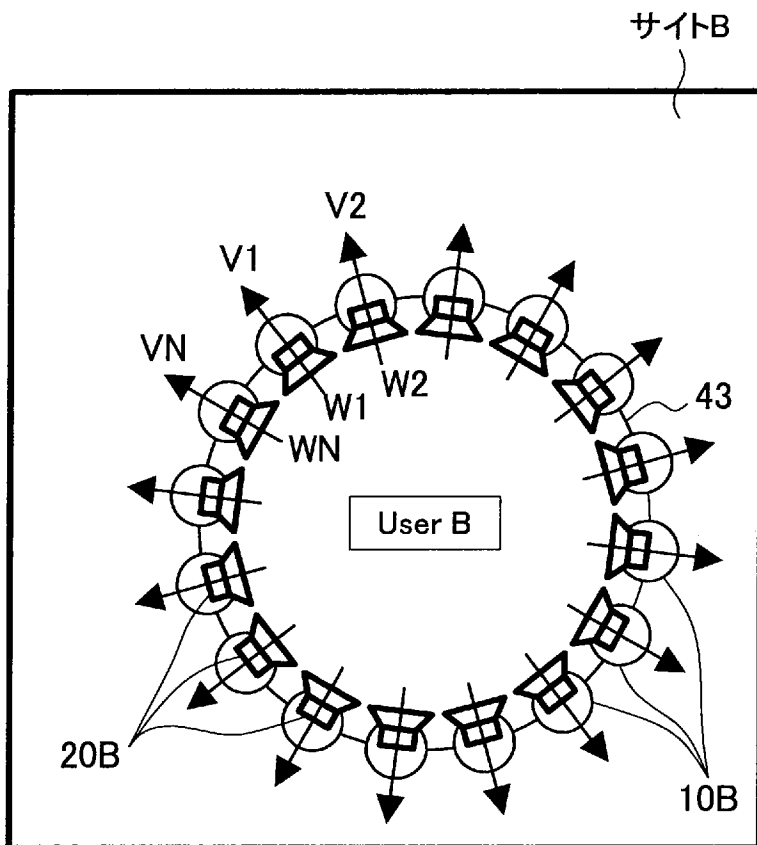
.....

RNのSPから →(サイト環境での反射)→R1,R2...RNのマイクまで

↓ 伝達関数Rを用いて行列化

$$\begin{pmatrix} R_{11} & R_{12} & \dots & R_{1N} \\ R_{21} & R_{22} & \dots & R_{2N} \\ \vdots & & \ddots & \vdots \\ R_{N1} & \dots & & R_{NN} \end{pmatrix} \quad \text{式(1)}$$

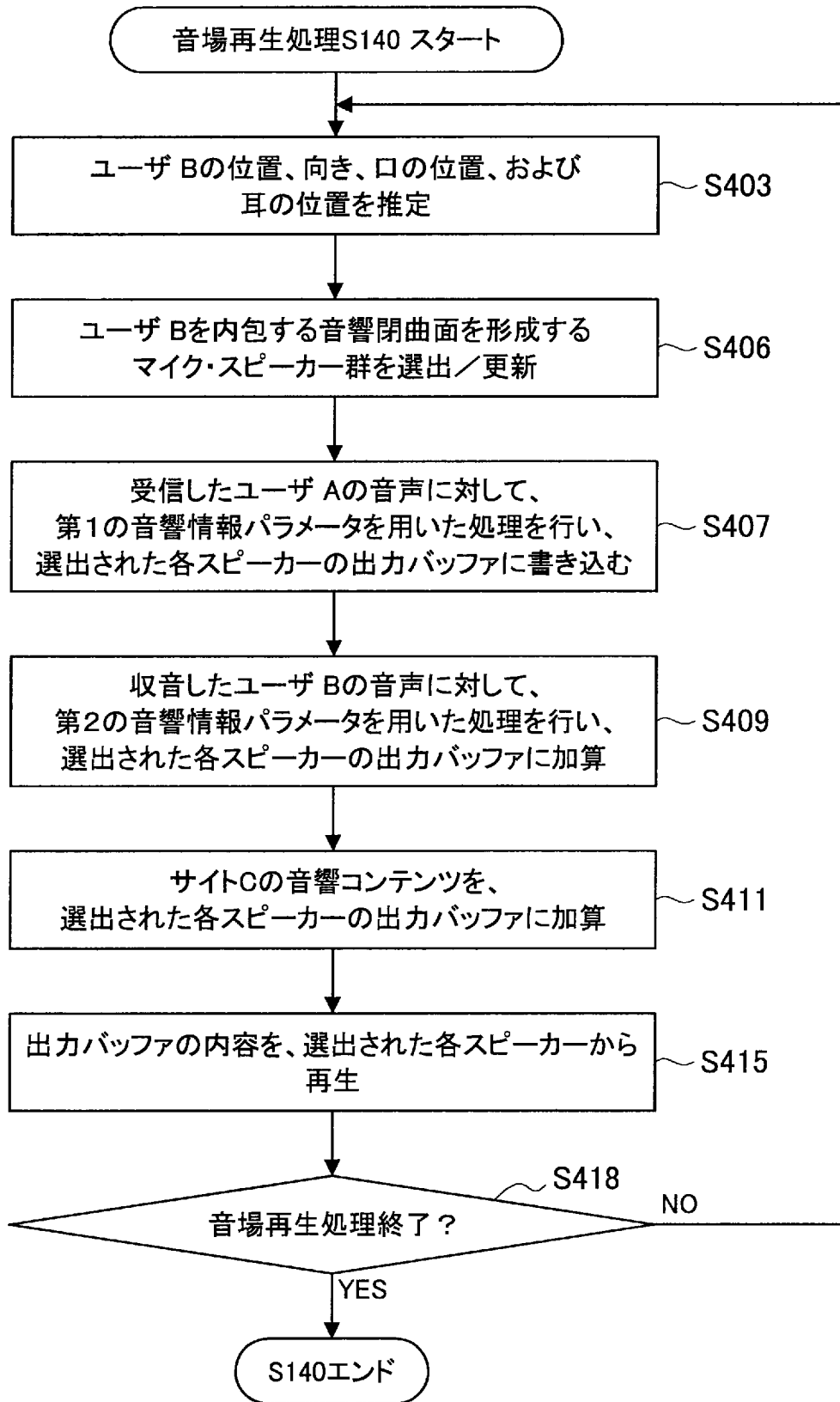
[図16B]



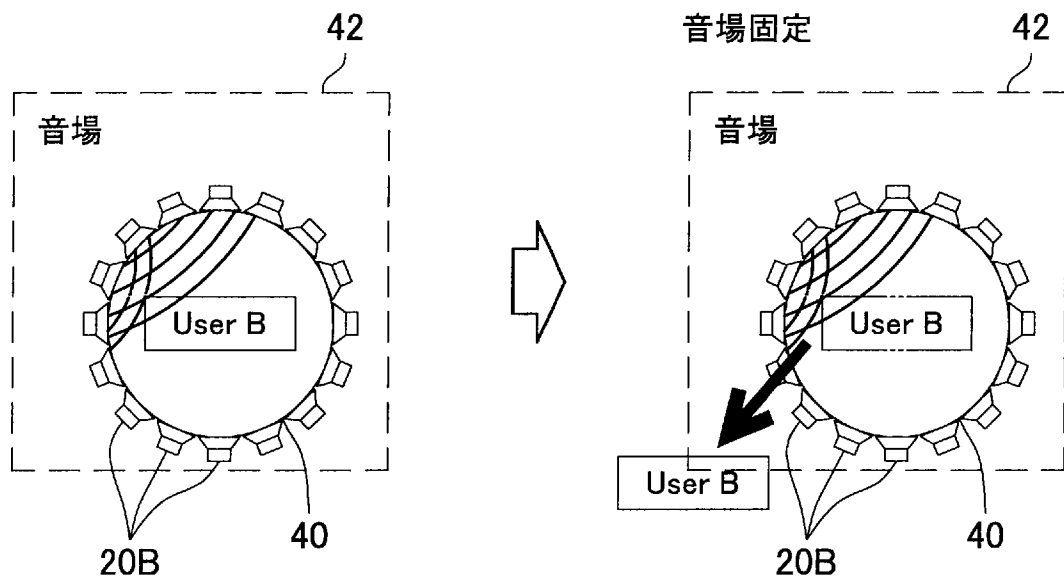
伝達関数群の行列(式(1))を用いて、
各スピーカ-の出力を求める

$$\begin{array}{c} \text{SP出力} \\ \left(\begin{array}{c} W_1 \\ W_2 \\ \vdots \\ W_N \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{cccc} R_{11} & R_{12} & \cdots & R_{1N} \\ R_{21} & R_{22} & \cdots & R_{2N} \\ \vdots & & \ddots & \vdots \\ R_{N1} & \cdots & & R_{NN} \end{array} \right) \end{array} \begin{array}{c} \text{Mic入力} \\ \left(\begin{array}{c} V_1 \\ V_2 \\ \vdots \\ V_N \end{array} \right) \end{array} \quad \text{式(2)}$$

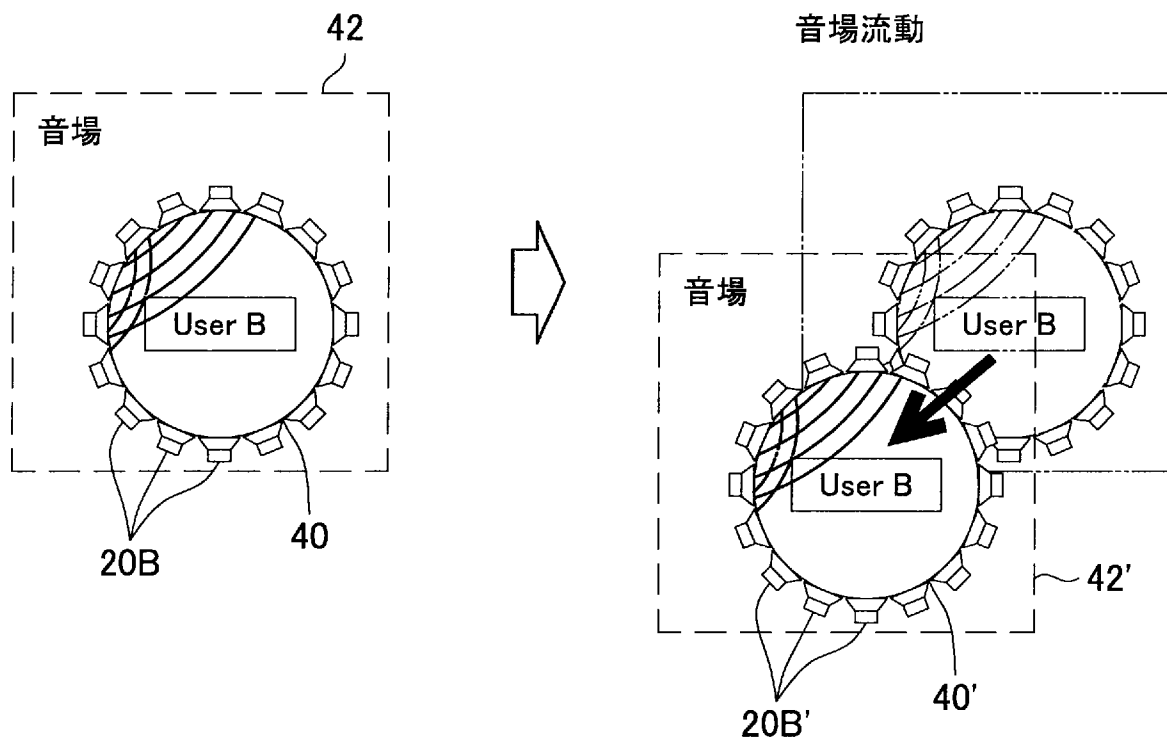
[図17]



[図18A]

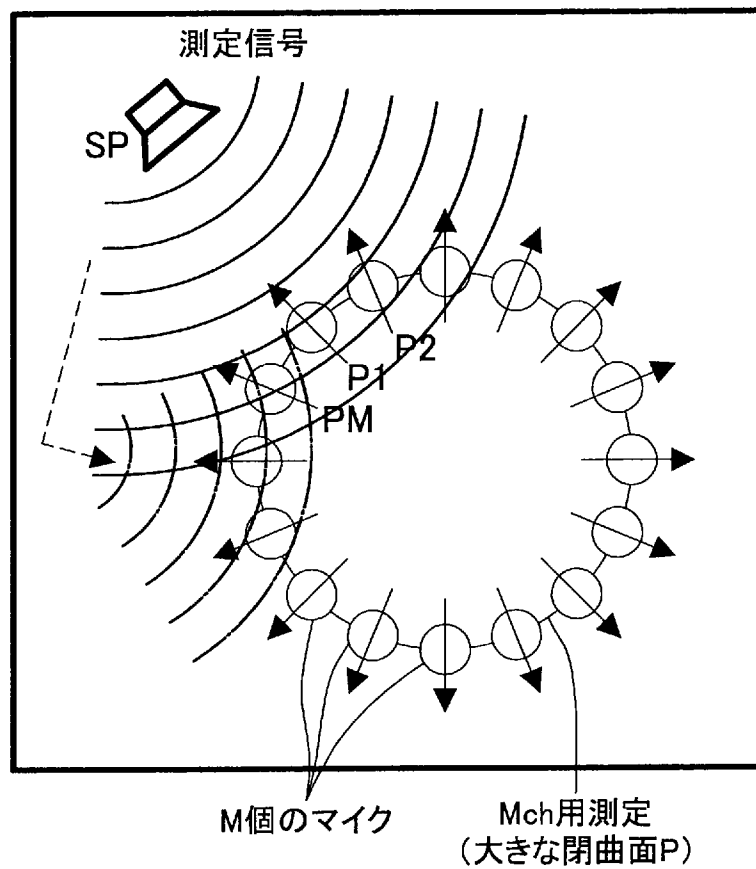


[図18B]



[図19A]

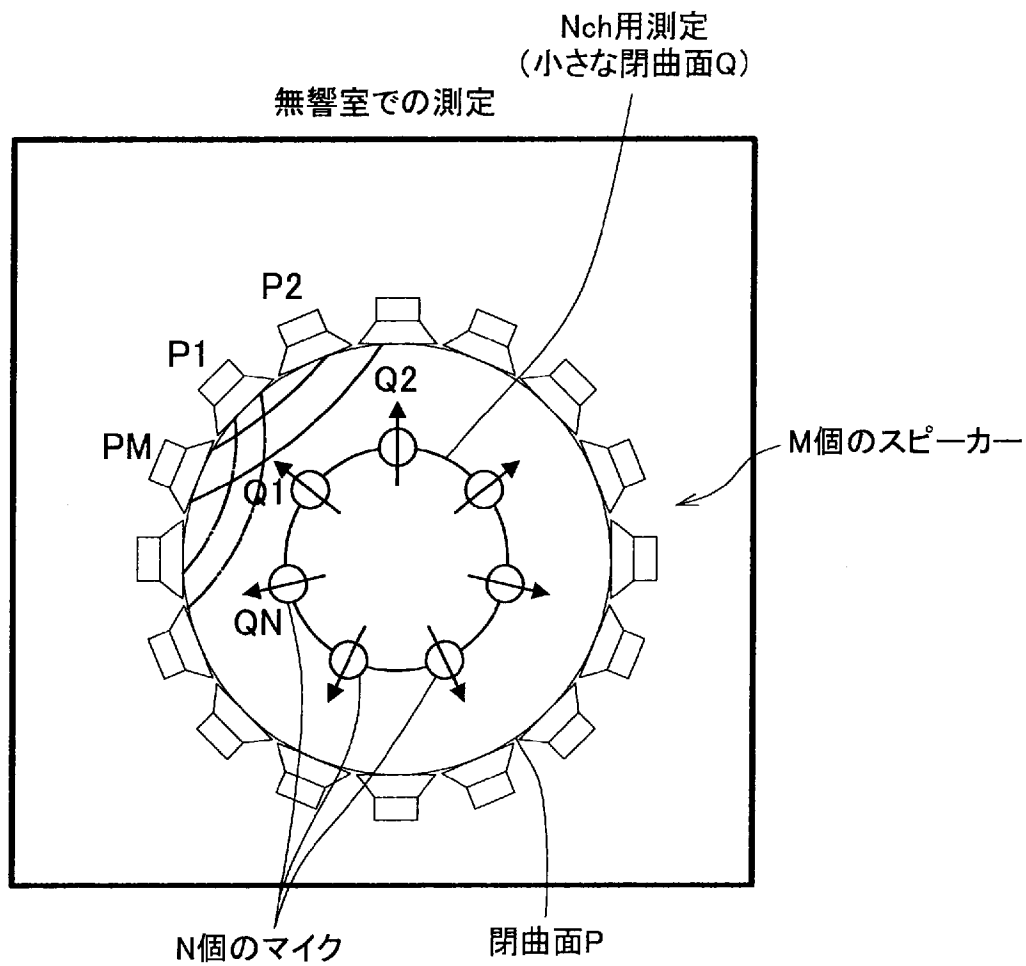
測定対象空間での測定



SPから→P1,P2...PMのマイクまでの
インパルス応答を測定

$$\begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_M \end{pmatrix} \text{式(3)}$$

[図19B]



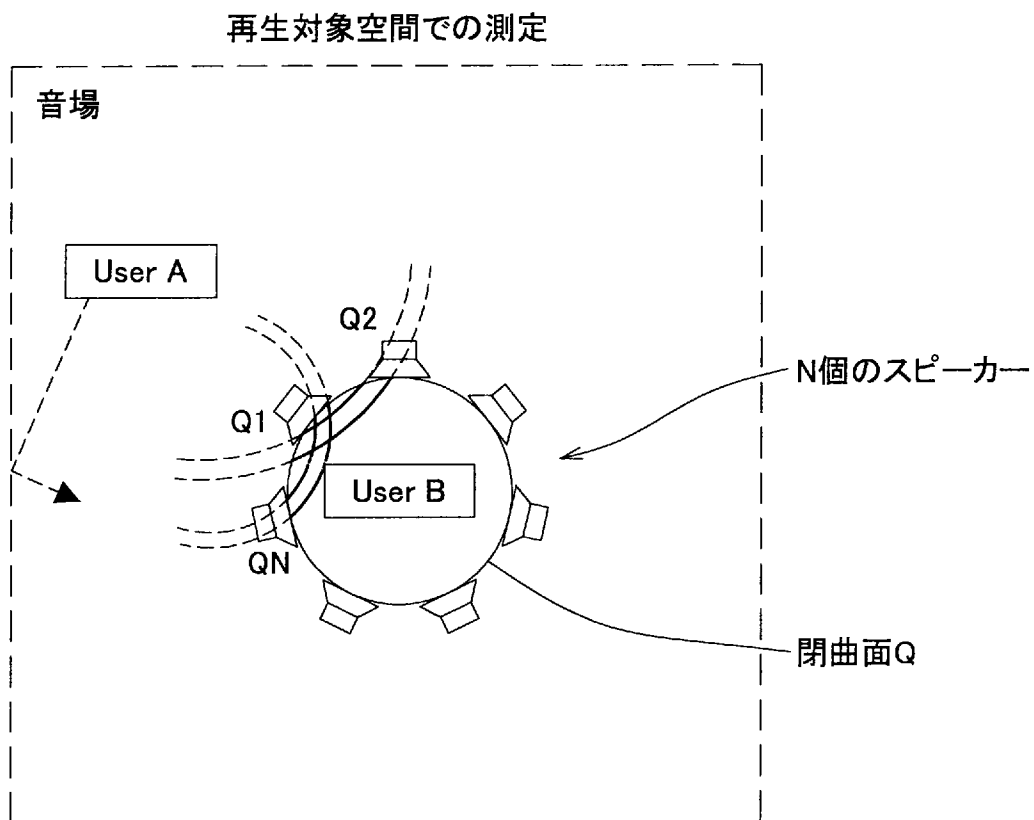
P1のスピーカーから→Q1,Q2...QNのマイクまで
 P2のスピーカーから→Q1,Q2...QNのマイクまで

 PMのスピーカーから→Q1,Q2...QNのマイクまで
 のインパルス応答を測定

↓ 行列化

$$\begin{pmatrix} PQ_{11} & PQ_{12} & \dots & PQ_{1N} \\ PQ_{21} & PQ_{22} & \dots & PQ_{2N} \\ \vdots & & \ddots & \vdots \\ PQ_{M1} & \dots & & PQ_{MN} \end{pmatrix} \quad \text{式(4)}$$

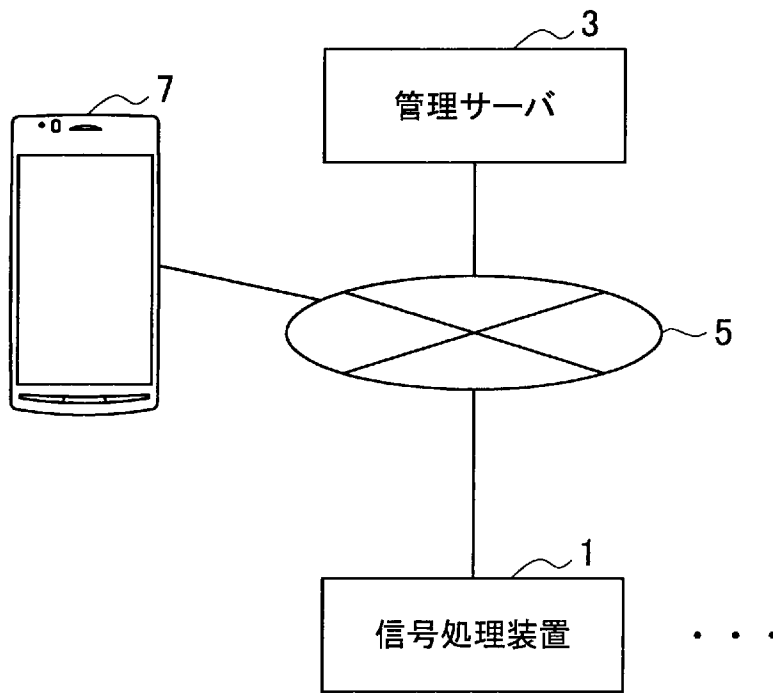
[図19C]



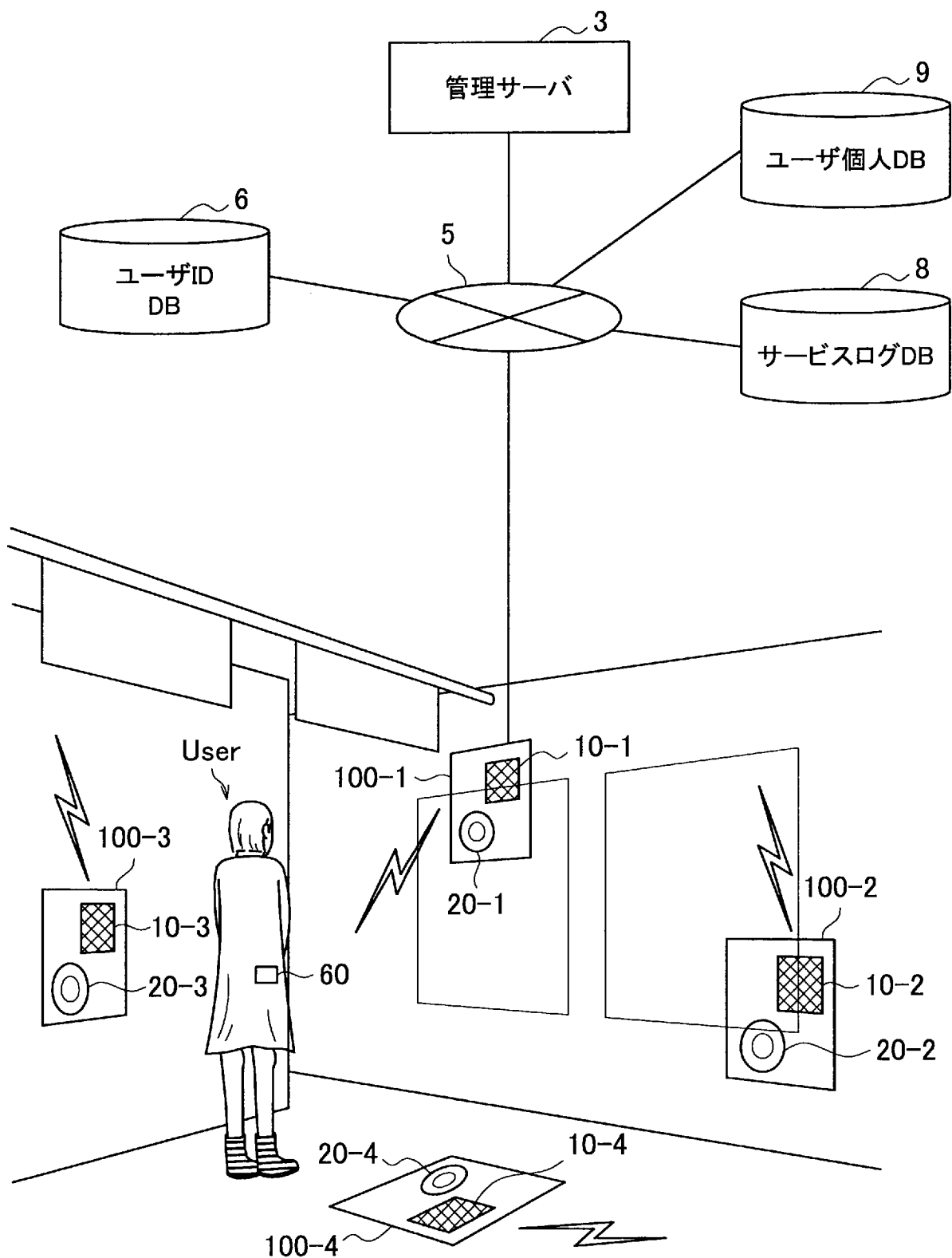
SP出力

$$\begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_N \end{pmatrix} = \begin{pmatrix} PQ_{11} & PQ_{12} & \dots & PQ_{1N} \\ PQ_{21} & PQ_{22} & \dots & PQ_{2N} \\ \vdots & & \ddots & \vdots \\ PQ_{M1} & \dots & & PQ_{MN} \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_M \end{pmatrix} \quad \begin{matrix} \text{音声信号} \\ \text{(モノラル)} \\ S \end{matrix} \quad \text{式(5)}$$

[図20]

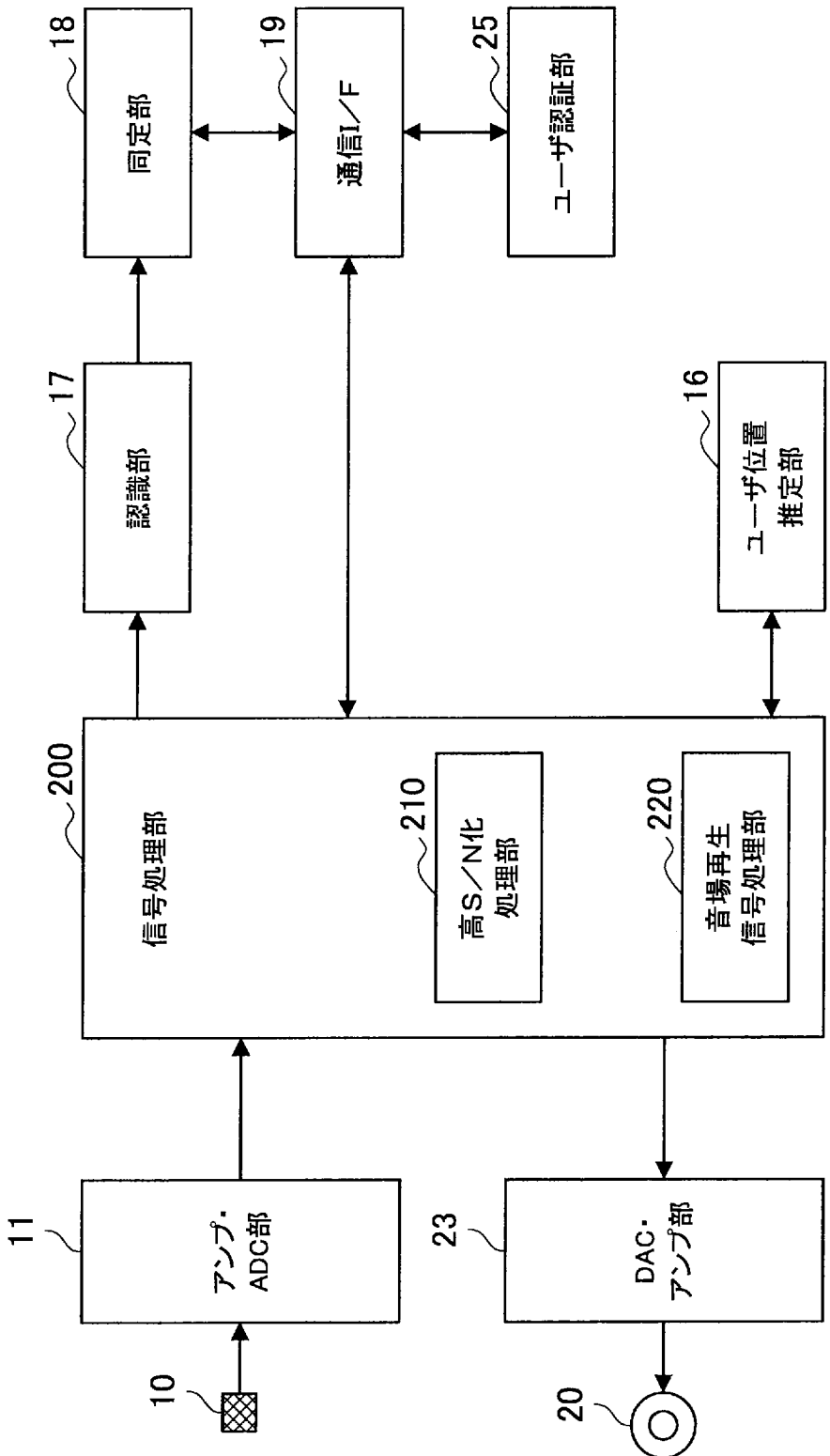


[図21]

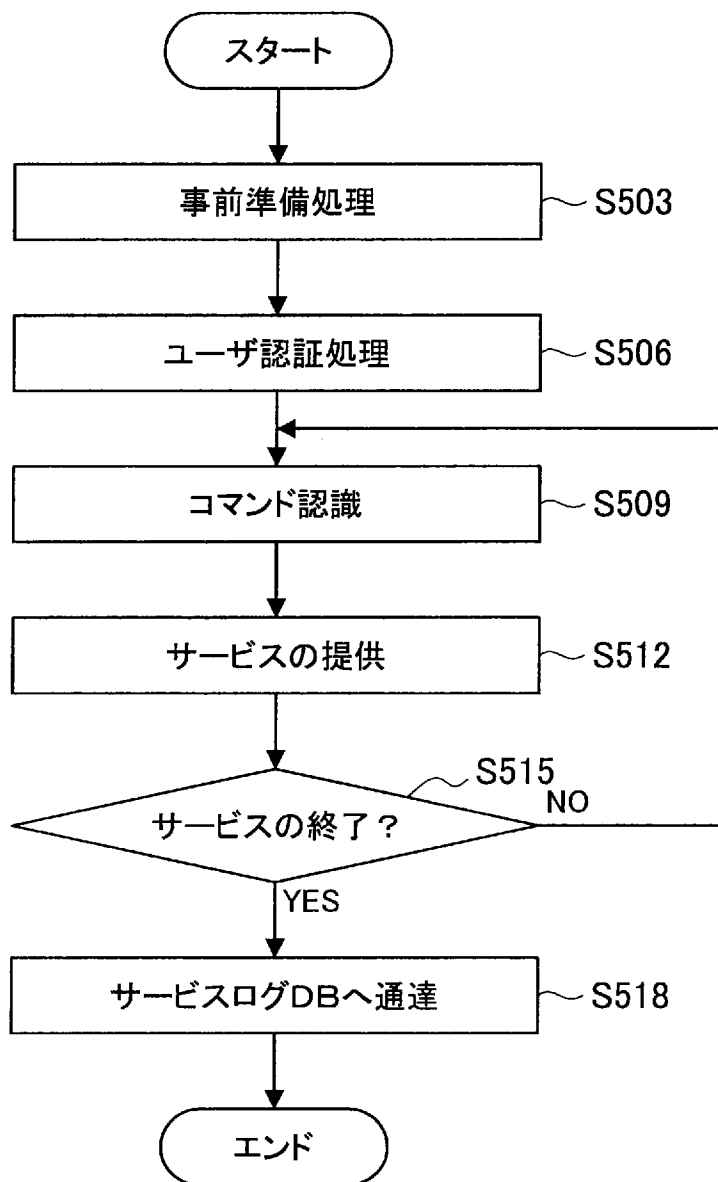


[図22]

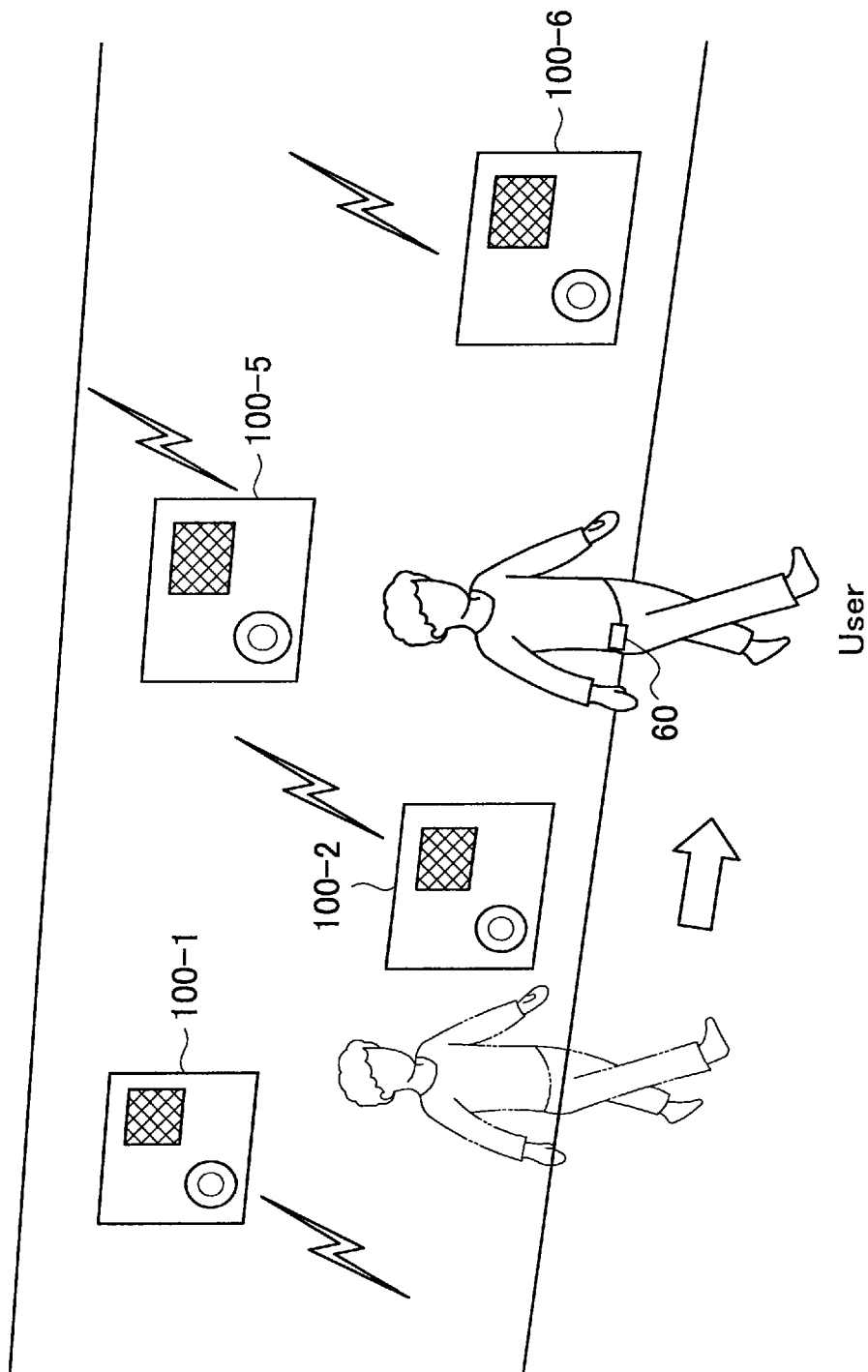
100: デバイス



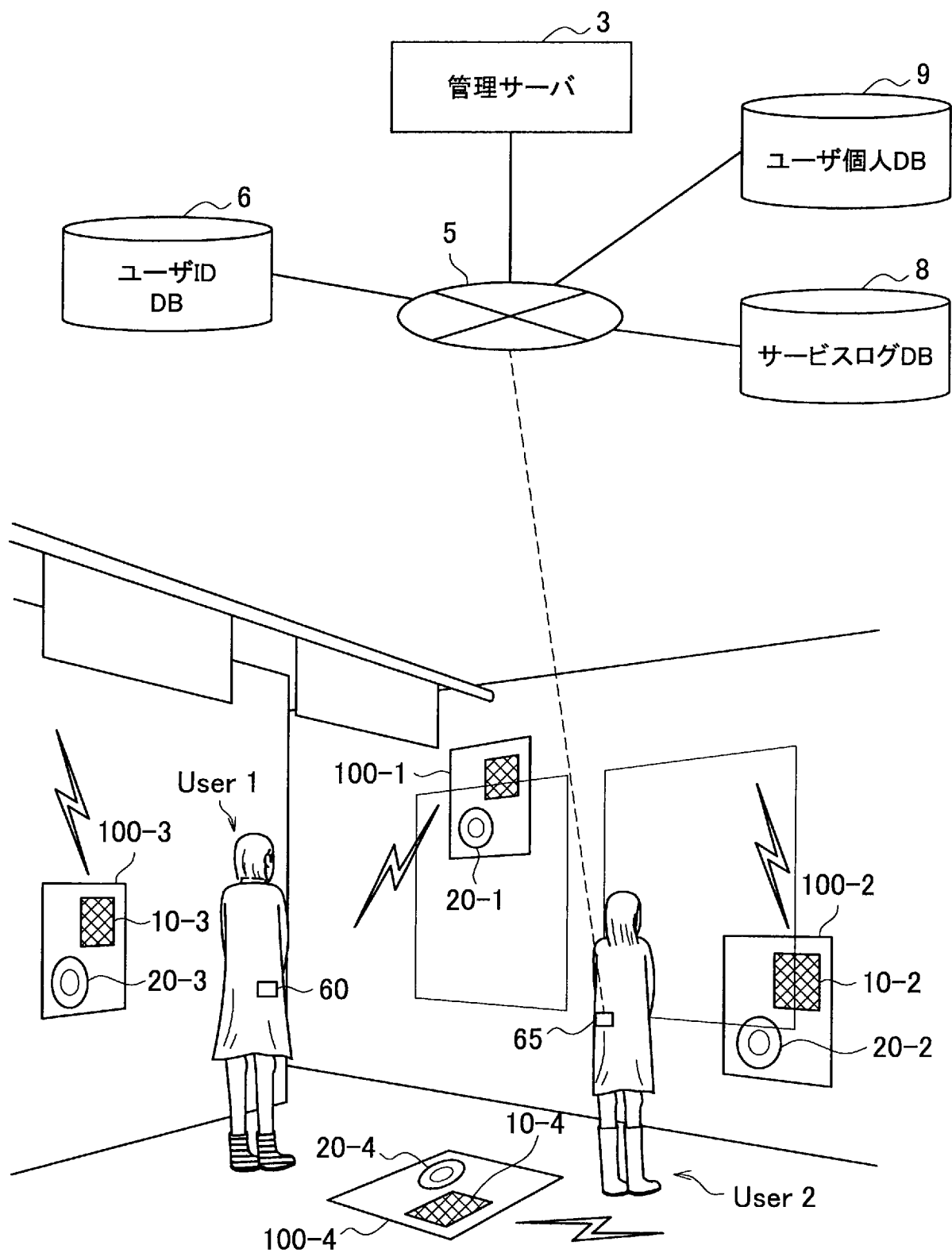
[図23]



[図24]



[図25]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/061894

A. CLASSIFICATION OF SUBJECT MATTER

H04R3/00(2006.01) i, G10L25/51(2013.01) i, H04R1/40(2006.01) i, H04S5/02(2006.01) i

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, G10L25/51, H04R1/40, H04S5/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2008-227773 A (Advanced Telecommunications Research Institute International et al.), 25 September 2008 (25.09.2008), paragraphs [0001] to [0088]; fig. 1 to 11 (Family: none)	1-16
Y	JP 9-261351 A (Nippon Telegraph and Telephone Corp.), 03 October 1997 (03.10.1997), paragraphs [0011] to [0014]; fig. 2, 6 to 8 (Family: none)	1-16
Y	JP 2006-155472 A (Kyocera Corp.), 15 June 2006 (15.06.2006), paragraphs [0019] to [0029]; fig. 1 to 6 (Family: none)	1-16

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
28 May, 2013 (28.05.13)Date of mailing of the international search report
11 June, 2013 (11.06.13)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/061894

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 2010-130411 A (Nippon Telegraph and Telephone Corp.), 10 June 2010 (10.06.2010), paragraph [0004]; fig. 1 (Family: none)	6 1-5, 7-16
Y A	JP 2010-170166 A (Sony Corp.), 05 August 2010 (05.08.2010), paragraphs [0275] to [0276]; fig. 23 & US 2010/0182231 A1 & CN 101794180 A	11 1-10, 12-16

A. 発明の属する分野の分類（国際特許分類（I P C））
 Int.Cl. H04R3/00(2006.01)i, G10L25/51(2013.01)i, H04R1/40(2006.01)i, H04S5/02(2006.01)i

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（I P C））
 Int.Cl. H04R3/00, G10L25/51, H04R1/40, H04S5/02

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 1 9 2 2 - 1 9 9 6 年
 日本国公開実用新案公報 1 9 7 1 - 2 0 1 3 年
 日本国実用新案登録公報 1 9 9 6 - 2 0 1 3 年
 日本国登録実用新案公報 1 9 9 4 - 2 0 1 3 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	JP 2008-227773 A（株式会社国際電気通信基礎技術研究所 外1名） 2008.09.25, 段落[0001]-[0088], [図1]-[図11]（ファミリーなし）	1-16
Y	JP 9-261351 A（日本電信電話株式会社）1997.10.03, 段落[0011]-[0014], [図2], [図6]-[図8]（ファミリーなし）	1-16
Y	JP 2006-155472 A（京セラ株式会社）2006.06.15, 段落[0019]-[0029], [図1]-[図6]（ファミリーなし）	1-16

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	の日の後に公表された文献
「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	「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矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「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	「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
「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）	「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
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日 2 8 . 0 5 . 2 0 1 3	国際調査報告の発送日 1 1 . 0 6 . 2 0 1 3
-----------------------------------	-----------------------------------

国際調査機関の名称及びあて先 日本国特許庁（I S A / J P） 郵便番号1 0 0 - 8 9 1 5 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 菊池 充 電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 8 9	5 Z	4 5 4 5
-----------------------------------------------------------------------------------	--------------------------------------------------------------------	-----	---------

C (続き) . 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 A	JP 2010-130411 A (日本電信電話株式会社) 2010. 06. 10, 段落[0004], [図 1] (ファミリーなし)	6 1-5, 7-16
Y A	JP 2010-170166 A (ソニー株式会社) 2010. 08. 05, 段落[0275]-[0276], [図 23] & US 2010/0182231 A1 & CN 101794180 A	11 1-10, 12-16