

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6841277号
(P6841277)

(45) 発行日 令和3年3月10日(2021.3.10)

(24) 登録日 令和3年2月22日(2021.2.22)

(51) Int.Cl.		F I			
G08B 25/04	(2006.01)	G08B	25/04	E	
H04N 7/18	(2006.01)	H04N	7/18	D	
G10L 25/51	(2013.01)	G10L	25/51	400	
G08B 25/00	(2006.01)	G08B	25/00	510M	

請求項の数 6 (全 15 頁)

(21) 出願番号	特願2018-509106 (P2018-509106)	(73) 特許権者	000004237
(86) (22) 出願日	平成29年3月22日 (2017.3.22)		日本電気株式会社
(86) 国際出願番号	PCT/JP2017/011316		東京都港区芝五丁目7番1号
(87) 国際公開番号	W02017/170007	(74) 代理人	100109313
(87) 国際公開日	平成29年10月5日 (2017.10.5)		弁理士 机 昌彦
審査請求日	令和2年2月17日 (2020.2.17)	(74) 代理人	100124154
(31) 優先権主張番号	特願2016-67021 (P2016-67021)		弁理士 下坂 直樹
(32) 優先日	平成28年3月30日 (2016.3.30)	(72) 発明者	荒井 友督
(33) 優先権主張国・地域又は機関	日本国 (JP)		東京都港区芝五丁目7番1号 日本電気株式会社内
		(72) 発明者	近藤 玲史
			東京都港区芝五丁目7番1号 日本電気株式会社内

最終頁に続く

(54) 【発明の名称】 監視装置、監視方法、及びプログラム

(57) 【特許請求の範囲】

【請求項1】

音響信号を取得する音響信号取得手段と、
取得した前記音響信号から音響イベントの発生方向を算出する音響イベント検知手段と

、
画像信号を取得する画像信号取得手段と、
取得した前記画像信号から画像イベントの発生方向を算出する画像イベント検知手段と

、
算出した前記音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した前記画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出する写像関係算出手段と、

算出した前記写像関係と算出した前記音響イベントのうちの検知対象イベントの発生方向とを用いて前記画像イベントのうちの検知対象イベントの発生方向を推定する座標変換手段と

を備える監視装置。

【請求項2】

前記写像関係算出手段は、

前記写像関係の変化が閾値以上の場合に前記画像信号取得手段と前記音響信号取得手段の方向関係が変化すると判断する

請求項1に記載した監視装置。

【請求項 3】

前記写像関係算出手段は、

算出した前記音響イベントのうちのキャリブレーション対象イベントの発生方向と算出した前記画像イベントのうちのキャリブレーション対象イベントの発生方向との間の写像変換行列と、写像変換誤差の期待値と信頼度とを算出する、

請求項 1 又は 2 に記載した監視装置。

【請求項 4】

前記画像信号を表示させる表示手段をさらに備え、

前記表示手段は、

前記画像イベント検知手段で算出した前記画像イベントのうちの検知対象イベントの位置と、前記座標変換手段で推定した前記画像イベントのうちの検知対象イベントの発生方向と、前記写像変換誤差の期待値と、前記信頼度とのいずれか又は全てを表示する、

請求項 3 に記載した監視装置。

10

【請求項 5】

音響信号を取得し、

取得した前記音響信号から音響イベントの発生方向を算出し、

画像信号を取得し、

取得した前記画像信号から画像イベントの発生方向を算出し、

算出した前記音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した前記画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出し、

20

算出した前記写像関係と算出した前記音響イベントのうちの検知対象イベントの発生方向とを用いて前記画像イベントのうちの検知対象イベントの発生方向を推定する、

監視方法。

【請求項 6】

音響信号を取得する処理と、

取得した前記音響信号から音響イベントの発生方向を算出する手段と、

画像信号を取得する処理と、

取得した前記画像信号から画像イベントの発生方向を算出する処理と、

算出した前記音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した前記画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出する処理と、

30

算出した前記写像関係と算出した前記音響イベントのうちの検知対象イベントの発生方向とを用いて前記画像イベントのうちの検知対象イベントの発生方向を推定する処理と、

をコンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数種類のセンサを用いてイベントを検知する技術に関し、特に、複数種類センサ間の方向関係のキャリブレーションを行う監視装置、監視方法、及び記録媒体に関する。

40

【背景技術】

【0002】

複数種類のセンサを用いてイベントを検知する監視技術が広く知られており、実用化されている。このような監視技術の一例が、特許文献 1 に記載されている。特許文献 1 に記載された監視装置は、カメラ又はマイクのいずれか一方で不審な変化（異常イベント）を検知し、異常イベントの発生方向を推定する。さらに、特許文献 1 に記載された監視装置は、推定した発生方向とカメラとマイクの配置位置の差とに基づいて、もう一方から見たときの異常イベントの発生方向を推定する。そして、特許文献 1 に記載された監視装置は、推定したもう一方から見たときの異常イベントの発生方向へ、もう一方のカメラ又はマ

50

イクを向ける。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2006-339741号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

特許文献1に記載された監視装置は、カメラとマイクとの配置位置の差に基づいて、制御を行う。このため、特許文献1に記載された監視装置においては、カメラ及びマイクの設置時の位置及び方向について、極めて高い正確度が要求されていた。このため、特許文献1に記載された監視装置においては、カメラとマイクの設置時に行う方向の較正（キャリブレーション）が高コストとなっていた。

10

【0005】

つまり、上述した監視装置においては、センサの設置時の位置及び方向について、極めて高い正確度が要求されるという問題があった。

【0006】

本発明の目的は、上述した課題である、センサの設置時の位置及び方向について、極めて高い正確度が要求されることを解決する監視装置、監視方法及び記録媒体を提供することにある。

20

【課題を解決するための手段】

【0007】

本発明の一形態における監視装置は、音響信号を取得する音響信号取得手段と、取得した音響信号から音響イベントの発生方向を算出する音響イベント検知手段と、画像信号を取得する画像信号取得手段と、取得した画像信号から画像イベントの発生方向を算出する画像イベント検知手段と、算出した音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出する写像関係算出手段と、算出した写像関係と算出した音響イベントのうちの検知対象イベントの発生方向とを用いて画像イベントのうちの検知対象イベントの発生方向を推定する座標変換手段とを備える。

30

【0008】

本発明の一形態における監視方法は、音響信号を取得し、取得した音響信号から音響イベントの発生方向を算出し、画像信号を取得し、取得した画像信号から画像イベントの発生方向を算出し、算出した音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出し、算出した写像関係と算出した音響イベントのうちの検知対象イベントの発生方向とを用いて画像イベントのうちの検知対象イベントの発生方向を推定する。

【0009】

本発明の一形態における記録媒体は、音響信号を取得する処理と、取得した音響信号から音響イベントの発生方向を算出する処理と、画像信号を取得する処理と、取得した画像信号から画像イベントの発生方向を算出する処理と、算出した音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出する処理と、算出した写像関係と算出した音響イベントのうちの検知対象イベントの発生方向とを用いて前記画像イベントのうちの検知対象イベントの発生方向を推定する処理とをコンピュータの実行させる監視プログラムを記録する。

40

【発明の効果】

【0010】

本発明によれば、センサの設置時の位置及び方向について、極めて高い正確度が要求さ

50

れることを解決することができる。

【図面の簡単な説明】

【0011】

【図1】本発明の実施形態に係る監視装置の構成を表すブロック図である。

【図2】本発明の実施形態に係る P_m と P_c の関係を説明するための概念的なイメージ図である。

【図3】本発明の実施形態に係る P_m 、 P_c 空間を P_m の値で部分領域に分けたときの概念的なイメージ図である。

【図4】本発明の実施形態に係る音響イベント検知部で検知される方向の座標と、画像イベント検知部で検知される方向の座標とを表す概念的なイメージ図である。

【図5】実施形態の概要を示すブロック図である。

【図6】ハードウェア構成の一例を示すブロック図である。

【発明を実施するための形態】

【0012】

以下に、図面を参照しながら、本発明の実施形態について詳細に説明する。なお、以下の説明では、同じ機能を有するものには同じ符号をつけ、その説明を省略する場合がある。また、図面中の矢印の方向は、一例を示すものであり、ブロック間の信号の向きを限定するものではない。

【0013】

(構成)

図1は、本発明の第一の実施形態に係る監視装置の構成を表すブロック図である。本発明の第一の実施形態における監視装置は、音響信号取得部1と、画像信号取得部2と、音響イベント検知部3と、画像イベント検知部4と、写像関係算出部5と、座標変換部6と、表示部7とを備える。

【0014】

音響信号取得部1は、複数のマイクロフォン(マイク)を備えるマイクロフォンアレイを用いて音響信号を外部から取得し、音響イベント検知部3に出力する。あるいは、音響信号取得部1は、複数の方向に向けた複数の指向性マイクロフォンを用いて音響信号を取得し、取得した音響信号を音響イベント検知部3に出力する。音響信号は、音を表す時系列信号であり、複数の時間フレームと各時間フレームに対応する音圧の情報とを持つ。以降、音響信号を可聴域音響信号であるものとして説明する。音響信号取得部1は、他にも、超音波、超低周波、又は振動情報などの時系列信号を音響信号として扱うことができる。

【0015】

音響イベント検知部3は、音響信号取得部1で取得された音響信号から音源の方向(音源方向)を算出(推定)する。音響信号取得部1が複数のマイクロフォンを備えるマイクロフォンアレイを用いる場合は、複数のマイクロフォンが測定した音響信号間の到達時間差が、音源方向の算出に利用できる。この場合、音響イベント検知部3は、例えば、相互相関法又はビームフォーミング法を用いて音源方向を算出(推定)する。あるいは、音響信号取得部1が複数の方向に向けた複数の指向性マイクロフォンを用いる場合は、音響イベント検知部3は、各指向性マイクロフォンに音響イベントの発生時刻を検知することで、音響イベントが検知された指向性マイクロフォンを特定できる。このため、音響イベント検知部3は、その特定した指向性マイクロフォンが向いている方向を、音源方向とする。若しくは、複数の指向性マイクロフォンで音響イベントが検知された場合には、音響イベント検知部3は、それら複数の指向性マイクロフォンが向いている方向の中間を音源方向と特定できる。この場合、音響イベント検知部3は、各指向性マイクロフォンが取得した音響信号の強度を用いて中間方向の強度を内挿する。

【0016】

また、音響イベント検知部3は、音響信号取得部1から受信した音響信号から音響イベントを検知するとともに、検知した音響イベントをキャリブレーション対象イベントと検

10

20

30

40

50

知対象イベントとに分類する。

【0017】

音響イベントは、予め定義された音響上の特徴である。例えば、音響イベントは、受信した音響信号のある時間フレーム（現時間フレーム）の音圧とそれ以前の過去の時間フレームの音圧とを比較した時に、音圧の差が事前に設定した閾値を超えている状態のことである。音響イベントの発生時刻は、現時間フレームに対応する時刻である。あるいは、音響イベントは、受信した音響信号を周波数スペクトルに変換した時に、周波数の分布（スペクトルパタン）が事前に設定した形状のまま一定時間継続している状態のことである。

【0018】

キャリブレーション対象イベントは、ユーザが予め定義する、音響信号と画像信号との両方から検出できる事象である。キャリブレーション対象イベントの例としては、歩く動作、話す動作、逃げる動作、暴れる動作、手をたたく動作、又は、外部の装置等に基づく音と光のパターン等がある。

【0019】

検知対象イベントは、ユーザが予め定義する音響信号と画像信号とのどちらかから検出できる事象である。検知対象イベントの例としては、例えば、犯罪又は事故の発生時刻における暴れる動作又は逃げる動作がある。検知対象イベントのうち、音響信号で検知可能な例としては、多人数が集まった場所で暴れる人の罵声又は悲鳴がある。また、同様に、画像信号で検知可能な例としては、群衆が単一方向に逃げる動作、又は、人が道路中に寝そべる動作がある。

【0020】

一つの音響イベントは、キャリブレーション対象イベントと検知対象イベントとのどちらか一方に分類される場合だけでなく、両方に分類される場合もある。

【0021】

画像信号取得部2は、カメラを用いて画像信号を取得し、画像信号を画像イベント検知部4及び表示部7に出力する。ここで、画像信号は、画像を表す電氣的な時系列信号であり、複数の時間フレームと各時間フレームに対応する輝度の情報とを持つ。以降、画像信号を可視光画像信号であるものとして説明する。カメラの数は、複数としてもよい。カメラの数を複数とした場合は、画像信号取得部2は、複数のカメラが取得した画像を1つの画像に合成する。画像信号取得部2は、他にも、赤外線画像、熱画像（サーモグラフィー）、又はレーダー画像などの時系列信号を画像信号として扱うことができる。また、画像信号は、2次元画像としてもよいし、立体画像（深度画像）としてもよい。

【0022】

画像イベント検知部4は、画像信号取得部2で取得された画像信号から画像イベントが発生した方向（画像イベント発生方向）を算出（推定）する。

【0023】

また画像イベント検知部4は、画像信号取得部2で取得された画像信号から画像イベントを検知するとともに、検知した画像イベントをキャリブレーション対象イベントと検知対象イベントとに分類する。画像イベントは、取得された画像信号のある時間フレーム（現時間フレーム）の輝度とそれ以前の過去の時間フレームの輝度とを比較した時に、輝度の差が事前に設定した閾値を超えている状態のことである。画像イベントの発生時刻は、現時間フレームに対応する時刻である。一つの画像イベントは、キャリブレーション対象イベントと検知対象イベントとのどちらか一方に分類される場合だけでなく、両方に分類される場合もある。

【0024】

キャリブレーション対象イベントと検知対象イベントとの定義は、例えば、予め専用の装置に検知したい事象の音響信号又は画像信号を取得し、機械学習させてモデル化させたものを辞書的に用いる。この場合、音響イベント検知部3又は画像イベント検知部4は、モデルの尤度が閾値を超えた場合にそのモデルに対応するキャリブレーション対象イベント又は検知対象イベントが発生したと判断する。より具体的には、例えば、音響イベント

10

20

30

40

50

検知部 3 又は画像イベント検知部 4 は、専用の装置に、検知したい事象の音響信号又は画像信号から特徴量を抽出させる。特徴量は、メル周波数ケプストラム係数 (M F C C、M e l - F r e q u e n c y C e p s t r u m C o e f f i c i e n t s) 等である。さらに、音響イベント検知部 3 又は画像イベント検知部 4 は、抽出させた特徴量を、モデル化手法を用いて専用の装置にモデル化させる。モデル化手法は、ガウス混合モデル (G M M、G a u s s i a n M i x t u r e M o d e l) や隠れマルコフモデル (H M M、H i d d e n M a r k o v M o d e l) 等である。

【 0 0 2 5 】

写像関係算出部 5 は、音響イベント検知部 3 と画像イベント検知部 4 とで同時に検知されたキャリブレーション対象イベントについて、音響イベント検知部 3 と画像イベント検知部 4 とのそれぞれが算出した発生方向から、写像関係を算出する。この発生方向とは、音源方向及び画像イベント発生方向のことである。さらに、写像関係算出部 5 は、算出した現時刻の写像関係と直前に算出した過去の時刻の写像関係とを比較し、その差が閾値以上である場合に、カメラ又はマイクの方向関係が変わったと判断する。同時に検知とは、必ずしも同時刻に検知されたことを意味するのではない。キャリブレーション対象イベントの種別ごとに設定した一定時間内の違いは、同時とみなす。

【 0 0 2 6 】

写像関係算出部 5 は、写像関係の算出方法として、例えば、写像関係として多項式を仮定し、最小二乗法等を用いて多項式のパラメータを推定する、多項式回帰と呼ばれる手法を利用する。あるいは、写像関係算出部 5 は、ノイズ対策として、サポートベクトル回帰と呼ばれる手法、又は、R A N S A C (r a n d o m s a m p l e c o n s e n s u s) と呼ばれる手法を利用してもよい。

【 0 0 2 7 】

ここでは、R A N S A C を利用する写像関係の算出方法を説明する。音響イベント検知部 3 で検知されたキャリブレーション対象イベントの発生方向 (第 1 の発生方向) を P_m とする。音響イベント検知部 3 で検知されたのと同時に画像イベント検知部 4 で検知されたキャリブレーション対象イベントの発生方向 (第 2 の発生方向) を P_c とする。写像関係算出部 5 は、第 1 の発生方向 P_m から第 2 の発生方向 P_c への写像変換行列 H を算出する。この写像変換行列 H の算出にあたって、写像関係算出部 5 は、数式 1 に示す変換誤差 E が予め設定した閾値以下 (条件 1) となるように算出する。 P_m と P_c の組が複数組存在する場合、写像関係算出部 5 は、条件 1 を満たす組 (以降、I n l i e r と呼ぶ) の数が最大となるような写像変換行列 H を算出する。

【 0 0 2 8 】

【 数 1 】

$$E = \|HP_m - P_c\| \dots (1)$$

【 0 0 2 9 】

図 2 は、 P_m と P_c との関係を示すための概念的なイメージ図である。横軸が P_m 、縦軸が P_c である。黒点は、音響イベント検知部 3 と画像イベント検知部 4 とで同時に検知されたキャリブレーション対象イベントの発生方向 (P_m と P_c の組) を示す。写像変換行列 H に基づく写像が点線 ($P_c = HP_m$) のように表されるとき、点線から縦軸方向に予め設定した閾値以内の範囲が (条件 1) を満たす領域である。つまり、写像関係算出部 5 は、この領域に含まれる黒点 (すなわち I n l i e r) の数が最大となるように写像変換行列 H を算出する。さらに、写像関係算出部 5 は、 P_m - P_c 空間を P_m の値で部分領域に分けたときの各部分領域 a_i^m (i は 1 以上 N 以下の自然数) について、変換誤差の期待値 μ_i を算出する。 N は部分領域の総数である。また、写像関係算出部 5 は、各部分領域 a_i^m について信頼度 R_i^m を算出する。

【 0 0 3 0 】

図 3 は、 P_m - P_c 空間を P_m の値で部分領域に分けたときの概念的なイメージ図である。点線で囲った領域が、それぞれ P_m の i 番目の部分領域 a_i^m を示す。

10

20

30

40

50

【0031】

図4は、音響イベント検知部3で検知される方向の座標と、画像イベント検知部4で検知される方向の座標を表す概念的なイメージ図である。黒点は、部分領域 a_i^m 内の音響イベント検知部3と画像イベント検知部4で同時に検知されたキャリブレーション対象イベントの発生方向(P_m と P_c の組)である。図4左におけるグレーの領域は、音響イベント検知部3で検知される部分領域 a_i^m を示している。図4右におけるグレーの領域は、音響イベント検知部3で検知される部分領域 a_i^m を画像イベント検知部4で検知される方向に置き換えた場合の部分領域(対応する部分領域) $a^{\wedge}_i^m$ を示している。

【0032】

なお、ハット記号「 \wedge 」は、各変数の上に記載される記号であるが、本明細書では変数の右側(例えば、「 a^{\wedge} 」)に記載する。

10

【0033】

対応する部分領域 $a^{\wedge}_i^m$ における変換誤差の期待値 σ_i は、例えば、数式2に示すように、対応する部分領域 $a^{\wedge}_i^m$ 内に存在するInlierの、変換誤差 E の標準偏差の定数倍とする。

【0034】

【数2】

$$\sigma_i = c \sqrt{\frac{1}{n_i^m} \sum_{e=1}^{n_i^m} E_e^2} \dots (2)$$

20

【0035】

n_i^m は部分領域 $a^{\wedge}_i^m$ 内に存在するInlierの数、 E_e は e 番目のInlierの変換誤差、 c は定数である。また、部分領域 $a^{\wedge}_i^m$ における信頼度 R_i^m は、例えば、部分領域 $a^{\wedge}_i^m$ 内に存在するキャリブレーション対象イベントの数に対する、対応する部分領域 $a^{\wedge}_i^m$ 内に存在するInlierの数の割合とする。あるいは、部分領域 $a^{\wedge}_i^m$ における信頼度 R_i^m は、対応する部分領域 $a^{\wedge}_i^m$ 内に存在するInlierの変換誤差 E の総和の逆数としてもよい。

【0036】

写像関係算出部5は、信頼度 R_i^m が閾値以下となる部分領域 a_i^m の数が部分領域の全数 N に対して一定割合(例えば80%以上)を占める場合には、カメラとマイクの間関係が適切でないと判断する。

30

【0037】

座標変換部6は、音響イベント検知部3で検知された検知対象イベントの発生方向 P_m^A (第3の発生方向)から画像イベント検知部4で検知されるであろう検知対象イベントの発生方向 $P^{\wedge}_m^A$ を推定する。言い換えると、座標変換部6は、 P_m^A に対応する画像イベント検知部4で検知される検知対象イベントの発生方向 $P^{\wedge}_m^A$ を算出する。第3の発生方向 $P^{\wedge}_m^A$ は、写像関係算出部5で算出された写像変換行列 H を用いて数式3から算出される。

【0038】

【数3】

$$\hat{P}_m^A = HP_m^A \dots (3)$$

40

【0039】

また、座標変換部6は、 P_m^A を含む部分領域 a_j^m を特定する。さらに、座標変換部6は、特定した部分領域 a_j^m における変換誤差の期待値 σ_j を算出する。

【0040】

表示部7は、画像信号取得部2から出力された画像信号を画像としてディスプレイに表示させる。同時に、表示部7は、音響信号取得部1から出力された音響信号を音声として再生する。また、表示部7は、画像イベント検知部4で算出した検知対象イベントの発生方向 P_m^A 又は座標変換部6で算出した検知対象イベントの発生方向($P^{\wedge}_m^A$)をそれ

50

ぞれ、ディスプレイに逐次表示させる。座標変換部 6 で算出した検知対象イベントの発生方向 P^m を表示する際、表示部 7 は、座標変換部 6 で算出した変換誤差の期待値 j を同時に表示させてもよい。変換誤差の期待値 j の表示の仕方としては例えば、 P^m を中心とし、変換誤差の期待値 j を半径とした円で表示する。さらに、表示部 7 は、円及びその周辺を拡大表示させてもよい。

【0041】

また、写像関係算出部 5 がカメラ若しくはマイクの方向関係が変わったと判断した場合、又は、写像関係算出部 5 がカメラとマイクとの方向関係が適切でないと判断した場合には、表示部 7 は、次のような内容をディスプレイ等に表示させる。すなわち、表示部 7 は、カメラとマイクの方向関係が適切でないことを示す警告、又は、確認若しくは注意を促すメッセージをディスプレイ等に表示させる。さらに、表示部 7 は、写像関係算出部 5 で算出した信頼度 R_i^m を、部分領域 a_i^m ごとにカラーマップ等でディスプレイ上に表示させる。

10

【0042】

以上のように、本発明の実施形態に係る監視装置は、マイクが外部から取得した音響信号とカメラが外部から取得した画像信号との両方からそれぞれキャリブレーション対象イベントの発生方向の算出を行う。さらに、本発明の実施形態に係る監視装置は、算出したキャリブレーション対象イベントの発生方向に基づいて写像関係を算出し、算出した写像関係とマイクが外部から取得した音響信号とに基づいて算出した検知対象イベントの発生方向の座標変換を行う。

20

【0043】

このため、本発明の実施形態に係る監視装置は、カメラとマイクの設置時の位置と方向について、高い正確度を要しない。このため、本発明の実施形態に係る監視装置においては、カメラとマイクの設置時に行う方向の較正（キャリブレーション）は、低コストである。

【0044】

また、本発明の実施形態に係る監視装置は、キャリブレーション対象イベントが検知される毎に写像関係を逐次算出し、算出した現時刻の写像関係と直前に算出した過去の時刻の写像関係を比較する。そして、マイクとカメラの方向関係が変わったと判断した場合には、本発明の実施形態に係る監視装置は、警告や確認、注意を促すメッセージをディスプレイ等に表示させる。このため、ユーザはカメラ又はマイクの方向関係が変わったことに気付くことができる。

30

【0045】

また、本発明の実施形態に係る監視装置は、音響信号取得部 1 が実際に外部から取得した音響信号を用いている。これにより、本発明の実施形態に係る監視装置は、音響信号取得部 1 が外部から取得する音響信号にマイクの設置場所毎に異なる反射音が含まれることにより生じる音源方向の算出における誤差を低減することができる。さらに、本発明の実施形態に係る監視装置は、気温や気圧等の要因で音速が変化することによる音源方向の算出における誤差を低減することができる。

【0046】

40

以上、音響イベント検知部 3 が検知するキャリブレーション対象イベント及び画像イベント検知部 4 が検知するキャリブレーション対象イベントが、位置的にも時間的にも断絶して発生する例について説明した。本実施形態は、このほかに、キャリブレーション対象イベントが、位置を変えながら一定時間継続する場合にも対応可能である。さらに、本実施形態は、キャリブレーション対象イベントだけでなく、検知対象イベントについても同様に対応可能である。このような場合としては例えば、自動車、飛行体、又は船舶の駆動音と移動する様子、自転車のブレーキ音と移動する様子、台車の振動と移動する様子、等がある。

【0047】

キャリブレーション対象イベントの始端時刻を t_s 、終端時刻を t_e とする。さらに、

50

時刻 t の時に音響イベント検知部 3 が検知するキャリブレーション対象イベントの発生方向を $P_m(t)$ 、時刻 t の時に画像イベント検知部 4 が検知するキャリブレーション対象イベントの発生方向を $P_c(t)$ とする。

【0048】

音響イベント検知部 3 又は画像イベント検知部 4 は、キャリブレーション対象イベントの始端時刻 t_s 及び終端時刻 t_e を求める。

【0049】

始端時刻 t_s は、音響イベント検知部 3 又は画像イベント検知部 4 が検知したキャリブレーション対象イベント（音響イベント又は画像イベント）の発生時刻と同様である。すなわち、始端時刻 t_s は、次の場合の現時間フレームに対応する時刻である。この場合とは、音響イベント検知部 3 が、受信した音響信号のある時間フレーム（現時間フレーム）の音圧とそれ以前の過去の時間フレームの音圧とを比較した時に、音圧の差が事前に設定した閾値を超えて大きくなっている場合である。あるいは始端時刻 t_s は、次の場合の現時間フレームに対応する時刻である。この場合とは、画像イベント検知部 4 が、受信した画像信号のある時間フレーム（現時間フレーム）の輝度とそれ以前の過去の時間フレームの輝度とを比較した時に、輝度の差が事前に設定した閾値を超えて大きくなっている場合である。

【0050】

終端時刻 t_e は、次の場合の現時間フレームに対応する時刻である。この場合とは、音響イベント検知部 3 が、受信した音響信号のある時間フレーム（現時間フレーム）の音圧とそれ以前の過去の時間フレームの音圧とを比較した時に、音圧の差が事前に設定した閾値より小さくなっている場合である。あるいは、終端時刻 t_e は、次の場合の現時間フレームに対応する時刻である。この場合とは、画像イベント検知部 4 が、受信した画像信号のある時間フレーム（現時間フレーム）の輝度とそれ以前の過去の時間フレームの輝度とを比較した時に、輝度の差が事前に設定した閾値より小さくなっている場合である。

【0051】

音響イベント検知部 3 又は画像イベント検知部 4 は、求めた始端時刻 t_s 及び終端時刻 t_e 間の時間を任意の数に分割する。この分割にあたって、音響イベント検知部 3 又は画像イベント検知部 4 は、例えば、1 秒毎などのように等時間毎に分割を行ってもよい、等時間でなくても分割を行ってもよい。

【0052】

音響イベント検知部 3 又は画像イベント検知部 4 は、分割した時間それぞれの開始時刻について、キャリブレーション対象イベントの発生方向を求める。

【0053】

以上により、複数の時刻 $t = \{t_s, t_1, t_2, \dots, t_e\}$ について P_m と P_c の組 $\{P_m(t_s), P_c(t_s)\}, \{P_m(t_1), P_c(t_1)\}, \{P_m(t_2), P_c(t_2)\}, \dots, \{P_m(t_e), P_c(t_e)\}$ が求まる。この $P_m(t)$ と $P_c(t)$ の組を用いて、写像関係算出部 5 は、写像関係を算出する。

【0054】

また、写像関係算出部 5 は、時刻のずれを Δt として、 $\{P_m(t), P_c(t + \Delta t)\}$ の相関が最大になるような Δt の値を求め、 $P_c(t)$ の代わりに $P_c(t + \Delta t)$ を用いて写像関係を算出することとしてもよい。

【0055】

本実施形態によれば、センサの設置時の位置及び方向について、極めて高い正確度が要求される問題を解決することができる。

【0056】

つまり、本実施形態に係る監視装置は、センサの設置時の位置及び方向について、極めて高い正確度を実現できるとの効果を奏する。

【0057】

その理由は、次のとおりである。音響信号取得部 1 は、音響信号を取得する。音響イベ

10

20

30

40

50

ント検知部 3 は、取得した音響信号から音響イベントの発生方向を算出する。画像信号取得部 2 は、画像信号を取得する。画像イベント検知部 4 は、取得した画像信号から画像イベントの発生方向を算出する。写像関係算出部 5 は、算出した音響イベントのうちのキャリブレーション対象イベントの発生方向と、算出した画像イベントのうちのキャリブレーション対象イベントの発生方向とに基づいて、写像関係を算出する。座標変換部 6 は、算出した写像関係と算出した音響イベントのうちの検知対象イベントの発生方向とを用いて画像イベントのうちの検知対象イベントの発生方向を推定する。

【 0 0 5 8 】

このように、本実施形態の監視装置は、音響イベント及び画像イベントのうちのキャリブレーション対象イベントの発生方向を用いて写像関係を算出する。そして、監視装置は、音響イベント及び画像イベントの発生方向だけではなく、算出した写像関係と音響イベントの検知対象イベントの発生方向とを基に画像イベントのうちの検知対象イベントを推定する。そのため、監視装置は、センサの設置時の位置及び方向について、極めて高い正確度を実現できる。

10

【 0 0 5 9 】

図 5 は、上記の効果の説明に関連する構成を含む監視装置の概要の構成を示すブロック図である。すなわち、図 5 に示されている監視装置は、音響信号取得部 1 と、画像信号取得部 2 と、音響イベント検知部 3 と、画像イベント検知部 4 と、写像関係算出部 5 と、座標変換部 6 とを含む。そして、図 5 に示されている監視装置は、図 1 に示されている監視装置と同様の効果を実現できる。なお、図 5 に示されている監視装置は、本実施形態の最小構成である。ただし、各構成部は、適宜、表示部 7 に相当する構成に情報を送信する。

20

【 0 0 6 0 】

次に、ハードウェア構成について説明する。以上の説明した監視装置は、次のように構成される。例えば、監視装置の各構成部は、ハードウェア回路で構成されてもよい。また、監視装置において、各構成部は、ネットワークを介して接続した複数の装置を用いて、構成されてもよい。また、監視装置において、複数の構成部は、1つのハードウェアで構成されてもよい。また、監視装置は、CPU (Central Processing Unit) と、ROM (Read Only Memory) と、RAM (Random Access Memory) とを含むコンピュータ装置として実現されてもよい。監視装置は、上記構成に加え、さらに、入出力接続回路 (I/O C : Input / Output Circuit) と、ネットワークインターフェース回路 (N I C : Network In 30
terface Circuit) とを含むコンピュータ装置として実現されてもよい。

30

【 0 0 6 1 】

図 6 は、監視装置のハードウェア構成の一例を示すブロック図である。監視装置 6 0 0 は、CPU 6 1 0 と、ROM 6 2 0 と、RAM 6 3 0 と、内部記憶装置 6 4 0 と、I/O C 6 5 0 と、N I C 6 8 0 とを含み、コンピュータ装置を構成している。

【 0 0 6 2 】

CPU 6 1 0 は、ROM 6 2 0 からプログラムを読み込む。そして、CPU 6 1 0 は、読み込んだプログラムに基づいて、RAM 6 3 0 と、内部記憶装置 6 4 0 と、I/O C 6 5 0 と、N I C 6 8 0 とを制御する。そして、CPU 6 1 0 を含むコンピュータは、これらの構成を制御し、図 1 に示す、音響信号取得部 1 と、画像信号取得部 2 と、音響イベント 40
検知部 3 と、画像イベント検知部 4 と、写像関係算出部 5 と、座標変換部 6 としての各機能を実現する。CPU 6 1 0 は、各機能を実現する際に、RAM 6 3 0 又は内部記憶装置 6 4 0 を、プログラムの一時記憶媒体として使用してもよい。また、CPU 6 1 0 は、コンピュータで読み取り可能にプログラムを記憶した記録媒体 7 0 0 が含むプログラムを、図示しない記録媒体読み取り装置を用いて読み込んでもよい。あるいは、CPU 6 1 0 は、N I C 6 8 0 を介して、図示しない外部の装置からプログラムを受け取り、RAM 6 3 0 に保存して、保存したプログラムを基に動作してもよい。

40

【 0 0 6 3 】

ROM 6 2 0 は、CPU 6 1 0 が実行するプログラム及び固定的なデータを記憶する。ROM 6 2 0 は、例えば、P - ROM (Programmable-ROM) 又はフラッシュROMである 50

50

。

【0064】

R A M 6 3 0 は、C P U 6 1 0 が実行するプログラム及びデータを一時的に記憶する。R A M 6 3 0 は、例えば、D - R A M (Dynamic-RAM) である。

【0065】

内部記憶装置 6 4 0 は、監視装置 6 0 0 が長期的に保存するデータ及びプログラムを記憶する。また、内部記憶装置 6 4 0 は、C P U 6 1 0 の一時記憶装置として動作してもよい。内部記憶装置 6 4 0 は、例えば、ハードディスク装置、光磁気ディスク装置、S S D (Solid State Drive) 又はディスクアレイ装置である。

【0066】

ここで、R O M 6 2 0 と内部記憶装置 6 4 0 は、不揮発性 (non-transitory) の記録媒体である。一方、R A M 6 3 0 は、揮発性 (transitory) の記録媒体である。そして、C P U 6 1 0 は、R O M 6 2 0、内部記憶装置 6 4 0、又は、R A M 6 3 0 に記憶されているプログラムを基に動作可能である。つまり、C P U 6 1 0 は、不揮発性記録媒体又は揮発性記録媒体を用いて動作可能である。

【0067】

I O C 6 5 0 は、C P U 6 1 0 と、入力機器 6 6 0 及び表示機器 6 7 0 とのデータを仲介する。I O C 6 5 0 は、例えば、I O インターフェースカード又は U S B (Universal Serial Bus) カードである。さらに、I O C 6 5 0 は、U S B のような有線に限らず、無線を用いてもよい。

【0068】

入力機器 6 6 0 は、監視装置 6 0 0 の操作者からの入力指示を受け取る機器である。入力機器 6 6 0 は、音響信号取得部 1 及び / 又は画像信号取得部 2 として動作してもよい。入力機器 6 6 0 は、例えば、キーボード、マウス又はタッチパネルである。

【0069】

表示機器 6 7 0 は、監視装置 6 0 0 の操作者に情報を表示する機器である。表示機器 6 7 0 は、表示部 7 として動作してもよい。表示機器 6 7 0 は、例えば、液晶ディスプレイである。

【0070】

N I C 6 8 0 は、ネットワークを介した図示しない外部の装置とのデータのやり取りを中継する。N I C 6 8 0 は、音響信号取得部 1 及び / 又は画像信号取得部 2 として動作してもよい。N I C 6 8 0 は、例えば、L A N (Local Area Network) カードである。さらに、N I C 6 8 0 は、有線に限らず、無線を用いてもよい。

【0071】

このように構成された監視装置 6 0 0 は、図 1 に示されている監視装置と同様の効果を得ることができる。その理由は、監視装置 6 0 0 の C P U 6 1 0 が、プログラムに基づいて図 1 に示されている監視装置と同様の機能を実現できるためである。

【0072】

以上、実施形態を参照して本願発明を説明したが、本願発明は上記実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明の Scope 内で当業者が理解し得る様々な変更をすることができる。

【0073】

上記の実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限られない。

【0074】

(付記 1) カメラが取得した画像信号とマイクが取得した音響信号の両方で同時に検知されたイベントの発生位置に基づいて、前記発生位置間の写像関係を算出する写像関係算出部を備えたことを特徴とする監視装置。

【0075】

(付記 2) カメラが取得した前記画像信号からイベントの発生時刻と位置を検知する画

10

20

30

40

50

像イベント検知部と、マイクが取得した前記音響信号からイベントの発生時刻と位置を検知する音響イベント検知部と、前記画像信号によるイベント発生時刻と前記音響信号によるイベント発生時刻の差が閾値以下の場合に同時とみなし、前記画像信号によるイベント発生位置と前記音響信号によるイベント発生位置の間の写像関係を算出する写像関係算出部を備える付記 1 に記載の監視装置。

【 0 0 7 6 】

(付記 3) 前記画像イベント検知部は、前記画像信号によるイベントを事前に定義したキャリブレーション対象イベントと検知対象イベントに分類し、前記音響イベント検知部は、前記音響信号によるイベントを事前に定義した前記キャリブレーション対象イベントと検知対象イベントに分類し、前記画像信号によるイベント発生時刻と前記音響信号によるイベント発生時刻が同時とみなされた前記キャリブレーション対象イベントの発生位置から、前記キャリブレーション対象イベントの発生位置間の写像関係を算出する写像関係算出部と、前記写像関係に基づいて前記音響イベント検知部で算出した検知対象イベントの位置を座標変換する座標変換部を備える付記 2 に記載の監視装置。

10

【 0 0 7 7 】

(付記 4) 前記写像関係算出部は、前記写像関係の変化が閾値以上の場合に前記カメラと前記マイクの方向関係が変化すると判断する付記 3 に記載の監視装置。

【 0 0 7 8 】

(付記 5) 前記写像関係算出部は、前記写像関係として、前記キャリブレーション対象イベントの発生位置間の写像変換行列を、前記写像変換行列の変換誤差が閾値以下のイベントである *In l i e r* の数が最大となるように算出する付記 3 に記載の監視装置。

20

【 0 0 7 9 】

(付記 6) 前記写像関係算出部は、前記写像関係として、前記キャリブレーション対象イベントの発生位置間の写像変換行列と、写像変換誤差の期待値と信頼度を算出する付記 3 に記載の監視装置。

【 0 0 8 0 】

(付記 7) 前記写像関係算出部は、前記写像変換誤差の期待値として、前記音響イベント検知部で算出した前記キャリブレーション対象イベントの位置が部分領域に含まれるイベントのうち、*In l i e r* の前記写像変換誤差の標準偏差の定数倍を算出する付記 6 に記載の監視装置。

30

【 0 0 8 1 】

(付記 8) 前記写像関係算出部は、前記信頼度として、前記音響イベント検知部で算出した前記キャリブレーション対象イベントの位置が部分領域に含まれるイベントのうち、前記 *In l i e r* の変換誤差の総和の逆数又は、前記部分領域に含まれる前記キャリブレーション対象イベントに対する前記 *In l i e r* の割合に基づいて算出する付記 6 に記載の監視装置。

【 0 0 8 2 】

(付記 9) 前記座標変換部は、前記写像関係算出部で算出した前記写像変換誤差の期待値から、前記画像イベント検知部で算出した前記検知対象イベントの位置を含む部分領域における前記写像変換誤差の期待値を算出する付記 5 から 7 に記載の監視装置。

40

【 0 0 8 3 】

(付記 10) 前記画像イベント検知部で算出した検知対象イベントの位置と、前記座標変換部で算出した座標変換された検知対象イベントの位置と、前記写像関係算出部で算出した前記信頼度と、前記座標変換部で算出した前記検知対象イベントの位置を含む部分領域における前記写像変換誤差の期待値と、前記写像関係算出部で前記カメラと前記マイクの方向関係が変化すると判断した場合にアラームの、いずれか又は全てを表示する表示部を備える付記 3 から 9 に記載の監視装置。

【 0 0 8 4 】

この出願は、2016年 3月30日に提出された日本出願特願 2016 - 067021 を基礎とする優先権を主張し、その開示の全てをここに取り込む。

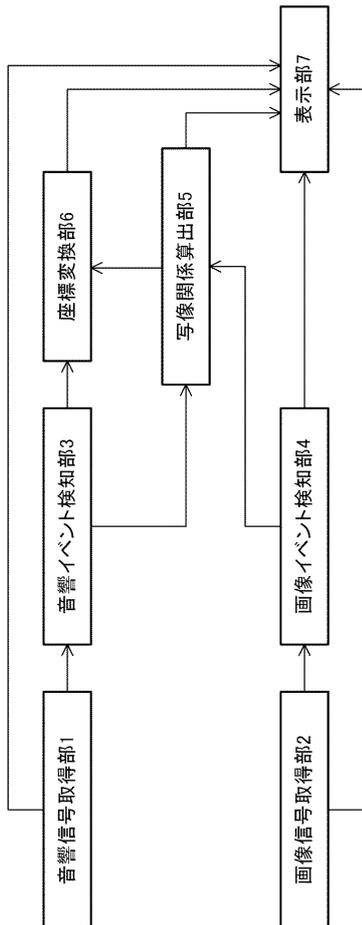
50

【符号の説明】

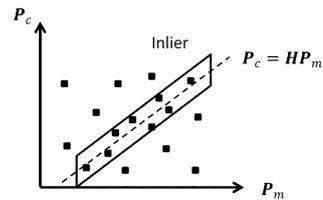
【0085】

- 1 音響信号取得部
 - 2 画像信号取得部
 - 3 音響イベント検知部
 - 4 画像イベント検知部
 - 5 写像関係算出部
 - 6 座標変換部
 - 7 表示部
- 600 監視装置
 - 610 CPU
 - 620 ROM
 - 630 RAM
 - 640 内部記憶装置
 - 650 IOC
 - 660 入力機器
 - 670 表示機器
 - 680 NIC
 - 700 記録媒体

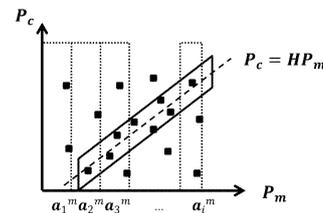
【図1】



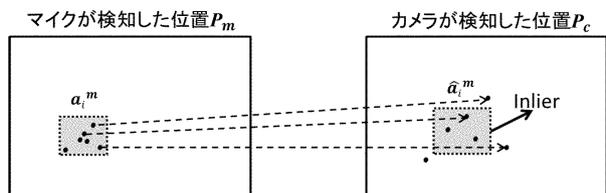
【図2】



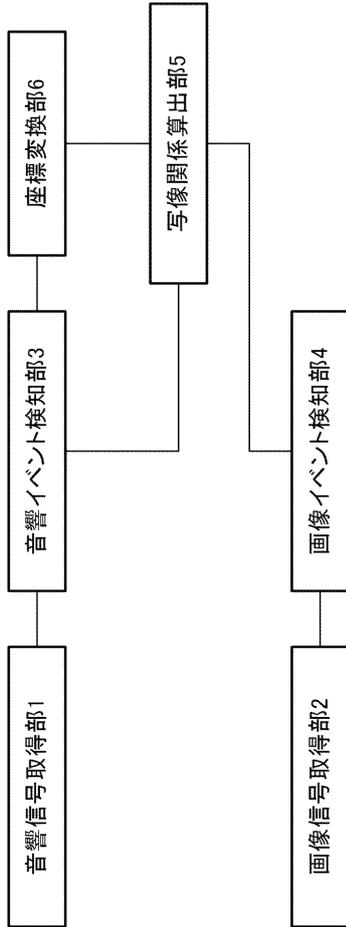
【図3】



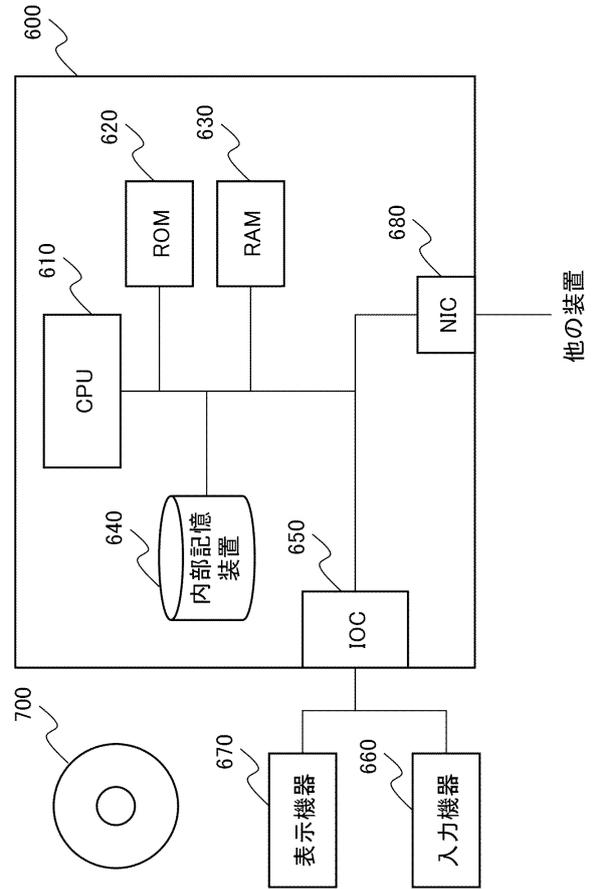
【図4】



【図5】



【図6】



フロントページの続き

(72)発明者 仙田 裕三
東京都港区芝五丁目7番1号

日本電気株式会社内

審査官 松原 徳久

(56)参考文献 特開2006-339741(JP,A)
特開2006-174216(JP,A)
特開2014-175996(JP,A)
国際公開第2014/125835(WO,A1)

(58)調査した分野(Int.Cl., DB名)

G08B23/00-31/00

G10L13/00-13/10

19/00-99/00

H04N7/18