

(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
2010 年 9 月 30 日(30.09.2010)

PCT

(10) 国際公開番号
WO 2010/109708 A1

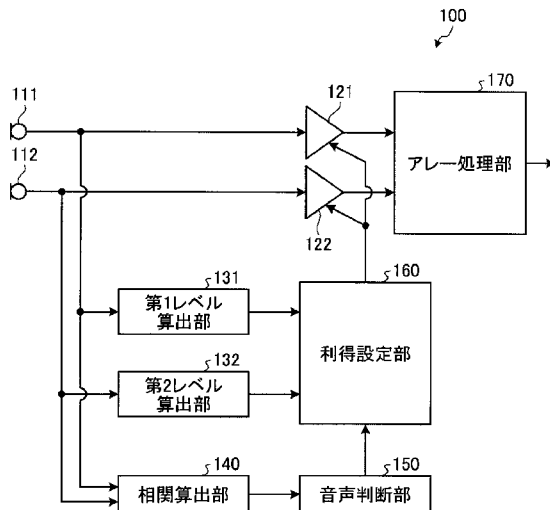
- (51) 国際特許分類:
H04R 3/00 (2006.01) H04R 1/40 (2006.01)
G10L 11/02 (2006.01)
- (21) 国際出願番号: PCT/JP2009/067709
- (22) 国際出願日: 2009 年 10 月 13 日(13.10.2009)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2009-074900 2009 年 3 月 25 日(25.03.2009) JP
- (71) 出願人 (米国を除く全ての指定国について): 株式会社東芝 (Kabushiki Kaisha Toshiba) [JP/JP]; 〒1058001 東京都港区芝浦一丁目 1 番 1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 天田 皇 (AMADA, Tadashi) [JP/JP]; 〒1058001 東京都港区芝浦一丁目 1 番 1 号 株式会社東芝 知的財産部内 Tokyo (JP).
- (74) 代理人: 酒井 宏明, 外(SAKAI, Hiroaki et al.); 〒1006020 東京都千代田区霞が関三丁目 2 番 5 号 霞が関ビルディング 酒井国際特許事務所 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (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,

[続葉有]

(54) Title: PICKUP SIGNAL PROCESSING APPARATUS, METHOD, AND PROGRAM

(54) 発明の名称: 受信信号処理装置、方法およびプログラム

[図1]



(57) Abstract: A pickup signal processing apparatus, provided with: a voice judgment unit (150) which judges whether pickup signals picked up by microphones (111, 112) are voice signals which include voice from a nearby sound source or are background noise signals which do not include voice; signal level calculation units (131, 132) which calculate the respective signal levels of the plurality of pickup signals that the microphones pick up; a setting unit (160) which, in the case that the pickup signals are judged to be background noise signals, determines from the signal levels of the plurality of pickup signals, a gain value reducing the difference of signal levels between the microphones, and which sets that value as the gain value of the pickup signal of at least one of the microphones, said gain value being a value that the pickup signal of at least one microphone among the microphones (111, 112) is to be multiplied by; and computing units (121, 122) which multiply the pickup signal of at least one of the microphones by the gain value.

(57) 要約:

[続葉有]

- 170 array processing unit
131 first level calculation unit
132 second level calculation unit
140 correlation calculation unit
150 voice judgment unit
160 gain setting unit

WO 2010/109708 A1



NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— 補正された請求の範囲及び説明書（条約第 19 条(1)）

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

受音信号処理装置は、マイクロホン 111, 112 が受音した受音信号が、近接音源からの音声を含む音声信号であるか、音声を含まない背景雑音信号かを判断する音声判断部 150 と、マイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部 131, 132 と、受音信号が背景雑音信号であると判断された場合に、複数の受音信号それぞれの信号レベルに基づいて、マイクロホン 111, 112 のうち少なくとも 1 つのマイクロホンの受音信号に乗じるべき利得値であって、マイクロホンの間の信号レベルの差を減少させる利得値を決定し、少なくとも 1 つのマイクロホンの受音信号の利得値として設定する設定部 160 と、少なくとも 1 つのマイクロホンの受音信号に、利得値に乗じる演算部 121, 122 とを備える。

明 細 書

発明の名称： 受音信号処理装置、方法およびプログラム

技術分野

[0001] 本発明は、複数のマイクロホンが取得した受音信号を処理する受音信号処理装置、方法およびプログラムに関する。

背景技術

[0002] 近年、複数のマイクロホンを用いて、特定の方向から到来する信号を強調し、その他の音を抑圧する技術や、音源の方向を検出する技術の研究が盛んである。代表的なマイクロホンアレー方式として、遅延和アレーがあげられる（非特許文献１）。この方法は、各マイクロホンの信号に所定の遅延を挿入し加算処理を行うと、事前に設定された方向から到来した信号のみが同位相で足し合わされ強調されるのに対し、その他の方向から到来した信号は位相が揃わず弱め合うという原理に基づいている。遅延和アレーでは、この原理に基づき加算処理を行うことにより、特定の方向からの信号を強調する。すなわち、特定の方向に指向性を形成する。遅延和アレーにより得られる出力信号 $Y(t)$ は、（式１）で表される。

[数1]

$$Y(t) = \sum_{n=1}^N X_n(t + n\tau) \quad \dots (\text{式 } 1)$$

（式１）において、 N はマイクロホンの個数、 $X_n(t)$ は、各マイクロホンで得られた受音信号であり、 $n = 1 \sim N$ である。マイクロホンは等間隔に添え字 n の順に配置されているものとする。また、 τ は、目的音の到来方向に受音信号を同相化するための遅延時間である。

[0003] マイクロホンアレー方式の別の例としては、Griffith-Jim型アレーがあげられる（非特許文献２）。Griffith-Jim型アレーは適応フィルタを用いて妨害音を除去する方式である。例えば、２つのマイクを利用したGriffith-Jim型アレーにおいて、目的音のアレー

の正面から到来し、妨害音がアレーの側方から到来するとする。この場合、正面から到来する目的音は左右のマイクに同相で受音される。その結果、加算部では前述の遅延和アレーと同じ原理で目的音は強調される。一方、減算部では目的音は同相で減算されるため消去される。妨害音はマイク間で位相がそろっていないため、加算部、減算部のいずれでも強調もされなければ消去もされずに出力される。ここで、ポイントになるのが減算部の出力信号が目的音を除いた、いわゆる雑音成分のみから成る点である。Griffiths-Jim型アレーではこの出力信号を参照信号として適応フィルタを駆動し、加算部の出力に残留している雑音成分を除去することにより、目的音の強調を行う。

先行技術文献

非特許文献

- [0004] 非特許文献1: J. L. Flanagan, J. D. Johnston, R. Zahn and G. W. Elko, "Computer-steered microphone arrays for sound transduction in large rooms," J. Acoust. Soc. Am., vol. 78, no. 5, pp. 1508-1518, 1985

非特許文献2: L. J. Griffiths and G. W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," IEEE Trans. Antennas & Propagation, Vol. AP-30, No. 1, Jan., 1982

発明の概要

発明が解決しようとする課題

- [0005] このようなアレー処理においては、複数のマイクロホンの感度が同一であることが前提となっている。しかしながら、実際にはマイクロホンの感度にはバラつきがあり、また経時変化も無視できない。このため、常に同一感度を維持することは困難である。感度が不揃いなマイクロホンを用いてアレーを構成すると設計通りの指向性を形成することができない。例えばGriffiths-Jim型アレーでは、減算部で目的音を除去する構成になってい

るが、2つのマイクロホンの感度が異なると同相で減算しても振幅の差分が消し残ってしまう。この消し残しは適応フィルタに供給される。この適応フィルタを用いた場合には、加算部の出力から目的音成分を一部除去することとなり、最終的な出力信号に歪みを生じる「目的音除去」という致命的な問題が発生してしまう。

[0006] 本発明は、上記に鑑みてなされたものであって、マイクロホンアレーを構成するマイクロホンの感度を補正することのできる受信信号処理装置、方法およびプログラムを提供することを目的とする。

課題を解決するための手段

[0007] 上述した課題を解決し、目的を達成するために、本発明は、音声を受音する複数のマイクロホンと、前記複数のマイクロホンが受音した受信信号が、前記マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを、受信信号に基づいて判断する音声判断部と、前記複数のマイクロホンが受音した複数の受信信号それぞれの信号レベルを算出する信号レベル算出部と、前記音声判断部において前記受信信号が前記背景雑音信号であると判断された場合に、前記複数の受信信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受信信号に乗じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受信信号の前記利得値として設定する設定部と、前記少なくとも1つのマイクロホンの前記受信信号に、前記設定部によって設定された前記利得値に乗じる演算部とを備えたことを特徴とする。

[0008] また、本発明の他の形態は、予め定められた規定位置に設置され、音声を受音する複数のマイクロホンと、前記複数のマイクロホンが受信した受信信号が、マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを、受信信号に基づいて判断する音声判断部と、前記複数のマイクロホンが受音した複数の受信信号それぞれの信

号レベルを算出する信号レベル算出部と、前記音声判断部において前記受信信号が音声信号であると判断された場合に、前記複数の受信信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受信信号に乘じるべき利得値であって、前記複数のマイクロホンそれぞれが受信する複数の受信信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマイクロホンによる前記複数の受信信号の理想的なレベルバランスに近づける利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受信信号の利得値として設定する設定部と、前記少なくとも1つのマイクロホンの前記受信信号に、前記設定部によって設定された前記利得値を乗じる演算部とを備えたことを特徴とする。

発明の効果

[0009] 本発明によれば、各マイクロホンの受信信号に乘じるべき利得値を自動的に継続して更新することができる。さらに、受信信号が背景雑音信号である場合に限り、利得値の調整を行うので、音声信号を利用することにより、不適切な利得値の調整を行うことがなく適切な利得値を設定することができるという効果を奏する。

図面の簡単な説明

- [0010] [図1] 受信信号処理装置 100 の構成を示すブロック図。
[図2] マイクロホンと音源の配置例を示す図。
[図3] マイクロホンと音源の配置例を示す図。
[図4] 受信信号処理装置 100 における受信信号処理を示すフローチャート。
[図5] 第5の変更例にかかる受信信号処理装置 101 の構成を示すブロック図。
[図6] 受信信号処理装置 102 の構成を示すブロック図。
[図7] 第1処理部 211 の構成を示すブロック図。
[図8] 第3の実施の形態にかかる受信信号処理装置 103 の構成を示すブロック図。

[図9] 受信信号処理装置 104 の構成を示すブロック図。

[図10] 受信信号処理装置 105 の構成を示すブロック図。

[図11] 受信信号処理装置 106 の構成を示すブロック図。

発明を実施するための形態

[0011] 以下に添付図面を参照して、この発明にかかる受信信号処理装置、方法およびプログラムの最良な実施の形態を詳細に説明する。

[0012] 図1は、本発明の第1の実施の形態にかかる受信信号処理装置100の構成を示すブロック図である。本実施の形態にかかる受信信号処理装置100は、2つのマイクロホンを含むマイクロホンアレーにおける受信信号処理を行う。なお、マイクロホンアレーを構成するマイクロホンの個数は2つに限定されるものではなく、3つ以上のマイクロホンを有してもよい。

[0013] 受信信号処理装置100は、第1マイクロホン111と、第2マイクロホン112と、第1利得演算部121と、第2利得演算部122と、第1レベル算出部131と、第2レベル算出部132と、相関算出部140と、音声判断部150と、利得設定部160と、アレー処理部170とを備えている。

[0014] 第1マイクロホン111および第2マイクロホン112は、マイクロホンアレーを構成し、それぞれ受信信号を取得する。第1マイクロホン111が取得した受信信号は、第1利得演算部121、第1レベル算出部131および相関算出部140に入力される。第2マイクロホン112が取得した受信信号は、第2利得演算部122、第2レベル算出部132および相関算出部140に入力される。

[0015] 第1利得演算部121は、第1マイクロホン111が取得した受信信号に対し利得値を乗じる。第2利得演算部122は、第1マイクロホン111が取得した受信信号に対し利得値を乗じる。これにより、マイクロホンアレーを構成する複数のマイクロホンの感度の差を補正することができる。なお、第1利得演算部121および第2利得演算部122が利用する利得値は、利得設定部160により設定される。

[0016] 第1レベル算出部131は、第1マイクロホン111が取得した受信信号の信号レベルを算出する。第2レベル算出部132は、第2マイクロホン112が取得した受信信号の信号レベルを算出する。具体的には、第1レベル算出部131および第2レベル算出部132は、それぞれ(式2)により信号パワーの平均値 L_n を信号レベルとして算出する。

[数2]

$$L_n = E\{X_n(t)^2\} \quad (n=1,2) \quad \dots (式2)$$

(式2)において、 $E\{\}$ は、期待値を表し、時間平均により算出する。 X は、受信信号、 t は時間インデックス、 n はマイクロホンを識別する識別情報、すなわちチャネル番号を表している。なお、第1レベル算出部131および第2レベル算出部132は、それぞれ予め設定されているレベル算出時間周期で定期的に信号レベル算出を行う。

[0017] 他の例としては、(式3)により再帰平均 $L_n(t)$ を信号レベルとして算出してもよい。

[数3]

$$L_n(t) = (1-\alpha)L_n(t-1) + \alpha X_n(t)^2 \quad \dots (式3)$$

(式3)において、 α は1より小さな正の値である。

[0018] また、他の例としては、信号パワーの平均値と再帰平均とを組み合わせ、時間窓の平均パワーに対して再帰平均を適用してもよい。また、受信信号の2乗に替えて、振幅を用いることとしてもよい。また、平均値に替えて、最大値を用いてもよい。このように、受信信号の信号レベルは既存の技術を用いて算出すればよく、その方法は本実施の形態に限定されるものではない。

[0019] 相関算出部140は、予め設定された相関算出時間周期で定期的に、第1マイクロホン111および第2マイクロホン112から受信信号を取得し、これらの相関を求める。第1マイクロホン111および第2マイクロホン112から取得した受信信号をそれぞれ $X_1(t)$ 、 $X_2(t)$ とすると、 X

1 (t) と X 2 (t) の相互相関 R 1 2 は、(式 4) で定義される。

[数4]

$$R_{12}(\tau) = E\{X_1(t) * X_2(t + \tau)\} \quad \dots (式 4)$$

相関算出部 1 4 0 は、窓幅 T での相関を信号のパワーで正規化した正規化相互相関関数 r 1 2 により X 1 (t) と X 2 (t) の相関を算出する。r の添え字 1, 2 は、それぞれチャネル番号を表している。相関算出部 1 4 0 は具体的には、(式 5) により時刻 t 0 における X 1 (t) と X 2 (t) の相関 r 1 2 を算出する。

[数5]

$$r_{12}(t_0, \tau) = \phi_{12}(t_0, \tau) / \text{sqrt}(P_{11}(t_0) * P_{22}(t_0 + \tau)) \quad \dots (式 5)$$

ここで、φ 1 2 は、(式 6) により算出される。また、P i i は、(式 7) により算出される。

[数6]

$$\phi_{12}(t_0, \tau) = \sum_{t=t_0-T/2}^{t_0+T/2} X_1(t) * X_2(t + \tau) \quad \dots (式 6)$$

[数7]

$$P_{ii}(t_0) = \sum_{t=t_0-T/2}^{t_0+T/2} X_i(t)^2 \quad \dots (式 7)$$

なお、φ の添え字 1, 2 および P の添え字 i はそれぞれチャネル番号を表している。正規化相互相関関数では値が 0 ~ 1 に正規化される。このため、相関の強さを表す指標として用いるのに便利である。なお、マイクロホンの数が 3 以上の場合、すなわち 3 チャネル以上の場合には、2 つのマイクロホン、すなわち 2 チャネルの相関値の統合により求めることができる。

[0020] 相関算出部 1 4 0 は、3 以上のチャネルにおける全チャネルの組み合わせを用いる場合は、(式 8) により相関 r m (t 0, τ) を算出する。

[数8]

$$rm(t_0, \tau) = \sum_{i < j} \phi_{ij}(t_0, \tau) / \sum_{i < j} \text{sqrt}(P_{ii}(t_0) * P_{jj}(t_0 + \tau)) \quad \dots (式 8)$$

[0021] 他の例としては、全チャネルの統合（ $i < j$ ）に替えて、隣接チャネルの統合（ $j = i + 1$ ）のように、他の統合方法を用いても良い。なお、以下では、簡単のため2チャネルの正規化相互相関関数 $r_{12}(t_0, \tau)$ を用いる場合について説明するが、3チャネル以上の場合も同様である。

[0022] 相関算出部 140 は、異なる τ の値に対する複数の相関値を算出し、 τ に関する相関値の最大値 $r_{12_max}(t_0, \tau_max)$ を特定する。相関値が大きいことは、相関の大きい信号が到来していることを意味し、またこのときの τ_max は、これらの信号が2つのマイクロホンに到達するまでの時間差、すなわち音源方向を示している。なお、相関算出部 140 は、算出規定時間周期で観測時刻 t_0 を設定し、各時刻 t_0 に対して算出された相関値の最大値 r_{12_max} を特定し、特定するごとに音声判断部 150 に出力する。

[0023] なお、第1レベル算出部 131 および第2レベル算出部 132 による信号レベル算出のタイミングであるレベル算出時間周期と、相関算出部 140 による相関算出のタイミングである相関算出時間周期は等しいことが望ましいが、互いに近いタイミングで信号レベルおよび相関が算出されていればよく、必ずしも一致する必要はない。

[0024] 一般的に音源がマイクロホンアレーから遠ざかるに従い、チャネル間の相関は減少する。このため、チャネル間の相関を手がかりに近接音源の存在を検出することが可能である。音声信号のように時間的に不連続な信号を扱う場合、音声信号が存在する音声信号区間と、音声信号の存在しない区間、すなわち背景雑音信号の区間である背景雑音区間とが存在する。ここで、音声信号とは近接音源から発せられた音声を含む信号である。すなわち、近接音源とは、マイクロホンアレーが音声として認識可能な音を発する音源である。背景雑音信号とは、近接音源からの音声信号が存在しない場合に、マイク

ロホンアレーが受音する雑音信号である。例えば、ドライバーの声を受音することを目的として設定されたマイクロホンアレーにおいて、助手席に座っている人物の声の信号も、マイクロホンアレーに対する近接音源からの信号であり、音声信号である。一方、例えば遠くを走行する救急車のサイレンの信号は、近接音源からの信号ではなく、背景雑音信号である。

[0025] 受音信号がマイクロホンアレーに近接する近接音源から発せられた音声信号である場合には、チャンネル間の相関は大きくなる。一方、受音信号が背景雑音のみを含む背景雑音信号である場合には、チャンネル間の相関は小さくなる。そこで、本実施の形態においては、相関の最大値 r_{12_max} を算出し、相関の最大値 r_{12_max} を用いて受音信号が音声信号であるか背景雑音信号であるかを判断する。

[0026] 音声判断部 150 は、相関算出部 140 から相関の最大値 r_{12_max} を取得する。そして、予め設定された相関値の閾値 r_{12_th} と比較し、最大値 r_{12_max} が閾値 r_{12_th} に比べて小さい場合には、相関が小さく、受音信号は背景雑音信号であると判断する。また、最大値 r_{12_max} が閾値 r_{12_th} 以上である場合には、相関が大きく、受音信号は音声信号であると判断する。なお、閾値 r_{12_th} は、実験により求めた値である。実験においては、背景雑音および音声に対する受音信号を測定し、これらの測定結果から閾値を算出する。なお、受音信号が背景雑音信号であるか音声信号であるかをより正確に判断するためには、受音信号処理装置 100 が設置される環境にできるだけ近い環境において測定を行うのが望ましい。

[0027] 利得設定部 160 は、予め設定された利得設定時間周期で音声判断部 150 から受音信号が音声信号であるか背景雑音信号であるかの判断結果を取得する。利得設定部 160 は、また第 1 レベル算出部 131 および第 2 レベル算出部 132 から第 1 マイクロホン 111 および第 2 マイクロホン 112 の受音信号の信号レベルを取得する。利得設定部 160 は、受音信号が背景雑音信号である場合には、第 1 マイクロホン 111 および第 2 マイクロホン 1

12それぞれが取得した受音信号の信号レベルに基づいて、各受音信号に乘じるべき利得値を決定する。利得設定部160は、第1マイクロホン111が取得した受音信号に対して決定した利得値を第1利得演算部121に設定し、第2マイクロホン112が取得した受音信号に対して決定した利得値を第2利得演算部122に設定する。

[0028] 例えば、受音信号の平均パワーが $L_1 < L_2$ の場合、第2利得演算部122に設定されているチャンネル2の利得を減少させ、第1利得演算部121に設定されているチャンネル1の利得を増加させる。これにより、2つのマイクロホンの感度差を減少させる方向に利得値を更新することができる。具体的には、利得設定部160は、(式9)および(式10)に示す利得を各チャンネルの利得演算部に設定する。なお、チャンネル n に現在設定している利得値を G_{n_old} 、利得設定部160が新たにチャンネル n の利得演算部に設定する利得値を G_{n_new} とする。

[数9]

$$G_{1_new} = G_{1_old} * \sqrt{L_x/L_1} \quad \dots (式9)$$

[数10]

$$G_{2_new} = G_{2_old} * \sqrt{L_x/L_2} \quad \dots (式10)$$

なお、 L_x は、平均パワーの目標値であり、(式11)で表される。

[数11]

$$L_x = (L_1 + L_2) / 2 \quad \dots (式11)$$

[0029] 利得設定部160は、第1レベル算出部131および第2レベル算出部132から取得した受音信号の信号レベルに基づいて算出した新たな利得値 G_{1_new} 、 G_{2_new} をそれぞれ第1利得演算部121および第2利得演算部122に設定する。これにより、第1マイクロホン111および第2マイクロホン112が取得した受音信号の感度、すなわち信号レベルの差が小さくなるように、より好ましくは等しくなるように信号レベルを調整することができる。

- [0030] 受音信号の利得を調整して感度補正を行うだけならば、目標レベル（例えば基準マイクのレベル）になるように、各マイクロホンの利得を独立に制御する方法が考えられる。しかしながら、この方法には問題がある。図2に示す配置例では、マイクロホンアレー111, 112の正面、すなわち各マイクロホン111, 112からの距離が等しい位置に音源11, 12がある。この場合、各音源11, 12と2つのマイクロホン111, 112の間の距離の比（ d_{11}/d_{12} および d_{21}/d_{22} ）は音源11, 12とマイクロホン111, 112の間の距離によらず1である。
- [0031] 図3に示す例では、マイクロホンアレー111, 112の斜め方向に音源13, 14がある。この場合には、2つのマイクロホン111, 112までの距離の比（ d_{31}/d_{32} および d_{41}/d_{42} ）は音源距離によって異なる。すなわち、マイクロホン111, 112と音源13, 14の間の距離が大きくなるほど音源13, 14からマイクロホン111, 112までの距離の比が1に近づくのに対し、マイクロホン111, 112と音源13, 14の間の距離が小さくなるほど音源13, 14からマイクロホン111, 112までの距離の比は1よりも大きくなる。
- [0032] 一般に、マイクロホンで受音した音波のエネルギーは、音源からの距離の2乗に反比例する。したがって、距離の比が大きくなるにつれて受音信号のパワーの比も大きくなる。すなわち、音源がマイクロホンアレーの近くであって、かつ斜め方向に存在する場合には、複数のマイクロホンの感度が等しければ各マイクロホンは異なる信号パワー、すなわち信号レベルの受音信号を取得するはずである。このようにマイクロホン毎に異なるべき信号レベルをすべて等しくなるように利得調整を行うことは、感度の等しいマイクロホンを用いた場合に得られる受音信号とは異なる受音信号に調整することになってしまう。
- [0033] 例えば、自動車内でドライバーの声を受音するために、ルームミラーにマイクロホンアレーを設置する場合がある。この場合、主な音源であるドライバーはマイクロホンアレーに対し斜め方向に存在する。単純にマイクロホン

間の信号パワーが等しくなるように利得を調整すると、ドライバーの発話時に、ドライバーにより近いマイクロホンほど大きな信号を出力するという現象と一致しなくなってしまう。また、使用中に同乗者など他の方向に音源が現れると、その都度、音源方向に逆らうように利得調整を行うことになる。しかしながら、これはマイクロホンの感度をそろえることにはならず、適切な利得調整を行うことはできない。

[0034] そこで、上述のように利得設定部 160 は、近接音源が存在しない場合、すなわち受信信号が背景雑音信号である場合に限り、新たな利得値を算出し、これを第 1 利得演算部 121 および第 2 利得演算部 122 に設定する。これにより、本来異なるべき信号パワーを等しくするような、不適切な利得調整を行うのを防ぐことができる。

[0035] アレー処理部 170 は、利得設定部 160 により設定された利得値により第 1 利得演算部 121 および第 2 利得演算部 122 において調整された後の受信信号を用いてアレー処理を行う。なお、アレー処理としては、Griffith-Jim 型アレーによる処理を行う。なお、他の例としては、アレー処理部 170 は、遅延和アレーや ICA など、複数のマイクロホンを用いた信号処理を行ってもよい。アレー処理部 170 は、第 1 利得演算部 121 および第 2 利得演算部 122 により信号レベルが調整された受信信号を利用して処理を行うので、設計通りの指向性を形成することができる。

[0036] 図 4 は、受信信号処理装置 100 における受信信号処理を示すフローチャートである。まず、マイクロホンアレーを形成する第 1 マイクロホン 111 および第 2 マイクロホン 112 は、受信信号を取得する（ステップ S100）。次に、第 1 レベル算出部 131 および第 2 レベル算出部 132 は、それぞれレベル算出時間が経過する度に、第 1 マイクロホン 111 および第 2 マイクロホン 112 が取得した受信信号の信号レベルを算出する（ステップ S102）。相関算出部 140 は、相関算出時間が経過する度に、第 1 マイクロホン 111 が取得した受信信号および第 2 マイクロホン 112 が取得した受信信号の相関値を算出し、相関の最大値 r_{12_max} を音声判断部 15

0に出力する（ステップS104）。

[0037] 音声判断部150は、相関算出部140から取得した最大値 r_{12_max} と、予め設定されている閾値 r_{12_th} とを比較する。最大値 r_{12_max} が閾値 r_{12_th} よりも小さい場合には（ステップS106, Yes）、受信信号は背景雑音信号であると判断する。一方、最大値 r_{12_max} が閾値 r_{12_th} 以上である場合には（ステップS106, No）、受信信号は音声信号であると判断する。

[0038] 利得設定部160は、利得設定時間が経過する度に、音声判断部150から判断結果を取得する。算出された相関の最大値 r_{12_max} が閾値 r_{12_th} よりも小さい場合には（ステップS106, Yes）、受信信号は背景雑音信号であるとの判断結果を取得する。この場合、利得設定部160は、第1利得演算部121および第2利得演算部122に設定されている利得値を更新する（ステップS108）。

[0039] 具体的には、利得設定部160は、第1レベル算出部131および第2レベル算出部132が算出した受信信号の信号レベルに基づいて、第1利得演算部121および第2利得演算部122に設定する新たな利得値 G_{1_new} , G_{2_new} を算出する。そして、算出した新たな利得値を第1利得演算部121および第2利得演算部122にそれぞれ設定する。

[0040] 一方、ステップS106において、最大値 r_{12_max} が閾値 r_{12_th} 以上である場合、すなわち受信信号が音声信号である場合には（ステップS106, No）、利得設定部160は、利得を更新しない。そして第1マイクロホン111および第2マイクロホン112による受信信号の取得が終了していなければ（ステップS110, No）、再びステップS102に戻り、更新処理を継続し、第1マイクロホン111および第2マイクロホン112による受信信号の取得が終了すると（ステップS110, Yes）、処理は完了する。

[0041] このように、第1の実施の形態にかかる受信信号処理装置100においては、背景雑音区間においてのみ利得値の更新を行うので、近接斜方向の音源

が存在する環境下において、音声信号を用いた利得調整により、異なるべき信号パワーを等しい信号パワーに調整するような不適切な利得調整を行うことなく、正しくマイクロホンの感度を合わせることができる。

[0042] また、受信信号処理装置 100 においては、受信信号が背景雑音信号である場合には、利得設定部 160 は、予め設定された利得設定時間が経過する度に必要に応じて利得を更新するので、マイクロホンアレーが作動している間継続して自動的に利得調整を行うことができる。したがって、マイクロホンの経時変化にも対応した利得調整を行うことができる。

[0043] 実施の形態の第 1 の変更例としては、音声判断部 150 は、所定の時間間隔内の複数の t_0 に対して得られた複数の相関値の最大値それぞれと閾値との比較を行い、予め設定された規定連続時間以上の間連続して相関値の最大値が閾値以下である場合に、受信信号が背景雑音であると判断してもよい。これにより、相関値の一時的な変動の影響を受けにくくすることができる。

[0044] 第 2 の変更例としては、利得設定部 160 は、既に第 1 利得演算部 121 および第 2 利得演算部 122 に設定されている利得値 G_{1_old} , G_{2_old} から 1 回の調整量を比較的小さい値にし、算出した新たな利得値である目標利得値まで徐々に更新していくこととしてもよい。これにより、急な感度調整により聴覚的な違和感を与えるのを避けることができる。

[0045] この場合、利得設定部 160 が、設定時間周期で第 1 利得演算部 121 および第 2 利得演算部 122 に設定する、新たな利得値は、(式 12) および (式 13) により示される。

[数12]

$$G_{1_new} = G_{1_old} * G_{up} \quad \dots (式 12)$$

[数13]

$$G_{2_new} = G_{2_old} * G_{down} \quad \dots (式 13)$$

ここで、 G_{up} , G_{down} はそれぞれ、 $G_{up} > 1$, $G_{down} < 1$ なる値である。例えば 1 回の更新時の利得値の変化量が 1 dB up , 1

d B d o w n 程度であれば更新による変化が知覚されることはまずない。このように、1回に変更する調整幅（ステップサイズ）を制限することにより、緩やかにゲイン調整を行うことができる。

[0046] さらに、チャンネル間の信号レベルの差が大きいほど大きい調整幅を設定し、この調整幅ずつ利得値を更新してもよい。これにより、新たな利得値 $G1_new$, $G2_new$ を設定するまでの収束時間を短縮することができる。また、他の例としては、チャンネル間の信号レベルの差が大きいほど利得値の更新を行う時間間隔、すなわち設定時間周期を短くしてもよい。なお、いずれの場合にも、緩やかに利得値を変更している間も、目標利得値の算出を行い、目標利得値を定期的に更新する。

[0047] また、第3の変更例としては、第1の実施の形態においては、受信信号が背景雑音信号である場合には、利得更新を行わないこととしたが、これにかえて更新時のステップサイズを小さくし、利得更新の程度を小さくすることとしてもよい。これにより、緩やかにゲイン調整を行うことができる。

[0048] 第4の変更例について説明する。図2および図3を参照しつつ説明したように、マイクロホンアレーの正面に音源が存在する場合には、音源とマイクロホンアレーの距離によらず、音源と各マイクロホンの間の距離は等しくなる。そこで、受信信号が音声信号であっても、音源がマイクロホンアレーの正面に位置する場合には、利得の更新を行うこととしてもよい。

[0049] 例えば、音声判断部150は、さらに最大の相関値を与える時間差の絶対値 $|\tau_max|$ と所定の閾値 τ_th とを比較する。そして、利得設定部160は、 $|\tau_max| < \tau_th$ の関係にある場合、すなわち、音源がマイクロホンアレーのほぼ正面付近に存在する場合には、利得の更新を行う。なお、ここで閾値 τ_th は、音源がマイクロホンアレーのほぼ正面に位置する場合に得られる τ を実測して求めたものである。

[0050] 図5は、第5の変更例にかかる受信信号処理装置101の構成を示すブロック図である。第5の変更例にかかる受信信号処理装置101においては、第1レベル算出部133および第2レベル算出部134はそれぞれ第1利得

演算部 1 2 3 および第 2 利得演算部 1 2 4 により利得値の演算が施された後の受信信号を取得する。そして、これらの受信信号の信号レベルを算出する。また、相関算出部 1 4 2 は、第 1 利得演算部 1 2 3 および第 2 利得演算部 1 2 4 から受信信号を取得し、これらの受信信号に基づいて相関値を算出し、音声判断部 1 5 2 に送出する。このように、利得調整後の受信信号の信号レベルを利用するので、利得設定部 1 6 2 による（式 9）および（式 10）を利用した相対的な更新の実装を簡単にすることができる。

[0051] さらに、他の例としては、信号レベル算出には利得調整前の受信信号を利用し、相関算出には利得調整後の受信信号を利用してもよい。またこれとは逆に、信号レベル算出には利得調整後の受信信号を利用し、相関算出には、利得調整後の受信信号を利用してもよい。なお、上記変更例は、いずれも他の実施の形態においても同様に適用することができることはいうまでもない。

[0052] 図 6 は、第 2 の実施の形態にかかる受信信号処理装置 1 0 2 の構成を示すブロック図である。第 2 の実施の形態にかかる受信信号処理装置 1 0 2 は、時間信号である受信信号を周波数領域の信号に変換する。そして、各周波数成分に対し、利得調整を行う。

[0053] 受信信号処理装置 1 0 2 は、第 1 マイクロホン 1 1 1 と、第 2 マイクロホン 1 1 2 と、第 1 D F T 2 0 1 と、第 2 D F T 2 0 2 と、第 1 処理部 2 1 1 ~ 第 L 処理部 2 2 0 と、I D F T 2 3 0 とを備えている。第 1 D F T 2 0 1 は、第 1 マイクロホン 1 1 1 が取得した受信信号を周波数領域の信号に変換する。第 2 D F T 2 0 2 は、第 2 マイクロホン 1 1 2 が取得した受信信号を周波数領域の信号に変換する。第 1 D F T 2 0 1 および第 2 D F T 2 0 2 は、受信信号を周波数領域の信号に変換する処理として、具体的には離散フーリエ変換（D F T）を行う。D F T では、所定の時間幅の時間窓を設定する。そして、この時間窓をシフトしながら連続時間信号を処理する。以下、時間窓により切り出される信号の単位をフレームと称する。フレーム毎に L 個の周波数成分が得られる。各周波数成分は、それぞれ第 1 処理部 2 1 1 ~ 第

L 処理部 220 に入力される。

[0054] 第 1 処理部 211 ~ 第 L 処理部 220 は、それぞれ各周波数成分に対する処理を行い、処理後の信号を出力する。なお、第 1 処理部 211 ~ 第 L 処理部 220 は同一の構成であり、第 1 処理部 211 ~ 第 L 処理部 220 には、それぞれ第 1 マイクロホン 111 および第 2 マイクロホン 112 が取得した受信信号の第 1 周波数成分 ~ 第 L 周波数成分が入力される。第 1 処理部 211 ~ 第 L 処理部 220 は、取得した周波数信号に対して利得調整処理を行う。IDFT230 は、各処理部から取得した周波数成分を時間信号に変換し出力する。IDFT230 は、具体的には、逆離散フーリエ変換 (IDFT) を行う。

[0055] 図 7 は、第 1 処理部 211 の構成を示すブロック図である。第 1 処理部 211 には、第 1 DFT201 から、第 1 マイクロホン 111 の受信信号の第 1 周波数成分が入力される。第 1 処理部 211 には、また第 2 DFT202 から第 2 マイクロホン 112 の受信信号の第 1 周波数成分が入力される。第 1 処理部 211 は、これらの周波数信号に対して利得調整処理を行う。

[0056] 第 1 処理部 211 は、第 1 利得演算部 241 と、第 2 利得演算部 242 と、第 1 レベル算出部 251 と、第 2 レベル算出部 252 と、相関算出部 260 と、音声判断部 270 と、利得設定部 280 と、アレー処理部 290 とを備えている。

[0057] 第 1 利得演算部 241 および第 2 利得演算部 242 は、それぞれ第 1 DFT201 および第 2 DFT202 から第 1 周波数成分を取得する。そして、第 1 利得演算部 241 および第 2 利得演算部 242 は、各第 1 周波数成分に対し、利得値を乗じる。なお、第 1 利得演算部 241 および第 2 利得演算部 242 が利用する利得値は、利得設定部 280 により設定される。

[0058] 第 1 レベル算出部 251 および第 2 レベル算出部 252 はそれぞれ第 1 DFT201 および第 2 DFT202 から第 1 周波数成分を取得する。そして、これらの周波数成分の信号レベルを算出する。具体的には、第 1 レベル算出部 251 および第 2 レベル算出部 252 は、それぞれ (式 14) により第

l 周波数成分の信号パワーの平均値 $L_n(1)$ を算出する。ここで、l は、周波数成分番号である。

[数14]

$$L_n(l) = E\{|X_n(l)|^2\} \quad l = 1, 2, \dots, L \quad \dots (式 14)$$

なお、期待値はフレーム平均として算出する。 $X_n(1)$ は複素数であるので、信号パワーの算出には絶対値の 2 乗を用いる。

[0059] 相関算出部 260 は、第 1 DFT 201 および第 2 DFT 202 から第 1 周波数成分を取得し、これらの相関を求める。相関算出部 260 は、周波数成分毎の相関を表す代表的な指標である、コヒーレンスを用いて相関を算出する。具体的には、(式 15) により第 l 周波数成分におけるチャンネル 1, 2 間のコヒーレンスを相関として算出する。ここで、 $\text{conj}()$ は共役複素数を、 $\text{sqrt}()$ は平方根を表している。

[数15]

$$\gamma_{12}(l) = E\{\text{conj}(X_1(l)) * (X_2(l))\} / \text{sqrt}(E\{|X_1(l)|^2\} * E\{|X_2(l)|^2\}) \quad \dots (式 15)$$

コヒーレンスは複素数であり、その絶対値は、0 ~ 1 の範囲の値をとる。絶対値が 1 に近いほど相関が高いことを意味する。

[0060] 音声判断部 270 は、相関算出部 260 により算出された相関値と、予め定めた閾値 r_{12_th} とを比較し、相関算出部 260 により算出された相関値 r_{12} が閾値 r_{12_th} に比べて小さい場合には、相関が小さく、受信信号は背景雑音信号であると判断する。また、相関値 r_{12} が閾値 r_{12_th} 以上である場合には、相関が大きく、受信信号は音声信号であると判断する。なお、閾値 r_{12_th} は実験により求めた値である。このように、コヒーレンスの絶対値が大きいことは、近接音源の存在を示唆しているので、コヒーレンスの絶対値に基づいて、受信信号が背景雑音信号か音声信号かを判断することができる。

[0061] 利得設定部 280 は、音声判断部 270 から受信信号が音声信号であるか

背景雑音信号であるかの判断結果を取得する。利得設定部 280 はまた、第 1 レベル算出部 251 および第 2 レベル算出部 252 から第 1 マイクロホン 111 および第 2 マイクロホン 112 が取得した受信信号の第 1 周波数成分の信号レベルを算出する。利得設定部 280 は、受信信号が背景雑音信号である場合には、第 1 マイクロホン 111 および第 2 マイクロホン 112 の受信信号の第 1 周波数成分の信号レベルに基づいて、各マイクロホンに対応する第 1 周波数成分に対して乗じる利得値を決定し、この値を第 1 利得演算部 241 および第 2 利得演算部 242 に設定する。

[0062] アレー処理部 290 は、第 1 利得演算部 241 および第 2 利得演算部 242 から利得調整後の第 1 周波数成分を取得し、第 1 周波数成分に対するアレー処理を行い、処理後の第 1 周波数成分を IDFT 230 に出力する。

[0063] このように、本実施の形態にかかる受信信号処理装置 102 においては、 L 個の周波数成分それぞれに対して、利得の調整を行うことができる。これにより、マイクロホンの感度差が周波数領域毎に異なる場合には、周波数成分毎にそれぞれ適した値に利得値を調整することができる。

[0064] なお、第 2 の実施の形態にかかる受信信号処理装置 102 のこれ以外の構成および処理は、第 1 の実施の形態にかかる受信信号処理装置 100 の構成および処理と同様である。

[0065] 第 2 の実施の形態にかかる受信信号処理装置 102 の第 1 の変更例としては、所定の周波数成分に対して求めた相関値を用いて、音声信号が背景雑音信号であるか音声信号であるかを判定し、この判定結果を他の周波数成分においても利用することとしてもよい。例えば、特定の周波数に大きなノイズが存在する場合、その周波数で求めた相関値を利用して音声信号か雑音信号かを判定するのは困難である。例えば、音声のような広帯域信号の近接音源が存在する場合には、この存在を検出するために、所定の周波数成分により算出した相関値を利用することができる。

[0066] さらに、低い周波数成分は近接音源の有無に関わらず相関が高くなる。このため、受信信号が音声信号であるか雑音信号であるかの判定精度が低下す

る可能性がある。そこで、比較的低い周波数成分に対応する処理部においては、相関算出部および音声判断部による処理を行わず、比較的高い周波数成分に対する処理部において得られた判断結果を利用することとする。これにより、受信信号が音声信号であるか雑音信号であるかの判断精度を向上させることができる。

[0067] また、第2の変更例としては、受信信号処理装置102は、IDFT230を備えなくともよい。例えば、音声認識などの用途でスペクトル情報のみが必要な場合は、IDFTを行わず周波数成分を出力してもよい。

[0068] 図8は、第3の実施の形態にかかる受信信号処理装置103の構成を示すブロック図である。第3の実施の形態にかかる受信信号処理装置103は、第2の実施の形態にかかる受信信号処理装置102と同様に、各周波数成分に対する利得調整を行う複数の処理部、すなわち第1処理部311～第L処理部320を備えている。ただし、受信信号処理装置103は、各周波数成分に対応する複数の相関算出部および音声判断部を有するのではなく、1つの相関算出部340および1つの音声判断部350を有している。

[0069] 相関算出部340は、第1IDFT201により得られたすべての周波数成分を取得する。さらに、第2IDFT202により得られたすべての周波数成分を取得する。相関算出部340は、取得したすべての周波数成分から、第1マイクロホン111が取得した受信信号と第2マイクロホン112が取得した受信信号の相関を算出する。相関算出部340は、すべての周波数成分を用いて（式16）により、一般化相互相関関数（GCC）を相関値として算出する。

[数16]

$$GCC(\tau) = IDFT\{w(l) * G_{12}(l)\} \quad \dots (式 16)$$

ここで、 $G_{12}(l)$ は $X_1(l)$ と $X_2(l)$ のクロススペクトルである。 $w(l)$ は周波数ごとの重みである。また、クロススペクトルは $E\{conj(X_1(l) * X_2(l))\}$ として期待値を用いる。フレーム毎に独立に求めても良いが、前者のほうが高い精度で得ることができる。 $w(l)$ は、

(式 17) により算出する。 $w(l)$ の決め方により異なる種類の相互相関関数が得られる点が一般化相互相関関数の特徴であり、詳細は、C. H. Knapp and G. C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," IEEE Trans, Acoust., Speech, Signal Processing, Vol. ASSP-24, No. 4, pp. 320-327, 1976に記載されている。

[数17]

$$w(l) = 1/\sqrt{G_{11}(l) * G_{22}(l)} \quad \dots (式 17)$$

GCC (τ) は周波数ごとに重み付けされている点を除いては第 1 の実施の形態において説明した相互相関関数 $R_{12}(\tau)$ と同じ性質の関数である。したがって、第 1 の実施の形態にかかる $R_{12}(\tau)$ と同様に扱うことができる。例えば、GCC (τ) のピークは相関の強さを表し、ピークを与える時間は音源方向に対応する。

[0070] なお、GCC と類似した相関関数として CSP (Cross Spectral Phase) と呼ばれるものがある。また、これに重みを付けた重みつき CSP も提案されている。これらは GCC の一形態と考えられ、相関算出部 340 はこれらの関数により相関値を算出してもよい。

[0071] 音声判断部 350 は、相関算出部 340 から相関値 GCC (τ) を取得する。そして、予め設定された閾値 GCC (τ) $_th$ と比較する。相関算出部 340 が算出した相関値 GCC (τ) が閾値 GCC (τ) $_th$ よりも小さい場合には、受信信号は背景雑音信号であると判断する。相関算出部 340 が算出した相関値 GCC (τ) が閾値 GCC (τ) $_th$ 以上である場合には、受信信号は音声信号であると判断する。音声判断部 350 は、判断結果を各処理部 311 ~ 320 の利得設定部に出力する。

[0072] 第 1 処理部 311 は、第 1 利得演算部 361 と、第 2 利得演算部 362 と、第 1 レベル算出部 371 と、第 2 レベル算出部 372 と、利得設定部 380 と、アレー処理部 390 とを備えている。なお、第 1 処理部 311 は、相関算出部および音声判断部は備えない。利得設定部 380 は、音声判断部 3

50から受信信号が音声信号であるか背景雑音信号であるかの判断結果を取得する。利得設定部380は、さらに第1レベル算出部371および第2レベル算出部372からそれぞれ受信信号の第1周波数成分の信号レベルを取得する。利得設定部380は、背景雑音信号区間である場合に、第1レベル算出部371および第2レベル算出部372から取得した信号レベルに基づいて、第1利得演算部361および第2利得演算部362に設定すべき利得値を決定し、これを第1利得演算部361および第2利得演算部362に設定する。

[0073] なお、第2処理部312～第L処理部320の構成および処理は、第1処理部311の構成および処理と同様である。また、第3の実施の形態にかかる受信信号処理装置103のこれ以外の構成は、第2の実施の形態にかかる受信信号処理装置102の構成と同様である。

[0074] このように、第3の実施の形態にかかる受信信号処理装置103においては、利得設定部は周波数毎に設けられているので、周波数毎に独立に利得設定を行うことができる。したがって、周波数毎にマイクロホンの感度が異なる場合には、周波数毎に適切な利得調整を行うことができる。

[0075] 図9は、第4の実施の形態にかかる受信信号処理装置104の構成を示すブロック図である。受信信号処理装置104は、第2、3の実施の形態にかかる受信信号処理装置と同様に、各周波数成分に対する利得調整を行う複数の処理部、すなわち第1処理部411～第L処理部420を備えている。ただし、本実施の形態にかかる受信信号処理装置104においては、アレー処理部は、入力信号の処理に加えて、音源方向の推定と受信信号の強度の推定を行う。音声判断部は、アレー処理部による推定結果に基づいて、受信信号が音声信号であるか背景雑音信号であるかの判断を行う。

[0076] 他の実施の形態において述べた相関の大きさは、本実施の形態において述べた信号の強度に対応する。また、コヒーレンスの位相や相関値の時間差 τ が音源方向に対応する。

[0077] アレー処理部480は、ビームフォーミング法により、アレーの指向性をスキ

ランしながら各方向の出力パワーを測定し、高い出力パワーを与える方向に音源が存在すると判定する。ビームフォーマ法では方向 θ における出力パワーは、(式 18) で表される。

[数18]

$$\text{Pow}(\theta) = \mathbf{a}'(\theta) \mathbf{R}_{xx} * \mathbf{a}(\theta) / \mathbf{a}'(\theta) \mathbf{a}(\theta) \quad \dots (\text{式 } 18)$$

ここで、 $\mathbf{a}(\theta)$ は音源方向に対応する縦ベクトルであり、方向ベクトルまたはモードベクトル等と呼ばれる。 $\mathbf{a}(\theta)$ の次元は、マイクロホンの数に相当する。すなわち、マイクロホンの数が N 個である場合には、 $\mathbf{a}(\theta)$ は、 N 次元となる。 $\mathbf{a}'(\theta)$ は、 $\mathbf{a}(\theta)$ を転置した横ベクトルである。 $\mathbf{R}_{xx}$ は空間相関行列であり、チャンネル間の相互相関を行列で表したものである。2チャンネルの場合の周波数領域で $\mathbf{R}_{xx}$ は (式 19) で表現される。

[数19]

$$\mathbf{R}_{xx}(l) = \begin{bmatrix} G_{11}(l) & G_{12}(l) \\ G_{21}(l) & G_{22}(l) \end{bmatrix} \quad \dots (\text{式 } 19)$$

ここで、 l は周波数成分番号である。(式 19) の成分 G_{xx} は、第 3 の実施の形態において説明したクロススペクトルであり、チャンネル間の相関を表している。

[0078] (式 18) において方向ベクトル $\mathbf{a}(\theta)$ は入力信号によらないベクトルである。したがって、 $\text{Pow}(\theta)$ が大きな値をとるためには、 $\mathbf{R}_{xx}(l)$ の成分が大きな値となる必要がある。つまり、他の実施の形態において説明した、受信信号間の相関が大きくなることと、アレー処理においてある方向に強い方向性が観測されることは等価なことである。

[0079] 音声判断部 460 は、アレー処理部 480 により算出された $\text{Pow}(\theta)$ の最大値と予め設定された閾値 Pow_th とを比較する。そして、 $\text{Pow}(\theta)$ が閾値よりも小さい場合には、相関が低く受信信号は背景雑音信号であると判断する。また、 $\text{Pow}(\theta)$ が閾値 Pow_th 以上である場合には、相関が高く受信信号は音声信号であると判断する。

- [0080] 利得設定部 470 は、受信信号が背景雑音信号であると判断される区間である背景雑音区間において第 1 レベル算出部 451 および第 2 レベル算出部 452 から取得した信号レベルに基づいて、利得値を決定し、これを第 1 利得演算部 441 および第 2 利得演算部 442 に設定する。
- [0081] なお、第 2 処理部 412 ~ 第 L 処理部 420 における処理および構成は、図 9 を参照しつつ説明した第 1 処理部 411 の処理および構成と同様である。また、受信信号処理装置 104 のこれ以外の構成および処理は、他の実施の形態にかかる受信信号処理装置の構成および処理と同様である。
- [0082] 本実施の形態の変更例としては、アレー処理部 480 は、例えば空間相関行列の固有値分解を利用した MUSIC 法など、従来から知られている他の方法を用いて音源方向を推定してもよい。方向推定の詳細な方法については、M. Brandstein and D. Ward, "Microphone Arrays," Springer, Part II, 2001 に記載されている。ビームフォーマ法以外の方向探索アルゴリズムを用いた場合でも、大抵の場合、強い方向性が観測されることと、大きな相関値が得られることは同じことであり、表現方法の違いに過ぎない。
- [0083] 図 10 は、第 5 の実施の形態にかかる受信信号処理装置 105 の構成を示すブロック図である。受信信号処理装置 105 は、第 1 の実施の形態にかかる受信信号処理装置 100 の相関算出部 140 にかえて音声検出部 500 を備えている。音声検出部 500 は、例えば VAD (Voice Activity Detector) 等の音声検出器であり、音声の存在の有無を検出する。音声判断部 510 は、音声が存在する場合には、受信信号は音声信号であると判断する。また、音声が存在しない場合には、受信信号は雑音信号であると判断する。
- [0084] 例えば、受信信号処理装置 105 が設置された周辺環境において想定され得る近接音源が音声信号に限られている場合には、本実施の形態にかかる受信信号処理装置 105 のように、音声検出部 500 による検出結果に基づいて、受信信号が音声信号であるか背景雑音信号であるかを判定することにより、精度よく受信信号の判断を行うことができる。
- [0085] なお、受信信号処理装置 105 のこれ以外の構成および処理は、第 1 の実

施の形態にかかる受信信号処理装置 100 の構成および処理と同様である。

[0086] なお音声検出部 500 による音声検出の方法は、本実施の形態に限定されるものではない。音声検出は、信号のパワー情報を用いる手法、スペクトル情報を用いる手法、信号対雑音比に基づく手法など様々な方法が提案されており、音声検出部 500 はこれらの方法により音声を検出してもよい。

[0087] 図 11 は、第 6 の実施の形態にかかる受信信号処理装置 106 の構成を示すブロック図である。受信信号処理装置 106 は、背景雑音区間ではなく音声区間において、マイクロホンアレーの理想的な利得バランスに近づくように利得値を調整する。受信信号処理装置 106 は、第 1 の実施の形態にかかる受信信号処理装置 100 の音声判断部 150 にかえて相関判断部 600 を備えている。また、第 1 の実施の形態にかかる受信信号処理装置 100 の構成に加えて利得データ記憶部 610 を備えている。

[0088] 相関判断部 600 は、相関算出部 140 から相関値の最大値 r_{12_max} と、このときの位相 τ_{12} 、すなわち τ_{12_max} の組を取得する。相関判断部 600 は、予め相関値およびこのときの位相の設定値の組を記憶しており、これと取得した最大値の組とを比較する。なお、設定値は、近接音源が存在する場合に得られる相関値の最大値 r_{12_max} と、このときの位相 τ_{12} であり、予め実験等により求めたものである。相関算出部 140 により算出された r_{12_max} と τ_{12_max} の値がそれぞれ r_{12_max} と τ_{12_max} の設定値と一致した場合には、利得設定部 620 に対し利得調整を行う旨の指示を出力する。なお、相関算出部 140 により算出された r_{12_max} と τ_{12_max} の値がそれぞれ r_{12_max} と τ_{12_max} の設定値を基準としたある範囲内の値であれば、一致したと判断することとする。

[0089] 利得データ記憶部 610 は、利得データを記憶している。ここで、利得データとは、相関が相関判断部 600 に記憶されている設定値になるような状況において感度のそろった複数のマイクロホンを用いて受信した場合の理想的な利得バランスを示す情報である。すなわち、利得データには、理想的な

状況での各マイクロホンの信号パワーが示されている。利得設定部 620 は、利得データに基づいて、第 1 マイクロホン 111 および第 2 マイクロホン 112 の受音信号に乗じるべき利得値を決定する。具体的には、利得値を乗じた受音信号のパワーが理想的な利得バランスと一致するような利得値を決定する。そして、決定した利得値を第 1 利得演算部 121 および第 2 利得演算部 122 に設定する。なお、この場合にも、利得設定部 620 は目標値を理想的な利得バランスとして段階的に利得値を設定してもよい。

[0090] 本実施の形態にかかる受音信号処理装置 106 においては、固定位置に音源が存在し、かつ、その音源から音が発せられている時間帯が長い場合において、効率良く利得調整を行うことが可能となる。

[0091] なお、本実施の形態にかかる受音信号処理装置 106 の構成および処理は、他の実施の形態にかかる受音信号処理装置の構成および処理と同様である。

[0092] 本実施の形態の受音信号処理装置は、CPU などの制御装置と、ROM (Read Only Memory) や RAM などの記憶装置と、HDD、CD ドライブ装置などの外部記憶装置と、ディスプレイ装置などの表示装置と、キーボードやマウスなどの入力装置を備えており、通常のコンピュータを利用したハードウェア構成となっている。

[0093] 本実施の形態の受音信号処理装置で実行される受音信号処理プログラムは、インストール可能な形式又は実行可能な形式のファイルで CD-ROM、フレキシブルディスク (FD)、CD-R、DVD (Digital Versatile Disk) 等のコンピュータで読み取り可能な記録媒体に記録されて提供される。

[0094] また、本実施の形態の受音信号処理装置で実行される受音信号処理プログラムを、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するように構成しても良い。また、本実施形態の受音信号処理装置で実行される受音信号処理プログラムをインターネット等のネットワーク経由で提供または配布す

るように構成しても良い。また、本実施形態の受信信号処理プログラムを、ROM等に予め組み込んで提供するように構成してもよい。

[0095] 本実施の形態の受信信号処理装置で実行される受信信号処理プログラムは、上述した各部（第1利得演算部、第2利得演算部、第1レベル算出部、第2レベル算出部、相関算出部、音声判断部、利得設定部、アレー処理部など）を含むモジュール構成となっており、実際のハードウェアとしてはCPU（プロセッサ）が上記記憶媒体から受信信号処理プログラムを読み出して実行することにより上記各部が主記憶装置上にロードされ、各部が主記憶装置上に生成されるようになっている。

[0096] なお、本発明は、上記実施の形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化することができる。また、上記実施の形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成することができる。例えば、実施の形態に示される全構成要素からいくつかの構成要素を削除してもよい。さらに、異なる実施の形態にわたる構成要素を適宜組み合わせても良い。

符号の説明

- [0097] 100～106 受信信号処理装置
- 111 第1マイクロホン
 - 112 第2マイクロホン
 - 121 第1利得演算部
 - 122 第2利得演算部
 - 131 第1レベル算出部
 - 132 第2レベル算出部
 - 150 音声判断部
 - 160 利得設定部

請求の範囲

[請求項1]

音声を受音する複数のマイクロホンと、

前記複数のマイクロホンが受音した受音信号が、前記マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを、受音信号に基づいて判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が前記背景雑音信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乗じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の前記利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値に乗じる演算部と
を備えたことを特徴とする受音信号処理装置。

[請求項2]

前記設定部は、現在設定されている利得値を前記複数のマイクロホンの信号レベルが等しくなる目標利得値まで変更させる際の利得値の調整幅を決定し、予め設定された第1規定時間が経過する度に、前記調整幅だけ既に設定されている前記利得値を変更した値を、新たな利得値として設定することを特徴とする請求項1に記載の受音信号処理装置。

[請求項3]

前記複数のマイクロホンが受音した複数の受音信号の相関を算出する相関算出部をさらに備え、

前記音声判断部は、前記相関算出部により算出された前記相関が予め定められた閾値に比べて小さい場合に、前記背景雑音信号であると

判断することを特徴とする請求項 1 に記載の受信信号処理装置。

[請求項4]

前記受信信号を周波数成分に変換する変換部をさらに備え、

前記信号レベル算出部は、前記変換部により得られた周波数成分毎に前記受信信号それぞれの信号レベルを算出し、

前記相関算出部は、前記周波数成分の相関を算出し、

前記設定部は、前記周波数成分毎に前記利得値を決定し、前記周波数成分毎に前記受信信号の前記利得値を設定し、

前記演算部は、前記受信信号の前記周波数成分それぞれに対し、各周波数成分に対して設定された前記利得値を乗じることを特徴とする請求項 3 に記載の受信信号処理装置。

[請求項5]

前記音声判断部は、予め設定された第 2 規定時間が経過する度に、前記受信信号が前記音声信号であるか前記背景雑音信号であるかを判断し、

前記決定部は、予め設定された第 3 規定時間の間、前記受信信号が前記背景雑音信号であるとの判断が連続して得られた場合に、前記受信信号の前記利得値を決定することを特徴とする請求項 1 に記載の受信信号処理装置。

[請求項6]

前記受信信号から発話を検出する音声検出部をさらに備え、

前記音声判断部は、前記音声検出部により発話を検出されない場合に、前記背景雑音信号であると判断することを特徴とする請求項 1 に記載の受信信号処理装置。

[請求項7]

予め定められた規定位置に設置され、音声を受音する複数のマイクロホンと、

前記複数のマイクロホンが受信した受信信号が、マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを判断する音声判断部と、

前記複数のマイクロホンが受信した複数の受信信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が音声信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乗じるべき利得値であって、前記複数のマイクロホンそれぞれが受音する複数の受音信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマイクロホンによる前記複数の受音信号の理想的なレベルバランスに近づける利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値に乗じる演算部と
を備えたことを特徴とする受音信号処理装置。

[請求項8]

複数のマイクロホンの受音信号を処理する受音信号処理をコンピュータに実行させるための受音信号処理プログラムであって、

前記コンピュータを、

前記複数のマイクロホンから前記受音信号を取得する取得部と、

前記受音信号が、前記マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを、受音信号に基づいて判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が前記背景雑音信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乗じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の前記利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値を乗じる演算部として機能させるためのプログラム。

[請求項9]

予め定められた規定位置に設置された複数のマイクロホンの受音信号を処理する受音信号処理をコンピュータに実行させるための受音信号処理プログラムであって、

前記コンピュータを、

前記複数のマイクロホンから前記受音信号を取得する取得部と、

前記受音信号が、マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを、受音信号に基づいて判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が音声信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乘じるべき利得値であって、前記複数のマイクロホンそれぞれが受音する複数の受音信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマイクロホンによる前記複数の受音信号の理想的なレベルバランスに近づける利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値を乗じる演算部として機能させるためのプログラム。

[請求項10]

複数のマイクロホンが音声を受音する受音ステップと、

音声判断部が、前記複数のマイクロホンが受音した受音信号が、前記マイクロホンに近接する近接音源からの音声を含む音声信号である

か、前記音声を含まない背景雑音信号かを、受信信号に基づいて判断する音声判断ステップと、

信号レベル算出部が、前記複数のマイクロホンが受信した複数の受信信号それぞれの信号レベルを算出する信号レベル算出ステップと、

設定部が、前記音声判断ステップにおいて前記受信信号が前記背景雑音信号であると判断された場合に、前記複数の受信信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受信信号に乗じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受信信号の前記利得値として設定する設定ステップと、

演算部が、前記少なくとも1つのマイクロホンの前記受信信号に、前記設定部によって設定された前記利得値に乗じる演算ステップとを有することを特徴とする受信信号処理方法。

[請求項11]

予め定められた規定位置に設置された複数のマイクロホンが音声を受音する受音ステップと、

音声判断部が、前記複数のマイクロホンが受信した受信信号が、マイクロホンに近接する近接音源からの音声を含む音声信号であるか、前記音声を含まない背景雑音信号かを、受信信号に基づいて判断する音声判断ステップと、

信号レベル算出部が、前記複数のマイクロホンが受信した複数の受信信号それぞれの信号レベルを算出する信号レベル算出ステップと、

設定部が、前記音声判断ステップにおいて前記受信信号が音声信号であると判断された場合に、前記複数の受信信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受信信号に乗じるべき利得値であって、前記複数のマイクロホンそれぞれが受音する複数の受信信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマ

マイクロホンによる前記複数の受信信号の理想的なレベルバランスに近づける利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受信信号の利得値として設定する設定ステップと、

演算部が、前記少なくとも1つのマイクロホンの前記受信信号に、前記設定部によって設定された前記利得値を乗じる演算ステップとを有することを特徴とする受信信号処理方法。

補正された請求の範囲
[2010年7月20日(20.07.2010)国際事務局受理]

[1] (補正後) 音声を受音する複数のマイクロホンと、

前記複数のマイクロホンが受音した受音信号が、前記マイクロホンに近接する近接音源からの信号であるか、背景雑音信号かを判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が前記背景雑音信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乗じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の前記利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値を乗じる演算部と
を備えたことを特徴とする受音信号処理装置。

[2] 前記設定部は、現在設定されている利得値を前記複数のマイクロホンの信号レベルが等しくなる目標利得値まで変更させる際の利得値の調整幅を決定し、予め設定された第1規定時間が経過する度に、前記調整幅だけ既に設定されている前記利得値を変更した値を、新たな利得値として設定することを特徴とする請求項1に記載の受音信号処理装置。

[3] 前記複数のマイクロホンが受音した複数の受音信号の相関を算出する相関算出部をさらに備え、

前記音声判断部は、前記相関算出部により算出された前記相関が予め定められた閾値に比べて小さい場合に、前記背景雑音信号であると判断することを特徴とする請求項1に記載の受音信号処理装置。

[4] 前記受音信号を周波数成分に変換する変換部をさらに備え、

前記信号レベル算出部は、前記変換部により得られた周波数成分毎に前記受音信号それぞれの信号レベルを算出し、

前記相関算出部は、前記周波数成分の相関を算出し、

前記設定部は、前記周波数成分毎に前記利得値を決定し、前記周波数成分毎に前記受音信号の前記利得値を設定し、

前記演算部は、前記受音信号の前記周波数成分それぞれに対し、各周波数成分に対して設定された前記利得値を乗じることの特徴とする請求項 3 に記載の受音信号処理装置。

[5] (補正後) 前記音声判断部は、予め設定された第 2 規定時間が経過する度に、前記受音信号が前記近接音源からの信号であるか前記背景雑音信号であるかを判断し、

前記決定部は、予め設定された第 3 規定時間の間、前記受音信号が前記背景雑音信号であるとの判断が連続して得られた場合に、前記受音信号の前記利得値を決定することの特徴とする請求項 1 に記載の受音信号処理装置。

[6] 前記受音信号から発話を検出する音声検出部をさらに備え、

前記音声判断部は、前記音声検出部により発話を検出されない場合に、前記背景雑音信号であると判断することの特徴とする請求項 1 に記載の受音信号処理装置。

[7] (補正後) 予め定められた規定位置に設置され、音声を受音する複数のマイクロホンと、

前記複数のマイクロホンが受信した受音信号が、マイクロホンに近接する近接音源からの信号であるか、前記近接音源からの信号を含まない雑音信号かを判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が前記近接音源からの信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乘じるべき利得値であって、前記複数のマイクロホンそれぞれが受音する複数の受音信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマイクロホンによる前記複数の受音信号の理想的なレベルバランスに近づける利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値を乗じる演算部と
を備えたことを特徴とする受音信号処理装置。

[8] (補正後) 複数のマイクロホンの受音信号を処理する受音信号処理をコンピュータに実行させるための受音信号処理プログラムであって、

前記コンピュータを、

前記複数のマイクロホンから前記受音信号を取得する取得部と、

前記受音信号が、前記マイクロホンに近接する近接音源からの信号であるか、背景雑音信号かを判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が前記背景雑音信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乘じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の前記利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によっ

て設定された前記利得値を乗じる演算部として機能させるためのプログラム。

[9] (補正後) 予め定められた規定位置に設置された複数のマイクロホンの受音信号を処理する受音信号処理をコンピュータに実行させるための受音信号処理プログラムであって、

前記コンピュータを、

前記複数のマイクロホンから前記受音信号を取得する取得部と、

前記受音信号が、マイクロホンに近接する近接音源からの信号であるか、前記近接音源からの信号を含まない雑音信号かを判断する音声判断部と、

前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出部と、

前記音声判断部において前記受音信号が前記近接音源からの信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乘じるべき利得値であって、前記複数のマイクロホンそれぞれが受音する複数の受音信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマイクロホンによる前記複数の受音信号の理想的なレベルバランスに近づける利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の利得値として設定する設定部と、

前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値を乗じる演算部として機能させるためのプログラム。

[10] (補正後) 複数のマイクロホンが音声を受音する受音ステップと、

音声判断部が、前記複数のマイクロホンが受音した受音信号が、前記マイクロホンに近接する近接音源からの信号であるか、背景雑音信号かを判断する音声判断ステップと、

信号レベル算出部が、前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出ステップと、

設定部が、前記音声判断ステップにおいて前記受音信号が前記背景雑音信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、前記複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乗じるべき利得値であって、前記複数のマイクロホンの間の信号レベルの差を減少させる利得値を決定し、前記利得値を、前記少なくとも1つのマイクロホンの前記受音信号の前記利得値として設定する設定ステップと、

演算部が、前記少なくとも1つのマイクロホンの前記受音信号に、前記設定部によって設定された前記利得値に乗じる演算ステップと
を有することを特徴とする受音信号処理方法。

[11] (補正後) 予め定められた規定位置に設置された複数のマイクロホンが音声を受音する受音ステップと、

音声判断部が、前記複数のマイクロホンが受信した受音信号が、マイクロホンに近接する近接音源からの信号であるか、前記近接音源からの信号を含まない雑音信号かを判断する音声判断ステップと、

信号レベル算出部が、前記複数のマイクロホンが受音した複数の受音信号それぞれの信号レベルを算出する信号レベル算出ステップと、

設定部が、前記音声判断ステップにおいて前記受音信号が前記近接音源からの信号であると判断された場合に、前記複数の受音信号それぞれの信号レベルに基づいて、複数のマイクロホンのうち少なくとも1つのマイクロホンの前記受音信号に乗じるべき利得値であって、前記複数のマイクロホンそれぞれが受音する複数の受音信号の信号レベルのバランスを、予め記憶部に記憶されている、前記規定位置に設置された複数のマイクロホンによる前記複数の受音信号の理想的なレベルバランスに近づける利得値を決定し、前記利

得値を、前記少なくとも 1 つのマイクロホンの前記受信信号の利得値として設定する設定ステップと、

演算部が、前記少なくとも 1 つのマイクロホンの前記受信信号に、前記設定部によって設定された前記利得値を乗じる演算ステップとを有することを特徴とする受信信号処理方法。

条約第19条(1)に基づく説明書

審査官は、文献1に「相互相関度に基づいて音声信号か雑音信号かを特定し、…」との記載を指摘しているが、そのような記載は見当たらない。

文献1の請求項2に「相互相関度に基づいて各音声信号の時間差を検出する」と記載されているように、文献1は、音声信号が存在する場合にその時間差を検出する方法である。

文献1の段落[0012]の最後に、「音声信号は時間差($t_a - t_b$)の位置に相関度が高く現れ、一方、不規則性の騒音は相関度が低い」と記載されている。この性質により、文献1では、騒音のある環境においても正しく音声信号の時間差($t_a - t_b$)を得ることができる。

しかし、文献1には、そもそも音声信号(本願の近接音源に相当)が存在しなかったときに、($t_a - t_b$)に生じた相関度をどう扱うのかについては記載がない。すなわち本願で扱っている音声信号の存在の有無を検出する方法については何ら記載がない。段落[0013]に音声信号と規則性騒音を区別する方法の記載があるが、これも音声信号が存在することが前提になっている。

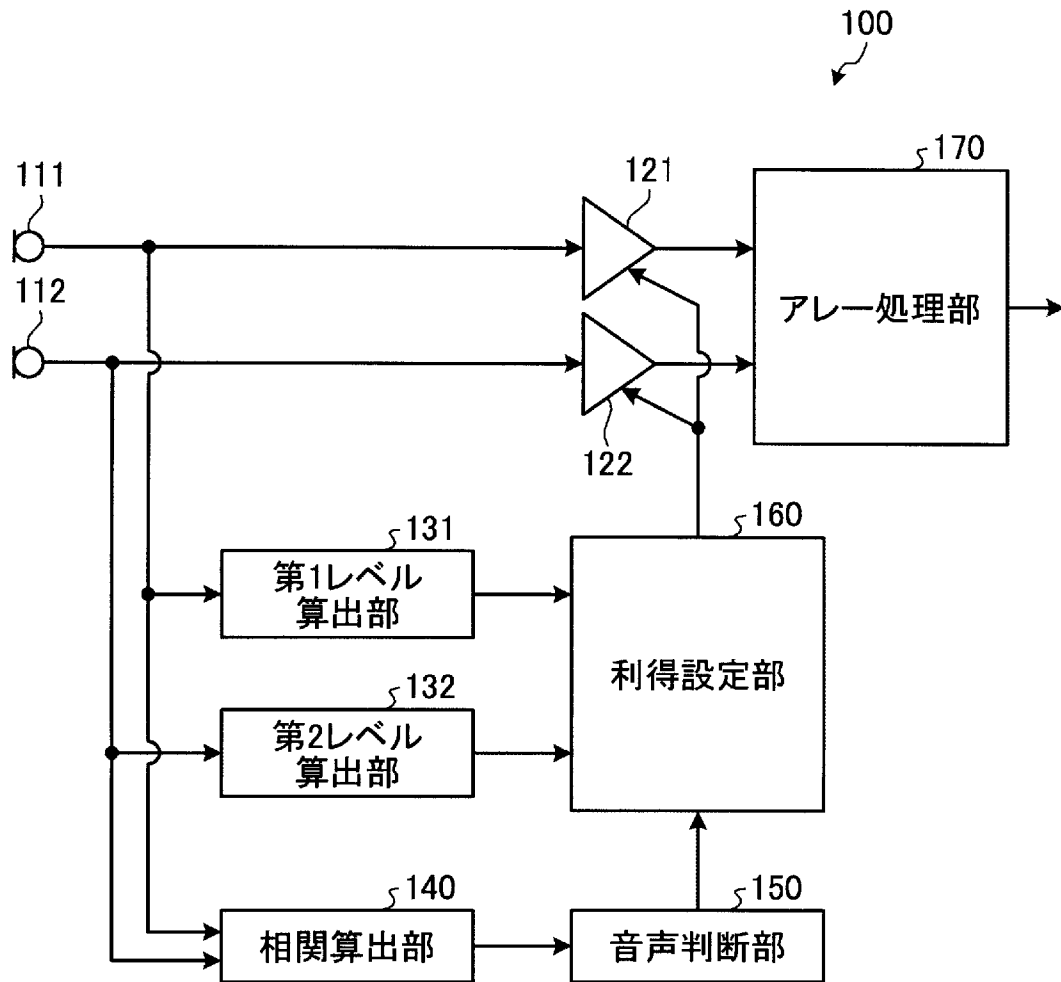
審査官は、文献2に「集音信号が雑音信号であった場合にはマイクロホンのキャリブレーションを行なう」との記載を指摘している。文献2のキャリブレーションと本願のキャリブレーション(調整)とは意味が異なる。

文献2のキャリブレーションは実施例に「雑音の到来方向により、マイクロホンで集音されたレベルが異なるので、そのレベル合わせをレベル設定部18で行なう。」と記載され、この動作を「キャリブレーション」と定義している。本願では到来方向によって生じたレベル差は調整すべきものではないという立場である。このことは、本願[0033]に具体例をあげて記載されている。マイクロホン間でレベル差が生じていることは近接音源である、すなわち、音源からマイクロホンまでの距離がマイクごとに異なることを意味しておりこのような場合に調整しないような仕組みを提供するのが本願の意図であり、文献2とはキャリブレーションを行なうタイミングが逆である。

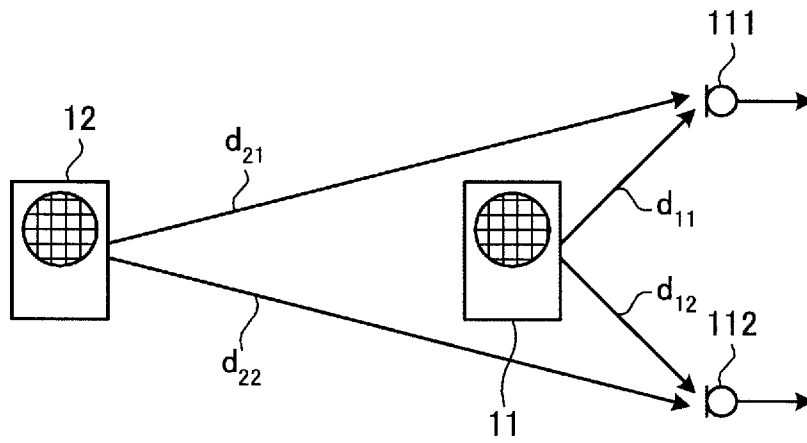
審査官は、文献3に「あらかじめマイクロホン同士の帯域別のスペクトル比を記憶しておくことで集音時にレベル調整を行う技術」の記載を指摘している。

文献3の方法では校正に43スピーカも用いるシステムになっている点が大きく異なる。また、文献3では必然的に、校正モードと使用モードを区別して使うことになる。一方本願は、マイクロホンだけからの入力により調整を行う構成であり、使用モード中に調整を行うタイミングをシステムが自動的に判定するものであり、構成や効果が異なる。

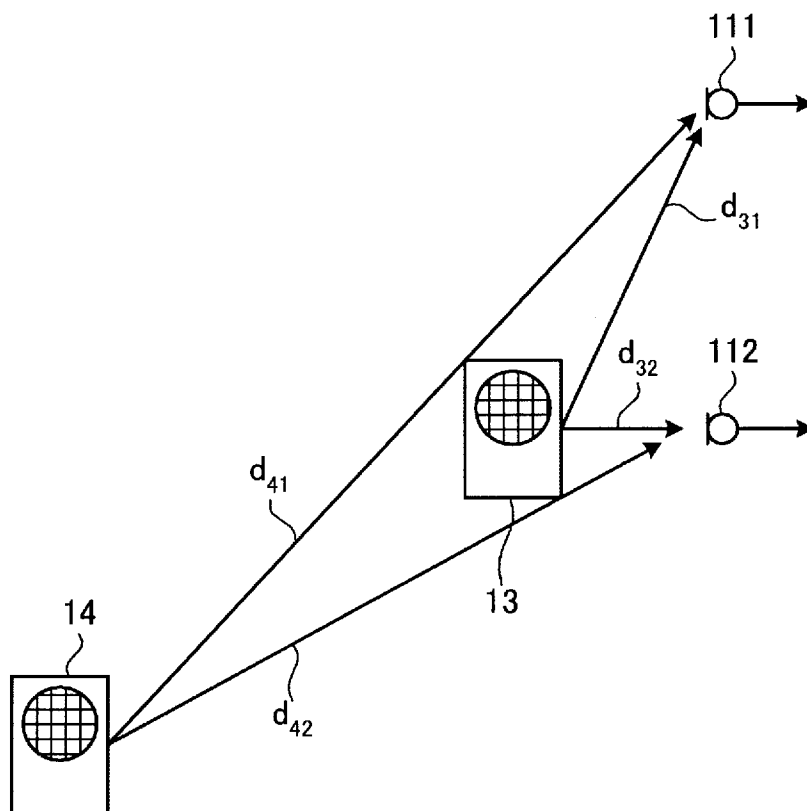
[図1]



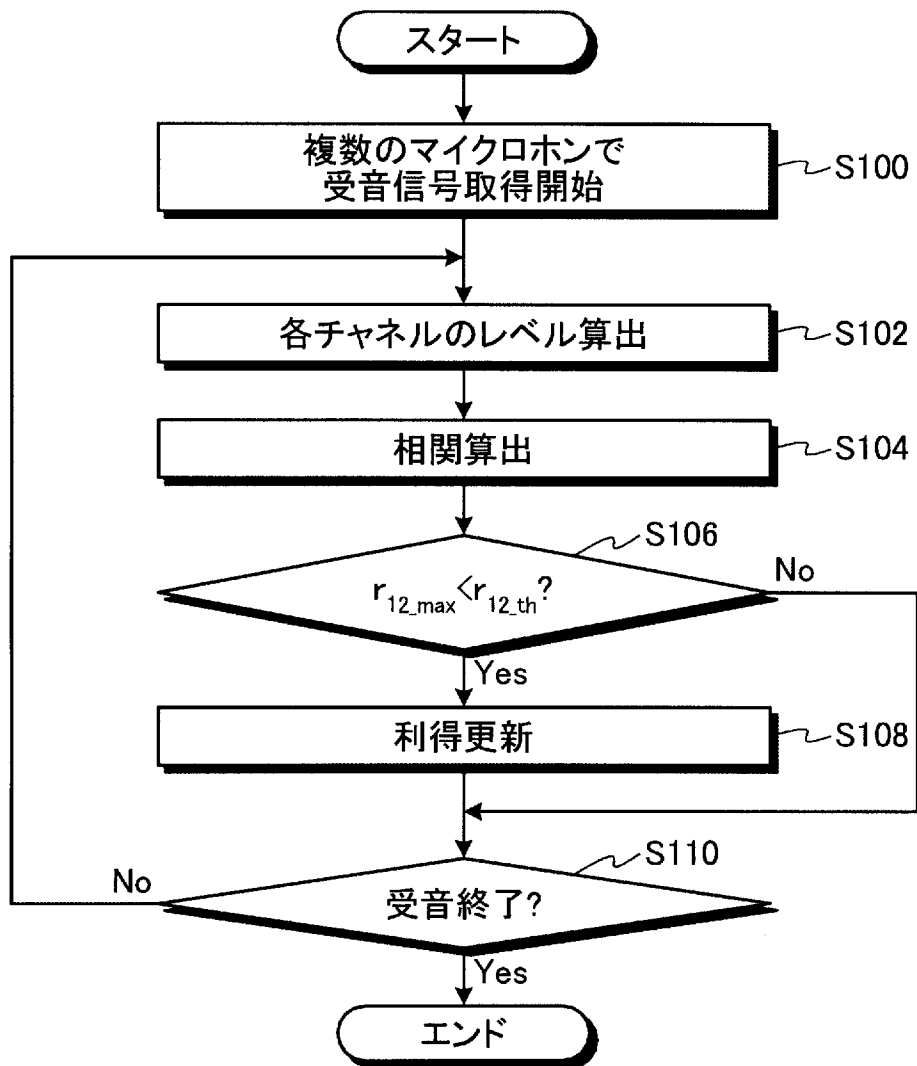
[図2]



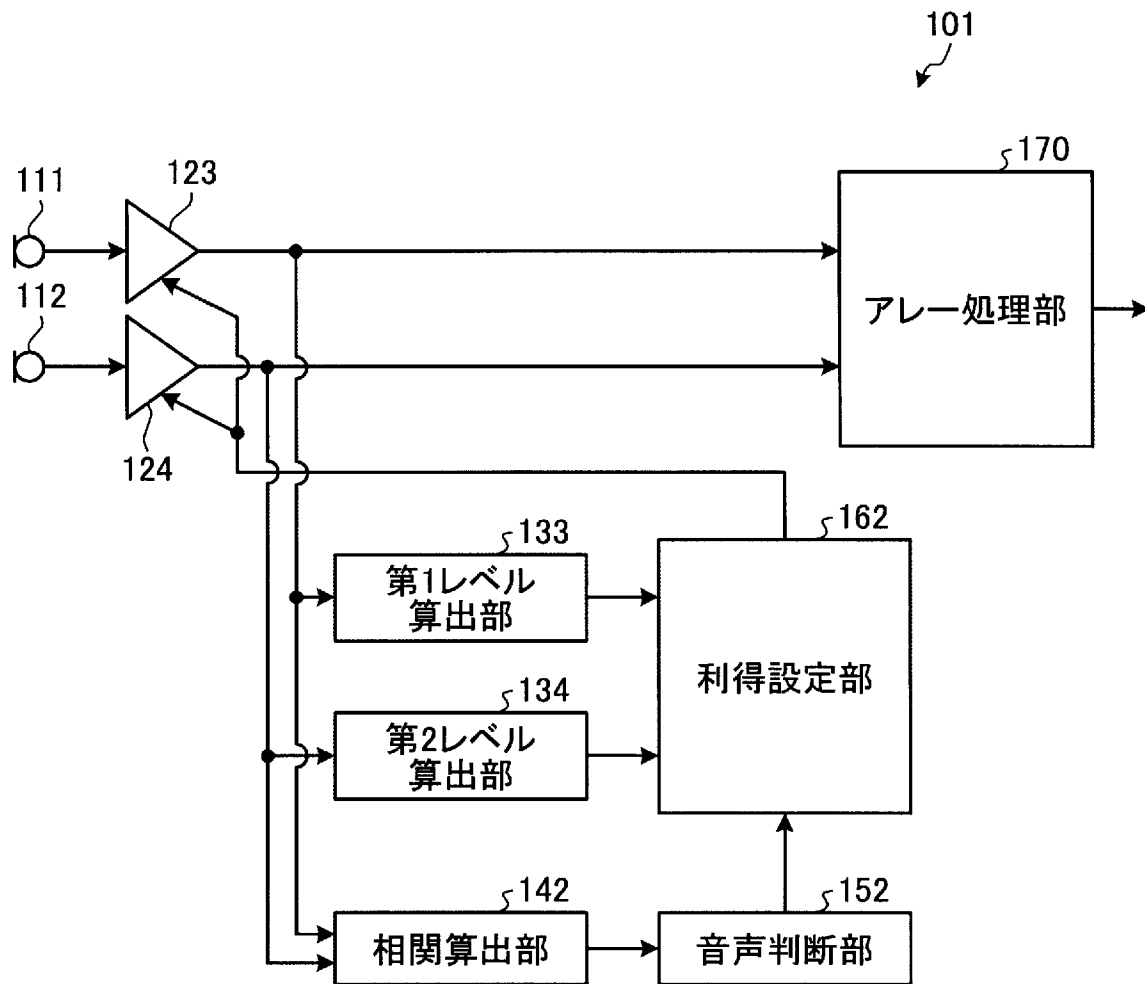
[図3]



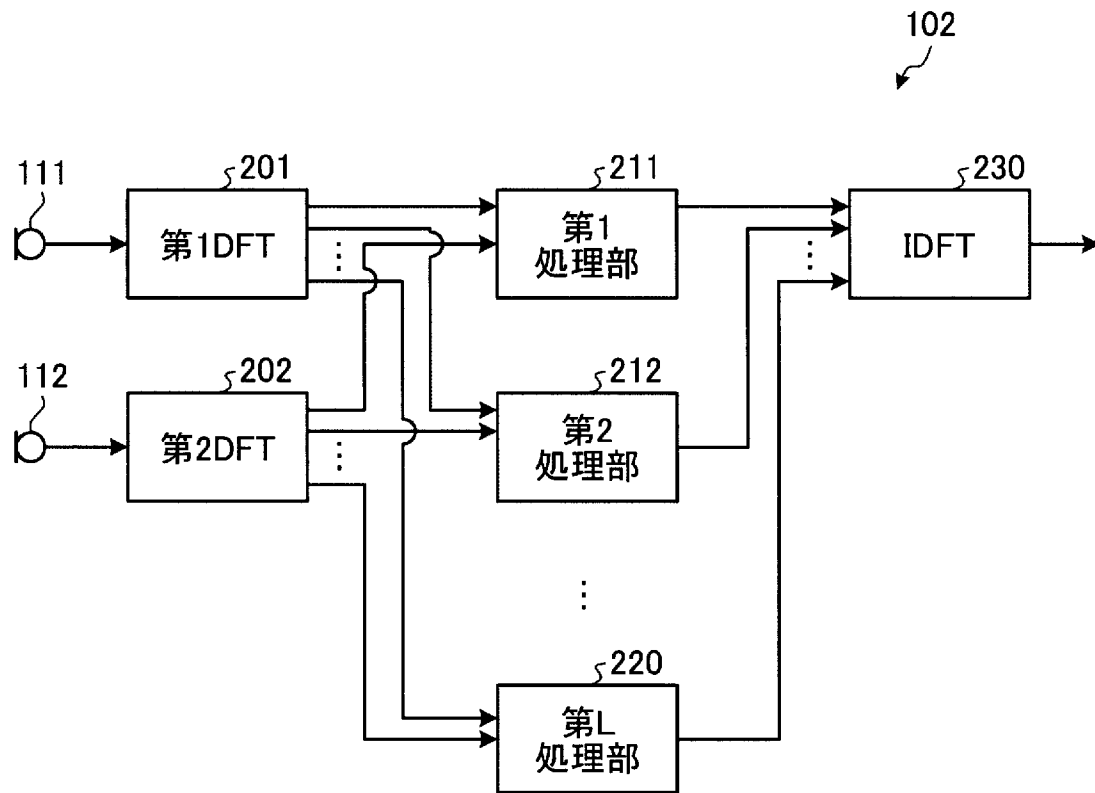
[図4]



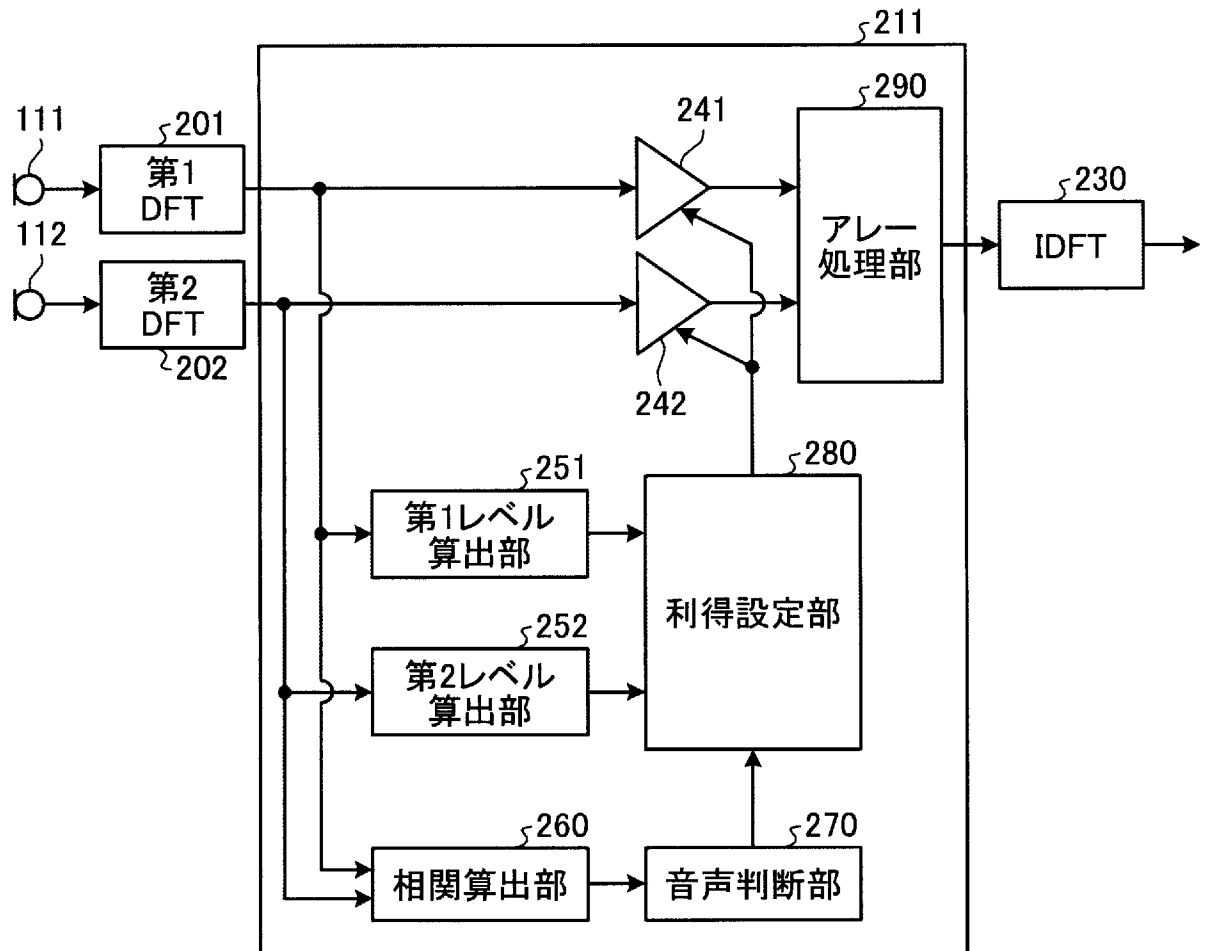
[図5]



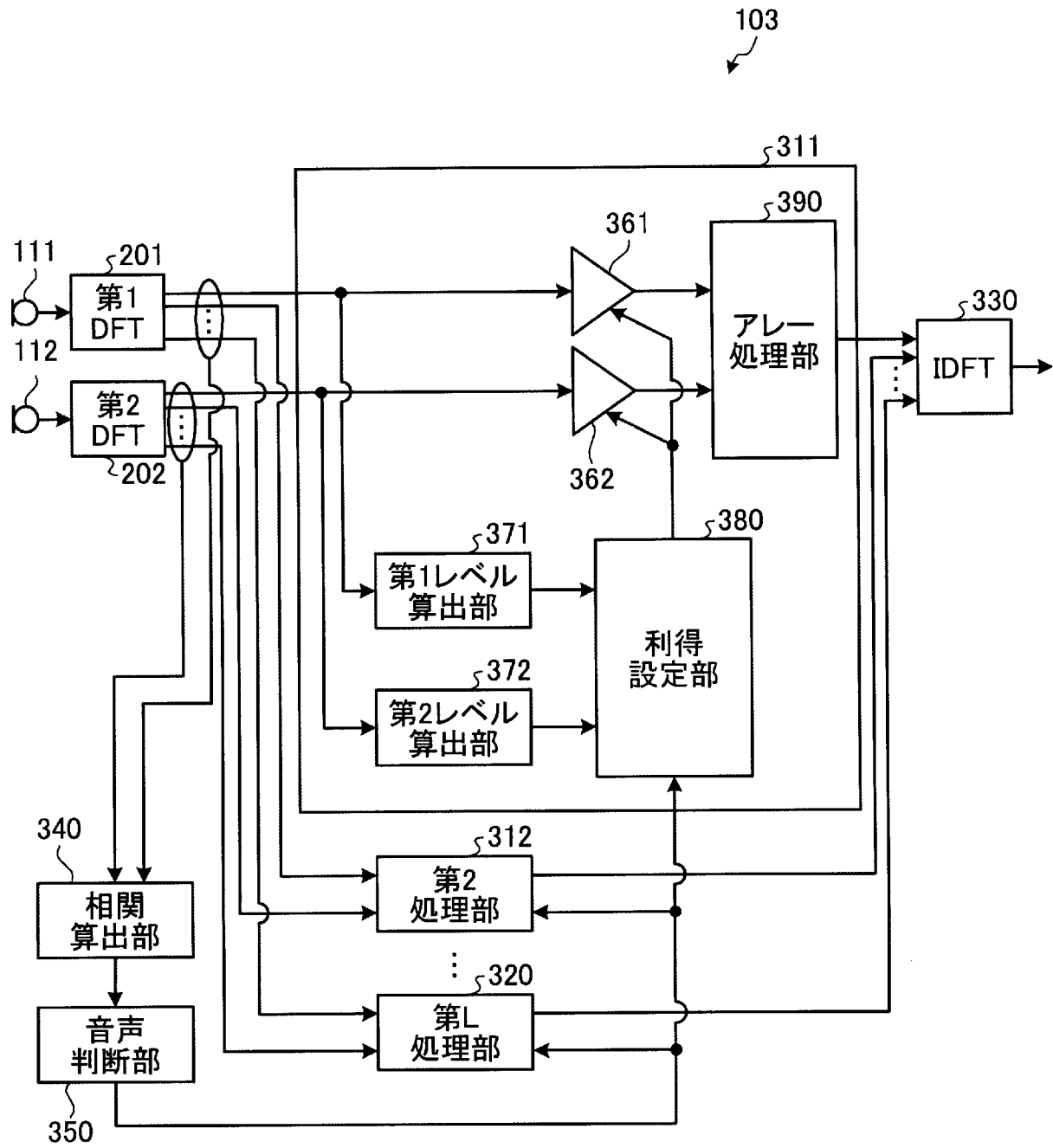
[図6]



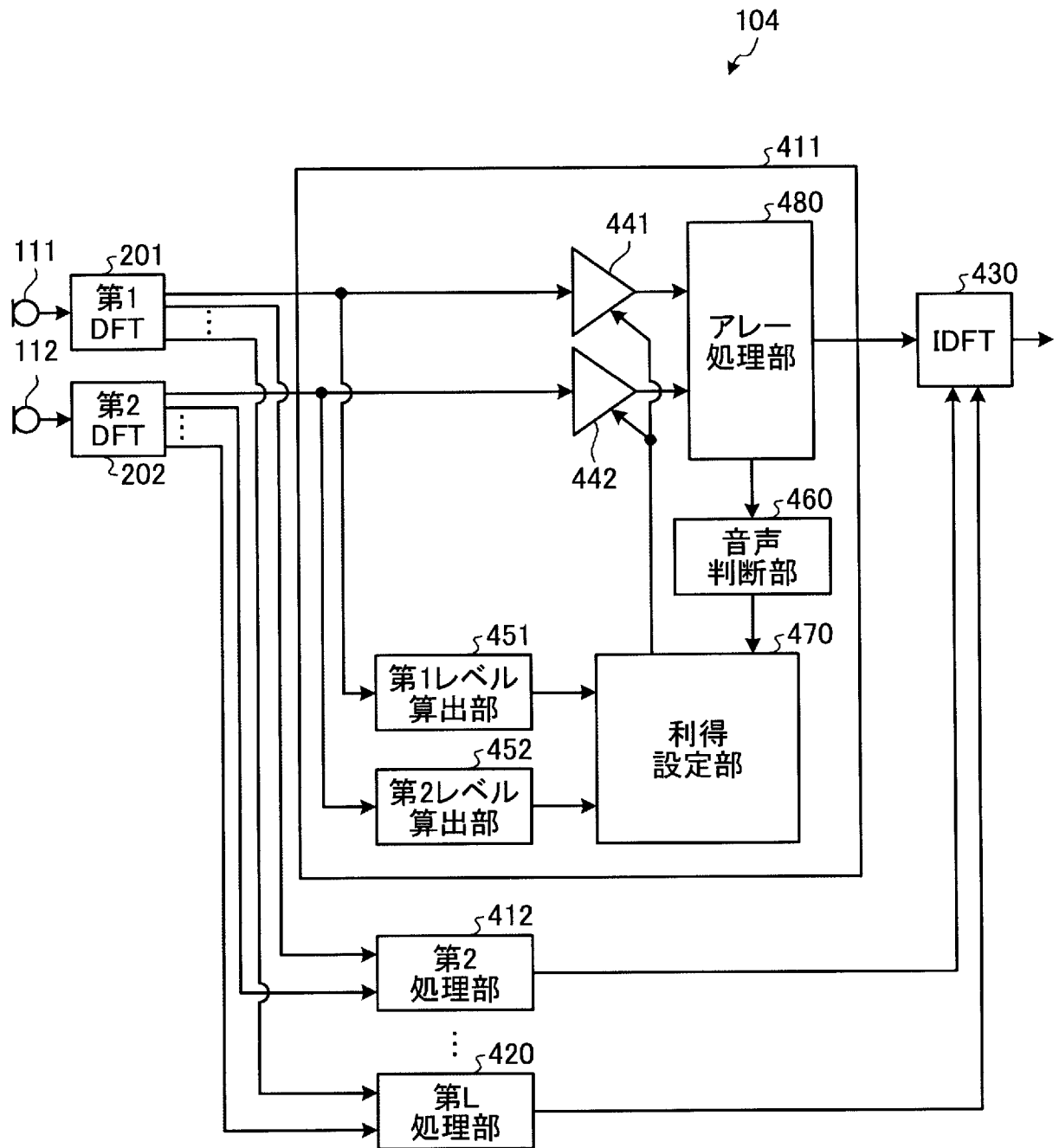
[図7]



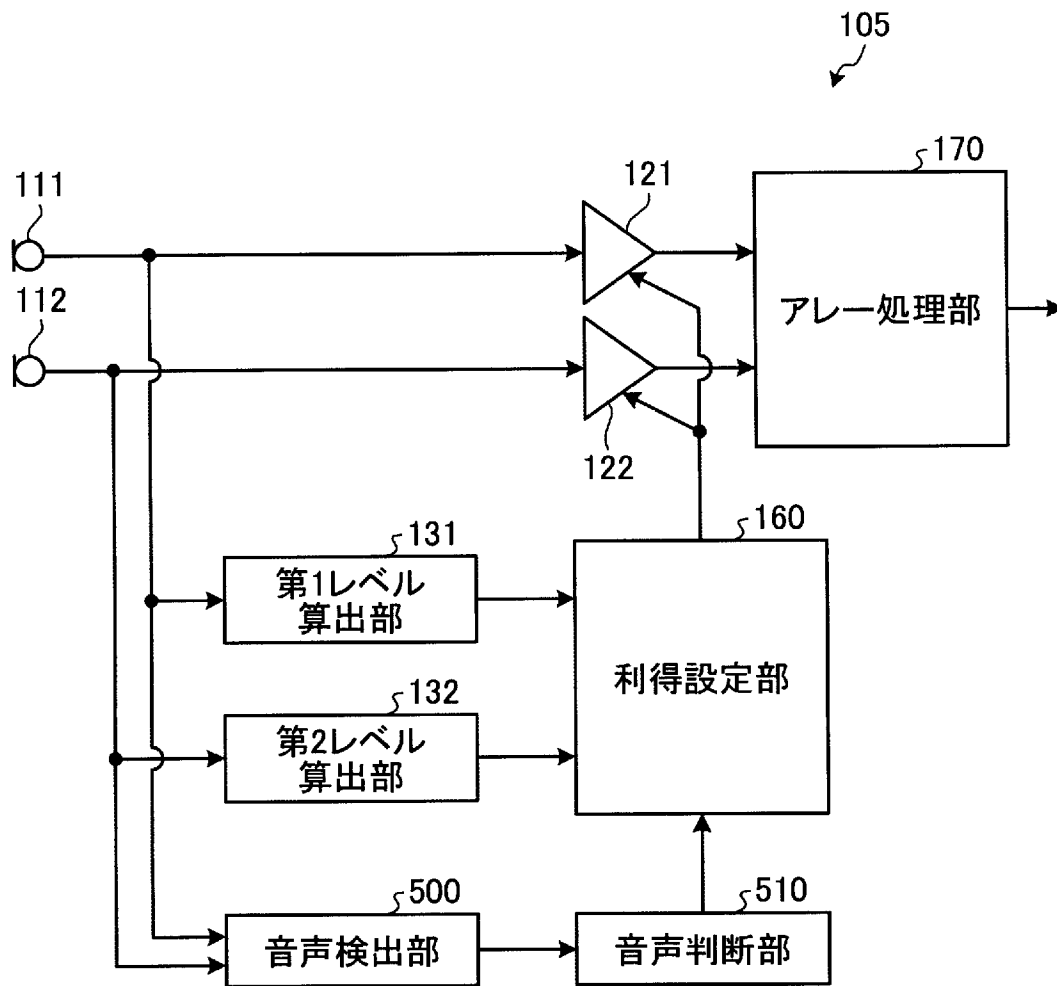
[図8]



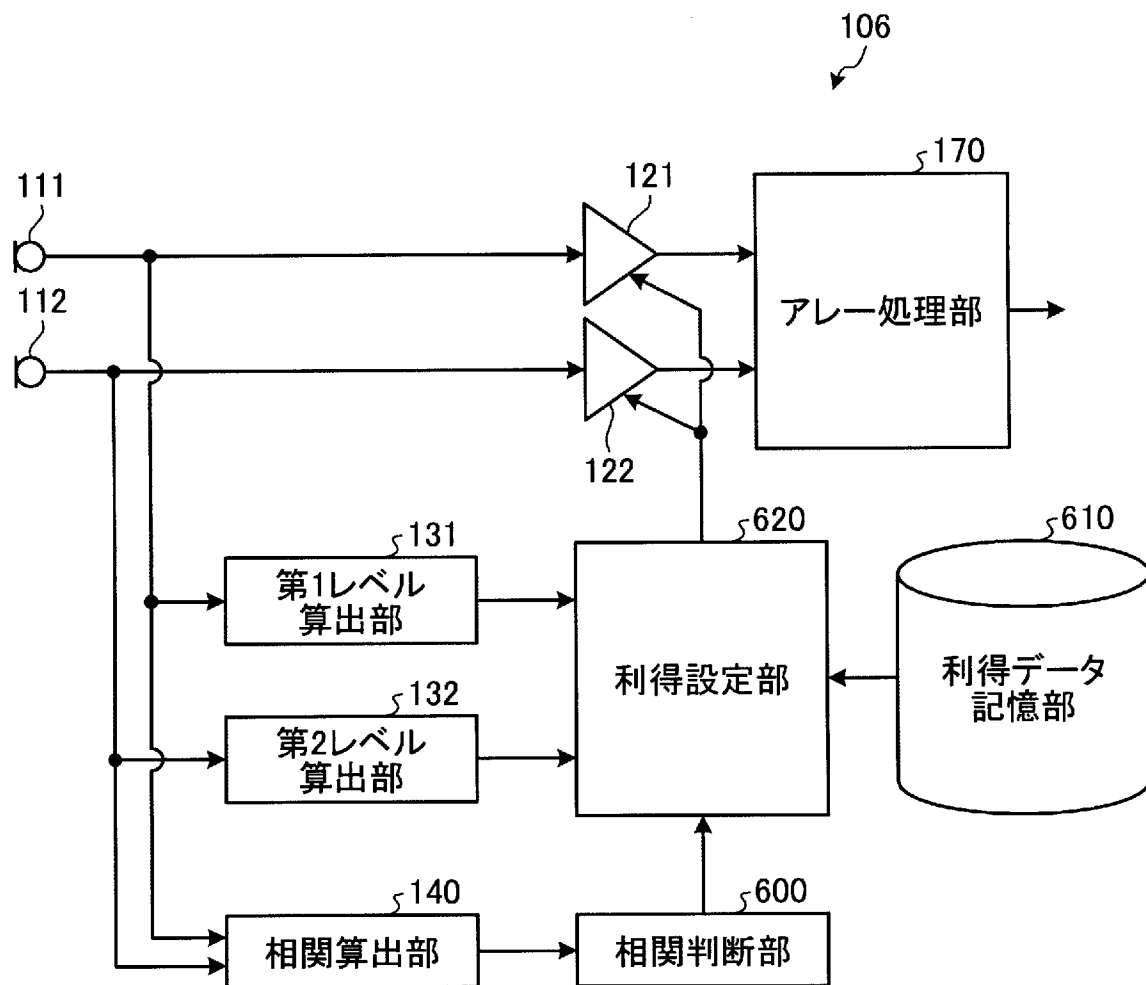
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/067709

A. CLASSIFICATION OF SUBJECT MATTER

H04R3/00(2006.01)i, G10L11/02(2006.01)i, H04R1/40(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04R3/00, G10L11/02, H04R1/40

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 1996-2010

Kokai Jitsuyo Shinan Koho 1971-2010 Toroku Jitsuyo Shinan Koho 1994-2010

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 05-095596 A (Mazda Motor Corp.), 16 April 1993 (16.04.1993), paragraphs [0008] to [0018]; fig. 1 to 4 (Family: none)	1, 3, 4, 6-11 2, 5
Y A	JP 04-016900 A (Clarion Co., Ltd.), 21 January 1992 (21.01.1992), page 2, lower right column, line 5 to page 3, lower right column, line 19; fig. 1 to 6 (Family: none)	1, 3, 4, 6, 8, 10 2, 5
Y A	JP 2007-129373 A (Waseda University), 24 May 2007 (24.05.2007), paragraphs [0054] to [0112]; fig. 1 to 5 (Family: none)	7, 9, 11 2, 5

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

* Special categories of cited documents:

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

"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
22 January, 2010 (22.01.10)Date of mailing of the international search report
09 February, 2010 (09.02.10)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/067709

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2005-195955 A (Toshiba Corp.), 21 July 2005 (21.07.2005), entire text; fig. 1 to 10 & US 2005/0152563 A1	1-11
A	JP 58-181099 A (Mitsubishi Electric Corp.), 22 October 1983 (22.10.1983), entire text; fig. 1 to 15 (Family: none)	1-11
A	JP 2006-270949 A (Microsoft Corp.), 05 October 2006 (05.10.2006), entire text; fig. 1 to 8 & US 2006/0210096 A1 & EP 1703774 A2	1-11

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. H04R3/00(2006.01)i, G10L11/02(2006.01)i, H04R1/40(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. H04R3/00, G10L11/02, H04R1/40

最小限資料以外の資料で調査を行った分野に含まれるもの

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

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

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	JP 05-095596 A (マツダ株式会社) 1993.04.16, 【0008】－【0018】及び第1－4図（ファミリーなし）	1, 3, 4, 6-11 2, 5
Y A	JP 04-016900 A (クラリオン株式会社) 1992.01.21, 第2頁右下欄第5行－第3頁右下欄第19行及び第1－6図（ファミリーなし）	1, 3, 4, 6, 8, 10 2, 5
Y A	JP 2007-129373 A (学校法人早稲田大学) 2007.05.24, 【0054】－【0112】及び第1－5図（ファミリーなし）	7, 9, 11 2, 5

☒ 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

2 2 . 0 1 . 2 0 1 0

国際調査報告の発送日

0 9 . 0 2 . 2 0 1 0

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号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）

境 周一

5 Z

3 6 5 4

電話番号 03-3581-1101 内線 3541

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
A	JP 2005-195955 A (株式会社東芝) 2005. 07. 21, 全文及び第 1 - 1 0 図 & US 2005/0152563 A1	1-11
A	JP 58-181099 A (三菱電機株式会社) 1983. 10. 22, 全文及び第 1 - 1 5 図 (ファミリーなし)	1-11
A	JP 2006-270949 A (マイクロソフト コーポレーション) 2006. 10. 05, 全文及び第 1 - 8 図 & US 2006/0210096 A1 & EP 1703774 A2	1-11