

(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6月20日(20.06.2024)



(10) 国際公開番号

WO 2024/127986 A1

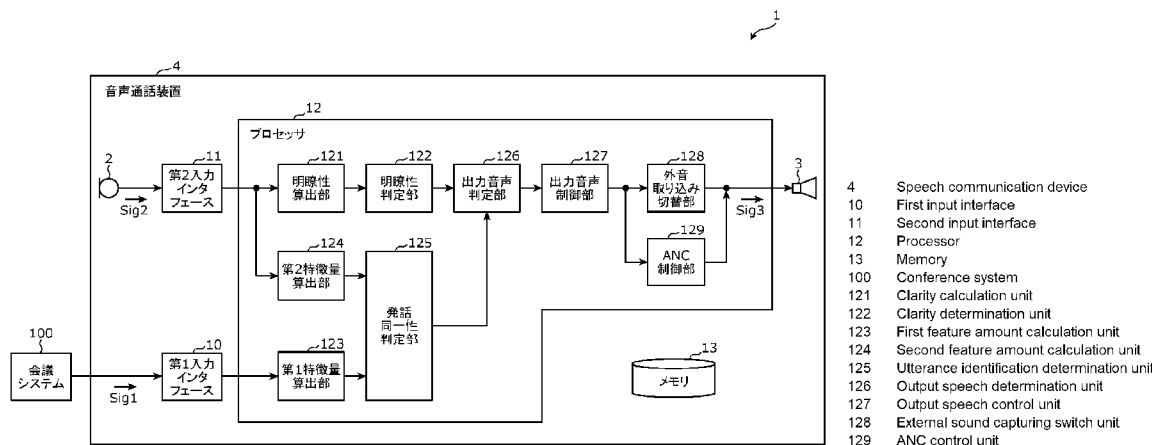
- (51) 国際特許分類:
H04R 3/02 (2006.01)
- (21) 国際出願番号: PCT/JP2023/042673
- (22) 国際出願日: 2023年11月29日(29.11.2023)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2022-198122 2022年12月12日(12.12.2022) JP
- (71) 出願人: パナソニックIPマネジメント株式会社 (PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.) [JP/JP]; 〒5710057 大阪府門真市元町2番6号 Osaka (JP).
- (72) 発明者: 横田 健治 (YOKOTA, Kenji). 岡崎 隆義 (OKAZAKI, Takayosi).
- (74) 代理人: 新居 広守, 外 (NII, Hiromori et al.); 〒5320011 大阪府大阪市淀川区西中島5丁

目3番10号イトーピア新大阪ビル6階新居国際特許事務所内 Osaka (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.
- (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,

(54) Title: SPEECH PROCESSING SYSTEM, SPEECH PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 音声処理システム、音声処理方法、及びプログラム



(57) Abstract: A speech processing system (1) comprises a first input interface (10), a second input interface (11), and a processor (12). The first input interface (10) acquires a first speech signal (Sig1) via a communication line. The second input interface (11) acquires a second speech signal (Sig2) based on speech collected by means of a microphone (2). The processor (12) outputs, to a loudspeaker (3), an output speech signal (Sig3) based on the first speech signal (Sig1) and the second speech signal (Sig2). When both a first condition and a second condition are satisfied, where the first condition is that the first speech signal (Sig1) and the second speech signal (Sig2) both include a speech signal based on speech uttered by the same person, and the second condition is that the second speech signal (Sig2) is clear, the processor (12) includes, in the output speech signal (Sig3), a signal in which a component corresponding to the first speech signal (Sig1) is reduced.

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：音声処理システム（1）は、第1入力I/F（10）と、第2入力I/F（11）と、プロセッサ（12）と、を備える。第1入力I/F（10）は、通信回線を介した第1音声信号（Sig1）を取得する。第2入力I/F（11）は、マイクロホン（2）で収音された音声に基づく第2音声信号（Sig2）を取得する。プロセッサ（12）は、第1音声信号（Sig1）及び第2音声信号（Sig2）に基づく出力音声信号（Sig3）をスピーカ（3）に出力する。プロセッサ（12）は、第1音声信号（Sig1）及び第2音声信号（Sig2）がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、第2音声信号（Sig2）が明瞭であるという第2条件をいずれも満たす場合、第1音声信号（Sig1）に対応する成分を低減した信号を出力音声信号（Sig3）に含める。

明 細 書

発明の名称：音声処理システム、音声処理方法、及びプログラム
技術分野

[0001] 本開示は、スピーカの発する音声进行处理するための音声処理システム等に関する。

背景技術

[0002] 例えば、特許文献1には、音声コミュニケーション端末が開示されている。この音声コミュニケーション端末は、多地点音声通信システムに参加する複数の端末のうちの少なくとも1つの音声出力を制御する装置であって、音声配置部と、対話者管理部と、を有する。音声配置部は、他の端末からの音声が出力される際の音源配置を設定する。対話者管理部は、複数の端末の中から、発話者とその相手である対話者とを検出し、検出された発話者及び対話者の組み合わせに基づいて会話グループを検出する。音声配置部は、検出された会話グループの変化に応じて音源配置の設定を変更する。

先行技術文献

特許文献

[0003] 特許文献1：特開2012-108587号公報

発明の概要

発明が解決しようとする課題

[0004] 本開示は、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい音声処理システム等を提供する。

課題を解決するための手段

[0005] 本開示の一態様に係る音声処理システムは、第1入力インタフェースと、第2入力インタフェースと、信号処理回路と、を備える。前記第1入力インタフェースは、通信回線を介した第1音声信号を取得する。前記第2入力インタフェースは、マイクロホンで収音された音声に基づく第2音声信号を取得する。前記信号処理回路は、前記第1音声信号及び前記第2音声信号に基

づく出力音声信号をスピーカに出力する。前記信号処理回路は、前記第1音声信号及び前記第2音声信号がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、前記第2音声信号が明瞭であるという第2条件をいずれも満たす場合、前記第1音声信号に対応する成分を低減した信号を前記出力音声信号に含める。

[0006] 本開示の一態様に係る音声処理方法では、通信回線を介した第1音声信号を取得する。前記音声処理方法では、マイクロホンで収録された音声に基づく第2音声信号を取得する。前記音声処理方法では、前記第1音声信号及び前記第2音声信号がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、前記第2音声信号が明瞭であるという第2条件をいずれも満たす場合、前記第1音声信号に対応する成分を低減した信号を出力音声信号に含めてスピーカに出力する。

[0007] 本開示の一態様に係るプログラムは、1以上のプロセッサに、前記音声処理方法を実行させる。

発明の効果

[0008] 本開示における音声処理システム等によれば、オフラインでの会話、及びオンラインでの会話の両方が混在した環境においても、会話の快適性が損なわれにくい、という利点がある。

図面の簡単な説明

[0009] [図1]図1は、会議システムを用いたコミュニケーションの課題の説明図である。

[図2]図2は、実施の形態に係る音声処理システムを含む全体構成の一例を示すブロック図である。

[図3]図3は、第2音声信号の明瞭性を判定するための第1判定動作の説明図である。

[図4]図4は、第2音声信号の明瞭性を判定するための第2判定動作の説明図である。

[図5]図5は、実施の形態に係る音声処理システムの動作例を示すフローチャ

ートである。

[図6]図6は、第2音声信号の明瞭性を判定するために必要なパラメータの算出例を示すフローチャートである。

[図7]図7は、発話同一性を判定するために必要なパラメータの算出例を示すフローチャートである。

[図8]図8は、実施の形態に係る音声処理システムの動作例の概要の説明図である。

[図9]図9は、実施の形態に係る音声処理システムの利点の説明図である。

発明を実施するための形態

[0010] [1. 本開示の基礎となった知見]

まず、発明者の着眼点が、下記に説明される。

[0011] 従来、例えばMCU (Multipoint Control Unit : 多地点制御装置) を介した会議システム、又はZoom (登録商標) 等のWeb会議サービスを用いて、複数の拠点間で同時に会議等のコミュニケーションを行う技術が知られている。このようなコミュニケーションにおいては、各参加者は、マイクロホン及びスピーカを備えた機器 (例えば、ヘッドセット等) を装着することで、他の参加者との間で会話する。また、近年では、各参加者は、XR (x Reality) 技術を用いた機器 (例えば、ヘッドマウントディスプレイ、又はスマートグラス等) を装着することで、同じ仮想空間上で、又は同じ仮想空間を見ながら、他の参加者との間で会話することも可能である。このようなコミュニケーションにおいては、同じ拠点に複数の参加者が存在する場合があります、以下のような課題がある。

[0012] 図1は、会議システムを用いたコミュニケーションの課題の説明図である。図1において、会議システム100は、上述のMCUを介した会議システム、又はWeb会議サービスが提供するサーバである。図1に示す例では、第1拠点A1にいる2人のユーザU1、U2と、第2拠点A2にいるユーザU3と、第3拠点A3にいるユーザU4とが、会議システム100を用いてオンラインで会議を行っていることを表している。ユーザU1~U4は、い

いずれも会議システム100を介して送信される音声信号に基づく音声スピーカーから出力されることで、他のユーザが発した音声を聴くことが可能である。一例として、ユーザU1が「こんにちは」という音声を発した場合、他のユーザU2、U3、U4は、会議システム100を介して送信される音声信号に基づく当該音声スピーカーから出力されることで、「こんにちは」というユーザU1が発した音声を聴くことが可能である。

[0013] ここで、図1に示す例では、第1拠点A1には、2人のユーザU1、U2がいる。このため、第1拠点A1では、2人のユーザU1、U2のうち一方のユーザが発した音声を、他方のユーザが会議システム100を介さずに直接的に聴くことが可能である。この場合、例えば第1拠点A1にいるユーザU2が「こんにちは」という音声V2を発した場合、ユーザU1は、ユーザU2が発した音声V2を直接的に聴くと共に、会議システム100を介して送信されるユーザU2が発した音声V1を聴くことになる。

[0014] 上述のように、同じ拠点に複数のユーザが存在する場合、当該拠点にいるユーザは、当該拠点にいる他のユーザからの直接的な音声と会議システム100を介した音声との両方を聴くことになるため、他のユーザが発した音声を聴き取りづらい、という課題がある。また、当該拠点にいる他のユーザからの会議システム100を介した音声は、他のユーザからの直接的な音声から遅れてユーザの耳に届く。このため、ユーザが他のユーザからの直接的な音声を聴いて発言しようとするタイミングで、他のユーザからの会議システム100を介した音声ユーザの耳に届くことで、ユーザの発言が妨害されて話しづらい、という課題がある。このように、同じ拠点に複数のユーザが存在する環境においては、会話の快適性が損なわれやすい、という課題がある。

[0015] 以上を鑑み、発明者は本開示を創作するに至った。

[0016] 以下、実施の形態について、図面を参照しながら具体的に説明する。なお、以下で説明する実施の形態は、いずれも包括的又は具体的な例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要

素の配置位置及び接続形態、ステップ、ステップの順序等は、一例であり、本開示を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、独立請求項に記載されていない構成要素については、任意の構成要素として説明される。

[0017] なお、各図は模式図であり、必ずしも厳密に図示されたものではない。また、各図において、実質的に同一の構成に対しては同一の符号を付し、重複する説明は省略又は簡略化される場合がある。

[0018] (実施の形態)

[2. 構成]

[2-1. 全体構成]

まず、実施の形態に係る音声処理システムを含む全体構成について図2を用いて説明する。図2は、実施の形態に係る音声処理システムを含む全体構成を示すブロック図である。音声処理システム1は、外部から音声信号を取得した場合に、音声信号に基づく音声をスピーカ3から出力させるためのシステムである。実施の形態では、音声処理システム1は、音声通話装置4で実現される。

[0019] 音声通話装置4は、インターネット等のネットワークを介して、会議システム100と通信可能である。なお、音声通話装置4は、LAN (Local Area Network) を介して会議システム100と通信してもよい。

[0020] 音声通話装置4は、ユーザの頭部又は首に装着される機器であって、密閉型の音声通話装置と、開放型の音声通話装置と、密閉型と開放型とを切り替え可能な音声通話装置と、に分けられる。密閉型の音声通話装置は、ユーザの耳穴(鼓膜)を覆う機器であって、例えばイヤホン型のヘッドセット、又はヘッドホン型のヘッドセット等を含む。開放型の音声通話装置は、ユーザの耳穴を覆わない機器であって、例えばネックスピーカ、又はXR用のゴーグル型のウェアラブルデバイス等を含む。密閉型と開放型とを切り替え可能な音声通話装置は、ユーザの耳穴を覆う機能と、ユーザの耳穴を覆わない機

能との2つを切り替え可能な機器であって、例えばハウジング部のプレート開閉による切り替えが可能なイヤホン型のヘッドセット、又はヘッドホン型のヘッドセット等を含む。なお、音声通話装置4は、音声処理等を実行する本体部と、マイクロホン及びスピーカを含むヘッドセット部とが一体に構成されていてもよいし、それぞれ別体で構成されていてもよい。

[0021] 音声処理システム1は、密閉型の音声通話装置、開放型の音声通話装置、及び密閉型と開放型とを切り替え可能な音声通話装置のいずれにも適用可能である。以下では、一例として、音声通話装置4が密閉型の音声通話装置であることとして説明する。

[0022] 会議システム100は、既に述べたように、例えばMCUを介した会議システム、又はWeb会議サービスが提供するサーバである。会議システム100は、任意のユーザが装着する音声通話装置4から出力される音声信号を受信すると、受信した音声信号に対して適宜の補正処理を実行した後に、補正後の音声信号を1以上の他のユーザがそれぞれ装着する1以上の音声通話装置4へ送信する。補正処理は、例えば受信した音声信号に含まれる雑音を低減するための雑音抑圧処理を含み得る。また、補正処理は、例えば受信した音声信号における人の可聴周波数領域を強調する周波数補正処理を含み得る。なお、会議システム100は、受信した音声信号に対して補正処理を実行しなくてもよい。

[0023] [2-2. 音声通話装置（音声処理システム）の構成]

次に、音声通話装置4（音声処理システム1）の構成について具体的に説明する。音声通話装置4は、図2に示すように、マイクロホン2と、第1入力インタフェース（以下、「第1入力I/F（Interface）」と記載）10と、第2入力インタフェース（以下、「第2入力I/F」と記載）11と、プロセッサ12と、メモリ13と、スピーカ3と、を備えている。

[0024] マイクロホン2は、音声通話装置4の周囲の音声を取得し、取得した音声に基づいて第2音声信号Sig2を出力する収音デバイスである。具体的には、マイクロホン2は、コンデンサマイク、ダイナミックマイク、又はME

MS (Micro Electro Mechanical Systems) マイク等であるが、特に限定されない。また、マイクロホンは、無指向性であってもよいし、指向性を有していてもよい。

[0025] スピーカ3は、プロセッサ12から出力される出力音声信号Sig3に基づく音声を出力する。スピーカ3は、音声通話装置4を装着したユーザの耳穴へ向けて音波を発するスピーカであるが、例えば骨伝導スピーカであってもよい。

[0026] 第1入力I/F10は、例えば無線通信インタフェースであり、Wi-Fi（登録商標）等の無線通信規格に基づいて、ネットワークを介して会議システム100と通信を行うことにより、会議システム100から送信される第1音声信号Sig1を受信する。言い換えれば、第1入力I/F10は、通信回線を介して第1音声信号Sig1を取得する。第1音声信号Sig1は、主として他のユーザが発した音声に基づく音声信号である。第1入力I/F10は、取得した第1音声信号Sig1をプロセッサ12に出力する。

[0027] 第2入力I/F11は、マイクロホン2から出力される第2音声信号Sig2を受信するインタフェースである。言い換えれば、第2入力I/F11は、マイクロホン2で収音された音声に基づく第2音声信号Sig2を取得する。第2入力I/F11は、取得した第2音声信号Sig2をプロセッサ12に出力する。

[0028] プロセッサ12は、例えばCPU (Central Processing Unit) 又はDSP (Digital Signal Processor) 等である。プロセッサ12は、第1入力I/F10で取得した第1音声信号Sig1、及び第2入力I/F11で取得した第2音声信号Sig2に基づく出力音声信号Sig3をスピーカ3に出力する情報処理を行う。上述の情報処理は、プロセッサ12がメモリ13に記憶されたコンピュータプログラムを実行することにより実現される。プロセッサ12は、音声処理システム1の信号処理回路の一例である。

[0029] プロセッサ12には、機能的な構成要素として、明瞭性算出部121、明

瞭性判定部122、第1特徴量算出部123、第2特徴量算出部124、発話同一性判定部125、出力音声判定部126、出力音声制御部127、外音取り込み切替部128、及びANC (Active Noise Cancelling) 制御部129が含まれる。上述の各機能は、例えばプロセッサ12がメモリ13に記憶されたコンピュータプログラムを実行することにより実現される。

[0030] 明瞭性算出部121は、第2音声信号Sig2が明瞭であるか否かを明瞭性判定部122で判定する際に用いる第2音声信号Sig2の特徴量を算出する。ここで、第2音声信号Sig2が明瞭であるとは、第2音声信号Sig2における人の音声に対応する周波数帯域（以下、「音声帯域」という）のSNR (Signal to Noise Ratio) が閾値よりも高く、かつ、人の音声の特徴が明確であることをいう。言い換えれば、第2音声信号Sig2が明瞭であるとは、スピーカ3から出力される第2音声信号Sig2に基づく音声を人が聴いた場合に、当該人がその内容を理解することができる程度であることをいう。

[0031] 明瞭性算出部121は、第2音声信号Sig2の特徴量として、上記SNRと、第2音声信号Sig2のスペクトル包絡と、を算出する。具体的には、明瞭性算出部121は、第2音声信号Sig2に対して適宜の信号処理を行うことにより、第2音声信号Sig2のスペクトルコントラストを算出し、算出したスペクトルコントラストに基づいて、第2音声信号Sig2における音声帯域のSNRを算出する。また、明瞭性算出部121は、第2音声信号Sig2のMFCC (Mel-Frequency Cepstral Coefficient、メル周波数ケプストラム係数) を算出する。MFCCは、音声認識等で特徴量として用いられるケプストラムの係数であり、メルフィルタバンクを用いて圧縮されたパワースペクトルを対数パワースペクトルに変換し、対数パワースペクトルに逆離散コサイン変換を適用することで得られる。MFCCは、スペクトル包絡に相当する。

[0032] 明瞭性判定部122は、明瞭性算出部121で算出した第2音声信号Si

g 2 の特徴量を用いて、第 2 音声信号 S i g 2 が明瞭であるという第 2 条件を満たすか否かを判定する。明瞭性判定部 1 2 2 による判定動作については、後述する [2 - 3 . 明瞭性の判定] にて詳細に説明する。

[0033] 第 1 特徴量算出部 1 2 3 は、第 1 音声信号 S i g 1 及び第 2 音声信号 S i g 2 がいずれも同一の人物の発する音声に基づく音声信号であるか否かを発話同一性判定部 1 2 5 で判定する際に用いる、第 1 音声信号 S i g 1 の特徴量を算出する。第 1 特徴量算出部 1 2 3 は、第 1 音声信号 S i g 1 の特徴量として、第 1 音声信号 S i g 1 の基本周波数と、第 1 音声信号 S i g 1 のスペクトル包絡と、を算出する。具体的には、第 1 特徴量算出部 1 2 3 は、第 1 音声信号 S i g 1 のケプストラムを算出し、算出したケプストラムから第 1 音声信号 S i g 1 の基本周波数を算出する。ケプストラムは、フーリエ変換を適用することで第 1 音声信号 S i g 1 のパワースペクトルを算出し、算出したパワースペクトルを対数パワースペクトルに変換し、対数パワースペクトルに更にフーリエ変換を適用することで得られる。また、第 1 特徴量算出部 1 2 3 は、第 1 音声信号 S i g 1 の MFCC を算出することにより、スペクトル包絡を算出する。また、第 1 特徴量算出部 1 2 3 は、算出したスペクトル包絡から、第 1 音声信号 S i g 1 において母音が出現するタイミングを算出する。

[0034] 第 2 特徴量算出部 1 2 4 は、第 1 音声信号 S i g 1 及び第 2 音声信号 S i g 2 がいずれも同一の人物の発する音声に基づく音声信号であるか否かを発話同一性判定部 1 2 5 で判定する際に用いる、第 2 音声信号 S i g 2 の特徴量を算出する。第 2 特徴量算出部 1 2 4 は、第 2 音声信号 S i g 2 の特徴量として、第 2 音声信号 S i g 2 の基本周波数と、第 2 音声信号 S i g 2 のスペクトル包絡と、を算出する。具体的には、第 2 特徴量算出部 1 2 4 は、第 2 音声信号 S i g 2 のケプストラムを算出し、算出したケプストラムから第 2 音声信号 S i g 2 の基本周波数を算出する。また、第 2 特徴量算出部 1 2 4 は、第 2 音声信号 S i g 2 の MFCC を算出することにより、スペクトル包絡を算出する。また、第 2 特徴量算出部 1 2 4 は、算出したスペクトル包

絡から、第2音声信号S i g 2において母音が出現するタイミングを算出する。

[0035] なお、第2音声信号S i g 2のスペクトル包絡は、明瞭性算出部121及び第2特徴量算出部124のいずれか一方でのみ算出すればよい。実施の形態では、第2音声信号S i g 2のスペクトル包絡は、第2特徴量算出部124で算出することとして説明する。したがって、明瞭性算出部121は、第2音声信号S i g 2のスペクトル包絡は算出しなくてもよい。なお、明瞭性算出部121及び第2特徴量算出部124のいずれか一方のみで第2音声信号S i g 2のスペクトル包絡を算出する場合、算出したスペクトル包絡は、他方で共有される。

[0036] 発話同一性判定部125は、第1特徴量算出部123で算出した第1音声信号S i g 1の特徴量と、第2特徴量算出部124で算出した第2音声信号S i g 2の特徴量とを用いて、第1音声信号S i g 1及び第2音声信号S i g 2がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件を満たすか否かを判定する。実施の形態では、発話同一性判定部125は、(i)第1音声信号S i g 1の基本周波数と第2音声信号S i g 2の基本周波数とが同じであり、かつ、(ii)第1音声信号S i g 1において母音が出現するタイミング及び種類と第2音声信号S i g 2において母音が出現するタイミング及び種類とが同じである場合、第1条件を満たすと判定する。一方、発話同一性判定部125は、上記の(i)及び(ii)のうちの少なくとも一方を満たさない場合、第1条件を満たさないと判定する。

[0037] 各音声信号において母音が出現するタイミングは、既に述べたように、各音声信号のスペクトル包絡から検出することが可能である。なお、第1音声信号S i g 1は、通信回線を経由するため、第2音声信号S i g 2よりも遅延して第1入力I/F10で取得される。このため、発話同一性判定部125は、上記の遅延を考慮して、(ii)について判定する。

[0038] このように、プロセッサ12(発話同一性判定部125)は、第1音声信号S i g 1における母音に対応する成分と、第2音声信号S i g 2における

母音に対応する成分との相関に基づいて、第1条件を満たすか否かを判定する。具体的には、発話同一性判定部125は、(i)第1音声信号S i g 1の基本周波数と第2音声信号S i g 2の基本周波数との差分を算出し、算出した差分が閾値以下であり、かつ、(i i)第1音声信号S i g 1において母音が出現する時間と第2音声信号S i g 2において母音が出現する時間との差分を算出し、算出した差分が閾値以下であり、かつ第1音声信号S i g 1において出現する母音の種類と第2音声信号S i g 2において出現する母音の種類とが同じある場合、第1条件を満たすと判定する。一方、発話同一性判定部125は、(i)又は(i i)において算出した差分が閾値を上回った場合、第1条件を満たしていないと判定する。なお、発話同一性判定部125は、(i i)の判定において、第1音声信号S i g 1において算出されたスペクトル包絡と第2音声信号S i g 2において算出されたスペクトル包絡との相関係数を算出し、算出した相関係数が閾値以下であるか否かを判定してもよい。また、発話同一性判定部125は、(i i)の判定において、いずれか1つの条件を満たした場合、(i i)を満たすとしてもよい。

[0039] なお、発話同一性判定部125は、(i)を満たすか否かのみで、第1条件を満たすか否かを判定してもよいし、(i i)を満たすか否かのみで第1条件を満たすか否かを判定してもよい。また、発話同一性判定部125は、(i)及び(i i)の少なくとも一方を満たす場合に第1条件を満たすと判定し、(i)及び(i i)の両方を満たさない場合に第1条件を満たさないと判定してもよい。

[0040] また、発話同一性判定部125は、(i i)に代えて、第1音声信号S i g 1において母音が連続して出現するパターンと、第2音声信号S i g 2において母音が連続して出現するパターンとが同じであるか否かを判定してもよい。この場合、発話同一性判定部125は、上記の遅延を考慮しなくてもよい。

[0041] ここで、第1音声信号S i g 1の波形と、第2音声信号S i g 2の波形との類似度に基づいて、波形の類似度が閾値以上であれば第1条件を満たし、

波形の類似度が閾値未満であれば第1条件を満たさない、と判定する方法も考えられる。ここでいう「波形」は、信号の振幅の波形、つまり音圧レベルの波形である。しかしながら、第1音声信号S i g 1は、既に述べたように、会議システム100において補正処理が実行された音声信号であるため、第1音声信号S i g 1の波形と第2音声信号S i g 2の波形とが異なってしまふ。このため、発話同一性判定部125は、上述のように、波形の類似度に基づく方法とは異なる方法で、第1条件を満たすか否かを判定している。

[0042] なお、会議システム100において第1音声信号S i g 1に対して補正処理が実行されない場合であれば、発話同一性判定部125は、波形の類似度に基づいて第1条件を満たすか否かを判定してもよい。例えば、発話同一性判定部125は、第1音声信号S i g 1の音圧レベルの変化と、第2音声信号S i g 2の音圧レベルの変化とが概ね一致するか否かに基づいて、言い換えれば第1音声信号S i g 1の振幅の包絡線と第2音声信号S i g 2の振幅の包絡線との相関に基づいて、第1条件を満たすか否かを判定してもよい。

[0043] 出力音声判定部126は、明瞭性判定部122における第2条件を満たすか否かの判定結果、及び発話同一性判定部125における第1条件を満たすか否かの判定結果に基づいて、第1状況及び第2状況のいずれの状況にあるかを判定する。第1状況は、ユーザと他のユーザとの間の距離が比較的近く、ユーザが他のユーザの発した音声を直接的に聴き取りやすい状況である。第2状況は、第1状況以外の状況である。第2状況は、例えばユーザと他のユーザとの間の距離が比較的遠く、ユーザが他のユーザの発した音声を直接的に聞き取りにくい状況を含む。出力音声判定部126は、第1条件及び第2条件をいずれも満たす場合、第1状況にあると判定する。一方、出力音声判定部126は、第1条件及び第2条件の少なくとも一方を満たさない場合、第2状況にあると判定する。

[0044] 出力音声制御部127は、出力音声判定部126の判定結果に基づいて、出力音声信号S i g 3に含める音声信号を制御する。具体的には、出力音声制御部127は、出力音声判定部126において第1状況にあると判定した

場合、スピーカ3から出力される第1音声信号S i g 1に基づく音声の音量を下げるように制御する。ここでいう「第1音声信号S i g 1に基づく音声の音量を下げる」とは、第1音声信号S i g 1に基づく音声の音量を、第2状況における第1音声信号S i g 1に基づく音声の音量（言い換えれば、デフォルトの音量）よりも低い音量にすることをいう。また、出力音声制御部127は、外音取り込み切替部128を制御することで外音取り込み機能をオンし、かつ、ANC制御部129を制御することでノイズキャンセル機能をオフする。

[0045] つまり、プロセッサ12（出力音声制御部127）は、第1条件及び第2条件をいずれも満たす場合、第1音声信号S i g 1に対応する成分を低減した信号を出力音声信号S i g 3に含める。なお、実施の形態では、プロセッサ12は、第1音声信号S i g 1に基づく音声の音量を下げることで、第1音声信号S i g 1に対応する成分を低減させているが、これに限られない。例えば、プロセッサ12は、第1条件及び第2条件をいずれも満たす場合、出力音声信号S i g 3に第1音声信号S i g 1を含めなくてもよい。また、例えば、プロセッサ12は、第1条件及び第2条件をいずれも満たす場合、第2音声信号S i g 2に基づいて第1音声信号S i g 1に対してサプレッション処理を実行し、処理後の第1音声信号S i g 1を出力音声信号S i g 3に含めてもよい。

[0046] また、プロセッサ12（出力音声制御部127）は、第1条件及び第2条件をいずれも満たす場合、外音取り込み機能をオンにする、つまり第2音声信号S i g 2を出力音声信号S i g 3に含める。なお、出力音声信号S i g 3に含める第2音声信号S i g 2は、ノイズ低減処理又はイコライジング処理等の音声処理により処理された信号であってもよい。

[0047] 一方、出力音声制御部127は、出力音声判定部126において第2状況にあると判定した場合、スピーカ3から出力される第1音声信号S i g 1に基づく音声の音量をデフォルトの音量とするように制御する。また、出力音声制御部127は、外音取り込み切替部128を制御することで外音取り込

み機能をオフし、かつ、ANC制御部129を制御することでノイズキャンセル機能をオンする。

[0048] つまり、プロセッサ12（出力音声制御部127）は、第1条件及び第2条件のうちの少なくとも一方を満たさない場合、第1音声信号S i g 1を出力音声信号S i g 3に含め、かつ、外音取り込み機能をオフにする、つまり第2音声信号S i g 2を出力音声信号S i g 3に含めない。なお、出力音声信号S i g 3に含める第1音声信号S i g 1は、ノイズ低減処理又はイコライジング処理等の音声処理により処理された信号であってもよい。

[0049] また、プロセッサ12（出力音声制御部127）は、第1条件及び第2条件のうちの少なくとも一方を満たさない場合、ノイズキャンセル機能をオンにする、つまり更に第2音声信号S i g 2とは逆位相の音声信号を出力音声信号S i g 3に含める。

[0050] 外音取り込み切替部128は、出力音声制御部127に制御されることにより、ユーザの周囲の音声を取り込む外音取り込み機能のオン／オフを切り替える。外音取り込み機能がオンの場合、スピーカ3は、第2音声信号S i g 2を含む出力音声信号S i g 3に基づく音声を出力する。一方、外音取り込み機能がオフの場合、スピーカ3は、第2音声信号S i g 2を含まない出力音声信号S i g 3に基づく音声を出力する。

[0051] ANC制御部129は、出力音声制御部127に制御されることにより、ノイズキャンセル機能のオン／オフを切り替える。ノイズキャンセル機能がオンの場合、ANC制御部129は、第2音声信号S i g 2とは逆位相の音声信号を生成し、生成した音声信号を出力音声信号S i g 3に含める。この場合、スピーカ3は、第2音声信号S i g 2とは逆位相の音声信号に基づく音声を出力する。これにより、ユーザの耳元において、第2音声信号S i g 2の元になる音声と、第2音声信号S i g 2とは逆位相の音声信号に基づく音声とが互いに打ち消しあうため、ユーザには、これらの音声が概ね聴こえなくなる。一方、ノイズキャンセル機能がオフの場合、ANC制御部129は、第2音声信号S i g 2とは逆位相の音声信号を生成しない。

[0052] メモリ 13 は、プロセッサ 12 が実行するコンピュータプログラム、及び各種機能の実現に必要な情報等が記憶される記憶装置である。メモリ 13 は、例えば半導体メモリ等によって実現される。なお、メモリ 13 は、プロセッサ 12 の外付けメモリではなく、プロセッサ 12 の内蔵メモリとして実現されてもよい。

[0053] [2-3. 明瞭性の判定]

以下、明瞭性判定部 122 による第 2 音声信号 S i g 2 が明瞭であるか否かの判定動作について詳細に説明する。実施の形態では、明瞭性判定部 122 は、第 1 判定動作及び第 2 判定動作を実行し、いずれの判定動作においても明瞭であると判定した場合に、第 2 音声信号 S i g 2 が明瞭である、つまり第 2 条件を満たすと判定する。一方、明瞭性判定部 122 は、第 1 判定動作及び第 2 判定動作のうちの少なくとも一方で明瞭でないと判定した場合、第 2 音声信号 S i g 2 が明瞭でない、つまり第 2 条件を満たさないと判定する。

[0054] 図 3 は、第 2 音声信号 S i g 2 の明瞭性を判定するための第 1 判定動作の説明図である。図 3 は、第 2 音声信号 S i g 2 のスペクトルコントラストを表している。図 3 において、縦軸は第 2 音声信号 S i g 2 の周波数帯域、横軸は時間（単位は「秒」）を表している。また、図 3 において、明暗は S N R の高低を表しており、明るい程 S N R が高く、暗い程 S N R が低いことを表している。

[0055] 第 1 判定動作では、明瞭性判定部 122 は、音声区間（図 3 における矩形枠で囲まれた区間。例えば、コンマ数秒の区間）において、第 2 音声信号 S i g 2 における音声帯域の S N R と閾値とを比較する。そして、第 1 判定動作では、明瞭性判定部 122 は、S N R が閾値よりも高ければ第 2 音声信号 S i g 2 が明瞭であると判定し、S N R が閾値よりも低ければ第 2 音声信号 S i g 2 が明瞭でないと判定する。

[0056] ここで、音声信号における音声帯域の S N R は、例えば音声信号のうちの音声帯域に含まれる各周波数帯域の S N R の代表値として算出することが可

能である。代表値は、例えば平均値、中央値、最大値、又は最頻値等である。また、音声信号における音声帯域のSNRは、例えば音声帯域に含まれる各周波数帯域のSNRの代表値と、音声帯域以外の各周波数帯域のSNRの代表値との比として算出することも可能である。後者では、例えば換気扇の動作音が大きい等によりユーザの周囲が比較的騒がしく、全ての周波数帯域でSNRが比較的高くなるような場合でも、明瞭性判定部122が第2音声信号S i g 2が明瞭であるか否かを判定しやすい。

[0057] 図3の(a)は、矩形枠で囲まれた音声区間において、第2音声信号S i g 2における音声帯域(図3の(a)において矢印で示す帯域)のSNRが閾値よりも低いことを表している。このため、図3の(a)に示す例では、明瞭性判定部122は、第1判定動作において第2音声信号S i g 2が明瞭でないとして判定する。

[0058] 一方、図3の(b)は、矩形枠で囲まれた音声区間において、第2音声信号S i g 2における音声帯域(図3の(b)において矢印で示す帯域)のSNRが閾値よりも高いことを表している。このため、図3の(b)に示す例では、明瞭性判定部122は、第1判定動作において第2音声信号S i g 2が明瞭であると判定する。

[0059] 図4は、第2音声信号S i g 2の明瞭性を判定するための第2判定動作の説明図である。図4は、上述の音声区間における第2音声信号S i g 2のスペクトルを表している。図4において、縦軸は第2音声信号S i g 2の振幅値、横軸は第2音声信号S i g 2の周波数を表している。また、図4において、実線L1はスペクトル包絡を、一点鎖線はスペクトル包絡の傾向を表している。

[0060] 第2判定動作では、明瞭性判定部122は、音声区間において、第1周波数帯域B1、第2周波数帯域B2、及び第3周波数帯域B3の各々でのスペクトル包絡の尖度を算出し、算出した尖度と閾値とを比較する。そして、第2判定動作では、明瞭性判定部122は、いずれの周波数帯域B1、B2、B3においても尖度が閾値よりも高ければ第2音声信号S i g 2が明瞭であ

ると判定し、少なくとも1つの周波数帯域で尖度が閾値よりも低ければ、第2音声信号S i g 2が明瞭でないとして判定する。

[0061] 第1周波数帯域B 1は、人の音声における母音の第1フォルマントに対応する周波数帯域である。第2周波数帯域B 2は、人の音声における母音の第2フォルマントに対応する周波数帯域である。第3周波数帯域B 3は、人の音声における母音の第2フォルマント以降のフォルマントに対応する周波数帯域である。

[0062] ここでは、各周波数帯域B 1～B 3は、日本語における母音のフォルマントに対応する周波数帯域である。したがって、例えば英語等の日本語以外の言語に関して第2音声信号S i g 2が明瞭であるか否かを判定する場合、明瞭性判定部1 2 2は、当該言語の母音のフォルマントに対応する1以上の周波数帯域の各々でスペクトル包絡の尖度を算出し、算出した尖度と閾値とを比較すればよい。

[0063] 尖度は、確率変数の確率密度関数又は頻度分布の鋭さを表す指標である。尖度が高ければ高い程、正規分布と比べて鋭いピークと長く太い裾を有する分布（言い換えれば、スペクトル包絡のピーク周りの変化が急峻）となり、尖度が低ければ低い程、正規分布と比べて丸みがかかったピークと短く細い裾を有する分布（言い換えれば、スペクトル包絡の変化がなだらか）となる。

[0064] 第2判定動作では、明瞭性判定部1 2 2は、上述のように各周波数帯域B 1～B 3で尖度が閾値よりも高い場合、人の音声における母音の特徴が顕著に現われている、つまり母音を聴き取ることができる程度に人の音声は明瞭である、と判定する。

[0065] 図4の(a)及び図4の(b)は、いずれも音声区間において人が母音の「o」を発した場合を表している。図4の(a)は、実線L 1及び一点鎖線で示されるように、各周波数帯域B 1～B 3においてスペクトル包絡がなだらかである、つまり各周波数帯域B 1～B 3において尖度が閾値よりも低いことを表している。このため、図4の(a)に示す例では、明瞭性判定部1

22は、第2判定動作において第2音声信号S i g 2が明瞭でないとは判定する。

[0066] 一方、図4の(b)は、実線及び一点鎖線で示されるように、各周波数帯域B1～B3においてスペクトル包絡のピークが現れており、かつ、ピーク周りの変化が急峻となっている、つまり各周波数帯域B1～B3において尖度が閾値よりも高いことを表している。このため、図4の(b)に示す例では、明瞭性判定部122は、第2判定動作において第2音声信号S i g 2が明瞭であると判定する。

[0067] このように、プロセッサ12(明瞭性判定部122)は、少なくとも第2音声信号S i g 2における母音に対応する成分に基づいて、第2条件を満たすか否かを判定する。

[0068] なお、実施の形態では、明瞭性判定部122は、第1判定動作及び第2判定動作の両方を実行しているが、これに限られず、例えば第2判定動作のみで第2条件を満たすか否かを判定してもよい。ただし、空間における残響等の背景雑音の影響を考慮すると、明瞭性判定部122は、第1判定動作及び第2判定動作の両方を実行した方が、第2音声信号S i g 2の明瞭性を精度高く判定することが可能である。

[0069] [3. 動作]

以下、実施の形態に係る音声通話装置4(音声処理システム1)の動作、つまり音声処理方法の一例について図5を用いて説明する。図5は、実施の形態に係る音声処理システム1の動作例を示すフローチャートである。

[0070] まず、プロセッサ12は、第2入力I/F11が第2音声信号S i g 2を取得すると(S101: Yes)、取得した第2音声信号S i g 2をバッファに保持する。以下、特に断りの無い限り、「第2音声信号S i g 2」は、バッファに保持された第2音声信号S i g 2に相当する。その後、第1入力I/F10が第1音声信号S i g 1を取得すると(S102: Yes)、プロセッサ12は、遅延時間の算出及び更新を行う(S103)。具体的には、プロセッサ12は、第1入力I/F10が第1音声信号S i g 1を取得し

た時点と、第2入力I/F11が第2音声信号Sig2を取得した時点との差分を算出することにより、遅延時間を算出し、従前の遅延時間を算出した遅延時間に更新する。なお、プロセッサ12は、算出した遅延時間が従前の遅延時間と同じであれば、更新しない。

[0071] 次に、プロセッサ12は、遅延時間に基づいて、第1音声信号Sig1の開始時点と、第2音声信号Sig2の開始時点とを一致させるように、第1音声信号Sig1と第2音声信号Sig2との時間ずれの補正を行う(S104)。

[0072] 次に、プロセッサ12は、第2音声信号Sig2に基づいて、第2音声信号Sig2の明瞭性を判定するために必要なパラメータを算出する(S105)。以下、ステップS105について、図6を用いて詳細に説明する。

[0073] 図6は、第2音声信号Sig2の明瞭性を判定するために必要なパラメータの算出例を示すフローチャートである。まず、プロセッサ12は、第2音声信号Sig2の音声区間を検出する(S201)。例えば、プロセッサ12は、第2音声信号Sig2の開始時点から所定時間が経過した時点を開始点として、音声区間を検出する。音声区間は、例えばコンマ数秒の区間である。

[0074] 次に、プロセッサ12は、検出した音声区間におけるスペクトルコントラストを算出する(S202)。そして、プロセッサ12は、算出したスペクトルコントラストに基づいて、第2音声信号Sig2における音声帯域のSNRを算出する(S203)。

[0075] また、プロセッサ12は、ステップS202、S203と並行して、又はステップS202、203の前後のいずれかにおいて、検出した音声区間における第2音声信号Sig2の特徴量を算出する(S204)。ここでは、プロセッサ12は、第2音声信号Sig2の特徴量として、第2音声信号Sig2の基本周波数と、第2音声信号Sig2のスペクトル包絡と、を算出する。次に、プロセッサ12は、算出した第2音声信号Sig2の特徴量をメモリ13に保存する(S205)。

[0076] 次に、プロセッサ12は、検出した音声区間における第2音声信号S i g 2のスペクトル包絡の尖度を算出する(S 2 0 6)。具体的には、プロセッサ12は、検出した音声区間において、第1周波数帯域B 1、第2周波数帯域B 2、及び第3周波数帯域B 3の各々でのスペクトル包絡の尖度を算出する。

[0077] 図5に戻り、プロセッサ12は、第2音声信号S i g 2の明瞭性を判定する(S 1 0 6)。具体的には、プロセッサ12は、検出した音声区間において、第2音声信号S i g 2の音声帯域のSNRと閾値とを比較する第1判定動作を実行する。また、プロセッサ12は、検出した音声区間において、各周波数帯域B 1～B 3のスペクトル包絡の尖度と閾値とを比較する第2判定動作を実行する。そして、プロセッサ12は、第1判定動作及び第2判定動作のいずれの判定動作においても明瞭であると判定した場合に、第2音声信号S i g 2が明瞭である、つまり第2条件を満たすと判定する。一方、プロセッサ12は、第1判定動作及び第2判定動作のうちの少なくとも一方の判定動作において明瞭でないと判定した場合、第2音声信号S i g 2が明瞭でない、つまり第2条件を満たさないと判定する。

[0078] 第2音声信号S i g 2が明瞭である、つまり第2条件を満たすと判定された場合(S 1 0 6 : Y e s)、プロセッサ12は、次に第1音声信号S i g 1及び第2音声信号S i g 2に基づいて、発話同一性を判定するために必要なパラメータを算出する(S 1 0 7)。以下、ステップS 1 0 7について、図7を用いて詳細に説明する。

[0079] 図7は、発話同一性を判定するために必要なパラメータの算出例を示すフローチャートである。まず、プロセッサ12は、第1音声信号S i g 1の音声区間を検出する(S 3 0 1)。例えば、プロセッサ12は、第1音声信号S i g 1の開始時点から所定時間が経過した時点を開始点として、音声区間を検出する。検出される音声区間は、第2音声信号S i g 2の音声区間と同じ区間である。

[0080] 次に、プロセッサ12は、メモリ13に保存している第2音声信号S i g

2の特徴量を読み込む（S302）。また、プロセッサ12は、ステップS302と並行して、又はステップS302の前後のいずれかにおいて、検出した音声区間における第1音声信号S i g 1の特徴量を算出する（S303）。ここでは、プロセッサ12は、第1音声信号S i g 1の特徴量として、第1音声信号S i g 1の基本周波数と、第1音声信号S i g 1のスペクトル包絡と、を算出する。

[0081] 図5に戻り、プロセッサ12は、発話同一性を判定する（S108）。具体的には、プロセッサ12は、（i）第1音声信号S i g 1の基本周波数と第2音声信号S i g 2の基本周波数とが同じであり、かつ、（ii）第1音声信号S i g 1において母音が出現するタイミングと第2音声信号S i g 2において母音が出現するタイミングとが同じである場合、話者が同じである、つまり第1条件を満たすと判定する。一方、プロセッサ12は、上記の（i）及び（ii）のうちの少なくとも一方を満たさない場合、話者が同じでない、つまり第1条件を満たさないと判定する。なお、ここでは、プロセッサ12は、2つの比較対象の差分が閾値以下であれば、2つの比較対象が同じであると判定する。

[0082] 話者が同じである、つまり第1条件を満たすと判定された場合（S108：Yes）、第1条件及び第2条件のいずれも満たされているので、プロセッサ12は、第1状況にあると判定し、スピーカ3から出力される第1音声信号S i g 1に基づく音声（つまり、通信音声）の音量を下げる（S109）。また、プロセッサ12は、ノイズキャンセル機能をオフにし（S110）、外音取り込み機能をオンにする（S111）。なお、ステップS109～S111を実行する順番は、この順に限られない。

[0083] 一方、プロセッサ12は、第1条件及び第2条件のうちの少なくとも一方が満たされない場合、つまり第2音声信号S i g 2が明瞭でないと判定された場合（S106：No）、又は話者が同一でないと判定された場合（S108：No）、第2状況にあると判定する。そして、プロセッサ12は、通信音声の音量をデフォルトの音量に設定する（S112）。また、プロセッサ

サ12は、ノイズキャンセル機能をオンにし（S113）、外音取り込み機能をオフにする（S114）。なお、ステップS112～S114を実行する順番は、この順に限られない。

[0084] また、ステップS112～S114は、第2入力I/F11が第2音声信号Sig2を取得していない場合（S101：No）、又は第1入力I/F10が第1音声信号Sig1を取得していない場合（S102：No）においても実行される。

[0085] そして、プロセッサ12は、通話が終了するまでの間（S115：No）、上記の一連の処理を繰り返す。一方、プロセッサ12は、通話が終了すると（S115：Yes）、動作を終了する。

[0086] 図8は、実施の形態に係る音声処理システム1の動作例の概要の説明図である。図8は、同一の拠点に2人のユーザU1、U2が存在する場合における、ユーザU1が装着する音声通話装置4（音声処理システム1）の一連の動作を示している。

[0087] 図8の（a）に示すように、他のユーザU2が音声V2を発すると、マイクロホン2で当該音声を第2音声信号Sig2に変換することで、第2入力I/F11が第2音声信号Sig2を取得する。そして、プロセッサ12は、音声区間を検出し、検出した音声区間において、第2音声信号Sig2の特徴量である第2音声信号Sig2の基本周波数及びMFCC（スペクトル包絡）を算出する。また、プロセッサ12は、算出した第2音声信号Sig2の基本周波数及びMFCCをメモリ13に保存する。また、他のユーザU2が発した音声V2は、第1音声信号Sig1として会議システム100へ送信される。

[0088] 次に、図8の（b）に示すように、プロセッサ12は、検出した音声区間における第2音声信号Sig2の音声帯域のSNRと、スペクトル包絡の尖度とを算出する。そして、プロセッサ12は、算出した第2音声信号Sig2の音声帯域のSNRとスペクトル包絡の尖度とを用いて、第2音声信号Sig2が明瞭であるか否か、つまり第2条件を満たすか否かを判定する。

[0089] また、図8の(c)に示すように、プロセッサ12は、会議システム100から送信される第1音声信号S i g 1を第1入力I / F 10が取得すると、第1音声信号S i g 1の音声区間を検出し、検出した音声区間において、第1音声信号S i g 1の特徴量である第1音声信号S i g 1の基本周波数及びMFCC(スペクトル包絡)を算出する。そして、プロセッサ12は、メモリ13から第2音声信号S i g 2の基本周波数及びMFCCを読み込み、第1音声信号S i g 1の基本周波数及びMFCCとの照合を行うことで、話者が同一であるか否か、つまり第1条件を満たすか否かを判定する。

[0090] 第1条件及び第2条件のいずれもが満たされた場合、つまり第1状況にあると判定された場合、プロセッサ12は、図8の(d)に示すように、スピーカ3から出力される通信音声(第1音声信号S i g 1に基づく音声)の音量を下げるか、又はスピーカ3から通信音声を再生させない。また、プロセッサ12は、ノイズキャンセル機能をオフにし、かつ、外音取り込み機能をオンにする。これにより、ユーザU 1は、他のユーザU 2が発する音声V 2について、会議システム100を介した音声を殆ど聴くことなく、他のユーザU 2からの直接的な音声であって、かつ、明瞭な音声を主として聴くことが可能である。

[0091] 一方、第1条件及び第2条件のうちの少なくとも一方が満たされない場合、つまり第2状況にあると判定された場合、プロセッサ12は、図8の(e)に示すように、スピーカ3から通信音声(第1音声信号S i g 1に基づく音声)を出力させる。また、プロセッサ12は、ノイズキャンセル機能をオンにし、かつ、外音取り込み機能をオフにする。これにより、ユーザU 1は、他のユーザU 2が発する音声V 2について、他のユーザU 2からの直接的な音声であって、かつ、不明瞭な音声を殆ど聴くことなく、会議システム100を介した音声を主として聴くことが可能である。

[0092] [4. 効果等]

以下、実施の形態に係る音声処理システム1の利点について図9を用いて説明する。図9は、実施の形態に係る音声処理システム1の利点の説明図で

ある。図9は、第1拠点A1にいる2人のユーザU1、U2と、第2拠点A2にいるユーザU3と、第3拠点A3にいるユーザU4とが、会議システム100を用いてオンラインで会議を行っていることを表している。

[0093] 図9の(a)は、第1拠点A1にいる2人のユーザU1、U2が互いに比較的近い位置に存在し、他のユーザU2が発した「こんにちは」という音声V2をユーザU1が直接的に聴きやすい状況を表している。このような状況の場合、ユーザU1が装着する音声通話装置4（音声処理システム1）は、第1条件及び第2条件をいずれも満たす、つまり第1状況にあると判定し、第1音声信号Sig1に対応する成分を低減した信号を出力音声信号Sig3に含める。ここでは、音声処理システム1は、第1音声信号Sig1を出力音声信号Sig3に含めない、つまりスピーカ3から第1音声信号Sig1に基づく音声を再生させない。

[0094] このため、ユーザU1は、他のユーザU2が発した「こんにちは」という明瞭な音声V2を直接的に聴くことができる一方、会議システム100を介して送信される他のユーザU2が発した音声V1を聴くことがない。つまり、ユーザU1は、他のユーザU2が発した音声について、他のユーザU2からの直接的な音声と会議システム100（通信回線）を介した音声との両方を聴くことにならないので、他のユーザU2が発した音声を聴き取りやすい。したがって、音声処理システム1では、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい、という利点がある。

[0095] また、実施の形態では、音声処理システム1は、第1状況にあると判定した場合、外音取り込み機能をオンにする、つまり第2音声信号Sig2を出力音声信号Sig3に含めるので、ユーザU1の周囲の音声を取り込むことで、他のユーザU2からの直接的な音声をより聴き取りやすくなる、という利点がある。

[0096] 図9の(b)は、第1拠点A1にいる2人のユーザU1、U2が互いに比較的離れた位置に存在し、他のユーザU2が発した「こんにちは」という音声V2をユーザU1が直接的に聴き取りにくい状況を表している。このよう

な状況の場合、ユーザU1が装着する音声通話装置4（音声処理システム1）は、少なくとも第2条件を満たさない、つまり第2状況にあると判定し、第1音声信号Sig1を出力音声信号Sig3に含め、かつ、第2音声信号Sig2を出力音声信号Sig3に含めない。

[0097] このため、ユーザU1は、会議システム100を介して送信される他のユーザU2が発した「こんにちは」という音声V1を聴くことができる一方、他のユーザU2からの直接的な音声であって、かつ、不明瞭な音声V2を殆ど聴くことがない。つまり、ユーザU1は、他のユーザU2が発した音声について、他のユーザU2からの直接的な音声と会議システム100（通信回線）を介した音声との両方を聴くことにならないので、他のユーザU2が発した音声を聴き取りやすい。したがって、音声処理システム1では、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい、という利点がある。

[0098] また、実施の形態では、音声処理システム1は、第2状況にあると判定した場合、ノイズキャンセル機能をオンにする、つまり第2音声信号Sig2とは逆位相の音声信号を出力音声信号Sig3に含める。このため、他のユーザU2からの直接的な音声を含むユーザU1の周囲の雑音を除去することで、ユーザU1が他のユーザU2からの会議システム100（通信回線）を介した音声をより聴き取りやすくなる、という利点がある。

[0099] ところで、複数のユーザ（ここでは、ユーザU1～U4）が同時に音声を発した場合、音声処理システム1は、第1条件及び第2条件のいずれも満たさない、つまり第2状況にあると判定する。この場合、スピーカ3からは、会議システム100を介した他のユーザU2～U4からの音声が出力される。このような場合、各ユーザU1～U4による会話が一時的に止まるため、音声処理システム1の利点が妨げられることはない。音声処理システム1を採用するユーザU1は、少なくとも各ユーザU1～U4が交互に音声を発する状況において、上記の利点を享受できればよい。

[0100] [5. その他の実施の形態]

以上、実施の形態について説明したが、本開示は、上記実施の形態に限定されるものではない。

[0101] 上記実施の形態において、プロセッサ12は、外音取り込み切替部128及びANC制御部129を備えていなくてもよい。この場合、音声処理システム1は、図5に示すフローチャートにおけるステップS110、S111、S113、S114を実行しなくてもよい。より具体的には、音声通話装置4が開放型の音声通話装置である場合、プロセッサ12は、外音取り込み切替部128及びANC制御部129を備えていてもよいが、備えていなくても差し支えない。なお、音声通話装置4が開放型の音声通話装置である場合も、プロセッサ12は、ANC制御部129を備えていてもよい。この場合、音声処理システム1は、図5に示すフローチャートにおけるステップS111、S114を実行しなくてもよい。また、音声通話装置4が密閉型の音声通話装置である場合、プロセッサ12は、外音取り込み切替部128を備えている方がよいが、例えば外音が多少漏れ聴こえるような場合であれば、外音取り込み切替部128を備えていなくてもよい。また、音声通話装置4が密閉型の音声通話装置である場合、プロセッサ12は、ANC制御部129を備えている方がよいが、ユーザの耳穴が塞がれることである程度外音が低減される場合であれば、ANC制御部129を備えていなくてもよい。

[0102] また、上記実施の形態において、例えば室内を移動しながら実施するVR会議、又は室内の環境雑音が変化する状態等の明瞭性の特徴量に変化しやすい状況では、出力音声制御部127、外音取り込み切替部128、及びANC制御部129の各々は常に制御されていたが、これらは一定時間制御されなくてもよい。より具体的には、音声処理システム1は、図5に示すフローチャートにおけるステップS112～S114、及びS109～S111を一定時間（例えば、数ミリ秒）実行しなくてもよい。この場合、出力音声制御部127、外音取り込み切替部128、及びANC制御部129の各々が一定時間ごとに制御され、高頻度で制御されることを防止してもよい。なお、出力音声制御部127を制御する時間、外音取り込み切替部128を制御

する時間、及びANC制御部129を制御する時間の各々は、同じでなくてもよい。

[0103] また、上記実施の形態において、音声処理システム1は、単一の装置（音声通話装置4）によって実現されたが、複数の装置として実現されてもよい。音声処理システム1が複数の装置によって実現される場合、音声処理システム1が備える機能的な構成要素は、複数の装置にどのように振り分けられてもよい。また、例えば、音声処理システム1は、第1入力I/F10と、第2入力I/F11と、プロセッサ12と、を備えたサーバにより実現されてもよい。この場合、音声処理システム1は、マイクロホン2及びスピーカ3を備える機器との間で通信することにより、マイクロホン2から第2音声信号Sig2を取得したり、スピーカ3から出力音声信号Sig3に基づく音声を出力させたりすることが可能である。

[0104] また、上記実施の形態における装置間の通信方法については特に限定されるものではない。上記実施の形態において2つの装置が通信を行う場合、2つの装置間には図示されない中継装置が介在してもよい。

[0105] また、上記実施の形態で説明された処理の順序は、一例である。複数の処理の順序は変更されてもよいし、複数の処理は並行して実行されてもよい。また、特定の処理部が実行する処理を別の処理部が実行してもよい。また、上記実施の形態で説明されたデジタル信号処理の一部がアナログ信号処理によって実現されてもよい。

[0106] また、上記実施の形態において、各構成要素は、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、CPU又はプロセッサ等のプログラム実行部が、ハードディスク又は半導体メモリ等の記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。

[0107] また、各構成要素は、ハードウェアによって実現されてもよい。例えば、各構成要素は、回路（又は集積回路）でもよい。これらの回路は、全体として1つの回路を構成してもよいし、それぞれ別々の回路でもよい。また、こ

これらの回路は、それぞれ、汎用的な回路でもよいし、専用の回路でもよい。

[0108] また、本開示の全般的又は具体的な態様は、システム、装置、方法、集積回路、コンピュータプログラム又はコンピュータ読み取り可能なCD-ROM等の記録媒体で実現されてもよい。また、システム、装置、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。例えば、本開示は、コンピュータが実行する音声処理方法として実行されてもよいし、このような音声処理方法をコンピュータに実行させるためのプログラムとして実現されてもよい。また、本開示は、このようなプログラムが記録されたコンピュータ読み取り可能な非一時的な記録媒体として実現されてもよい。なお、ここでのプログラムには、汎用の情報端末を上記実施の形態の音声処理システムとして機能させるためのアプリケーションプログラムが含まれる。

[0109] その他、各実施の形態に対して当業者が思いつく各種変形を施して得られる形態、又は、本開示の趣旨を逸脱しない範囲で各実施の形態における構成要素及び機能を任意に組み合わせることで実現される形態も本開示に含まれる。

[0110] (まとめ)

以上述べたように、第1の態様に係る音声処理システム1は、第1入力I/F10と、第2入力I/F11と、プロセッサ12と、を備える。プロセッサ12は、信号処理回路の一例である。第1入力I/F10は、通信回線を介した第1音声信号Sig1を取得する。第2入力I/F11は、マイクロホン2で収音された音声に基づく第2音声信号Sig2を取得する。プロセッサ12は、第1音声信号Sig1及び第2音声信号Sig2に基づく出力音声信号Sig3をスピーカ3に出力する。プロセッサ12は、第1音声信号Sig1及び第2音声信号Sig2がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、第2音声信号Sig2が明瞭であるという第2条件をいずれも満たす場合、第1音声信号Sig1に対応する成分を低減した信号を出力音声信号Sig3に含める。

[0111] これによれば、同じ拠点に複数のユーザが存在する環境において、ユーザは、他のユーザからの直接的な音声と通信回線を介した音声とのうち直接的な音声は明瞭である場合、直接的な音声を主に聴くことになるので、他のユーザが発した音声を聴き取りやすい。つまり、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい、という利点がある。

[0112] また、第2の態様に係る音声処理システム1では、第1の態様において、プロセッサ12は、第1条件及び第2条件のうちの少なくとも一方を満たさない場合、第1音声信号S i g 1を出力音声信号S i g 3に含め、かつ、第2音声信号S i g 2を出力音声信号S i g 3に含めない。

[0113] これによれば、同じ拠点に複数のユーザが存在する環境において、ユーザは、他のユーザからの直接的な音声と通信回線を介した音声とのうち直接的な音声よりも聴き取りやすい通信回線を介した音声を主に聴くことになるので、他のユーザが発した音声を聴き取りやすい。つまり、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい、という利点がある。

[0114] また、第3の態様に係る音声処理システム1では、第1又は第2の態様において、プロセッサ12は、第1音声信号S i g 1における母音に対応する成分と、第2音声信号S i g 2における母音に対応する成分との相関に基づいて、第1条件を満たすか否かを判定する。

[0115] これによれば、第1音声信号S i g 1の波形と第2音声信号S i g 2の波形との類似度に基づく方法と比較して、同一の人物が発する音声に基づく音声信号を含むか否かを判定しやすい、という利点がある。

[0116] また、第4の態様に係る音声処理システム1では、第1～第3のいずれか1つの態様において、プロセッサ12は、第2音声信号S i g 2における母音に対応する成分に基づいて、第2条件を満たすか否かを判定する。

[0117] これによれば、人の音声の聴き取りやすさの指標となる母音に対応する成分に基づいて判定するので、第2音声信号S i g 2が明瞭であるか否かを判

定しやすい、という利点がある。

[0118] また、第5の態様に係る音声処理システム1では、第1～第4のいずれか1つの態様において、プロセッサ12は、第1条件及び第2条件をいずれも満たす場合、第2音声信号S i g 2を出力音声信号S i g 3に含める。

[0119] これによれば、ユーザの周囲の音声を取り込むことで、ユーザが他のユーザからの直接的な音声をより聴き取りやすくなる、という利点がある。

[0120] また、第6の態様に係る音声処理システム1では、第2の態様において、プロセッサ12は、第1条件及び第2条件のうちの少なくとも一方を満たさない場合、更に第2音声信号S i g 2とは逆位相の音声信号を出力音声信号S i g 3に含める。

[0121] これによれば、他のユーザからの直接的な音声を含むユーザの周囲の雑音を除去することで、ユーザが他のユーザからの通信回線を介した音声をより聴き取りやすくなる、という利点がある。

[0122] また、第7の態様に係る音声処理方法では、通信回線を介した第1音声信号S i g 1を取得し（S 1 0 2 : Y e s）、マイクロホン2で収音された音声に基づく第2音声信号S i g 2を取得し（S 1 0 1 : Y e s）、第1音声信号S i g 1及び第2音声信号S i g 2がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、第2音声信号S i g 2が明瞭であるという第2条件をいずれも満たす場合（S 1 0 6 : Y e s、S 1 0 8 : Y e s）、第1音声信号S i g 1に対応する成分を低減した信号を出力音声信号S i g 3に含めてスピーカ3に出力する（S 1 0 9）。

[0123] これによれば、同じ拠点に複数のユーザが存在する環境において、ユーザは、他のユーザからの直接的な音声と通信回線を介した音声とのうち直接的な音声は明瞭である場合、直接的な音声を主に聴くことになるので、他のユーザが発した音声を聴き取りやすい。つまり、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい、という利点がある。

[0124] また、第8の態様に係るプログラムは、1以上のプロセッサに、第7の態

様に係る音声処理方法を実行させる。

- [0125] これによれば、同じ拠点に複数のユーザが存在する環境において、ユーザは、他のユーザからの直接的な音声と通信回線を介した音声とのうち直接的な音声は明瞭である場合、直接的な音声を主に聴くことになるので、他のユーザが発した音声を聴き取りやすい。つまり、同じ拠点に複数のユーザが存在する環境においても、会話の快適性が損なわれにくい、という利点がある。

産業上の利用可能性

- [0126] 本開示の音声処理システム等は、スピーカの発する音声を処理するシステム等に適用可能である。

符号の説明

- [0127]
- 1 音声処理システム
 - 10 第1入力I/F
 - 11 第2入力I/F
 - 12 プロセッサ
 - 121 明瞭性算出部
 - 122 明瞭性判定部
 - 123 第1特徴量算出部
 - 124 第2特徴量算出部
 - 125 発話同一性判定部
 - 126 出力音声判定部
 - 127 出力音声制御部
 - 128 外音取り込み切替部
 - 129 ANC制御部
 - 13 メモリ
 - 2 マイクロホン
 - 3 スピーカ
 - 100 会議システム

S i g 1 第1音声信号

S i g 2 第2音声信号

S i g 3 出力音声信号

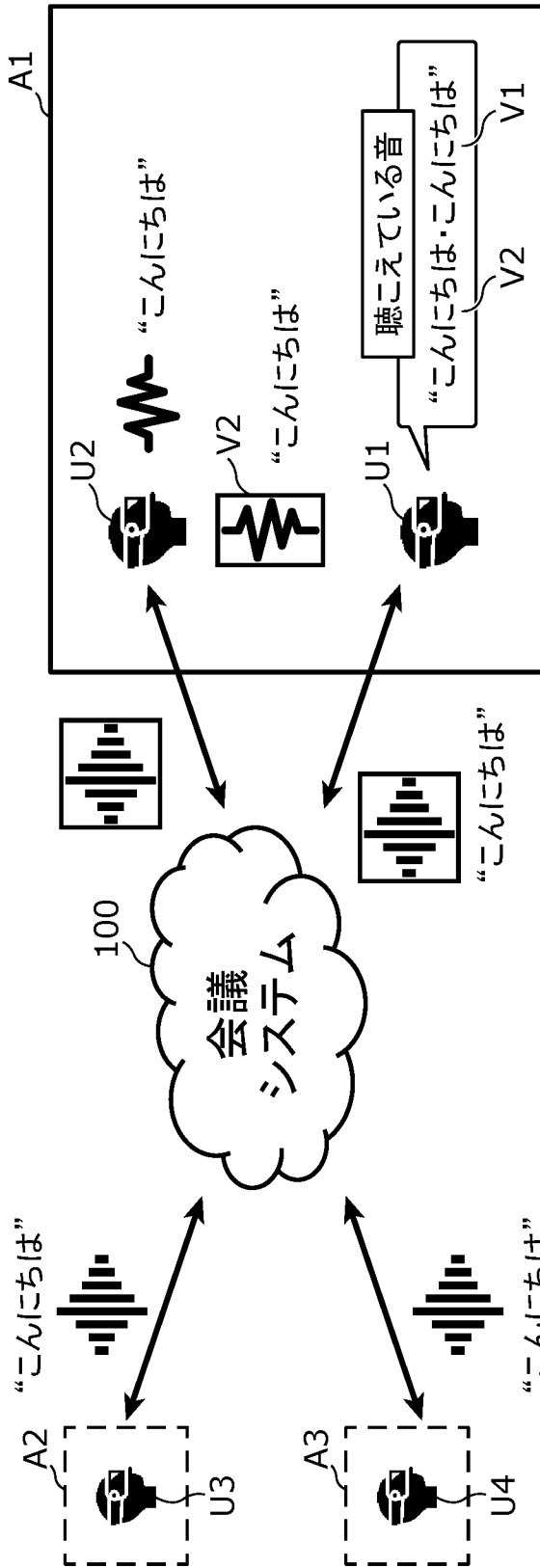
V 1、V 2 音声

請求の範囲

- [請求項1] 通信回線を介した第1音声信号を取得する第1入力インタフェースと、
マイクロホンで収音された音声に基づく第2音声信号を取得する第2入力インタフェースと、
前記第1音声信号及び前記第2音声信号に基づく出力音声信号をスピーカに出力する信号処理回路と、を備え、
前記信号処理回路は、
前記第1音声信号及び前記第2音声信号がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、前記第2音声信号が明瞭であるという第2条件をいずれも満たす場合、前記第1音声信号に対応する成分を低減した信号を前記出力音声信号に含める、
音声処理システム。
- [請求項2] 前記信号処理回路は、
前記第1条件及び前記第2条件のうちの少なくとも一方を満たさない場合、前記第1音声信号を前記出力音声信号に含め、かつ、前記第2音声信号を前記出力音声信号に含めない、
請求項1に記載の音声処理システム。
- [請求項3] 前記信号処理回路は、
前記第1音声信号における母音に対応する成分と、前記第2音声信号における母音に対応する成分との相関に基づいて、前記第1条件を満たすか否かを判定する、
請求項1又は2に記載の音声処理システム。
- [請求項4] 前記信号処理回路は、
前記第2音声信号における母音に対応する成分に基づいて、前記第2条件を満たすか否かを判定する、
請求項1又は2に記載の音声処理システム。

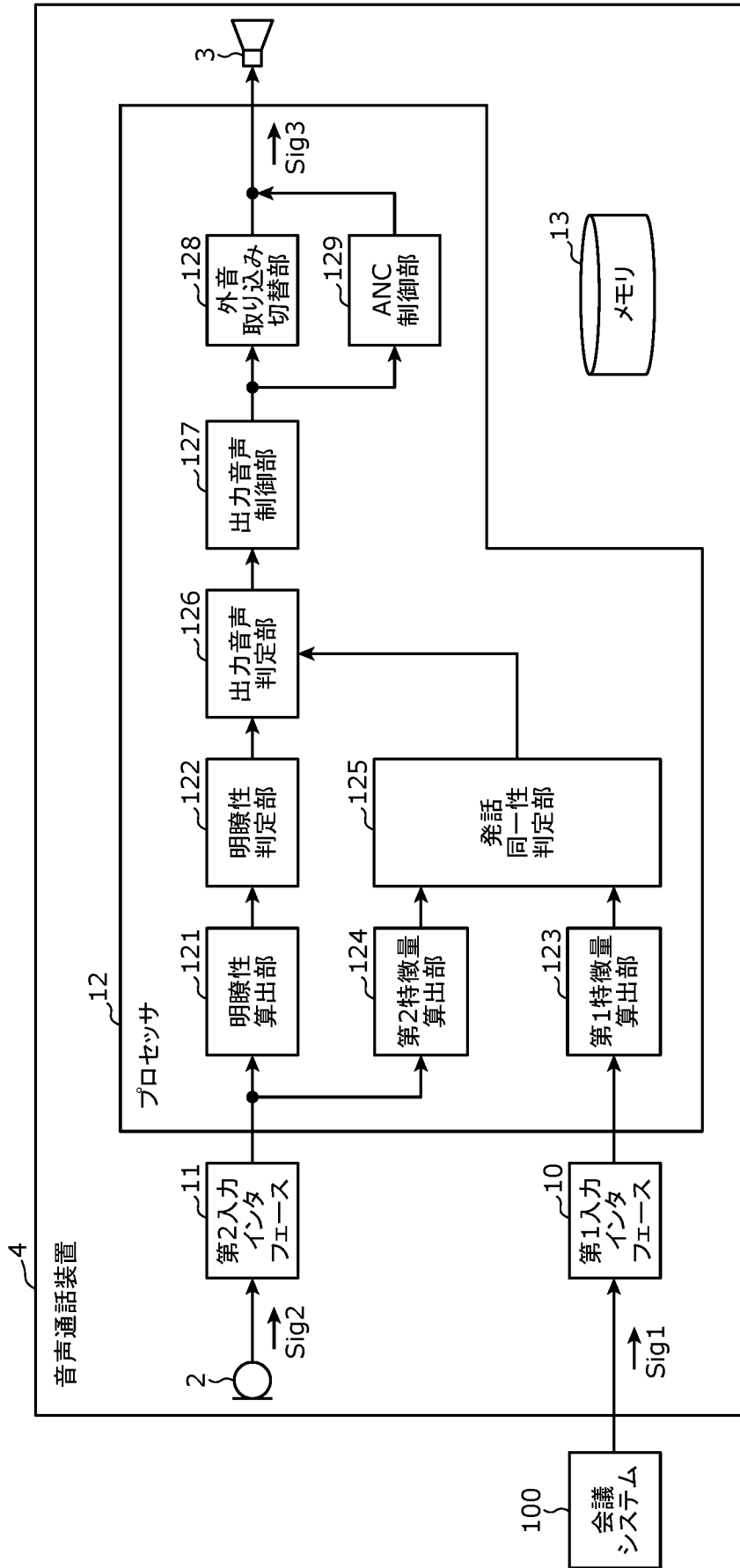
- [請求項5] 前記信号処理回路は、
前記第1条件及び前記第2条件をいずれも満たす場合、前記第2音声信号を前記出力音声信号に含める、
請求項1又は2に記載の音声処理システム。
- [請求項6] 前記信号処理回路は、
前記第1条件及び前記第2条件のうちの少なくとも一方を満たさない場合、更に前記第2音声信号とは逆位相の音声信号を前記出力音声信号に含める、
請求項2に記載の音声処理システム。
- [請求項7] 通信回線を介した第1音声信号を取得し、
マイクロホンで収音された音声に基づく第2音声信号を取得し、
前記第1音声信号及び前記第2音声信号がいずれも同一の人物の発する音声に基づく音声信号を含むという第1条件、及び、前記第2音声信号が明瞭であるという第2条件をいずれも満たす場合、前記第1音声信号に対応する成分を低減した信号を出力音声信号に含めてスピーカに出力する、
音声処理方法。
- [請求項8] 1以上のプロセッサに、
請求項7に記載の音声処理方法を実行させる、
プログラム。

[図1]

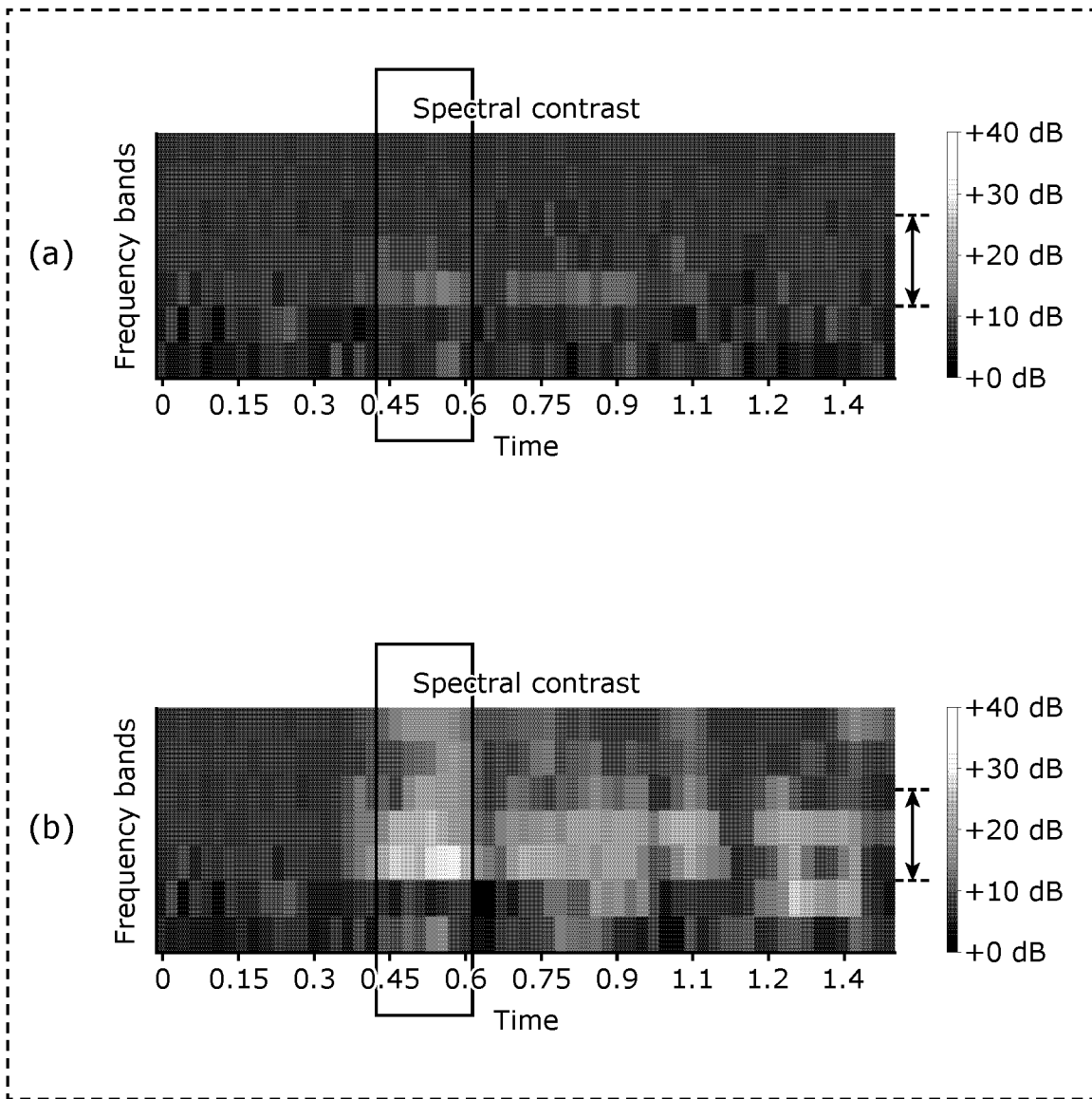


[図2]

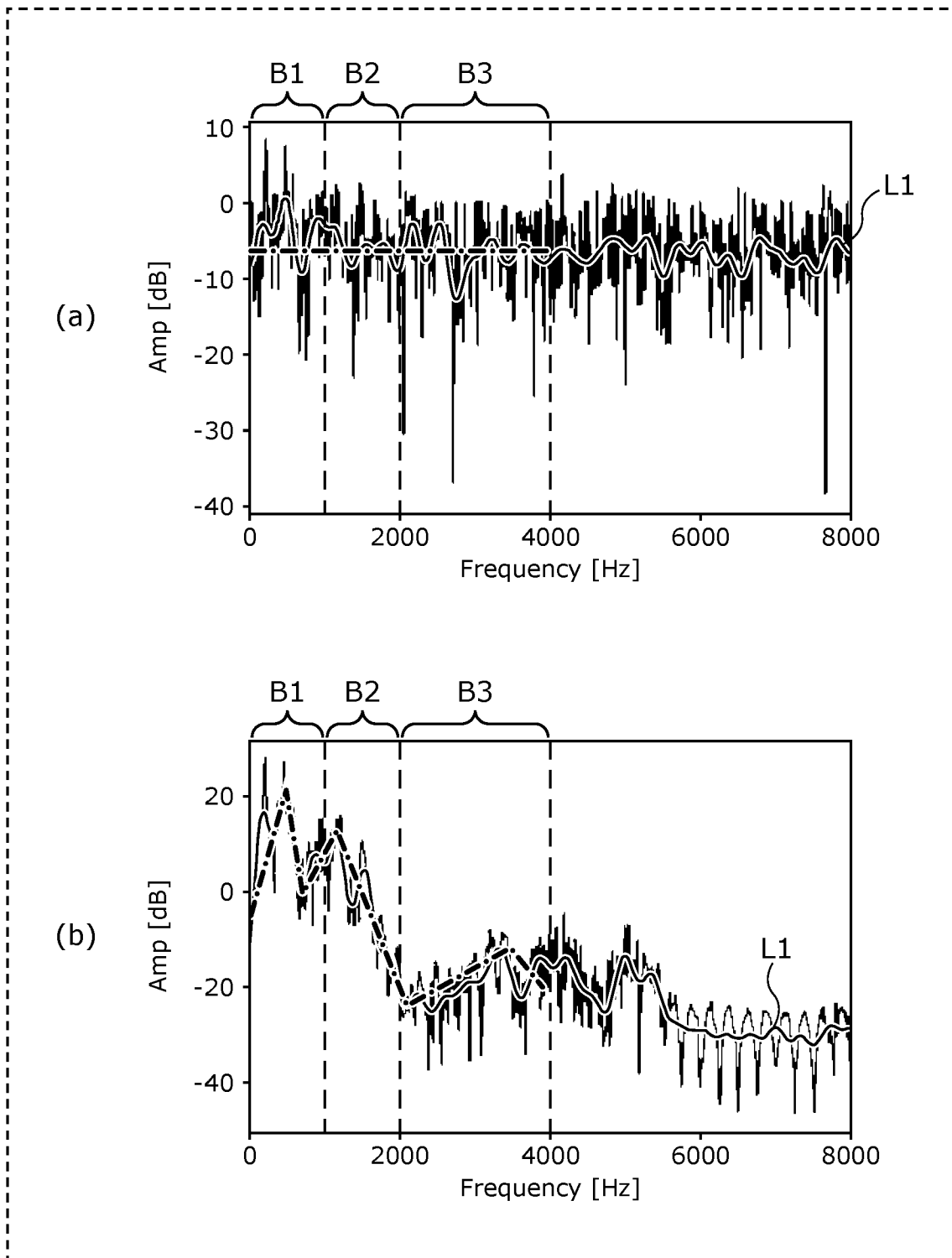
1



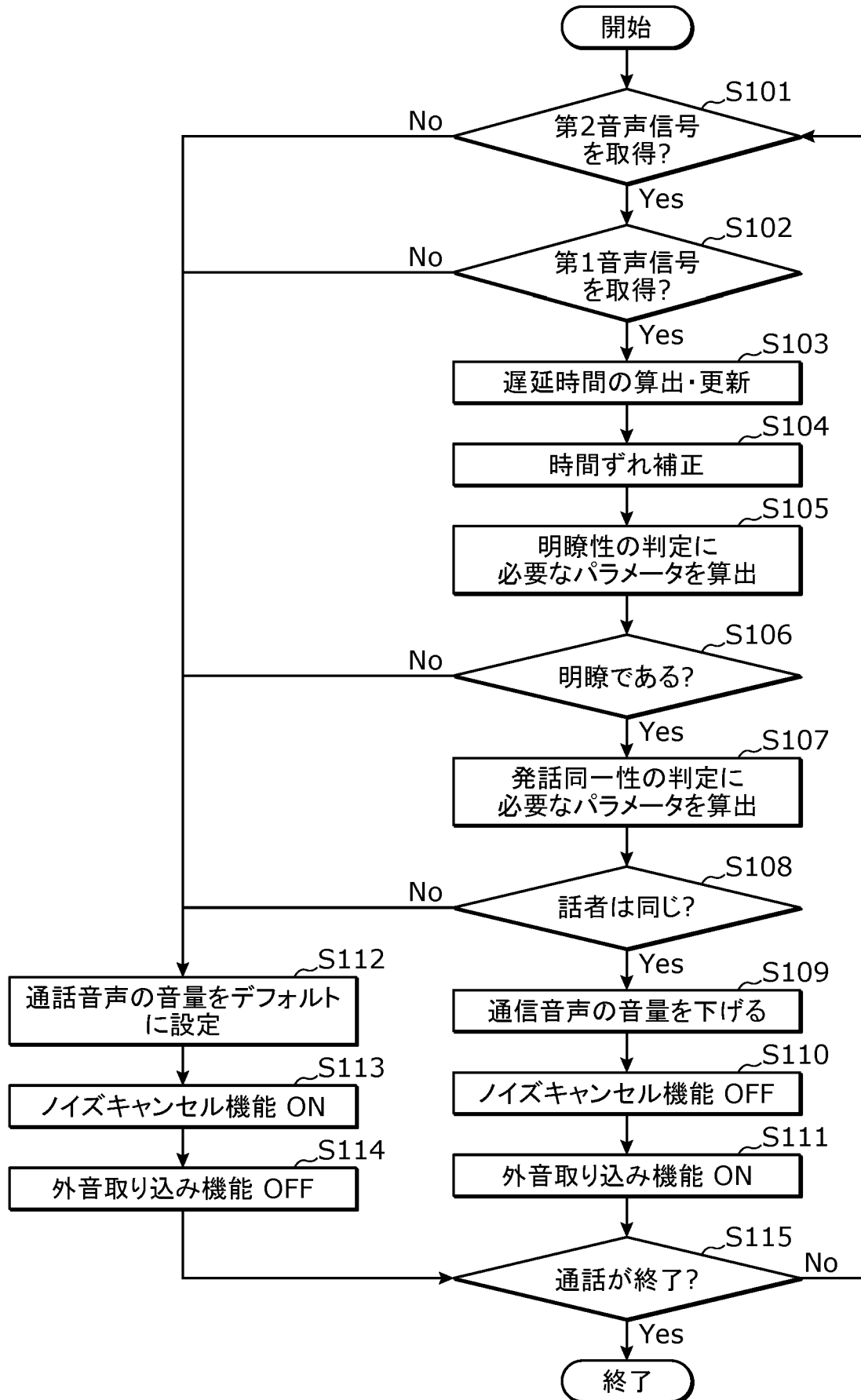
[図3]



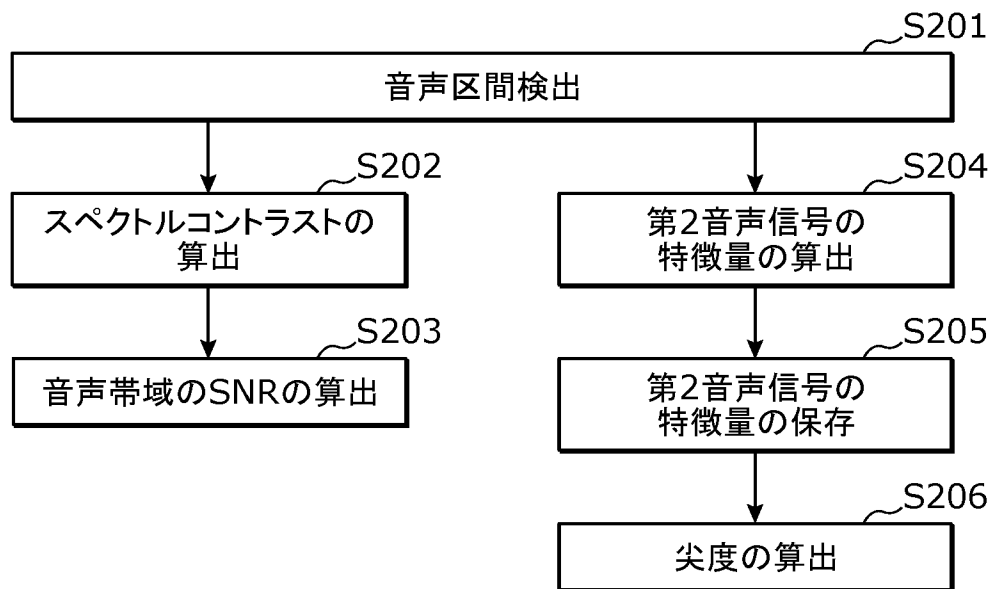
[図4]



[図5]



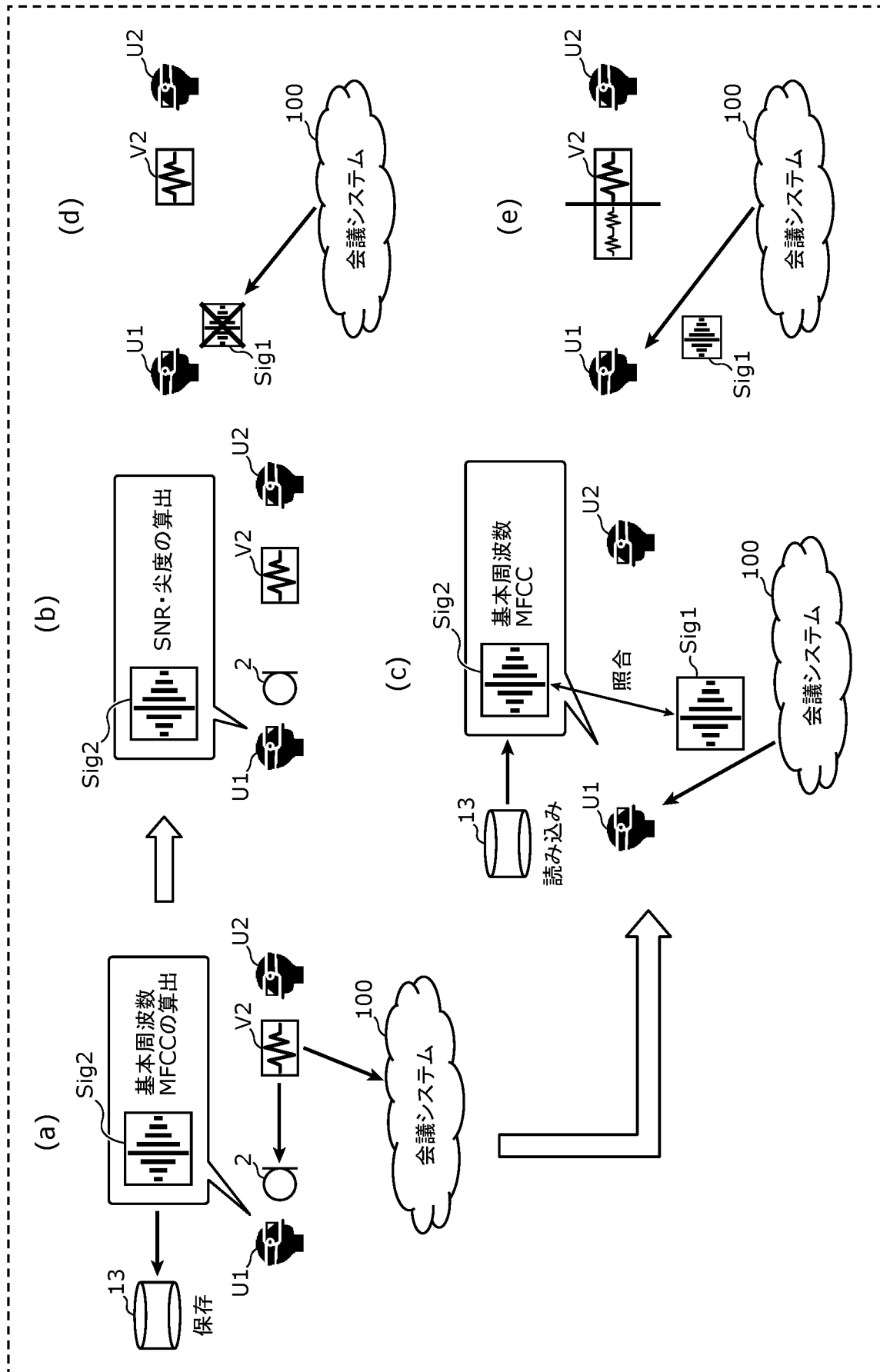
[図6]



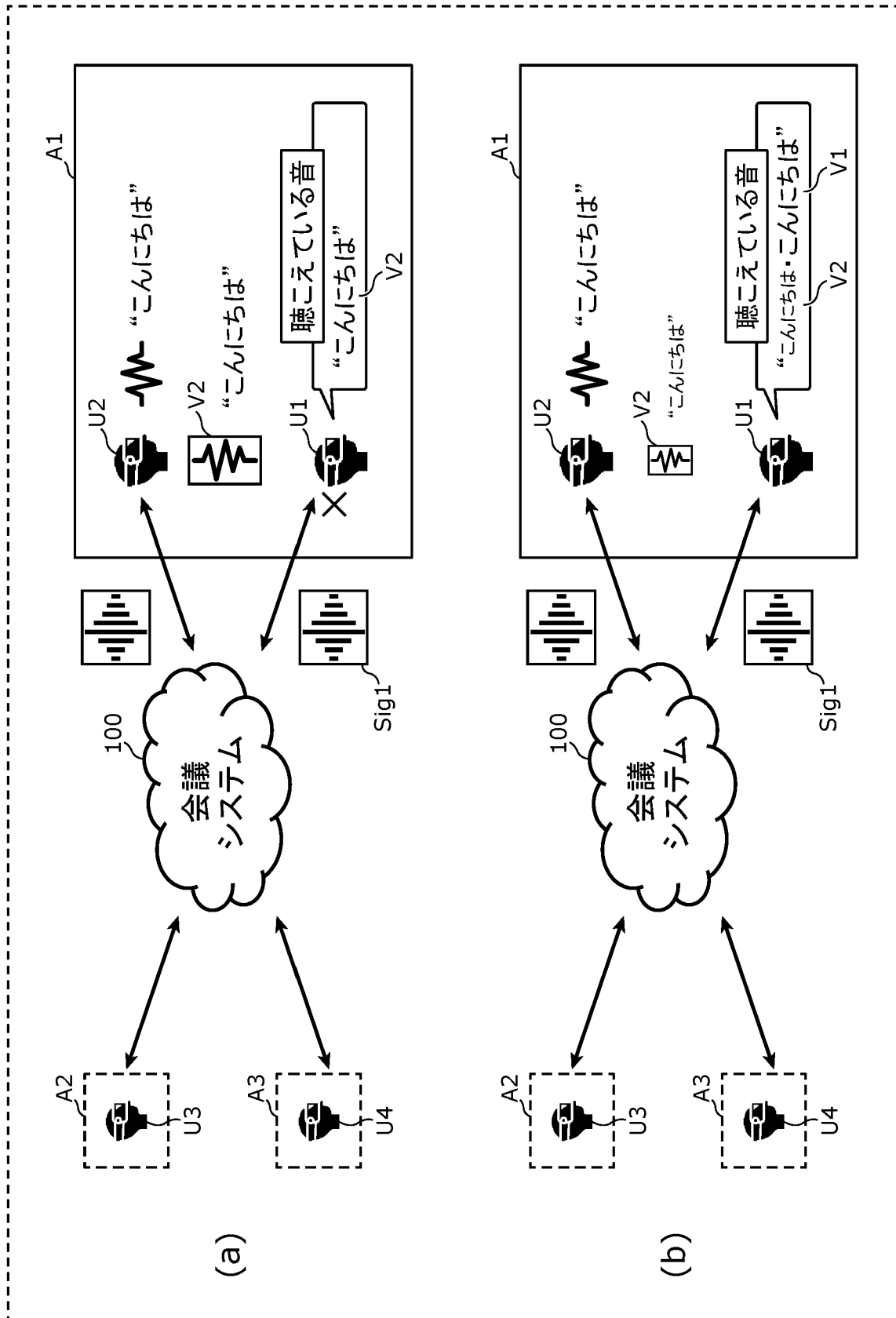
[図7]



[図8]



[図9]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/042673

A. CLASSIFICATION OF SUBJECT MATTER		
<i>H04R 3/02</i> (2006.01)i FI: H04R3/02		
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) G10L13/00-25/93; H04R1/00-3/14		
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.
Y	JP 2017-028351 A (FUJITSU LTD) 02 February 2017 (2017-02-02) paragraphs [0011]-[0025], fig. 5	1-3, 6-8
A		4-5
Y	JP 2019-140517 A (FUJI XEROX CO., LTD.) 22 August 2019 (2019-08-22) paragraphs [0018], [0034]-[0041], [0049]	1-3, 6-8
Y	JP 2021-140065 A (BANDAI NAMCO ENTERTAINMENT INC) 16 September 2021 (2021-09-16) paragraph [0094]	3
Y	JP 2006-023758 A (YAMAHA CORP) 26 January 2006 (2006-01-26) paragraph [0040]	3
A	JP 2022-142038 A (COTOPA DESIGN INC) 30 September 2022 (2022-09-30) entire text, all drawings	1-8
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "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 02 February 2024		Date of mailing of the international search report 13 February 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/JP2023/042673

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2022-138245 A (SHARP KK) 26 September 2022 (2022-09-26) entire text, all drawings	1-8
A	JP 2017-045180 A (FUJI XEROX CO., LTD.) 02 March 2017 (2017-03-02) entire text, all drawings	1-8
A	WO 2022/118671 A1 (SONY GROUP CORPORATION) 09 June 2022 (2022-06-09) entire text, all drawings	1-8

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2023/042673

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2017-028351 A	02 February 2017	(Family: none)	
JP 2019-140517 A	22 August 2019	(Family: none)	
JP 2021-140065 A	16 September 2021	(Family: none)	
JP 2006-023758 A	26 January 2006	(Family: none)	
JP 2022-142038 A	30 September 2022	(Family: none)	
JP 2022-138245 A	26 September 2022	US 2022/0295184 A1 entire text, all drawings	
JP 2017-045180 A	02 March 2017	(Family: none)	
WO 2022/118671 A1	09 June 2022	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） H04R 3/02(2006.01)i FI: H04R3/02		
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）） G10L13/00-25/93; H04R1/00-3/14 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
Y A	JP 2017-028351 A（富士通株式会社）02.02.2017（2017-02-02） [0011]-[0025], 図5	1-3, 6-8 4-5
Y	JP 2019-140517 A（富士ゼロックス株式会社）22.08.2019（2019-08-22） [0018], [0034]-[0041], [0049]	1-3, 6-8
Y	JP 2021-140065 A（株式会社バンダイナムコエンターテインメント）16.09.2021 （2021-09-16） [0094]	3
Y	JP 2006-023758 A（ヤマハ株式会社）26.01.2006（2006-01-26） [0040]	3
A	JP 2022-142038 A（株式会社コトバデザイン）30.09.2022（2022-09-30） 全文, 全図	1-8
A	JP 2022-138245 A（シャープ株式会社）26.09.2022（2022-09-26） 全文, 全図	1-8
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的な技術水準を示すもの “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 02.02.2024	国際調査報告の発送日 13.02.2024	
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） 中村 天真 5Z 1786 電話番号 03-3581-1101 内線 3591	

C. 関連すると認められる文献		
引用文献の カテゴリ*	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
A	JP 2017-045180 A (富士ゼロックス株式会社) 02.03.2017 (2017 - 03 - 02) 全文, 全図	1-8
A	WO 2022/118671 A1 (ソニーグループ株式会社) 09.06.2022 (2022 - 06 - 09) 全文, 全図	1-8

国際調査報告
 パテントファミリーに関する情報

国際出願番号
 PCT/JP2023/042673

引用文献	公表日	パテントファミリー文献	公表日
JP 2017-028351 A	02.02.2017	(ファミリーなし)	
JP 2019-140517 A	22.08.2019	(ファミリーなし)	
JP 2021-140065 A	16.09.2021	(ファミリーなし)	
JP 2006-023758 A	26.01.2006	(ファミリーなし)	
JP 2022-142038 A	30.09.2022	(ファミリーなし)	
JP 2022-138245 A	26.09.2022	US 2022/0295184 A1 全文, 全図	
JP 2017-045180 A	02.03.2017	(ファミリーなし)	
WO 2022/118671 A1	09.06.2022	(ファミリーなし)	