

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-74033

(P2007-74033A)

(43) 公開日 平成19年3月22日(2007.3.22)

(51) Int.CI.	F 1	テーマコード (参考)
H04N 5/232 (2006.01)	H04N 5/232	C 2H044
G03B 15/00 (2006.01)	H04N 5/232	B 2H105
G03B 17/56 (2006.01)	G03B 15/00	R 5C122
G02B 7/08 (2006.01)	G03B 15/00	P
	GO3B 15/00	Q

審査請求 未請求 請求項の数 23 O L (全 25 頁) 最終頁に続く

(21) 出願番号	特願2005-255620 (P2005-255620)	(71) 出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成17年9月2日 (2005.9.2)	(74) 代理人	100076428 弁理士 大塚 康徳
		(74) 代理人	100112508 弁理士 高柳 司郎
		(74) 代理人	100115071 弁理士 大塚 康弘
		(74) 代理人	100116894 弁理士 木村 秀二
		(72) 発明者	伊藤 嘉則 東京都大田区下丸子3丁目30番2号 キ ヤノン株式会社内
			F ターム (参考) 2H044 DA02 DB02 DC01 DE01
			最終頁に続く

(54) 【発明の名称】撮像装置及びその制御方法及びプログラム及び記憶媒体

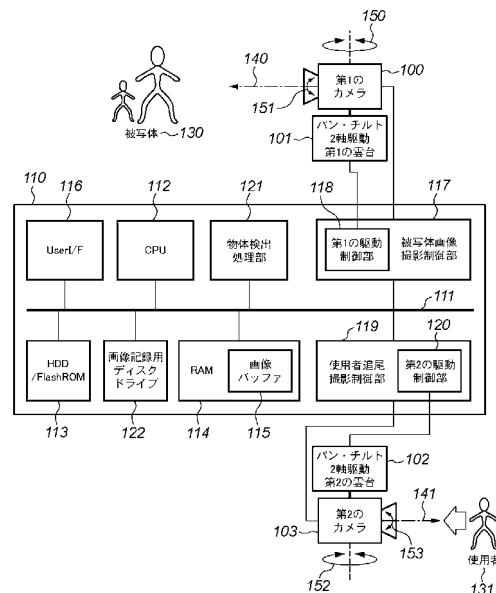
(57) 【要約】

【課題】使用者の自由度を飛躍的に向上させ、撮影操作に専念しなくとも所望の画像・映像を撮影し記録できるようとする。

【解決手段】被写体となる第1の対象を撮影するための第1のカメラ100と、使用者となる第2の対象を撮影するための第2のカメラ103と、第2のカメラにより撮影された画像に基づいて、第2の対象131の状態情報を検出する状態検出部112と、状態検出部により検出された第2の対象の状態情報に基づいて、第1の対象が存在すると予想される位置に関する情報である予想位置情報を検出する予想位置検出部112と、予想位置検出部により検出された予想位置情報に基づいて、第1の対象を撮影するように第1のカメラを制御する被写体撮影制御部117とを具備する。

【選択図】

図1



【特許請求の範囲】**【請求項 1】**

第1の対象を撮影するための少なくとも1つの第1の撮像手段と、

第2の対象を撮影するための少なくとも1つの第2の撮像手段と、

前記第2の撮像手段により撮影された画像に基づいて、前記第2の対象の状態情報を検出する状態検出手段と、

前記状態検出手段により検出された前記第2の対象の状態情報に基づいて、前記第1の対象が存在すると予想される位置に関する情報である予想位置情報を検出する予想位置検出手段と、

前記予想位置検出手段により検出された前記予想位置情報をに基づいて、前記第1の対象を撮影するように前記第1の撮像手段を制御する第1の撮影制御手段と、10

を具備することを特徴とする撮像装置。

【請求項 2】

前記予想位置検出手段により検出された前記予想位置情報をに基づいて、前記第1の撮像手段の撮像光軸の方向を変更する撮像方向変更手段を備えることを特徴とする請求項1に記載の撮像装置。

【請求項 3】

前記第2の撮像手段により撮影された画像から前記第2の対象を検出する第2の対象検出手段と、前記第2の対象検出手段の検出結果に基づいて、前記第2の撮像手段が前記第2の対象を撮像し続けるように、前記第2の撮像手段の撮像光軸の方向を変更する撮像方向変更手段とをさらに備えることを特徴とする請求項1に記載の撮像装置。20

【請求項 4】

前記予想位置検出手段により検出された前記予想位置情報をに基づいて、前記第1の撮像手段の撮像倍率を制御する拡大撮影手段をさらに備えることを特徴とする請求項1に記載の撮像装置。

【請求項 5】

前記状態検出手段により検出された前記第2の対象の状態情報をに基づいて、前記第1の撮像手段の撮像倍率を制御する拡大撮影手段をさらに備えることを特徴とする請求項1に記載の撮像装置。

【請求項 6】

前記状態検出手段による前記第2の対象の状態情報の検出において必要とされる場合に、前記第2の撮像手段の撮像倍率を制御する拡大撮影手段をさらに備えることを特徴とする請求項1に記載の撮像装置。30

【請求項 7】

複数の前記第1の撮像手段を備え、前記第1の撮影制御手段は、前記予想位置情報をに基づいて、前記第1の対象の最適な画像を撮影することのできる前記第1の撮像手段を選択して前記第1の対象を撮影することを特徴とする請求項1に記載の撮像装置。

【請求項 8】

複数の前記第2の撮像手段を備えるとともに、前記第2の撮像手段により撮影された画像から前記第2の対象を検出する第2の対象検出手段と、該第2の対象検出手段による検出結果に基づいて、前記第2の対象の最適な画像を撮影することのできる前記第2の撮像手段を選択して前記第2の対象を撮影する第2の撮影制御手段とをさらに具備することを特徴とする請求項1に記載の撮像装置。40

【請求項 9】

前記状態検出手段により検出される前記第2の対象の状態情報は、前記第2の対象の移動方向の情報、第2の対象の顔が向いている方向の情報、及び第2の対象の視線の方向の情報の少なくとも一つを含むことを特徴とする請求項1に記載の撮像装置。

【請求項 10】

前記状態検出手段により検出される前記第2の対象の状態情報は、前記第2の対象の眼の開閉度の情報を含むことを特徴とする請求項1に記載の撮像装置。50

【請求項 1 1】

前記予想位置検出手段により検出される前記予想位置情報は、第1の対象が存在すると予想される方向の情報及び予想される距離の情報の少なくとも一つを含むことを特徴とする請求項1に記載の撮像装置。

【請求項 1 2】

前記第1の撮影制御手段は、前記状態検出手段により検出される前記第2の対象の状態情報に基づいて、前記第1の対象の画像の記憶手段への記録の開始および中断を行うよう制御することを特徴とする請求項1に記載の撮像装置。

【請求項 1 3】

第1の対象を撮影するための少なくとも1つの第1の撮像手段と、第2の対象を撮影するための少なくとも1つの第2の撮像手段とを備える撮像装置を制御する方法であつて、

前記第2の撮像手段により撮影された画像に基づいて、前記第2の対象の状態情報を検出する状態検出工程と、

前記状態検出工程において検出された前記第2の対象の状態情報に基づいて、前記第1の対象が存在すると予想される位置に関する情報である予想位置情報を検出する予想位置検出工程と、

前記予想位置検出工程において検出された前記予想位置情報に基づいて、前記第1の対象を撮影するように前記第1の撮像手段を制御する第1の撮影制御工程と、
を具備することを特徴とする撮像装置の制御方法。

【請求項 1 4】

前記予想位置検出工程において検出された前記予想位置情報に基づいて、前記第1の撮像手段の撮像光軸の方向を変更する撮像方向変更工程をさらに備えることを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項 1 5】

前記第2の撮像手段により撮影された画像から前記第2の対象を検出する第2の対象検出工程と、前記第2の対象検出工程における検出結果に基づいて、前記第2の撮像手段が前記第2の対象を撮像し続けるように、前記第2の撮像手段の撮像光軸の方向を変更する撮像方向変更工程をさらに備えることを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項 1 6】

前記予想位置検出工程において検出された前記予想位置情報に基づいて、前記第1の撮像手段の撮像倍率を制御する拡大撮影工程をさらに備えることを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項 1 7】

前記状態検出工程において検出された前記第2の対象の状態情報に基づいて、前記第1の撮像手段の撮像倍率を制御する拡大撮影工程をさらに備えることを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項 1 8】

前記状態検出工程における前記第2の対象の状態情報の検出において必要とされる場合に、前記第2の撮像手段の撮像倍率を制御する拡大撮影工程をさらに備えることを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項 1 9】

前記撮像装置は複数の前記第1の撮像手段を備え、前記第1の撮影制御工程では、前記予想位置情報に基づいて、前記第1の対象の最適な画像を撮影することのできる前記第1の撮像手段を選択して前記第1の対象を撮影することを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項 2 0】

前記撮像装置は複数の前記第2の撮像手段を備え、前記第2の撮像手段により撮影された画像から前記第2の対象を検出する第2の対象検出工程と、該第2の対象検出工程における検出結果に基づいて、前記第2の対象の最適な画像を撮影することのできる前記第2

10

20

30

40

50

の撮像手段を選択して前記第2の対象を撮影する第2の撮影制御工程とをさらに具備することを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項21】

前記状態検出工程において検出される前記第2の対象の状態情報は、前記第2の対象の移動方向の情報、第2の対象の顔が向いている方向の情報、及び第2の対象の視線の方向の情報の少なくとも一つを含むことを特徴とする請求項13に記載の撮像装置の制御方法。

【請求項22】

請求項13乃至21のいずれか1項に記載の制御方法をコンピュータに実行させることを特徴とするプログラム。

10

【請求項23】

請求項22に記載のプログラムを記憶したことを特徴とするコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、撮像装置の操作の煩わしさから使用者を解放し、撮影していることを意識しなくとも、所望する被写体の画像・映像を自動的に撮影記録可能とする技術に関するものである。

【背景技術】

20

【0002】

従来、被写体を静止画または動画にて撮影する撮像装置あるいはシステムにおいて、使用者（ユーザ）の負担を軽減するための提案がいくつかなされている。

【0003】

例えば特開平1-241511号公報（特許文献1）、特開2001-188164号公報（特許文献2）等には視線検出装置を備えたカメラが開示されている。これらのカメラにおいては、ファインダ部を覗き込んだ使用者の眼球部を、光分割器やハーフミラー等を介して光電変換素子により撮影し、画像処理を行うことによって瞳孔部の向いている方向すなわち、撮影者の視線方向を自動的に検知する。そして撮像画面中の視線方向に被写体が存在するものとみなし、オートフォーカス等の制御を行う構成となっている。

30

【0004】

特開平7-309577号公報（特許文献3）に開示されるクレーン用テレビモニタ装置においては、クレーンの先端に方向制御装置を介して取り付けられたテレビカメラによって撮影される画像を表示するテレビモニタを備える。そして、そのテレビモニタの前に座ったオペレーターに対し、その視線方向を検出する検出装置を取り付け、視線方向と対応する位置にテレビカメラの撮影方向を追尾させる駆動制御を行っている。

【0005】

また、特開平7-199278号公報（特許文献4）には撮影装置が開示されており、特開平7-209700号公報（特許文献5）には光学機器の立体リンク機構が開示されている。これらの装置は、カメラとカメラから分離した眼鏡型ファインダ装置を備え、眼鏡型ファインダの向いている方向を検知する。そして、検知した方向に一致するようにカメラの撮影方向を一致させる駆動機構を備えることによって、眼鏡型ファインダを装着した使用者が向いた方向を撮影可能としている。

40

【0006】

一方で、使用者を必要とせず、撮影する被写体を自動的に検知し、追尾撮影を行う撮像装置も存在し、監視カメラ等に用いられている。特許第3440916号公報（特許文献6）等に開示されている自動追尾装置においては、パン・チルトの2軸に旋回動作可能な駆動装置を備え、撮像する方向（光軸方向）を所望の方向に移動させることを可能としている。そして、撮像した画像信号より例えば人物等の移動物体を検出する追尾物体抽出手段を備え、抽出した追尾物体が常に撮像範囲内に収まるようにパン・チルト駆動制御を行

50

っている。

【0007】

また、パン・チルト駆動機構を備えないビデオカメラやデジタルカメラ等においても、撮像された画面内における移動物体を追尾し、撮像領域内で追尾視野領域を移動させるものも多数存在する。この追尾視野領域に被写体が存在するものと判断して、自動焦点検出（オートフォーカス）処理等を行う。このような技術は、特公平5-60712号公報（特許文献7）、特許第3449489号公報（特許文献8）等に開示されている。尚、追尾物体抽出処理としては、処理負荷の少なさから、異なった時点での2つ以上の撮像画像フレームの相關情報を利用しているものが多い。

【特許文献1】特開平1-241511号公報

10

【特許文献2】特開2001-188164号公報

【特許文献3】特開平7-309577号公報

【特許文献4】特開平7-199278号公報

【特許文献5】特開平7-209700号公報

【特許文献6】特許第3440916号公報

【特許文献7】特公平5-60712号公報

【特許文献8】特許第3449489号公報

【発明の開示】

【発明が解決しようとする課題】

【0008】

20

上述したような従来の撮像装置においては、以下に示すような問題点があった。

【0009】

特許文献1、特許文献2、特許文献3に開示されている構成では、使用者が明確に被写体を撮影しようとする意思を持っている必要がある。すなわち、特許文献1、特許文献2においては、撮像装置のファインダ部を覗き込む必要があり、特許文献3においては、テレビモニタの前に座る必要があって、使用者の自由を制限するものとなってしまっている。

【0010】

また、特許文献4および特許文献5に開示されている構成では、使用者は専用の眼鏡型ファインダ装置を装着する必要あって、やはり自由度を制限するものくなっている。

30

【0011】

一方で特許文献6、特許文献7、特許文献8に開示される、撮像された画像から自動的に追尾対象物体を抽出する構成では、撮影者の意図を全く組み入れることができないため、自動抽出された被写体が意図するものでないということが頻繁に有り得る。

【0012】

特にコンシューマ用途においては、例えば、運動会や学芸会のビデオ撮影や、旅行に行った際の風景写真など、その場においてはできれば撮影操作には専念したくはないが、そのとき注目していたもの等は記録保存しておきたい、という要求は非常に強い。

【0013】

従って、本発明は上述した課題に鑑みてなされたものであり、その目的は、使用者の自由度を飛躍的に向上させ、撮影操作に専念しなくとも所望の画像・映像を撮影し記録できるようにすることである。

【課題を解決するための手段】

【0014】

上述した課題を解決し、目的を達成するために、本発明に係わる撮像装置は、第1の対象を撮影するための少なくとも1つの第1の撮像手段と、第2の対象を撮影するための少なくとも1つの第2の撮像手段と、前記第2の撮像手段により撮影された画像に基づいて、前記第2の対象の状態情報を検出する状態検出手段と、前記状態検出手段により検出された前記第2の対象の状態情報を基づいて、前記第1の対象が存在すると予想される位置に関する情報である予想位置情報を検出する予想位置検出手段と、前記予想位置検出手段

40

50

により検出された前記予想位置情報に基づいて、前記第1の対象を撮影するように前記第1の撮像手段を制御する第1の撮影制御手段と、を具備することを特徴とする。

【0015】

また、本発明に係わる撮像装置の制御方法は、第1の対象を撮影するための少なくとも1つの第1の撮像手段と、第2の対象を撮影するための少なくとも1つの第2の撮像手段とを備える撮像装置を制御する方法であって、前記第2の撮像手段により撮影された画像に基づいて、前記第2の対象の状態情報を検出する状態検出工程と、前記状態検出工程において検出された前記第2の対象の状態情報を基づいて、前記第1の対象が存在すると予想される位置に関する情報を予想位置情報を検出する予想位置検出工程と、前記予想位置検出工程において検出された前記予想位置情報を基づいて、前記第1の対象を撮影するように前記第1の撮像手段を制御する第1の撮影制御工程と、を具備することを特徴とする。10

【0016】

また、本発明に係わるプログラムは、上記の制御方法をコンピュータに実行させることを特徴とする。

【0017】

また、本発明に係わる記憶媒体は、上記のプログラムを記憶したことを特徴とする。

【発明の効果】

【0018】

本発明によれば、使用者の自由度を飛躍的に向上させ、撮影操作に専念しなくとも所望の画像・映像を撮影し記録することが可能となる。20

【発明を実施するための最良の形態】

【0019】

以下、本発明の好適な実施形態について、添付図面を参照して詳細に説明する。

【0020】

(第1の実施形態)

図1は、本発明の第1の実施形態に係わる撮像装置のハードウェア構成を示すブロック図である。

【0021】

図1において、100は目的とする被写体130を光学的に撮影し光電変換して画像データ信号として出力するための第1のカメラであって、パン・チルト駆動機構を備えた第1の雲台101の上に設置されている。第1の雲台101を駆動させることにより、第1のカメラ100の撮像光軸140を水平方向(パン)150および垂直方向(チルト)151の2軸方向に回転駆動することが可能である。30

【0022】

103は、被写体撮影のための操作を行う人物すなわち使用者(ユーザ)131を撮影し同様に画像データとして出力するための第2のカメラである。第2のカメラ103も、第1のカメラ100と同様に、パン・チルト駆動機構を備えた第2の雲台102の上に設置され、撮像光軸141を2軸方向(152, 153)に回転駆動することが可能である。40

【0023】

尚、第1の雲台101および第2の雲台102のパン軸およびチルト軸の各駆動モータには、ステッピングモータや超音波モータ等が用いられ、後述するコントローラ部110から制御される。また、第1のカメラ100および第2のカメラ103はズーム機構を備え、必要に応じて対象を所望の倍率で拡大して撮影することが可能となっている。

【0024】

110は、本実施形態の撮像装置のコントローラ部であって、上記の第1のカメラ100、第1の雲台101、第2のカメラ103、第2の雲台102が接続されている。コントローラ部110の内部には、CPU112と、これに接続されたシステムバス111が配置されており、システムバス111に各装置や処理回路が接続され、CPU112で動50

作するプログラムで統括的に制御される。113はハードディスクドライブやフラッシュROMなどの不揮発性メモリ装置であって、CPU112で動作するプログラムや、各処理を行う上で必要となる固定値パラメータ等をデータとして格納している。

【0025】

114は揮発性メモリ(RAM)であって、CPU112がプログラムを実行する際に作業領域として使用する。また、一部領域を第1のカメラ100および第2のカメラ103にて撮影された画像データを一時的に格納するためのバッファ領域である画像バッファ115として使用している。尚、本実施形態では、同一のRAM装置をCPU112の作業領域と画像バッファ領域の両方に使用する構成としているが、これらを別々のRAM装置に分割する構成としてももちろん構わない。また、この画像バッファ115を、後述する被写体画像撮影制御部117や使用者追尾撮影制御部119あるいは物体検出処理部121に分散して内蔵する構成とすることも可能である。

【0026】

116はユーザI/F装置であって、不図示のディスプレイ装置とキー入力装置を備え、CPU112の指令に従って、ユーザに対して本装置の状態等の情報を通知したり、ユーザが撮像の開始・停止指令や、後述する使用者登録操作等を行う際に使用する。キー入力信号は、システムバス111を介してCPU112に伝えられ適切に処理される。

【0027】

117は被写体画像撮影制御部であって、第1のカメラ100に接続されている。また内蔵する第1の駆動制御部118は、第1の雲台101に接続されている。被写体画像撮影制御部117は、CPU112からの指令を受けて第1のカメラ100および第1の雲台101の光軸駆動および撮像制御を行い、光軸方向の画像を取得する。取得した画像信号はデジタル画像データとして、画像バッファ115に格納される。

【0028】

尚、画像バッファ115は複数枚の撮像範囲画像を格納することが可能な容量となっており、それぞれをフレームと呼ぶ。もちろん画像バッファ115は、被写体画像撮影制御部117からだけでなく、CPU112や後述する物体検出処理部121からもアクセス可能であり、これらの複数の装置からのアクセスは、バス・アービタによって排他制御される。

【0029】

118は、被写体画像撮影制御部117に内蔵される第1の駆動制御部であって、第1の雲台101のパン軸およびチルト軸駆動モータに接続されている。そして、不図示のパルス発生回路やドライバ回路などを内蔵し、被写体画像撮影制御部117を介したCPU112からの指令に基づいて、所望のパン・チルト駆動を行うことを可能とする。また、第1の雲台101に内蔵される不図示の位置センサ(エンコーダ等)の出力に基づき、現在のパンおよびチルト角の位置情報を取得し、被写体画像撮影制御部117を介してCPU112に通知することも可能である。

【0030】

尚、第1の駆動制御部118を本実施形態のようにコントローラ110に内蔵する構成とするのではなく、外部に独立した筐体として配置し、シリアルやネットワーク回線等の所定のI/Fを介してCPU112と通信する構成をとることもできる。

【0031】

119は使用者追尾撮影制御部であって、第2の駆動制御部120を備え、第2のカメラ103および第2の雲台102に対して、被写体画像撮影制御部117および第1の駆動制御部118と同様の処理を行う。

【0032】

121は画像データ中から、対象とする物体を検出する処理を行う物体検出処理部であって、画像処理ハードウェアにより構成されている。検出対象とする物体の特性に応じて、例えば人物を検出する処理回路など、最適の検出アルゴリズムを搭載する。本実施形態では、後述するように、人物の顔部を検出可能なニューラルネットワークアルゴリズムを

実行する処理回路を実装している。この回路には、不図示のレジスタやCPU112に接続された割り込み出力が備えられている。そして、CPU112から検出の開始指令を受け付け、検出完了を割り込みで通知し、検出結果を画像中の座標値としてレジスタに格納しCPU112に知らせる、といった機能を持つ。尚、もちろんCPU112の処理能力が十分にある場合には、CPU112で動作させるソフトウェア・プログラムとしてこの検出処理部を実装しても構わないし、DSPなどの比較的汎用的な画像処理用チップを用いて構成してももちろん構わない。また、他のアルゴリズムを用いて、顔でなく人物の全身、あるいは人物の他の一部を検出するようにしたり、これら複数の検出器を組み合わせるようにもしても、もちろん構わない。

【0033】

10

CPU122は画像記録用ディスクドライブであって、ドライブにセットされた画像記録用の光学ディスクメディアに、第1のカメラ100で撮影された被写体画像を記録する。記録の開始と停止は、CPU112からの指令に基づいて行われ、画像バッファ115に一時的に保存されている被写体画像データを読み出して記録する。尚、本実施形態の装置においては、画像データの記録は、動画も静止画も可能である。

【0034】

[撮影処理プロセス]

次に図2および図3、図4を用いて、本実施形態におけるCPU112で動作するプログラムの処理について説明する。

【0035】

20

本実施形態のプログラムは、図2に示すように大きく2つの処理プロセスにより構成されている。ここでプロセスとは、プログラム中の独立した並行処理体である。具体的には、オペレーティングシステム(OS)による1つのCPUでの時分割処理、あるいは複数のCPUでの分配処理により実現されている。

【0036】

本実施形態におけるプログラムにおいては、2つのプロセスがある。すなわち、使用者追尾撮影制御部119により使用者を追尾撮像し続ける処理を行う使用者追尾プロセス201と、被写体画像撮影制御部117により目的とする被写体を撮影し画像信号を記録する処理を行う被写体撮影プロセス202である。また、使用者追尾プロセス201で検出した使用者に関する情報203は、OSのプロセス間共有メモリ機能を用いて被写体撮影プロセス202に渡される。

30

【0037】

図3は使用者追尾プロセス201の処理の流れを説明するフローチャートである。

【0038】

このプロセスは、第2のカメラ103により撮影された画像から、撮影の操作を行うものとみなされる人物(使用者と呼ぶ)を抽出し、その挙動や状態を検出するとともに、使用者を常に撮影し続けられるように追尾制御を行うプロセスである。

【0039】

40

ステップS301は、初期設定処理である。この処理は、使用者が撮影開始前に上述したユーザI/F116を操作することによって行う、使用者に関する情報や撮影条件等を設定する処理である。

【0040】

尚、使用者に関する情報や撮影条件等の初期設定処理は、基本的に一度行っておけば不揮発性メモリ113に保存されるため、使用者の変更が無い限りは省略可能である。

【0041】

50

上述したように、本実施形態においては、物体検出処理部121として顔検出を行うニューラルネットワークを搭載しており、使用者の検出にもこれを用いている。人物の顔であることのみを条件とする使用者検出を行い、その後状態判別等を行わないならば、このステップにおいて使用者に関する情報として設定すべき項目は無い。しかし、顔から後述するような使用者状態の判別を行う場合や、検出した顔に対し使用者であるかどうかを判

別する個人認識を行う場合には、使用者の顔に関する基準情報を予め登録する必要がある。ここで基準情報とは、例えば正面を向いたときの両目間の間隔、正面と所定角度を向いたときの瞳孔の位置、通常状態の場合の眼球部の輝度分布や、使用者個人を識別するための特徴情報等である。

【0042】

ステップS302からステップS308までの処理は、撮影を開始した状態での繰り返し処理である。ユーザI/F116の操作によりステップS301の初期設定処理のステップが完了して撮影開始指令が出されると、使用者追尾プロセス201はこの繰り返し処理に移行する。以後使用者は撮影していることを殆ど意識することなく、自然に過ごすだけで、その時々に注目していた所望の被写体を自動的に記録させてゆくことができる。

10

【0043】

まずステップS302において、使用者追尾プロセス201は、第2のカメラ103による撮影画像を取得する。この画像データは、上述したとおり画像バッファ115に一旦格納される。

【0044】

次にステップS303において、画像バッファ115に格納された画像データに対し、物体検出処理部121による処理を実行し、画像中より使用者を検出する。本実施形態における物体検出処理である顔検出処理に関しては、後ほど詳細に説明する。

20

【0045】

ステップS304においては、使用者が見つかったかどうかを判断し、見つからなかつた場合にはステップS302の画像取得から繰り返す。見つかった場合にはステップS305へ進む。

30

【0046】

ステップS305においては、検出した使用者の位置や、興味を示していると考えられる方向の判別および使用者の状態判別を行う。興味を示していると考えられる方向（興味方向と呼ぶ）とは、顔の向いている方向や進んでいる方向、視線の向いている方向であって、その方向に所望の被写体が存在すると考えられる。興味方向のみでなく、興味のある距離も合わせて判別するようにすることも可能である。また、使用者の状態とは、例えば表情であって、注視していたり眠っていたりしているかどうかを判別する。これらの判別処理に関しては後述する。

30

【0047】

ステップS305で抽出した使用者に関する情報203は、ステップS306において共有メモリに書き出され、被写体撮影プロセス202に通知される。

40

【0048】

そして、ステップS307において、使用者の位置情報と、必要ならば興味を示している方向の情報を用いて、第2の駆動制御部120を介して第2の雲台102および第2のカメラ103のズーム機能を駆動し、常に使用者を最適に撮影できるように制御する。

【0049】

そしてステップS308において、検出した使用者情報を履歴としてRAM144の所定領域に記録し、ステップS302に戻って処理を繰り返す。

40

【0050】

次に図4は被写体撮影プロセス202の処理の流れを説明するフローチャートである。

【0051】

このプロセスは、使用者追尾プロセス201より通知される使用者情報203に基づき、第1のカメラ100および第1の雲台101を制御し、使用者が興味を示している被写体を撮影して画像データを記録する処理を行うプロセスである。

【0052】

まずステップS401において、必要に応じて対象被写体に関する情報の設定等の初期設定処理を行う。これは例えば、物体検出処理部121が複数種類の物体を検出可能である場合には、対象とすべき物体の種類を決定し、その物体を検出可能なようにパラメータ

50

を設定する処理である。本実施形態においては、物体検出処理部の対象物体は顔のみであるので、検出処理に関しては特にここで設定すべきパラメータはない。ここで設定するのは、大まかな画像中の顔の位置と倍率、すなわち構図に関するパラメータのみで、ステップ S 3 0 1 の設定と同様、使用者がユーザ I / F 1 1 6 を用いて予め設定する。

【 0 0 5 3 】

ステップ S 4 0 2 以下は、撮影開始後の定常状態である。ステップ S 4 0 2 の処理は、共有メモリより使用者情報 2 0 3 を読み出す処理である。

【 0 0 5 4 】

そしてステップ S 4 0 3 において、これらの情報が前回から変化しているかどうかを確認する。尚、撮影開始直後の初回時においては、無条件に変化していることとなり、ステップ S 4 0 4 へと進む。変化していない場合には、ステップ S 4 0 9 へと進み、前回と同条件のまま被写体撮影を継続する。

【 0 0 5 5 】

初回あるいは使用者情報が変化すると、ステップ S 4 0 4 へ進む。ステップ S 4 0 4 では、まず使用者の位置とステップ S 3 0 5 で抽出した興味方向・距離の変化をチェックする。

【 0 0 5 6 】

変化した場合には、これらの情報に基づきステップ S 4 0 5 の被写体発見処理を行う。これについては後述する。変化していなければ、次の処理へと進む。

【 0 0 5 7 】

次に使用者の状態の変化をチェックする。例えば、白目部分を含む目の大きさにより開き度合いを判別し、通常の大きさに比べて目が大きくなった場合には、目を見開くような表情すなわち被写体に注目したものと判断し、被写体撮影のためのズーム倍率設定値を上げる。目が半分閉じたような状態となった場合には、被写体にあまり注目していないものと判断し、ズーム倍率をワイド側に切り替える。目を閉じた状態が一定時間続いていると判断される場合には、使用者が眠ってしまったものと判断し、被写体画像のメディアへの記録を一時的に停止するか、もしくは直前に見ていた被写体をロックし、その被写体を追尾して撮影し続けるようにモードを変更する。尚、目の開閉度の判別方法については後述する。

【 0 0 5 8 】

そしてステップ S 4 0 8 において、これらの使用した使用者情報を履歴として使用可能のように記録すると共に、ステップ S 4 0 9 においてステップ S 4 0 7 での設定に基づいて被写体撮影を行い、ステップ S 4 0 2 の処理に戻る。

【 0 0 5 9 】

[顔検出]

対象物体検出方法の一例として、本実施形態で使用している顔検出処理方法について、図 5、図 6、図 7 を用いて説明する。尚、本実施形態で採用している顔検出技術は、特開 2 0 0 4 - 1 9 9 2 0 号公報に開示されているものをベースとしているが、別の顔検出技術を用いてももちろんかまわない。

【 0 0 6 0 】

図 5 は、検出すべき人物の顔を模式的に表した図である。人物の顔は、目や口、鼻といったいくつかの特徴点が、およそ定まった位置に配置されている。本実施形態は、このうち特に目と口の特徴点に注目し、局所的な低次元特徴から、階層的により高次元の特徴抽出を行うことによって、画像中の顔検出処理を行うものである。

【 0 0 6 1 】

具体的な階層特徴は、図 6 に示すようになっている。各々の四角形の大きさは入力画像信号のある 1 点に注目したとき、その周辺部分を切り出して入力する範囲を示す（受容野）。

【 0 0 6 2 】

図 6 の (a) は、もっとも原始的な微少な受容野範囲における特徴を示している。上か

10

20

30

40

50

ら順に、縦方向特徴、横方向特徴、右上がり斜め方向特徴、左上がり斜め方向特徴となっている。これは例えばガボールウェーブレット変換等のフィルタによって得られる微小線分の方向と考えればよく、画像中のどの位置にこれらそれぞれの特徴が存在するかを抽出するのが第1次特徴抽出処理である。もちろんこの4種類だけでなく、必要に応じて同程度のサイズの受容野の特徴を追加することもできる。

【0063】

これらの第1次特徴抽出処理の出力を用いて、第2次特徴の抽出を行う。第2次特徴の例は図6(b)に示すように、第1次特徴を複数集め特定の空間的配置に並べたものとなっている。上から順に右開きVエッジ、左開きVエッジ、水平エッジ、垂直エッジとなっている。この他、色成分を含む特徴としてもよい。

10

【0064】

さらに、第2次特徴抽出出力を同様に用いて、第3次特徴を抽出する。本実施形態では、図6(c)に示すように、第3次特徴が先に述べた注目すべき特徴点であって、目および口としている。

【0065】

同様にして、この第3次特徴点が、おおよそ図6(d)に示すような配置にあるとき、これを第4次特徴すなわち顔として判定するのである。つまり、顔検出処理部としては、ある程度の位置ロバスト性を持たせつつ、第1次から第4次までの特徴抽出処理を階層的に行うことができればよい。

【0066】

この処理は、例えば、図7に示すConvolutional Neural Network構造(LeCun, Y. and Bengio, Y., 1995, "Convolutional Networks for Images Speech, and Time Series" in Handbook of Brain Theory and Neural Networks (M. Arbib, Ed.), MIT Press, pp. 255-258)により実現できる。その最終層の出力は、認識された検出対象物体の入力データ上の位置情報である。

20

【0067】

図7において、801はデータ入力層であって、画像バッファ中に格納されている各カメラで撮影された画像データを入力する層である。

【0068】

最初の特徴検出層S1(802)は、局所的な低次の特徴である第1次特徴を抽出するフィルタとして機能するようにパラメータ設定されている。図6の(a)に示した第1次特徴1つにつき、これを検出するための細胞平面が1つ割り当てられている。これにより、入力画像面の各位置を中心とする局所領域における複数種類の特徴を抽出することが可能である。尚、各細胞平面は、入力画像信号の画素数と等しいかこれを所定レートで間引いた数に等しい数のニューロンにより構成されている。これは他の層に関しても同様である。

30

【0069】

続く特徴統合層C1(803)は、S1層の各特徴検出のための細胞平面毎に、同サイズの細胞平面を1つずつ用意しており、所定の受容野構造をもって、それぞれの出力を統合する。これにより、各特徴についてある程度の位置ロバスト性を持たせることを可能としている。

40

【0070】

以下、Sn層とCn層を最終出力層であるN層まで、 $n = 1, 2, \dots, N$ の順に並べたネットワーク構造となっており、各々のSn層の各特徴抽出平面への入力は、 $n = 1$ を除いて、その前のCn-1層の全特徴統合平面の出力を用いるものとなっている。このような構造により、低次の特徴を順次検出・統合していき、最終層の出力として、所望の検出物体すなわち本実施形態の撮像装置においては、顔の存在する位置を出力することが可能となる。

【0071】

ここで、特徴検出層における各細胞平面を構成するニューロンの入出力特性は、次のよ

50

うに表される。第1層目の特徴検出層の第k番目の特徴を検出する細胞平面の位置nにあるニューロンの出力を $usl(n, k)$ とすると、 $usl(n, k)$ は次式で表わされる。

【0072】

【数1】

$$usl(n, k) = f \left(\sum_{\kappa}^{K_{d-1}} \sum_{v \in W_i} w_i(v, \kappa, k) \cdot u_{Cl-1}(n+v, k) \right)$$

【0073】

ここで $ucl(n, k)$ は、第1層目の特徴統合層の第k番目の細胞平面の位置nにあるニューロンの出力を示す。 Kcl は、第1層目の特徴統合層の細胞平面の数であって、同一層の特徴検出層の細胞平面の数すなわち検出する特徴の種類の数に等しい。 $w1(, , k)$ は、第1層目の特徴検出層の第k番目の細胞平面の位置nにあるニューロンの、第1-1層目の特徴統合層の第1番目の細胞平面の位置 $n + |$ にあるニューロンからの、入力結合パラメータである。また、 $W1$ は、検出細胞の受容野を示し、その大きさは有限である。

【0074】

受容野 $W1$ の各位置からの結合加重値 $w1$ を適切に設定してやることにより、所望の特徴を検出することができる。この設定は、図6に示した各層の各特徴ごとに、十分な数のサンプル（教師データ）を用意し、特徴平面毎に例えば誤差逆伝播法などの学習アルゴリズムを用いて設定すればよい。本実施形態では、左右開きのVエッジや、目および口、顔のサンプルを多数用意し、これを学習させることによって、結合加重パラメータを定めた。

【0075】

また、上記式において、 $f(x)$ は積和演算結果に対しての非線形処理を示し、例えば次式で表わされるロジスティック関数を使用する。

【0076】

$$f(x) = 1 / (1 + e^{-x})$$

次に、第1層目の特徴統合層の第k番目の細胞平面の位置nにあるニューロンの出力を $ucl(n, k)$ とすると、 $ucl(n, k)$ は次式で表わされる。

【0077】

【数2】

$$u_{Cl}(n, k) = \sum_{v \in D_i} d_i(v) \cdot u_{Sl-1}(n+v, k)$$

【0078】

この式において、 $d1()$ は第1層目の特徴検出層（S層）ニューロンから、同一特徴に対する第1層目の特徴統合層（C層）の細胞平面のニューロンへの入力結合であり、 $|$ に関して単純に減少する関数で設定されている。また、 $D1$ は統合細胞の受容野を示す。 $d1()$ および $D1$ を適切に調整することにより、前層の特徴に対する位置ロバスト性を調整することが可能である。

【0079】

以上Convolutional Neural Network構造により、第4次特徴として、画像中の人物の顔がある位置を検出する検出器が構成される。また、顔検出位置を基準に所定の範囲の第3次特徴を調べることにより、目および口特徴点の位置も特定することが可能となる。この構成によると、例えば目特徴の検出は、第2次特徴であるVエッジおよび横方向エッジのおおよその位置に基づいて検出されるので、目を瞑るなど少々の形の変形は関係なく検出を行うことが可能である。

【0080】

尚、本実施形態においては、このConvolutional Neural Network検出処理を専用のハ

10

20

30

40

50

ドウェア回路を用いて高速に行っている。しかし、CPUが十分な処理能力を持つものであれば、ソフトウェアプログラムにより、同一のアルゴリズムを実行するようにすることも、もちろん可能である。

【0081】

[使用者追尾と移動方向の検出]

使用者の興味方向として、フレーム間画像より移動量および移動方向を検出する方法と、これに基づき使用者を追尾撮像するためのパン・チルト駆動量の計算方法について、図8を用いて説明する。この処理は、図3におけるステップS305の興味方向判別処理と、ステップS307における追尾処理の具体的な一例である。

【0082】

まず、図8(a)は、対象物体検出処理部により検出された使用者の画像中の位置を表している。本実施形態においては、ステップS307の追尾処理は、撮像画像の中心位置に使用者がくるようにパン・チルト制御を行うものであるので、図8(a)の中心位置黒点は、前フレームにおける使用者の存在していた位置とみなせる。従って、黒点位置から現在位置を指すべきトルD_xを考えると、D_xは前フレームから今回のフレームが撮影されるまでに、使用者が動いた方向および距離の、第2のカメラ103のセンサ平面への射影を表していることになる。なお、ここでD_xはベクトルを表わすものとする。

【0083】

使用者を追尾し続けるためには、次のフレームの撮影までに、D_xの位置が画像中心となるようにパン・チルト駆動してやればよい。

【0084】

このとき、D_xの第2のカメラ103のセンサ平面の水平方向の成分をD_{x'}とすると、図8(b)のように表せる。ただし、ここで第2のカメラ103はレンズ中心をPの位置とし、Pがパン軸と一致するように回転駆動するものとする。カメラの撮像画像中心すなわち光軸方向はC'の方向に一致する。D_{x'}は実空間における検出物体の移動量のセンサ平面の水平方向成分であるが、全フレームにおける検出物体までの距離Zが不定の状態ではD_{x'}も定まらない。しかしそのセンサ平面上への射影であるD_xに関しては、センサ平面上で中心Cからの画素数で表すことが可能である。尚、S_xはセンサの水平方向の画素数、xは同画角であって、第2のカメラ103に付随するパラメータである。

【0085】

今、レンズ中心からセンサ面までの距離(焦点距離)をfとすると、fは次のように表わされる。

【0086】

$$f = S_x / \{ 2 \tan(x/2) \}$$

従って、パン方向の駆動量xは、次のように求めることができる。

【0087】

【数3】

$$\theta_x = \tan^{-1} \left(\frac{|\vec{D}_x|}{f} \right) = \tan^{-1} \left(\frac{2|\vec{D}_x| \tan(\alpha_x/2)}{S_x} \right)$$

【0088】

同様に、センサ垂直方向成分D_yを用いてチルト方向駆動量yを求め、この駆動量をもってステップS307における追尾処理の第2の雲台への駆動指令量とする。これにより、常に使用者を画面の中心付近に捕え続けるように、追尾制御を行うことができる。

【0089】

尚、過去の数フレームにおけるD_xを履歴として記憶しておくことにより、使用者の動きを予測してより精度良く追尾するようにしてもよい。

10

20

30

40

50

【0090】

また、対象物体検出処理部の出力結果に基づき、その時点での対象物体の大きさをパラメータとしてこの変化量を調べることにより、奥行き方向すなわちカメラ光軸方向の移動量 D_z および対象物体までの距離 Z を推測することも可能である。すなわち、図 8 (c) に示すように第 2 のカメラ 103 の光軸方向を基準として、実空間における対象物体の移動方向を推測することが可能なのは自明である。

【0091】

[被写体発見処理]

使用者の位置と興味方向情報に基づき、図 4 のステップ S405 の被写体発見処理を行う。この処理の一例を、図 9 を用いて説明する。

10

【0092】

本実施形態の一例としては、被写体撮影用の第 1 のカメラ 100 と使用者撮影用の第 2 のカメラ 103 は、極めて近い位置に設置されている。例えばコントローラ部 110 を含めて一体化されているものとする。

【0093】

コントローラ部 110 では、第 2 の雲台 102 の制御位置情報を知ることができるので、これと、検出された使用者の位置情報に基づき、第 1 のカメラ 100 が第 2 のカメラ 103 と同じ方向に使用者を撮影するように制御する。

【0094】

そして、興味方向を使用者に近い位置から順次撮影（スキャン駆動撮影）できるように、第 1 の雲台 101 を一定速度で動かす。ここで、一定速度とは、使用者の移動速度より速く、被写体を所望の時間で発見できる程度の速度である。尚、上述した移動方向の検出処理での奥行き方向のように、第 2 のカメラ 103 のセンサ面水平および垂直軸方向の情報だけでなく、光軸方向の情報も使用者の興味方向として推定できる場合もある。その場合には、この情報に基づき第 1 のカメラ 100 が注目すべき位置 = スキャン駆動撮影する際に被写体が存在するかもしれない位置までの距離情報を算出する。そして、この距離情報に基づいて、フォーカスや、対象とする被写体を最適の構図に捉えられるズーム倍率を適時設定して、撮影するようとする。

20

【0095】

被写体を撮影する第 1 のカメラ 100 が撮影する画像は、順次、対象物体検出部 121 で処理される。対象物体が見つかったとき、これを被写体とみなして、スキャン駆動を停止して、その物体の画像の記録を開始する。

30

【0096】

尚、本実施形態では、同一の顔検出手段を使用者検出処理にも被写体発見処理にも用いているようにしているが、複数の検出手段を備え、それぞれ別の検出手段を用いるようにしてももちろんかまわない。あるいは状況によって切り替えるようにしてももちろんかまわない。また、本被写体発見処理は、必ずしも毎回行う必要は無い。ステップ S404 で使用者の位置と興味方向が、今まで被写体を撮影していた条件から大きく変化したときのみを行うことによって、スキャン駆動撮影の回数を減らすこともできる。スキャン駆動撮影をしない場合には、第 1 のカメラ 100 の映像のみから被写体を検出して追尾撮像すればよい。

40

【0097】

[目の開閉度による使用者状態判別]

図 3 のステップ S305 における使用者の状態判別の一例として、目の開閉度合に基づく表情判別方法について、図 10 を用いて説明する。

【0098】

図 10 の上段の 3 つの図は、物体検出処理部 121 で検出された顔の目特徴点付近を切り出した画像を示している。切り出し範囲は、水平方向は目の両端左右開き V エッジ特徴点の間で、垂直方向は最大限目を開いたときに白目部分が全て含まれるような一定の値である。尚、この処理に先立ち、切り出し元となる画像は、目の両端の V エッジ間距離で正

50

規化してあるため、水平方向サイズも常に一定である。

【0099】

図10(a)の上段の図は、ほぼ目を閉じている状態のときの切り出し画像である。このとき、切り出し領域内の輝度ヒストグラムを取ると、図10(a)の下段の図のようになる。ここで横軸は輝度であり、縦軸は画素数である。

【0100】

同様に、通常状態の目の画像を切り出したものが図10(b)の上段の図で、その輝度ヒストグラムが図10(b)の下段の図である。さらに、目を大きく見開いたとき(すなわち注目しているとき)の画像が図10(c)の上段の図で、その輝度ヒストグラムが図10(c)の下段の図である。

【0101】

輝度ヒストグラムを見て分かることおり、特に肌色付近の輝度の部分に、目の開閉度合の特徴が現れている。従って、本実施形態では、肌色付近の画素数を2つの閾値で区別することにより、目の開閉具合を「目を瞑っている」、「通常」、「注目」の3段階に判別する処理を行っている。

【0102】

尚、画像全体の明るさの変化等に対応するために、これら3つのヒストグラムの形状をテンプレートとして記憶しておく、どれに近いかを判定することによって、開閉度合を判定する処理とすることも可能である。

【0103】

また、ここで述べた目の開閉度合のみでなく、他の画像情報を組み合わせて使用者の状態判別を行なってももちろんよい。他の情報の例としては、顔色、特徴点間の距離の変化による表情認識情報などがある。

【0104】

これらの判別した使用者状態は、共有メモリを介して被写体撮影プロセス202に伝えられる。また、「目を瞑っている」が一定時間以上続いた場合には、眠ってしまったものと判断している。使用者状態が眠っていると伝えられた被写体撮影プロセス202は、ステップS407の処理において被写体記録を一次停止する、あるいは直前まで撮影していた被写体を撮り続けるように撮影条件を変更する。

【0105】

[第2の実施形態]

[顔方向検出]

使用者の興味方向の抽出の別の一例として、使用者の顔の向いている方向を推定する手段を用いることも可能である。

【0106】

使用者の顔の向いている方向を推定する手段としては、例えば「頑健性と精緻性を備えた顔特徴点追跡による顔方向推定(電子情報通信学会論文誌2001/8 Vol. J84-D-I No.8 pp.1762-1771)」に紹介されている手法を応用することができる。但し、本実施形態では、上記の文献の鼻特徴点の代わりに口特徴点を用いている。

【0107】

まず使用者は、予め使用者を撮像する第2のカメラ103に正対し、正面顔画像を撮影して、その特徴点位置を登録しておくものとする。これは図3におけるステップS301の処理の一部に相当する。ここで、右目、左目、口の符号をそれぞれ順に、0, 1, 2とし、正面顔画像上でその特徴点の位置をx0, x1, x2とする。これを図11の(a)に示す。

【0108】

そして、ある時点における撮影者顔画像の各特徴点位置をx0', x1', x2'とする。これは例えば、図11の(b)や(c)に示される目および口特徴点の位置である。このとき、動きベクトルIを以下の通り定義する。

【0109】

10

20

30

40

50

$$| = x | - x |'$$

並進成分を除くこの式の、点 x_0 を中心とする動きベクトル場は、アフィン変換行列と相対位置 $x_1 - x_0$ の積によって与えられるとする。この動きベクトルが拡大される量を div とすれば、 div はアフィン変換行列の対角成分の和によって与えられる (R.Cipolla, Active Visual Inference of Surface Shape, Springer, 1992)。この div を用いることにより、基準正面顔 (顔特徴点位置 x_0, x_1, x_2) の画像に対して、特徴点位置が x_0', x_1', x_2' である画像の顔部 (頭部) の大きさを正規化することができる。もちろんこの正規化は、特徴点付近の顔の存在する領域のみに対して行えばよい。

【0110】

10

ここで、基準正面顔における両目間距離を d_{eye} 、特徴点位置 x_0', x_1', x_2' のときの顔画像を div で正規化した両目間距離を d'_{eye} とする (図 11 (a) および (b) 参照)。図 11 (e) の左図に示すように、縦軸 (背骨方向の軸) 方向の顔の回転量 (左右回転量) を θ とすると、以下のようになる。

【0111】

$$\theta = \cos^{-1}(d'_{\text{eye}} / d_{\text{eye}})$$

同様に、口特徴点から両目を結ぶ線分へ降ろした垂線の足の長さをそれぞれ d_{mouth} , d'_{mouth} として (図 11 (a) および (b) 参照)、横軸 (左右の耳の穴を結ぶ方向の軸) 方向の顔の回転量 (上下回転量) を ϕ とすれば (図 11 (e) の右図参照)、次式を得ることができる。

20

【0112】

$$\phi = \cos^{-1}(d'_{\text{mouth}} / d_{\text{mouth}})$$

また、これに加えて、両目間を結ぶ線分が、例えば、図 11 の (c) に示すように、水平方向線分と成す角度 α も顔方向を表す一つの角度パラメータである。ここで、基準正面顔の両目間線分は撮影画像の水平方向と一致するように記録している。図 11 の (c) のように、 α が 0 でない顔を扱う場合には、先に線分 d'_{eye} と水平方向 (または線分 d_{eye}) とを比較して α の値を求め、(b) の画像となるように回転補正してから θ , ϕ を求めようすればよい。

【0113】

30

ところで、 θ , ϕ に関して上記で求まるのは、あくまで回転「量」であって、これだけでは左右 (または上下) どちらの方向に回転したのかは分からぬ。これを判別するためには、特徴点が頭部輪郭中のどちらに寄っているかを判定する必要がある。すなわち、まず頭部輪郭の位置を得ることが必要となる。この頭部輪郭抽出処理を、図 12 を用いて説明する。

【0114】

図 12 (a) は、頭部輪郭を抽出すべき顔画像の一例である。この顔画像をエッジ抽出処理したもの (b) に示している。尚このエッジ抽出処理は、側抑制構造のエッジ抽出フィルタによって処理しても良いし、先に説明した顔検出処理部のニューラルネットワークの低次特徴出力を統合して生成しても良い。

【0115】

40

図 12 の (b) に示すように、頭部輪郭は、ほぼ円形に近いものとみなせるので、エッジ抽出フィルタによって得られたエッジ画像を、図 12 (c) に示すような、ある程度幅を持たせた円形テンプレートでマッチングする。これにより、その位置を特定することができる。

【0116】

通常、この方法で撮影された画像から直接頭部輪郭を抽出しようとすると、サイズの異なる多数のテンプレートを用意して、これら全てに対して計算を行う必要があるため、その計算コストは膨大なものとなり、また誤検出も非常に多くなってしまう。しかし、本実施形態においては、上述したとおり、別の検出手段により顔の位置はほぼ分かっており、その大きさも正規化されているので、一つのテンプレートで画像の一部領域に探索領域を

50

絞って計算することができる。そのため、極めて少ない計算コストでマッチングを取ることが可能である。

【0117】

もちろん頭部輪郭抽出はこの方法に限らず、他の方法、例えば「HeadFinder: フレーム間差分をベースにした人物追跡 馬場 功淳, 大橋 健, 乃万 司, 松尾 英明, 江島 俊朗 (画像センシングシンポジウム 2000, pp329-334, 2000)」に見られるようなフレーム間差分を利用する方法を用いても良い。また、肌色や髪の色を利用する方法も考えられる。

【0118】

検出された頭部輪郭を簡単のため真円で表したとき、図11の(a)の基準正面顔に対し、図11の(b)の特徴点は全体的に右上によっている。顔輪郭の中心位置を原点として x_0 , x_1 , x_2 の位置ベクトルを表すとき、 x_0' , x_1' , x_2' を同様の位置ベクトルで表して x_0 , x_1 , x_2 からの変化の方向を調べれば、およびの符号を決定できる。

【0119】

以上のように , , とその符号が求まれば、これらの値を変換処理して移動方向の抽出時と同様に使用者の興味方向を求めることができる。すなわち、図11(d)に示すように、第2のカメラ103のセンサ平面水平方向をx軸、垂直方向をy軸、奥行き方向をz軸とする座標系で、使用者の興味方向を求めることができる。

【0120】

もちろん、この顔方向検出処理だけを使用者の興味方向の検出処理として用いても良いが、第1の実施形態で説明した移動方向抽出処理を組み合わせて、双方の情報より興味方向を決定する処理としてもよいのは言うまでもない。

【0121】

このようにして決定した興味方向を第1の実施形態と同様、被写体の存在する方向と見なし、図9で説明したのと同様のスキャン駆動撮影を行う。以後の処理は第1の実施形態と同様である。

【0122】

[第3の実施の形態]

[視線方向の検出]

使用者の興味方向の抽出のまた別の一例として、使用者の眼球部の画像を処理して、視線の方向を推定する手段を用いることも可能である。尚、上述したとおり顔検出処理において目特徴点の位置は判明しているので、この情報に基づき、眼球部をズーム撮影してより詳細な眼部の画像情報を得るようにしてもらおう。

【0123】

図13は、本実施形態における視線検出処理を説明する図である。この視線方向の検出方法は、目特徴の両端の左右開きVエッジ間ににおける瞳孔の位置に基づく。瞳孔の詳細な位置検出には、顔検出で用いたのと同様の階層的な特徴抽出アルゴリズムを用いている。

【0124】

ステップS301の処理の一部として、基準正面注視眼画像における左右Vエッジ間ににおける瞳孔の位置を予め登録しておく。このときの眼部の画像の一例としては、図13の(a)の下段に示したものとなる。また、ユーザI/Fのディスプレイ部を用いるなどして、正面より予め定めてある所定角度 の方向を向いた眼部画像における瞳孔の位置も登録しておく。このときの眼部の画像の一例としては、図13の(b)の下段に示したものとなる。

【0125】

また、眼球は図13の上段に示すように、半径rの球形としてモデル化する。

【0126】

視線方向が所定角度 の方向を向いたときの瞳孔位置の、基準正面注視位置からの水平方向のずれ量をx とすると、

10

20

30

40

50

$$r = x / \sin t$$

となる。そのため、ある時点における瞳孔位置の基準正面注視位置からの水平方向のずれ量を x_t とすれば、このときの視線方向 t は、次のように求めることができる。

【0127】

$$t = \sin^{-1} (x_t / r) = \sin^{-1} \{ x_t / (x / \sin t) \}$$

垂直方向の視線角度に関して同様に求めることができる。第2の実施形態で説明した顔方向検出処理に、ここで求めた視線角度を加えることにより、より正確に、使用者の見ている方向を推測することが可能となる。

【0128】

また、図14に示すとおり、両眼それぞれの視線方向を求め、その交差する地点を割り出すことにより、被写体までの距離 Z を推測することも可能である。左右の目のそれぞれの視線方向を r, l 、両眼距離を d_{eye} とすれば、距離 Z は、

$$Z = d_{eye} / (\tan r + \tan l)$$

として求めることができる。

【0129】

尚、眼球の動きは顔の向きの変化に比べると激しいので、ある程度の時間間隔で、平均化した視線方向を用いるようにする。

【0130】

また、第2の実施形態と同様、このようにして検出した情報を、顔方向や移動方向等の他の興味方向および距離の検出手段と組み合わせても良いことは自明である。

【0131】

このようにして決定した興味方向を第1および第2の実施形態と同様、被写体の存在する方向と見なし、図9で説明したのと同様のスキャン駆動撮影を行う。以後の処理は第1の実施形態と同様である。

【0132】

[第4の実施形態]

図15は、第4の実施形態に係わる撮像装置のハードウェア構成を示すブロック図である。本実施形態では、被写体撮影用の第1のカメラおよび使用者撮影用の第2のカメラを、第1の実施の形態のようにそれぞれ1台のみとするのではなく、どちらか一方あるいは両方を複数台にて構成する。

【0133】

図15の510～522は、第1の実施形態を示す図1の110～122の百の桁を除いて対応する部位と同様の処理部あるいは装置であるため、詳細な説明は省略する。ただし、図1の118および120に相当する駆動制御部は、図15上では記載していない。もし各カメラをそれぞれ1つの雲台に搭載したとすると、駆動制御部は雲台毎に1つずつ必要となる。これらは、被写体画像撮影制御部517もしくは使用者追尾撮影制御部内519に搭載されることになる。しかしながら、各カメラを雲台には取り付けずに撮像光軸方向固定の構成とすることも可能である。

【0134】

図15において、500, 501, 502は、被写体を撮影するための3台の第1のカメラを示している。それぞれのカメラの撮影光軸方向は異なって設置されており、被写体が存在すると想定される範囲を全て撮影可能に配置されている。

【0135】

また、504, 505, 506は、使用者を追尾撮影するための3台の第2のカメラを示している。この図の例では、使用者が移動すると想定される範囲の周りを取り囲むように配置している。

【0136】

もちろん、これらの複数の第1のカメラおよび第2のカメラを、空間的に離れた位置に配置した撮像システムとしてもかまわない。

【0137】

10

20

30

40

50

図15の撮像装置のプログラムの処理フローは図2で説明した第1の実施形態の処理フローと同様である。

【0138】

複数の使用者撮像用の第2のカメラ504～506を、使用者の行動範囲の周りを取り囲むように配置すれば、常に使用者のほぼ正面の顔画像を撮り続けるようにすることができる。従って、第1の実施形態で説明したように、物体検出処理部に人物の正面顔検出処理のみを用いていたとしても、すべてのカメラ映像に対して検出処理を行うことにより、使用者を見失う可能性は低くなる。

【0139】

このように図3におけるステップS307の追尾処理相当の処理を、最適に配置された複数のカメラで行うことにより、パン・チルト駆動機構を省略することも可能となる。さらに、上述した顔方向検出処理を行い、カメラの撮影方向情報と合わせて、使用者の興味方向や距離を判別することももちろん可能である。もちろん、一部のカメラにあえてパン・チルト駆動機構を備える構成にしてもよいし、ズーム機構を備えて、前述したのと同様、眼部の拡大撮影を行って視線方向や開閉度合の判別をし易くしてもよい。

【0140】

また、撮影方向の異なる被写体撮影用の第1のカメラを複数台設置した場合、使用者の意図する方向より特定される被写体を、最も望ましい構図で撮影することのできるカメラが選択されることになる。このとき図4のステップS405の被写体発見処理に相当する処理は、複数の第1のカメラの内、まず使用者の存在する地点に最も近い位置を撮影可能な1台を選択して撮影を開始する。そして、そこに被写体が存在しなければ、使用者の興味方向の次に近い位置を撮影可能なカメラで撮影する、というように、順次撮影者から遠ざかる位置を撮影可能なカメラを選択してゆく処理を行う。従って、個々のカメラにパン・チルト駆動機構が無くとも、あるいはパン・チルト駆動機構はあってもその駆動範囲が限られていたとしても、複数台を最適に配置することにより、第1の実施形態と同様の処理を行うことが可能となる。そして、使用者の所望する被写体画像を得ることができる。

【0141】

第1の実施形態をとるか第4の実施形態とするかは、撮像装置あるいはシステムが対象とする被写体や使用者の動き回る範囲やコスト等の要因に基づき、より適した方を選べば良いのは明らかである。

【0142】

以上述べてきたように、上記の実施形態によれば、撮影の操作担当者（使用者）は、必要な初期設定を行うのみで、以後はほとんど撮影を意識することなく過ごしているだけでも、所望の画像・映像を撮影し記録してゆくことが可能になる。これにより、使用者の自由度は飛躍的に向上する。また、例えば、特定の客や見学者を使用者と見なすことにより、その人物そのものの映像ではなく、興味を示した物体等を記録することができるため、店舗や展示会等での調査目的に応用することができる。

【0143】

もちろん本発明は上述した実施形態の構成のみに限ったものではなく、例えば第1のカメラや第2のカメラのどちらか一方あるいは両方を、1台のカメラであるが全方位カメラを用いるようにしても良い。また、第1のカメラと第2のカメラを兼用とする構成にして、複数台のカメラを状況に応じて使用者撮影用と被写体撮影用に切り替えて使用する構成としてもよいことは言うまでもない。

【0144】

（他の実施形態）

また、各実施形態の目的は、次のような方法によっても達成される。すなわち、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給する。そして、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行する。この場合、記憶媒体から読み出されたプログラムコード自体が前述した

10

20

30

40

50

実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、本発明には次のような場合も含まれる。すなわち、プログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(ΟＳ)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される。

【0145】

さらに、次のような場合も本発明に含まれる。すなわち、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれる。その後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される。10

【0146】

本発明を上記記憶媒体に適用する場合、その記憶媒体には、先に説明した手順に対応するプログラムコードが格納されることになる。

【図面の簡単な説明】

【0147】

【図1】本発明の第1の実施形態に係わる撮像装置のハードウェア構成を示すブロック図である。

【図2】本発明の第1の実施形態に係わる撮像装置のCPUで動作させる制御プログラムの基本的な処理プロセス構成を示すブロック図である。20

【図3】本発明の第1の実施形態に係わる撮像装置における制御プログラムの使用者追尾プロセスの流れを説明するフローチャートである。

【図4】本発明の第1の実施形態に係わる撮像装置における制御プログラムの被写体撮影プロセスの流れを説明するフローチャートである。

【図5】本発明の第1の実施形態に係わる撮像処理において、物体検出処理の一例として行う顔検出処理で検出される顔画像の一例を示す図である。

【図6】本発明の第1の実施形態に係わる撮像処理において、物体検出処理の一例として行う顔検出処理で用いている低次から高次の特徴を示す図である。

【図7】本発明の第1の実施形態に係わる撮像処理において、物体検出処理として行う顔検出処理を実行するConvolutional Neural Network構造を示す図である。30

【図8】物体検出処理結果に基づくパン・チルト駆動を行う駆動量の算出方法を説明し、かつ使用者の移動方向の推定方法を説明するための図である。

【図9】被写体の探索動作の一例を説明する図である。

【図10】本発明の第1の実施形態に係わる撮像処理において、使用者の状態判別のための一手法として導入している、眼球部の開閉度を判定する判定方法の一例を説明する図である。

【図11】検出された顔特徴点より、顔の向いている方向を推定する推定方法の一例を説明する図である。

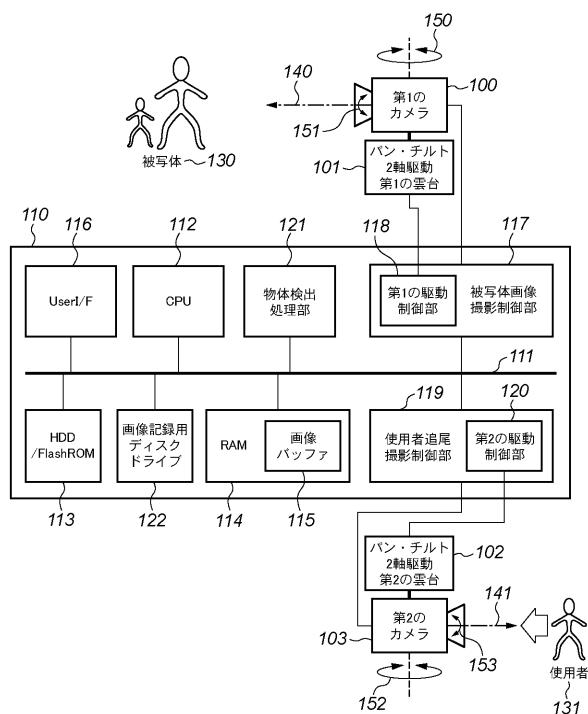
【図12】顔輪郭の位置の抽出方法の一例を説明する図である。

【図13】顔の眼部の画像に基づき、視線方向を推定する推定方法の一例を説明する図である。

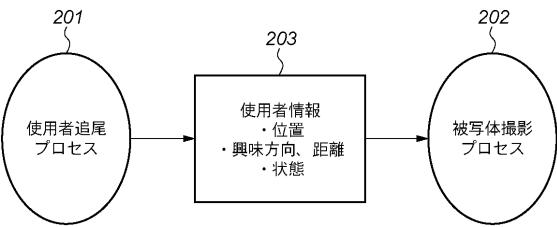
【図14】検出した両眼の視線方向より、対象とする被写体の存在する位置を推定する推定方法を説明する図である。

【図15】本発明の第4の実施形態に係わる撮像装置のハードウェア構成を示すブロック図である。

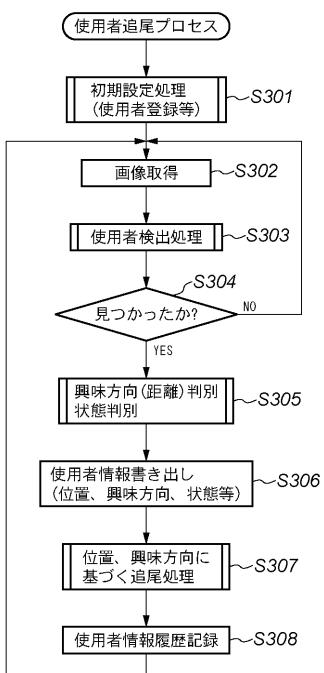
【図1】



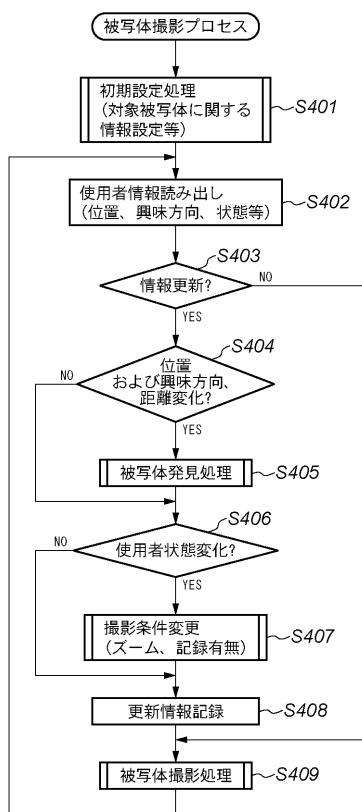
【図2】



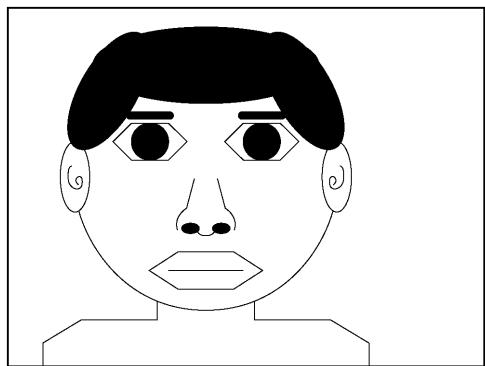
【図3】



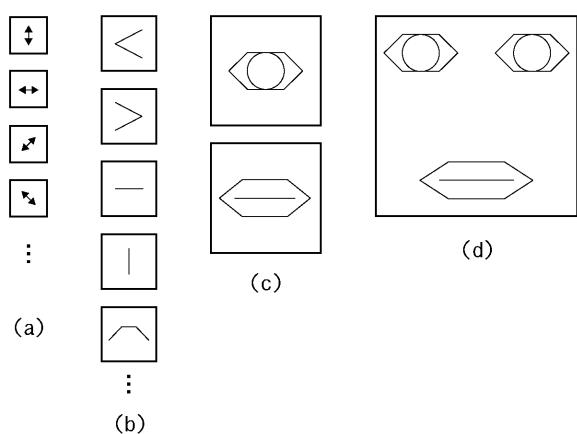
【図4】



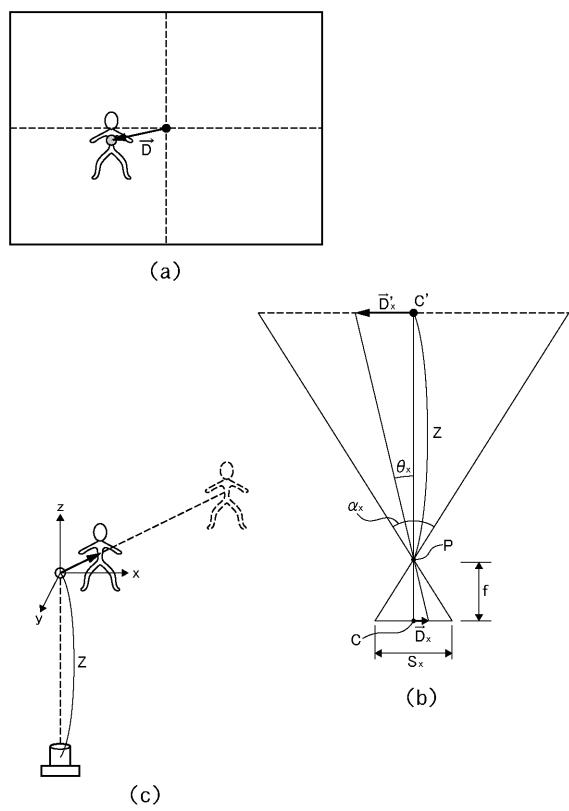
【 図 5 】



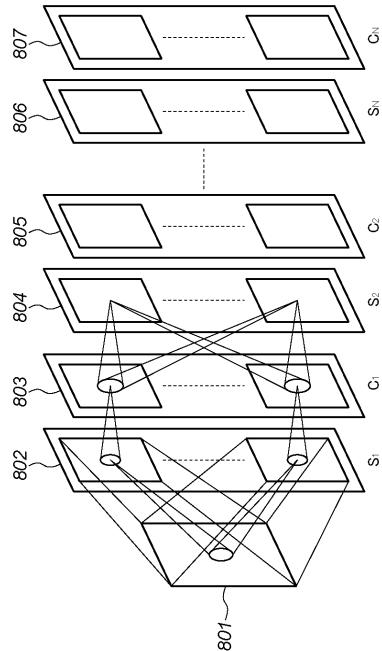
【図6】



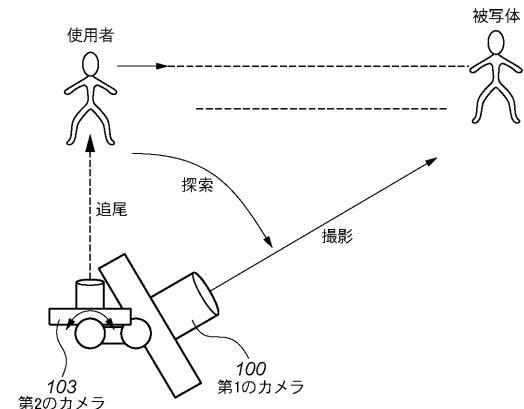
【図8】



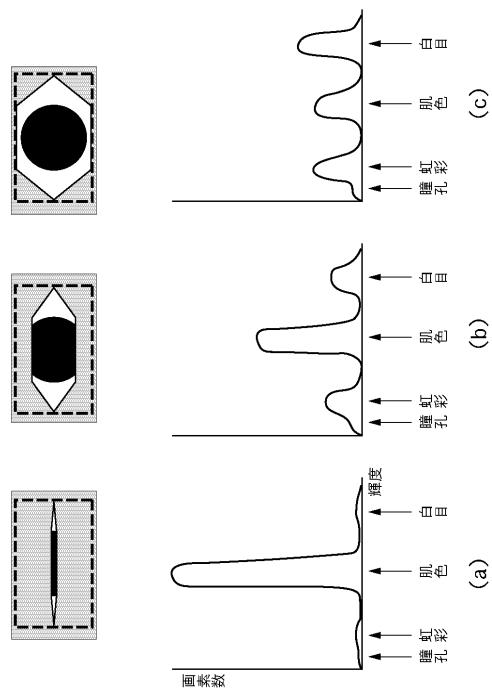
【 四 7 】



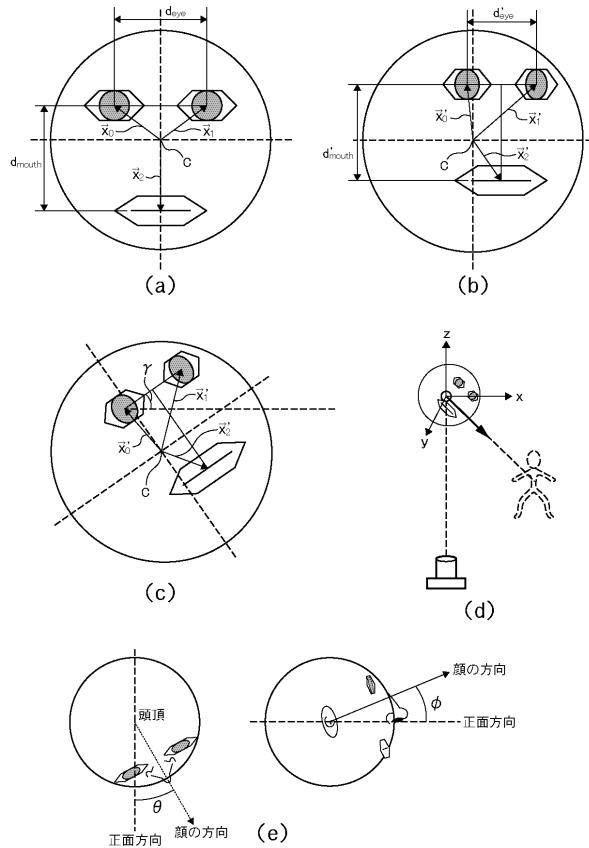
【 図 9 】



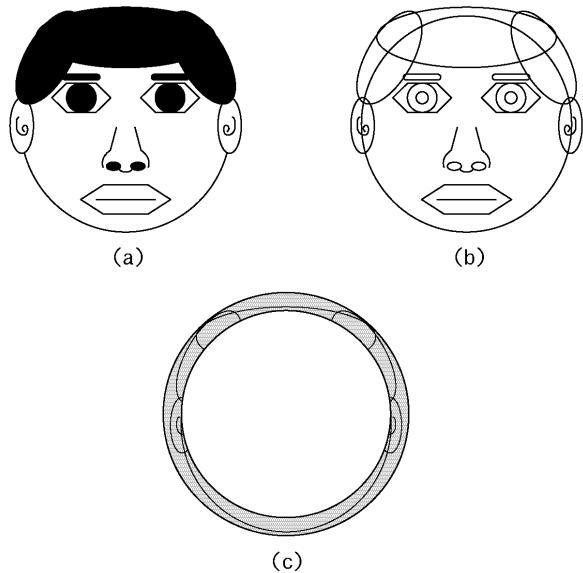
【図10】



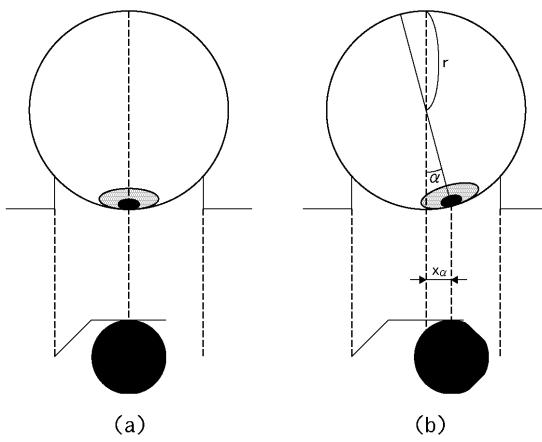
【図11】



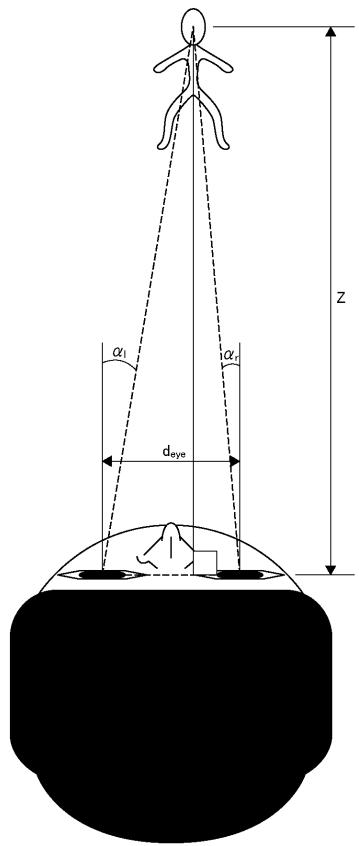
【図12】



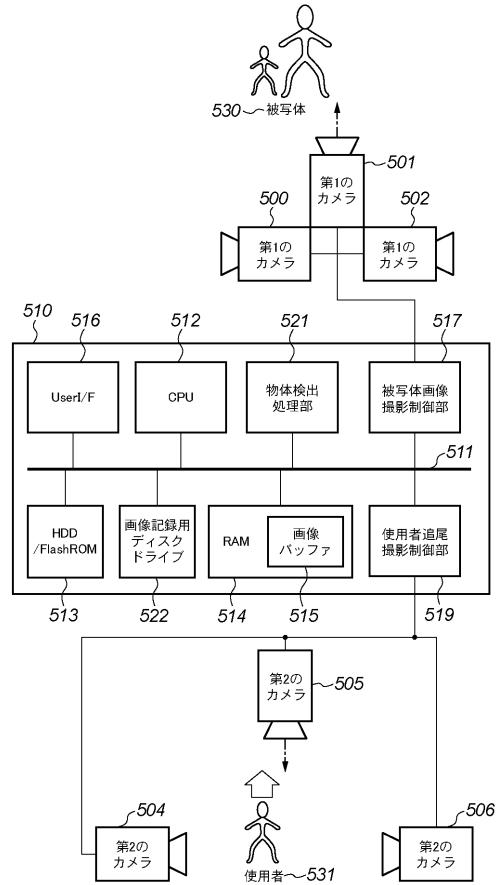
【図13】



【図14】



【 図 1 5 】



フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 3 B	17/56	B
G 0 3 B	15/00	U
G 0 2 B	7/08	C

F ターム(参考) 2H105 AA13 AA14

5C122	DA02	DA03	DA19	EA42	EA63	EA65	EA66	EA67	FA11	FE02
FH11	FH14	GA23	GD04	GD06	HA26	HA82	HB01	HB10		