

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014 年 9 月 4 日(04.09.2014)



(10) 国際公開番号

WO 2014/132533 A1

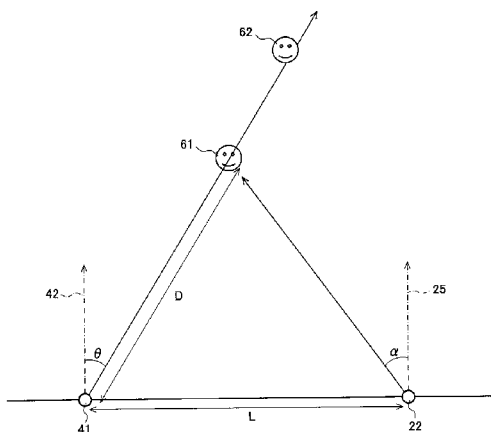
- (51) 国際特許分類:
H04R 3/00 (2006.01) H04R 1/40 (2006.01)
G10L 15/28 (2013.01)
- (21) 国際出願番号: PCT/JP2013/084034
- (22) 国際出願日: 2013 年 12 月 19 日(19.12.2013)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2013-040307 2013 年 3 月 1 日(01.03.2013) JP
- (71) 出願人: シャープ株式会社(SHARP KABUSHIKI KAISHA) [JP/JP]; 〒5458522 大阪府大阪市阿倍野区長池町 2 2 番 2 2 号 Osaka (JP).
- (72) 発明者: 田丸 貴行(TAMARU, Takayuki). 徳井 圭(TOKUI, Kei). 有田 真一(ARITA, Shinichi).
- (74) 代理人: 岡田 宏之(OKADA, Hiroyuki); 〒2310041 神奈川県横浜市中区吉田町 7 2 番地サリュートビル 9 F なぎさ特許事務所 Kanagawa (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, 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, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, 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))

(54) Title: VOICE INPUT DEVICE AND IMAGE DISPLAY DEVICE EQUIPPED WITH VOICE INPUT DEVICE

(54) 発明の名称: 音声入力装置およびその音声入力装置を備えた画像表示装置



(57) Abstract: Provided is a voice input device that is capable of reducing false recognition of the voice of a user from whom the voice is to be obtained in a situation where a voice source not intended by a recording person is present, for example, when a speaking person is present in the same direction as the user and at a different distance. In the case of this voice input device, a camera and a microphone array are arranged such that a camera reference point (41) and a microphone array reference point (22) are separated from each other by a predetermined distance (L), and a camera-user angle (θ) and a camera-user distance (D) with reference to the camera are calculated on the basis of an image input from the camera. Then, the voice input device calculates a microphone array-user angle (α) with reference to the microphone array based on the camera-user angle (θ) and the camera-user distance (D) and controls the directional angle of the microphone array so as to match the microphone array-user angle (α).

(57) 要約: 音声取得対象となるユーザと同じ方向で違う距離に発話中の人物など、録音者の意図しない音源がある場合に、ユーザ音声の誤認識を低減することが可能な音声入力装置を提供する。音声入力装置は、カメラの基準点(41)とマイクアレイの基準点(22)が所定距離(L)だけ離間するようにカメラとマイクアレイを配置し、カメラで入力した画像に基づいてカメラを基準としたカメラ基準ユーザ角度 θ およびカメラ基準ユーザ距離Dを算出する。そして、音声入力装置は、カメラ基準ユーザ角度 θ およびカメラ基準ユーザ距離Dに基づいて、マイクアレイを基準としたマイクアレイ基準ユーザ角度 α を算出し、マイクアレイ基準ユーザ角度 α になるようマイクアレイの指向角度を制御する。

明 細 書

発明の名称：

音声入力装置およびその音声入力装置を備えた画像表示装置

技術分野

[0001] 本発明は、複数のマイクロフォンを有するマイクアレイを備え、特定の方向からの音声を入力することが可能な音声入力装置、およびその音声入力装置を備えた画像表示装置に関する。

背景技術

[0002] 複数のマイクロフォンを有するマイクアレイを利用して音声入力の指向角度を変化させ、特定の方向からの音声情報を強調して、それ以外の方向からの音声情報を抑制する方法が知られている。これを音声入力装置に搭載し、マイクアレイの指向角度を音声が発生している方向へ向けることで、音声認識の認識精度を向上することができる。このような方法により音声入力を行うときは、ユーザによって発話された音声の方向を特定して、発話された音声の方向へマイクアレイの指向角度を向ける。

[0003] しかし、音声情報のみから方向を特定するため、ユーザ以外の人物の音声情報など、ユーザが意図していない音声情報の発生方向に対しても指向角度が向けられてしまい、ユーザが意図しない音声情報が入力され認識されてしまうことがある。

[0004] そこで、カメラで撮影した画像情報からユーザの方向を検出して、その方向にのみ音声入力の指向角度を向けることで、意図しない音声が入力されることを低減する技術が提案されている（特許文献1を参照）。特許文献1に記載の技術では、カメラにより撮影された撮像画像を用いて動的に変化する話者方向を特定し、マイクロフォンの指向性方向を話者方向になるよう制御することで、音声認識の精度を向上させている。

先行技術文献

特許文献

[0005] 特許文献1：特開2009-225379号公報

発明の概要

発明が解決しようとする課題

[0006] しかしながら、特許文献1に記載の技術では次のような課題を有する。

カメラに対して音声取得対象のユーザの後ろに別の発話者が存在するときなど、カメラで検出したユーザの方向と同じ方向に録音者の意図しない音源が存在する場合、マイクアレイの指向角度をユーザの方向に向けてもユーザの意図しない音声が入力されてしまい、音声認識処理において誤認識が発生してしまうことがある。つまり、この技術では、例えば対象となるユーザと違う距離にいる人物の音声など、ユーザと同じ方向からのユーザの音声でない音も拾って認識してしまい、ユーザ音声の誤認識を起こしてしまう。

[0007] 本発明は、上述のような実状に鑑みてなされたものであり、その目的は、音声取得対象となるユーザと同じ方向で違う距離に発話中の人物など、録音者の意図しない音源がある場合に、ユーザ音声の誤認識を低減することが可能な音声入力装置、およびその音声入力装置を備えた画像表示装置を提供することにある。

課題を解決するための手段

[0008] 上記の課題を解決するために、本発明の第1の技術手段は、音声を取得するマイクロフォンを複数有するマイクアレイと、撮像素子を有し撮影により画像を入力するカメラと、該カメラで入力された該画像に基づいて前記マイクアレイの指向角度を制御する指向角度制御部と、を備えた音声入力装置であって、前記カメラと前記マイクアレイが所定距離だけ離間して配置され、前記音声入力装置は、前記画像に基づいて、前記カメラを基準とした音声取得対象ユーザのユーザ位置の角度であるカメラ基準ユーザ角度を算出するカメラ基準ユーザ角度算出部と、前記画像に基づいて、前記カメラを基準とした前記ユーザ位置までの距離であるカメラ基準ユーザ距離を算出するカメラ基準ユーザ距離算出部と、前記所定距離、前記カメラ基準ユーザ角度、および前記カメラ基準ユーザ距離に基づいて、前記マイクアレイを基準としたユ

ーザ位置の角度であるマイクアレイ基準ユーザ角度を算出するマイクアレイ基準ユーザ角度算出部と、をさらに備え、前記指向角度制御部は、前記マイクアレイ基準ユーザ角度になるよう前記マイクアレイの指向角度を制御することを特徴としたものである。

[0009] 本発明の第2の技術手段は、第1の技術手段において、前記マイクアレイ基準ユーザ角度算出部は、前記所定距離と、前記カメラ基準ユーザ距離と、前記カメラ基準ユーザ角度と、前記カメラ基準ユーザ角度に応じた角度算出誤差範囲とに基づいて、前記マイクアレイ基準ユーザ角度範囲を算出し、前記指向角度制御部は、前記マイクアレイ基準ユーザ角度範囲になるよう前記マイクアレイの指向角度を制御することを特徴としたものである。

[0010] 本発明の第3の技術手段は、第1の技術手段において、前記マイクアレイ基準ユーザ角度算出部は、前記所定距離と、前記カメラ基準ユーザ角度と、前記カメラ基準ユーザ距離と、前記カメラ基準ユーザ角度および／または前記カメラ基準ユーザ距離に応じた距離算出誤差範囲とに基づいて、前記マイクアレイ基準ユーザ角度範囲を算出し、前記指向角度制御部は、前記マイクアレイ基準ユーザ角度範囲になるよう前記マイクアレイの指向角度を制御することを特徴としたものである。

[0011] 本発明の第4の技術手段は、第1～第3のいずれか1の技術手段において、前記画像における被写体の顔領域を検出し、該顔領域の位置を前記ユーザ位置として、前記カメラ基準ユーザ角度および／または前記カメラ基準ユーザ距離を算出することを特徴としたものである。

[0012] 本発明の第5の技術手段は、画像表示装置において、第1～第4のいずれか1の技術手段における音声入力装置を備えたことを特徴としたものである。

発明の効果

[0013] 本発明によれば、指向角度を制御できる音声入力装置において、音声取得対象となるユーザと同じ方向で違う距離に発話中の人物など、録音者の意図しない音源がある場合に、その距離を判定して目的の距離にいるユーザの音

声のみを入力することでユーザ音声の誤認識を低減し、低ノイズでユーザ音声を取得することができる。

図面の簡単な説明

[0014] [図1]本発明の実施形態1に係る音声入力装置の一構成例を示すブロック図である。

[図2]図1の音声入力装置におけるマイクアレイの一構成例を示す俯瞰図である。

[図3]図1の音声入力装置におけるマイクアレイの他の構成例を示す俯瞰図である。

[図4]図1の音声入力装置におけるカメラとマイクアレイの配置関係の一例を示す俯瞰図である。

[図5]図1の音声入力装置におけるカメラ基準ユーザ角度算出方法の一例を説明するための透視投影カメラモデルの図である。

[図6]図1の音声入力装置におけるマイクアレイ基準ユーザ角度の算出方法を説明するための図で、カメラとマイクアレイの中間にユーザが位置する場合の一例を示す俯瞰図である。

[図7]図1の音声入力装置におけるマイクアレイ基準ユーザ角度の算出方法を説明するための図で、カメラとマイクアレイから見て同じ方向かつマイクアレイに近い場所にユーザが位置する場合の一例を示す俯瞰図である。

[図8]図1の音声入力装置におけるマイクアレイ基準ユーザ角度の算出方法を説明するための図で、カメラとマイクアレイから見て同じ方向かつカメラに近い場所にユーザが位置する場合の一例を示す俯瞰図である。

[図9]本発明の実施形態2に係る音声入力装置におけるマイクアレイの一構成例を示す図である。

[図10]本発明の実施形態2に係る音声入力装置におけるマイクアレイ基準ユーザ角度範囲の算出方法の一例を説明するための図で、算出誤差範囲が大きい場合の一例を示す俯瞰図である。

[図11]本発明の実施形態2に係る音声入力装置におけるマイクアレイ基準ユ

一ザ角度範囲の算出方法の一例を説明するための図で、算出誤差範囲が小さい場合の一例を示す俯瞰図である。

[図12]本発明の実施形態4に係る音声認識装置の一構成例を示すブロック図である。

[図13]本発明の実施形態5に係る画像表示装置の一構成例を示すブロック図である。

発明を実施するための形態

[0015] (実施形態1)

以下、図1～図8を参照し、本発明の実施形態1について具体的に説明する。

図1は、本実施形態に係る音声入力装置の一構成例を示すブロック図で、図2は、図1の音声入力装置におけるマイクアレイの一構成例を示す俯瞰図で、図3は、図1の音声入力装置におけるマイクアレイの他の構成例を示す俯瞰図である。

[0016] 図1に示すように、本実施形態における音声入力装置1は、マイクアレイ11、カメラ12、画像内ユーザ座標検出部13、カメラ基準ユーザ角度算出部14、カメラ基準ユーザ距離算出部15、マイクアレイ基準ユーザ角度算出部16、および指向角度制御部17を備える。

[0017] マイクアレイ11は、音声を取得するマイクロフォンを複数有する。各マイクロフォンは音声入力素子であればよいが、図2では、マイクアレイ11は複数の超指向性マイクロフォン21を配置して構成した例を挙げている。各超指向性マイクロフォン21は、ガンマイクなどの鋭い指向性をもつマイクロフォンであり、空間上の任意の点を中心に、各超指向性マイクロフォン21の集音方向23が略放射状になるように配置される。本実施形態では、集音方向23の中心となる点をマイクアレイ基準点22と呼ぶ。ただし、マイクアレイ基準点22は正確に1点である必要はなく、各集音方向が略同一領域で交差するようにすればよい。

[0018] また、本実施形態では、マイクアレイ11の指向角度の基準とする方向を

マイクアレイ 1 1 が配列する中央の方向（配列の中心とマイクアレイ基準点 2 2 とを結ぶ方向）に決め、以下、その方向をマイクアレイ正面方向 2 4 と呼ぶ。また、マイクアレイ基準点 2 2 を通りマイクアレイ正面方向 2 4 に伸びる仮想的な直線をマイクアレイ軸 2 5 と呼ぶ。つまり、配列の中心とマイクアレイ基準点とを結ぶ直線がマイクアレイ軸 2 5 となる。マイクアレイ 1 1 をこのように構成にしておくことで、入力される音声情報の指向角度を指向角度制御部 1 7 により制御することができる。

[0019] なお、本実施形態では図 2 に示す構成により指向角度の制御を実現しているが、指向角度を制御する方法はこれに限定されない。例えば、マイクアレイ 1 1 として指向性の弱いマイクロフォンを複数利用し、各マイクロフォンに到達する音声の時間や音量の差を計算することで指向角度の制御を行うマイクアレイを利用してもよい。図 3 で例示するマイクロフォンの配置は、各マイクロフォン 3 1 に到達する音声の時間差（音声到達時間の差）を計算することで指向角度の制御を行い、マイクアレイ 1 1 を実現する場合の配置例である。マイクロフォン 3 1 は、指向性が弱いマイクロフォンか、指向性が無いマイクロフォンである。指向方向 3 2（指向角度）を決定すると、各マイクロフォン 3 1 への音声の到達時間に差ができる。各マイクロフォン 3 1 の音声をこの到達時間の差だけずらして重ね合わせることで、指向方向 3 2 からの音声成分のみを強調することができる。指向方向 3 2 を決定する際の基準とする点をマイクアレイ基準点 2 2 とし、図 2 の構成と同様にマイクアレイ正面方向 2 4 とマイクアレイ軸 2 5 を定義する。マイクアレイ 1 1 をこのような構成にしておくことで、入力される音声情報の指向角度を指向角度制御部 1 7 により制御することができる。

[0020] カメラ 1 2 は、撮像素子を有し、撮影により画像、または画像およびその画像についての撮影情報（例えばその画像における焦点距離などを含む）を入力する画像入力部である。上記の撮像素子の例としては、CCD（Charge Coupled Device）センサやCMOS（Complementary Metal Oxide Semiconductor）センサなどの固体撮像素子が挙げられる。カメラ 1 2 は、撮像素子の

ほかに、レンズなどの光学部品などを備え、カメラモジュールとして音声入力装置 1 に搭載することができる。カメラ 1 2 としては、超広角レンズや魚眼レンズを装着したカメラなど、撮影画角が広いカメラを利用すると、広い範囲で音声取得対象ユーザ（以下、単に「ユーザ」と呼ぶ）を撮影できるため好適である。

[0021] そして、本実施形態におけるカメラ 1 2 は、撮影した画像（および撮影情報）を画像内ユーザ座標検出部 1 3 に伝達する。なお、カメラ 1 2 は、静止画撮影専用である必要はなく、動画撮影専用であっても、動画・静止画兼用であってもよい。また、画像内ユーザ座標検出部 1 3 で使用する画像は、基本的に静止画であり、動画像を利用する場合には音声取得時直前のフレームを利用するなどすればよい。

[0022] 図 4 は、本実施形態におけるカメラ 1 2 とマイクアレイ 1 1 の配置関係の一例を示す俯瞰図である。図 4 に示すように、カメラ 1 2 とマイクアレイ 1 1 はそれらの基準点 4 1, 2 2 間が所定距離 L だけ離間して配置されている。ユーザ位置の変化は上下方向である鉛直方向への変化よりも左右方向である水平方向への変化が大きいため、所定距離 L が水平方向に向くようにカメラ 1 2 とマイクアレイ 1 1 とを水平方向に離間させて配置させるとよい。これは、離間した方向に対して、後述するように指向方向を制御することにより、ノイズを低減した所望の音声を入力音声として出力することが可能となるためである。カメラ基準点 4 1 は、カメラ 1 2 の撮影範囲 4 0 内においてカメラ 1 2 に対してユーザが存在する角度（方向）を算出する際に基準とする点である。本実施形態ではカメラ焦点をカメラ基準点 4 1 とした例を挙げるが、これに限ったものではない。無論、カメラ 1 2 とマイクアレイ 1 1 とは水平方向以外の方向に離間させて配置させ、それに応じて指向方向の制御を行うこともできる。

[0023] カメラ 1 2 とマイクアレイ 1 1 は、カメラ基準点 4 1 とマイクアレイ基準点 2 2 が水平方向に同一位置にならないように配置する。これは、上述したようにユーザ位置の変化が鉛直方向より水平方向への方が大きいためである

。

[0024] また、カメラ１２とマイクアレイ１１との位置関係は、両者が音声入力装置１に内蔵されている場合、その位置関係は設計や製造の段階で既知である。一方、カメラ１２とマイクアレイ１１とが独立して可動できるような構造の場合、それらの位置関係が変化してしまうため、カメラ１２とマイクアレイ１１との位置関係をあらかじめ計測しておき、計測後に変化しないように固定する。本実施形態では、カメラ光軸４２とマイクアレイ軸２５が略平行になり、カメラ撮影方向４３とマイクアレイ正面方向２４が同じ方向になり、かつ、カメラ基準点４１を通りカメラ光軸４２に対して略垂直な直線４４上にマイクアレイ基準点２２が存在するような位置関係でマイクアレイ１１とカメラ１２を設置する。また、マイクアレイ１１の向きは集音方向２３（図２）が放射状に広がっている同一平面上に直線４４が存在するように設置する。

[0025] 画像内ユーザ座標検出部１３は、カメラ１２で撮影された画像（入力画像）内から、ユーザの位置情報の検出を行う。ユーザ位置は、例えば、画像情報から顔領域を検出することで検出することが可能である。画像内のユーザの顔の二次元座標、および、画像内のユーザの顔の大きさを算出する。顔を検出する技術は一般に利用されている方法を使用することができる。例えば、標準的な人物の顔が撮影された画像を、参照データとしてあらかじめ保有しておき、その参照データとの相関値から顔を検出する方法がある。ここで、人物の顔の画像としては、標準的な男性、女性の双方の顔の画像を参照データとして保有しておいてもよいし、さらに世代別の標準的な男性、女性の顔の画像を参照データとして保有しておいてもよい。なお、特定のユーザの顔を保有しておくこともできるが、そのような例については実施形態３として後述する。

[0026] 画像内ユーザ座標検出部１３は、画像内のユーザの顔が存在する二次元座標をカメラ基準ユーザ角度算出部１４に伝達し、画像内のユーザの顔の大きさをカメラ基準ユーザ距離算出部１５へ伝達する。顔領域の検出は、声を発

している口が存在するため、ユーザ位置を検出するための処理として好ましい。無論、顔中の口の領域を代わりに検出してもよい。

[0027] 画像内に複数のユーザが存在する場合、画像内ユーザ座標検出部 13 は、画像内のそれぞれのユーザの顔が存在する二次元座標をカメラ基準ユーザ角度算出部 14 に伝達し、画像内のそれぞれのユーザの顔の大きさをカメラ基準ユーザ距離算出部 15 へ伝達する。このとき、多数のユーザが検出されるとデータ量が多くなるので、検出された顔の二次元座標位置や検出された顔の大きさによって、特定の数の検出結果を伝達するようにしてもよい。

[0028] カメラ基準ユーザ角度算出部 14 は、入力画像に基づいてカメラ基準ユーザ角度を算出する。ここで、カメラ基準ユーザ角度とは、カメラ 12 を基準とした音声取得対象ユーザのユーザ位置の角度である。

[0029] 本実施形態におけるカメラ基準ユーザ角度算出部 14 は、画像内ユーザ座標検出部 13 から伝達された画像内の二次元座標である画像内ユーザ顔座標から、カメラ基準点 41 を基準とするユーザが存在する方向を示す角度であるカメラ基準ユーザ角度の算出を行う。カメラ基準ユーザ角度は、透視投影カメラモデルを利用することで算出される。

[0030] 図 5 を参照してこの算出方法について説明する。図 5 は、音声入力装置 1 におけるカメラ基準ユーザ角度算出方法の一例を説明するための透視投影カメラモデルの図である。この透視投影カメラモデルは、カメラ光軸 42 と直線 44 に垂直な方向から見たカメラ 12 のモデルである。カメラ 12 の焦点距離、センサ（撮像素子）の画素ピッチ、センサ画素と光軸 42 の交差点座標が既知ならば、投影面 51 の位置を算出することができる。カメラ基準点 41 から投影面 51 上の画像内ユーザ顔座標 52 の方向 53 が示す角度が、カメラ基準ユーザ角度 θ となる。このようにして、カメラ基準ユーザ角度 θ が算出できる。カメラ基準ユーザ角度算出部 14 は、算出したカメラ基準ユーザ角度 θ をマイクアレイ基準ユーザ角度算出部 16 へ伝達する。

[0031] なお、カメラ 12 が魚眼カメラであった場合には、投影面 51 がカメラ基準点 41 を中心とした球である投影カメラモデルとなり、魚眼カメラの画角

およびイメージサークル半径の画素数、センサ画素と光軸42の交差点座標が既知ならば、投影面51の位置を算出することができるため、同じ方法でカメラ基準ユーザ角度 θ を算出することができる。

[0032] カメラ基準ユーザ距離算出部15は、入力画像に基づいて、カメラ基準ユーザ距離を算出する。ここで、カメラ基準ユーザ距離とは、カメラ12を基準としたユーザ位置までの距離である。

[0033] 本実施形態におけるカメラ基準ユーザ距離算出部15は、画像内ユーザ座標検出部13から伝達された、画像内のユーザの顔の大きさから、カメラ基準点からユーザまでのカメラ基準ユーザ距離を算出する。カメラ基準ユーザ距離は、画像内ユーザ座標検出部13で参照データとして利用する人物の顔が撮影された画像を、カメラ12を利用して撮影し、撮影時の距離を計測しておくことで、次の式(1)の計算により算出することができる。

$$D = d \cdot (t / t') \quad \dots (1)$$

[0034] 式(1)の各変数は、Dが算出するカメラ基準ユーザ距離、dが参照データ撮影時のカメラ基準ユーザ距離、tが参照データの縦の画素数、t'が画像内ユーザ座標検出部13から送られた抜き出し範囲の縦の画素数である。

[0035] ここで、顔の大きさは年齢や性別により異なるため、性別や世代別の参照データ（顔の大きさが異なった参照データ）を保持していない場合には、画像内ユーザ座標検出部13で検出する顔から年齢や性別を推定して、カメラ基準ユーザ距離Dを算出すると、距離算出の誤差を低減できるため好適である。これは、基準の年齢の顔の大きさと距離の関係を設定しておき、推定される年齢によって距離を増減させることにより実現することができる。また、年齢ごとに顔の大きさと距離の関係を設定することでも実現することができる。例えば、同じ大きさの顔が検出された場合、子供と大人では距離が異なり、子供の方が大人よりも近くに存在していることになる。そこで、例えば検出された顔が子供であるときは、算出された距離をあらかじめ定められた割合で大きくすればよい。カメラ基準ユーザ距離算出部15は、算出したカメラ基準ユーザ距離Dをマイクアレイ基準ユーザ角度算出部16へ伝達する

。

[0036] また、本実施形態では、画像中のユーザ顔の大きさからカメラ基準点41からユーザまでの距離 D を算出しているが、カメラ基準点41からユーザまでの距離を算出する方法はこの方法に限定されず、カメラ12でその画像の取得時に取得した各種の撮影情報を利用することもできる。例えば、カメラ12がオートフォーカス機能を搭載していれば、顔領域を検出し検出された顔領域にフォーカスを合わせたときのフォーカス距離をユーザ位置の距離とすることで、ユーザ位置の距離を算出することができる。また、赤外線を利用したTOF (Time Of Flight) や、カメラ12として多視点カメラを利用し、ユーザ顔領域の視差情報を撮影情報として得て、その視差情報からカメラ基準点41からユーザまでの距離を算出してもよい。さらに、上述した様々な距離算出方法を組み合わせて適用することもできる。

[0037] 本実施形態における画像内ユーザ座標検出部13では、入力画像における被写体（主に人物であるが動物であってもよい）の顔領域を検出し、カメラ基準ユーザ角度算出部14やカメラ基準ユーザ距離算出部15において、その顔領域の位置を上記ユーザ位置として、カメラ基準ユーザ角度 θ やカメラ基準ユーザ距離 D を算出している。ただし、カメラ基準ユーザ角度 θ またはカメラ基準ユーザ距離 D の一方だけを顔領域の検出に基づいて算出するようにしてもよい。もしくは、カメラ基準ユーザ角度 θ またはカメラ基準ユーザ距離 D のいずれも、顔領域検出を利用せずにユーザ位置を検出することで算出することもできる。なお、この点については、他の実施形態でも同様である。

[0038] なお、カメラ基準ユーザ角度算出部14およびカメラ基準ユーザ距離算出部15とは別に画像内ユーザ座標検出部13を設けた構成例を挙げたが、この構成例は、画像内ユーザ座標検出部13のようなユーザ位置を特定する機能をカメラ基準ユーザ角度算出部14およびカメラ基準ユーザ距離算出部15の双方にもたせることと同義である。

[0039] マイクアレイ基準ユーザ角度算出部16は、所定距離 L 、カメラ基準ユー

ザ角度算出部 14 から伝達されたカメラ基準ユーザ角度 θ 、およびカメラ基準ユーザ距離算出部 15 から伝達されたカメラ基準ユーザ距離 D に基づいて、マイクアレイ基準ユーザ角度の算出を行う。ここで、マイクアレイ基準ユーザ角度とは、マイクアレイ 11 を基準とした上記ユーザ位置の角度であり、この例ではマイクアレイ基準点 22 を基準としたユーザ角度である。

[0040] 図 6 は、このマイクアレイ基準ユーザ角度の算出方法を説明するための図で、カメラ 12 とマイクアレイ 11 の間にユーザが位置する場合の一例を示す俯瞰図である。カメラ基準ユーザ角度 θ とカメラ基準ユーザ距離 D が算出されれば、マイクアレイ基準点 22 に対するユーザ 61 の位置を幾何学的な計算により算出することができる。具体的には、カメラ基準点 41 とマイクアレイ基準点 22 の位置関係は、音声入力装置 1 を設計するときに上記所定距離 L を含めてあらかじめ設定されるため、マイクアレイ基準点 22 を基準としたユーザのマイクアレイ基準ユーザ角度 α を算出することができる。例えば、カメラ基準点 41 から見てユーザの後ろに別の発話者 62 が発話しており、カメラで別の発話者 62 を検出できない場合にも、ユーザ 61 のみにマイクアレイ 11 の指向角度を向けて適切にユーザ 61 の音声を入力することができる。マイクアレイ基準ユーザ角度算出部 16 で算出したマイクアレイ基準ユーザ角度 α は、指向角度制御部 17 へ伝達される。

[0041] また、図 6 ではユーザ 61 がカメラ 12 とマイクアレイ 11 の間に位置する場合で説明したが、各基準点から同じ方向にユーザがいる場合でも同様に適用することができる。図 7 および図 8 を参照してこの点について説明する。図 7 は、カメラ 12 とマイクアレイ 11 から見て同じ方向かつマイクアレイ 11 に近い場所にユーザが位置する場合の一例を示す俯瞰図で、図 8 は、カメラ 12 とマイクアレイ 11 から見て同じ方向かつカメラ 12 に近い場所にユーザが位置する場合の一例を示す俯瞰図である。

[0042] 図 7 で例示するように各基準点 22, 41 から見て右側にユーザ 61 が位置する場合でも、図 8 で例示するように各基準点 22, 41 から見て左側にユーザ 61 が位置する場合でも、図 6 の場合と同様に、カメラ基準ユーザ角

度 θ とカメラ基準ユーザ距離 D から、マイクアレイ基準点22に対するユーザ61の位置を幾何学的な計算により算出することができる。

[0043] このように、カメラ基準点41とマイクアレイ基準点22の位置関係を利用して、マイクアレイ基準点22を基準としたユーザ61のマイクアレイ基準ユーザ角度 α を算出することができる。ただし、カメラ基準点41とマイクアレイ基準点22が同じ位置になるようにカメラ12とマイクアレイ11を配置すると、別の発話者62がユーザ61の後方にいる場合など、カメラから見てほぼ同じ角度にいる場合、指向角度制御によりユーザの音声を分離し独立に音声入力することができない。また、カメラ12の撮影画像もユーザが重なり検出することができない。したがって、カメラ基準ユーザ角度算出部14の算出した角度が変化する方向に、所定距離 L だけ離れた状態でカメラ12とマイクアレイ11を配置する。これにより、別の発話者62がカメラ12から見てユーザ61とほぼ同じ角度にいる場合に、各発話者の音声を独立に音声入力することが可能となる。なお、所定距離 L はどのような値でもよい。

[0044] 指向角度制御部17は、カメラ12で入力された画像に基づいてマイクアレイ11の指向角度を制御するが、本発明の主たる特徴として、上記のマイクアレイ基準ユーザ角度 α になるようマイクアレイ11の指向角度を制御する。つまり、指向角度制御部17は、マイクアレイ基準ユーザ角度算出部16から伝達されたマイクアレイ基準ユーザ角度 α に向けてマイクアレイ11の指向角度制御を行い、マイクアレイ11から伝達された音声を出力する。なお、図6～図8においては便宜上、正負を問わず、角度 θ 、 α がそれぞれ軸42、25とのなす角度であるものとして表現しているが、角度が0度の位置やいずれが正方向であるかは任意に決めておけばよい。

[0045] 本実施形態においては、マイクアレイ11から送られた超指向性マイクロフォン21（または図3のマイクロフォン31）の音声情報から、マイクアレイ基準ユーザ角度 α に該当する音声を選択することで、このような制御が実現できる。ここで、マイクアレイ11が指向性の弱いマイクロフォンを利

用している場合には、音声の到達時間の差を考慮することで、特定角度からの音声を音声情報として取得することができる。

[0046] また、本実施形態では、角度 θ や角度 α をそれぞれ一つの値として説明したが、算出される値には誤差や幅が存在するので、各角度には所定の許容範囲を設定することができる。例えば、算出される角度の許容範囲を例えば $\pm 5^\circ$ などの一定範囲としてあらかじめ設定しておくことができる。なお、この許容範囲についてはあらかじめ角度として定める必要や一定範囲とする必要はなく、そのような例については実施形態2として後述する。

[0047] 以上で説明したように、カメラ基準ユーザ角度 θ が変化する面において、所定距離 L だけ離れた状態でカメラ12とマイクアレイ11を配置し、カメラ12から目的の距離（上記の距離 D ）だけ離れた位置にマイクアレイ11の指向角度を向けることで、別のユーザ（別の発話者）62がカメラ12から見てユーザ61とほぼ同じ角度にいる場合に、各ユーザの音声を独立に音声入力することが可能となり、音声取得対象となるユーザの音声のみを別のユーザの音声と分けて入力することでユーザ音声の誤認識を低減することができ、好適である。なお、入力画像として動画像を適用することで、逐次、そのユーザの位置に合った方向に指向角度が切り替わるため、ユーザが移動したとしても誤認識を低減できるといった効果を奏する。

[0048] このように、本発明によれば、指向角度を制御できる音声入力装置において、音声取得対象となる発話者（ユーザ）と同じ方向で違う距離に発話中の人物など、録音者（音声入力装置を使用するユーザであって、基本的には上記発話者とは異なる者）の意図しない音源がある場合に、その距離を判定して目的の距離にいるユーザの音声のみを入力することでユーザ音声の誤認識を低減し、低ノイズでユーザ音声を取得することができる。

[0049] （実施形態2）

以下、図9～図11を併せて参照し、本発明の実施形態2について具体的に説明する。

図9は、本実施形態に係る音声入力装置におけるマイクアレイの一構成例

を示す図である。図 10 および図 11 は、本実施形態に係る音声入力装置におけるマイクアレイ基準ユーザ角度範囲の算出方法の一例を説明するための図である。図 10 には算出誤差範囲が大きい場合の一例を示し、図 11 には算出誤差範囲が小さい場合の一例を示している。

[0050] 本実施形態における音声入力装置の構成例を示すブロック図は、実施形態 1 で示した図 1 と同様であり、各部の詳細な説明は省略する。また、マイクアレイ 11 やカメラ 12 の動作も実施形態 1 で示したものと同様である。

[0051] 本実施形態における指向角度制御部 17 は、図 1 のマイクアレイ 11 の指向範囲（指向角度が示す範囲）を制御することが可能に構成されており、マイクアレイ 11 もそのような制御が可能な複数のマイクロフォンで構成されている。本実施形態では、マイクアレイ 11 として、図 9 で例示するように、指向性が弱いまたはない複数のマイクロフォン 31 を利用することが望ましく、そのような例を挙げる。ただし、指向性が強い複数のマイクロフォンであっても利用できる。

[0052] 指向角度制御部 17 では、音声を取得したい指向範囲（指向領域）91 に入っている各指向方向（指向角度）の音声を、実施形態 1 と同様の方法で強調し、その他の角度の音声を抑制する。ここでは、図 3 の例と同様に複数のマイクロフォン 31 への音声到達時間の差を利用することで指向範囲の制御を行っている。これにより、指向角度制御部 17 は、マイクアレイ 11 の指向範囲を制御することができる。以下、具体的にこの制御に関して説明する。なお、指向性が強い複数のマイクロフォンであっても、指向角度が示す指向範囲に向いているような 1 または複数のマイクロフォンを駆動させるなどして、このような制御は可能である。

[0053] カメラ 12 は、実施形態 1 と同様に撮影した画像（および撮影情報）を画像内ユーザ座標検出部 13 に伝達する。ここで、例えば、カメラ 12 として魚眼カメラを利用する場合、撮影画角が広いため広範囲でユーザの顔を検出することができるが、画像が円形になるため画像中心と画像周辺とで解像度が異なる。そのため、カメラ光軸を基準としたカメラ基準ユーザ角度（図 6

の角度 θ)の角度値が小さいときと大きいときとは、カメラ12の撮影画像の解像度が異なるため角度算出精度が異なる。そこで、本実施形態では、角度の算出精度が異なっても所望の効果が適切に得られるように、カメラ基準ユーザ角度算出部14が角度算出誤差範囲を考慮した動作をする。

[0054] カメラ基準ユーザ角度算出部14は、画像内ユーザ顔座標からカメラ基準ユーザ角度 θ を算出するときに許容範囲を設定しておくことで、角度算出誤差による動作不良を低減することができる。そこで、カメラ基準ユーザ角度算出部14は、画像内ユーザ顔座標からカメラ基準ユーザ角度 θ を算出するときに、算出された角度 θ における角度算出誤差範囲の算出も行う。

[0055] 角度算出誤差範囲は、例えば、カメラ12において、画像内ユーザ顔座標が左右に1画素ずれたときのカメラ基準ユーザ角度 θ_a 、 θ_b (図10を参照)を算出し、そのカメラ基準ユーザ角度 θ_a 、 θ_b を境界とした範囲を角度算出誤差範囲101とする。このとき、あるカメラ基準ユーザ角度において、撮影画像の解像度が高く角度算出誤差が小さい場合は、画像内ユーザ顔座標が1画素ずれたときの角度変化が小さい。一方、撮影画像の解像度が低く角度算出誤差が大きい場合は、画像内ユーザ顔座標が1画素ずれたときの角度変化が大きくなる。つまり、算出されるカメラ基準ユーザ角度 θ により角度算出誤差範囲が異なってくる。

[0056] 比較のために、実施形態1で説明したように、算出されるカメラ基準ユーザ角度 θ に依らずに誤差範囲が一定であるとして誤差の許容範囲を設定した場合について説明する。このような設定の場合、入力画像中心の解像度が高い領域の角度算出誤差範囲を基準に設定すると、入力画像の周辺領域(縁側)では解像度が低いため、角度算出誤差が大きくなり所望の音声除外されてしまう可能性がある。一方、入力画像の周辺に位置する解像度が低い領域の角度算出誤差範囲を基準に設定すると、画像中心の領域では解像度が高く角度算出誤差が小さいにも拘わらず過度の許容範囲を設けたために、所望の音声にノイズが含まれてしまう可能性がある。

[0057] このような事態を防ぐために、本実施形態では、算出されるカメラ基準ユ

一ザ角度 θ に依って角度算出誤差範囲が異なるものとし、誤差の許容範囲をその角度 θ に依存して設定し、入力画像の中心では角度算出誤差範囲を小さく、入力画像の周辺領域では角度算出誤差範囲を大きくする。これにより、結果的にノイズ低減と所望音声取得を適切に行うことが可能となる。

[0058] ここで、角度算出誤差範囲を左右に1画素ずれた場合の範囲として説明したが、カメラの解像度など撮影する条件により角度算出誤差範囲を適宜設定することができる。また、算出された角度 θ から角度算出誤差範囲を毎回算出することもできるが、算出された角度 θ ごとに角度算出誤差範囲を定めたLUT (Look Up Table) を保持してもよい。

[0059] 本実施形態におけるカメラ基準ユーザ角度算出部14は、算出したカメラ基準ユーザ角度 θ および角度算出誤差範囲101を、マイクアレイ基準ユーザ角度算出部16へ伝達する。

[0060] カメラ基準ユーザ距離算出部15は、画像内ユーザ座標検出部13から伝達された画像内ユーザの顔の大きさからカメラ基準ユーザ距離Dを算出するときに、許容範囲を設定しておくことで、距離算出誤差による動作不良を低減することができる。そこで、カメラ基準ユーザ距離算出部15は、画像内ユーザ顔座標からカメラ基準ユーザ距離Dを算出するときに、距離算出誤差範囲の算出も行う。

[0061] 本実施形態では、距離算出誤差範囲を例えば、画像内のユーザの顔の大きさが左右1画素ずつ大きくなった場合と左右1画素ずつ小さくなった場合の距離 D_a 、 D_b (図10を参照) を算出し、算出した距離 D_a 、 D_b を境界とした範囲を距離算出誤差範囲102とする。このとき、カメラ基準ユーザ距離Dが短い場合、画像内ユーザ顔の大きさが1画素ずれたときのユーザ距離変化が小さい。一方、距離Dが長い場合は、画像内ユーザ顔の大きさが1画素ずれたときのユーザ距離変化が大きくなる。つまり、算出されるカメラ基準ユーザ距離Dにより距離算出誤差が異なってくる。

[0062] 比較のために、算出されるカメラ基準ユーザ距離Dに依らずに距離算出誤差範囲が一定であるとして誤差の許容範囲を設定した場合について説明する

。このような設定の場合、距離Dが短い場合の距離算出誤差範囲を基準に設定すると、距離Dが長い場合に距離算出誤差が大きくなり所望のユーザ音声除外されてしまう可能性がある。一方、距離Dが長い場合の距離算出誤差範囲を基準に設定すると、距離Dが短い場合に距離算出誤差が小さいにも拘わらず過度の許容範囲を設けたために、所望のユーザ音声にノイズが含まれてしまう可能性がある。

[0063] また、距離算出誤差範囲102は、算出時のカメラ基準ユーザ角度 θ によっても変化する。例えば、カメラ12として魚眼レンズを搭載したカメラを用いる場合、カメラ基準ユーザ角度 θ が大きくなると、顔領域の解像度が低くなる。撮影画像の解像度が高く距離算出誤差が小さい場合は、画像内ユーザ顔の大きさが1画素ずれたときのユーザ距離変化が小さい。一方、撮影画像の解像度が低く距離算出誤差が大きい場合は、画像内ユーザ顔の大きさが1画素ずれたときのユーザ距離変化が大きくなる。つまり、算出されるカメラ基準ユーザ角度 θ により距離算出誤差が異なってくる。

[0064] 比較のために、算出されるカメラ基準ユーザ角度 θ に依らずに距離算出誤差範囲が一定であるとして誤差の許容範囲を設定した場合について説明する。このような設定の場合、入力画像中心の解像度が高い領域（角度 θ が小さい領域に相当）の距離算出誤差範囲を基準に設定すると、入力画像の周辺領域では解像度が低いため、距離算出誤差が大きくなり所望のユーザ音声除外されてしまう可能性がある。一方、入力画像の周辺に位置する解像度が低い領域（角度 θ が大きい領域に相当）の距離算出誤差範囲を基準に設定すると、画像中心の領域では解像度が高く距離算出誤差が小さいにも拘わらず過度の許容範囲を設けたために、所望のユーザ音声にノイズが含まれてしまう可能性がある。

[0065] このような事態を防ぐために、本実施形態では、カメラ基準ユーザ距離Dおよびカメラ基準ユーザ角度 θ に依って距離算出誤差範囲が異なるものとし、誤差の許容範囲を距離Dおよび角度 θ に依存して設定する。画像内のユーザの顔の大きさが左右1画素ずつ大きくなった場合と左右1画素ずつ小さく

なった場合の距離 D_a 、 D_b を算出し、算出した距離 D_a 、 D_b を境界とした範囲を距離算出誤差範囲102とすることで、入力画像の中心（角度 θ が小さい領域に相当）では距離算出誤差範囲102が小さく、入力画像の周辺領域（縁側であって、角度 θ が大きい領域に相当）では距離算出誤差範囲102が大きく、また、距離 D が短い場合には距離算出誤差範囲102が小さく、距離 D が長い場合には距離算出誤差範囲102が大きくなる。これにより、結果的にノイズ低減と所望音声取得を適切に行うことが可能となり好適である。

[0066] ここで、距離算出誤差範囲を画像内のユーザの顔の大きさが左右に1画素ずれた場合の範囲として説明したが、カメラの解像度など撮影する条件により適宜設定することができる。また、算出時の角度 θ および算出された距離 D から距離算出誤差範囲を毎回算出することもできるが、算出時の角度 θ および算出された距離 D ごとに距離算出誤差範囲を定めたLUTを保持してもよい。

[0067] また、本実施形態では、距離 D および角度 θ に依って距離算出誤差範囲が異なるものとし、誤差の許容範囲を距離 D および角度 θ に依存して設定した例を挙げた。しかし、距離算出誤差範囲を距離 D または角度 θ のいずれか一方のみに依って異なるものとし、誤差の許容範囲を距離 D または角度 θ に依存して設定するようにしてもよい。この場合にも、算出された距離 D ごとに距離算出誤差範囲を定めたLUT、または算出時の角度 θ ごとに距離算出誤差範囲を定めたLUTを保持して、LUTを参照して距離算出誤差を求めてもよい。

[0068] 本実施形態のカメラ基準ユーザ距離算出部15は、算出したカメラ基準点41からユーザまでの距離であるカメラ基準ユーザ距離 D および距離算出誤差範囲102を、マイクアレイ基準ユーザ角度算出部16へ伝達する。

[0069] マイクアレイ基準ユーザ角度算出部16は、所定距離 L 、カメラ基準ユーザ角度 θ 、角度算出誤差範囲101、カメラ基準ユーザ距離 D 、および距離算出誤差範囲102に基づいて、マイクアレイ基準点22を基準としたユー

ザ角度の範囲であるマイクアレイ基準ユーザ角度範囲の算出を行う。

[0070] マイクアレイ基準ユーザ角度範囲の算出方法について、図10および図11を参照しながら説明する。マイクアレイ基準ユーザ角度算出部16は、所定距離 L と、カメラ基準ユーザ角度算出部14から伝達されたカメラ基準ユーザ角度 θ および角度算出誤差範囲101と、カメラ基準ユーザ距離算出部15から伝達されたカメラ基準ユーザ距離 D および距離算出誤差範囲102とから、ユーザ61が存在すると推定できる範囲103を幾何学的な計算により算出する。ここで、範囲103は、角度算出誤差範囲101と距離算出誤差範囲102との共通部分となる共通範囲である。

[0071] 具体的には、カメラ基準点41とマイクアレイ基準点22の位置関係は、音声入力装置1を設計するときに上記所定距離 L を含めてあらかじめ設定されるため、ユーザ61が存在すると推定できる範囲103を含むような、マイクアレイ基準点22を基準としたマイクアレイ基準ユーザ角度範囲104を算出することができる。このとき、図10のようにユーザ61の顔の大きさに対して大きい範囲103が検出されることも、図11のようにユーザ61の顔の大きさに対して小さい範囲103が検出されることもある。しかし、距離 D や角度 θ に基づく適切な誤差を鑑みて範囲103が算出されているため、この範囲103は誤差も鑑みた適切な範囲と言え、この範囲103から得られるマイクアレイ基準ユーザ角度範囲104も適切な範囲と言える。

[0072] マイクアレイ基準ユーザ角度算出部16で算出したマイクアレイ基準ユーザ角度範囲104は、指向角度制御部17へ伝達される。なお、このとき、同時にマイクアレイ基準ユーザ角度（範囲104の中央値の角度）も伝達されるようにしてもよい。

[0073] 本実施形態における指向角度制御部17は、マイクアレイ基準ユーザ角度算出部16から伝達されたマイクアレイ基準ユーザ角度範囲104の音声を取得できるようマイクアレイ11の指向範囲91を制御する。つまり、本実施形態における指向角度制御部17は、マイクアレイ11の指向角度である指向範囲91が、マイクアレイ基準ユーザ角度範囲104になるように制御

する。これにより、マイクアレイ基準ユーザ角度範囲 104 に応じて適切にマイクアレイ 11 の指向範囲を制御することができる。

[0074] 本実施形態の音声入力装置 1 によれば、カメラ 12 のカメラ基準ユーザ角度 θ ごとの解像度やカメラ基準ユーザ距離 D ごとの解像度に応じて、適切なマイクアレイ基準ユーザ角度範囲 104 を算出できるため、マイクアレイの指向範囲を適切に設定することができる。また、マイクアレイ基準ユーザ角度（範囲 104 の中央値の角度）も伝達するか、マイクアレイ基準ユーザ角度範囲 104 からその中央値の角度として求めることで、マイクアレイ基準ユーザ角度（上記の角度 α ）が示す指向方向の音声入力レベルを上げて、範囲 104 の端に近づくほど音声入力レベルを下げるなどの制御を行うこともできる。

[0075] 以上、角度算出誤差範囲 101 および距離算出誤差範囲 102 が算出された結果に対してマイクアレイ基準ユーザ角度範囲 104 を設定したが例を挙げたが、角度算出誤差範囲 101 のみまたは距離算出誤差範囲 102 のみでこのような設定を実施することもでき、いずれの場合にもノイズ低減と所望音声取得を実現する効果が得られる。角度算出誤差範囲 101 のみでこのような設定を実施した場合には、距離 D でかつ角度算出誤差範囲 101 となる円弧を挟むように、距離算出誤差範囲 102 のみでこのような設定を実施した場合には、角度 θ でかつ距離算出誤差範囲 102 となる線分を挟むように、それぞれマイクアレイ基準ユーザ角度範囲 104 が設定されることになる。

[0076] つまり、マイクアレイ基準ユーザ角度算出部 16 は、所定距離、カメラ基準ユーザ距離、カメラ基準ユーザ角度、および角度算出誤差範囲に基づいて、マイクアレイ基準ユーザ角度範囲を算出するようにしてもよい。もしくは、マイクアレイ基準ユーザ角度算出部 16 は、所定距離、カメラ基準ユーザ角度、カメラ基準ユーザ距離、および距離算出誤差範囲に基づいて、マイクアレイ基準ユーザ角度範囲を算出するようにしてもよい。

[0077] また、角度算出誤差範囲や距離算出誤差範囲を算出する例は、入力画像の

縁側の方が中央部分に比べてフォーカスが合っていない場合が多いなどの理由から、カメラ基準ユーザ角度やカメラ基準ユーザ距離の算出誤差が大きくなることが多いと言えるため、顔領域の検出以外の方法でユーザ位置を検出した場合でも有益である。

[0078] 以上で説明したように、本実施形態では、マイクアレイ基準ユーザ角度範囲を、カメラ基準ユーザ角度および／またはカメラ基準ユーザ距離に応じて、より具体的にはカメラ基準ユーザ角度ごとの解像度および／またはカメラ基準ユーザ距離ごとの解像度に応じて算出しており、これにより、マイクアレイの指向範囲を適切に制御することができる。

[0079] (実施形態3)

以下、本発明の実施形態3を具体的に説明する。

本実施形態における音声入力装置の構成例を示すブロック図は、実施形態1で示した図1と同様であり、各部の詳細な説明は省略する。また、マイクアレイ11やカメラ12の動作も実施形態1で示したものと同様である。ただし、本実施形態の特徴は実施形態2に適用することもできる。

[0080] 本実施形態における画像内ユーザ座標検出部13は、カメラ12で撮影された画像内から、ユーザ個人を特定し、そのユーザの位置情報の検出を行う。ユーザの特定は、例えば、あらかじめ保存したユーザの顔（以下、ユーザ顔）が撮影された画像について、数段階に拡大率を変えた画像をテンプレート画像としてテンプレートマッチングを行うことで実現できる。

[0081] より具体的に説明すると、ユーザ顔を検出する画像からテンプレート画像と同じ縦横の画素数の範囲をウィンドウ画像として抜き出して、ウィンドウ画像全画素について対応するテンプレート画像の画素との画素値の差分の合計値を計算する。ユーザ顔を検出する画像から抜き出せるすべてのウィンドウ画像について差分を計算し、差分が最も小さくなるウィンドウ画像が抜き出される範囲を探す。差分が最も小さくなりかつ差分が閾値以下となる抜き出し範囲のユーザ顔を検出する画像に対する二次元座標をユーザの顔の二次元座標とし、差分が最も小さくなりかつ差分が閾値以下となる抜き出し範囲

の例えば縦の画素数をユーザ顔の大きさとする。差分が最も小さくなりかつ差分が閾値以下となる抜き出し範囲が存在しないときは画像内にユーザ顔が存在しないと判定する。

[0082] 画像内ユーザ座標検出部 13 は、画像内のそれぞれの画像内ユーザ顔座標およびユーザ特定情報をカメラ基準ユーザ角度算出部 14 に伝達し、画像内のそれぞれのユーザ顔について、大きさおよびユーザ特定情報（どのユーザであるかを示す ID または氏名など）をカメラ基準ユーザ距離算出部 15 へ伝達する。

[0083] カメラ基準ユーザ角度算出部 14 は、特定されたユーザの画像内ユーザ顔座標から、カメラ基準ユーザ角度の算出を行い、カメラ基準ユーザ角度がどのユーザのものかユーザ特定情報と関連付ける。カメラ基準ユーザ角度の算出は実施形態 1 と同様の方法で実現できる。カメラ基準ユーザ角度算出部 14 は、ユーザ特定情報が付与されたカメラ基準ユーザ角度を、マイクアレイ基準ユーザ角度算出部 16 へ伝達する。

[0084] カメラ基準ユーザ距離算出部 15 は、画像内ユーザ座標検出部 13 から伝達された、画像内のユーザの顔の大きさから、カメラ基準点からそれぞれのユーザまでのカメラ基準ユーザ距離の算出を行い、カメラ基準ユーザ距離がどのユーザのものかユーザ特定情報と関連付ける。カメラ基準ユーザ距離は、画像内ユーザ座標検出部 13 でテンプレート画像として利用するユーザ顔が撮影された画像を、カメラ 12 を利用して撮影し、撮影時の距離を計測しておくことで、実施形態 1 と同様に式（1）の計算により算出することができる。

[0085] 本実施形態における式（1）の各変数は、 D が算出するカメラ基準ユーザ距離、 d がテンプレート画像撮影時のカメラ基準ユーザ距離、 t がテンプレート画像の縦の画素数、 t' が画像内ユーザ座標検出部 13 から送られた抜き出し範囲の縦の画素数である。ここで、顔の大きさはユーザにより異なるため、ユーザのテンプレート画像撮影時の距離を利用してカメラ基準ユーザ距離を算出することで、画像内ユーザ座標検出部 13 で特定したユーザの顔

の大きさを反映したカメラ基準ユーザ距離が算出され、距離算出の誤差を低減できる。カメラ基準ユーザ距離算出部 15 は、ユーザ特定情報が付与されたカメラ基準ユーザ距離を、マイクアレイ基準ユーザ角度算出部 16 へ伝達する。

[0086] マイクアレイ基準ユーザ角度算出部 16 は、所定距離 L 、カメラ基準ユーザ角度算出部 14 から伝達されたカメラ基準ユーザ角度 θ 、およびカメラ基準ユーザ距離算出部 15 から伝達されたカメラ基準ユーザ距離 D に基づいて、マイクアレイ基準点 22 を基準としたユーザ角度であるマイクアレイ基準ユーザ角度の算出を行う。マイクアレイ基準ユーザ角度の算出は、所定距離 L に加え、同じユーザ特定情報が付与されたカメラ基準ユーザ角度 θ およびカメラ基準ユーザ距離 D から、実施形態 1 と同様の方法で算出することができる。

[0087] マイクアレイ基準ユーザ角度算出部 16 は、カメラ基準ユーザ角度算出部 14 から伝達されたユーザ特定情報、および、カメラ基準ユーザ距離算出部 15 から伝達されたユーザ特定情報を、マイクアレイ基準ユーザ角度 α に付与する。マイクアレイ基準ユーザ角度算出部 16 は、ユーザ特定情報とユーザ特定情報とが付与されたマイクアレイ基準ユーザ角度 α を、指向角度制御部 17 へ伝達する。

[0088] 指向角度制御部 17 は、マイクアレイ基準ユーザ角度算出部 16 から伝達されたマイクアレイ基準ユーザ角度 α に向けてマイクアレイ 11 の指向角度制御を行い、マイクアレイ 11 から伝達された音声を出力する。これにより、ユーザの実際の顔の大きさを考慮した距離算出ができ、指向角度制御部 17 での指向角制御が精度良く行えるようになる。

[0089] ここで、本実施形態では、ユーザ顔をあらかじめ撮影するときに、顔撮影データと距離とを関連付けて参照データとしたが、画像内ユーザ座標検出部 13 において検出された顔属性をユーザ特定情報とし、カメラ基準ユーザ距離を補正するようにしてもよい。顔属性とは、性別、年齢などで、女性より男性、子供より大人の方が大きい、などで設定することができる。各属性の

判定には従来の方法を使用することができ、参照データの類似度などから推定することが可能である。このように、顔属性を考慮したユーザ距離算出により、指向角度制御部 17 での指向角制御が精度良く行えるようになる。

[0090] (実施形態 4)

以下、図 12 を併せて参照し、本発明の実施形態 4 について具体的に説明する。

図 12 は、本発明の実施形態 4 に係る音声認識装置の一構成例を示すブロック図である。図 12 に示すように、本実施形態における音声認識装置は、実施形態 3 で説明した音声入力装置 1 と音声認識部 121 とを備えた音声認識装置である。ただし、本実施形態は、実施形態 1, 2 のいずれにも適用できる。

[0091] 音声認識部 121 は、音声入力装置 1 から伝達される音声入力結果を解析し、解析結果を出力する。解析方法は従来の方法を使用することができ、例えば、参照データと比較することにより音声を解析して認識する。出力された結果は、テキストデータに変換したり、電気機器の ON、OFF などの制御に使用したりする。本実施形態の音声認識装置は、適切に所望の音声が高ノイズで取得できる音声入力装置 1 を備えるため、高精度で音声を認識することが可能である。

[0092] さらに、本実施形態における指向角度制御部 17 は、指向角度制御に利用したマイクアレイ基準ユーザ角度 α に付与されたユーザ特定情報を、マイクアレイ 11 から伝達された音声に付与して出力することができる。これにより、音声認識において、ユーザに合わせて音声認識精度の高い音声認識方法を利用することが可能になり、音声認識の誤認識を低減することができるため好適である。例えば、ユーザ特定情報と音声入力結果から音声認識の参照データを更新していき、ユーザごとの参照データにより音声認識を行う。これによりユーザ固有の音声の特徴を考慮した音声認識を実現することができる。特徴量としては、音声のパラメータとして一般的なものを対象とすることができ、イントネーション、周波数、などがある。したがって、ユーザ特

定情報を音声入力に付与することで、音声認識の参照データをユーザに合わせて変更することができ、誤認識を低減することができる。

[0093] (実施形態5)

以下、図13を併せて参照し、本発明の実施形態5について具体的に説明する。

図13は、本発明の実施形態5に係る画像表示装置の一構成例を示すブロック図である。図13に示すように、本実施形態における画像表示装置は、実施形態4で説明した音声認識装置、制御部131、および画像表示部132を備えた画像表示装置である。本実施形態に係る画像表示装置が音声認識部121を備えることを前提に説明するが、音声認識部121を別装置で構成することも可能であり、本実施形態に係る画像表示装置は、実施形態1～3で説明したような音声入力装置1を備えていればよい。

[0094] 制御部131は、音声認識部121から伝達されるユーザごとの音声認識結果に基づき、画像表示部132の表示内容や、電源のON/OFFや、音量などの制御を行う。これにより、本実施形態の画像表示装置は、音声による操作を行うことが可能となる。音声入力装置1は、画像表示装置のリモートコントローラに搭載しておくことが好ましい。

[0095] 操作の内容の例としては、テレビ放送のチャンネル変更や、番組表の呼び出し、番組の指定などが挙げられる。本実施形態の画像表示装置は、適切に所望の音声が高ノイズで取得できる音声入力装置1を備え、高精度で音声を認識することが可能であるため、音声を利用したより正確な操作を行うことができる。

[0096] カメラ12による撮影のタイミングについては、このような操作がなされた時に行えばよい。また、実施形態1～5のいずれにおいても、音声入力装置1でのカメラ12による撮影のタイミングは、音声入力を行っている最中か音声入力直前または直後であればよく、所定の操作（例えば音声入力ボタンの押下、もしくは音声入力操作とは別のシャッターボタンの押下）がなされたときとすればよい。

[0097] また、本実施形態に限らず、音声入力装置 1 に画像表示部を備えていれば、カメラ 1 2 による撮影画像表示中にそのユーザの顔をタッチするなどして、ユーザ位置を録音者が指定し、その指定された位置に基づきカメラ基準ユーザ角度やカメラ基準ユーザ距離を算出することもできる。

[0098] このように、本実施形態によれば、マイクアレイ 1 1 の指向範囲を適切に制御できる音声入力装置を備えた画像表示装置が提供できる。さらに、本実施形態に係る画像表示装置は、カメラ 1 2 や画像内ユーザ座標検出部 1 3 等のユーザ位置検出機能を備えているため、複数のユーザの画像内ユーザ座標を検出することが可能であるうえに、カメラ 1 2 の画像を画像表示部 1 3 2 に表示させることもできる。加えて、本実施形態では、音声入力装置 1 により、複数のユーザの音声を別々に低ノイズで入力することができる。そのため、本実施形態に係る画像表示装置を用いれば、例えば、複数のユーザの画像をそれぞれ切り出して表示し、各ユーザの画像および音声をユーザごとに別々に配信することが可能なテレビ会議システムなどを構築することができる。また、本実施形態における画像表示装置としては、携帯電話機（スマートフォンと呼ばれるものも含む）、タブレットなども挙げられる。

[0099] （他の実施形態）

例えば図 1、図 1 2、図 1 3 で例示したマイクアレイ 1 1、カメラ 1 2、画像表示部 1 3 2 を除く部位など、本発明に係る装置の各構成要素は、例えばマイクロプロセッサ（または DSP : Digital Signal Processor）、メモリ、バス、インターフェイス、リモコン等の周辺装置などのハードウェアと、これらのハードウェア上にて実行可能なソフトウェアとにより実現できる。上記ハードウェアの一部は集積回路／IC（Integrated Circuit）チップセットとして搭載することができ、その場合、上記ソフトウェアは上記メモリに記憶しておければよい。また、本発明の各構成要素の全てをハードウェアで構成してもよく、その場合についても同様に、そのハードウェアの一部を集積回路／ICチップセットとして搭載することも可能である。

[0100] また、上述した様々な構成例における機能を実現するためのソフトウェア

のプログラムコードを記録した記録媒体を、音声入力装置またはそれを備えた装置（汎用コンピュータであってもよい）であって、マイクアレイ１１およびカメラ１２が内蔵または接続された装置に供給し、その装置内のマイクロプロセッサまたはDSPによりプログラムコードが実行されることによっても、本発明の目的が達成される。この場合、ソフトウェアのプログラムコード自体が上述した様々な構成例の機能を実現することになり、このプログラムコード自体や、プログラムコードを記録した記録媒体（外部記録媒体や内部記憶装置）であっても、そのコードを制御側が読み出して実行することで、本発明を構成することができる。外部記録媒体としては、例えばCD-ROMまたはDVD-ROMなどの光ディスクやメモリカード等の不揮発性の半導体メモリなど、様々なものが挙げられる。内部記憶装置としては、ハードディスクや半導体メモリなど様々なものが挙げられる。また、プログラムコードはインターネットからダウンロードして実行することや、放送波から受信して実行することもできる。

[0101] 以上、本発明に係る音声入力装置、およびその音声入力装置を備えた音声認識装置や画像表示装置について説明したが、その処理の手順を説明したように、本発明は、上述したような配置のマイクアレイとカメラを備えた音声入力装置における音声入力方向変更方法としての形態も採り得る。この音声入力方向変更方法は、カメラ基準ユーザ角度算出部が、カメラで入力した画像に基づいて、カメラを基準とした音声取得対象ユーザのユーザ位置の角度であるカメラ基準ユーザ角度を算出するステップと、カメラ基準ユーザ距離算出部が、上記画像に基づいて、カメラを基準とした上記ユーザ位置までの距離であるカメラ基準ユーザ距離を算出するステップと、マイクアレイ基準ユーザ角度算出部が、上記所定距離、上記カメラ基準ユーザ角度、および上記カメラ基準ユーザ距離に基づいて、マイクアレイを基準としたユーザ位置の角度であるマイクアレイ基準ユーザ角度を算出するステップと、指向角度制御部が、マイクアレイ基準ユーザ角度になるようマイクアレイの指向角度を制御するステップと、を有する。その他の応用例については、音声入力装

置について説明した通りであり、その説明を省略する。

[0102] なお、上記プログラムコード自体は、換言すると、この音声入力方向変更方法を、コンピュータに実行させるためのプログラムである。すなわち、このプログラムは、コンピュータに、カメラで入力した画像に基づいて、カメラを基準とした音声取得対象ユーザのユーザ位置の角度であるカメラ基準ユーザ角度を算出するステップと、上記画像に基づいて、カメラを基準とした上記ユーザ位置までの距離であるカメラ基準ユーザ距離を算出するステップと、上記所定距離、上記カメラ基準ユーザ角度、および上記カメラ基準ユーザ距離に基づいて、マイクアレイを基準としたユーザ位置の角度であるマイクアレイ基準ユーザ角度を算出するステップと、マイクアレイ基準ユーザ角度になるようマイクアレイの指向角度を制御するステップと、を実行させるためのものである。その他の応用例については、音声入力装置について説明した通りであり、その説明を省略する。

[0103] 以上のように、本発明に係る音声入力装置は、音声を取得するマイクロフォンを複数有するマイクアレイと、撮像素子を有し撮影により画像を入力するカメラと、該カメラで入力された該画像に基づいて前記マイクアレイの指向角度を制御する指向角度制御部と、を備えた音声入力装置であって、前記カメラと前記マイクアレイが所定距離だけ離間して配置され、前記音声入力装置は、前記画像に基づいて、前記カメラを基準とした音声取得対象ユーザのユーザ位置の角度であるカメラ基準ユーザ角度を算出するカメラ基準ユーザ角度算出部と、前記画像に基づいて、前記カメラを基準とした前記ユーザ位置までの距離であるカメラ基準ユーザ距離を算出するカメラ基準ユーザ距離算出部と、前記所定距離、前記カメラ基準ユーザ角度、および前記カメラ基準ユーザ距離に基づいて、前記マイクアレイを基準としたユーザ位置の角度であるマイクアレイ基準ユーザ角度を算出するマイクアレイ基準ユーザ角度算出部と、をさらに備え、前記指向角度制御部は、前記マイクアレイ基準ユーザ角度になるよう前記マイクアレイの指向角度を制御することを特徴とする。

- [0104] これにより、指向角度を制御できる音声入力装置において、音声取得対象となる発話者（ユーザ）と同じ方向で違う距離に発話中の人物など、録音者（音声入力装置を使用するユーザであって、基本的には上記発話者とは異なる者）の意図しない音源がある場合に、その距離を判定して目的の距離にいるユーザの音声のみを入力することでユーザ音声の誤認識を低減し、低ノイズでユーザ音声を取得することができる。
- [0105] また、前記マイクアレイ基準ユーザ角度算出部は、前記所定距離と、前記カメラ基準ユーザ距離と、前記カメラ基準ユーザ角度と、前記カメラ基準ユーザ角度に応じた角度算出誤差範囲とに基づいて、前記マイクアレイ基準ユーザ角度範囲を算出し、前記指向角度制御部は、前記マイクアレイ基準ユーザ角度範囲になるよう前記マイクアレイの指向角度を制御するようにしてもよい。カメラ基準ユーザ角度に応じてマイクアレイ基準ユーザ角度範囲を算出しているため、マイクアレイの指向範囲を適切に制御することができる。
- [0106] もしくは、前記マイクアレイ基準ユーザ角度算出部は、前記所定距離と、前記カメラ基準ユーザ角度と、前記カメラ基準ユーザ距離と、前記カメラ基準ユーザ角度および／または前記カメラ基準ユーザ距離に応じた距離算出誤差範囲とに基づいて、前記マイクアレイ基準ユーザ角度範囲を算出し、前記指向角度制御部は、前記マイクアレイ基準ユーザ角度範囲になるよう前記マイクアレイの指向角度を制御するようにしてもよい。カメラ基準ユーザ距離に応じてマイクアレイ基準ユーザ角度範囲を算出しているため、マイクアレイの指向範囲を適切に制御することができる。
- [0107] また、前記画像における被写体の顔領域を検出し、該顔領域の位置を前記ユーザ位置として、前記カメラ基準ユーザ角度および／または前記カメラ基準ユーザ距離を算出するようにしてもよい。顔領域の検出は、声を発している口が存在するため、ユーザ位置を検出するための処理として好ましい。
- [0108] また、本発明に係る画像表示装置は、上述のような音声入力装置を備えたことを特徴とする。これにより、マイクアレイの指向範囲を適切に制御できる音声入力装置を備えた画像表示装置が提供できる。

符号の説明

[0109] D…カメラ基準ユーザ距離、L…カメラとマイクアレイ間の距離（所定距離）、 θ …カメラ基準ユーザ角度、 α …指向方向、1…音声入力装置、11…マイクアレイ、12…カメラ、13…画像内ユーザ座標検出部、14…カメラ基準ユーザ角度算出部、15…カメラ基準ユーザ距離算出部、16…マイクアレイ基準ユーザ角度算出部、17…指向角度制御部、21…超指向性マイクロフォン、22…マイクアレイ基準点、23…集音方向、24…マイクアレイ正面方向、25…マイクアレイ軸、31…マイクロフォン、32…カメラ基準点、40…撮影範囲、41…カメラ基準点、42…カメラ光軸、43…カメラ撮影方向、44…直線、51…カメラの投影面、52…画像内ユーザ顔座標、61…ユーザ、62…別の発話者、91…指向範囲（指向領域）、101…角度算出誤差範囲、102…距離算出誤差範囲、103…共通範囲、104…マイクアレイ基準ユーザ角度範囲、121…音声認識部、131…制御部、132…画像表示部。

請求の範囲

[請求項1]

音声を取得するマイクロフォンを複数有するマイクアレイと、
撮像素子を有し撮影により画像を入力するカメラと、
該カメラで入力された該画像に基づいて前記マイクアレイの指向角度を制御する指向角度制御部と、を備えた音声入力装置であって、
前記カメラと前記マイクアレイが所定距離だけ離間して配置され、
前記音声入力装置は、
前記画像に基づいて、前記カメラを基準とした音声取得対象ユーザのユーザ位置の角度であるカメラ基準ユーザ角度を算出するカメラ基準ユーザ角度算出部と、
前記画像に基づいて、前記カメラを基準とした前記ユーザ位置までの距離であるカメラ基準ユーザ距離を算出するカメラ基準ユーザ距離算出部と、
前記所定距離、前記カメラ基準ユーザ角度、および前記カメラ基準ユーザ距離に基づいて、前記マイクアレイを基準としたユーザ位置の角度であるマイクアレイ基準ユーザ角度を算出するマイクアレイ基準ユーザ角度算出部と、をさらに備え、
前記指向角度制御部は、前記マイクアレイ基準ユーザ角度になるよう前記マイクアレイの指向角度を制御することを特徴とする音声入力装置。

[請求項2]

前記マイクアレイ基準ユーザ角度算出部は、前記所定距離と、前記カメラ基準ユーザ距離と、前記カメラ基準ユーザ角度と、前記カメラ基準ユーザ角度に応じた角度算出誤差範囲とに基づいて、前記マイクアレイ基準ユーザ角度範囲を算出し、
前記指向角度制御部は、前記マイクアレイ基準ユーザ角度範囲になるよう前記マイクアレイの指向角度を制御することを特徴とする請求項1に記載の音声入力装置。

[請求項3]

前記マイクアレイ基準ユーザ角度算出部は、前記所定距離と、前記

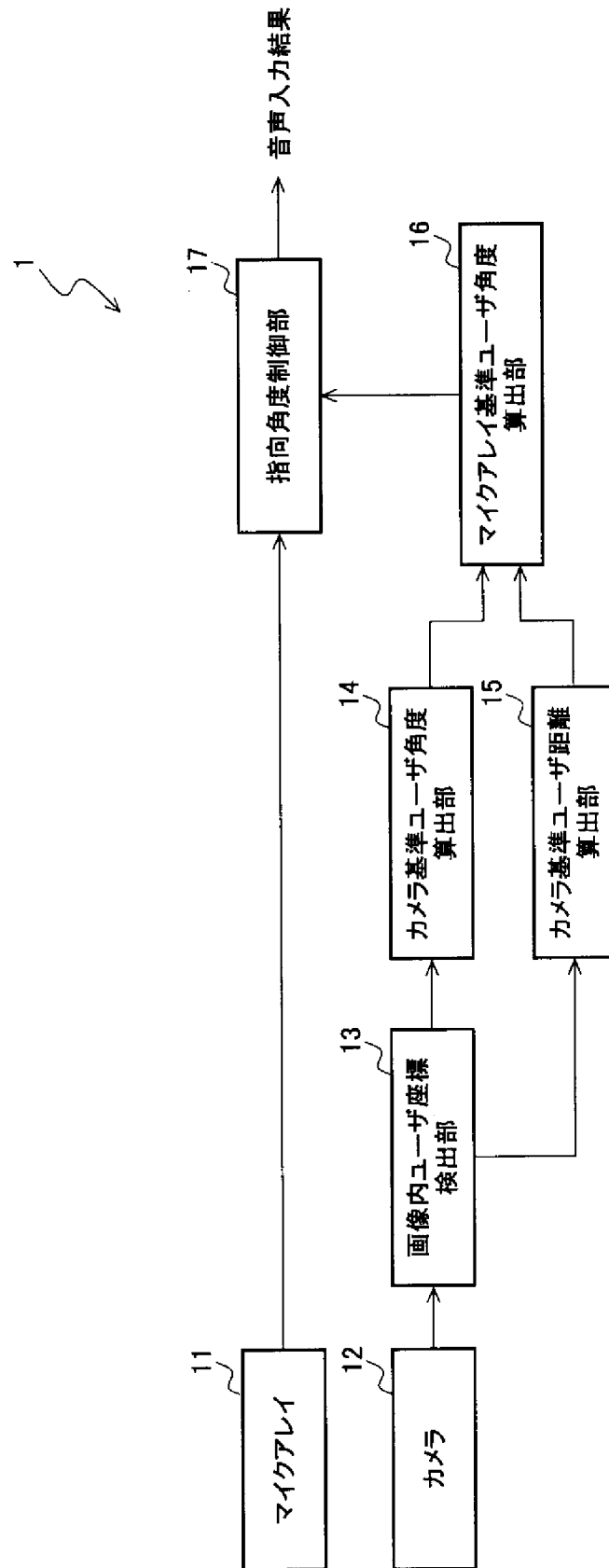
カメラ基準ユーザ角度と、前記カメラ基準ユーザ距離と、前記カメラ基準ユーザ角度および／または前記カメラ基準ユーザ距離に応じた距離算出誤差範囲とに基づいて、前記マイクアレイ基準ユーザ角度範囲を算出し、

前記指向角度制御部は、前記マイクアレイ基準ユーザ角度範囲になるよう前記マイクアレイの指向角度を制御することを特徴とする請求項 1 に記載の音声入力装置。

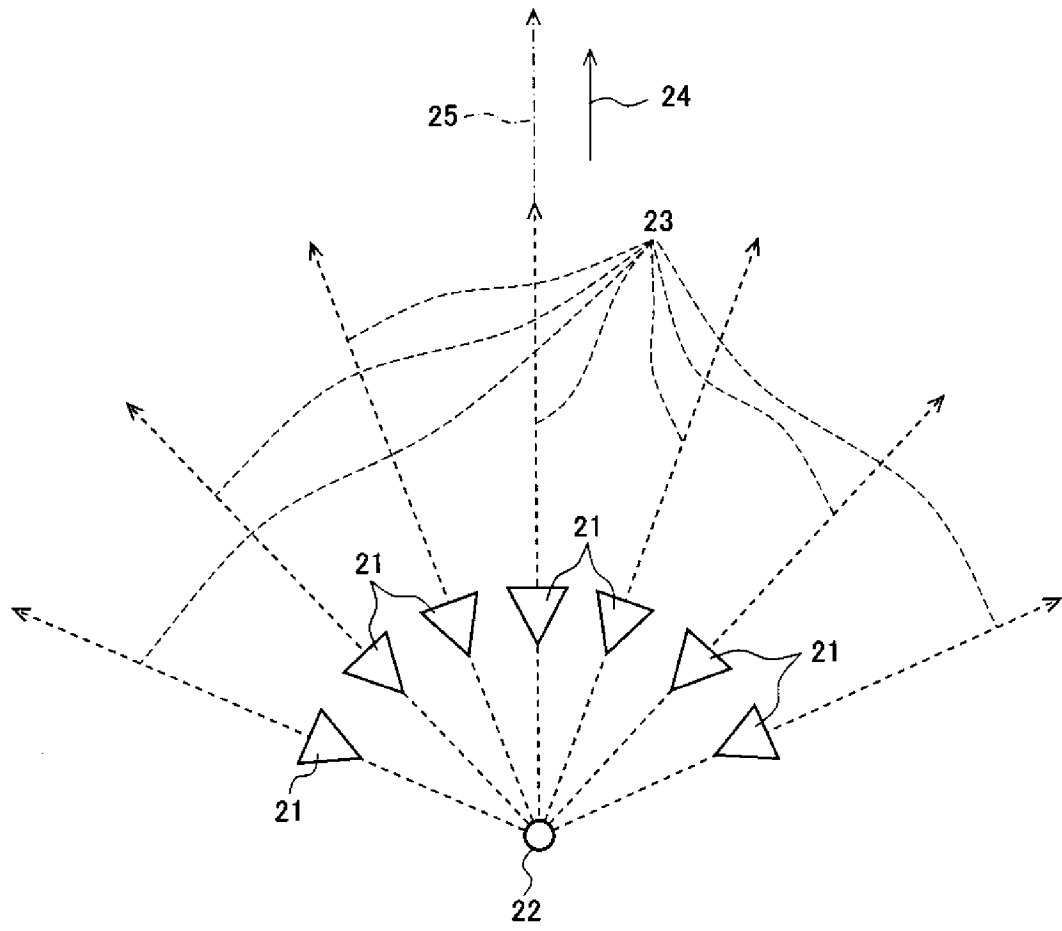
[請求項4] 前記画像における被写体の顔領域を検出し、該顔領域の位置を前記ユーザ位置として、前記カメラ基準ユーザ角度および／または前記カメラ基準ユーザ距離を算出することを特徴とする請求項 1 ～ 3 のいずれか 1 項に記載の音声入力装置。

[請求項5] 請求項 1 ～ 4 のいずれか 1 項に記載の音声入力装置を備えた画像表示装置。

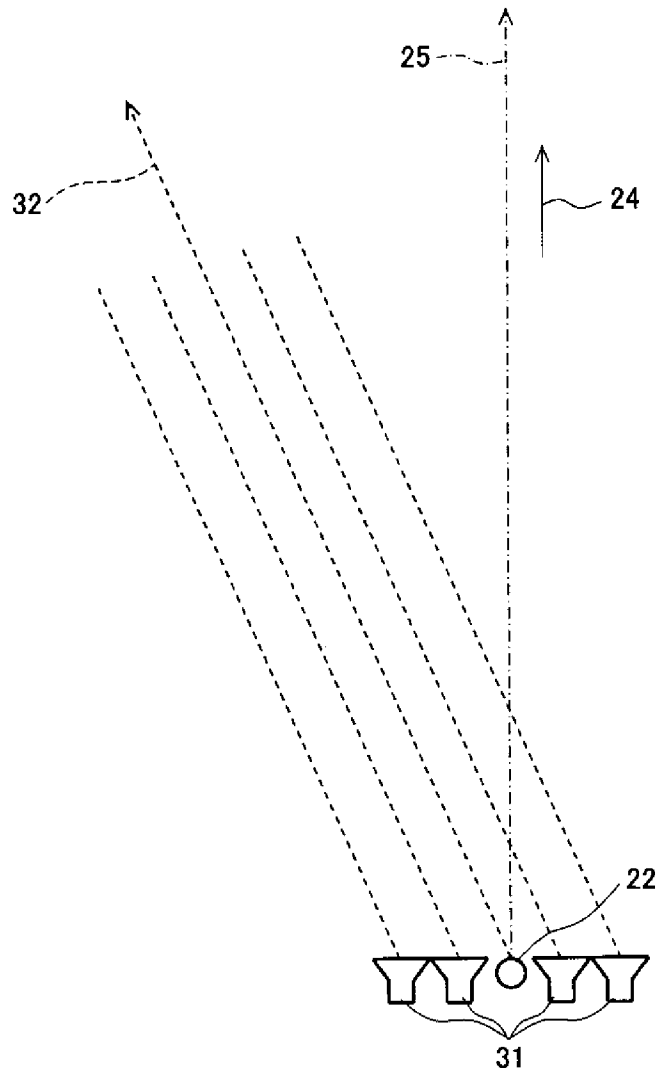
[図1]



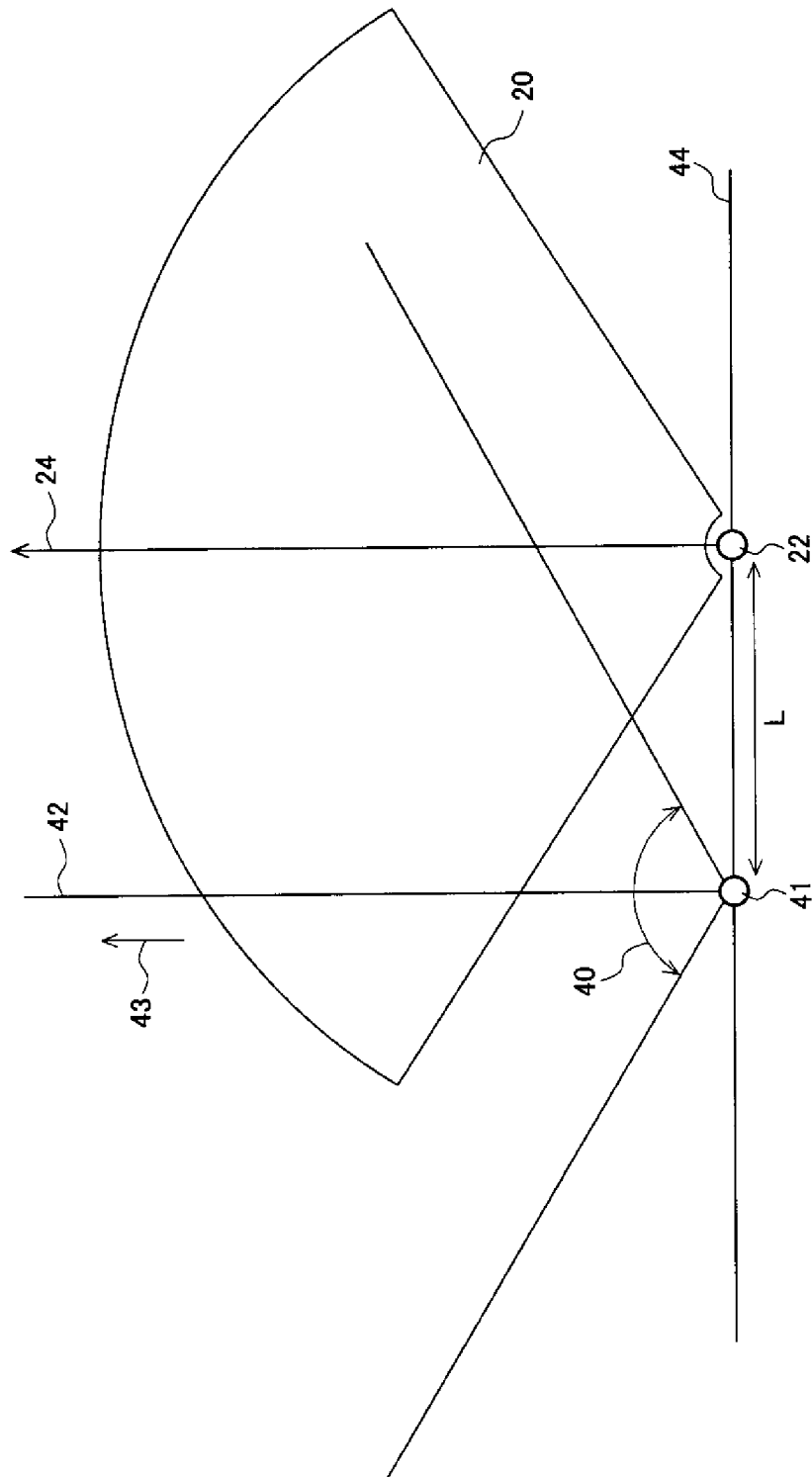
[図2]



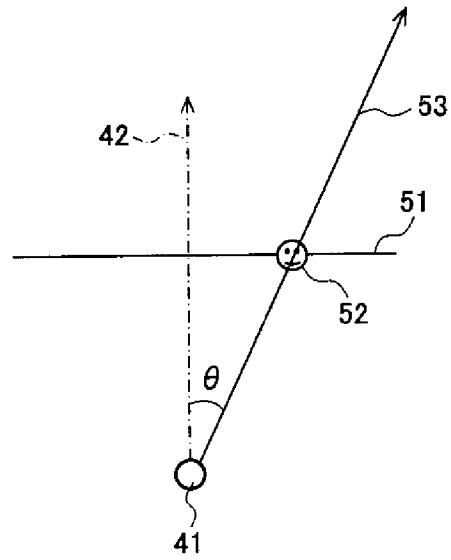
[図3]



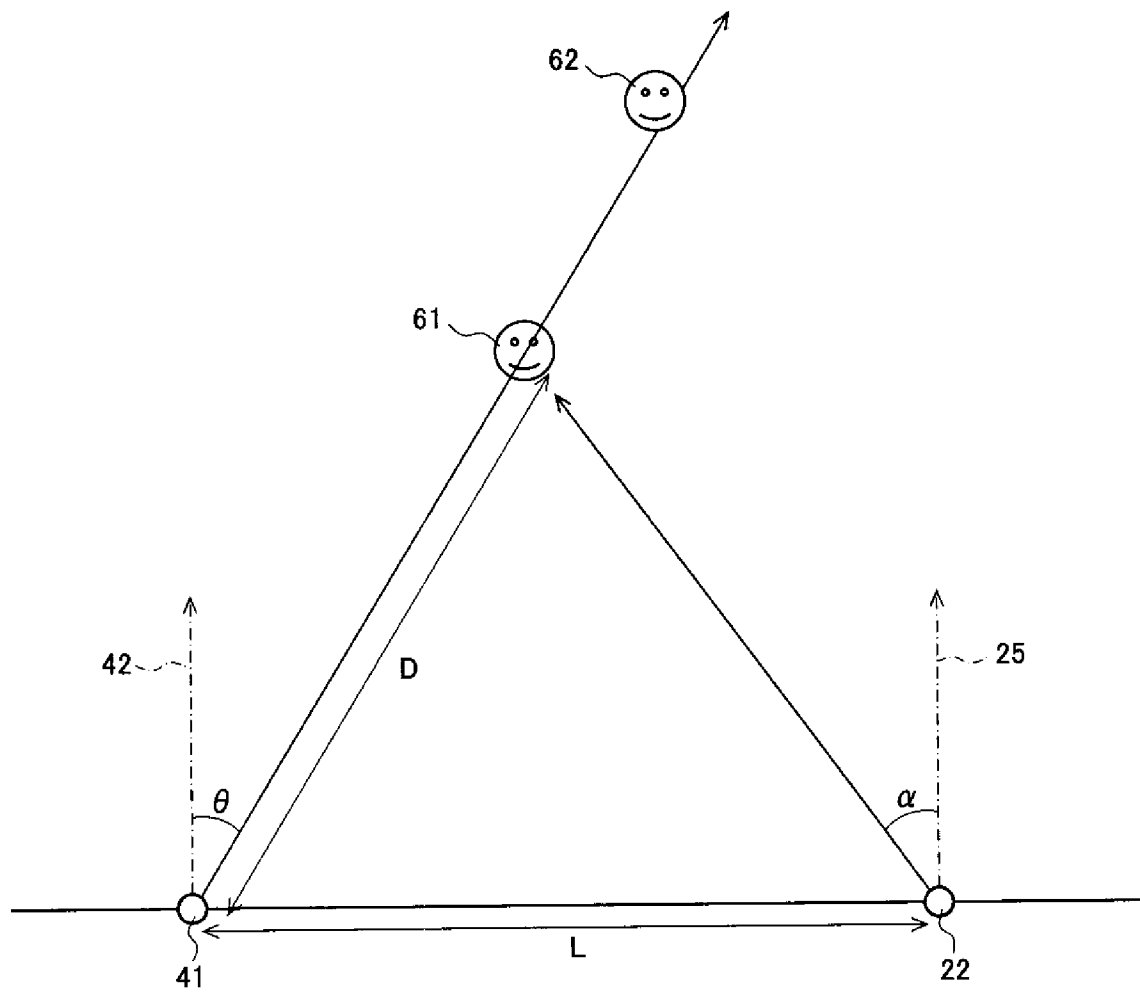
[図4]



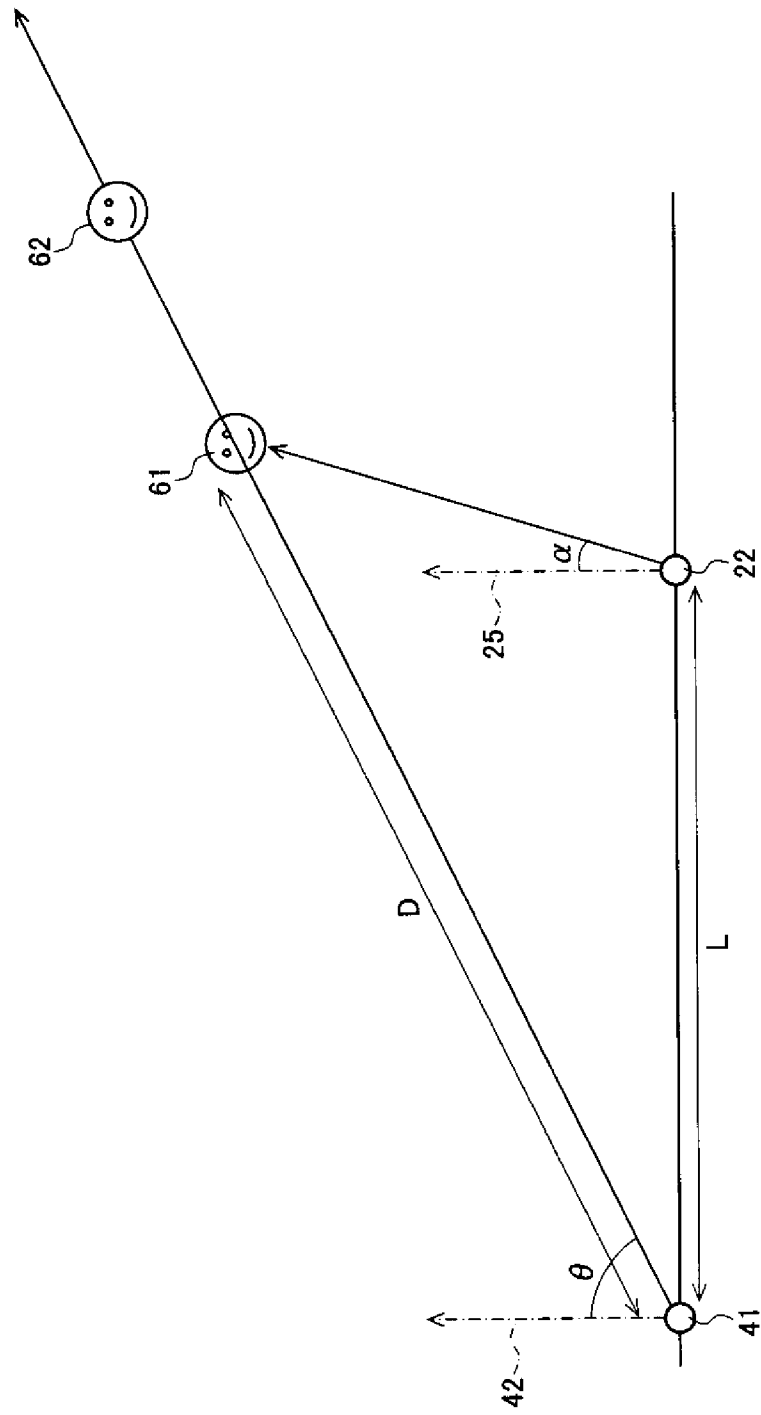
[図5]



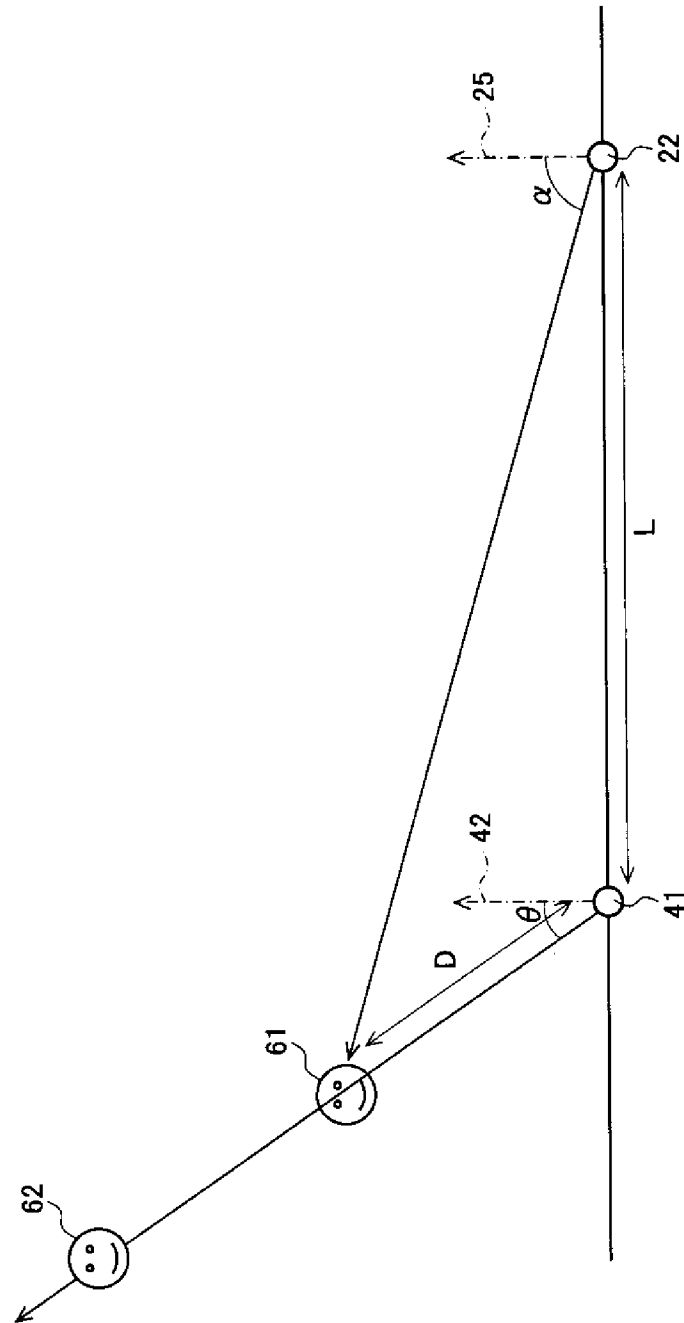
[図6]



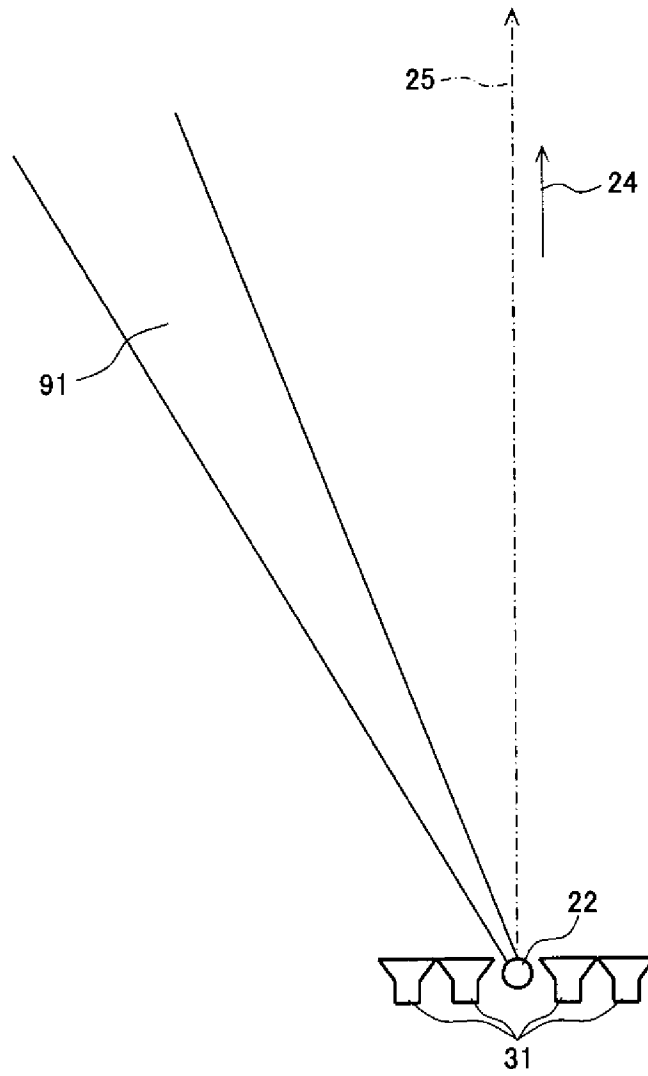
[図7]



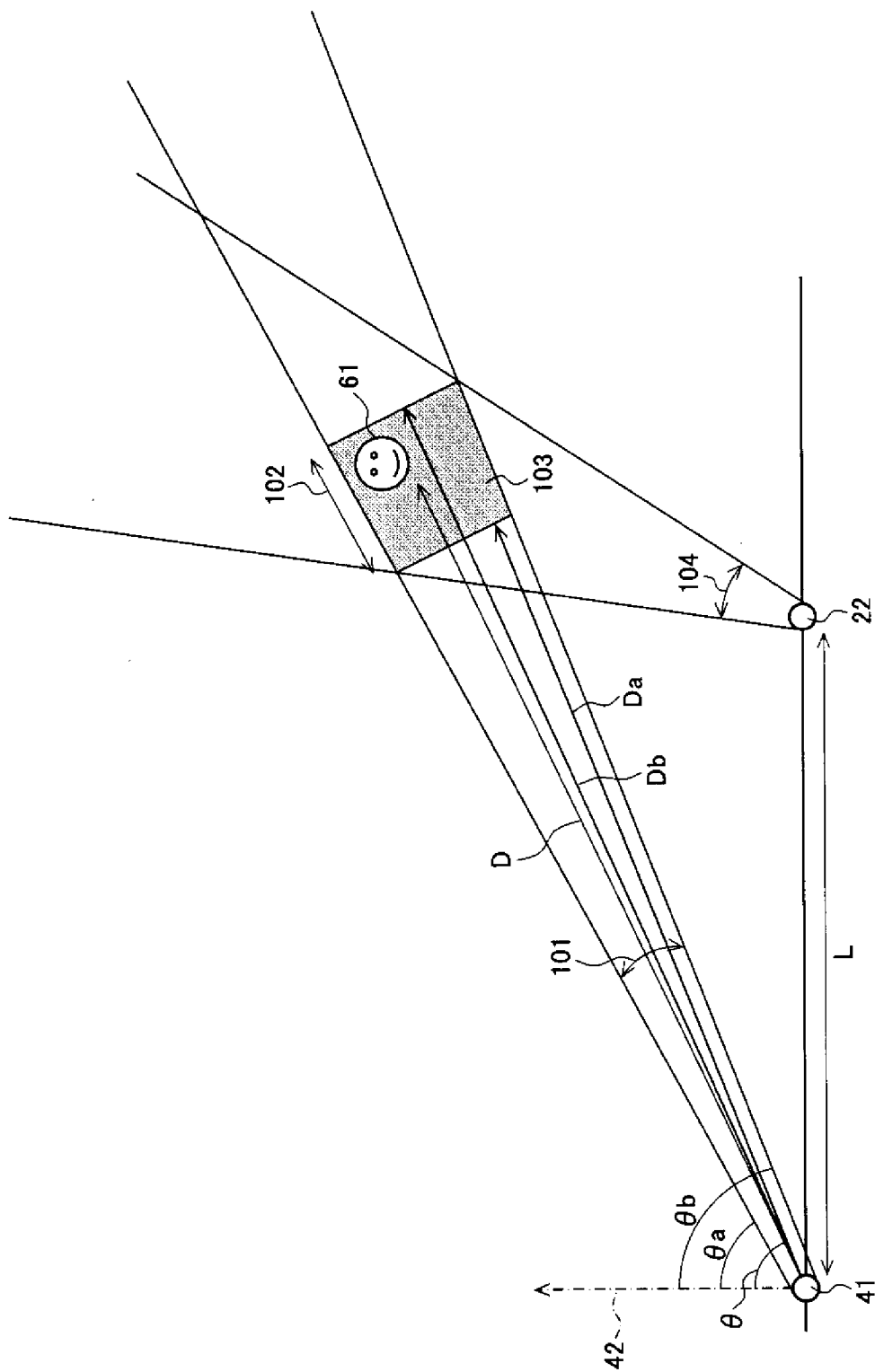
[図8]



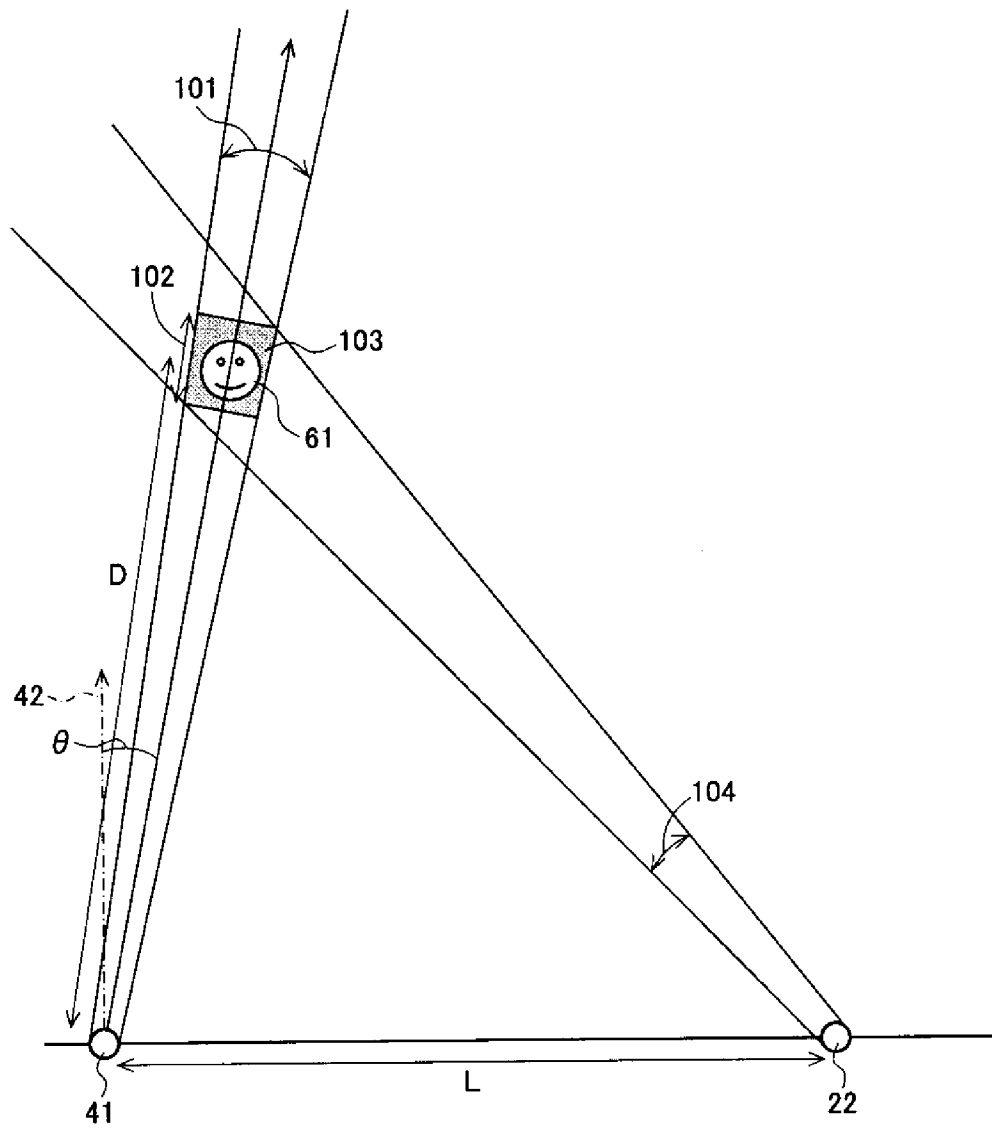
[図9]



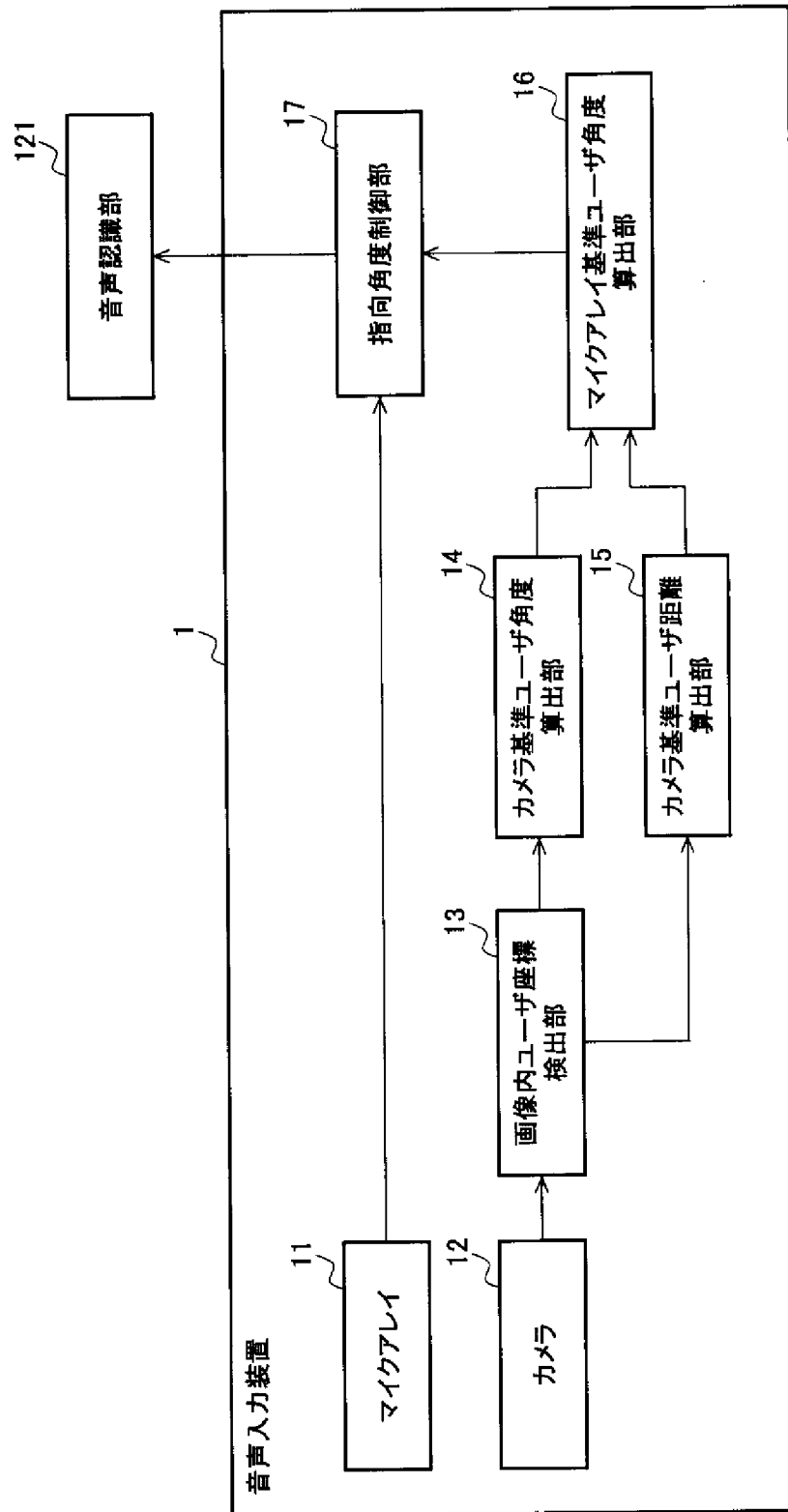
[図10]



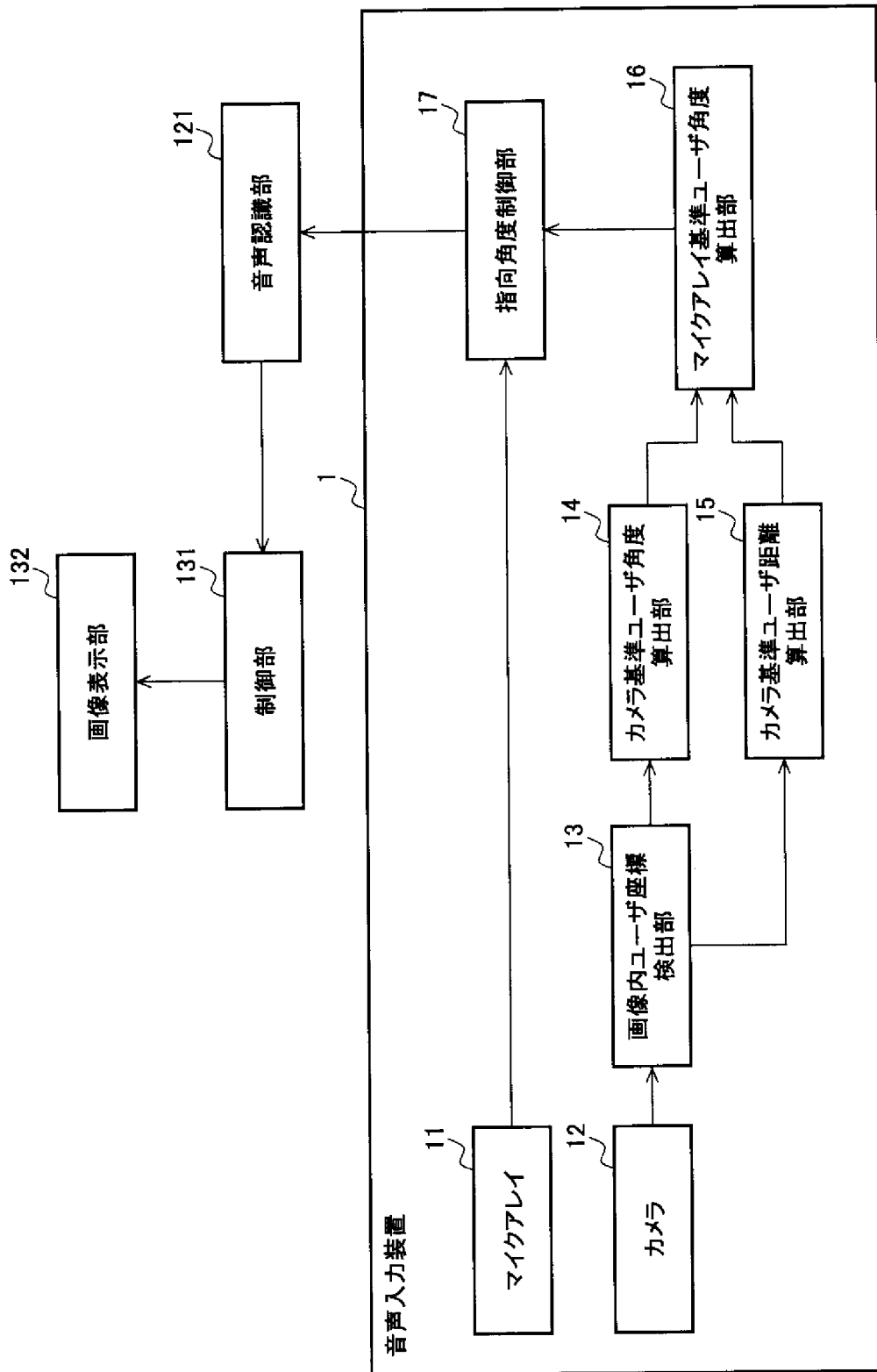
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/084034

A. CLASSIFICATION OF SUBJECT MATTER

H04R3/00(2006.01)i, G10L15/28(2013.01)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, G10L15/28, 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	1996-2014
Kokai Jitsuyo Shinan Koho	1971-2014	Toroku Jitsuyo Shinan Koho	1994-2014

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 A	JP 2012-186551 A (Hitachi, Ltd.), 27 September 2012 (27.09.2012), paragraphs [0012] to [0112]; fig. 1 to 12 (Family: none)	1, 5 3, 4 2
Y	JP 2011-071702 A (Fujitsu Ltd.), 07 April 2011 (07.04.2011), paragraphs [0054] to [0056]; fig. 6 (Family: none)	3
Y	JP 2009-225379 A (Fujitsu Ltd.), 01 October 2009 (01.10.2009), paragraphs [0037] to [0059]; fig. 7 to 8 (Family: none)	4

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier application or patent but published on or after the international filing date
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 "&" document member of the same patent family

Date of the actual completion of the international search
20 February, 2014 (20.02.14)

Date of mailing of the international search report
04 March, 2014 (04.03.14)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/084034

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 2010-154259 A (Victor Company of Japan, Ltd.), 08 July 2010 (08.07.2010), entire text; all drawings (Family: none)	1-5
A	JP 2008-193196 A (Casio Computer Co., Ltd.), 21 August 2008 (21.08.2008), entire text; all drawings (Family: none)	1-5

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, G10L15/28(2013.01)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, G10L15/28, 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 4 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 4 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 4 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 A	JP 2012-186551 A（株式会社日立製作所）2012.09.27, 段落【0012】－【0112】，第1－12図（ファミリーなし）	1, 5 3, 4 2
Y	JP 2011-071702 A（富士通株式会社）2011.04.07, 段落【0054】－【0056】，第6図（ファミリーなし）	3
Y	JP 2009-225379 A（富士通株式会社）2009.10.01, 段落【0037】－【0059】，第7－8図（ファミリーなし）	4

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 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 0 . 0 2 . 2 0 1 4

国際調査報告の発送日

0 4 . 0 3 . 2 0 1 4

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）

武田 裕司

5 Z

8 9 4 7

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
A	JP 2010-154259 A (日本ビクター株式会社) 2010.07.08, 全文, 全図 (ファミリーなし)	1 - 5
A	JP 2008-193196 A (カシオ計算機株式会社) 2008.08.21, 全文, 全図 (ファミリーなし)	1 - 5