

(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
2017 年 4 月 20 日(20.04.2017)



(10) 国際公開番号

WO 2017/065092 A1

- (51) 国際特許分類:
H04R 1/02 (2006.01) H04M 1/02 (2006.01)
G10L 21/0208 (2013.01) H04R 1/40 (2006.01)
H04M 1/00 (2006.01) H04R 3/00 (2006.01)
- (21) 国際出願番号: PCT/JP2016/079855
- (22) 国際出願日: 2016 年 10 月 6 日(06.10.2016)
- (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 2015-201723 2015 年 10 月 13 日(13.10.2015) JP
特願 2016-133593 2016 年 7 月 5 日(05.07.2016) JP
PCT/JP2016/073655 2016 年 8 月 10 日(10.08.2016) JP
特願 2016-177402 2016 年 9 月 12 日(12.09.2016) 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: 関矢 俊之 (SEKIYA, Toshiyuki); 〒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). 小山 裕一郎(KOY-

AMA, Yuichiro); 〒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). 平野雄哉(HIRANO, Yuya); 〒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).

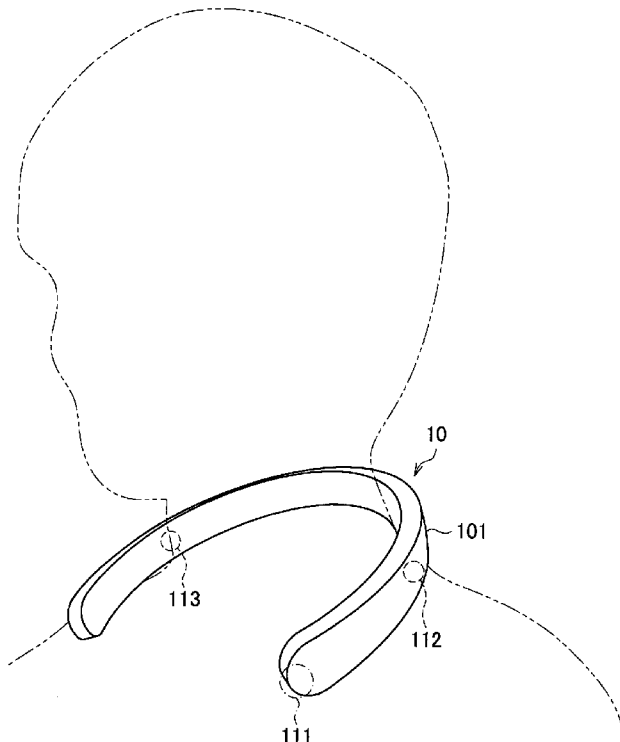
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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,

[続葉有]

(54) Title: INFORMATION PROCESSING DEVICE

(54) 発明の名称: 情報処理装置



(57) Abstract: [Problem] Desired sound is collected in a more suitable form even in an environment in which noises randomly occur. [Solution] An information processing device according to the present invention includes a sound collector and a support member which has a protrusion with a streamlined shape on a part thereof and which supports the sound collector at the tip end of the protrusion or at a position in the vicinity of the tip end.

(57) 要約: 【課題】雑音がランダムに発生するような環境下においても、目的音をより好適な態様で集音する。【解決手段】集音部と、流線形の形状を有する凸部を少なくとも一部に備え、前記凸部の先端、または、当該先端の近傍に位置するように前記集音部を支持する支持部材と、を備える、情報処理装置。

WO 2017/065092 A1



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:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称： 情報処理装置

技術分野

[0001] 本開示は、情報処理装置に関する。

背景技術

[0002] 近年では、通信技術の進歩や各種デバイスの小型化に伴い、所謂情報処理装置と呼ばれる機器の種別も多様化してきており、P C (Personal Computer) 等に限らず、スマートフォンやタブレット端末のように、ユーザが携行可能に構成された情報処理装置も普及してきている。特に、近年では、ユーザが身体の一部に装着することで携行しながら使用可能に構成された、所謂ウェアラブルデバイスも提案されている。

[0003] また、近年では、所謂音声認識技術や自然言語処理技術の発展に伴い、ユーザが音声入力により、各種処理を指示することが可能なユーザインタフェース (U I : User Interface) を有する情報処理装置も普及してきている。

先行技術文献

特許文献

[0004] 特許文献1：特開2012-203122号公報

発明の概要

発明が解決しようとする課題

[0005] また、音声認識や音声通話等のためにユーザが発する音声を集音可能に構成された情報処理装置においては、集音対象となる音声以外の他の音響（即ち、雑音）を抑圧することで、当該音声の集音品質をより向上させることが可能な仕組みが検討されている。例えば、特許文献1には、雑音を抑圧するための仕組みの一例が開示されている。

[0006] 一方で、情報処理装置が屋外で 사용되는場合等のように、当該情報処理装置の利用シーンの多様化に伴い、当該情報処理装置の周囲の環境が動的に変化するような状況が想定される。このような状況下においては、風切音や

振動に伴う音のように、情報処理装置上で発生する音が、雑音として集音される場合も想定され得る。このような音は、発生する箇所、発生する時刻が不規則でランダムに発生する雑音となる。

[0007] そこで、本開示では、雑音がランダムに発生するような環境下においても、目的音をより好適な態様で集音することが可能な情報処理装置を提案する。

課題を解決するための手段

[0008] 本開示によれば、集音部と、流線形の形状を有する凸部を少なくとも一部に備え、前記凸部の先端、または、当該先端の近傍に位置するように前記集音部を支持する支持部材と、を備える、情報処理装置が提供される。

発明の効果

[0009] 以上説明したように本開示によれば、雑音がランダムに発生するような環境下においても、目的音をより好適な態様で集音することが可能な情報処理装置が提供される。

[0010] なお、上記の効果は必ずしも限定的なものではなく、上記の効果とともに、または上記の効果に代えて、本明細書に示されたいずれかの効果、または本明細書から把握され得る他の効果が奏されてもよい。

図面の簡単な説明

[0011] [図1]本開示の第1の実施形態に係る情報処理装置の概略的な構成の一例について説明するための説明図である。

[図2]同実施形態に係る情報処理装置の概略的な構成の一例について説明するための説明図

[図3]風切音の影響を観測するための観測環境の一例について説明するための説明図である。

[図4]情報処理装置に設けられた複数の集音部それぞれの設置位置の一例を示している。

[図5]情報処理装置に対して、互いに異なる角度から風を当てた場合における、当該集音部それぞれによる風切音の観測結果の一例について説明するため

の説明図である。

[図6]同実施形態に係る情報処理装置の機能構成の一例を示したブロック図である。

[図7]同実施形態に係る情報処理装置が、複数の集音部それぞれの集音結果に基づき目的音を取得する処理の一例を示している。

[図8]同実施形態に係る情報処理装置の一連の処理の流れの一例を示したフローチャートである。

[図9]実施例 1 に係る情報処理装置の一例について説明するための説明図である。

[図10]実施例 1 に係る情報処理装置の他の一例について説明するための説明図である。

[図11]実施例 1 に係る情報処理装置の他の一例について説明するための説明図である。

[図12]実施例 2 に係る情報処理装置の一例について説明するための説明図である。

[図13]実施例 2 に係る情報処理装置の他の一例について説明するための説明図である。

[図14]実施例 2 に係る情報処理装置の他の一例について説明するための説明図である。

[図15]実施例 2 に係る情報処理装置の他の一例について説明するための説明図である。

[図16]実施例 3 に係る情報処理装置の一例について説明するための説明図である。

[図17]変形例 3 に係る情報処理装置 30 の利用形態の一例について説明するための説明図である。

[図18]実施例 4 に係る情報処理装置の一例について説明するための説明図である。

[図19]実施例 4 に係る情報処理装置の他の一例について説明するための説明

図である。

[図20]実施例5に係る情報処理装置の一例について説明するための説明図である。

[図21]実施例5に係る情報処理装置における撮像部のレンズの近傍の概略的な構成の一例について説明するための説明図である。

[図22]本開示の第2の実施形態に係る情報処理装置の機能構成の一例を示したブロック図である。

[図23]非相関成分パワー推定部の処理の基本原理について説明するための説明図である。

[図24]本開示の第3の実施形態に係る情報処理装置の機能構成の一例を示したブロック図である。

[図25]同実施形態に係る信号処理装置のハードウェア構成の一例を示した図である。

発明を実施するための形態

[0012] 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。

[0013] なお、説明は以下の順序で行うものとする。

1. 第1の実施形態

1. 1. 概要

1. 2. 集音部の設置位置の検討

1. 3. 機能構成

1. 4. 処理

1. 5. 実施例

1. 5. 1. 実施例1：首に装着されるウェアラブルデバイスの一例

1. 5. 2. 実施例2：頭部に装着されるウェアラブルデバイスの一

例

1. 5. 3. 実施例 3 : 携帯型情報端末への適用例

1. 5. 4. 実施例 4 : 時計型のウェアラブルデバイスへの適用例

1. 5. 5. 実施例 5 : 撮像装置への適用例

2. 第 2 の実施形態

2. 1. 概要

2. 2. 機能構成

2. 3. 非相関成分パワー推定部の詳細

2. 4. ランダムノイズパワー推定部の詳細

2. 5. 評価

3. 第 3 の実施形態

3. 1. 概要

3. 2. 機能構成

3. 3. マルチチャネルウィナーフィルタの算出方法の詳細

3. 4. 評価

4. ハードウェア構成

5. むすび

[0014] << 1. 第 1 の実施形態 >>

< 1. 1. 概要 >

まず、図 1 を参照して、本開示の第 1 の実施形態に係る情報処理装置の概略的な構成の一例について説明し、次いで、本実施形態に係る情報処理装置の技術的課題について説明する。図 1 は、本開示の第 1 の実施形態に係る情報処理装置の概略的な構成の一例について説明するための説明図である。

[0015] 図 1 に示す例では、情報処理装置 10 は、所謂ウェアラブルデバイスとして構成されている。より具体的には、情報処理装置 10 は、一部が開口したリング状の形状（換言すると、カチューシャ状、または、U 字状の形状）を成し、リング状の部分の内面のうち少なくとも一部が、ユーザの首の一部に当接するように（即ち、首に掛けるように）当該ユーザに装着される。

[0016] また、情報処理装置 10 は、所謂マイクロフォンのような集音部を備え、

当該集音部からユーザが発する音声を音響情報として集音する。例えば、図 1 に示す例では、情報処理装置 10 は、参照符号 111～113 で示される複数の集音部を備える。より具体的には、集音部 111～113 は、例えば、情報処理装置 10 の筐体 101 により支持される。

[0017] 例えば、図 2 は、本実施形態に係る情報処理装置 10 の概略的な構成の一例について説明するための説明図であり、当該情報処理装置 10 のうち集音部 111 が設けられた部分の構成の一例を示した図である。図 1 及び図 2 に示すように、情報処理装置 10 は、ユーザの首に装着されたときに、当該ユーザの口の近傍に、当該ユーザの前方に向かって突出するように流線形の形状を有する凸部が設けられており、当該凸部の先端（またな、当該先端の近傍）に、当該凸部が突出する方向を向くように集音部 111 が設けられている。また、集音部 111 は、情報処理装置 10 とは個別のデバイスとして構成され、当該凸部の先端（またな、当該先端の近傍）に対して、当該凸部が突出する方向を向くように支持されていてもよい。なお、以降の説明においては、情報処理装置 10 に対して集音部 110 が設けられていると記載した場合には、当該集音部 110 が情報処理装置 10 とは別体として、当該情報処理装置 10 の少なくとも一部に支持されている場合も含み得るものとする。

[0018] また、図 1 に示すように、集音部 112 及び 113 は、情報処理装置 10 に対して、互いに異なる方向を向くように設けられている。より具体的には、集音部 112 及び 113 は、情報処理装置 10 がユーザの首に装着されたときに、当該ユーザの首を基準として、互いに略対称となる位置に設けられている。なお、各集音部が設けられる位置については、詳細を別途後述する。また、図 1 に示す例では、集音部 112 及び 113 は、リング状の形状の筐体 101 に対して、当該リングの外側（即ち、リングの中心とは逆側）を向くように設けられている。即ち、集音部 112 と集音部 113 とは、互いに逆側の方向を向くように設けられていることとなる。

[0019] このような構成に基づき、例えば、情報処理装置 10 は、集音部（例えば

、集音部 111～113）により集音されたユーザの音声（音響情報）を、音声認識技術や自然言語処理技術に基づく解析を施すことで、ユーザが発話した内容を認識してもよい。これにより、情報処理装置 10 は、例えば、ユーザからの指示内容を認識し、認識結果に応じて各種処理（アプリケーション）を実行することが可能となる。

[0020] また、他の一例として、情報処理装置 10 は、所謂通話機能を備えていてもよい。この場合には、情報処理装置 10 は、集音部（例えば、集音部 111～113）により集音された音声を、通話の相手である他の情報処理装置に転送してもよい。

[0021] 一方で、図 1 に示すような所謂ウェアラブルデバイスのように、ユーザが携帯可能に構成された情報処理装置 10 は、例えば、屋外で利用される場合等のように利用シーンが多岐にわたり、当該情報処理装置 10 の周囲の環境が動的に変化する状況が想定される。このような状況下においては、例えば、風切音、振動に伴う雑音、及び装置の装着に伴う衣擦れ等のようなランダムに発生する雑音が、情報処理装置 10 の集音部により集音される場合がある。

[0022] そこで、本開示では、雑音がランダムに発生するような環境下においても、目的音をより好適な態様で集音することが可能な仕組みの一例として、各集音部の設置位置や、当該集音部による集音結果に基づく信号処理の一例について詳細に説明する。

[0023] <1. 2. 集音部の設置位置の検討>

まず、本実施形態に係る情報処理装置 10 が、図 1 に示すように、ユーザの首に装着されるウェアラブルデバイスとして構成される場合を例に、当該ユーザの音声をより好適な態様で集音することが可能な集音部の設置位置に関する検討の結果について説明する。より具体的には、所謂風切音を雑音として想定し、複数箇所に集音部が設置された情報処理装置 10 に対して、互いに異なる角度から風を当てた場合における、当該集音部それぞれによる風切音の観測結果の一例について説明する。

[0024] 例えば、図3は、風切音の影響を観測するための観測環境の一例について説明するための説明図である。本観測では、図3に示すように、ユーザの胸より上の部位を模したダミー人形U1の首に情報処理装置10を装着し、当該ダミー人形U1の正面にサーキュレーターU2を配置する。そして、ダミー人形U1の鉛直方向を軸として、当該ダミー人形U1を、0度～360度の範囲内において10度刻みで回転させることで、情報処理装置10に対してサーキュレーターU2からの風が到来する角度を変更し、各集音部により集音される風切音のレベルを観測した。

[0025] 図4は、今回の観測において、情報処理装置10に設けられた複数の集音部それぞれの設置位置の一例を示している。具体的には、図4に示す例では、情報処理装置10に対して、集音部M1～M6が設置されている。情報処理装置10に付されたマーカは、集音部M1～M6それぞれが設置された位置を模式的に示している。なお、矢印が付されたマーカにおいて、当該矢印は、当該マーカに対応する集音部の向きを示している。また、矢印が付されていないマーカについては、当該マーカに対応する集音部（即ち、集音部M3及びM6）は、情報処理装置10の鉛直上方向（即ち、図面に対して奥行き方向の手前側）を向くように設定されているものとする。

[0026] 具体的には、集音部M1は、図1を参照して説明した情報処理装置10における集音部111に相当する。即ち、集音部M1は、情報処理装置10がユーザに装着されたときに、当該ユーザの口の近傍に相当する位置に、当該ユーザの前方に向けて突出するように設けられた凸部の先端に設けられている。また、集音部M5は、図1を参照して説明した情報処理装置10における集音部112に相当する。即ち、集音部M5は、情報処理装置10がユーザに装着されたときに、当該ユーザの左側（図3における略270度の方向）に相当する位置に、当該情報処理装置10の筐体101の外側に、当該筐体101の外側（換言すると、図3における略270度の方向）を向くように設けられている。

[0027] また、集音部M2～M4、及びM6は、情報処理装置10がユーザに装着

されたときに、当該ユーザの右前方（換言すると、図3における略45度の方向）の領域に相当する位置に設けられている。このとき、集音部M2は、情報処理装置10の筐体101とユーザの首との間に介在し、当該筐体101の内側を向くように設置されている。また、集音部M4は、情報処理装置10の筐体101の外側に、当該筐体101の外側（換言すると、図3における略45度の方向）を向くように設けられている。なお、集音部M3及びM6については、前述したように、鉛直上方向を向くように設けられている。

[0028] また、図5は、情報処理装置10に対して、互いに異なる角度から風を当てた場合における、当該集音部それぞれによる風切音の観測結果の一例について説明するための説明図である。即ち、図5は、図3を参照して説明した観測環境における、図4を参照して説明した集音部M1～M6それぞれによる風切音の集音結果の一例を示している。なお、図5に示した、集音部M1～M6それぞれの集音結果を示すグラフにおいて、円周方向に記載された数値は、サーキュレーターU2からの風が到来する方向を示している。また、グラフの半径方向に沿って記載された数値は、対応する集音部により集音された音響のレベル（即ち、当該集音部の観測レベル）を示している。即ち、図5に示す集音部M1～M6それぞれの集音結果を示すグラフにおいて、観測レベルが小さいほど（換言すると、観測値がグラフの内側に位置するほど）、風切音（即ち、雑音）の影響が小さいことを意味している。

[0029] ここで、集音部M1の観測結果に着目すると、特に、ユーザの正面（即ち、0度の方向）から風が到来する状況下において風切音の影響が小さいことがわかる。また、集音部M1については、正面以外の方向から風が到来する場合においても、他の集音部に比べて、風切音の影響が小さいことがわかる。

[0030] このことから、例えば、図1に示す集音部111のように、流線形の凸部の先端（または、当該先端の近傍）に当該凸部が突出する方向に向くように集音部を設けることで、風切音のようなランダムに発生する雑音の影響を小

さくすることが可能となるものと推定される。

[0031] また、集音部M 5 及びM 6 の観測結果に着目すると、当該集音部に対してユーザの首側から風が到来する場合において、風切音の影響が小さいことがわかる。これは、ユーザの首や頭部により風が遮蔽されることで、風切音の影響が小さくなったものと推定される。

[0032] このことから、例えば、図1 に示す集音部1 1 2 及び1 1 3 のように、情報処理装置1 0 が装着されるユーザの部位（例えば、首や頭部）を、風等に対する遮蔽物として利用可能な集音部を設けることで、他の集音部（例えば、図1 に示す集音部1 1 1 ）の特性を補うことが可能であるものと推定される。

[0033] 以上、図3 ～図5 を参照して、本実施形態に係る情報処理装置1 0 が、ユーザの首に装着されるウェアラブルデバイスとして構成される場合を例に、当該ユーザの音声をより好適な態様で集音する（即ち、風切音等の雑音の影響をより小さくする）ことが可能な集音部の設置位置に関する検討の結果について説明した。

[0034] < 1. 3. 機能構成 >

次いで、図6 を参照して、本実施形態に係る情報処理装置1 0 の機能構成の一例について、特に、情報処理装置1 0 が、複数の集音部それぞれの集音結果に基づき目的音（例えば、ユーザの音声）を取得する処理に着目して説明する。図6 は、本実施形態に係る情報処理装置1 0 の機能構成の一例を示したブロック図である。

[0035] 図6 に示すように、情報処理装置1 0 は、複数の集音部1 1 1 ～1 1 M（Mは正の整数）と、周波数分解部1 3 と、チャンネルパワー推定部1 5 と、フィルタ推定部1 6 と、フィルタ処理部1 7 と、周波数合成部1 8 とを含む。なお、以降の説明では、集音部1 1 1 ～1 1 Mを特に区別しない場合には、「集音部1 1 0」と称する場合がある。また、集音部1 1 0 の数（即ち、M）は、複数であれば特に限定はされないが、3 以上であることがより望ましい。

- [0036] 集音部 110 は、所謂マイクロフォンのように、外部環境の音響（即ち、外部環境を伝搬して到達する音響）を集音するための集音デバイスとして構成される。なお、ユーザからの音声入力についても、集音部 110 により集音されることで、情報処理装置 10 に取り込まれることとなる。また、集音部 110 は、例えば、マイクロフォンアレイのように、複数の集音デバイスを含んでもよい。集音部 110 は、外部環境の音響の集音結果に基づく音響信号を周波数分解部 13 に出力する。なお、集音部 110 から出力される音響信号は、例えば、アンプ等によりゲインが調整され、A/D 変換によりアナログ信号からデジタル信号に変換されたうえで、周波数分解部 13 に入力されてもよい。なお、以降の説明では、集音部 110 のチャンネル番号を m ($1 \leq m \leq M$)、離散時間を n とした場合に、当該集音部 110 から出力される音響信号を $x_m(n)$ で表すものとする。
- [0037] 周波数分解部 13 は、集音部 110 から出力される音響信号 $x_m(n)$ を周波数成分に分解して出力するための構成である。具体的には、周波数分解部 13 は、取得した音響信号 $x_m(n)$ に対して、フレーム分割、所定の窓関数の適用、及び、時間一周波数変換（例えば、FFT (Fast Fourier Transform)、DFT (Discrete Fourier Transform) 等）等の処理を施すことで、当該音響信号 $x_m(n)$ を周波数成分に分解する。なお、以降の説明では、音響信号 $x_m(n)$ の周波数成分を、 $X_m(i, k)$ と記載する場合がある。ここで、 i はフレーム番号を示し、 k は離散周波数番号を示すものとする。そして、周波数分解部 13 は、取得した音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を、後段に位置するフィルタ処理部 17 と、チャンネルパワー推定部 15 とのそれぞれに出力する。これにより、集音部 111 ~ 11M それぞれについて、音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ が、フィルタ処理部 17 と、チャンネルパワー推定部 15 とのそれぞれに出力されることとなる。
- [0038] チャンネルパワー推定部 15 は、周波数分解部 13 から集音部 110 ごとに（即ち、集音部 111 ~ 11M それぞれについて）音響信号 $x_m(n)$ の各周

波数成分 $X_m(i, k)$ を取得する。次いで、チャンネルパワー推定部 15 は、各集音部 110 それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ に基づき、周波数ごとに、各集音部 110 のパワースペクトルを推定する。ここで、 m 番目の集音部 110（即ち、集音部 11 m ）における、 i フレーム、周波数 k に対応するパワースペクトルを $P_m(i, k)$ とした場合に、パワースペクトル $P_m(i, k)$ は、以下に（式 1）として示す計算式で表される。なお、以下に示す（式 1）において、 $X_m^*(i, k)$ は、 $X_m(i, k)$ の共役複素数を示している。また、（式 1）において、 r は、急激なパワースペクトルの変化を抑制するためのフレーム方向の平滑化係数を表すものとする（ $0 \leq r < 1$ ）。

[0039] [数1]

$$P_m(i, k) = r \cdot P_m(i-1, k) + (1-r) \cdot X_m(i, k) \cdot X_m^*(i, k)$$

・・・（式 1）

[0040] そして、チャンネルパワー推定部 15 は、周波数ごとに、各集音部 110 のパワースペクトル $P_m(i, k)$ の推定結果をフィルタ推定部 16 に出力する。

[0041] フィルタ推定部 16 は、チャンネルパワー推定部 15 から出力される、周波数ごとの、各集音部 110 のパワースペクトル $P_m(i, k)$ の推定結果に基づき、後述するフィルタ処理部 17 が、フィルタリング処理を実行するためのフィルタ係数を算出する。

[0042] 具体的には、フィルタ推定部 16 は、周波数ごとに、チャンネルパワー推定部 15 から取得した各集音部 110 のパワースペクトル $P_m(i, k)$ の推定結果に基づき、以下に（式 2）として示す行列 $R(i, k)$ を生成する。

[0043]

[数2]

$$R(i, k) = \begin{pmatrix} P_1(i, k) & 0 & 0 & 0 \\ 0 & P_2(i, k) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & P_M(i, k) \end{pmatrix}$$

・・・ (式2)

[0044] また、フィルタ推定部16は、各集音部110それぞれについて、周波数ごとに、当該集音部110と目的音の音源（例えば、ユーザの口元等）との間の距離に基づき、当該集音部110までの減衰及び遅延特性を示すアレイマニフォールドベクトル $a(k)$ を算出する。なお、目的音の音源と各集音部110との間の距離については、情報処理装置10がユーザに装着されたときに、情報処理装置10（ひいては、当該情報処理装置10に設けられた各集音部110）と当該音源との間の相対的な位置関係に基づき、あらかじめ特定することが可能である。

[0045] ここで、アレイマニフォールドベクトル $a(k)$ は、以下に（式3）及び（式4）として示す計算式で表される。なお、以下に示す計算式において、 d_m は、目的音の音源（例えば、口元）と、 m 番目の集音部110（即ち、集音部11 m ）との間の距離を示している。また、 g_m は、目的音が集音部11 m に到達するまでの減衰量を示している。また、 ω_k は、離散周波数番号 k に対応する角周波数を示している。また、 C は、音速を示している。また、上付き文字の T が付された行列は、当該行列の転置を示すものとする。なお、以降の説明では、上付き文字の T が付された行列を、「転置ベクトル行列」と称する場合がある。

[0046]

[数3]

$$a_m(k) = g_m \cdot \exp\left(-j \frac{\omega_k \cdot d_m}{c}\right)$$

・・・(式3)

$$a(k) = [a_1(k), a_2(k), \dots, a_M(k)]^T$$

・・・(式4)

[0047] そして、フィルタ推定部16は、生成した行列 $R(i, k)$ と、算出したアレイマニフォールドベクトル $a(k)$ と、以下に(式5)として示す条件とに基づき、後述するフィルタ処理部17が、フィルタリング処理を実行するためのフィルタ係数 $w(i, k)$ を算出する。ここで、上付き文字の H が付された行列は、当該行列の複素共役転置を示すものとする。なお、以降の説明では、上付き文字の H が付された行列を、「複素共役転置ベクトル行列」と称する場合がある。

[0048] [数4]

$$\min w(i, k)^H R(i, k) w(i, k) \quad \text{subject to} \quad w^H(i, k) a(k) = 1$$

・・・(式5)

[0049] 周波数ごとのフィルタ係数 $w(i, k)$ は、以下に(式6)として示す計算式で表される。なお、 i はフレーム番号を示し、 k は離散周波数番号を示すものとする。

[0050] [数5]

$$w(i, k) = \frac{R(i, k)^{-1} a(k)}{a^H(k) R(i, k)^{-1} a(k)} = [w_1(i, k), w_2(i, k), \dots, w_M(i, k)]^T$$

・・・(式6)

[0051] なお、上記（式 6）として示したフィルタ係数 $w(i, k)$ は、上記（式 5）に示すように、目的音の音源（例えば、口元）から到来する成分 $a(k)$ の利得を 1 に保ち、かつ、ノイズ成分（例えば、風切音等）を最小にする係数となっている。そして、フィルタ推定部 16 は、周波数ごとに算出したフィルタ係数 $w(i, k)$ を、フィルタ処理部 17 に出力する。

[0052] フィルタ処理部 17 は、周波数分解部 13 から集音部 110 ごとに（即ち、集音部 111 ~ 11M それぞれについて）音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を取得する。また、フィルタ処理部 17 は、フィルタ推定部 16 から周波数ごとに算出されたフィルタ係数 $w(i, k)$ を取得する。フィルタ処理部 17 は、集音部 110 ごとの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を入力信号として、取得した周波数ごとのフィルタ係数 $w(i, k)$ に基づくフィルタリング処理を施すことで、出力信号 $Y(i, k)$ を生成する。

[0053] 具体的には、フィルタ処理部 17 は、集音部 110 ごとの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を入力信号として、取得した周波数ごとのフィルタ係数 $w(i, k)$ に基づき当該入力信号を重み付け加算することで、周波数ごとの出力信号 $Y(i, k)$ を生成する。例えば、出力信号 $Y(i, k)$ は、以下に（式 7）として示す計算式で表される。なお、 i はフレーム番号を示し、 k は離散周波数番号を示すものとする。

[0054] [数 6]

$$Y(i, k) = \sum_{m=1}^M w_m^*(i, k) \cdot X_m(i, k)$$

・・・（式 7）

[0055] そして、フィルタ処理部 17 は、周波数ごとに生成した出力信号 $Y(i, k)$ を、周波数合成部 18 に出力する。

[0056] 周波数合成部 18 は、フィルタ処理部 17 から周波数ごとに生成された出

力信号 $Y(i, k)$ を取得する。周波数合成部 18 は、取得した周波数ごとの出力信号 $Y(i, k)$ を合成することで、音響信号 $y(n)$ を生成する。即ち、周波数合成部 18 は、前述した周波数分解部 13 とは逆の処理を実行することとなる。具体的には、周波数合成部 18 は、周波数ごとの出力信号 $Y(i, k)$ に対して、周波数－時間変換（例えば、IFFT (Inverse FFT)、IDFT (Inverse DFT) 等）、所定の窓関数の適用、及びフレーム合成等の処理を施すことで、当該周波数ごとの出力信号 $Y(i, k)$ が合成された音響信号 $y(n)$ を生成する。

[0057] 例えば、図 7 は、本実施形態に係る情報処理装置 10 が、複数の集音部 110 それぞれの集音結果に基づき目的音を取得する処理の一例を示している。図 7 に示す例では、複数の集音部 110 として、集音部 111～114 の 4 つのマイクロフォンを使用した場合の一例を示している。即ち、図 7 に示す例では、集音部 111～114 それぞれによる集音結果の一例（即ち、集音された音響信号）と、情報処理装置 10 による信号処理により、集音部 111～114 それぞれの集音結果が合成された音響信号（合成音）の一例を示している。

[0058] 前述したように、複数の集音部 110 それぞれによる集音結果（より具体的には、音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ ）を合成するためのフィルタリング処理の係数 $w(i, k)$ は、目的音の音源（例えば、口元）から到来する成分 $a(k)$ の利得を 1 に保ち、かつ、ノイズ成分（例えば、風切音等）を最小にする特性を有する。このような構成により、ノイズ成分のレベルがより小さい集音部 110（換言するとノイズ成分の影響がより小さい集音部 110）の入力がより優先されるように重み付けされて、各集音部 110 の集音結果が合成される。このような処理により、風切音のような雑音がランダムに発生するような環境下においても、当該雑音の影響を抑圧し、目的音をより好適な態様で集音することが可能となる。

[0059] なお、上記に説明したように本実施形態に係る情報処理装置 10 は、複数の集音部 110 それぞれの集音結果から目的音を合成する構成となっており

、複数の集音部 110 の中から集音結果を取得する集音部 110 を単に切り替える構成とは異なる。より具体的には、集音結果を取得する集音部 110 を単に切り替えるような構成の場合には、当該切り替えの前後で音響信号に劣化が生じる場合があり、特に、風切音等の雑音が到来する方向が動的に変化するような状況下では当該音響信号の劣化がより顕在化しやすい傾向にある。これに対して、本実施形態に係る情報処理装置 10 は、上述した信号処理により目的音が合成されるため、風切音等の雑音が到来する方向が動的に変化するような状況下においても音響信号の劣化が発生せずに、より自然なかたちで目的音を取得することが可能となる。

[0060] なお、上記に説明した各集音部 110 の集音結果に対する信号処理はあくまで一例であり、ノイズ成分のレベルがより小さい集音部 110 の入力により優先されるように重み付けされて、各集音部 110 の集音結果を合成することが可能であれば、その内容は特に限定されない。

[0061] そして、周波数合成部 18 は、生成した音響信号 $y(n)$ を目的音の集音結果として出力する。周波数合成部 18 から出力される音響信号 $y(n)$ は、例えば、情報処理装置 10 により実行される各種処理（例えば、音声認識や音声通話等）に利用されることとなる。

[0062] なお、図 6 に示す構成はあくまで一例であり、上記に説明した各種処理が実現可能であれば、情報処理装置 10 の構成は、必ずしも図 6 に示す例には限定されない。例えば、図 6 に示す例では、集音部 111 ~ 11m それぞれについて周波数分解部 13 が設けられているが、複数の集音部 110 それぞれから出力される音響信号を、1つの周波数分解部 13 が処理する構成としてもよい。また、一部の構成が、情報処理装置 10 に対して外付けされてもよい。具体的な一例として、複数の集音部 110 のうち、少なくとも一部が、情報処理装置 10 に対して着脱可能に構成されていてもよい。

[0063] 以上、図 6 及び図 7 を参照して、本実施形態に係る情報処理装置 10 の機能構成の一例について、特に、情報処理装置 10 が、複数の集音部それぞれの集音結果に基づき目的音を取得する処理に着目して説明した。

[0064] < 1. 4. 処理 >

次に、図 8 を参照して、本実施形態に係る情報処理装置 10 の一連の処理の流れの一例について、特に、情報処理装置 10 が、複数の集音部それぞれの集音結果に基づき目的音（例えば、ユーザの音声）を取得する処理に着目して説明する。図 8 は、本実施形態に係る情報処理装置 10 の一連の処理の流れの一例を示したフローチャートである。

[0065] (ステップ S 101)

外部環境の音響は、複数の集音部 110 により集音されることで、情報処理装置 10 に取り込まれる。集音部 110 は、集音結果に基づく音響信号（アナログ信号）のゲインの調整し、A/D 変換によりアナログ信号からデジタル信号に変換したうえで、変換後の音響信号（デジタル信号） $x_m(n)$ を周波数分解部 13 に出力する。

[0066] (ステップ S 103)

周波数分解部 13 は、集音部 110 から出力される音響信号 $x_m(n)$ に対して、フレーム分割、所定の窓関数の適用、及び、時間一周波数変換等の処理を施すことで、当該音響信号 $x_m(n)$ を周波数成分に分解する。そして、周波数分解部 13 は、音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を、後段に位置するフィルタ処理部 17 と、チャンネルパワー推定部 15 とのそれぞれに出力する。これにより、複数の集音部 110 それぞれについて、音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ が、フィルタ処理部 17 と、チャンネルパワー推定部 15 とのそれぞれに出力されることとなる。

[0067] (ステップ S 105)

チャンネルパワー推定部 15 は、周波数分解部 13 から集音部 110 ごとに音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を取得する。次いで、チャンネルパワー推定部 15 は、各集音部 110 それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ に基づき、周波数ごとに、各集音部 110 のパワースペクトルを推定する。そして、チャンネルパワー推定部 15 は、周波数ごとに、各集音部 110 のパワースペクトル $P_m(i, k)$ の推定結果を

フィルタ推定部 16 に出力する。

[0068] (ステップ S 107)

フィルタ推定部 16 は、チャンネルパワー推定部 15 から出力される、周波数ごとの、各集音部 110 のパワースペクトル $P_m(i, k)$ の推定結果に基づき、後述するフィルタ処理部 17 が、フィルタリング処理を実行するためのフィルタ係数 $w(i, k)$ を算出する。

[0069] 具体的には、フィルタ推定部 16 は、各集音部 110 のパワースペクトル $P_m(i, k)$ に基づき、行列 $R(i, k)$ を生成する。また、フィルタ推定部 16 は、各集音部 110 それぞれについて、周波数ごとに、当該集音部 110 と目的音の音源との間の距離に基づき、当該集音部 110 までの減衰及び遅延特性を示すアレイマニフォールドベクトル $a(k)$ を算出する。そして、フィルタ推定部 16 は、生成した行列 $R(i, k)$ と、算出したアレイマニフォールドベクトル $a(k)$ とに基づき、フィルタ係数 $w(i, k)$ を算出し、当該フィルタ係数 $w(i, k)$ をフィルタ処理部 17 に出力する。

[0070] (ステップ S 109)

フィルタ処理部 17 は、周波数分解部 13 から集音部 110 ごとに音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を取得する。また、フィルタ処理部 17 は、フィルタ推定部 16 から周波数ごとに算出されたフィルタ係数 $w(i, k)$ を取得する。フィルタ処理部 17 は、集音部 110 ごとの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を入力信号として、取得した周波数ごとのフィルタ係数 $w(i, k)$ に基づき当該入力信号を重み付け加算することで、周波数ごとの出力信号 $Y(i, k)$ を生成する。そして、フィルタ処理部 17 は、周波数ごとに生成した出力信号 $Y(i, k)$ を、周波数合成部 18 に出力する。

[0071] (ステップ S 111)

周波数合成部 18 は、フィルタ処理部 17 から出力される周波数ごとの出力信号 $Y(i, k)$ に対して、周波数－時間変換、所定の窓関数の適用、及びフレーム合成等の処理を施すことで、当該周波数ごとの出力信号 $Y(i,$

k) を合成する。これにより、各集音部 110 による集音結果が合成された音響信号 $y(n)$ が生成される。なお、周波数合成部 18 により生成された音響信号 $y(n)$ は、集音結果として、情報処理装置 10 により実行される各種処理（例えば、音声認識や音声通話等）に利用されることとなる。

[0072] 以上、図 8 を参照して、本実施形態に係る情報処理装置 10 の一連の処理の流れの一例について、特に、情報処理装置 10 が、複数の集音部それぞれの集音結果に基づき目的音を取得する処理に着目して説明した。

[0073] <1. 5. 実施例>

次に、実施例として、本実施形態に係る情報処理装置 10 の他の一態様について説明する。

[0074] <1. 5. 1. 実施例 1：首に装着されるウェアラブルデバイスの一例>

まず、実施例 1 として、図 9～図 11 を参照して、図 1 に示した所謂ネットワークバンド型のウェアラブルデバイスのように、ユーザの首に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例について説明する。

[0075] 例えば、図 9 は、実施例 1 に係る情報処理装置の一例について説明するための説明図であり、ユーザの首に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例を示している。なお、本説明では、図 9 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 10a」と称する場合がある。

[0076] 図 9 に示すように、情報処理装置 10a は、集音部 111～114 を備える。集音部 111～113 は、図 1 を参照して前述した情報処理装置 10 における集音部 111～113 に対応している。また、集音部 114 は、情報処理装置 10a がユーザの首に装着された場合に、当該ユーザの後方の位置に、当該ユーザの後方側を向くように設けられている。このような構成とすることで、例えば、ユーザの後方から到来する雑音の影響をより緩和するこ

とが可能となる。

[0077] また、情報処理装置 10a は、集音部 112～114 が設置される位置に、当該集音部 112～114 それぞれが向く方向に突出し流線形の形状を有する凸部が設けられ、各凸部の先端に集音部 112～114 のそれぞれが設けられている。このような構成とすることで、集音部 112～114 についても、集音部 111 と同様に、風切音等のような雑音の影響を緩和し、凸部が突出する方向（即ち、集音部が向いた方向）から到来する音響をより好適な態様で集音することが可能となる。

[0078] なお、凸部を設ける位置（即ち、集音部 110 を設ける位置）については、特に限定はされない。そのため、例えば、ドライバ等の各種回路やバッテリー等が設けられることで筐体 101 にふくらみが生じ得る箇所に凸部を設け、当該凸部の先端（または、当該先端の近傍）に集音部 110 を設ける構成としてもよい。

[0079] また、図 10 は、実施例 1 に係る情報処理装置の他の一例について説明するための説明図であり、ユーザの首に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例を示している。なお、本説明では、図 10 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 10b」と称する場合がある。

[0080] 図 10 に示すように、情報処理装置 10b は、リング状の形状を成し、参照符号 19 で示された部分が開口可能に構成されている。なお、参照符号 19 で示された部分が開口することで互いに離間する端部間は着脱可能に構成されている。このような構成により、情報処理装置 10b は、リング状の部分の内面がユーザの首に当接するように（即ち、首に巻きつけるように）当該ユーザに装着される。

[0081] また、情報処理装置 10b は、集音部 115～118 が、リング状に形成された筐体の円周に沿って互いに異なる位置に、当該リングの外側（即ち、リングの中心とは逆側）を向くように設けられている。なお、情報処理装置

10bにおいては、集音部115～118が、前述した実施形態に係る情報処理装置10における、集音部110（例えば、図1に示す集音部111～113等）に相当する。

[0082] このような構成により、集音部115～118のそれぞれは、自身が向いている方向とは逆側から到来する雑音が、情報処理装置10bが装着されるユーザの部位（即ち、首）により遮蔽されるため、当該雑音の影響が緩和されることとなる。特に、図10に示す情報処理装置10bにおいては、集音部115～118のそれぞれが、図1に示した情報処理装置10に比べてよりユーザの首に近接するように支持されるため、風切音等の雑音（特に、ユーザの首側から到来する雑音）の影響がより緩和される。このことは、図5を参照して説明した、集音部M5及びM6（即ち、よりユーザの部位に近接する集音部）において、ユーザの部位側から到来する雑音の影響がより緩和されていることから明らかである。また、集音部115～118のそれぞれは、互いに異なる方向を向くように設けられているため、例えば、一部の集音部の集音結果に基づき他の集音部の特性を補うことも可能となる。

[0083] なお、図10に示す情報処理装置10bにおいても、筐体の少なくとも一部に流線形の形状の凸部を設け、当該凸部の先端（または、当該先端の近傍）に集音部110（例えば、集音部115～118のうちの少なくとも一部）を設けてもよい。

[0084] また、図11は、実施例1に係る情報処理装置の他の一例について説明するための説明図であり、所謂ネックレスのような形状を成すウェアラブルデバイスとして構成された情報処理装置の一例を示している。なお、本説明では、図11に示す情報処理装置を、前述した実施形態に係る情報処理装置10や、他の実施例に係る情報処理装置と区別するために、「情報処理装置10c」と称する場合がある。

[0085] 図11において、参照符号119は、前述した実施形態に係る情報処理装置10における集音部110の一例を示している。即ち、ネックレスのような形状を成す情報処理装置10cにおいては、例えば、所謂ペンダントに相

当する部分に、ユーザが装着したときに当該ユーザの前方を向くように流線型の凸部を設け、当該凸部の先端（または、当該先端の近傍）に集音部 119 を設ければよい。

[0086] なお、図 11 に示す例では、情報処理装置 10c に対して集音部 110 を 1 つ設けているが、集音部 110 が複数設けられていてもよい。また、情報処理装置 10c に対して集音部 110 を複数設ける場合には、当該複数の集音部 110 のそれぞれが、互いに異なる方向を向くように設けられているとよい。

[0087] 以上、実施例 1 として、図 9～図 11 を参照して、図 1 に示した所謂ネットワークバンド型のウェアラブルデバイスのように、ユーザの首に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例について説明した。

[0088] <1. 5. 2. 実施例 2：頭部に装着されるウェアラブルデバイスの一例>

次に、実施例 2 として、図 12～図 15 を参照して、頭部に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例について説明する。

[0089] 例えば、図 12 は、実施例 2 に係る情報処理装置の一例について説明するための説明図であり、ユーザの頭部に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例を示している。なお、本説明では、図 12 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 20a」と称する場合がある。

[0090] 図 12 に示すように、情報処理装置 20a は、ユーザの頭部に装着されることで、各種機能を実現するための回路等が組み込まれた筐体が、ユーザの耳の近傍に保持される。具体的な一例として、図 12 に示す例では、情報処理装置 20a は、ユーザの耳孔内に挿入されるイヤフォン部と、ユーザの耳に掛けることで筐体を支持するケーブル状の支持部材とを備える。情報処理

装置 20a は、このイヤフォン部とケーブル状の支持部材とにより、筐体がユーザの耳の近傍に保持される。

[0091] また、図 12 に示すように、情報処理装置 20a は、集音部 211 及び 212 を備える。なお、情報処理装置 20a おいては、集音部 211 及び 212 が、前述した実施形態に係る情報処理装置 10 における、集音部 110（例えば、図 1 に示す集音部 111～113 等）に相当する。

[0092] 具体的には、情報処理装置 20a は、ユーザの頭部に装着された状態において、当該ユーザの耳の近傍に保持される筐体のユーザの前方側に位置する端部に、当該前方側を向くように突出した流線形の形状を有する凸部を備える。そして、当該凸部の先端には、集音部 211 が、当該凸部が突出する方向（即ち、ユーザの前方）を向くように設けられている。また、情報処理装置 20a は、ユーザの頭部に装着された場合に、当該筐体の外側（即ち、当該頭部とは逆側）に位置する側面の少なくとも一部に、当該外側の方向（即ち、ユーザの横方向）を向くように集音部 212 が設けられている。また、情報処理装置 20a は、筐体の側面に対して、当該筐体の外側の方向に向けて突出した流線形の形状を有する凸部を備え、当該凸部の先端に集音部 212 が設けられていてもよい。

[0093] なお、図 12 に示す例では、ユーザの左耳の近傍に保持される筐体に着目して説明したが、当該ユーザの右耳の近傍に保持される筐体についても、左耳の近傍に保持される筐体と類似する構成をとり得る。具体的には、右耳側に保持される筐体には、集音部 212 に相当する構成のみが設けられていてもよいし、集音部 211 及び 212 に相当する構成が設けられていてもよい。

[0094] また、図 13 は、実施例 2 に係る情報処理装置の他の一例について説明するための説明図であり、ユーザの頭部に装着される所謂メガネ型のウェアラブルデバイスとして構成された情報処理装置の一例を示している。なお、本説明では、図 13 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理

装置 20b」と称する場合がある。

[0095] 図 13 に示すように、情報処理装置 20b は、集音部 213～215 を備える。なお、情報処理装置 20b においては、集音部 213～215 が、前述した実施形態に係る情報処理装置 10 における、集音部 110（例えば、図 1 に示す集音部 111～113 等）に相当する。

[0096] 例えば、情報処理装置 20b は、メガネのフロントに相当する部分の少なくとも一部に集音部 213 が設けられている。より具体的な一例として、情報処理装置 20b は、メガネのブリッジに相当する部分に、前方に向けて突出した流線形の形状を有する凸部を備え、当該凸部の先端には、集音部 213 が、当該凸部が突出する方向を向くように設けられている。また、他の一例として、参照符号 213' として示すように、メガネのフロントに相当する部分のうち、ブリッジに相当する部分とは異なる他の部分に当該凸部及び当該集音部が設けられていてもよい。

[0097] また、情報処理装置 20b は、メガネのテンプルに相当する部分の少なくとも一部に集音部 214 及び 215 が設けられている。なお、集音部 214 及び 215 は、例えば、情報処理装置 20b がユーザの頭部に装着された場合に、当該頭部とは逆側の方向（即ち、ユーザの横方向）に向くように設けられているとよい。

[0098] また、図 14 は、実施例 2 に係る情報処理装置の他の一例について説明するための説明図であり、ユーザの頭部に装着され得るウェアラブルデバイスとして構成された情報処理装置の他の一例を示している。なお、本説明では、図 14 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 20c」と称する場合がある。

[0099] 図 14 に示すように、情報処理装置 20c は、集音部 216～218 を備える。なお、情報処理装置 20c においては、集音部 216～218 が、前述した実施形態に係る情報処理装置 10 における、集音部 110（例えば、図 1 に示す集音部 111～113 等）に相当する。

[0100] より具体的には、メガネのフレームに相当する部分（例えば、フロント及びテンプル）の互いに異なる位置に、集音部 216～218 が互いに異なる方向を向くように設けられている。より具体的には、集音部 216～218 のそれぞれは、情報処理装置 20c がユーザの頭部に装着された場合に、当該頭部とは逆側の方向を向くように設けられる。

[0101] このような構成とすることで、集音部 216～218 のそれぞれは、自身が向いている方向とは逆側から到来する雑音が、ユーザの頭部により遮蔽されるため、当該雑音の影響が緩和される。また、集音部 216～218 のそれぞれは、互いに異なる方向を向くように設けられているため、例えば、一部の集音部の集音結果に基づき他の集音部の特性を補うことも可能となる。

[0102] また、図 15 は、実施例 2 に係る情報処理装置の他の一例について説明するための説明図であり、所謂ヘッドフォン等のようなオーバヘッド型のウェアラブルデバイスとして構成された情報処理装置の一例を示している。なお、本説明では、図 15 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 20d」と称する場合がある。

[0103] 図 15 に示す例では、情報処理装置 20d は、撮像部 25 と、集音部 219 とを備える。なお、情報処理装置 20d おいては、集音部 219 が、前述した実施形態に係る情報処理装置 10 における、集音部 110（例えば、図 1 に示す集音部 111～113 等）に相当する。

[0104] 具体的には、撮像部 25 は、情報処理装置 20d がユーザの頭部に装着された場合に、当該情報処理装置 20d の筐体の、当該ユーザの前方を画角内に収めることが可能な位置に設けられている。例えば、図 15 に示す例では、撮像部 25 は、情報処理装置 20d の筐体に、ユーザの前方を向くように設けられている。

[0105] また、情報処理装置 20d は、筐体の少なくとも一部に、ユーザの頭部に装着された状態においてユーザの前方側を向くように突出した流線形の形状を有する凸部を備え、当該凸部の先端には、集音部 219 が、当該凸部が突

出する方向を向くように設けられている。例えば、図 15 に示す例では、集音部 219 は、撮像部 25 の近傍に設けられている。また、他の一例として、参照符号 219' に示すように、情報処理装置 20d をユーザの頭部に保持するための保持部材の少なくとも一部に、ユーザの前方側を向くように突出した流線形の形状を有する凸部を設け、当該凸部の先端に、当該凸部が突出する方向を向くように集音部を設けてもよい。

[0106] 以上、実施例 2 として、図 12～図 15 を参照して、頭部に装着され得るウェアラブルデバイスとして構成された情報処理装置の一例について説明した。なお、上記に説明した例はあくまで一例であり、必ずしも上記に示す例には限定されない。具体的な一例として、所謂カチューシャ状の形状を有する頭部装着型のウェアラブルデバイスとして構成された情報処理装置に対して、前述した実施形態に係る情報処理装置 10 における集音部 110 に相当する構成を設けてもよい。

[0107] <1. 5. 3. 実施例 3：携帯型情報端末への適用例>

次に、実施例 3 として、図 16 及び図 17 を参照して、所謂スマートフォン等のような携帯型情報端末として構成された情報処理装置の一例について説明する。

[0108] 例えば、図 16 は、実施例 3 に係る情報処理装置の一例について説明するための説明図である。なお、本説明では、図 16 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 30」と称する場合がある。

[0109] 図 16 に示すように、情報処理装置 30 は、集音部 311～314 を含む。なお、情報処理装置 30 においては、集音部 311～314 が、前述した実施形態に係る情報処理装置 10 における、集音部 110（例えば、図 1 に示す集音部 111～113 等）に相当する。

[0110] 具体的には、情報処理装置 30 の筐体は、少なくとも一部に略矩形状の面 36 を有し、当該面 36 の角を含む所定の領域（即ち、当該角、または、当該角の近傍）に、当該筐体の外側を向くように流線型の形状を有する凸部が

形成されている。換言すると、情報処理装置 30 の筐体は、略平面状の面 36 と、当該面 36 の端部に沿って異なる方向を向くように形成された複数の側面 371 ～ 374 とを備え、互いに隣り合う側面が接続する部分を含む所定の領域に流線形の形状を有する凸部が形成されている。なお、面 36 は、例えば、ディスプレイ等の表示部が設けられる面に相当し得る。また、情報処理装置 30 の筐体の角自体が、当該凸部であってもよい。そして、集音部 311 ～ 314 のそれぞれは、当該凸部のうちのいずれかの先端（または、当該先端の近傍）に、情報処理装置 30 の筐体の外側を向くように設けられている。

[0111] また、図 17 は、変形例 3 に係る情報処理装置 30 の利用形態の一例について説明するための説明図であり、ユーザが情報処理装置 30 を利用して音声通話を行っている場合の一例を示している。

[0112] 図 17 に示すように、例えば、ユーザが情報処理装置 30 を自身の右耳の近傍に保持しながら音声通話を行う場合には、集音部 312 が当該ユーザの略前方を向くように当該情報処理装置 30 が保持されることとなる。このような構成により、例えば、ユーザが移動しながら音声通話を行っているような状況下においても、集音部 312 は、ユーザの移動により前方から到来する風に伴う風切音の影響を受けにくくなる。なお、ユーザが情報処理装置 30 を自身の左耳の近傍に保持しながら音声通話を行う場合も想定され得る。この場合には、集音部 311 がユーザの略前方を向くように当該情報処理装置 30 が保持されることとなり、当該集音部 311 が、ユーザの移動により前方から到来する風に伴う風切音の影響を受けにくくなる。即ち、情報処理装置 30 は、上記に説明した構成に基づき、ユーザの移動により前方から到来する風に伴う風切音の影響を緩和することが可能となる。

[0113] また、情報処理装置 30 は、集音部 311 ～ 314 が互いに異なる方向を向くように設けられている。このような構成により、情報処理装置 30 は、少なくとも一部の集音部の集音結果に基づき、他の集音部の特性を補うことも可能となる。

[0114] 以上、実施例3として、図16及び図17を参照して、所謂スマートフォン等のような携帯型情報端末として構成された情報処理装置の一例について説明した。

[0115] <1. 5. 4. 実施例4：時計型のウェアラブルデバイスへの適用例>
次に、実施例4として、図18及び図19を参照して、腕に装着され得る、所謂時計型のウェアラブルデバイスとして構成された情報処理装置の一例について説明する。

[0116] 例えば、図18は、実施例4に係る情報処理装置の一例について説明するための説明図である。なお、本説明では、図18に示す情報処理装置を、前述した実施形態に係る情報処理装置10や、他の実施例に係る情報処理装置と区別するために、「情報処理装置40a」と称する場合がある。

[0117] 図18に示すように、情報処理装置40aは、集音部411～415を含む。なお、情報処理装置30においては、集音部411～415が、前述した実施形態に係る情報処理装置10における、集音部110（例えば、図1に示す集音部111～113等）に相当する。

[0118] 具体的には、情報処理装置40aは、各種機能を実現するための回路等が組み込まれた筐体481と、当該筐体481をユーザの腕に支持するベルト状の支持部材482とを含む。筐体481は、前述した実施例3に係る情報処理装置30と同様に、少なくとも一部に略矩形状の面を有し、当該略矩形状の面の角を含む所定の領域に、当該筐体481の外側を向くように流線型の形状を有する凸部が形成されている。なお、当該略矩形状の面は、所謂時計における文字盤が設けられる側の面に相当する。そして、集音部411～414のそれぞれは、当該凸部のうちのいずれかの先端（または、当該先端の近傍）に、筐体481の外側を向くように設けられている。

[0119] また、支持部材482には、情報処理装置40aが腕に装着された状態において、当該腕を基準として、筐体481とは略対称となる位置に集音部415が、当該腕とは逆側の方向を向くように設けられている。

[0120] このような構成により、情報処理装置40aは、例えば、ユーザが、情報

処理装置 40a が装着された側の腕を振っているような状況下においても、集音部 411～414 のうちの少なくともいずれかが、腕が振られる方向と略等しい方向を向いている状態となる。そのため、情報処理装置 40a は、集音部 411～414 による集音結果により、腕の振りに伴う風切音の影響を緩和することが可能となる。また、情報処理装置 40a は、集音部 411～415 が互いに異なる方向を向くように設けられている。特に、集音部 415 については、当該集音部 415 が向いている方向とは逆側から到来する雑音が、情報処理装置 40a が装着された腕により遮蔽される。このような構成により、情報処理装置 40a は、集音部 411～415 のうち、少なくとも一部の集音部の集音結果に基づき、他の集音部の特性を補うことも可能となる。

[0121] また、図 19 は、実施例 4 に係る情報処理装置の他の一例について説明するための説明図である。なお、本説明では、図 19 に示す情報処理装置を、前述した実施形態に係る情報処理装置 10 や、他の実施例に係る情報処理装置と区別するために、「情報処理装置 40b」と称する場合がある。

[0122] 図 19 に示すように、情報処理装置 40b は、参照符号 483 で示された所謂時計のネジ部に相当する部分（以降では、「ネジ部 483」と称する）に、集音部 416 を備える。具体的には、ネジ部 483 を、流線形の形状を有するように形成することで、当該ネジ部 483 を、集音部 416 を設けるための凸部として利用してもよい。なお、情報処理装置 40b おいては、集音部 416 が、前述した実施形態に係る情報処理装置 10 における、集音部 110（例えば、集音部 111）に相当する。

[0123] 以上、実施例 4 として、図 18 及び図 19 を参照して、腕に装着され得る、所謂時計型のウェアラブルデバイスとして構成された情報処理装置の一例について説明した。

[0124] <1. 5. 5. 実施例 5：撮像装置への適用例>

次に、実施例 5 として、図 20 及び図 21 を参照して、動画像や静止画像を撮像可能な撮像装置として構成された情報処理装置の一例について説明す

る。

[0125] 例えば、図20は、実施例5に係る情報処理装置の一例について説明するための説明図である。なお、本説明では、図20に示す情報処理装置を、前述した実施形態に係る情報処理装置10や、他の実施例に係る情報処理装置と区別するために、「情報処理装置50」と称する場合がある。

[0126] 図20において、参照符号53は、動画像や静止画像等の画像を撮像するための撮像部に相当する。また、参照符号511及び512は、情報処理装置50に設けられた集音部の一例に相当する。なお、情報処理装置50においては、集音部511及び512が、前述した実施形態に係る情報処理装置10における、集音部110（例えば、図1に示す集音部111～113等）に相当する。

[0127] 具体的には、図20に示すように、情報処理装置50は、例えば、撮像部53を支持する筐体の、当該撮像部53が画像を撮像する方向（以降では、「撮像方向」と称する場合がある）を向いた面の一部に、当該撮像方向に突出した流線形の形状を有する凸部を備える。そして、当該凸部の先端（または、当該先端の近傍）には、撮像部53の撮像方向（換言すると、前方）を向くように集音部511が設けられている。

[0128] また、撮像部53の近傍（例えば、撮像部53のレンズの近傍）に、集音部512が設けられていてもよい。例えば、図21は、実施例5に係る情報処理装置50における撮像部53のレンズの近傍の概略的な構成の一例について説明するための説明図である。図21に示す例では、情報処理装置50は、撮像部53のレンズの近傍に、当該情報処理装置50の筐体の外側に向けて突出する凸部551が設けられている。また、凸部551は、撮像部53の撮像方向（即ち、前方）に向けて突出した流線型の形状を有する凸部553を備え、当該凸部553の先端（または、当該先端の近傍）に集音部513が設けられている。

[0129] このような構成により、情報処理装置50は、例えば、ユーザが移動しながら画像を撮像するような状況下においても、ユーザの移動により前方から

到来する風に伴う風切音の影響を緩和することが可能となる。

[0130] また、図20及び図21には図示していないが、情報処理装置50は、集音部511及び512とは異なる他の集音部を備えていてもよい。この場合には、当該他の集音部は、集音部511及び512とは異なる方向を向くように設けられているとよい。より具体的な一例として、例えば、情報処理装置50の筐体の、撮像部53の撮像方向とは逆側の面に、当該撮像方向とは逆側の方向（即ち、後方）に向くように当該他の集音部が設けられていてもよい。このような構成とすることで、例えば、他の集音部による集音結果に基づき、集音部511及び512の特性を補うことも可能となる。

[0131] 以上、実施例5として、図20及び図21を参照して、動画像や静止画像を撮像可能な撮像装置として構成された情報処理装置の一例について説明した。

[0132] <<2. 第2の実施形態>>

<2. 1. 概要>

続いて、本開示の第2の実施形態について説明する。前述した第1の実施形態に係る情報処理装置10では、複数の集音部それぞれの集音結果に基づき、観測レベル（即ち、集音された音響のレベル）のより小さい集音部の入力が優先されるようにフィルタリング処理を施すことで、風切音のようなランダムに発生する雑音の影響を低減していた。このような制御により、特に、風切音のようなランダムに発生する雑音の影響がより大きい場合に、より好適な態様で当該雑音の影響を緩和することが可能となる。

[0133] 一方で、上述した制御のように各集音部の集音結果をそのまま評価する場合には、音声等の目的音が主な成分として集音されるような状況下において、当該目的音をより高いレベルで集音した集音部の集音結果が使用されない場合がある。即ち、風切音等のランダムに発生する雑音の影響が小さい状況下においては、例えば、S/N比（signal-to-noise ratio）の小さい集音部の集音結果が優先的に使用される場合がある。

[0134] そこで、本実施形態では、前述した第1の実施形態と同様に風切音等のラ

ンダムに発生する雑音の抑圧効果を維持し、さらに、ランダムに発生する雑音の影響が小さい場合において、より好適な態様で目的音を取得することが可能な仕組みの一例について提案する。

[0135] < 2. 2. 機能構成 >

まず、図 22 を参照して、本実施形態に係る情報処理装置の機能構成の一例について説明する。図 22 は、本実施形態に係る情報処理装置の機能構成の一例を示したブロック図である。なお、以降の説明では、本実施形態に係る情報処理装置を、前述した第 1 の実施形態に係る情報処理装置 10（図 6 参照）と明示的に区別するために、「情報処理装置 60」と称する場合がある。

[0136] 図 22 に示すように、本実施形態に係る情報処理装置 60 は、複数の集音部 111 ~ 11M（M は正の整数）と、周波数分解部 13 と、チャンネルパワー推定部 65 と、フィルタ推定部 66 と、フィルタ処理部 17 と、周波数合成部 18 とを含む。なお、複数の集音部 111 ~ 11M（M は正の整数）と、周波数分解部 13 と、フィルタ処理部 17 と、周波数合成部 18 とは、前述した第 1 の実施形態に係る情報処理装置 10（図 6 参照）において、同様の符号が付された構成に相当する。即ち、本実施形態に係る情報処理装置 60 は、チャンネルパワー推定部 65 及びフィルタ推定部 66 の処理内容が、前述した第 1 の実施形態に係る情報処理装置 10 と異なる。そこで、以降では、本実施系形態に係る情報処理装置 60 の機能構成について、特に、前述した第 1 の実施形態に係る情報処理装置 10 と異なる部分に着目して説明し、当該情報処理装置 10 と同様の構成については詳細な説明は省略する。

[0137] 図 22 に示すように、チャンネルパワー推定部 65 は、入力パワー推定部 651 と、非相関成分パワー推定部 653 と、ランダムノイズパワー推定部 655 とを含む。

[0138] 入力パワー推定部 651 は、前述した第 1 の実施形態に係る情報処理装置 10 におけるチャンネルパワー推定部 15 に相当する。即ち、入力パワー推定部 651 は、各集音部 110 それぞれに対応する音響信号 $x_m(n)$ の各周波

数成分 $X_m(i, k)$ に基づき、周波数ごとに、各集音部110のパワースペクトルを推定する。そして、入力パワー推定部651は、周波数ごとに、各集音部110のパワースペクトル $P_m(i, k)$ の推定結果を、ランダムノイズパワー推定部655に出力する。

[0139] 非相関成分パワー推定部653は、フィルタ処理部17によりフィルタリング処理が施されることで生成された出力信号 $Y(i, k)$ のフィードバックを受ける。なお、当該出力信号 $Y(i, k)$ は、従前に集音された集音部110ごとの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ において雑音（ランダムノイズ）の影響が抑圧された音響であり、例えば、ユーザが発話した音声等のような目的音の、集音部110ごとの周波数成分に相当する。次いで、非相関成分パワー推定部653は、各集音部110それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ と、フィードバックされた出力信号 $Y(i, k)$ との間の相関性に基づき、当該出力信号 $Y(i, k)$ と非相関の成分のパワースペクトル $Q_m(i, k)$ を推定する。なお、周波数成分 $X_m(i, k)$ のうち、出力信号 $Y(i, k)$ と非相関の成分（以降では、単に「非相関成分」とも称する）が、当該周波数成分 $X_m(i, k)$ に含まれるランダムノイズ等の雑音の成分に相当する。また、非相関成分パワー推定部653による信号処理の詳細については別途後述する。そして、非相関成分パワー推定部653は、周波数ごとに、各集音部110のパワースペクトル $Q_m(i, k)$ の推定結果をランダムノイズパワー推定部655に出力する。

[0140] ランダムノイズパワー推定部655は、入力パワー推定部651から、周波数ごとに、各集音部110のパワースペクトル $P_m(i, k)$ の推定結果を取得する。また、ランダムノイズパワー推定部655は、非相関成分パワー推定部653から、周波数ごとに、各集音部110に対応する非相関成分のパワースペクトル $Q_m(i, k)$ の推定結果を取得する。そして、ランダムノイズパワー推定部655は、取得したパワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ それぞれの推定結果に基づき、フィルタ推定部66がフィルタ係

数 $w(i, k)$ を算出するための、周波数ごとの、各集音部110のパワースペクトル $W_m(i, k)$ を決定する。なお、ランダムノイズパワー推定部655による、パワースペクトル $W_m(i, k)$ の決定に係る処理の詳細については別途後述する。そして、ランダムノイズパワー推定部655は、周波数ごとに、各集音部110のパワースペクトル $W_m(i, k)$ を示す情報を、フィルタ推定部66に出力する。

[0141] フィルタ推定部66は、チャネルパワー推定部65から出力される、周波数ごとの、各集音部110のパワースペクトル $W_m(i, k)$ を示す情報に基づき、フィルタ処理部17が、フィルタリング処理を実行するためのフィルタ係数 $w(i, k)$ を算出する。なお、このときフィルタ推定部66は、(式2)として示した前述した行列 $R(i, k)$ を生成する際に、パワースペクトル $P_m(i, k)$ に替えて、パワースペクトル $W_m(i, k)$ を適用する点で、前述した第1の実施形態に係るフィルタ推定部16と異なる。

[0142] 一方で、以降の処理、即ち、(式3)～(式6)に基づき前述した、アレイマニフォールドベクトル $a(k)$ と、生成した生成された行列 $R(i, k)$ とに基づきフィルタ係数 $w(i, k)$ の算出に係る処理については、前述した第1の実施形態に係るフィルタ推定部16と同様である。そのため、当該処理の内容については、詳細な説明は省略する。

[0143] 以上のようにして、フィルタ推定部66は、取得した周波数ごとの、各集音部110のパワースペクトル $W_m(i, k)$ を示す情報に基づきフィルタ係数 $w(i, k)$ を算出し、算出したフィルタ係数 $w(i, k)$ を、フィルタ処理部17に出力する。なお、以降の処理については、前述した第1の実施形態に係る情報処理装置10(図6参照)と同様である。

[0144] 以上、図22を参照して、本実施形態に係る情報処理装置の機能構成の一例について説明した。

[0145] <2. 3. 非相関成分パワー推定部の詳細>

続いて、非相関成分パワー推定部653が、周波数ごとに、各集音部110に対応する非相関成分のパワースペクトル $Q_m(i, k)$ を算出する処理の

詳細について説明する。

[0146] まず、非相関成分パワー推定部653が、パワースペクトル $Q_m(i, k)$ を算出するための基本原理について説明する。マイクロフォン等の集音部に入力される音響（信号）には、例えば、ユーザの音声等のような目的音 S_m と、所謂背景ノイズ N_m と、風切音等のようなランダムノイズ W_m とが含まれる。即ち、集音部110ごとの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ は、目的音 S_m 、背景ノイズ N_m 、及びランダムノイズ W_m に基づき、以下に（式8）として示す関係式で表される。

[0147] [数7]

$$X_m(i, k) = S_m(i, k) + N_m(i, k) + W_m(i, k)$$

・・・（式8）

[0148] ここで、M個の集音部それぞれの入力される音響（信号）をまとめると、以下に（式9）として示す関係式で表される。

[0149] [数8]

$$X = S + N + W = a_k S_{org} + N + W$$

・・・（式9）

[0150] 上記に示した（式9）において、Sは、目的音 S_m をM個の集音部についてまとめたものである。同様に、Nは、背景ノイズ N_m をM個の集音部についてまとめたものであり、Wは、ランダムノイズ W_m をM個の集音部についてまとめたものである。なお、S、N、及びWは、それぞれベクトルとして示される。また、 S_{org} は、音源から出力された目的音そのものを示しており、スカラー値で示される。また、 a_k は、前述したアレイマニフォールドベクトル $a(k)$ に相当する。即ち、Sは、音源から出力された目的音 S_{org} が、集音部に到達するまでの空間を伝搬する際に生じる、信号の劣化や遅延等の影響を考慮した目的音の成分を示している。

[0151] ここで、風切音等のようなランダムノイズWの発生タイミングはランダム

であり、本開示に係る情報処理装置では、複数の集音部（特に、図1に示すように分散配置された集音部）間において近似的には相関性の無い信号として定義することが可能である。

[0152] このような特性に基づき、上記（式9）は、図23に示すようなベクトル間の関係として規定することが可能である。図23は、非相関成分パワー推定部653の処理の基本原理について説明するための説明図である。なお、図23に示す例では、ユーザが発話した音声を目的音として集音する場合について示している。また、図23に示すベクトル空間は、マニフォールドベクトル a_k に基づき規定される。

[0153] 図23において、 X は、集音部により集音された音響（即ち、入力信号）を示しており、（式9）に示した X に相当する。また、 Y は、理想的には、入力信号 X に対する目的音 S_{rg} の推定結果に基づく成分（即ち、ユーザの発話成分）に相当する。即ち、成分 Y は、入力信号 X に含まれる各成分のうち、ユーザの発話成分（もしくは、ユーザの発話成分と相関性を有する成分）を模式的に示している。これに対して、 Z は、入力信号 X に含まれる各成分のうち、ユーザの発話成分との相関性が小さい（もしくは、相関性の無い）成分に相当する。

[0154] なお、背景ノイズ N とランダムノイズ W とを全て抑圧することが可能であれば、成分 Z は、背景ノイズ N 及びランダムノイズ W の成分のみとなる。しかしながら、本開示に係る情報処理装置（例えば、図1参照）のように首回りに各集音部が配置される構成では、集音部間が比較的近傍に位置するため、背景ノイズ N は、当該集音部間において相関性を有する成分として観測される。そのため、成分 Y には、ユーザの発話成分 S に加えて、背景ノイズ N の成分が含まれる。一方で、風切音等のようなランダムノイズ W は、ユーザの発話成分との相関性が小さいため、成分 Z として示される。

[0155] 以上のような特性を利用し、非相関成分パワー推定部653は、出力信号 Y （即ち、ユーザの発話成分）のフィードバックを利用することで、当該出力信号 Y と相関性の小さい（もしくは、相関性の無い）成分を、ランダムノ

イズWの成分として抽出する。なお、以降の説明では、成分Zを、「非相関成分Z」とも称する。

[0156] 例えば、集音部110の数が4個の場合には、(式4)として前述した計算式に基づき、アレイマニフォールドベクトル a_k は、以下に(式10)として示す計算式で表される。

[0157] [数9]

$$a_k = (a_1(k), a_2(k), a_3(k), a_4(k))^T$$

・・・ (式10)

[0158] ここで、入力信号Xと、マニフォールドベクトル a_k との内積に基づき、当該入力信号Xをマニフォールドベクトル a_k に射影した成分を抽出することが可能である。このような特性から、マニフォールドベクトル a_k に直交する成分として、非相関成分Zを、以下に(式11)として示す計算式に基づき抽出することが可能である。

[0159] [数10]

$$Z = X - a_k \cdot (a_k^H \cdot a_k)^{-1} \cdot a_k^H \cdot X$$

・・・ (式11)

[0160] ここで、上記(式11)において、 $(a_k^H \cdot a_k)^{-1} \cdot a_k^H \cdot X$ として示された成分が、図23に示したユーザの発話成分Yに相当する。即ち、上記(式11)は、以下に(式12)として示す計算式で表すことが可能となる。

[0161] [数11]

$$Z = X - a_k \cdot Y$$

・・・ (式12)

[0162] ここで、上記(式12)における成分Yとして、フィードバックされた出

力信号 Y（即ち、フィルタ処理部 17 によるフィルタリング処理後の出力信号）を適用すると、上記（式 12）は、前述した（式 6）に基づき、以下に（式 13）として示す計算式で表すことが可能となる。

[0163] [数12]

$$Z = X - a_k \cdot \left(\frac{a^H(k)R(i, k)^{-1}}{a^H(k)R(i, k)^{-1}a_k} \cdot X \right)$$

・・・（式 13）

[0164] 以上のようにして算出された非相関成分 Z に基づき信号のパワーを算出し、時間平滑化を行うことで、非相関成分 Z のパワースペクトルを推定することが可能となる。ここで、m 番目の集音部 110（即ち、集音部 11m）における、i フレーム、周波数 k に対応する非相関成分 Z のパワースペクトル $Q_m(i, k)$ は、以下に（式 14）として示す計算式で表される。なお、以下に示す（式 14）において、 $Z_m^*(i, k)$ は、 $Z_m(i, k)$ の共役複素数を示している。また、（式 14）において、r は、急激なパワースペクトルの変化を抑制するためのフレーム方向の平滑化係数を表すものとする（ $0 \leq r < 1$ ）。

[0165] [数13]

$$Q_m(i, k) = r \cdot Q_m(i - 1, k) + (1 - r) \cdot Z_m(i, k) \cdot Z_m^*(i, k)$$

・・・（式 14）

[0166] 以上のようにして、非相関成分パワー推定部 653 は、非相関成分のパワースペクトル $Q_m(i, k)$ を算出する。

[0167] なお、非相関成分パワー推定部 653 は、パワースペクトル $Q_m(i, k)$ を推定する際に、2 以上の集音部 110 の集音結果を使用できれば、必ずしも全ての集音部 110 の集音結果を使用する必要はない。具体的な一例として、非相関成分パワー推定部 653 は、ユーザの頭部に対して後方に位置す

る集音部 1 1 0 のように、音声等の目的音を集音しにくい位置に設置された集音部 1 1 0 の集音結果については、パワースペクトル $Q_m(i, k)$ の推定に使用しなくてもよい。

[0168] 以上、非相関成分パワー推定部 6 5 3 が、周波数ごとに、各集音部 1 1 0 に対応する非相関成分のパワースペクトル $Q_m(i, k)$ を算出する処理の詳細について説明した。

[0169] < 2. 4. ランダムノイズパワー推定部の詳細 >

続いて、ランダムノイズパワー推定部 6 5 5 が、フィルタ係数 $w(i, k)$ の算出に用いられる、周波数ごとの、各集音部 1 1 0 のパワースペクトル $W_m(i, k)$ を決定する処理の詳細について説明する。

[0170] 前述したように、ランダムノイズパワー推定部 6 5 5 は、入力パワー推定部 6 5 1 から取得されるパワースペクトル $P_m(i, k)$ と、非相関成分パワー推定部 6 5 3 から取得される非相関成分のパワースペクトル $Q_m(i, k)$ とのそれぞれの推定結果に基づき、パワースペクトル $W_m(i, k)$ を決定する。

[0171] (パワースペクトル Q_m を適用するケース)

例えば、ランダムノイズパワー推定部 6 5 5 は、非相関成分のパワースペクトル $Q_m(i, k)$ の推定結果を、パワースペクトル $W_m(i, k)$ としてフィルタ推定部 6 6 に出力してもよい。なお、この場合には、チャンネルパワー推定部 6 5 は、入力パワー推定部 6 5 1 を含まなくてもよい。

[0172] (パワースペクトル P_m 及び Q_m を選択的に切り替えるケース)

また、他の一例として、ランダムノイズパワー推定部 6 5 5 は、所定の条件に基づき、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ それぞれの推定結果のうちのいずれかを選択的に、パワースペクトル $W_m(i, k)$ としてフィルタ推定部 6 6 に出力してもよい。

[0173] (パワースペクトル W_m を適応的に算出するケース)

また、他の一例として、ランダムノイズパワー推定部 6 5 5 は、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ それぞれの推定結果に基づき、パ

ワースペクトル $W_m(i, k)$ を適応的に算出してもよい。

[0174] 例えば、ランダムノイズパワー推定部655は、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ を入力として、以下に(式15)として示す計算式に基づき、目的音(音声等)とランダムノイズとの関係を考慮したパワースペクトル $\widetilde{W}_m$ を算出する。なお、「 $\widetilde{W}_m$ 」は、「 W_m 」の上にチルダが付された文字を示すものとする。また、以下に示す P_m 及び Q_m は、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ を一般化して記載したものである。

[0175] [数14]

$$\widetilde{W}_m = F(P_m, Q_m)$$

・・・(式15)

[0176] 例えば、以下に示す(式16)は、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ を入力として、目的音とランダムノイズとの関係を考慮したパワースペクトル $\widetilde{W}_m$ を算出するための関数 F の具体的な一例を示している。

[0177] [数15]

$$\widetilde{W}_m = F(P_m, Q_m) = \frac{Q_m}{P_m + Q_m} P_m$$

・・・(式16)

[0178] そして、ランダムノイズパワー推定部655は、上述した目的音とランダムノイズとの関係を考慮したパワースペクトル $\widetilde{W}_m$ に基づき、パワースペクトル W_m を、以下に(式17)として示す計算式に基づき算出する。なお、(式17)において、 r は、急激なパワースペクトルの変化を抑制するためのフレーム方向の平滑化係数を表している($0 \leq r < 1$)。即ち、ランダムノイズパワー推定部655は、以下に(式17)として示す計算式に基づき算出されるパワースペクトル W_m を、係数 r の設定に基づきフレーム間で平滑化してもよい。

[0179] [数16]

$$W_m = r \cdot W_m(i-1, k) + (1-r) \cdot \widetilde{W}_m$$

・・・ (式17)

[0180] ここで、(式16)に示すパワースペクトル P_m 、即ち、入力パワー推定部651によるパワースペクトル $P_m(i, k)$ の推定結果は、前述したように、集音部110により集音された音響のレベルに相当する。これに対して、(式16)に示すパワースペクトル Q_m 、即ち、非相関成分パワー推定部653によるパワースペクトル $Q_m(i, k)$ の推定結果は、風切音等のようなランダムノイズのレベルに相当する。即ち、(式16)に示した重み $Q_m / (P_m + Q_m)$ は、音声等の目的音と、風切音等のランダムノイズとの間の関係に基づき変化する。

[0181] 具体的には、ランダムノイズに対して目的音の信号レベルが十分に大きい場合には、パワースペクトル P_m の影響が支配的となり、重み $Q_m / (P_m + Q_m)$ はより小さくなる。即ち、この場合における当該重み $Q_m / (P_m + Q_m)$ は、対応するチャネル(即ち、集音部110)の集音結果の使用をより抑制する制御を示している。ここで、フィルタ係数 $w(i, k)$ の算出には、上記重み $Q_m / (P_m + Q_m)$ の逆数が適用される。そのため、ランダムノイズに対して目的音の信号レベルが十分に大きい場合には、該当するチャネルによる集音結果の使用がより優先されるように、フィルタ係数 $w(i, k)$ が算出されることとなる。

[0182] 一方で、風切音等のようなランダムノイズの影響がより大きい場合には、パワースペクトル Q_m の影響がより支配的となり、重み $Q_m / (P_m + Q_m)$ はより大きくなる。即ち、この場合における当該重み $Q_m / (P_m + Q_m)$ は、対応するチャネル(即ち、集音部110)の集音結果の使用をより優先する制御を示している。なお、前述したようにフィルタ係数 $w(i, k)$ の算出には、上記重み $Q_m / (P_m + Q_m)$ の逆数が適用される。そのため、ランダムノ

イズの影響がより大きい場合には、該当するチャネルによる集音結果の使用が抑制されるように、フィルタ係数 $w(i, k)$ が算出されることとなる。

[0183] 即ち、上述した制御により、風切音等のランダムノイズの影響が小さく、主に音声が集音されるような状況下では、音声のレベルがより高く集音された集音部 110 の集音結果が、より優先的に使用されて、フィルタ係数 $w(i, k)$ が算出される。これに対して、風切音等のランダムノイズの影響が大きい状況下では、前述した第 1 の実施形態と同様に、観測レベルのより小さい集音部 110 の集音結果が、より優先的に使用されて、フィルタ係数 $w(i, k)$ が算出されることとなる。このように、ランダムノイズパワー推定部 655 は、フィルタ係数 $w(i, k)$ が算出するためのパワースペクトル $W_m(i, k)$ を、音声等のような目的音と、風切音等のようなランダムノイズとの関係に応じて、適応的に算出することが可能となる。

[0184] そして、ランダムノイズパワー推定部 655 は、上記（式 17）に基づき算出したパワースペクトル $W_m(i, k)$ を、フィルタ推定部 66 に出力してもよい。

[0185] 以上、ランダムノイズパワー推定部 655 が、フィルタ係数 $w(i, k)$ の算出に用いられる、周波数ごとの、各集音部 110 のパワースペクトル $W_m(i, k)$ を決定する処理の詳細について説明した。なお、上記に説明した例は、あくまで一例であり、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ の少なくともいずれかの推定結果に基づき、パワースペクトル $W_m(i, k)$ を決定することが可能であれば、その内容は特に限定されない。

[0186] <2. 5. 評価>

以上説明したように、本実施形態に係る情報処理装置 60 は、複数の集音部 110 のうち少なくとも 2 以上の集音部 110 による集音結果と、フィルタ処理部 17 の出力信号 $Y(i, k)$ のフィードバックとに基づき、非相関成分のパワースペクトル $Q_m(i, k)$ を推定する。そして、情報処理装置 60 は、非相関成分のパワースペクトル $Q_m(i, k)$ の推定結果を、フィルタ係数 $w(i, k)$ の推定に利用する。このような構成により、情報処理装置

60は、前述した第1の実施形態と同様に風切音等のランダムに発生する雑音の抑圧効果を維持し、さらに、ランダムに発生する雑音の影響が小さい場合において、より好適な態様で目的音を取得することが可能となる。

[0187] なお、上記では、本実施形態に係る信号処理を、例えば、図1に示した所謂ネックバンド型のウェアラブルデバイスに適用する場合に着目して説明した。一方で、本実施形態に係る信号処理の適用先は、必ずしも図1に示す例のみに限定されない。具体的には、本実施系に係る信号処理は、複数の集音部を備える装置であれば適用することが可能である。なお、より好適には、複数の集音部は、目的音の音源（例えば、音声が発話される口元）からの距離が互いに異なるように配置されているとよい。また、より好適には、複数の集音部は、目的音の音源に対して互いに異なる方向に位置するように配置されているとよい。

[0188] <<第3の実施形態>>

<3. 1. 概要>

続いて、本開示の第3の実施形態として、本開示に係る技術を、所謂マルチチャネルウィナーフィルタ（MWF : Multi Channel Wiener Filter）に適用した場合の一例について説明する。

[0189] まず、本実施形態に係る情報処理装置の特徴をよりわかりやすくするために、マルチチャネルウィナーフィルタについて概要を説明する。マルチチャネルウィナーフィルタは、背景ノイズの抑圧等に利用される技術である。例えば、マルチチャネルウィナーフィルタを W_{mwf} とした場合に、 W_{mwf} は、以下に（式18）として示す計算式に基づき算出される。

[0190] [数17]

$$W_{mwf} = \arg \min_W E \left[\left| S - W^H X \right|^2 \right]$$

・・・（式18）

[0191] 上記（式18）において、Sは、音声等のような目的音を示しており、前

述した（式 9）に示した S_{rg} に相当する。また、 X は、集音部により集音された音響（即ち、入力信号）を示しており、前述した（式 9）に示した X に相当する。また、 W は、マルチチャネルウィナーフィルタにおいて、入力信号 X に対して適用されるフィルタを模式的に示したものである。即ち、マルチチャネルウィナーフィルタは、理想的には上記（式 18）に示したコスト関数を最小化するフィルタに基づいて元となる信号を復元する。

[0192] しかしながら、一般的には、目的音 S を独立して観測することは困難であるため、複数のチャネル（即ち、集音部）のうち一部のチャネルをリファレンスとして、信号劣化と、ノイズの抑圧量と、をトレードオフとした最小化問題に置き換える。この場合には、例えば、マルチチャネルウィナーフィルタ W_{mwf} は、以下に（式 19）として示す計算式で表される。

[0193] [数18]

$$\begin{aligned} W_{mwf} &= \arg \min_W E \left[|X_i - W^H X|^2 + \mu |W^H N|^2 \right] \\ &= (R_x + \mu R_n)^{-1} R_x e_i \end{aligned}$$

・・・（式 19）

[0194] 上記（式 19）において、 X_i は、リファレンスとしたチャネルによる集音結果に基づく出力信号を示している。なお、以降の説明においては、リファレンスとしたチャネルを「リファレンスマイク」とも称する。また、 N は、背景ノイズを示しており、前述した（式 9）に示した N に相当する。また、係数 μ は、目的音に対するノイズの抑圧量に応じて設定される係数（重み）であり、例えば、事前の実験等の結果に基づきあらかじめ設定される。また、上記（式 19）における R_x 及び R_n は、以下に（式 20）及び（式 21）として示す計算式で表される。また、（式 19）において、リファレンスマイクは e_i により特定される。 e_i は、 i 番目の値のみが 1 をとり他の値は 0 を示す M 次ベクトルであり、例えば、以下に示す（式 22）のように表される。

[0195] [数19]

$$R_x = E[XX^H]$$

・・・(式20)

$$R_n = E[NN^H]$$

・・・(式21)

$$e_i = [0, \dots, 1, \dots, 0]^T$$

・・・(式22)

・・・(式22)

[0196] なお、上記(式20)において、 X は、複数のチャネルそれぞれによる集音結果に基づく入力信号をベクトルで表したものである。即ち、 R_x は、入力信号 X そのものから算出される。また、(式21)において、 N は、目的音が集音されていない区間（以下、「休止区間」とも称する）における、複数のチャネルそれぞれによる集音結果に基づく入力信号（即ち、雑音に相当する信号）をベクトルで表したものである。そのため、 R_n は、例えば、目的音が集音される区間（以下、「目的音区間」とも称する）の検出結果に基づき特定される休止区間において算出される。なお、以降では、 R_x を「入力相関行列」とも称し、 R_n を「ノイズ相関行列」とも称する。

[0197] ここで、他の実施形態において前述したように、屋外で使用される場合等のような周囲の環境が動的に変化する状況下では、風切音、振動に伴う雑音、及び装置の装着に伴う衣擦れ等のようなランダムノイズが集音部により集音される場合がある。その一方で、従来のマルチチャネルウィナーフィルタにおいては、(式19)において e_i により固定的に決定される集音部（即ち、リファレンスマイク）に上記ランダムノイズが混入した場合には、当該ランダムノイズを抑制することが困難である。

[0198] そこで、本実施形態では、他の実施形態において上述した技術をマルチチャネルウィナーフィルタに対して適用することで、当該マルチチャネルウィナーフィルタにおけるランダムノイズの影響をより低減する仕組みの一例について提案する。

[0199] 具体的には、本実施形態に係る情報処理装置では、前述したパワースペクトル P_m （または、パワースペクトル W_m ）の推定結果に基づきランダムノイズの影響がより小さい集音部をリファレンスマイクとして適応的に選択することで、ランダムノイズの影響をより低減する。例えば、以下に（式 23）として示す計算式は、本実施形態に係る情報処理装置におけるマルチチャネルウィナーフィルタ W_{mwf} の導出に係る基本原理を示している。

[0200] [数20]

$$W_{mwf} = \arg \min_W E \left[|Y - W^H X|^2 + \mu |W^H N|^2 \right]$$

・・・（式 23）

[0201] 上記（式 23）を前述した（式 19）と比較するとわかるように、本実施形態に係る情報処理装置では、リファレンスマイクによる集音結果に基づく出力信号 X_i として、他の実施形態において前述したパワースペクトル P_m （または、パワースペクトル W_m ）の推定結果に基づく出力信号 Y を適用する。このような構成に基づき、本実施形態に係る情報処理装置は、ランダムノイズの影響がより顕著に表れる状況下（即ち、周囲の環境が動的に変化するような状況下）においても、ランダムノイズの影響のより小さい集音部をリファレンスマイクとして動的に選択する。これにより、本実施形態に係る情報処理装置に依れば、従来のマルチチャネルウィナーフィルタを適用した装置に比べて、ランダムノイズの影響をより低減し、ひいては、より好適な態様で目的音を抽出することが可能となる。なお、以降では、本実施形態に係る情報処理装置についてさらに詳しく説明する。

[0202] <3. 2. 機能構成>

図 24 を参照して、本実施形態に係る情報処理装置の機能構成の一例につ

いて説明する。図24は、本実施形態に係る情報処理装置の機能構成の一例を示したブロック図である。なお、以降の説明では、本実施形態に係る情報処理装置を、前述した他の実施形態に係る情報処理装置（例えば、情報処理装置10及び60）と明示的に区別するために、「情報処理装置70」と称する場合がある。

[0203] 図24に示すように、本実施形態に係る情報処理装置70は、複数の集音部111～11M（Mは正の整数）と、周波数分解部13と、チャンネルパワー推定部71と、フィルタ推定部72と、入力相関行列算出部73と、ノイズ相関行列算出部74と、マルチチャンネルウィナーフィルタ算出部75と、マルチチャンネルウィナーフィルタ処理部76と、周波数合成部77と、を含む。なお、複数の集音部111～11M（Mは正の整数）と、周波数分解部13とは、前述した第1の実施形態に係る情報処理装置10（図6参照）において、同様の符号が付された構成に相当する。即ち、本実施形態に係る情報処理装置70は、チャンネルパワー推定部71と、フィルタ推定部72と、入力相関行列算出部73と、ノイズ相関行列算出部74と、マルチチャンネルウィナーフィルタ算出部75と、マルチチャンネルウィナーフィルタ処理部76と、周波数合成部77との処理内容が、前述した第1の実施形態に係る情報処理装置10と異なる。そこで、以降では、本実施系形態に係る情報処理装置70の機能構成について、特に、前述した第1の実施形態に係る情報処理装置10と異なる部分に着目して説明し、当該情報処理装置10と同様の構成については詳細な説明は省略する。

[0204] （マルチチャンネルウィナーフィルタ処理部76）

マルチチャンネルウィナーフィルタ処理部76は、周波数分解部13から集音部110ごとに音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を取得する。また、マルチチャンネルウィナーフィルタ処理部76は、後述するマルチチャンネルウィナーフィルタ算出部75から、周波数ごとにマルチチャンネルウィナーフィルタ $W_{mwf}(i, k)$ の算出結果を取得する。なお、マルチチャンネルウィナーフィルタ $W_{mwf}(i, k)$ の算出方法については詳細を別途後述す

る。そして、マルチチャネルウィナーフィルタ処理部 76 は、集音部 110 などの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を入力信号として、当該マルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ に基づくフィルタリング処理を施すことで、出力信号 S を周波数ごとに生成する。例えば、出力信号 S は、以下に (式 24) として示す計算式で表される。なお、以下に示す (式 24) では、フレーム番号 i 及び離散周波数番号 k の記載を省略している。

[0205] [数21]

$$S = W_{mwf}^H X$$

・・・ (式 24)

[0206] そして、マルチチャネルウィナーフィルタ処理部 76 は、周波数ごとに生成した出力信号 S を、周波数合成部 77 に出力する。

[0207] (周波数合成部 77)

周波数合成部 77 は、マルチチャネルウィナーフィルタ処理部 76 から周波数ごとに生成された出力信号 S を取得する。周波数合成部 77 は、取得した周波数ごとの出力信号 S を合成することで音響信号を生成する。なお、周波数合成部 77 による処理は、前述した第 1 及び第 2 の実施形態に係る周波数合成部 18 が、周波数ごとの出力信号 $Y(i, k)$ を合成することで音響信号 $y(n)$ を生成する処理と同様のため、詳細な説明は省略する。

[0208] (チャネルパワー推定部 71 及びフィルタ推定部 72)

続いて、チャネルパワー推定部 71 及びフィルタ推定部 72 の構成について説明する。チャネルパワー推定部 71 及びフィルタ推定部 72 は、前述した第 1 の実施形態に係る情報処理装置 10 と同様の仕組みを適用して構成してもよいし、第 2 の実施形態に係る情報処理装置 60 と同様の仕組みを適用して構成してもよい。そこで、以下にそれぞれの場合に着目して説明する。

[0209] (チャネルパワー推定部 71 及びフィルタ推定部 72 の構成例 1)

まず、前述した第 1 の実施形態に係る情報処理装置 10 (図 6 参照) と同

様の仕組みを適用して、チャンネルパワー推定部 7 1 及びフィルタ推定部 7 2 を構成する場合について説明する。この場合には、チャンネルパワー推定部 7 1 及びフィルタ推定部 7 2 は、第 1 の実施形態に係るチャンネルパワー推定部 1 5 及びフィルタ推定部 1 6 に相当することとなる。

[0210] 具体的には、チャンネルパワー推定部 7 1 は、各集音部 1 1 0 それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ に基づき、周波数ごとに、各集音部 1 1 0 のパワースペクトル $P_m(i, k)$ を推定する。

[0211] また、フィルタ推定部 7 2 は、パワースペクトル $P_m(i, k)$ の推定結果に基づき、フィルタ係数 $w(i, k)$ を算出する。そして、フィルタ推定部 7 2 は、フィルタ係数 $w(i, k)$ の算出結果に基づきフィルタ G を周波数ごとに算出し、当該フィルタ G の算出結果を後述するマルチチャンネルウィナーフィルタ算出部 7 5 に出力する。

[0212] (チャンネルパワー推定部 7 1 及びフィルタ推定部 7 2 の構成例 2)

続いて、前述した第 2 の実施形態に係る情報処理装置 6 0 (図 2 2 参照) と同様の仕組みを適用して、チャンネルパワー推定部 7 1 及びフィルタ推定部 7 2 を構成する場合について説明する。この場合には、チャンネルパワー推定部 7 1 及びフィルタ推定部 7 2 は、第 2 の実施形態に係るチャンネルパワー推定部 6 5 及びフィルタ推定部 6 6 に相当することとなる。

[0213] 具体的には、チャンネルパワー推定部 7 1 は、各集音部 1 1 0 それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ に基づき、周波数ごとに、各集音部 1 1 0 のパワースペクトル $P_m(i, k)$ を推定する。

[0214] また、チャンネルパワー推定部 7 1 は、フィルタリング処理に基づき雑音 (特に、ランダムノイズ) の影響が抑圧された周波数ごとの音響信号のフィードバックを受ける。

[0215] 具体的な一例として、フィルタ推定部 7 2 は、マルチチャンネルウィナーフィルタ処理部 7 6 によるフィルタリング処理の結果として出力される周波数ごとの音響信号 S のフィードバックを受けてもよい。

[0216] また、他の一例として、第 2 の実施形態に係る情報処理装置 6 0 における

フィルタ処理部 17 に相当する構成を別途設けることで、フィルタ推定部 72 は、当該構成からノイズ成分が抑圧された周波数ごとの音響信号のフィードバックを受けてもよい。この場合には、フィルタ推定部 72 は、第 2 の実施形態に係る情報処理装置 60 における出力信号 $Y(i, k)$ に相当する音響信号のフィードバックを受けることとなる。

[0217] 上記フィードバックを受けると、チャンネルパワー推定部 71 は、各集音部 110 それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ と、フィードバックされた音響信号との間の相関性に基づき、当該音響信号と非相関の成分のパワースペクトル $Q_m(i, k)$ を推定する。

[0218] そして、チャンネルパワー推定部 71 は、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ それぞれの推定結果に基づき、フィルタ推定部 72 がフィルタ係数 $w(i, k)$ を算出するための、周波数ごとの、各集音部 110 のパワースペクトル $W_m(i, k)$ を決定すればよい。なお、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ それぞれの算出方法や、パワースペクトル $W_m(i, k)$ の決定方法については、前述した第 2 の実施形態に係るチャンネルパワー推定部 65 と同様のため詳細な説明は省略する。

[0219] また、フィルタ推定部 72 は、パワースペクトル $W_m(i, k)$ の推定結果に基づき、フィルタ係数 $w(i, k)$ を算出する。そして、フィルタ推定部 72 は、フィルタ係数 $w(i, k)$ の算出結果に基づきフィルタ G を周波数ごとに算出し、当該フィルタ G の算出結果を後述するマルチチャンネルウィナーフィルタ算出部 75 に出力すればよい。

[0220] なお、上述した通り、フィルタ G は、フィルタ係数 $w(i, k)$ に基づき算出されるため、前述した(式 2)～(式 6)を鑑みると、例えば、以下に(式 25)及び(式 26)として示す計算式で表すことが可能である。

[0221]

[数22]

$$G = \gamma R_w^{-1} a$$

・・・(式25)

$$\gamma = 1 / (a^H R_w^{-1} a)$$

・・・(式26)

[0222] なお、上記(式25)及び(式26)において、 R_w は、前述した(式6)における行列 $R(i, k)$ に相当する。即ち、第1の実施形態と同様の思想に基づく場合には、 R_w は、パワースペクトル $P_m(i, k)$ の推定結果に基づく行列となる。また、第1の実施形態と同様の思想に基づく場合には、パワースペクトル $P_m(i, k)$ 及び $Q_m(i, k)$ それぞれの推定結果に応じて決定されたパワースペクトル $W_m(i, k)$ に基づく行列となる。

[0223] 次に、入力相関行列算出部73、ノイズ相関行列算出部74、及びマルチチャネルウィナーフィルタ算出部75について説明する。

[0224] (入力相関行列算出部73)

入力相関行列算出部73は、周波数分解部13から集音部110ごとに音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を取得する。次いで、入力相関行列算出部73は、取得した各周波数成分 $X_m(i, k)$ を入力として、前述した(式20)に基づき、入力相関行列 R_x を周波数ごとに算出する。そして、入力相関行列算出部73は、周波数ごとに算出した入力相関行列 R_x をマルチチャネルウィナーフィルタ算出部75に出力する。

[0225] (ノイズ相関行列算出部74)

ノイズ相関行列算出部74は、周波数分解部13から集音部110ごとに音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ を取得する。次いで、ノイズ相関行列算出部74は、音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ の取得結果に基づき休止区間を特定する。次いで、ノイズ相関行列算出部74は、特定した休止区間における当該各周波数成分 $X_m(i, k)$ を入力とし

て、前述した（式 2 1）に基づき、ノイズ相関行列 R_n を周波数ごとに算出する。そして、ノイズ相関行列算出部 7 4 は、周波数ごとに算出したノイズ相関行列 R_n をマルチチャネルウィナーフィルタ算出部 7 5 に出力する。

[0226] （マルチチャネルウィナーフィルタ算出部 7 5）

マルチチャネルウィナーフィルタ算出部 7 5 は、フィルタ推定部 7 2 から、フィルタ係数 $w(i, k)$ に基づくフィルタ G の算出結果を周波数ごとに取得する。また、マルチチャネルウィナーフィルタ算出部 7 5 は、入力相関行列算出部 7 3 から、入力相関行列 R_x の算出結果を周波数ごとに取得する。また、マルチチャネルウィナーフィルタ算出部 7 5 は、ノイズ相関行列算出部 7 4 から、ノイズ相関行列 R_n の算出結果を周波数ごとに取得する。次いで、マルチチャネルウィナーフィルタ算出部 7 5 は、周波数ごとに取得した、フィルタ G 、入力相関行列 R_x 、及びノイズ相関行列 R_n それぞれの算出結果に基づき、周波数ごとにマルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ を算出する。そして、マルチチャネルウィナーフィルタ算出部 7 5 は、マルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ の算出結果をマルチチャネルウィナーフィルタ処理部 7 6 に出力する。これにより、マルチチャネルウィナーフィルタ処理部 7 6 は、集音部 1 1 0 ごとの音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ に対して、当該マルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ に基づきフィルタリング処理を施すことが可能となる。なお、マルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ の算出方法については、より詳細な内容について別途後述する。

[0227] 以上、図 2 4 を参照して、本実施形態に係る情報処理装置の機能構成の一例について説明した。なお、上述した情報処理装置 7 0 の各構成のうち、例えば、マルチチャネルウィナーフィルタ算出部 7 5 及びマルチチャネルウィナーフィルタ処理部 7 6 が、「出力制御部」の一例に相当する。

[0228] < 3. 3. マルチチャネルウィナーフィルタの算出方法の詳細 >

続いて、マルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ の算出方法についてより詳しく説明する。なお、本説明では、フレーム番号 i 及び離散周波

番号 k の記載については省略するものとする。

[0229] まず、前述した（式 2 3）に着目する。前述したように、フィルタ G は、フィルタ係数 $w(i, k)$ に基づき算出される。そのため、（式 2 3）に示す出力信号 Y は、前述した（式 7）を鑑みると、フィルタ係数 $w(i, k)$ に基づくフィルタ G と、入力信号 X と、により、以下に（式 2 7）として示す計算式で表すことが可能である。

[0230] [数23]

$$Y = G^H X$$

・・・（式 2 7）

[0231] 即ち、前述した（式 2 3）は、上記（式 2 7）で示された出力信号 Y と、フィルタ G 、入力相関行列 R_x 、及びノイズ相関行列 R_n と、により、以下に（式 2 8）として示す計算式で表すことが可能である。

[0232] [数24]

$$\begin{aligned} W_{mwf} &= (R_x + \mu R_n)^{-1} E[XY^H] \\ &= (R_x + \mu R_n)^{-1} E[XX^H G] \\ &= (R_x + \mu R_n)^{-1} R_x G \end{aligned}$$

・・・（式 2 8）

[0233] ここで、上記（式 2 8）と前述した（式 1 9）とを比較するとわかるように、上記（式 2 8）は、前述した（式 1 9）において、リファレンスマイクを特定する行列 e_i を、フィルタ G に置き換えた計算式と等価であることがわかる。また、フィルタ G は、前述した（式 2 5）及び（式 2 6）を鑑みると、以下に（式 2 9）として示す計算式で表すことも可能である。

[0234]

[数25]

$$G = \begin{pmatrix} g_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & g_M \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_M \end{pmatrix} = [g_1 a_1, \dots, g_i a_i, \dots, g_M a_M]^T$$

・・・ (式29)

[0235] ここで、上記(式29)において、係数 g_1 、 $\dots$ 、 g_i 、 $\dots$ 、 g_M は、(式25)及び(式26)における行列 R_w に基づき決定される係数である。より具体的には、係数 g_1 、 $\dots$ 、 g_i 、 $\dots$ 、 g_M は、例えば、前述した第1の実施形態におけるパワースペクトル $P_m(i, k)$ の推定結果や、第2の実施形態におけるパワースペクトル $W_m(i, k)$ の推定結果に基づき決定されることとなる。即ち、係数 g_1 、 $\dots$ 、 g_i 、 $\dots$ 、 g_M は、各集音部110それぞれに対するランダムノイズの影響の大きさに応じた重みを示しており、換言すると、各集音部110それぞれの集音結果が、目的音の集音結果としてのどの程度信頼できるかを示している。

[0236] 即ち、本実施形態に係る情報処理装置70においては、各集音部110による音響(特に、ランダムノイズ)の集音結果に応じて、係数 g_1 、 $\dots$ 、 g_i 、 $\dots$ 、 g_M が変化し、当該係数に応じて、例えば、ランダムノイズの影響がより小さい集音部110がリファレンスマイクとして動的に選択されることとなる。

[0237] <3.4. 評価>

以上説明したように、本実施形態に係る情報処理装置70は、各集音部110それぞれに対応する音響信号 $x_m(n)$ の各周波数成分 $X_m(i, k)$ に基づき、周波数ごとに、各集音部110のパワースペクトル(例えば、前述したパワースペクトル $P_m(i, k)$ やパワースペクトル $W_m(i, k)$)を推定する。そして、情報処理装置70は、当該パワースペクトルの推定結果に基づきフィルタ係数 $w(i, k)$ を推定し、当該フィルタ係数 $w(i, k)$

）の推定結果をマルチチャネルウィナーフィルタ $W_{mwf}(i, k)$ の算出に利用する。このような構成により、本実施形態に係る情報処理装置70は、複数の集音部110のうち、ランダムノイズの影響がより小さい集音部110をリファレンスマイクとして動的に選択することが可能となる。即ち、本実施形態に係る情報処理装置70に依れば、リファレンスマイクが固定的に設定される従来のマルチチャネルウィナーフィルタを適用した場合に比べて、ランダムノイズの影響をより低減し、ひいては、より好適な態様で目的音を抽出することが可能となる。

[0238] なお、本実施形態に係る信号処理の適用先は、必ずしも図1に示すような所謂ネックバンド型のウェアラブルデバイスの例のみには限定されない。具体的には、本実施系に係る信号処理は、複数の集音部を備える装置であれば適用することが可能である。なお、より好適には、複数の集音部は、目的音の音源（例えば、音声が発話される口元）からの距離が互いに異なるように配置されているとよい。また、より好適には、複数の集音部は、目的音の音源に対して互いに異なる方向に位置するように配置されているとよい。

[0239] <<4. ハードウェア構成>>

次に、図25を参照して、本開示の各実施形態に係る情報処理装置10（即ち、上述した信号処理装置11～14）のハードウェア構成の一例について説明する。図25は、本開示の各実施形態に係る情報処理装置10のハードウェア構成の一例を示した図である。

[0240] 図25に示すように、本実施形態に係る情報処理装置10は、プロセッサ901と、メモリ903と、ストレージ905と、操作デバイス907と、報知デバイス909と、音響デバイス911と、集音デバイス913と、バス917とを含む。また、情報処理装置10は、通信デバイス915を含んでもよい。

[0241] プロセッサ901は、例えばCPU（Central Processing Unit）、GPU（Graphics Processing Unit）、DSP（Digital Signal Processor）又はSoC（System on Chip）であってよく、情報処理装置10の様々

な処理を実行する。プロセッサ 901 は、例えば、各種演算処理を実行するための電子回路により構成することが可能である。なお、前述した周波数分解部 13、チャネルパワー推定部 15、フィルタ推定部 16、フィルタ処理部 17、及び周波数合成部 18 は、プロセッサ 901 により実現され得る。

[0242] メモリ 903 は、RAM (Random Access Memory) 及び ROM (Read Only Memory) を含み、プロセッサ 901 により実行されるプログラム及びデータを記憶する。ストレージ 905 は、半導体メモリ又はハードディスクなどの記憶媒体を含み得る。

[0243] 操作デバイス 907 は、ユーザが所望の操作を行うための入力信号を生成する機能を有する。操作デバイス 907 は、例えば、タッチパネルとして構成され得る。また、他の一例として、操作デバイス 907 は、例えばボタン、スイッチ、及びキーボードなどユーザが情報を入力するための入力部と、ユーザによる入力に基づいて入力信号を生成し、プロセッサ 901 に供給する入力制御回路などから構成されてよい。

[0244] 報知デバイス 909 は、出力デバイスの一例であり、例えば、液晶ディスプレイ (LCD: Liquid Crystal Display) 装置、有機 EL (OLED: Organic Light Emitting Diode) ディスプレイなどのデバイスであってよい。この場合には、報知デバイス 909 は、画面を表示することにより、ユーザに対して所定の情報を報知することができる。

[0245] なお、上記に示した報知デバイス 909 の例はあくまで一例であり、ユーザに対して所定の情報を報知可能であれば、報知デバイス 909 の態様は特に限定されない。具体的な一例として、報知デバイス 909 は、LED (Light Emitting Diode) のように、点灯又は点滅のパターンにより、所定の情報をユーザに報知するデバイスであってもよい。また、報知デバイス 909 は、所謂バイブレータのように、振動することで、所定の情報をユーザに報知するデバイスであってもよい。

[0246] 音響デバイス 911 は、スピーカ等のように、所定の音響信号を出力することで、所定の情報をユーザに報知するデバイスである。

[0247] 集音デバイス 913 は、マイクロフォン等のような、ユーザから発せられた音声や周囲の環境の音響を集音し、音響情報（音響信号）として取得するためのデバイスである。また、集音デバイス 913 は、集音された音声や音響を示すアナログの音響信号を示すデータを音響情報として取得してもよいし、当該アナログの音響信号をデジタルの音響信号に変換し、変換後のデジタルの音響信号を示すデータを音響情報として取得してもよい。なお、前述した集音部 110（例えば、図 6 に示す集音部 111～11M）は、集音デバイス 913 により実現され得る。

[0248] 通信デバイス 915 は、情報処理装置 10 が備える通信手段であり、ネットワークを介して外部装置と通信する。通信デバイス 915 は、有線または無線用の通信インタフェースである。通信デバイス 915 を、無線通信インタフェースとして構成する場合には、当該通信デバイス 915 は、通信アンテナ、RF（Radio Frequency）回路、ベースバンドプロセッサなどを含んでもよい。

[0249] 通信デバイス 915 は、外部装置から受信した信号に各種の信号処理を行う機能を有し、受信したアナログ信号から生成したデジタル信号をプロセッサ 901 に供給することが可能である。

[0250] バス 917 は、プロセッサ 901、メモリ 903、ストレージ 905、操作デバイス 907、報知デバイス 909、音響デバイス 911、集音デバイス 913、及び通信デバイス 915 を相互に接続する。バス 917 は、複数の種類のバスを含んでもよい。

[0251] また、コンピュータに内蔵されるプロセッサ、メモリ、及びストレージなどのハードウェアを、上記した情報処理装置 10 が有する構成と同等の機能を発揮させるためのプログラムも作成可能である。また、当該プログラムを記録した、コンピュータに読み取り可能な記憶媒体も提供され得る。

[0252] <<5. むすび>>

以上、説明したように、本実施形態に係る情報処理装置 10 は、少なくとも一部に流線形の形状を有する凸部を備え、当該凸部の先端、または、当該

先端の近傍に位置するように集音部 110 が支持される。このような構成とすることで、例えば、風切音、振動に伴う雑音、及び装置の装着に伴う衣擦れ等のようなランダムに発生する雑音の影響を緩和し、より好適な態様で目的音（例えば、ユーザの音声）を集音することが可能となる。

[0253] また、本実施形態に係る情報処理装置 10 は、複数の集音部 110 を備え、当該複数の集音部 110 が互いに異なる方向を向くように支持されていてもよい。このような構成により、風切音、振動に伴う雑音、及び装置の装着に伴う衣擦れ等の雑音がランダムに発生するような状況下においても、一部の集音部（即ち、雑音の影響が少ない集音部）の集音結果に基づき、他の集音部の特性を補うことが可能となる。

[0254] 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、本開示の技術的範囲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。

[0255] また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

[0256] なお、以下のような構成も本開示の技術的範囲に属する。

(1)

集音部と、

流線形の形状を有する凸部を少なくとも一部に備え、前記凸部の先端、または、当該先端の近傍に位置するように前記集音部を支持する支持部材と、
を備える、情報処理装置。

(2)

前記集音部である第 1 の集音部に加えて、当該第 1 の集音部とは異なる 1

以上の第2の集音部を備える、前記（1）に記載の情報処理装置。

（3）

前記支持部材は、複数の前記第2の集音部のそれぞれが、互いに異なる方向を向くように支持する、前記（2）に記載の情報処理装置。

（4）

前記支持部材は、ユーザの所定の部位に装着され、前記集音部と当該部位とが所定の位置関係となるように当該集音部を支持する、前記（1）に記載の情報処理装置。

（5）

前記部位は、首であり、

前記支持部材は、前記首に装着された場合に、前記凸部の先端が当該ユーザの略前方を向くように当該凸部が設けられている、

前記（4）に記載の情報処理装置。

（6）

前記集音部である第1の集音部に加えて、当該第1の集音部とは異なる複数の第2の集音部を備え、

複数の前記第2の集音部のうち少なくとも2以上の第2の集音部を、前記部位を基準として互いに略対称となる位置に支持する、

前記（4）または（5）に記載の情報処理装置。

（7）

前記第1の集音部と1以上の前記第2の集音部とのそれぞれにより集音された音響に基づき、前記第1の集音部に対して所定の方向から到来する音響に対する雑音成分を抑圧する信号処理部を備える、前記（2）に記載の情報処理装置。

（8）

前記信号処理部は、前記第1の集音部と1以上の前記第2の集音部とのそれぞれにより集音された音響に基づき、当該音響の周波数成分それぞれの信号レベルを推定し、当該信号レベルの推定結果に基づき、前記雑音成分を抑

圧する、前記（７）に記載の情報処理装置。

（９）

前記信号処理部は、前記第１の集音部と１以上の前記第２の集音部とのうちの少なくとも複数の集音部それぞれにより集音された第１の音響と、従前の処理により前記雑音成分が抑圧された第２の音響との間の相関性に基づき、前記第１の音響に含まれる前記雑音成分を抑圧する、前記（７）に記載の情報処理装置。

（１０）

前記支持部材は、前記複数の集音部のうち少なくとも２以上の集音部のそれぞれと、所定の音源との間の距離が互いに異なるように、当該複数の集音部を支持する、前記（９）に記載の情報処理装置。

（１１）

前記支持部材は、前記複数の集音部のうち少なくとも２以上の集音部のそれぞれが、所定の音源に対して互いに異なる方向に位置するように、当該複数の集音部を支持する、前記（９）または（１０）に記載の情報処理装置。

（１２）

前記第１の集音部と１以上の前記第２の集音部とのうち一部の集音部により集音された音響を選択的に出力する出力制御部を備え、

前記出力制御部は、前記第１の集音部と１以上の前記第２の集音部とのそれぞれにより集音された音響に基づき、当該音響の周波数成分それぞれの信号レベルを推定し、当該信号レベルの推定結果に基づき、前記一部の集音部を選択する、

前記（２）に記載の情報処理装置。

（１３）

前記出力制御部は、マルチチャネルウィナーフィルタを含み、当該マルチチャネルウィナーフィルタのリファレンスマイクを、前記信号レベルの推定結果に基づき選択する、前記（１２）に記載の情報処理装置。

（１４）

前記支持部材は、少なくとも一部に略矩形状の面を有する筐体であり、
前記筐体は、前記略矩形状の面の角を含む所定の領域中に前記凸部を有し、
当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、
前記（１）に記載の情報処理装置。

（１５）

複数の前記集音部を備え、

前記筐体は、前記略矩形状の面の角のうち複数の角それぞれについて、当該角を含む所定の領域中に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、前記（１４）に記載の情報処理装置。

（１６）

ユーザの腕に対して前記筐体を支持するバンド部を備え、

前記バンド部は、前記腕に装着された場合に、当該腕を基準として前記筐体と略対称となる位置に前記集音部とは異なる他の集音部を備える、

前記（１４）または（１５）に記載の情報処理装置。

（１７）

前記支持部材は、ユーザの頭部に装着されるメガネ型のフレームであり、

前記フレームは、フロントの少なくとも一部に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、

前記（１）に記載の情報処理装置。

（１８）

前記フレームは、ブリッジ、または、当該ブリッジの近傍に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、前記（１７）に記載の情報処理装置。

符号の説明

[0257]	１０	情報処理装置
	１３	周波数分解部
	１５	チャネルパワー推定部
	１６	フィルタ推定部

- 1 7 フィルタ処理部
- 1 8 周波数合成部
- 1 1 0 ~ 1 1 3 集音部
- 6 0 情報処理装置
- 6 5 チャネルパワー推定部
- 6 5 1 入力パワー推定部
- 6 5 3 非相関成分パワー推定部
- 6 5 5 ランダムノイズパワー推定部
- 6 6 フィルタ推定部

請求の範囲

- [請求項1] 集音部と、
流線形の形状を有する凸部を少なくとも一部に備え、前記凸部の先端、または、当該先端の近傍に位置するように前記集音部を支持する支持部材と、
を備える、情報処理装置。
- [請求項2] 前記集音部である第1の集音部に加えて、当該第1の集音部とは異なる1以上の第2の集音部を備える、請求項1に記載の情報処理装置。
- [請求項3] 前記支持部材は、複数の前記第2の集音部のそれぞれが、互いに異なる方向を向くように支持する、請求項2に記載の情報処理装置。
- [請求項4] 前記支持部材は、ユーザの所定の部位に装着され、前記集音部と当該部位とが所定の位置関係となるように当該集音部を支持する、請求項1に記載の情報処理装置。
- [請求項5] 前記部位は、首であり、
前記支持部材は、前記首に装着された場合に、前記凸部の先端が当該ユーザの略前方を向くように当該凸部が設けられている、
請求項4に記載の情報処理装置。
- [請求項6] 前記集音部である第1の集音部に加えて、当該第1の集音部とは異なる複数の第2の集音部を備え、
複数の前記第2の集音部のうち少なくとも2以上の第2の集音部を、前記部位を基準として互いに略対称となる位置に支持する、
請求項4に記載の情報処理装置。
- [請求項7] 前記第1の集音部と1以上の前記第2の集音部とのそれぞれにより集音された音響に基づき、前記第1の集音部に対して所定の方向から到来する音響に対する雑音成分を抑圧する信号処理部を備える、請求項2に記載の情報処理装置。
- [請求項8] 前記信号処理部は、前記第1の集音部と1以上の前記第2の集音部

とのそれぞれにより集音された音響に基づき、当該音響の周波数成分それぞれの信号レベルを推定し、当該信号レベルの推定結果に基づき、前記雑音成分を抑圧する、請求項 7 に記載の情報処理装置。

[請求項9] 前記信号処理部は、前記第 1 の集音部と 1 以上の前記第 2 の集音部とのうちの少なくとも複数の集音部それぞれにより集音された第 1 の音響と、従前の処理により前記雑音成分が抑圧された第 2 の音響との間の相関性に基づき、前記第 1 の音響に含まれる前記雑音成分を抑圧する、請求項 7 に記載の情報処理装置。

[請求項10] 前記支持部材は、前記複数の集音部のうち少なくとも 2 以上の集音部のそれぞれと、所定の音源との間の距離が互いに異なるように、当該複数の集音部を支持する、請求項 9 に記載の情報処理装置。

[請求項11] 前記支持部材は、前記複数の集音部のうち少なくとも 2 以上の集音部のそれぞれが、所定の音源に対して互いに異なる方向に位置するように、当該複数の集音部を支持する、請求項 9 に記載の情報処理装置。

[請求項12] 前記第 1 の集音部と 1 以上の前記第 2 の集音部とを含む複数の集音部のうち、一部の集音部により集音された音響を選択的に出力する出力制御部を備え、

前記出力制御部は、前記第 1 の集音部と 1 以上の前記第 2 の集音部とのそれぞれにより集音された音響に基づき、当該音響の周波数成分それぞれの信号レベルを推定し、当該信号レベルの推定結果に基づき、前記一部の集音部を選択する、

請求項 2 に記載の情報処理装置。

[請求項13] 前記出力制御部は、マルチチャネルウィナーフィルタを含み、当該マルチチャネルウィナーフィルタのリファレンスマイクを、前記信号レベルの推定結果に基づき選択する、請求項 12 に記載の情報処理装置。

[請求項14] 前記支持部材は、少なくとも一部に略矩形状の面を有する筐体であ

り、

前記筐体は、前記略矩形状の面の角を含む所定の領域中に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、

請求項 1 に記載の情報処理装置。

[請求項15]

複数の前記集音部を備え、

前記筐体は、前記略矩形状の面の角のうち複数の角それぞれについて、当該角を含む所定の領域中に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、請求項 1 4 に記載の情報処理装置。

[請求項16]

ユーザの腕に対して前記筐体を支持するバンド部を備え、

前記バンド部は、前記腕に装着された場合に、当該腕を基準として前記筐体と略対称となる位置に前記集音部とは異なる他の集音部を備える、

請求項 1 4 に記載の情報処理装置。

[請求項17]

前記支持部材は、ユーザの頭部に装着されるメガネ型のフレームであり、

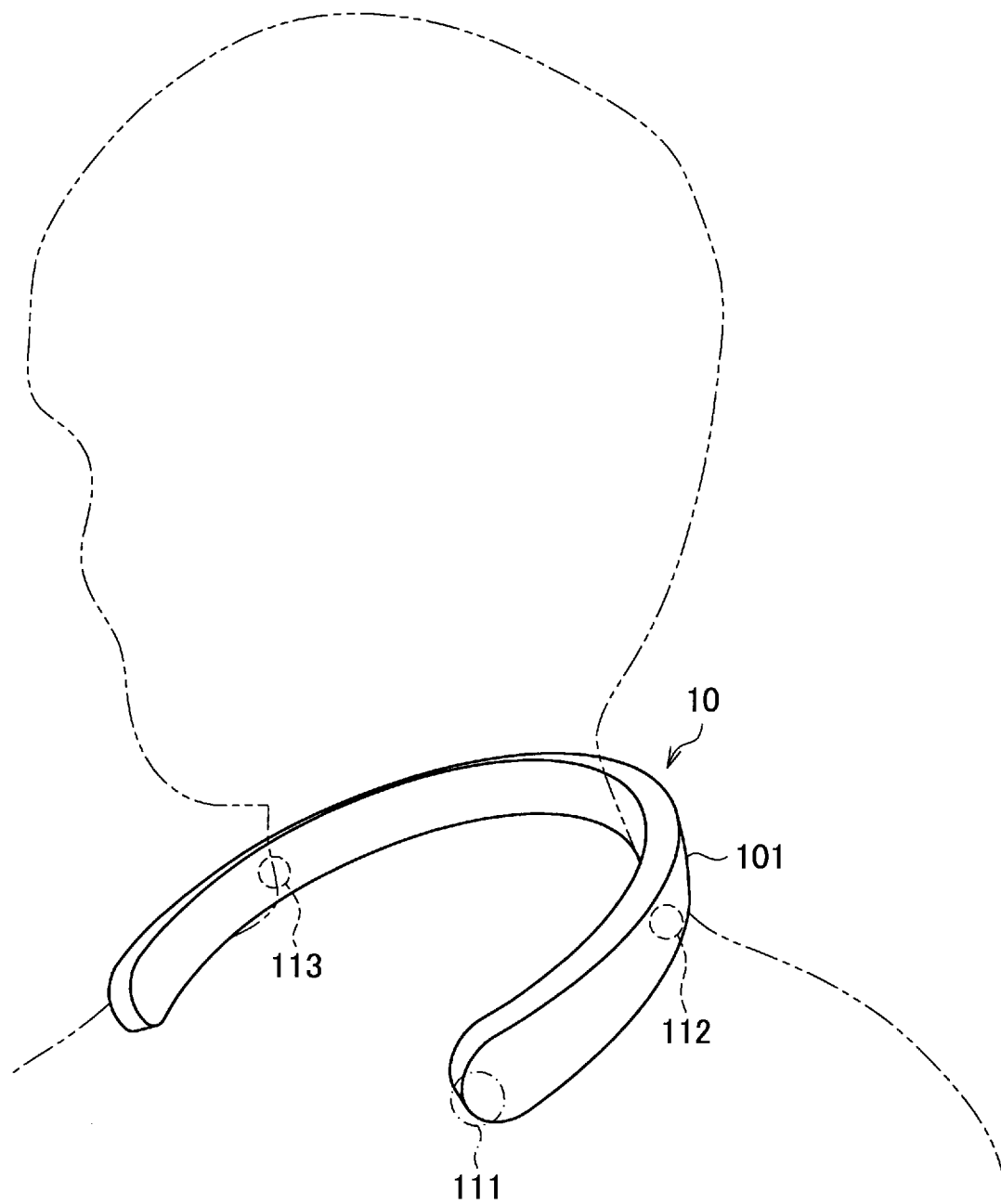
前記フレームは、フロントの少なくとも一部に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、

請求項 1 に記載の情報処理装置。

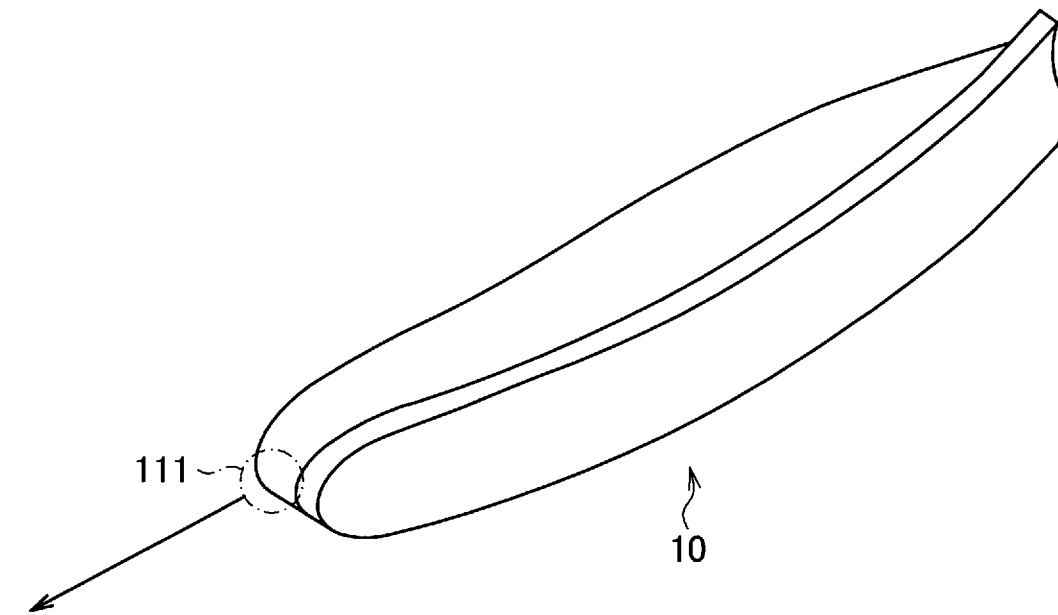
[請求項18]

前記フレームは、ブリッジ、または、当該ブリッジの近傍に前記凸部を有し、当該凸部の先端、または、当該先端の近傍に前記集音部を支持する、請求項 1 7 に記載の情報処理装置。

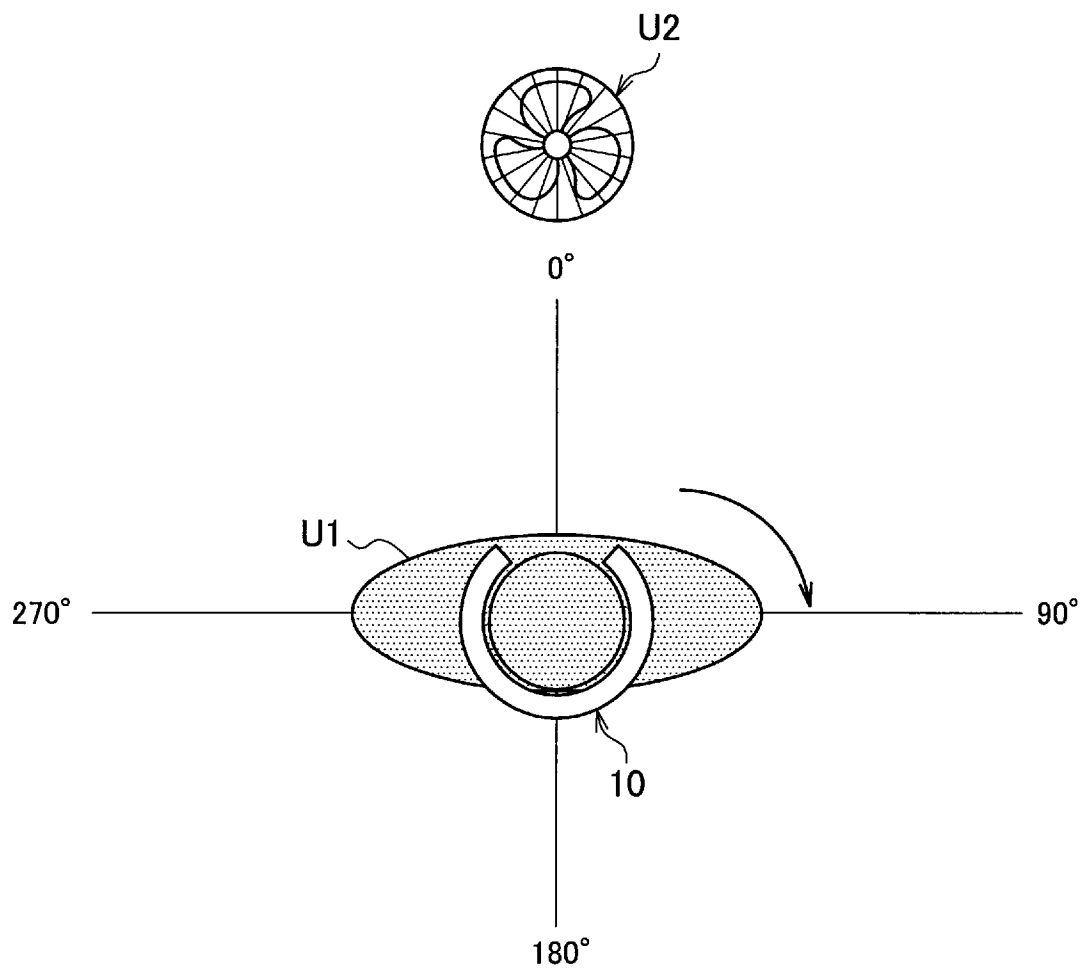
[図1]



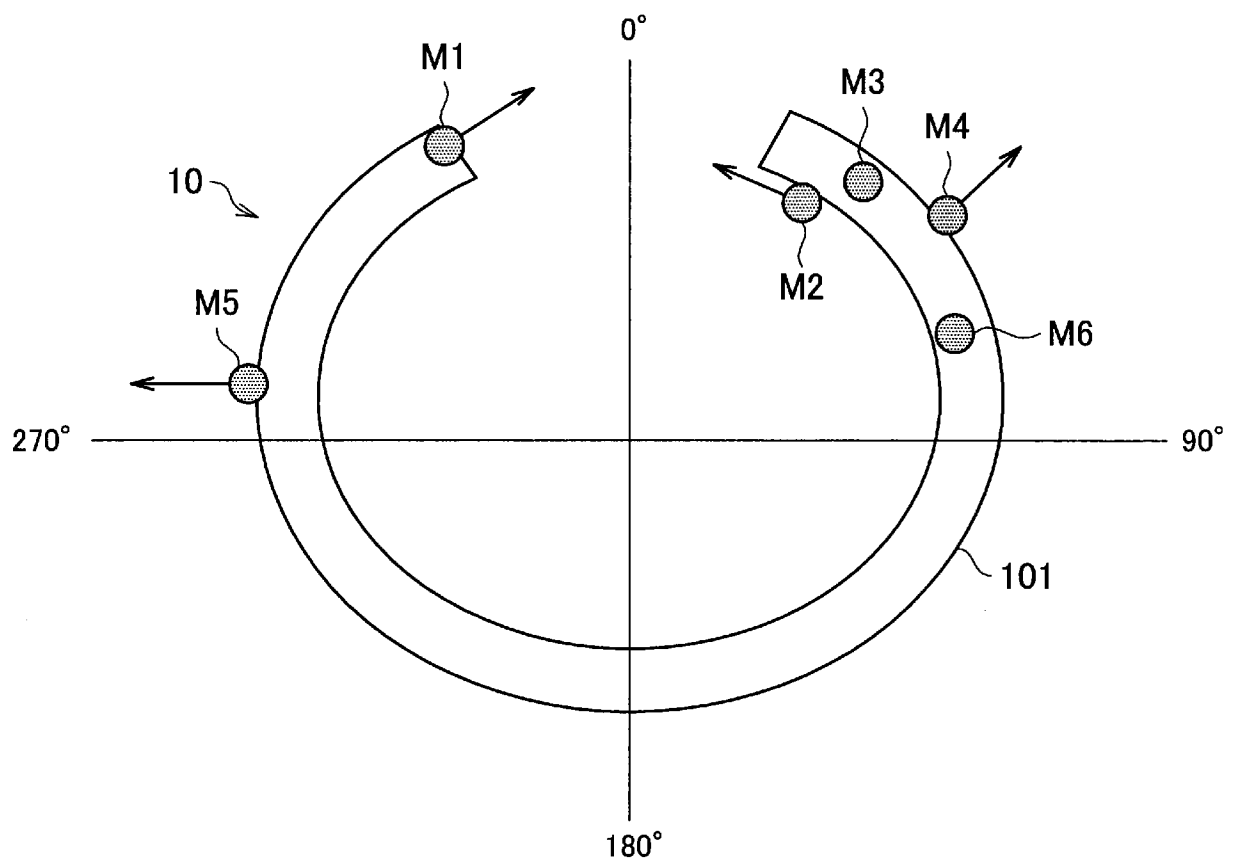
[図2]



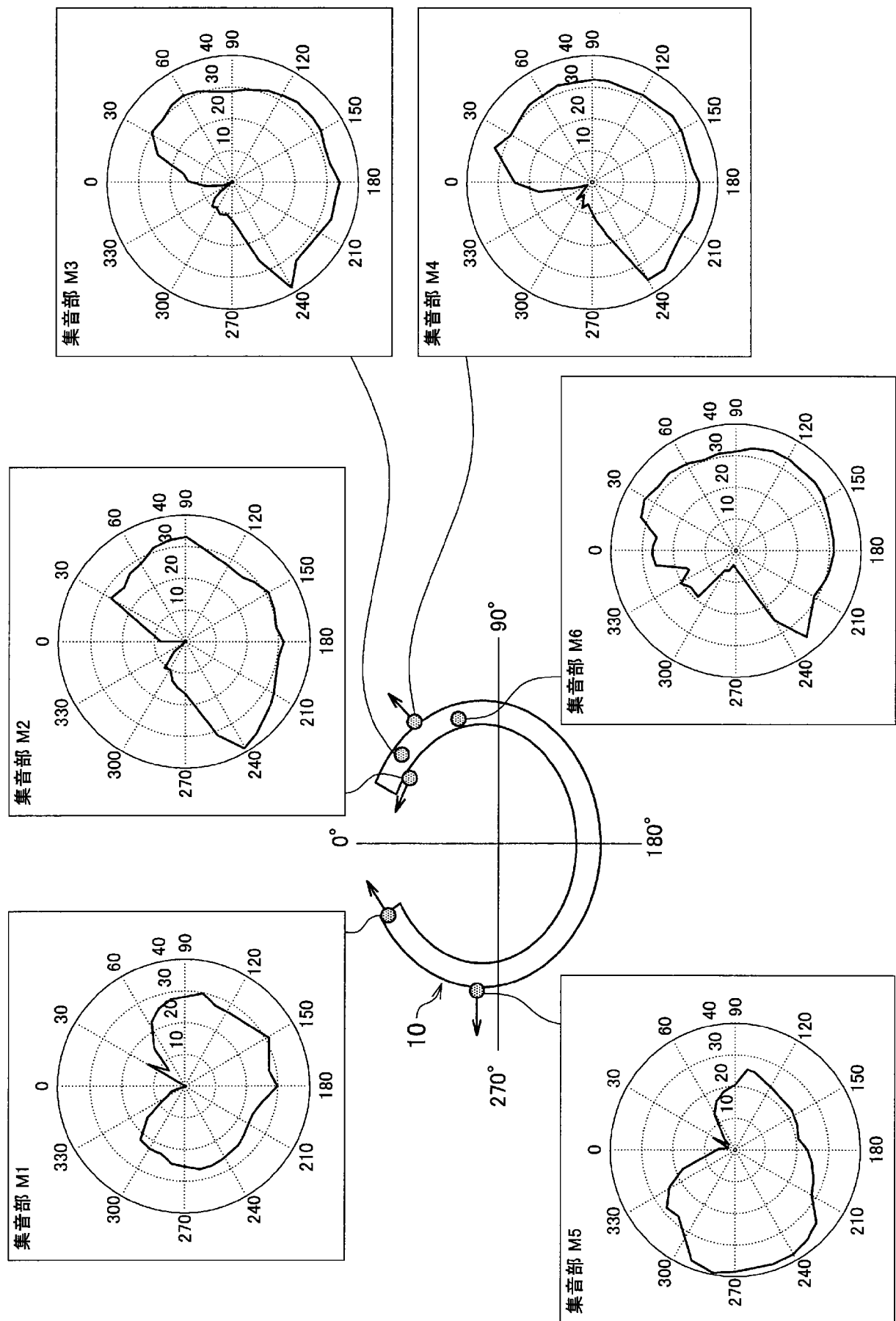
[図3]



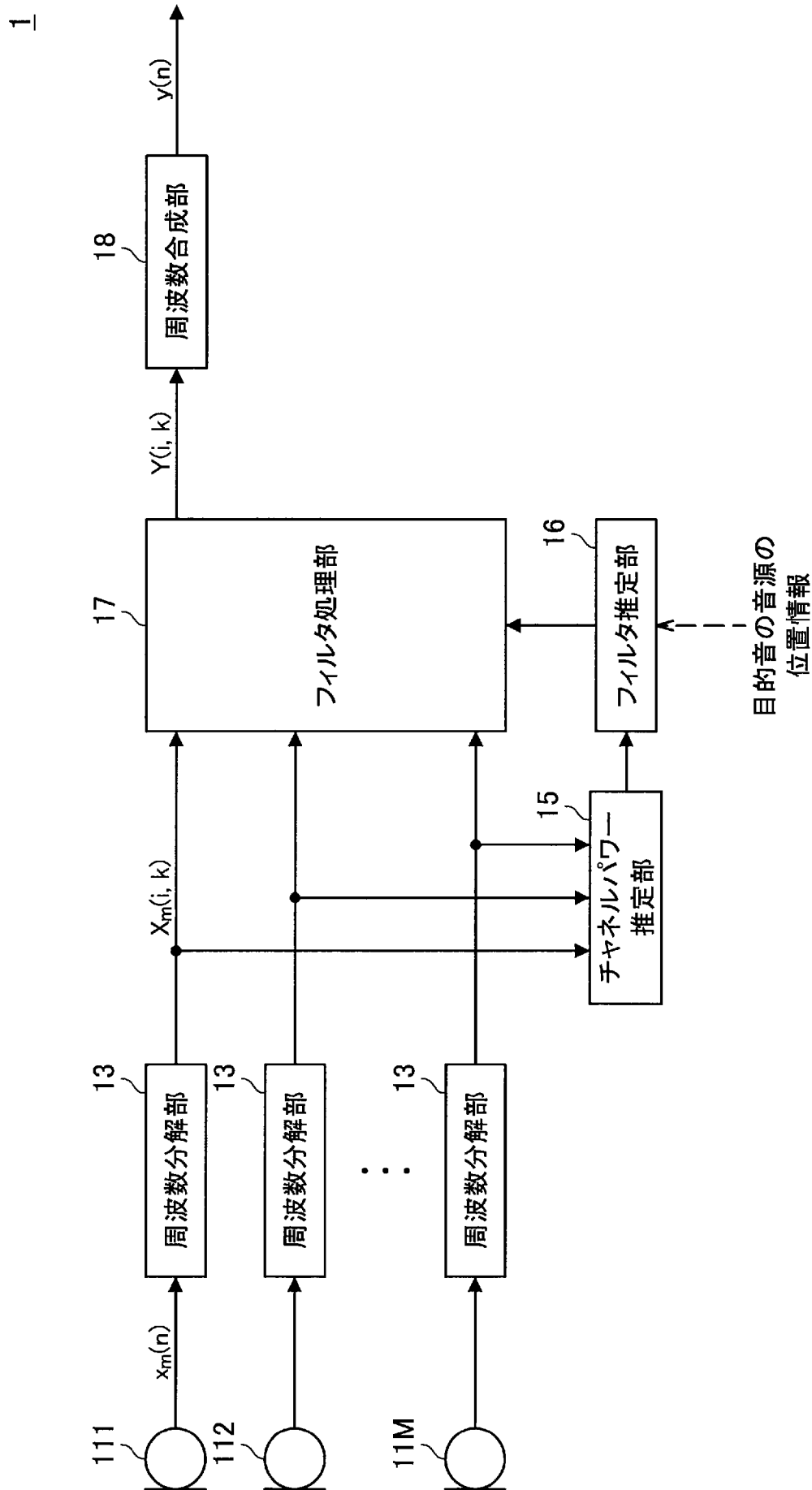
[図4]



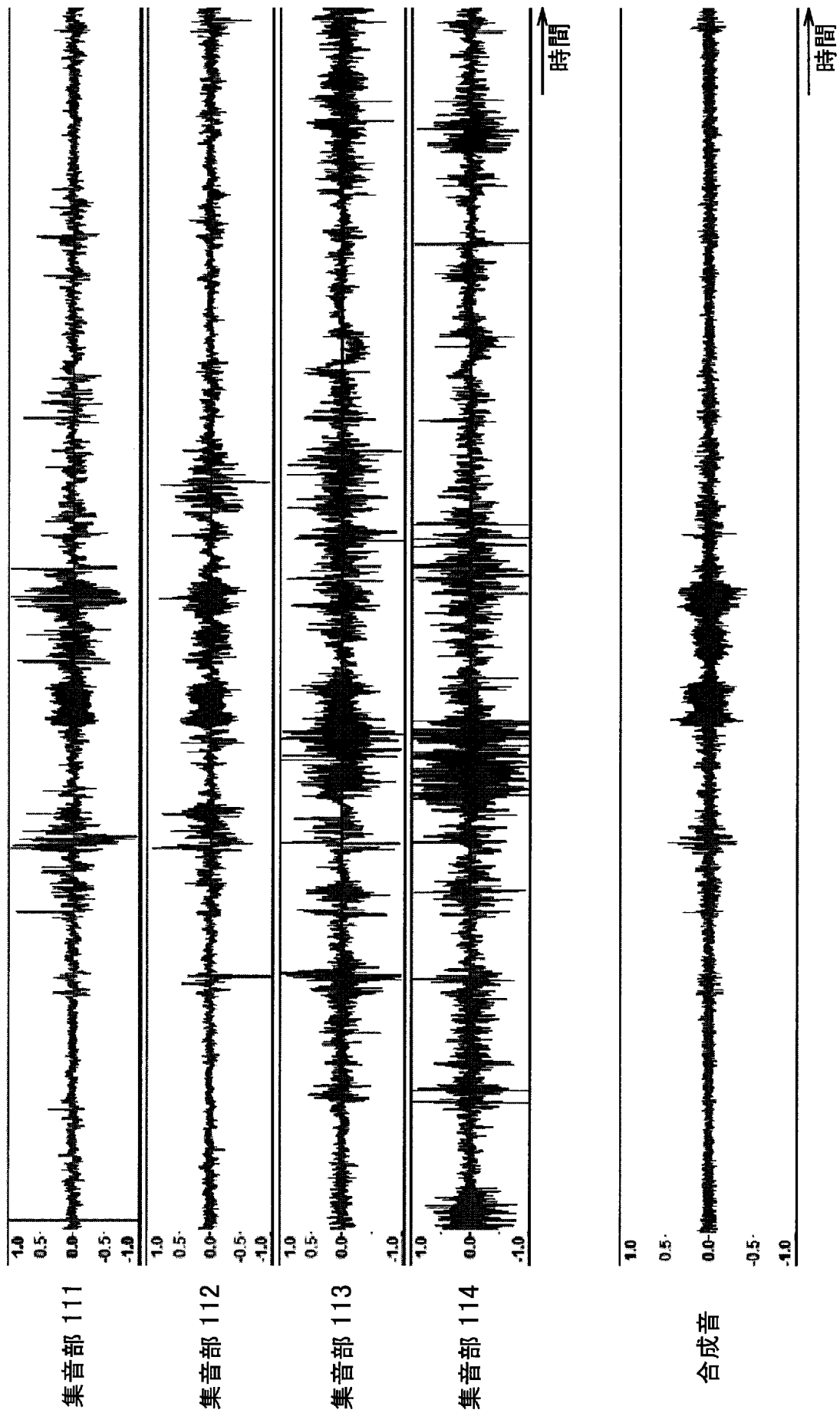
[図5]



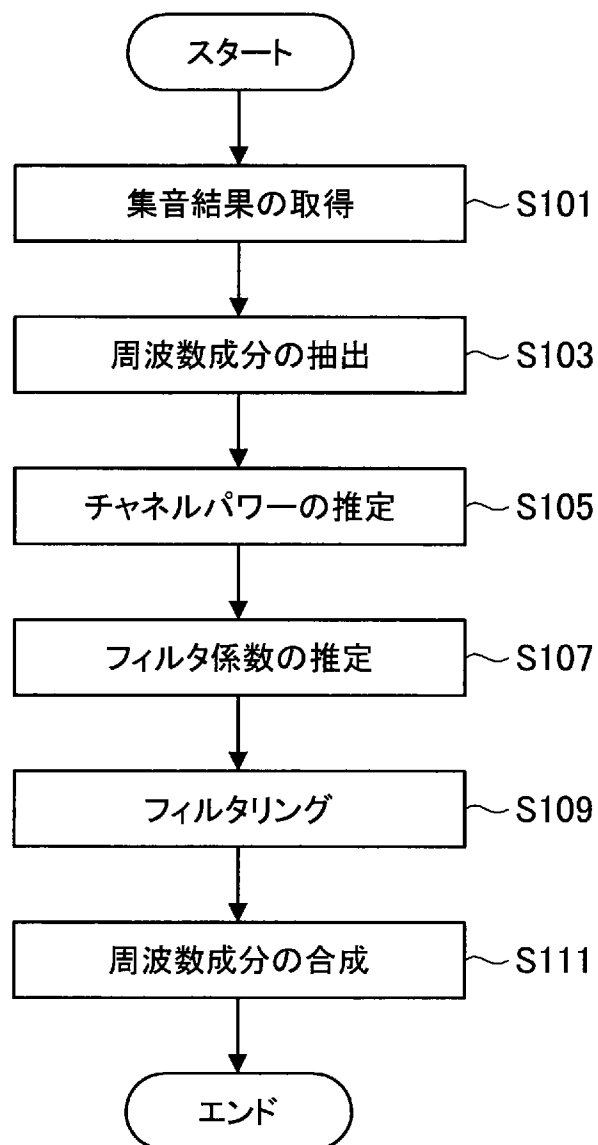
[図6]



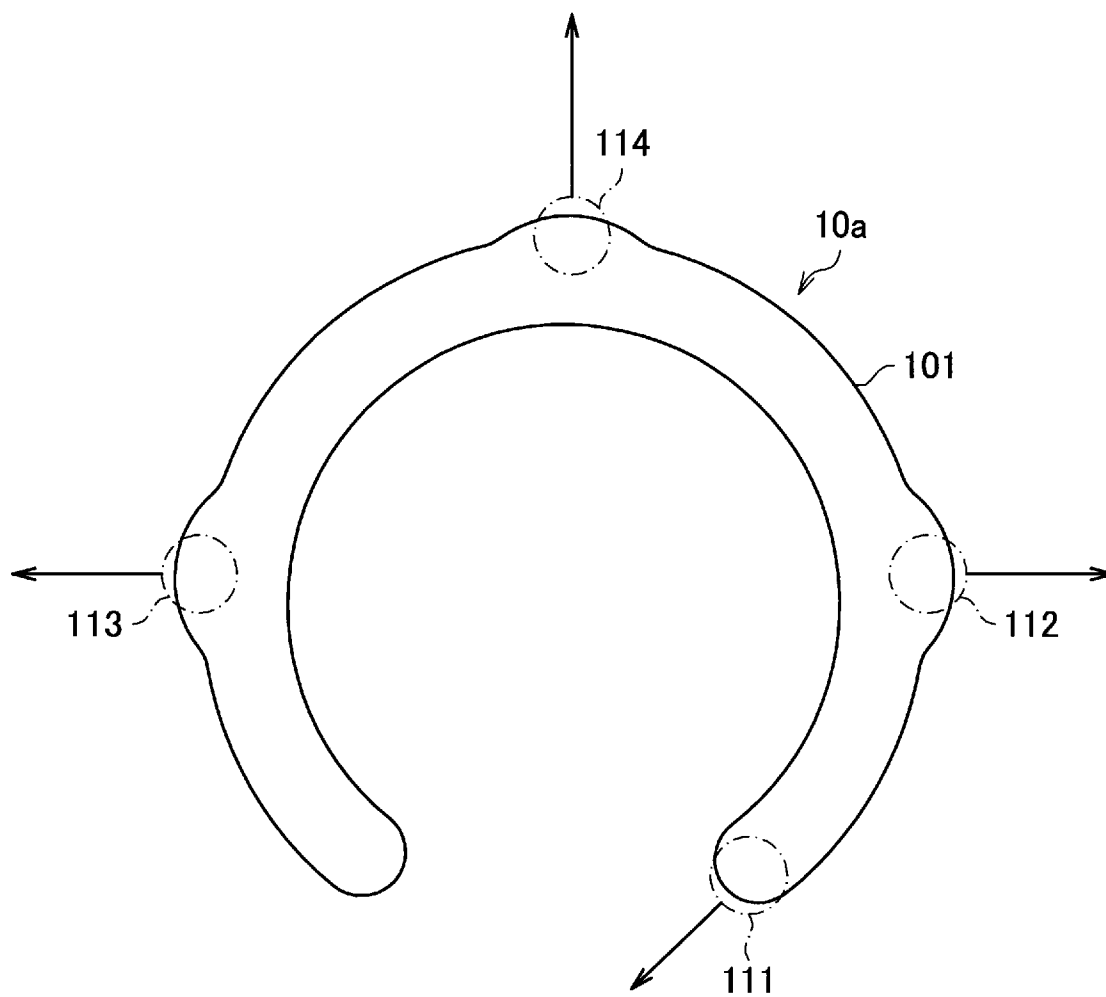
[図7]



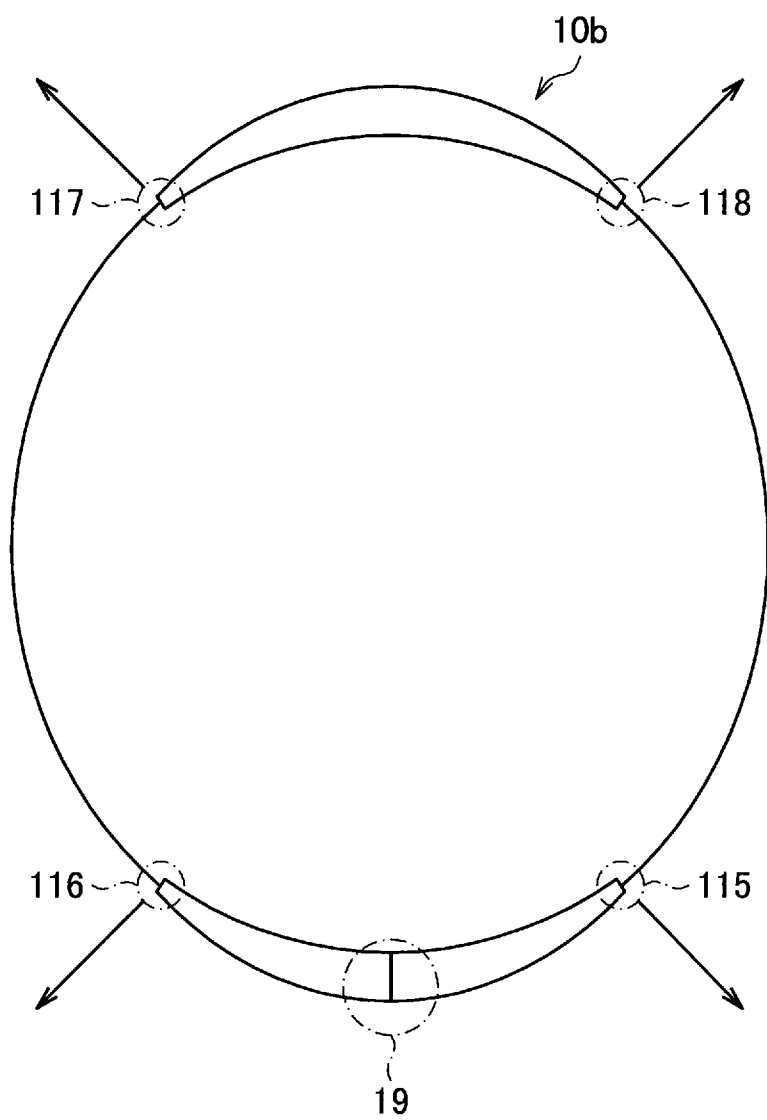
[図8]



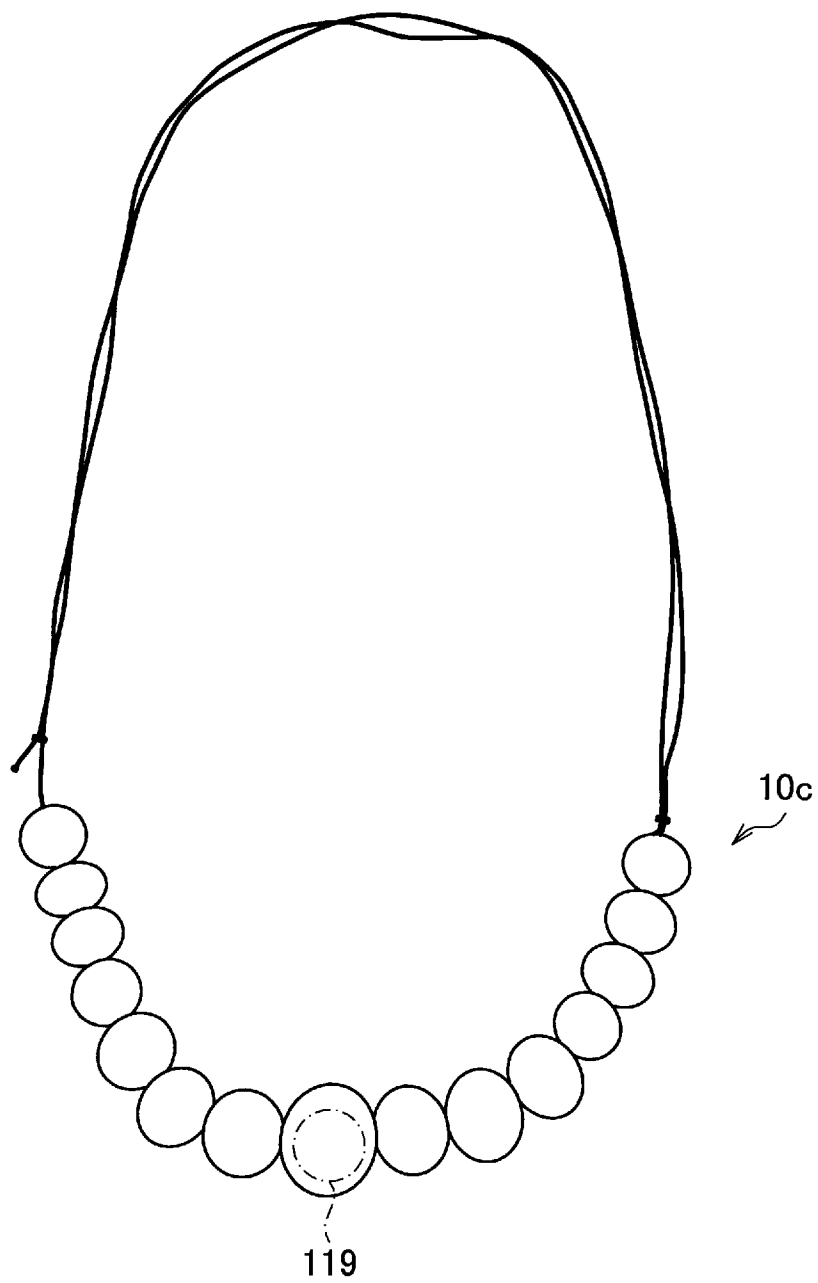
[図9]



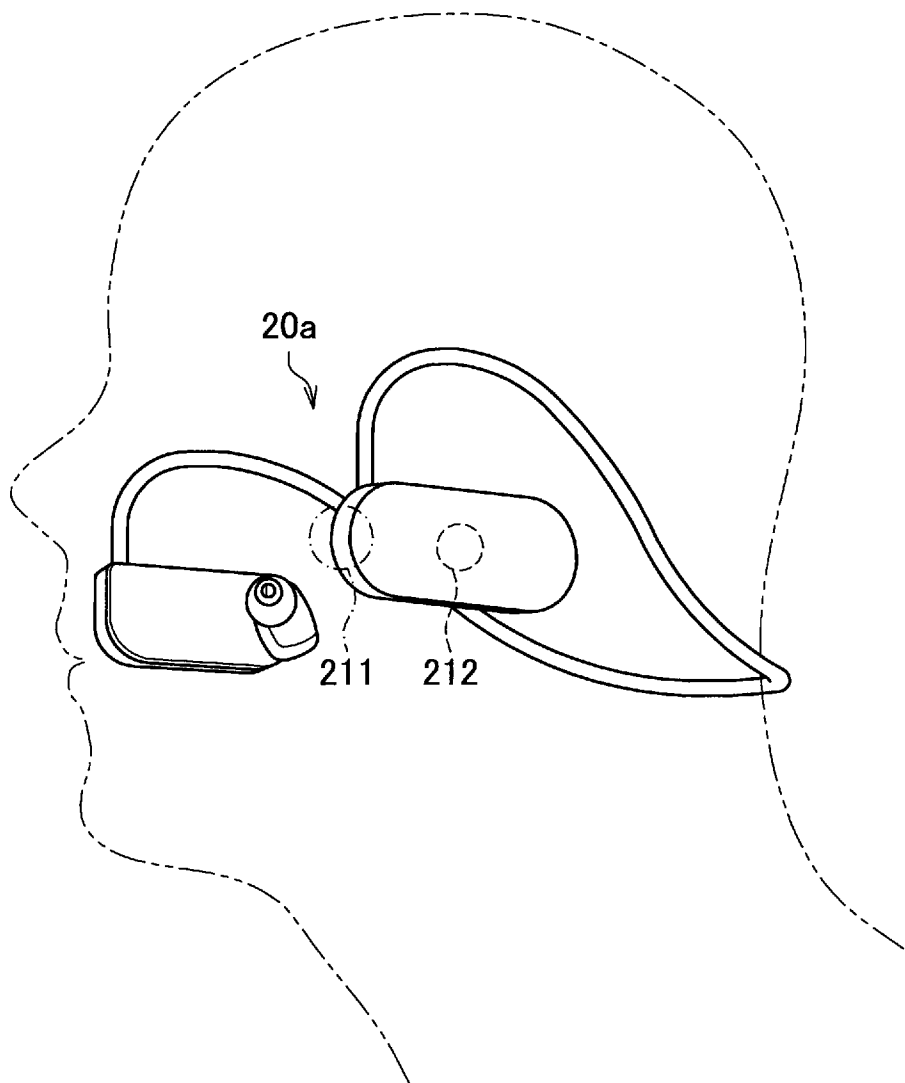
[図10]



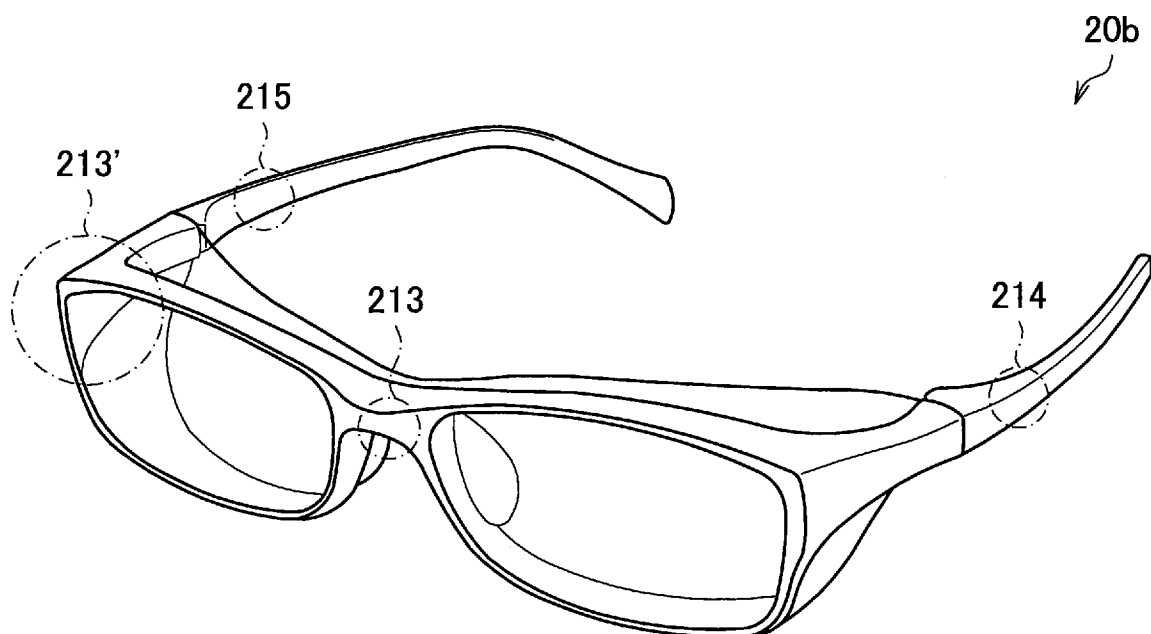
[図11]



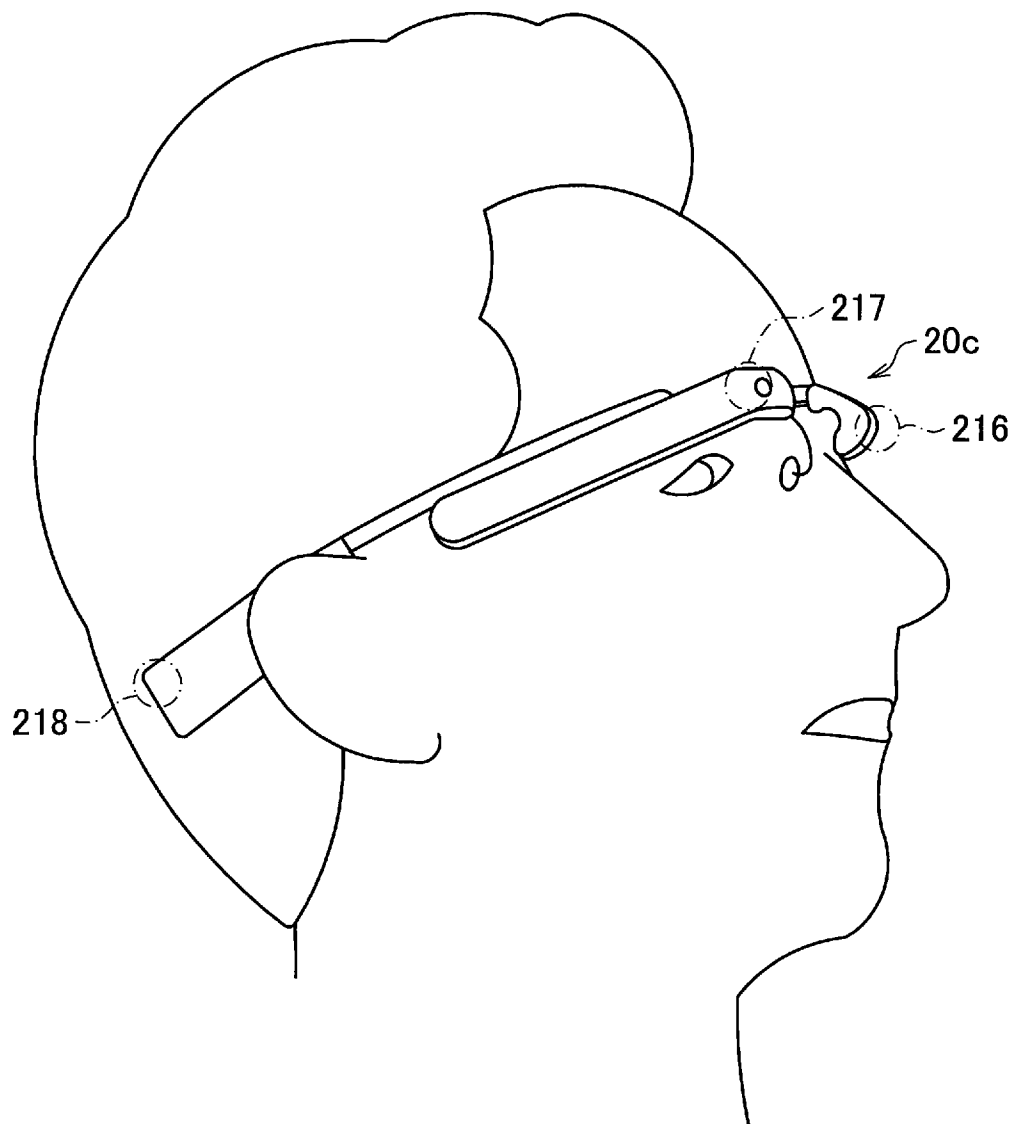
[図12]



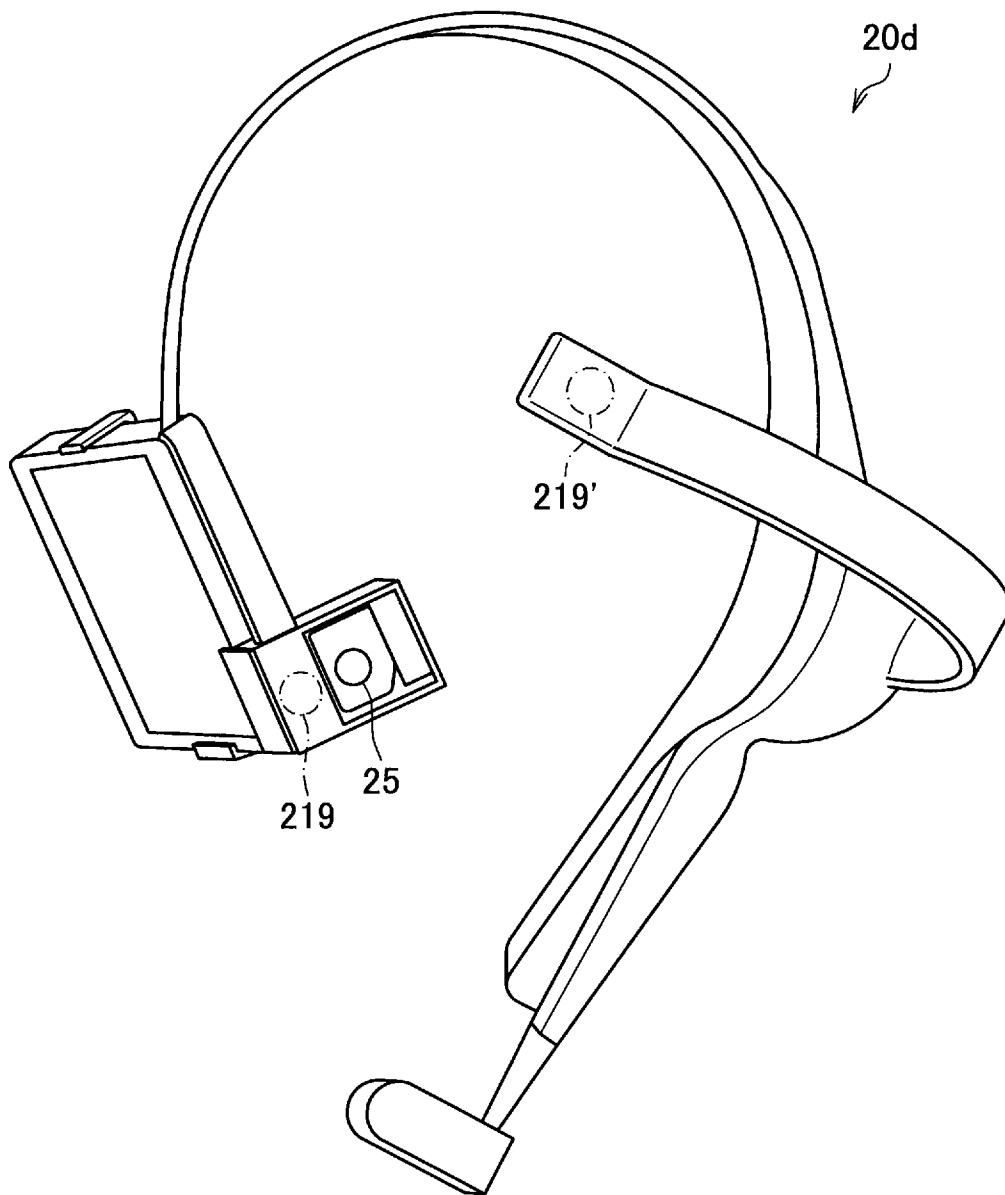
[図13]



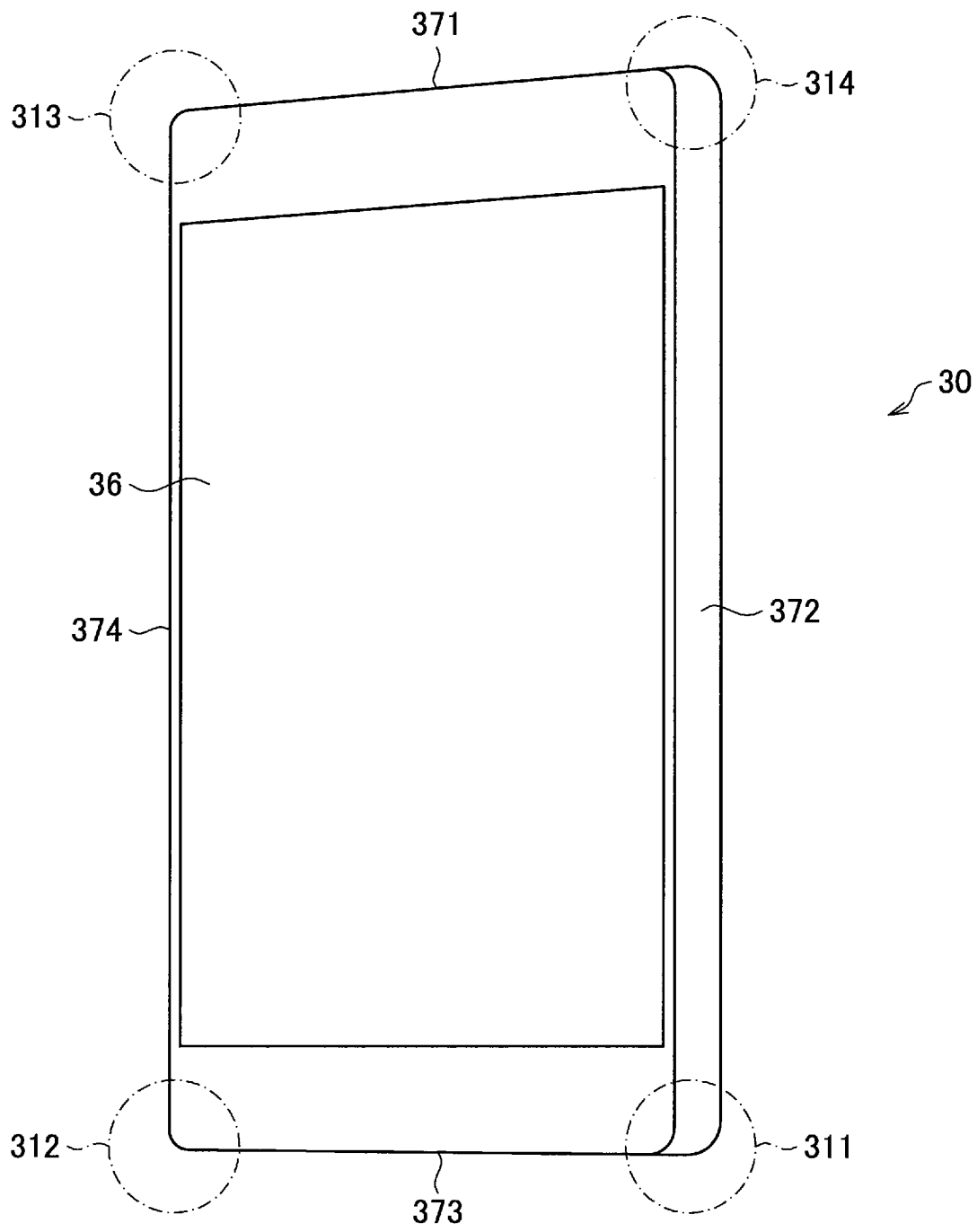
[図14]



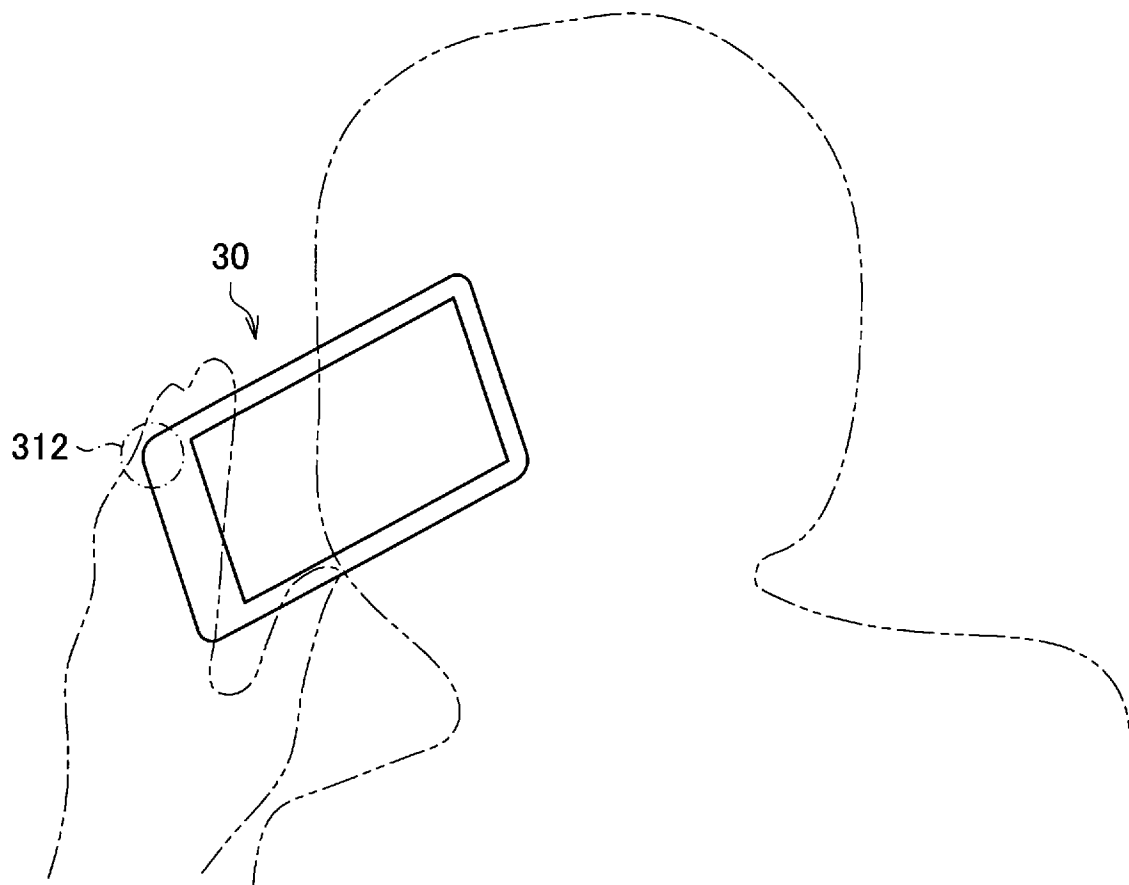
[図15]



[図16]

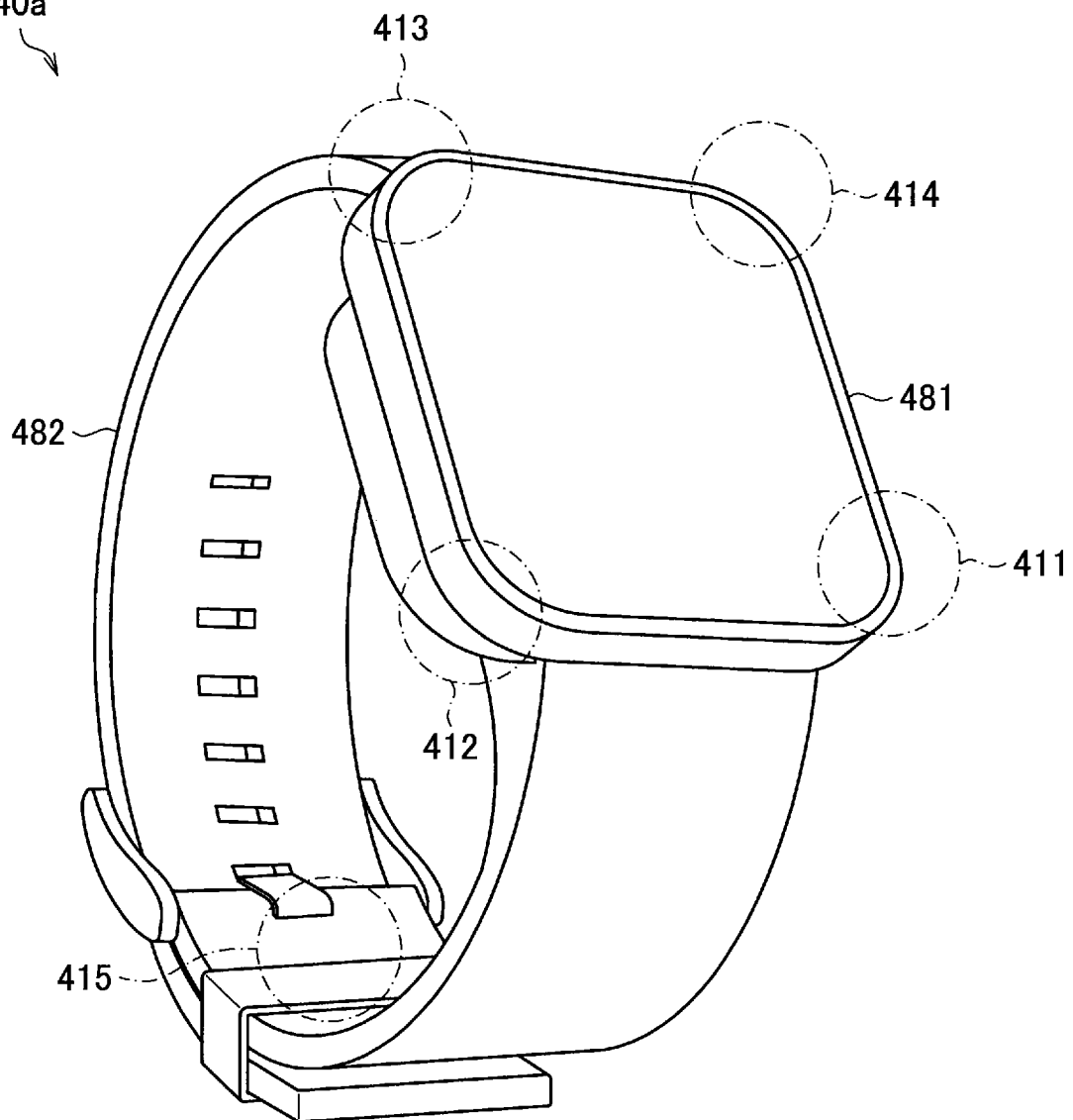


[図17]

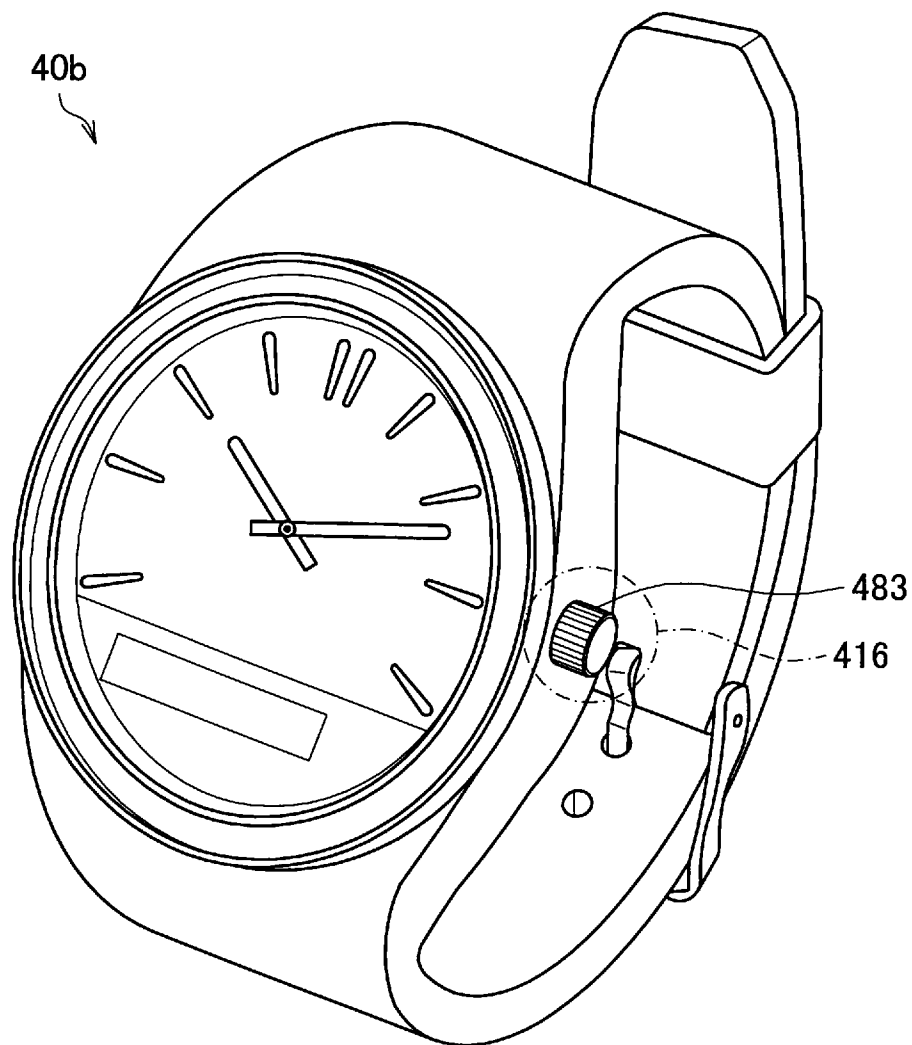


[図18]

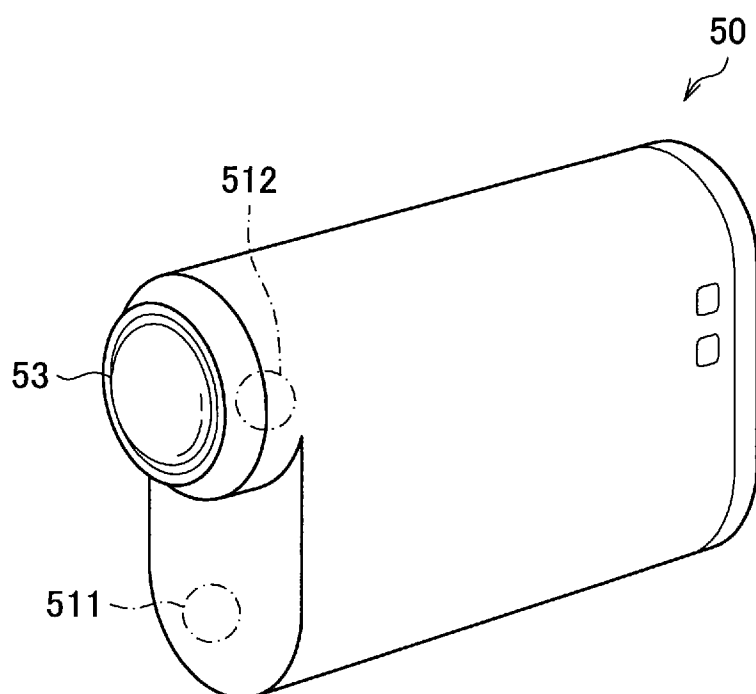
40a



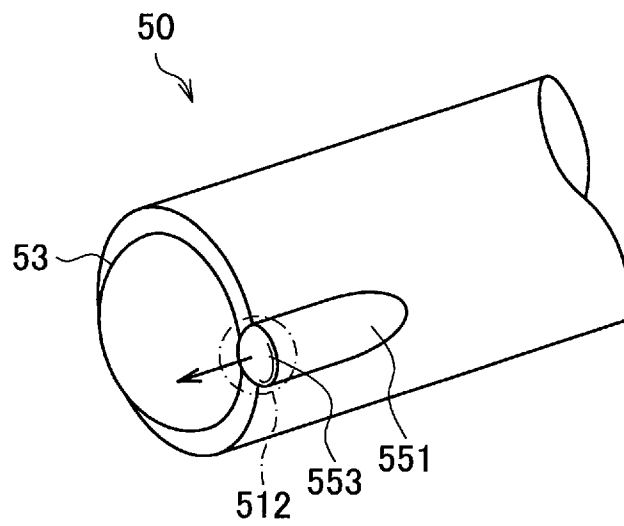
[図19]



[図20]

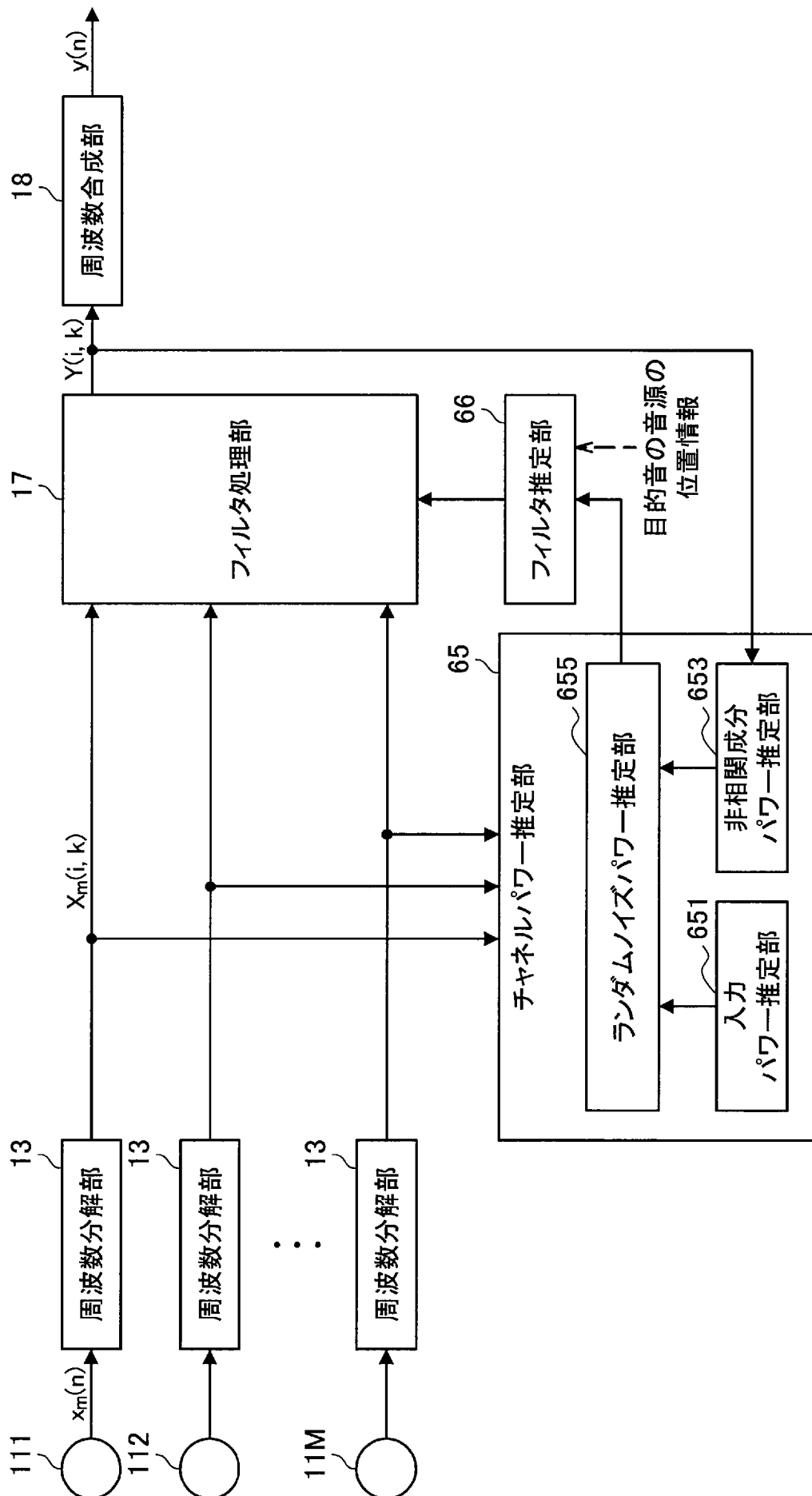


[図21]

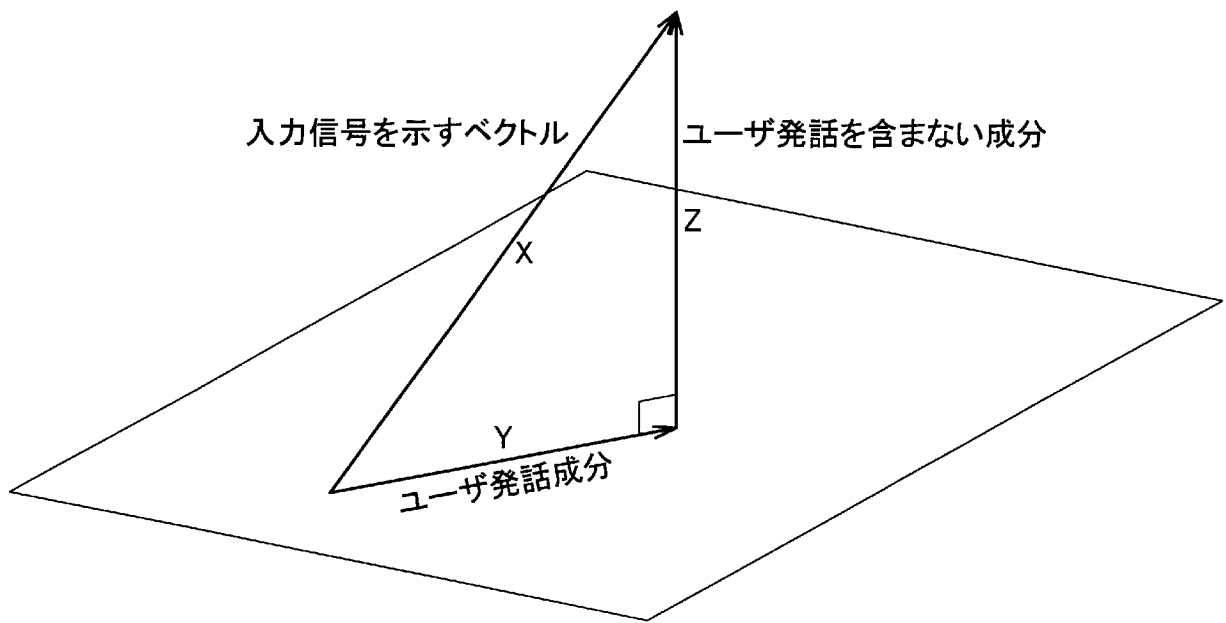


[図22]

60

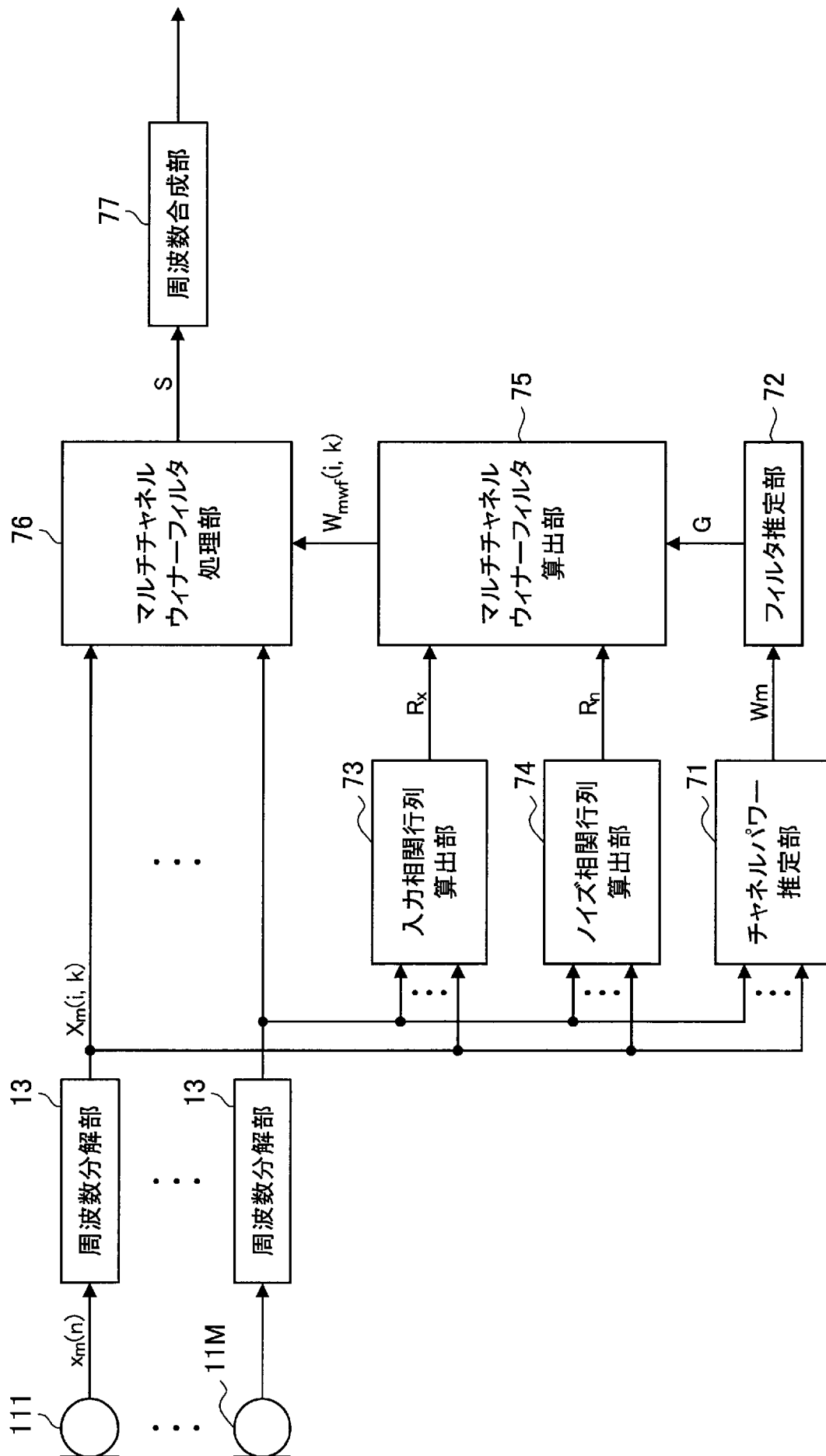


[図23]

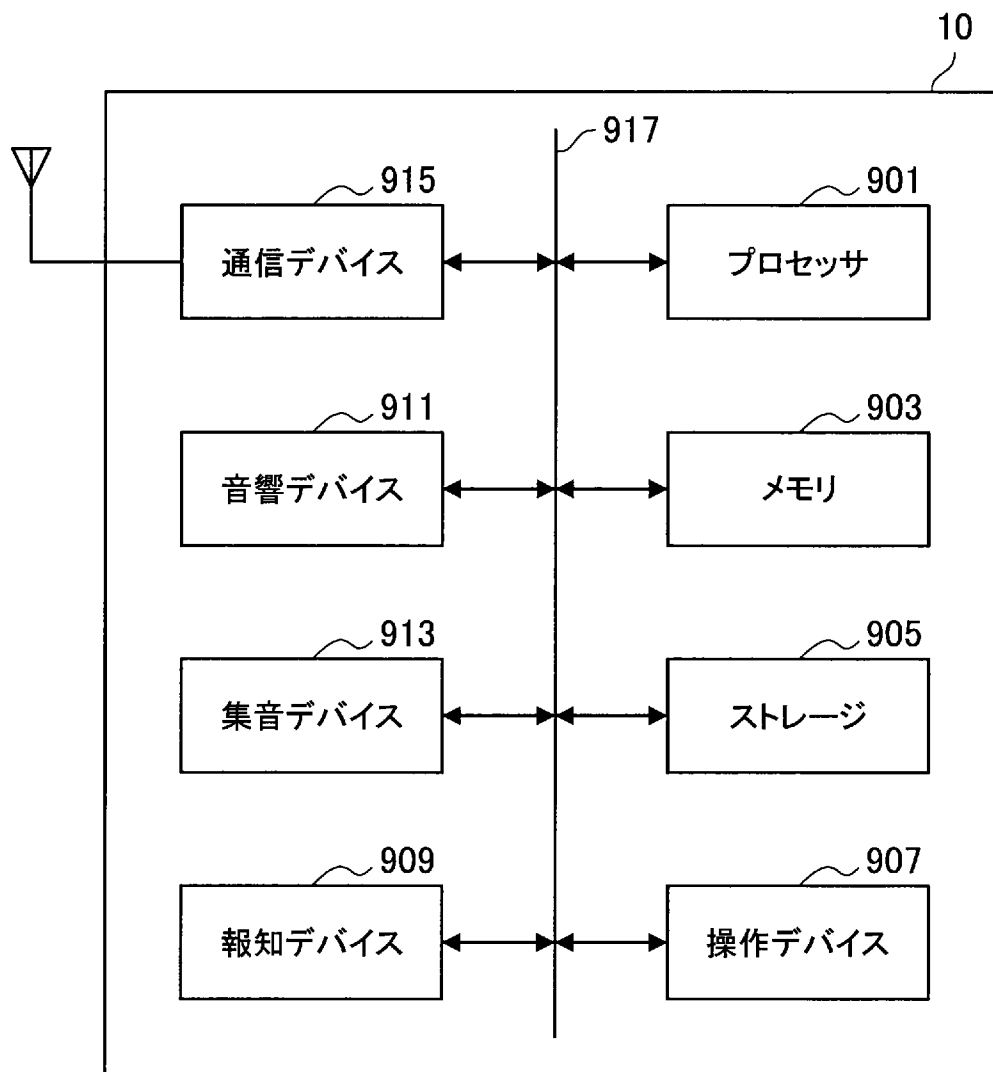


[図24]

70



[図25]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/079855

A. CLASSIFICATION OF SUBJECT MATTER

H04R1/02(2006.01)i, G10L21/0208(2013.01)i, H04M1/00(2006.01)i, H04M1/02(2006.01)i, H04R1/40(2006.01)i, H04R3/00(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)

H04R1/02, G10L21/0208, H04M1/00, H04M1/02, H04R1/40, H04R3/00

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	1996-2016
Kokai Jitsuyo Shinan Koho	1971-2016	Toroku Jitsuyo Shinan Koho	1994-2016

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 Y	JP 2008-507926 A (Softmax, Inc.), 13 March 2008 (13.03.2008), paragraphs [0030] to [0036]; fig. 5 to 8 & US 2008/0201138 A1 paragraphs [0048] to [0054]; fig. 5 to 8 & WO 2006/028587 A2 & CN 101031956 A	1-11, 17-18 12-13
X Y	WO 2015/003220 A1 (WOLFSON DYNAMIC HEARING PTY LTD.), 15 January 2015 (15.01.2015), paragraph [0030]; fig. 1b & US 2016/0155453 A1	1-3, 7-11, 14-16 12-13

☒ 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
02 December 2016 (02.12.16)

Date of mailing of the international search report
13 December 2016 (13.12.16)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/079855

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 51-000248 Y1 (Matsushita Electric Industrial Co., Ltd.), 07 January 1976 (07.01.1976), page 1, right side, the last line; fig. 4 (Family: none)	1
X	JP 2014-045236 A (Aco Co., Ltd.), 13 March 2014 (13.03.2014), paragraph [0017]; fig. 1 (Family: none)	1
A	US 5793875 A (CARDINAL SOUND LABS, INC.), 11 August 1998 (11.08.1998), fig. 12 & WO 1997/040645 A1	5-6
A	JP 2012-133250 A (Sony Corp.), 12 July 2012 (12.07.2012), fig. 8 to 15 & US 2012/0162259 A1 fig. 8 to 15 & EP 2469323 A1 & CN 102543099 A	17-18
A	JP 2005-534269 A (Oakley, Inc.), 10 November 2005 (10.11.2005), fig. 12 & US 2004/0160571 A1 & WO 2004/012477 A2	17-18
A	JP 2014-023141 A (Satoru KATSUMATA), 03 February 2014 (03.02.2014), fig. 2 to 4 (Family: none)	16
Y	Toby Christian Lawin-ore, Simon Doclo, Reference Microphone Selection for MWF-based Noise Reduction Using Distributed Microphone Arrays, Proc. 10.ITG Symposium on Speech Communication, 2012.09.25, ISBN: 978-3-8007-3455-9	12-13
A	JP 2007-519342 A (Koss Corp.), 12 July 2007 (12.07.2007), fig. 1 & US 2005/0153750 A1 fig. 1 & WO 2005/071929 A1	5-6

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. H04R1/02(2006.01)i, G10L21/0208(2013.01)i, H04M1/00(2006.01)i, H04M1/02(2006.01)i,
H04R1/40(2006.01)i, H04R3/00(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（IPC））

Int.Cl. H04R1/02, G10L21/0208, H04M1/00, H04M1/02, H04R1/40, H04R3/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2016年
日本国実用新案登録公報	1996-2016年
日本国登録実用新案公報	1994-2016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 Y	JP 2008-507926 A（ソフトマックス、インク）2008.03.13, [0030] - [0036], [図5] - [図8] & US 2008/0201138 A1, [0048]-[0054], Figs.5-8 & WO 2006/028587 A2 & CN 101031956 A	1-11, 17-18 12-13
X Y	WO 2015/003220 A1 (WOLFSON DYNAMIC HEARING PTY LTD) 2015.01.15, [0030], Fig.1b & US 2016/0155453 A1	1-3, 7-11, 14-16 12-13

☑ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「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
「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

02.12.2016

国際調査報告の発送日

13.12.2016

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

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）

千本 潤介

5Z

3457

電話番号 03-3581-1101 内線 3591

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
X	JP 51-000248 Y1 (松下電器産業株式会社) 1976.01.07, 第1頁右側の最終行、第4図 (ファミリーなし)	1
X	JP 2014-045236 A (株式会社アコー) 2014.03.13, [0017]、[図1] (ファミリーなし)	1
A	US 5793875 A (CARDINAL SOUND LABS, INC.) 1998.08.11, Fig.12 & WO 1997/040645 A1	5-6
A	JP 2012-133250 A (ソニー株式会社) 2012.07.12, [図8] - [図15] & US 2012/0162259 A1, Figs.8-15 & EP 2469323 A1 & CN 102543099 A	17-18
A	JP 2005-534269 A (オークレイ・インコーポレイテッド) 2005.11.10, Fig.12 & US 2004/0160571 A1 & WO 2004/012477 A2	17-18
A	JP 2014-023141 A (勝又 哲) 2014.02.03, [図2] - [図4] (ファミリーなし)	16
Y	Toby Christian Lawin-ore, Simon Doclo, Reference Microphone Selection for MWF-based Noise Reduction Using Distributed Microphone Arrays, Proc. 10. ITG Symposium on Speech Communication, 2012.09.25, ISBN: 978-3-8007-3455-9	12-13
A	JP 2007-519342 A (コス コーポレイション) 2007.07.12, [図1] & US 2005/0153750 A1, Fig.1 & WO 2005/071929 A1	5-6