

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4674316号  
(P4674316)

(45) 発行日 平成23年4月20日(2011.4.20)

(24) 登録日 平成23年2月4日(2011.2.4)

(51) Int.Cl.		F I			
<b>GO 6 T</b>	<b>7/60</b>	<b>(2006.01)</b>	GO 6 T	7/60	1 5 0 B
<b>GO 1 B</b>	<b>11/00</b>	<b>(2006.01)</b>	GO 6 T	7/60	1 5 0 J
<b>GO 1 B</b>	<b>11/02</b>	<b>(2006.01)</b>	GO 1 B	11/00	H
			GO 1 B	11/02	H

請求項の数 8 (全 13 頁)

(21) 出願番号	特願2005-18438 (P2005-18438)	(73) 特許権者	393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2
(22) 出願日	平成17年1月26日(2005.1.26)	(74) 代理人	100067828 弁理士 小谷 悦司
(65) 公開番号	特開2006-209334 (P2006-209334A)	(74) 代理人	100096150 弁理士 伊藤 孝夫
(43) 公開日	平成18年8月10日(2006.8.10)	(74) 代理人	100109438 弁理士 大月 伸介
審査請求日	平成19年11月9日(2007.11.9)	(72) 発明者	北原 格 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
(出願人による申告)平成16年度独立行政法人情報通信研究機構、研究テーマ「日常行動・状況理解に基づく知識共有システムの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願		(72) 発明者	小暮 潔 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
特許権者において、実施許諾の用意がある。			最終頁に続く

(54) 【発明の名称】 位置検出装置、位置検出方法及び位置検出プログラム

(57) 【特許請求の範囲】

【請求項1】

対象物を撮影して当該対象物を含む2次元画像を取得する一の撮影手段と、  
前記撮影手段により取得された2次元画像から対象物領域を抽出して当該対象物領域を3次元空間中に所定間隔で予め設定された複数の水平面上に投影する投影手段と、  
前記投影手段により各水平面上に投影された対象物の写像を積分して写像の積分値を算出する積分手段と、  
前記積分手段により算出された積分値のピークが位置する前記3次元空間中の水平位置を対象物の位置として検出する検出手段とを備えることを特徴とする位置検出装置。

【請求項2】

前記検出手段は、対象物の3次元形状を近似した近似フィルタと前記積分値との畳み込み値が最大となる前記3次元空間中の水平位置を対象物の位置として検出することを特徴とする請求項1記載の位置検出装置。

【請求項3】

前記検出手段は、前記ピークが位置する前記3次元空間中の水平位置に対象物の写像が存在する最も高い水平面の高さを対象物の高さとして検出することを特徴とする請求項1又は2記載の位置検出装置。

【請求項4】

前記検出手段は、前記積分手段により算出された積分値に複数のピークが存在する場合、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔以上のピーク毎に、当

該ピークが位置する前記3次元空間中の水平位置を対象物の位置として検出するとともに、当該ピークが位置する前記3次元空間中の水平位置に対象物の写像が存在する最も高い水平面の高さを対象物の高さとして検出することを特徴とする請求項3記載の位置検出装置。

【請求項5】

前記検出手段は、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔未満である場合、第1の対象物と、前記撮影手段から前記第1の対象物より離れている第2の対象物とが重なっていると判断し、基準高さ以上のピークが位置する前記3次元空間中の水平位置を第1の対象物の位置として検出するとともに、第1の対象物の高さとして所定時間前に検出された高さを用いることを特徴とする請求項3又は4記載の位置検出装置。

10

【請求項6】

前記検出手段は、前記3次元空間における前記撮影手段の撮影中心位置と前記3次元空間における前記基準高さ以上のピークの3次元位置とを結ぶ直線と、第2の対象物の高さとして所定時間前に検出された高さ位置する水平面とが交わる点が位置する前記3次元空間中の水平位置を第2の対象物の位置として検出するとともに、第2の対象物の高さとして所定時間前に検出された高さを用いることを特徴とする請求項5記載の位置検出装置。

【請求項7】

一の撮影手段、投影手段、積分手段、及び検出手段を備える位置検出装置の位置検出方法であって、

20

前記撮影手段が、対象物を撮影して当該対象物を含む2次元画像を取得する第1のステップと、

前記投影手段が、前記撮影手段により取得された2次元画像から対象物領域を抽出して当該対象物領域を3次元空間中に所定間隔で予め設定された複数の水平面上に投影する第2のステップと、

前記積分手段が、前記投影手段により各水平面上に投影された対象物の写像を積分して写像の積分値を算出する第3のステップと、

前記検出手段が、前記積分手段により算出された積分値のピークが位置する前記3次元空間中の水平位置を対象物の位置として検出する第4のステップとを含むことを特徴とする位置検出方法。

30

【請求項8】

対象物を撮影して当該対象物を含む2次元画像を取得する一の撮影手段により取得された2次元画像から対象物領域を抽出して当該対象物領域を3次元空間中に所定間隔で予め設定された複数の水平面上に投影する投影手段と、

前記投影手段により各水平面上に投影された対象物の写像を積分して写像の積分値を算出する積分手段と、

前記積分手段により算出された積分値のピークが位置する前記3次元空間中の水平位置を対象物の位置として検出する検出手段としてコンピュータを機能させることを特徴とする位置検出プログラム。

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、一の撮影手段を用いて、対象物、例えば、人間の3次元位置を検出する位置検出装置及び位置検出方法に関するものである。

【背景技術】

【0002】

近年、画像取得技術の進歩及びネットワーク化により、画像監視システムが市中に設置されるようになってきている。このような画像監視システムが、近い将来、安全面に問題のある人を監視したり、歩行者、特に老人及び子供に有益な情報を提供したり、種々の重要な目的のために使用されることが予想され、これらのサービスから提供されるデータを分析

50

することにより生活の質を向上させることができる。

【0003】

上記のような画像監視システムには、人間の位置を検出して追跡することが必要となり、例えば、複数のカメラを使用したステレオ技術を用いて人間の3次元位置を追跡することが報告されている（非特許文献1参照）。このステレオ技術を用いて人間を追跡する場合、人間が重なり合った状態でも、正確に各人間の3次元位置を検出することができる。

【非特許文献1】エイミタル(A. Mittal)他、「M2Tトラッカー：混雑した状態の人間の分割及び追跡に対する複数視点のアプローチ」(M2Tracker: A Multi-View Approach to Segmenting and Tracking People in a Cluttered Scene)、コンピュータビジョン国際ジャーナル(International Journal of Computer Vision)、2003年、Vol. 51(3), p. 189 - p. 203

10

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、上記のように複数のカメラを用いた場合、多数の画像データを処理する必要があり、計算コストが増加する。また、一つの撮影範囲に対して複数台のカメラが必要となり、広い範囲を撮影する場合、多数のカメラが必要となり、装置のコストが増大する。このため、1台のカメラを用いて撮影した画像を3次元空間中の所定高さに配置された仮想平面に投影して人間の位置を検出することも考えられるが、この場合、水平面の高さが固定されるため、測定物体の高さと想定した平面の高さが一致しない場合、正確な2次元位置を検出することは困難である。一方、本発明が測定対象としている人間の身長は、個人ごとに異なるのが一般的であるため、測定可能な高さが固定される問題の影響は深刻である。

20

【0005】

本発明の目的は、一の撮影手段を用いて簡略な検出処理により、人間のように様々な高さを有する物体の2次元位置を検出することができる位置検出装置、位置検出方法及び位置検出プログラムを提供することである。

【課題を解決するための手段】

【0006】

本発明に係る位置検出装置は、対象物を撮影して当該対象物を含む2次元画像を取得する一の撮影手段と、撮影手段により取得された2次元画像から対象物領域を抽出して当該対象物領域を3次元空間中に所定間隔で予め設定された複数の水平面上に投影する投影手段と、投影手段により各水平面上に投影された対象物の写像を積分して写像の積分値を算出する積分手段と、積分手段により算出された積分値のピークが位置する3次元空間中の水平位置を対象物の位置として検出する検出手段とを備えるものである。

30

【0007】

本発明に係る位置検出装置においては、一の撮影手段により対象物を撮影して当該対象物を含む2次元画像を取得され、取得された2次元画像から抽出された対象物領域が3次元空間中に所定間隔で予め設定された複数の水平面上に投影され、各水平面上に投影された対象物の写像を積分して写像の積分値が算出され、算出された積分値のピークが位置する3次元空間中の水平位置が対象物の位置として検出される。

40

【0008】

このように、対象物が直立していると仮定して、積分値のピークの位置を対象物の水平位置として検出するとともに、このピークの位置に対象物の写像が存在する最も高い水平面の高さを対象物の高さとして検出しているため、一の撮影手段により撮影された画像のみを用いて、人間のよう様々な高さを有する物体の2次元位置を検出することができる。また、一の撮影手段により撮影された画像のみを用いているため、検出処理を簡略化することができる。

【0009】

検出手段は、対象物の3次元形状を近似した近似フィルタと積分値との畳み込み値が最

50

大となる 3 次元空間中の水平位置を対象物の位置として検出することが好ましい。

【 0 0 1 0 】

この場合、近似フィルタにより対象物のみを抽出することができ、他の物体等による外乱に対してロバストな検出を行うことができる。

【 0 0 1 1 】

検出手段は、ピークが位置する 3 次元空間中の水平位置に対象物の写像が存在する最も高い水平面の高さを対象物の高さとして検出することが好ましい。

【 0 0 1 2 】

この場合、ピークが位置する 3 次元空間中の水平位置に対象物の写像が存在する最も高い水平面の高さを対象物の高さとして検出することができるので、3 次元物体の登頂部の位置（3 次元物体が地図上で存在する 2 次元座標及びその高さ）を物体の 3 次元位置として検出することができる。

10

【 0 0 1 3 】

検出手段は、積分手段により算出された積分値に複数のピークが存在する場合、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔以上のピーク毎に、当該ピークが位置する 3 次元空間中の水平位置を対象物の位置として検出するとともに、当該ピークが位置する 3 次元空間中の水平位置に対象物の写像が存在する最も高い水平面の高さを対象物の高さとして検出することが好ましい。

【 0 0 1 4 】

この場合、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔以上の各ピークが対象物によるものであると判断することができるので、複数の対象物が近接している状態でも、各対象物の 3 次元位置を高精度に検出することができる。

20

【 0 0 1 5 】

検出手段は、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔未満である場合、第 1 の対象物と、撮影手段から第 1 の対象物より離れている第 2 の対象物とが重なっていると判断し、基準高さ以上のピークが位置する 3 次元空間中の水平位置を第 1 の対象物の位置として検出するとともに、第 1 の対象物の高さとして所定時間前に検出された高さを用いることが好ましい。

【 0 0 1 6 】

この場合、第 1 及び第 2 の対象物が非常に近接して撮影画像上で重なり合っているも、撮影手段側に位置する第 1 の対象物の 3 次元位置を高精度に検出することができる。

30

【 0 0 1 7 】

検出手段は、3 次元空間における撮影手段の撮影中心位置と 3 次元空間における基準高さ以上のピークの 3 次元位置とを結ぶ直線と、第 2 の対象物の高さとして所定時間前に検出された高さに位置する水平面とが交わる点が位置する 3 次元空間中の水平位置を第 2 の対象物の位置として検出するとともに、第 2 の対象物の高さとして所定時間前に検出された高さを用いることが好ましい。

【 0 0 1 8 】

この場合、第 1 及び第 2 の対象物が非常に近接して撮影画像上で重なり合っているも、撮影手段から第 1 の対象物より離れている第 2 の対象物の 3 次元位置を高精度に検出することができる。

40

【 0 0 1 9 】

本発明に係る位置検出方法は、一の撮影手段、投影手段、積分手段、及び検出手段を備える位置検出装置の位置検出方法であって、前記撮影手段が、対象物を撮影して当該対象物を含む 2 次元画像を取得する第 1 のステップと、前記投影手段が、前記撮影手段により取得された 2 次元画像から対象物領域を抽出して当該対象物領域を 3 次元空間中に所定間隔で予め設定された複数の水平面上に投影する第 2 のステップと、前記積分手段が、前記投影手段により各水平面上に投影された対象物の写像を積分して写像の積分値を算出する第 3 のステップと、前記検出手段が、前記積分手段により算出された積分値のピークが位置する 3 次元空間中の水平位置を対象物の位置として検出する第 4 のステップとを含むも

50

のである。

【 0 0 2 0 】

本発明に係る位置検出プログラムは、対象物を撮影して当該対象物を含む2次元画像を取得する一の撮影手段により取得された2次元画像から対象物領域を抽出して当該対象物領域を3次元空間中に所定間隔で予め設定された複数の水平面上に投影する投影手段と、投影手段により各水平面上に投影された対象物の写像を積分して写像の積分値を算出する積分手段と、積分手段により算出された積分値のピークが位置する3次元空間中の水平位置を対象物の位置として検出する検出手段としてコンピュータを機能させるものである。

【 発明の効果 】

【 0 0 2 1 】

本発明によれば、対象物が直立していると仮定して、積分値のピークの位置を対象物の水平位置として検出しているので、一の撮影手段により撮影された画像のみを用いて、人間のように様々な高さを有する物体の2次元位置を検出することができるのと同時に、検出処理を簡略化することができる。

【 発明を実施するための最良の形態 】

【 0 0 2 2 】

以下、本発明の一実施の形態による位置検出装置について図面を参照しながら説明する。図1は、本発明の一実施の形態による位置検出装置の構成を示すブロック図である。

【 0 0 2 3 】

図1に示す位置検出装置は、1台のビデオカメラ11、投影部12、積分部13、検出部14及び追跡部15を備える。投影部12、積分部13、検出部14及び追跡部15は、入力装置、ROM(リードオンリメモリ)、CPU(中央演算処理装置)、RAM(ランダムアクセスメモリ)、画像I/F(インターフェース)部及び外部記憶装置等を備えるコンピュータを用いて後述する各処理を行うための位置検出プログラムをCPU等で実行することにより実現される。なお、投影部12、積分部13、検出部14及び追跡部15の構成例は、本例に特に限定されず、各ブロックを専用のハードウェアから構成したり、一部のブロック又はブロック内の一部の処理のみを専用のハードウェアで構成したりする等の種々の変更が可能である。

【 0 0 2 4 】

ビデオカメラ11は、通常の監視カメラ等から構成され、撮影空間の所定位置、例えば、天井の隅に取り付けられる。ビデオカメラ11は、後述する校正方法により校正されており、対象物である人間を撮影して人間を含む2次元画像を取得して投影部12へ出力する。なお、ビデオカメラ11としては、既設の監視カメラを用いてもよい。

【 0 0 2 5 】

投影部12は、取得された2次元画像から対象物領域となる人物領域を抽出し、抽出した人物領域を3次元空間中に所定間隔で予め設定された複数の水平面(各水平面の高さは既知)上に投影する。なお、2次元画像における背景領域(人間以外の領域)と人物領域(人間のシルエット)との分離には、背景差分等の公知の画像処理手法を用いることができる。

【 0 0 2 6 】

ここで、ビデオカメラ11の校正方法について説明する。3次元空間中に設定された複数の水平面上に人物領域画像を投影するためには、ビデオカメラ11の撮影画像と3次元空間中に設定された二つの基準面(水平面)との間の2次元射影変換行列が必要となる。まず、二つの基準面のうち一方の面が地面の高さ(高さ0)に位置し、他方の面は高さY<sub>h</sub>に位置するものとし、撮影空間となる3次元空間に普通の人間の身長より高い4本の校正用バーを地面に対して垂直に所定間隔に(4本の校正用バーが立方体をなすように)設置する。各校正用バーの頭部及び底部には色付けされたマーカーが設けられ、各基準面に対して4個のマーカーを観測することにより、2次元射影変換行列H<sub>0</sub>、H<sub>1</sub>を計算する。ここで、二つの基準面の間に位置する中間の高さの水平面の高さをY<sub>n</sub>とすると、この水平面の2次元射影変換行列H<sub>n</sub>は、下記の式(1)に示すように、2次元射影変換行列

10

20

30

40

50

H0、H1を用いた補間により推定することができる。

【0027】

$$H_n = ((Y_h - Y_n) H_0 + (Y_n) H_1) / Y_h \quad (1)$$

上記のようにして、各設置位置で校正用バーの3次元位置を正確に測定することにより、複数の水平面と3次元世界座標系との間の変換を容易に行うことができ、ビデオカメラ11の校正を簡略に行うことができる。

【0028】

積分部13は、各水平面上に投影された人間の写像を積分して写像の積分値を算出して検出部14へ出力する。検出部14は、算出された積分値のピークが位置する3次元空間中の水平位置を人間の位置として検出するとともに、このピークが位置する3次元空間中の水平位置に人物領域の写像が存在する最も高い水平面の高さを人物の高さとして検出する。追跡部15は、検出された3次元位置を用いて人間を追跡し、各観測時刻における人間の3次元位置(水平位置及び高さ)を記憶している。

10

【0029】

なお、ビデオカメラ11がパン方向及びチルト方向に移動可能な可動カメラから構成される場合、追跡部15は、追跡結果に応じてビデオカメラ11の撮影位置を制御するようにしてもよい。この場合、その撮影位置でカメラ校正を行うことにより、常に最適な観測位置で対象物を撮影することができ、3次元位置を高精度に検出することができる。

【0030】

具体的には、検出部14は、対象物の3次元形状を近似した近似フィルタと積分値との畳み込み値が最大となる3次元空間中の水平位置を対象物の位置として検出するとともに、この対象物の位置において人間の写像が存在する最も高い水平面の高さを人間の高さとして検出することにより、人間の3次元位置を検出する。

20

【0031】

また、検出部14は、近接する複数の人間が撮影され、積分値に複数のピークが存在する場合、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔以上、例えば、1m以上あるピーク毎に、当該ピークが位置する3次元空間中の水平位置を人間の位置として検出するとともに、当該ピークが位置する3次元空間中の水平位置に人間の写像が存在する最も高い水平面の高さを対象物の高さとして検出する。

【0032】

さらに、検出部14は、複数の人間が非常に近接して撮影画像上で重なり、隣接するピーク間の谷の高さとピークの高さとの間隔が所定間隔未満である場合、二人以上の人間が重なっていると判断し、基準高さ以上、例えば、2.5m以上のピークが位置する3次元空間中の水平位置をビデオカメラ11に近い方の人間の位置として検出するとともに、この人間の高さとして所定時間前、例えば、1又は数フレーム前に検出された高さを追跡部15から読み出し、読み出した高さを現在の高さとして用いる。

30

【0033】

このとき、検出部14は、ビデオカメラ11から遠い方の人間の高さとして所定時間前、例えば、1又は数フレーム前に検出された高さを追跡部15から読み出し、読み出した高さを現在の高さとして用いるとともに、3次元空間におけるビデオカメラ11のカメラ中心位置と3次元空間における基準高さ以上のピークの3次元位置とを結ぶ直線と、追跡部15から読み出した高さに位置する水平面との交点が位置する3次元空間中の水平位置をこの人間の位置として検出する。

40

【0034】

上記のように、本実施の形態では、3次元空間中に一つの水平面ではなく、複数の水平面を設定し、校正されたビデオカメラ11により撮影された単眼画像だけを使用している。この単眼画像すなわち2次元情報のみを使用して3次元位置を推定する場合、次元数が1次元不足するが、これを補うために人間が垂直に立っていることを仮定し、複数の水平面上に人間のシルエットである人物領域を投影し、この投影領域を垂直軸に沿って積分した積分値のピークを用いて人間の3次元位置(水平位置及び高さ)を検出し、3次元空間

50

において種々の高さの人間を追跡することができる。

【 0 0 3 5 】

なお、本実施の形態では、ビデオカメラ 1 1 が撮影手段の一例に相当し、投影部 1 2 が投影手段の一例に相当し、積分部 1 3 が積分手段の一例に相当し、検出部 1 4 が検出手段の一例に相当する。

【 0 0 3 6 】

次に、一人の人間を撮影した画像を用いて当該人間の 3 次元位置を検出する通常 3 次元位置検出処理について説明する。図 2 は、図 1 に示す位置検出装置による通常 3 次元位置検出処理を説明するための原理図である。

【 0 0 3 7 】

図 2 に示すように、3 次元空間中に複数の水平面 H P が設定され、各水平面 H P の 2 次元座標は、X 軸及び Y 軸によって規定され、高さ方向が Y 軸であり、対象物である人間 O B は、説明を容易にするため、図中にバー形状で示している。まず、投影部 1 2 は、高さ  $Y_n$  ( $0 < Y_n < Y_h$ ) を変更することにより、ビデオカメラ 1 1 により撮影された 2 次元画像から分割された人物領域 S を水平面 n (各水平面 H P) に投影する。この投影結果  $P(Y_n)$  は、高さ  $Y_n$  を用いて下記式 (2) により表すことができる。

【 0 0 3 8 】

$$P(Y_n) = H_n S \quad (2)$$

ここで、人間 O B が直立していると仮定すると、投影された人物領域 P R (図中の斜線部) は、常に、実際の人間が存在する水平面 n 上のある点 (X, Z) を含み、全ての水平面 n を統合 (マージ) することにより人物の 3 次元位置を推定することができる。

【 0 0 3 9 】

積分部 1 3 は、下記式 (3) を用いて、点 (X, Z) で垂直軸 Y に沿って対象物の写像である人物領域 P (Y) の積分値  $I_{tg}(X, Z)$  を算出する。

【 0 0 4 0 】

【数 1】

$$I_{tg}(X, Z) = \int_{Y=Y_0}^{Y_h} P(Y) dY \quad (3)$$

【 0 0 4 1 】

また、積分値のピークを検出するために、人間の形状を近似する近似フィルタである 3 次元凸フィルタが使用され、例えば、下記式 (4) に表される円筒形状フィルタ  $Cvx(X, Z)$  を用いることができ、ここで、r は円筒の底部の半径である。なお、近似フィルタとしては、本例に特に限定されず、対象物の 3 次元形状等に応じて種々のものを用いることができる。また、積分値のピークを検出する方法も、以下の方法に特に限定されず、種々の方法を用いることができ、例えば、上記式 (3) により表される積分値を微分してピークを検出する方法 (微分フィルタを用いる方法) を用いてもよい。

【 0 0 4 2 】

10

20

30

40

【数 2】

$$C_{vx}(X, Z) = \begin{cases} Y_h & \text{if } (\sqrt{X^2 + Z^2} \leq r) \\ \dots\dots & \\ 0 & \text{else} \end{cases} \quad (4)$$

10

【0043】

検出部14は、上記の凸フィルタ $C_{vx}(X, Z)$ と積分値 $I_{tg}(X, Z)$ との畳み込み値を算出し、畳み込み値が最大となる座標値 $(X_m, Z_m)$ を人間の水平位置として検出し、人間の水平位置 $(X_m, Z_m)$ における人物領域の最大高さ(人間の写像が存在する最も高い水平面の高さ)を人間の高さ $Y_m$ として検出する。図2に示す例では、畳み込み値 $CV$ が算出され、そのピーク $PI$ が位置する座標値 $(X_m, Z_m)$ が人間の水平位置として検出され、そのピーク $PI$ が位置する座標値 $(X_m, Z_m)$ における人物領域の最大高さが人間の高さ $Y_m$ として検出される。

20

【0044】

上記の積分処理の計算中に画像中の小さなホールやクラックなどの分割エラーがスムーズ化され、各領域を完全に分割する必要がなくなるとともに、凸フィルタ $C_{vx}(X, Z)$ により人間のみを抽出することができ、他の物体等による外乱に対してロバストな検出を行うことができる。

【0045】

次に、複数の人間が撮影された画像を用いて各人間の3次元位置を検出する複数3次元位置検出処理について説明する。まず、複数の人間が入力画像に撮影され、各人物領域を分割できる場合、第1の複数3次元位置検出処理として、投影部12は、2次元画像を一人の人物領域のみを含む分割領域に分割し、積分部13及び検出部14は、分割領域毎に一人に対して上記の通常3次元位置検出処理を繰り返すことにより、全ての人間の3次元位置を検出する。

30

【0046】

しかしながら、例えば、撮影された複数の人間が手をつないでいる場合、二つの人物領域の一部が連結され、二つの人物領域が弱く連結されるため、上記のように領域を分割することができない。この場合、第2の複数3次元位置検出処理として、複数の人物領域の一部が連結された連結領域を一つの領域として上記の通常3次元位置検出処理が実行される。

【0047】

すなわち、上記と同様に、投影部12は、複数の人物領域の一部が連結された連結領域を一つの領域として複数の水平面に投影し、積分部13は、投影領域の積分値 $I_{tg}(X_n, Z_n)$ を算出し、検出部14は、凸フィルタ $C_{vx}(X_n, Z_n)$ 及び積分値 $I_{tg}(X_n, Z_n)$ の畳み込み値を算出する。

40

【0048】

次に、検出部14は、隣接する畳み込み値のピーク間に位置する谷の高さと畳み込み値のピークの高さとの間隔と、予め記憶している所定間隔とを比較して所定間隔以上のピークを検出する。次に、検出部14は、検出した畳み込み値のピークと、予め記憶している所定値、例えば、1.5mとを比較して所定値以上のピークを抽出する。次に、検出部14は、抽出したピーク毎に当該ピークが位置する座標点 $(X_n, Z_n)$ を一人の人間の水平位置として検出するとともに、水平位置 $(X_n, Z_n)$ における人物領域の最大高さを

50



人間の高さ  $Y_n$  として検出する。上記の処理が 3 次元空間の全体に適用され、複数の人間の 3 次元位置が順次検出される。なお、上記のピークの高さ及び谷の高さは、人間の 3 次元位置の検出と同様に検出され、ピーク及び谷が位置する座標点をピーク及び谷の水平位置として検出し、検出された水平位置における人物領域の最大高さをピークの高さとして検出することができる。

【 0 0 4 9 】

上記の処理により、隣接する畳み込み値のピーク間の谷の高さと畳み込み値のピークの高さとの間隔が所定間隔以上の各ピークが人間によるものであると判断することができるので、複数の人間が近接している状態でも、各人間の 3 次元位置を高精度に検出することができる。

10

【 0 0 5 0 】

次に、ビデオカメラ 11 のカメラ中心と人間とを結ぶ直線上に他の人間がいるとき、二人の人間の人物領域が大きく重なり合い、二つの人物領域が強く連結されるため、隣接する畳み込み値のピーク間に位置する谷の高さと畳み込み値のピークの高さとの間隔が所定間隔未満となり、上記のフィルタ処理では連結している領域を分割することが困難なため、上記の第 2 の複数 3 次元位置検出処理を適用することができない。この場合、数フレーム前に既に検出された各人間の位置を用いて重なり合った人間の位置を検出する第 3 の複数 3 次元位置検出処理が実行される。なお、同じ投影領域に二人以上の人間がいる場合は、下記の処理を人間の数だけ繰り返す。

【 0 0 5 1 】

20

まず、上記と同様に、投影部 12 は、二人の人間を含む連結領域を複数の水平面に投影し、積分部 13 は、投影領域の積分値  $I_{tg}(X_n, Z_n)$  を算出する。検出部 14 は、凸フィルタ  $Cvx(X_n, Z_n)$  及び積分値  $I_{tg}(X_n, Z_n)$  の畳み込み値を算出し、最大となる座標値  $(X_0(t), Z_0(t))$  における人物領域の最大高さ  $Y_0(t)$  を検出する(ここで、 $t$  は時間パラメータ)。

【 0 0 5 2 】

次に、検出部 14 は、検出した高さ  $Y_0(t)$  と、予め記憶している基準値  $Y_r$  (例えば、2.5 m) とを比較して高さ  $Y_0(t)$  が基準値  $Y_r$  以上の場合、ビデオカメラ 11 に近い方の人間が水平位置  $(X_0(t), Z_0(t))$  にいると判断し、前のフレームを用いて検出された同一人物(ビデオカメラ 11 に近い方の人間)の高さ  $Y_0(t-1)$  を追跡部 15 から読み出し、読み出した高さ  $Y_0(t-1)$  を現在の高さ  $Y_0(t)$  として用いる。

30

【 0 0 5 3 】

上記の処理により、検出部 14 は、ビデオカメラ 11 に近い方の人間の 3 次元位置として、水平位置  $(X_0(t), Z_0(t))$  及び高さ  $Y_0(t-1)$  を取得することができるので、二人の人間が非常に近接して撮影画像上で重なり合っている場合、ビデオカメラ 11 に近い方の人間の 3 次元位置を高精度に検出することができる。

【 0 0 5 4 】

次に、検出部 14 は、ビデオカメラ 11 から遠い方の人間として前のフレームを用いて検出された同一人物の高さ  $Y_1(t-1)$  を追跡部 15 から読み出し、読み出した高さ  $Y_1(t-1)$  をビデオカメラ 11 から遠い方の人間の現在の高さ  $Y_1(t)$  として用いるとともに、基準値  $Y_r$  以上のピークとして検出された座標値  $(X_0(t), Y_0(t), Z_0(t))$  とビデオカメラ 11 のカメラ中心  $(X_c, Y_c, Z_c)$  とを結ぶ直線を設定し、読み出した高さ  $Y_1(t-1)$  に位置する水平面とこの直線との交点の座標値  $(X_1(t), Z_1(t))$  をビデオカメラ 11 から遠い方の人間の位置として検出する。

40

【 0 0 5 5 】

上記の処理により、検出部 14 は、ビデオカメラ 11 から遠い方の人間の 3 次元位置として、水平位置  $(X_1(t), Z_1(t))$  及び高さ  $Y_1(t-1)$  を取得することができるので、二人の人間が非常に近接して撮影画像上で重なり合っている場合、ビデオカメラ 11 から遠い方の人間の 3 次元位置を高精度に検出することができる。

50

## 【0056】

なお、上記の同一人物予測処理としては、例えば、追跡部15において、各フレームで検出された人間の3次元位置を記憶しておき、各3次元位置を用いて人間の移動（例えば、等速度移動）に対する速度予測及び/又は加速度予測を行うことにより現在の3次元位置を予測し、予測した3次元位置と検出された現在の水平位置とが最も近い人間が、現在の水平位置にいる人間と同一人物であると判定することにより行うことができる。この場合、速度予測及び/又は加速度予測を用いた人間の3次元位置予測を3次元空間において直接行っているため、投影による幾何学的な非線形性の影響を受けず、高精度に速度予測及び/又は加速度予測を行うことができ、同一人物を高精度に判定することができる。

## 【0057】

次に、上記の位置検出装置を用いて実際に人間の3次元位置を検出した結果について説明する。図3は、撮影空間を説明するための模式図であり、図4は、図1に示す位置検出装置による検出結果の一例を示す図である。

## 【0058】

図3に示すように、ビデオカメラ11をそのカメラ中心の高さが2.6mになる位置に取り付け、2.5×1.1mの撮影空間（廊下）に対して水平面において68度傾けるとともに撮影空間を見下ろすように配置し、撮影空間を移動する人間を撮影して320×240pixelの画像を1秒間隔で100枚取得した。撮影空間には、22個のランドマークを配置し、3次元レーザー測定装置を用いて全ての点の3次元位置を測定し、この3次元位置情報と撮影画像の2次元位置情報とを用いてビデオカメラ11を校正した。撮影画像の空間分解能は距離に依存し、廊下の端部に位置する対象物を撮影した場合の空間分解能は、約56mmであり、撮影された対象物がビデオカメラ11の正面に立っている場合の空間分解能は、約17mmであった。

## 【0059】

上記の測定条件の下、人間の3次元位置を検出した結果、図4に示す結果を得られた。まず、図4の(a)は一人の人間が歩いている場面、(b)は一人の人間が屈伸している場面、(c)は二人の人間がすれ違う場面（人物領域が弱く結合されている状態）、(d)は二人の人間が撮影画像中で重なり合った場面（人物領域が強く結合されている状態）をそれぞれ示している。また、図4の上段は撮影画像を示し、中段は撮影画像から抽出された人物領域を示し、下段は撮影空間の水平面への写像を積分した結果を示し、上段及び下段の四角及び丸は検出されたピークの位置すなわち人間の位置を示している。図4から、上記の4種類の場面において、一人又は二人の人間を正確に検出することができ、複数の人間が重なり合っても、正確に追跡することができることがわかった。

## 【0060】

上記のように、本発明では、入力センサとして1台のビデオカメラだけを必要とするので、既に広く設置されている監視カメラを用いて人間の3次元位置を容易に検出することができ、また、検出結果が3次元座標系で記述されているので、他のセンサ（例えば、GPS、RFID）と同じ内容を共有したり、ネットワークを介してロボットをナビゲートしたりする場合等に好適に用いることができる。

## 【0061】

なお、上記の説明では、検出される対象物として、人間を例に説明したが、他の動物、他の移動体等にも同様に適用することができ、同様の効果を得ることができる。

## 【図面の簡単な説明】

## 【0062】

【図1】本発明の一実施の形態による位置検出装置の構成を示すブロック図である。

【図2】図1に示す位置検出装置による通常3次元位置検出処理を説明するための原理図である。

【図3】撮影空間を説明するための模式図である。

【図4】図1に示す位置検出装置による検出結果の一例を示す図である。

## 【符号の説明】

10

20

30

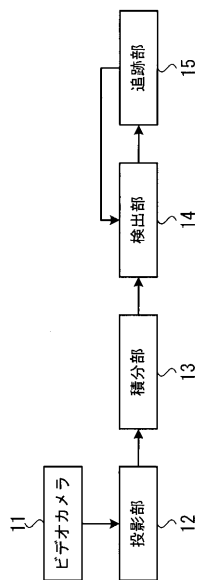
40

50

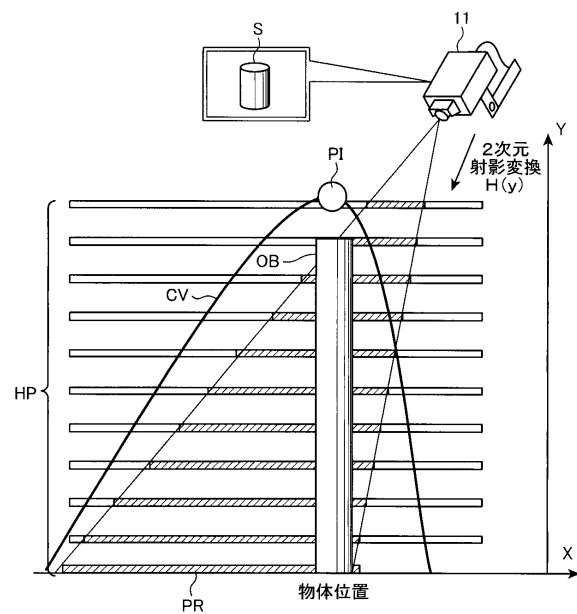
【 0 0 6 3 】

- 1 1 ビデオカメラ
- 1 2 投影部
- 1 3 積分部
- 1 4 検出部
- 1 5 追跡部

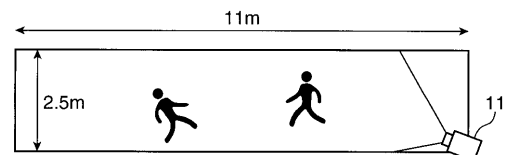
【 図 1 】



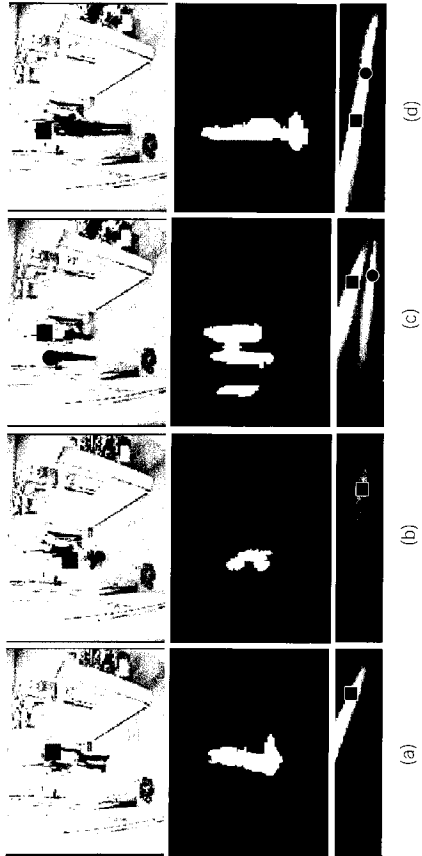
【 図 2 】



【 図 3 】



【 図 4 】



---

フロントページの続き

(72)発明者 萩田 紀博

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 松尾 俊介

(56)参考文献 特開2002-197463(JP,A)

澤井清明, 村上伸一, 単眼視による3次元人物動作の認識に関する一検討, 映像情報メディア学会技術報告2003年2月開催分(上巻), 日本, 社団法人映像情報メディア学会, 2003年2月4日, 第27巻, 第9号, p.59~64

(58)調査した分野(Int.Cl., DB名)

G06T 7/00~7/60

G01B 11/00

G01B 11/02