

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年3月31日(31.03.2022)



(10) 国際公開番号

WO 2022/065441 A1

(51) 国際特許分類:
H04R 3/00 (2006.01) *G10L 21/0364* (2013.01)
H04R 25/00 (2006.01)

(21) 国際出願番号: PCT/JP2021/035137

(22) 国際出願日: 2021年9月24日(24.09.2021)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2020-160162 2020年9月24日(24.09.2020) JP

(71) 出願人: 株式会社 J V C ケンウッド (JVCKENWOOD CORPORATION) [JP/JP];
〒2210022 神奈川県横浜市神奈川区守屋町
3丁目12番地 Kanagawa (JP).

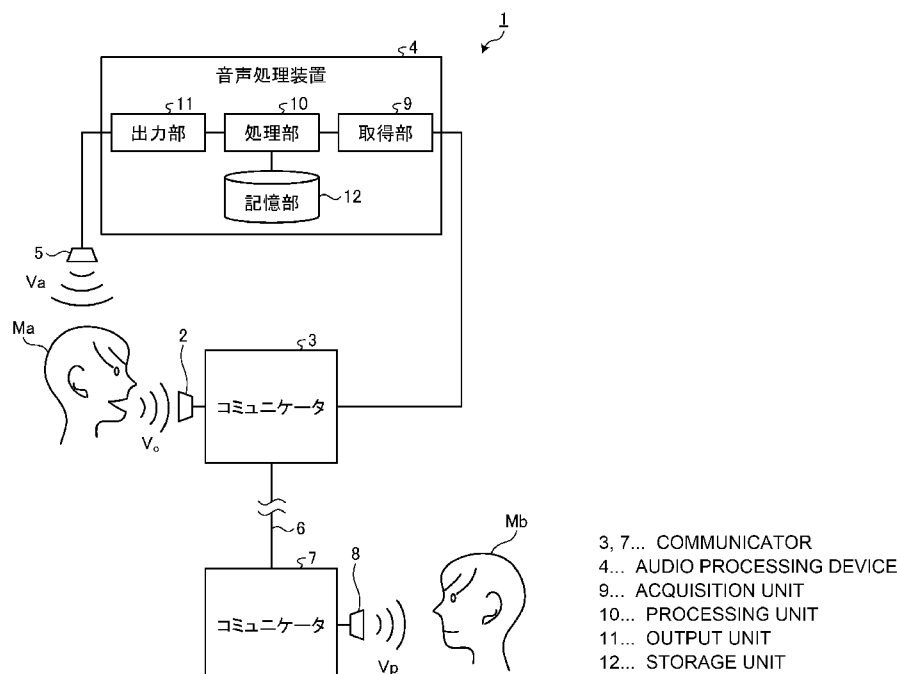
(72) 発明者: 鶴 秀生(TSURU, Hideo); 〒2210022 神奈川県横浜市神奈川区守屋町3丁目12番地 株式会社 J V C ケンウッド 知的財産部内 Kanagawa (JP). 高田 規(TAKADA, Motomu); 〒2210022 神奈川県横浜市神奈川区守屋町3丁目12番地 株式会社 J V C ケンウッド 知的財産部内 Kanagawa (JP). 辻井 秀弥(TSUJII, Hideya); 〒2210022 神奈川県横浜市神奈川区守屋町3丁目12番地 株式会社 J V C ケンウッド 知的財産部内 Kanagawa (JP).

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

(54) Title: PRONUNCIATION FEEDBACK DEVICE, PRONUNCIATION FEEDBACK METHOD, AND COMPUTER PROGRAM

(54) 発明の名称: 発音フィードバック装置、発音フィードバック方法、及びコンピュータプログラム

[図1]



(57) Abstract: Provided is a pronunciation feedback device that provides correct audio feedback to a person speaking. The pronunciation feedback device comprises: an acquisition unit that obtains original audio data indicating original audio spoken by the speaker; a processing unit that adjusts the acoustic features of the original audio data and generates playback audio data that indicates playback audio; and an output unit that outputs playback audio data.

[続葉有]

WO 2022/065441 A1

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

(57) 要約: 発話者に適正な音声をフィードバックする発音フィードバック装置を提供する。発音フィードバック装置は、発話者が発した原音声を示す原音声データを取得する取得部と、原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成する処理部と、再生音声データを出力する出力部と、を備える。

明 細 書

発明の名称：

発音フィードバック装置、発音フィードバック方法、及びコンピュータプログラム

技術分野

[0001] 本発明は、発音フィードバック装置、発音フィードバック方法、及びコンピュータプログラムに関する。

背景技術

[0002] 特許文献1には、スピーカ付きマイクが開示されている。発話者がマイクに発した音声は、スピーカから発話者にフィードバックされる。

先行技術文献

特許文献

[0003] 特許文献1：特開2006-197526号公報

発明の概要

発明が解決しようとする課題

[0004] 雑音環境下においては、発話者は普段よりも大きい声又は高い声で話してしまう可能性がある。また、発話者の吃音又は活舌の悪さに起因して、発話者の音声を認識し難くなる可能性がある。このような不具合を抑制するためには、発話者に適正な音声をフィードバックすることが有効である。

[0005] 本発明は、発話者に適正な音声をフィードバックすることを目的とする。

課題を解決するための手段

[0006] 本発明の一態様に係る発音フィードバック装置は、発話者が発した原音声を示す原音声データを取得する取得部と、前記原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成する処理部と、前記再生音声データを出力する出力部と、を備える。

[0007] 本発明の一態様に係る発音フィードバック方法は、発話者が発した原音声

を示す原音声データを取得するステップと、前記原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成するステップと、前記再生音声データを出力するステップと、を含む。

[0008] 本発明の一態様に係るコンピュータプログラムは、発話者が発した原音声を示す原音声データを取得するステップと、前記原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成するステップと、前記再生音声データを出力するステップと、を含む発音フィードバック方法を、コンピュータに実行させる。

発明の効果

[0009] 本発明によれば、発話者に適正な音声をフィードバックすることができる。

図面の簡単な説明

[0010] [図1]図1は、第1実施形態に係る発音フィードバック装置を示す模式図である。

[図2]図2は、第1実施形態に係る音声処理装置を示す機能ブロック図である。

[図3]図3は、第1実施形態に係るフォルマントシフトを説明するための図である。

[図4]図4は、第1実施形態に係るフォルマントシフトを説明するための図である。

[図5]図5は、第1実施形態に係る発音フィードバック方法を示すフローチャートである。

[図6]図6は、第1実施形態に係る発音フィードバック装置の第1適用例を説明するための図である。

[図7]図7は、第1実施形態に係る発音フィードバック装置の第2適用例を説明するための図である。

[図8]図8は、第1実施形態の変形例に係る発音フィードバック装置を示す模式図である。

[図9]図 9 は、第 2 実施形態に係る発音フィードバック装置を示す模式図である。

[図10]図 10 は、第 2 実施形態に係るイヤホンの周波数特性を説明するための図である。

発明を実施するための形態

[0011] 以下に、本発明の実施形態を図面に基づいて詳細に説明する。なお、以下に説明する実施形態により本発明が限定されるものではない。

[0012] [第 1 実施形態]

(発音フィードバック装置)

図 1 は、本実施形態に係る発音フィードバック装置 1 を示す模式図である。図 1 に示すように、発音フィードバック装置 1 は、マイクロホン 2 と、コミュニケーター 3 と、音声処理装置 4 と、スピーカ 5 とを備える。

[0013] 発話者 M a は、音声を発する。本実施形態において、発話者 M a が発した音声を適宜、原音声 V o、と称する。発話者 M a が発した原音声 V o は、気導音として、マイクロホン 2 に入力される。気導音とは、空中を伝播する音をいう。

[0014] マイクロホン 2 は、発話者 M a が発した原音声 V o を原音声データに変換する。マイクロホン 2 は、コミュニケーター 3 に接続される。マイクロホン 2 からの原音声データは、伝送装置 6 を介してコミュニケーター 3 から別のコミュニケーター 7 に伝送される。コミュニケーター 7 にスピーカ 8 が接続される。スピーカ 8 は、原音声データを再生音声 V p に変換する。再生音声 V p は、スピーカ 8 から出力される。原音声 V o の周波数特性と再生音声 V p の周波数特性とは類似する。視聴者 M b は、スピーカ 8 から出力された再生音声 V p を聞くことができる。

[0015] コミュニケーター 3 は、マイクロホン 2 からの原音声データを音声処理装置 4 に送信する。

[0016] 音声処理装置 4 は、取得部 9 と、処理部 10 と、出力部 11 と、記憶部 12 とを有する。

- [0017] 取得部 9 は、発話者 M a が発した原音声 V o を示す原音声データを取得する。原音声 V o は、気導音である。取得部 9 は、コミュニケータ 3 を介してマイクロホン 2 から原音声データを取得する。
- [0018] 処理部 10 は、取得部 9 により取得された原音声データの音響特徴量を調整して、再生音声 V a を示す再生音声データを生成する。
- [0019] 出力部 11 は、処理部 10 により生成された再生音声データをスピーカ 5 へ出力する。スピーカ 5 は、再生音声データを再生音声 V a に変換する。再生音声 V a は、スピーカ 5 から出力される。発話者 M a は、スピーカ 5 から出力された再生音声 V a を聞くことができる。
- [0020] 本実施形態において、再生音声 V a の音響特徴量は、発話者 M a の骨導音と気導音との混合音の音響特徴量と一致又は類似する。骨導音とは、発話者 M a の声帯の振動が発話者 M a の頭蓋骨を介して発話者 M a の聴覚神経に伝わる音をいう。気導音とは、発話者 M a が発した原音声 V o が空気及び発話者 M a の鼓膜を介して発話者 M a の聴覚神経に伝わる音をいう。通常、発話者 M a が原音声 V o を発したときに発話者 M a が知覚する音声は、骨導音と気導音との混合音である。処理部 10 は、スピーカ 5 から発話者 M a の骨導音と気導音との混合音の音響特徴量と一致又は類似する音響特徴量を示す再生音声 V a が出力されるように、再生音声データを生成する。
- [0021] 本実施形態において、発話者 M a の骨導音と気導音との混合音を適宜、自己知覚音声、と称する。自己知覚音声は、発話者 M a が原音声 V o を発したときに発話者 M a が知覚する音声である。処理部 10 は、自己知覚音声が発話者 M a から出力されるように、再生音声データを生成する。

[0022] (音声処理装置)

図 2 は、本実施形態に係る音声処理装置 4 を示す機能ブロック図である。音声処理装置 4 は、コンピュータを含む。音声処理装置 4 は、プロセッサ 41 と、メインメモリ 42 と、ストレージ 43 と、インタフェース 44 とを有する。プロセッサ 41 として、CPU (Central Processing Unit) 又は MPU (Micro Processing Unit) が例示される。メインメモリ 42 として、

不揮発性メモリ又は揮発性メモリが例示される。不揮発性メモリとして、ROM (Read Only Memory) が例示される。揮発性メモリとして、RAM (Random Access Memory) が例示される。ストレージ43として、ハードディスクドライブ (HDD : Hard Disk Drive) 又はソリッドステートドライブ (SSD : Solid State Drive) が例示される。インタフェース44として、入出力回路又は通信回路が例示される。

[0023] コンピュータプログラム45がメインメモリ42に展開される。プロセッサ41は、コンピュータプログラム45に従って、本実施形態に係る発音フィードバック方法を実行する。インタフェース44は、コミュニケータ3及びスピーカ5のそれぞれと接続される。

[0024] プロセッサ41は、処理部10として機能する。ストレージ43は、記憶部12として機能する。インタフェース44は、取得部9及び出力部11として機能する。

[0025] 本実施形態において、処理部10により調整される原音声データの音響特徴量は、原音声V_oの周波数帯域、原音声V_oのピッチ、及び原音声V_oのフォルマントを含む。処理部10は、原音声V_oの周波数帯域を可聴帯域に制限するフィルタ処理部13と、原音声V_oをピッチシフトするピッチシフト処理部14と、原音声V_oをフォルマントシフトするフォルマントシフト処理部15とを含む。

[0026] フィルタ処理部13は、取得部9により取得された原音声データから可聴帯域の原音声データのみを抽出する。可聴帯域とは、ヒトが知覚可能な音声の周波数範囲をいう。ヒトの可聴帯域は、例えば15 [Hz] 以上20 [kHz] 以下である。フィルタ処理部13は、20 [kHz] 以下の周波数の原音声データを通過させるローパスフィルタと、15 [Hz] 以上の周波数の原音声データを通過させるハイパスフィルタとを含む。なお、フィルタ処理部13は、15 [Hz] 以上20 [kHz] 以下の原音声データを通過させるバンドパスフィルタを含んでもよい。なお、処理部10は、フィルタ処理部13を含まなくてもよい。

- [0027] フィルタ処理部 13 を通過した原音声データは、ピッチシフト処理部 14 に入力される。ピッチシフト処理部 14 は、原音声データをピッチシフトする。
- [0028] ピッチとは、音声の基音の周波数をいう。ピッチは、音程に影響する。ピッチは、ヒトの声帯により作られる。男性のピッチは、例えば 100 [Hz] 以上 150 [Hz] 以下である。女性のピッチは、例えば 250 [Hz] 以上 300 [Hz] 以下である。
- [0029] ピッチシフトとは、所定のピッチシフト条件に基づいて、ピッチをシフトさせることをいう。ピッチシフト条件は、ピッチシフト方向及びピッチシフト量 D_p を含む。ピッチシフト方向は、高周波数側又は低周波数側を含む。すなわち、ピッチシフトとは、ピッチを高周波数側又は低周波数側に所定のピッチシフト量 D_p だけシフトさせることをいう。
- [0030] ピッチシフト処理部 14 を通過した原音声データは、フォルマントシフト処理部 15 に入力される。フォルマントシフト処理部 15 は、原音声データをフォルマントシフトする。
- [0031] フォルマントとは、声道の共鳴によって強調される周波数成分をいう。フォルマントは、音色に影響する。フォルマントは、ヒトの声道により作られる。フォルマントは、ヒトによって異なる。周波数が最も低いフォルマントは、第 1 フォルマントと呼ばれる。第 1 フォルマントに次いで周波数が低いフォルマントは、第 2 フォルマントと呼ばれる。第 1 フォルマント及び第 2 フォルマントは、母音を決定付ける要素である。第 3 フォルマントよりも高い周波数のフォルマントは、男女差又はヒトの声の特徴を形作る要素である。第 1 フォルマントは、例えば 600 [Hz] 以上 800 [Hz] 以下である。第 2 フォルマントは、例えば 1100 [Hz] 以上 1900 [Hz] 以下である。
- [0032] フォルマントシフトとは、所定のフォルマントシフト条件に基づいて、フォルマントをシフトさせることをいう。フォルマントシフト条件は、フォルマントシフト方向及びフォルマントシフト量 D_f を含む。フォルマントシフト

ト方向は、高周波数側又は低周波数側を含む。すなわち、フォルマントシフトとは、フォルマントを高周波数側又は低周波数側に所定のフォルマントシフト量 Df だけシフトさせることをいう。

[0033] 本実施形態において、ピッチシフト条件及びフォルマントシフト条件は、予め定められており、記憶部12に記憶されている。ピッチシフト処理部14は、記憶部12に記憶されているピッチシフト条件に基づいて、ピッチシフトする。フォルマントシフト処理部15は、記憶部12に記憶されているフォルマントシフト条件に基づいて、フォルマントシフトする。

[0034] 本実施形態において、ピッチシフト方向とフォルマントシフト方向とは、同一である。すなわち、フォルマントシフト処理部15が高周波数側にフォルマントシフトした場合、ピッチシフト処理部14は高周波数側にピッチシフトする。フォルマントシフト処理部15が低周波数側にフォルマントシフトした場合、ピッチシフト処理部14は低周波数側にピッチシフトする。

[0035] ピッチシフト処理部14は、ピッチの振幅を変化させることなく、ピッチシフトする。フォルマントシフト処理部15は、フォルマントの振幅を変化させることなく、フォルマントシフトする。なお、ピッチシフト処理部14及びフォルマントシフト処理部15は、それぞれピッチシフト及びフォルマントシフトのみを精度よく行うため、振幅を変化させないとしている。ピッチシフト処理部14及びフォルマントシフト処理部15は、振幅を変化させてもよい。出力部11が、所定の周波数の振幅を変化させてもよい。

[0036] 本実施形態において、処理部10は、発話者Maの骨導音と気導音との混合音である自己知覚音声スピーカ5から出力されるように、再生音声データを生成する。フォルマントシフト処理部15は、高周波数側にフォルマントシフトする。ピッチシフト処理部14は、高周波数側にピッチシフトする。

[0037] 原音声Voを自己知覚音声である再生音声Vaに変換するためのピッチシフト量 Dp 及びフォルマントシフト量 Df は、例えば統計的に導出することができ、記憶部12に予め記憶される。ピッチシフト処理部14は、記憶部

12に記憶されているピッチシフト量 D_p だけ高周波数側にピッチシフトする。フォルマントシフト処理部15は、記憶部12に記憶されているフォルマントシフト量 D_f だけ高周波数側にフォルマントシフトする。

[0038] なお、原音声 V_o を自己知覚音声である再生音声 V_a に変換するためのピッチシフト量 D_p 及びフォルマントシフト量 D_f が、発話者 M_a ごとに定められてもよい。ピッチシフト量 D_p は、可変でもよい。フォルマントシフト量 D_f は、可変でもよい。

[0039] 図3及び図4のそれぞれは、本実施形態に係るフォルマントシフトを説明するための図である。図3及び図4に示すグラフにおいて、横軸は周波数[Hz]を示し、縦軸は振幅[dB]を示す。横軸は線形スケールである。図3及び図4に示す例において、原音声データのフォルマントは、第1フォルマント F_1 と、第2フォルマント F_2 と、第3フォルマント F_3 と、第4フォルマント F_4 とを含む。

[0040] フォルマントシフト処理部15は、原音声データについて高速フーリエ変換(FFT: Fast Fourier Transform)のような直交変換処理を実施して、フォルマントの包絡線 L_0 を含む原音声データの周波数特性を算出する。包絡線 L_0 は、複数の周波数のそれぞれの最大振幅値(最大パワー値)を結ぶように形成される。フォルマントシフト処理部15は、フォルマントの包絡線 L_0 の少なくとも一部を高周波数側にフォルマントシフト量 D_f だけフォルマントシフトする。

[0041] 図3に示すように、フォルマントシフト処理部15は、第1フォルマント F_1 及び第2フォルマント F_2 の包絡線 L_0 を高周波数側にフォルマントシフト量 D_f だけフォルマントシフトしてもよい。包絡線 L_0 がフォルマントシフトされることにより、第1フォルマント F_1 及び第2フォルマント F_2 の包絡線 L_1 が生成される。第1フォルマント F_1 のフォルマントシフト方向と第2フォルマント F_2 のフォルマントシフト方向とは、同一である。第1フォルマント F_1 のフォルマントシフト量 D_f と第2フォルマント F_2 のフォルマントシフト量 D_f とは、同一である。フォルマントシフト処理部1

5は、第1フォルマントF1の振幅及び第2フォルマントF2の振幅を変化させることなく、第1フォルマントF1及び第2フォルマントF2をフォルマントシフトする。

[0042] なお、第1フォルマントF1のフォルマントシフト量 Df と第2フォルマントF2のフォルマントシフト量 Df とは、異なってもよい。

[0043] なお、図4に示すように、フォルマントシフト処理部15は、フォルマントの包絡線L0全体を高周波数側にフォルマントシフト量 Df だけフォルマントシフトしてもよい。図4におけるフォルマントの包絡線L0全体とは、第1フォルマントF1から第4フォルマントF4までを含む範囲である。

[0044] なお、フォルマントシフト量 Df は、第1フォルマントF1のピーク周波数 $P0$ に基づいて決定されてもよい。ピーク周波数 $P0$ における第1フォルマントF1の振幅が $A0$ である場合、フォルマントシフト処理部15は、振幅 $A0$ の80 [%] となる振幅 $A1$ と、振幅 $A1$ における第1フォルマントF1の周波数 $P1$ とを算出する。フォルマントシフト量 Df は、ピーク周波数 $P0$ と周波数 $P1$ との差を超えないように定められてもよい。フォルマントシフト量 Df は、ピーク周波数 $P0$ と周波数 $P1$ との差でもよい。なお、振幅 $A1$ は振幅 $A0$ の70 [%] 以上100 [%] 未満であればよく、振幅 $A0$ の80 [%] 程度とするのが好適である。

[0045] なお、フォルマントシフトにおいて、フォルマントの包絡線L0全体がシフトされなくてもよい。基本周波数の時間変化や振幅包絡の時間情報等が保持されていれば、包絡線L0においてピークを含む所定の周波数範囲だけをシフトさせてもよい。

[0046] 原音声データは、フィルタ処理部13、ピッチシフト処理部14、及びフォルマントシフト処理部15のそれぞれで処理されることにより、再生音声データに変換される。再生音声 Va は、スピーカ5によって再生される。発話者 Ma は、スピーカ5から出力された再生音声 Va を聞くことができる。

[0047] (発音フィードバック方法)

図5は、本実施形態に係る発音フィードバック方法を示すフローチャート

である。コンピュータプログラム45は、発音フィードバック方法を音声処理装置4に実行させることができる。

[0048] 発話者Maは、マイクロホン2に向かって原音声Voを発する。取得部9は、発話者Maが発した原音声Voを示す原音声データを取得する（ステップS1）。

[0049] フィルタ処理部13は、原音声データの周波数帯域を可聴帯域に制限する（ステップS2）。なお、ステップS2は任意の処理である。

[0050] ピッチシフト処理部14は、フィルタ処理部13を通過した原音声データをピッチシフトする（ステップS3）。

[0051] フォルマントシフト処理部15は、ピッチシフト処理部14を通過した原音声データをフォルマントシフトする（ステップS4）。

[0052] ステップS3及びステップS4により、再生音声Vaを示す再生音声データが生成される。処理部10は、ステップS3及びステップS4において、自己知覚音声スピーカ5から出力されるように、再生音声データを生成する。処理部10は、ステップS2、ステップS3、及びステップS4により、再生音声Vaを示す再生音声データを生成してもよい。なお、ステップS2、ステップS3、及びステップS4の順序は任意である。

[0053] 出力部11は、処理部10において生成された再生音声データをスピーカ5に出力する（ステップS5）。

[0054] スピーカ5は、再生音声Vaを発話者Maに出力する。スピーカ5から出力される再生音声Vaは、発話者Maの自己知覚音声と類似する。

[0055] （適用例）

図6は、本実施形態に係る発音フィードバック装置1の第1適用例を説明するための図である。図6は、発音フィードバック装置1が携帯電話20に適用された例を示す。携帯電話20は、送話口21と、受話口22とを有する。マイクロホン2が送話口21に配置される。スピーカ5が受話口22に配置される。音声処理装置4は、携帯電話20の内部に配置される。

[0056] 例えば雑音環境下で電話する場合、発話者Maは、発話者Maが発した音

声を聞き取り難いため、普段よりも大きい声で話したり高い声で話したりする可能性がある。本実施形態においては、発話者M aが送話口21に発した原音声V oが、音声処理装置4において再生音声V aに変換される。再生音声V aは、受話口22から出力される。発話者M aは、自己知覚音声である再生音声V aを聞きながら話すことができる。したがって、雑音環境下で電話する場合において、発話者M aが普段よりも大きい声で話したり高い声で話したりすることが抑制される。

[0057] 図7は、本実施形態に係る発音フィードバック装置1の第2適用例を説明するための図である。図7は、発音フィードバック装置1が歌唱練習装置30に適用された例を示す。歌唱練習装置30は、マイクスタンド31に支持されるマイクロホン2と、スピーカ5を含むモニタスピーカ32とを有する。音声処理装置4は、マイクロホン2とモニタスピーカ32との間に配置される。

[0058] 発話者M aが自己認識音声である再生音声V aを聞きながら歌唱すると、歌唱の音程が安定する場合が多い。本実施形態においては、発話者M aがマイクロホン2に発した歌唱音声である原音声V oが、音声処理装置4において再生音声V aに変換される。再生音声V aは、モニタスピーカ32から出力される。発話者M aは、自己知覚音声である再生音声V aを聞きながら歌唱することができる。これにより、発話者M aの歌唱の音程は安定する。

[0059] (効果)

以上説明したように、本実施形態によれば、発話者M aが発した原音声V oを示す原音声データの音響特徴量が処理部10により調整される。処理部10は、原音声データの音響特徴量を調整して、再生音声V aを示す再生音声データを生成する。出力部11は、再生音声データをスピーカ5に出力する。スピーカ5は、再生音声V aを発話者M aに出力する。これにより、発話者M aに適正な再生音声V aがフィードバックされる。発話者M aに適正な再生音声V aがフィードバックされるので、雑音環境下で発話者M aが普段よりも大きい声で話したり高い声で話したりする現象が抑制される。

[0060] 原音声 V_o を自己知覚音声である再生音声 V_a に変換する場合、ピッチシフト及びフォルマントシフトが有効である。また、原音声 V_o を自己知覚音声である再生音声 V_a に変換する場合、フォルマントシフト方向とピッチシフト方向とを一致させることが有効である。

[0061] また、原音声 V_o を自己知覚音声である再生音声 V_a に変換する場合、ピッチシフト及びフォルマントシフトの前に、原音声 V_o の周波数帯域を可聴帯域に制限するフィルタ処理を実施することが有効である。

[0062] (変形例)

図8は、本実施形態の変形例に係る発音フィードバック装置101を示す模式図である。上述の実施形態においては、ピッチシフト条件及びフォルマントシフト条件が予め記憶部12に記憶されていることとした。図8に示すように、発音フィードバック装置101は、ピッチシフト条件及びフォルマントシフト条件を調整する操作装置16を備えてもよい。図8に示すように、操作装置16は、音声処理装置4に接続される。操作装置16は、ピッチシフト条件を調整するピッチスライダ16Aと、フォルマントシフト条件を調整するフォルマントスライダ16Bとを有する。ピッチスライダ16Aがスライドされることにより、ピッチシフト方向及びピッチシフト量 D_p を含むピッチシフト条件が変更される。フォルマントスライダ16Bがスライドされることにより、フォルマントシフト方向及びフォルマントシフト量 D_f を含むフォルマントシフト条件が変更される。発話者Maは、スピーカ5から出力される再生音声 V_a を聞きながら、再生音声 V_a が自己知覚音声に近付くように、操作装置16を操作することができる。

[0063] 上述の実施形態においては、再生音声 V_a が自己知覚音声であることとした。再生音声 V_a は自己知覚音声でなくてもよい。また、ピッチシフト処理部14は、低周波数側にピッチシフトしてもよい。フォルマントシフト処理部15は、低周波数側にフォルマントシフトしてもよい。発話者Maがピッチの変化及びフォルマントの変化を知覚できる程度にピッチシフト及びフォルマントシフトが実施されればよい。再生音声 V_a が発話者Maにフィード

バックされることにより、発話者M aは、音声生成及び音声知覚を適正に行うことができる。音声生成及び音声知覚が適正に行われるので、発話者M aの吃音又は活舌の悪さに起因して、発話者M aの音声を認識し難くなる現象が抑制される。例えば英語学習において再生音声V aを聞きながら発音練習をする場合、スピーカ5から出力される再生音声V aを発話者M aが適正に認識できるように、原音声データの音響特徴量が調整されてもよい。

[0064] [第2実施形態]

第2実施形態について説明する。以下の説明において、上述の実施形態と同一又は同等の構成要素については同一の符号を付し、その構成要素の説明を簡略又は省略する。

[0065] 図9は、本実施形態に係る発音フィードバック装置102を示す模式図である。本実施形態において、出力部11は、スピーカ5を含むイヤホン50に再生音声データを出力する。イヤホン50は、発話者M aの外耳道に挿入されるインナイヤホンである。イヤホン50は、外耳道において再生音声V aを出力する。

[0066] イヤホン50は、外耳道の内面に接触するイヤピース51を含む。イヤピース51は、耳栓としても機能する。イヤピース51は、例えばゴム製、シリコーン製、及びウレタン製等である。なお、イヤピース51は、指で押すと変形する軟質材料で形成されていればよい。

[0067] イヤピース51の形状により、発話者M aの鼓膜に伝達される外部音声V nの周波数特性が変化する。外部音声V nとは、外耳道の外部から鼓膜に伝達される音声をいう。外部音声V nとして、発話者M aの周囲の雑音が例示される。

[0068] イヤピース51の形状は、イヤピース51が変形した場合の形状を含む。イヤピース51が変形することによっても、発話者M aの鼓膜に伝達される外部音声V nの周波数特性が変化する。

[0069] 図10は、本実施形態に係るイヤホン50の周波数特性を説明するための図である。図10に示すラインH A, H B, H C, H Dのように、イヤピー

ス51の形状が変化することにより、発話者Maの鼓膜に伝達される外部音声Vnの周波数特性が変化する。

[0070] ラインHAは、第1直径のイヤピース51に係る外部音声Vnの周波数特性を示す。ラインHBは、第1直径よりも大きい第2直径のイヤピース51に係る外部音声Vnの周波数特性を示す。ラインHCは、第2直径よりも大きい第3直径のイヤピース51に係る外部音声Vnの周波数特性を示す。ラインHDは、第3直径よりも大きい第4直径のイヤピース51に係る外部音声Vnの周波数特性を示す。4形態のイヤピース51のうち、第1直径のイヤピース51は、最も緩めに外耳道に挿入される。第1直径のイヤピース51は、ほぼ変形しない状態で外耳道に挿入される。4形態のイヤピース51のうち、第4直径のイヤピース51は、最もきつめに外耳道に挿入される。第4直径のイヤピース51は、最も変形した状態で外耳道に挿入される。4形態のイヤピース51のうち、第4直径のイヤピース51は、外耳道を最も密閉する。

[0071] なお、発話者Maの鼓膜の近傍にマイクを配置し、4形態のイヤピース51のそれぞれを外耳道に挿入した状態で、外部音声Vnを外耳道に入力することにより、発話者Maの鼓膜に伝達される外部音声Vnの周波数特性を測定することができる。外部音声Vnは、インパルス音等の所定の周波数パターンを備えた信号を用いてもよいし、列車や自動車の走行音等の各種騒音を用いてもよい。

[0072] 図10に示すように、イヤピース51の直径が大きくなるほど、鼓膜に対する外部音声Vnの遮断効果が高まり、特に高周波数帯域においてゲインが低下する。ここでいう高周波数帯域は、1000 [Hz] 以上20 [kHz] 以下の周波数帯域である。

[0073] 図9に示すように、本実施形態において、処理部10は、再生音声Vaの周波数特性を調整する調整部17を含む。調整部17は、イヤホン50のイヤピース51の形状を模擬するように、イヤホン50のスピーカ5から発話者Maの鼓膜に伝達される再生音声Vaの周波数特性を調整する。記憶部1

2には、ラインHA、HB、HC、HDで示した周波数特性が記憶されている。調整部17は、再生音声Vaのゲインを調整するゲインコントロール機能を有する。なお、記憶部12には、ラインHA、HB、HC、HDで示した周波数特性に限らず、所定の周波数帯域における複数の異なる周波数特性が記憶されていてもよい。

[0074] 上述の実施形態と同様、処理部10は、フィルタ処理部13、ピッチシフト処理部14、及びフォルマントシフト処理部15を含む。フォルマントシフト処理部15から再生音声データが出力される。調整部17は、フォルマントシフト処理部15から出力された再生音声データの周波数特性（ゲイン）を調整する。

[0075] 調整部17に操作装置18が接続される。操作装置18は、回転可能なノブを含む。調整部17は、操作装置18の操作量に基づいて、再生音声Vaの周波数特性を、ラインHAで示す周波数特性、ラインHBで示す周波数特性、ラインHCで示す周波数特性、及びラインHDで示す周波数特性のそれぞれに変化させる。

[0076] 本実施形態において、調整部17は、操作装置18の操作量に基づいて、スピーカ5から出力される再生音声Vaの音量を調整する。調整部17は、操作装置18の操作量に基づいて、再生音声Vaの周波数特性の変化と連動するように、再生音声Vaの音量を変化させる。再生音声Vaの周波数特性がラインHAで示す周波数特性に調整される場合、再生音声Vaの音量は、第1音量に調整される。再生音声Vaの周波数特性がラインHBで示す周波数特性に調整される場合、再生音声Vaの音量は、第1音量よりも小さい第2音量に調整される。再生音声Vaの周波数特性がラインHCで示す周波数特性に調整される場合、再生音声Vaの音量は、第2音量よりも小さい第3音量に調整される。再生音声Vaの周波数特性がラインHDで示す周波数特性に調整される場合、再生音声Vaの音量は、第3音量よりも小さい第4音量に調整される。

[0077] 例えば、再生音声Vaの周波数特性がラインHAで示す周波数特性に調整

されることにより、発話者 M_a は、クリアな再生音声 V_a を第 1 音量で聞けることができる。再生音声 V_a の周波数特性がライン $H D$ で示す周波数特性に調整されることにより、発話者 M_a は、こもった再生音声 V_a を第 4 音量で聞けることができる。発話者 M_a は、発話者 M_a の好みに合わせて再生音声 V_a の周波数特性及び再生音声 V_a の音量を調整することができる。

符号の説明

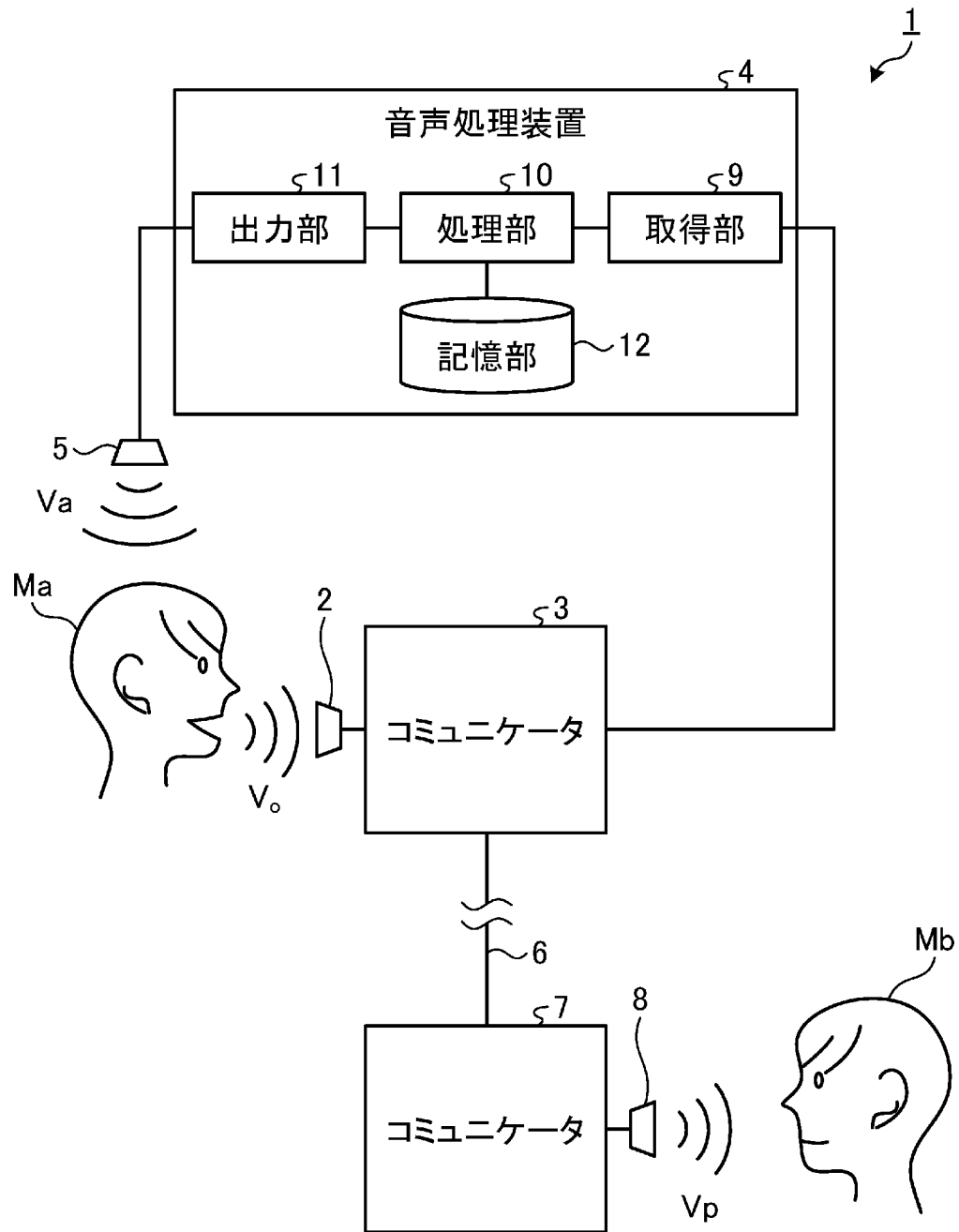
[0078] 1…発音フィードバック装置、2…マイクロホン、3…コミュニケータ、4…音声処理装置、5…スピーカ、6…伝送装置、7…コミュニケータ、8…スピーカ、9…取得部、10…処理部、11…出力部、12…記憶部、13…フィルタ処理部、14…ピッチシフト処理部、15…フォルマントシフト処理部、16…操作装置、16A…ピッチスライダ、16B…フォルマントスライダ、17…調整部、18…操作装置、20…携帯電話、21…送話口、22…受話口、30…歌唱練習装置、31…マイクスタンド、32…モニタスピーカ、41…プロセッサ、42…メインメモリ、43…ストレージ、44…インタフェース、45…コンピュータプログラム、50…イヤホン、51…イヤピース、101…発音フィードバック装置、102…発音フィードバック装置、A0…振幅、A1…振幅、 D_f …フォルマントシフト量、 D_p …ピッチシフト量、F1…第1フォルマント、F2…第2フォルマント、F3…第3フォルマント、F4…第4フォルマント、L0…包絡線、L1…包絡線、 M_a …発話者、 M_b …視聴者、P0…ピーク周波数、P1…周波数、 V_a …再生音声、 V_n …外部音声、 V_o …原音声、 V_p …再生音声。

請求の範囲

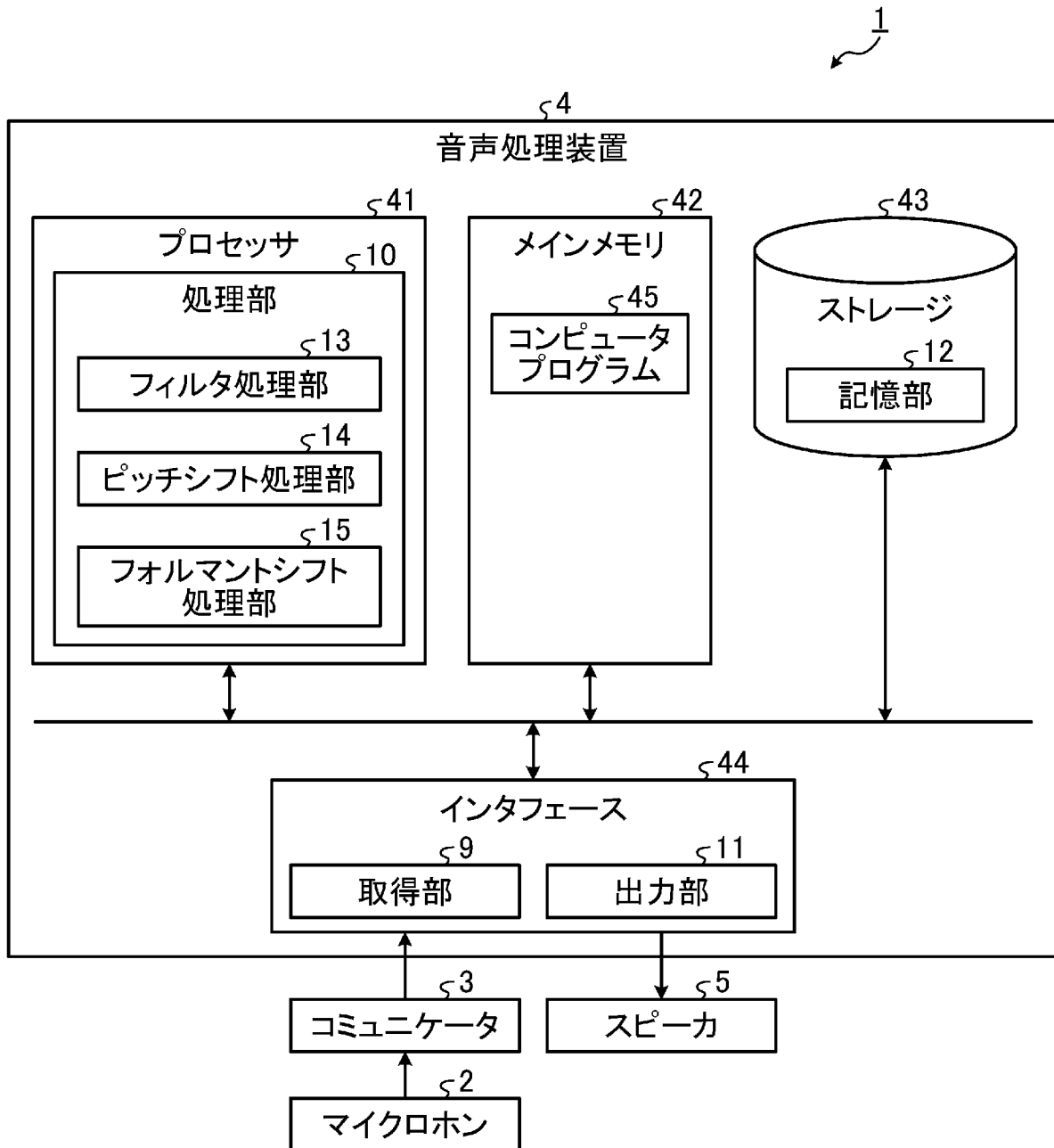
- [請求項1] 発話者が発した原音声を示す原音声データを取得する取得部と、
前記原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成する処理部と、
前記再生音声データを出力する出力部と、を備える、
発音フィードバック装置。
- [請求項2] 前記音響特徴量は、前記原音声のピッチ及びフォルマントを含み、
前記処理部は、ピッチシフトするピッチシフト処理部及びフォルマントシフトするフォルマントシフト処理部を含む、
請求項1に記載の発音フィードバック装置。
- [請求項3] 前記処理部は、高周波数側にフォルマントシフトした場合、高周波数側にピッチシフトし、低周波数側にフォルマントシフトした場合、低周波数側にピッチシフトする、
請求項2に記載の発音フィードバック装置。
- [請求項4] 前記音響特徴量は、前記原音声の周波数帯域を含み、
前記処理部は、前記周波数帯域を可聴帯域に制限するフィルタ処理部を含む、
請求項1から請求項3のいずれか一項に記載の発音フィードバック装置。
- [請求項5] 発話者が発した原音声を示す原音声データを取得するステップと、
前記原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成するステップと、
前記再生音声データを出力するステップと、を含む、
発音フィードバック方法。
- [請求項6] 発話者が発した原音声を示す原音声データを取得するステップと、
前記原音声データの音響特徴量を調整して再生音声を示す再生音声データを生成するステップと、
前記再生音声データを出力するステップと、を含む発音フィードバ

ック方法を、コンピュータに実行させる、
コンピュータプログラム。

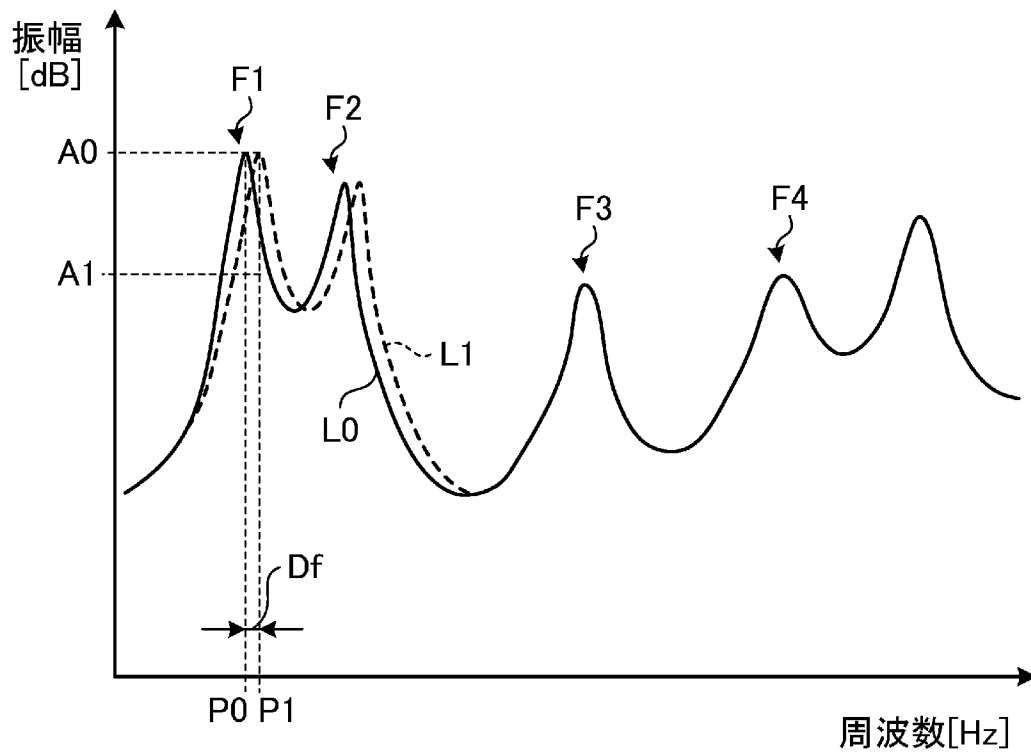
[図1]



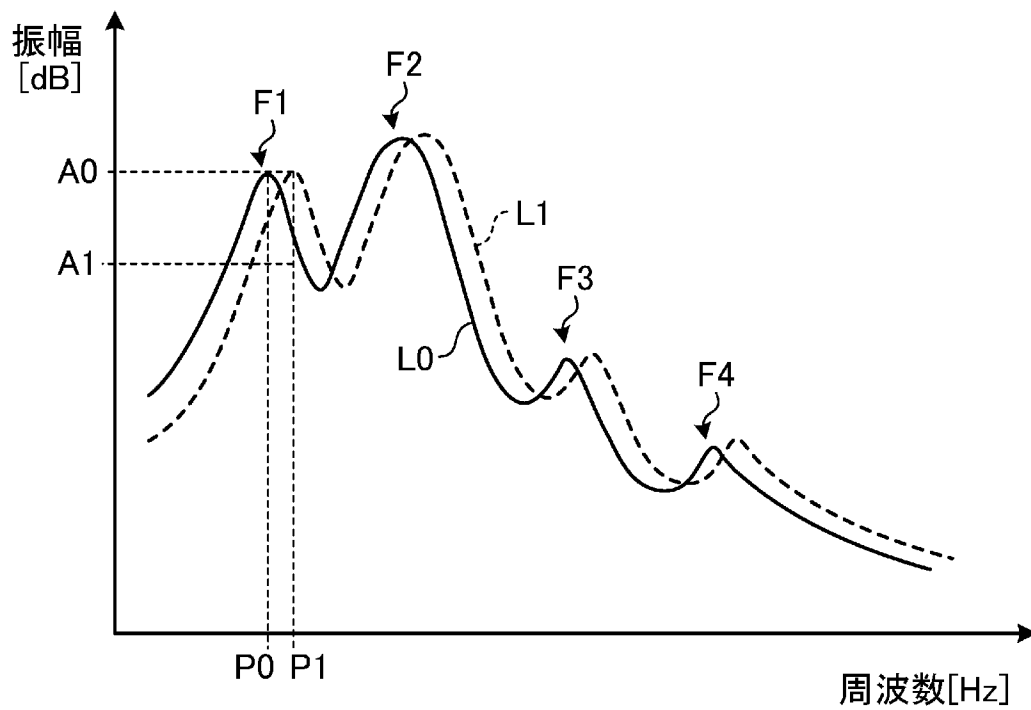
[図2]



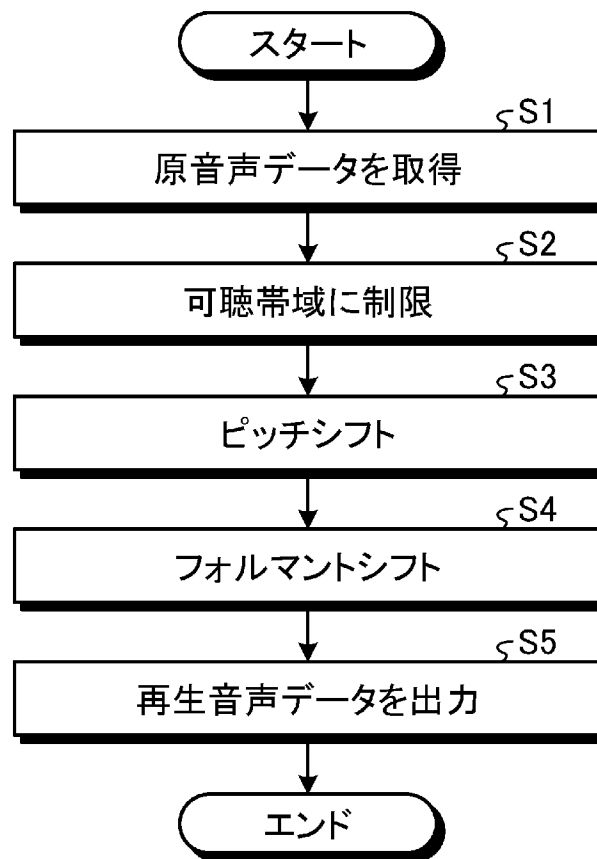
[図3]



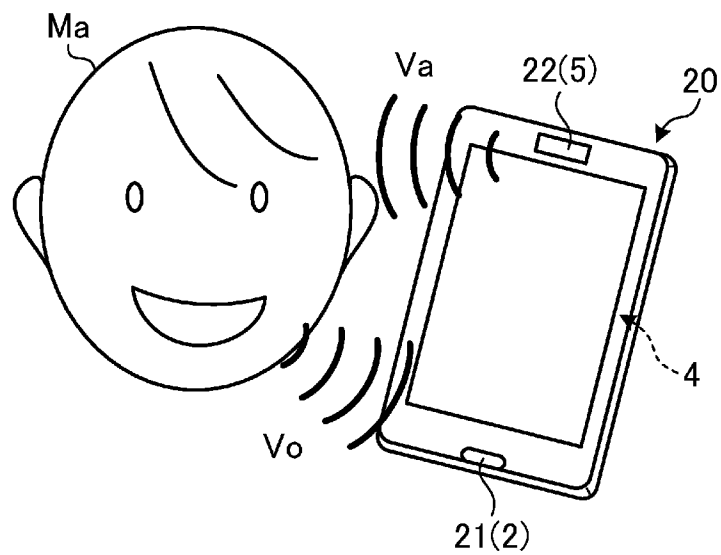
[図4]



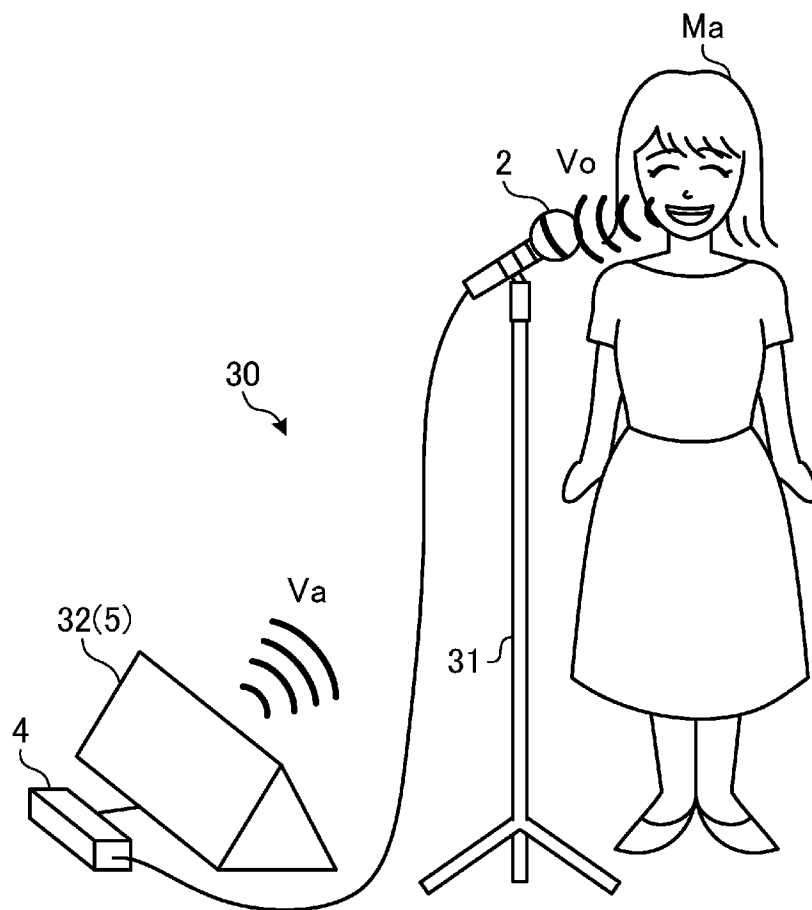
[図5]



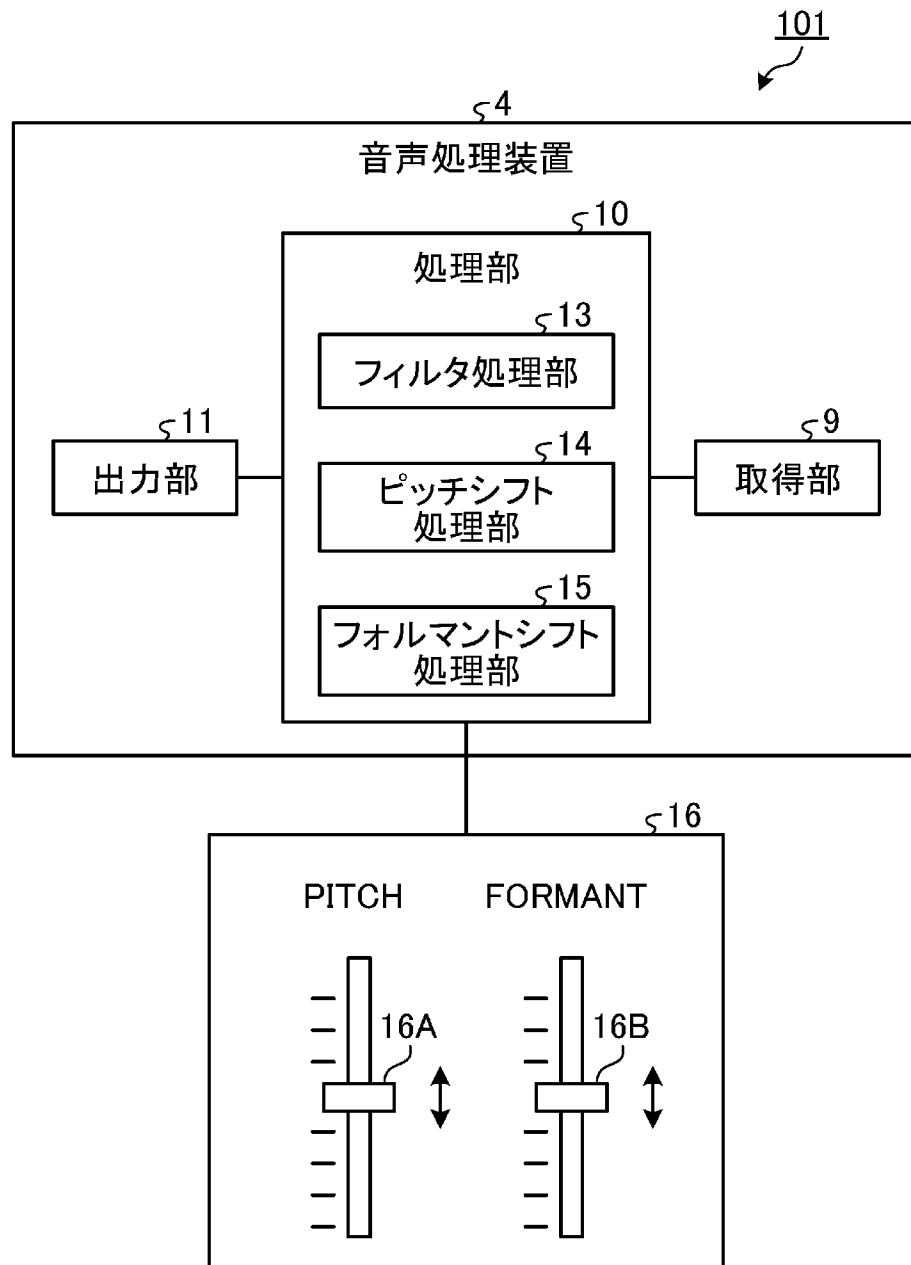
[図6]



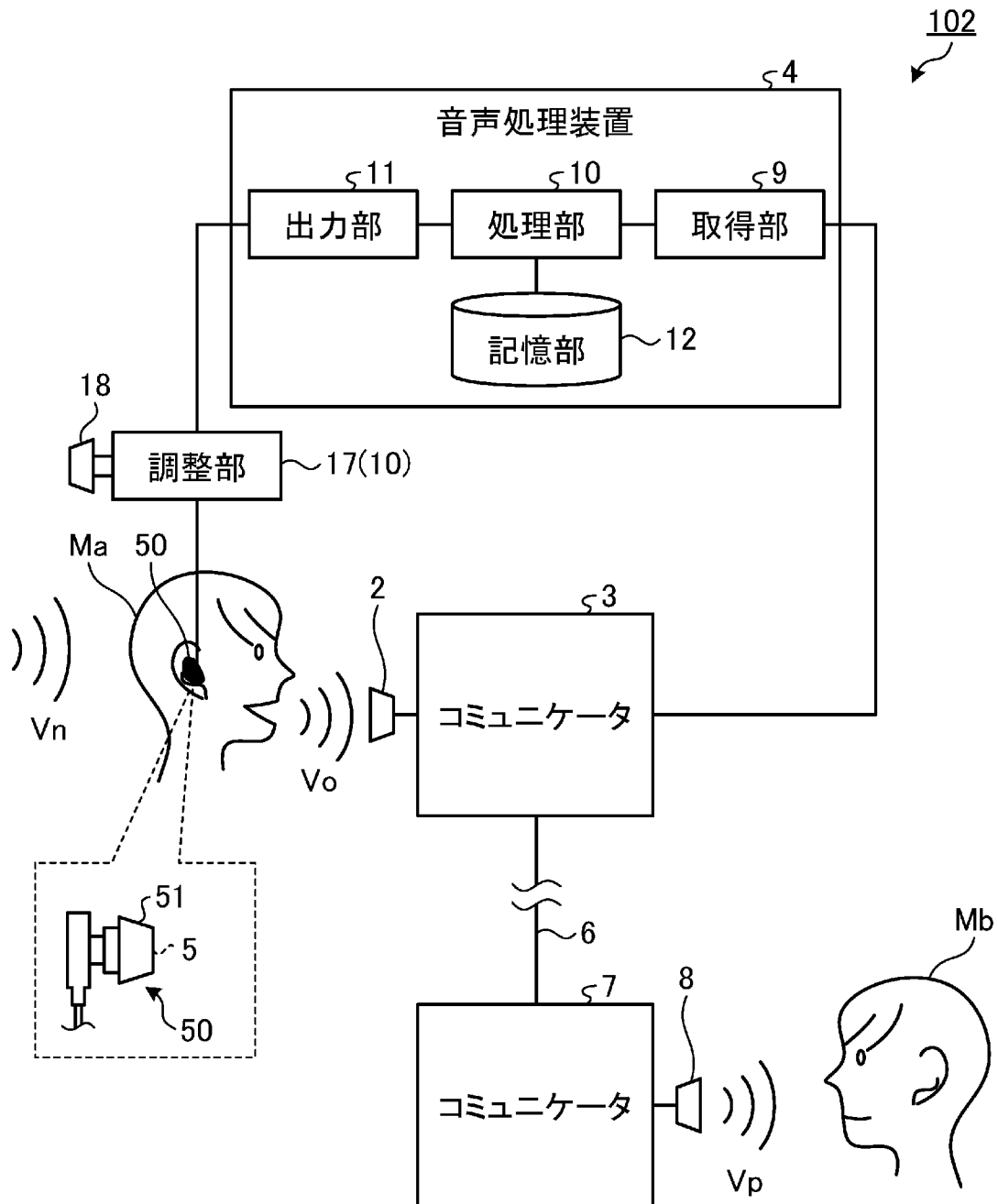
[図7]



[図8]

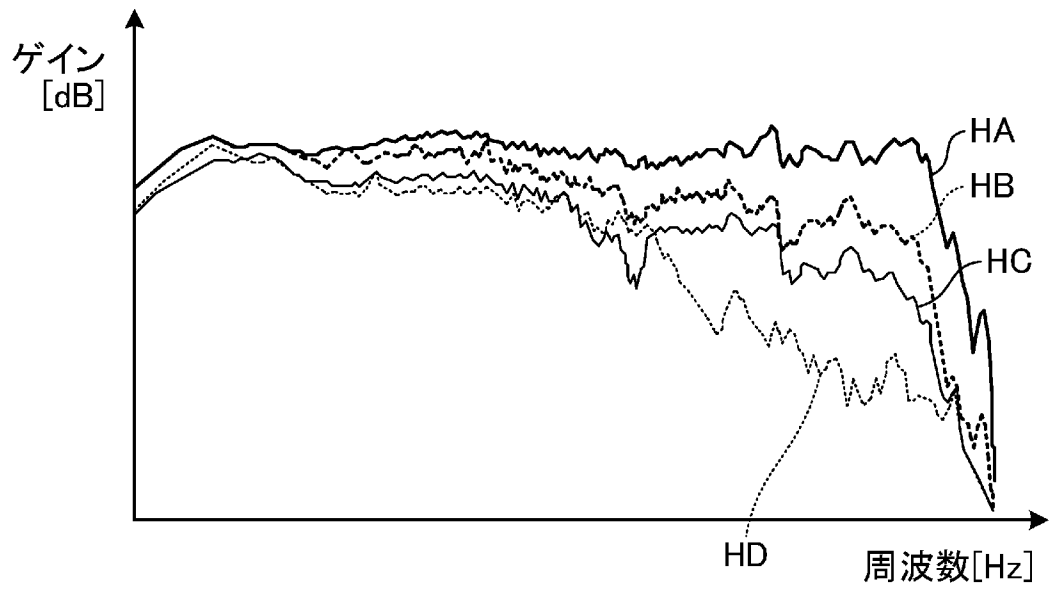


[図9]



[図10]

[イヤホンの周波数特性]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/035137

A. CLASSIFICATION OF SUBJECT MATTER**H04R 3/00**(2006.01)i; **H04R 25/00**(2006.01)i; **G10L 21/0364**(2013.01)i

FI: H04R3/00 310; G10L21/0364; H04R25/00 L

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; H04R25/00; G10L21/0364

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2021

Registered utility model specifications of Japan 1996-2021

Published registered utility model applications of Japan 1994-2021

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 10-224898 A (SANYO ELECTRIC CO., LTD.) 21 August 1998 (1998-08-21) paragraphs [0014]-[0102], fig. 1-10	1-3, 5, 6
Y		4
X	JP 2006-501929 A (UNIV. EAST CAROLINA) 19 January 2006 (2006-01-19) paragraphs [0056]-[0080], fig. 6	1-6
Y		4



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

01 November 2021

Date of mailing of the international search report

16 November 2021

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2021/035137

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	10-224898	A	21 August 1998	(Family: none)	
JP	2006-501929	A	19 January 2006	US 2006/0089522 A1	
				paragraphs [0076]-[0098], fig. 6	
				WO 2004/032816 A2	
				KR 10-2005-0053744 A	
				CN 1713873 A	

A. 発明の属する分野の分類（国際特許分類（IPC）） H04R 3/00(2006.01)i; H04R 25/00(2006.01)i; G10L 21/0364(2013.01)i FI: H04R3/00 310; G10L21/0364; H04R25/00 L														
B. 調査を行った分野														
調査を行った最小限資料（国際特許分類（IPC）） H04R3/00; H04R25/00; G10L21/0364														
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2021年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2021年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2021年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2021年	日本国実用新案登録公報	1996 - 2021年	日本国登録実用新案公報	1994 - 2021年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2021年													
日本国実用新案登録公報	1996 - 2021年													
日本国登録実用新案公報	1994 - 2021年													
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）														
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
X	JP 10-224898 A（三洋電機株式会社）21.08.1998（1998 - 08 - 21） 段落[0014]-[0102], 図1-10	1-3, 5, 6												
Y		4												
X	JP 2006-501929 A（イースト カロライナ ユニバーシティ）19.01.2006（2006 - 01 - 19） 段落[0056]-[0080], 図6	1-6												
Y		4												
☐ 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。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“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</td> </tr> <tr> <td>“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</td> <td>“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</td> </tr> <tr> <td>“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</td> <td>“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</td> </tr> <tr> <td>“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）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“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</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“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	“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	“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	“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）	“&” 同一パテントファミリー文献	“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	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“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													
“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	“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													
“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）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 01.11.2021	国際調査報告の発送日 16.11.2021													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 堀 洋介 5Z 3996 電話番号 03-3581-1101 内線 3591													

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2021/035137

引用文献			公表日	パテントファミリー文献	公表日
JP	10-224898	A	21.08.1998	(ファミリーなし)	
JP	2006-501929	A	19.01.2006	US 2006/0089522 A1	
				段落[0076]-[0098], 図6	
				WO 2004/032816 A2	
				KR 10-2005-0053744 A	
				CN 1713873 A	