

(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2017-33547

(P2017-33547A)

(43) 公開日 平成29年2月9日 (2017. 2. 9)

(51) Int.Cl.		F I				テーマコード (参考)
G06T 7/60	(2017.01)	G06T 7/60	150P			5B043
G06T 7/00	(2017.01)	G06T 7/00	510B			5L096
G06F 17/30	(2006.01)	G06F 17/30	170B			
		G06F 17/30	210A			

審査請求 未請求 請求項の数 15 O L (全 17 頁)

(21) 出願番号	特願2016-136297 (P2016-136297)	(71) 出願人	000001007
(22) 出願日	平成28年7月8日 (2016. 7. 8)		キヤノン株式会社
(31) 優先権主張番号	特願2015-155391 (P2015-155391)		東京都大田区下丸子3丁目30番2号
(32) 優先日	平成27年8月5日 (2015. 8. 5)	(74) 代理人	100076428
(33) 優先権主張国	日本国 (JP)		弁理士 大塚 康德
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

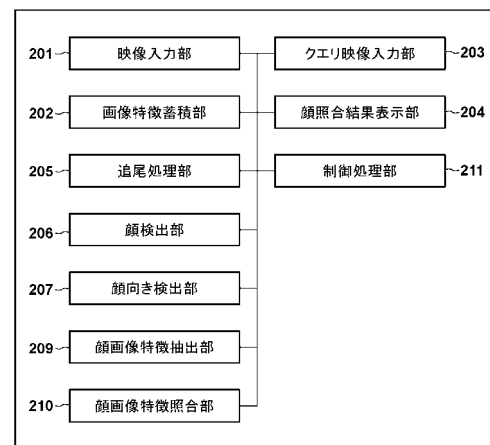
(54) 【発明の名称】 情報処理装置及びその制御方法及びプログラム

(57) 【要約】 (修正有)

【課題】照合データベースに登録する顔の特徴量を抑制することで、照合データベースが肥大することを抑制する技術を提供する。

【解決手段】情報処理装置は、映像を取得する取得部と、取得された映像の複数のフレームより同一人物の顔を少なくとも1つ検出する検出部と、検出された顔を、予め定められた複数のグループに分類する分類部と、前記グループごとに分類された顔の中から、前記グループごとに2以上の整数である第1の所定数以下の顔を選択する選択部と、前記選択した顔の特徴量をデータベースに登録する登録部とを有する。

【選択図】図2



【特許請求の範囲】**【請求項 1】**

映像を取得する取得手段と、
前記取得された映像の複数のフレームより同一人物の顔を少なくとも 1 つ検出する検出手段と、
前記検出された顔を、予め定められた複数のグループのいずれかに分類する分類手段と、
前記グループごとに分類された顔の中から、前記グループごとに 2 以上の整数である第 1 の所定数以下の顔を選択する選択手段と、
前記選択した顔の特徴量をデータベースに登録する登録手段と
を有することを特徴とする情報処理装置。

10

【請求項 2】

前記選択手段は、顔のサイズ、ブレ、目つぶり、口開けの少なくとも 1 つに基づいて、前記第 1 の所定数以下の顔を選択することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

前記検出手段は、あるフレームにおいて検出された顔に基づいて当該検出された顔を前記映像において追尾することにより、前記複数のフレームより同一人物の顔を検出することを特徴とする請求項 1 又は 2 に記載の情報処理装置。

【請求項 4】

前記グループは人物の顔向きであることを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の情報処理装置。

20

【請求項 5】

前記複数のグループのうちの第 1 のグループに対して定められた前記第 1 の所定数と、第 2 のグループに対して定められた前記第 1 の所定数とは異なる値であることを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載の情報処理装置。

【請求項 6】

前記選択手段は、優先度に基づいて、前記複数のグループから第 2 の所定数以下のグループを選択し、当該選択した第 2 の所定数以下のグループごとに前記第 1 の所定数以下の顔を選択し、

前記優先度は、前記複数のグループに対して予め設定されている

ことを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載の情報処理装置。

30

【請求項 7】

前記グループは人物の顔向きであり、

前記優先度は、前記顔向きが正面から外れるグループほど低い

ことを特徴とする請求項 6 に記載の情報処理装置。

【請求項 8】

対象人物のクエリ画像から抽出された特徴量に基づき、前記データベースに登録された特徴量との照合を行う照合手段を更に有することを特徴とする請求項 1 乃至 7 のいずれか 1 項に記載の情報処理装置。

【請求項 9】

対象人物のクエリ映像を入力する入力手段を更に有し、

前記照合手段は、前記クエリ映像から抽出された前記対象人物のクエリ画像の特徴量に基づいて照合を行うことを特徴とする請求項 8 に記載の情報処理装置。

40

【請求項 10】

前記入力手段は対象人物の複数のクエリ画像を入力し、

前記照合手段は、前記複数のクエリ画像を顔向きごとに分類し、1 つの顔向きに対して 1 つの特徴量を決定し、当該決定した顔向きごとの特徴量を用いて照合を行う

ことを特徴とする請求項 9 に記載の情報処理装置。

【請求項 11】

前記入力手段は、対象人物の顔を複数の方向から撮影して得られた複数のクエリ画像を

50

入力し、

前記照合手段は、前記入力された複数のクエリ画像から得られた特徴量を用いて照合を行う

ことを特徴とする請求項 9 に記載の情報処理装置。

【請求項 12】

前記照合手段による照合結果を表示装置に表示する表示手段を更に有することを特徴とする請求項 8 乃至 11 のいずれか 1 項に記載の情報処理装置。

【請求項 13】

前記登録手段は、顔器官点に関する特徴量をデータベースに登録し、

前記照合手段は、対象人物のクエリ画像から抽出された顔器官点に関する特徴量を、前記データベースに登録された顔器官点に関する特徴量と照合する

ことを特徴とする請求項 8 乃至 12 のいずれか 1 項に記載の情報処理装置。

【請求項 14】

情報処理装置の制御方法であって、

取得手段が、映像を取得する取得工程と、

検出手段が、前記取得された映像の複数のフレームより同一人物の顔を少なくとも 1 つ検出する検出工程と、

分類手段が、前記検出された顔を、予め定められた複数のグループのいずれかに分類する分類工程と、

選択手段が、前記グループごとに分類された顔の中から、前記グループごとに 2 以上の整数である第 1 の所定数以下の顔を選択する選択工程と、

登録手段が、前記選択した顔の特徴量をデータベースに登録する登録工程と

を有することを特徴とする情報処理装置の制御方法。

【請求項 15】

コンピュータが読み込み実行することで、前記コンピュータを、請求項 1 乃至 13 のいずれか 1 項に記載の情報処理装置の各手段として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像中の人物の照合技術に関するものである。

【背景技術】

【0002】

従来、監視映像の各フレームから人間の顔を検出し、その顔から画像特徴量を算出して映像のフレームと対応づけて蓄積し、クエリ画像で示された人物の顔と一致する映像を蓄積画像から探し出す技術が知られている。例えば、迷子の子供の顔をクエリ画像とし、蓄積画像から探し出すことで、その子供が映っている映像から場所と時間を特定する、ということが可能となっている。

【0003】

映像の全てのフレームから検出される顔を全て登録する方法が有るが、この場合、蓄積する画像特徴量が膨大となる。そこで、フレーム間引きを行う事により処理フレームレートを低下させ、登録する画像特徴量を削減する方法が考えられる。しかし、一般に、顔照合の精度は顔向きに依存するため、フレーム間引きを行う事で好ましい顔向きのフレームを捨ててしまう事が起きるため、照合漏れを起こす確率が高くなる。

【0004】

特許文献 1 は、顔認証システムに関するものである。この特許文献 1 は、予め登録してある人物 A に対し、新規登録人物 B のどの画像特徴量を登録するか、また、どの様に高精度な照合を行うかをその課題としている。このシステムでは登録のフェーズで、動画を撮影し複数のフレームから抽出される人物 B の顔から、登録にふさわしい顔およびその画像特徴量を決定する。映像フレームから検出された人物 B の顔の向きを判別し、顔向き毎に区分し、それらを総当たりで類似比較し本人の顔らしさを決定するための類似度の閾値を

10

20

30

40

50

決定する。更に人物 B の各顔向きの顔の画像特徴量と、人物 A の同じ顔向きの画像特徴量を比較し、類似度がその閾値以上であれば人物 B のその顔向きの画像特徴量は記憶を行わず破棄する。これにより、人物 A、B の顔を弁別するために必要な画像特徴量に絞り、蓄積する画像特徴量の量を削減し、且つ、弁別精度の劣化を防ぐ事を可能としている。

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特開 2012-37995 号公報

【特許文献 2】特開 2002-37333 号公報

【非特許文献】

【0006】

【非特許文献 1】Erik Murphy - Chutorian, "Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation," in Proc. IEEE Conf. Intelligent Transportation Systems, 2007, pp. 709-714.

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかしながら、この特許文献 1 の技術は、監視映像からの人物検索には利用するのは困難である。なぜなら、或る人物の顔を登録する際には、既に登録済みの全ての人物の顔との照合を行う必要があり、大量に人物が登場する監視映像中では、登録処理にかかる負荷が極端に大きくなるからである。

【0008】

本発明は、上記問題点に鑑みなされたものである。そして、本発明は、照合データベースに登録する顔の特徴量を抑制することで、照合データベースが肥大することを抑制する技術を提供しようとするものである。

【課題を解決するための手段】

【0009】

上記課題を解決するために本発明の情報処理装置は以下の構成を備える。すなわち、映像を取得する取得手段と、前記取得された映像の複数のフレームより同一人物の顔を少なくとも 1 つ検出する検出手段と、前記検出された顔を、予め定められた複数のグループのいずれかに分類する分類手段と、前記グループごとに分類された顔の中から、前記グループごとに 2 以上の整数である第 1 の所定数以下の顔を選択する選択手段と、前記選択した顔の特徴量をデータベースに登録する登録手段とを有する。

【発明の効果】

【0010】

本発明によれば、照合データベースに登録する顔の特徴量を抑制することで、照合データベースが肥大することを抑制することが可能になる。

【図面の簡単な説明】

【0011】

【図 1】実施形態の装置のブロック構成図。

【図 2】実施形態の装置の機能構成図。

【図 3】実施形態に顔向きに応じた分類を説明するための図。

【図 4】実施形態におけるオブジェクト追尾情報の内容を示す図。

【図 5】実施形態における画像特徴量蓄積部の処理を説明するための図。

【図 6】実施形態におけるオブジェクト追尾が開始と終了するタイミングの例を示す図。

10

20

30

40

50

【図 7】実施形態における照合データベースの構造の例を示す図である。

【図 8】第 1 の実施形態における映像蓄積処理を示すフローチャート。

【図 9】図 8 における DB 登録処理を示すフローチャート。

【図 10】第 1 の実施形態における照合処理を示すフローチャート。

【図 11】第 2 の実施形態における照合処理を示すフローチャート。

【発明を実施するための形態】

【0012】

以下、添付図面に従って本発明に係る実施形態を詳細に説明する。なお、以下で説明する各実施形態は、本発明を具体的に実施した例を示すもので、特許請求の範囲に記載の構成の具体的な実施形態の 1 つである。

【0013】

[第 1 の実施形態]

図 1 は、本実施形態における画像照合装置 100 のブロック構成図である。本装置 100 は、画像検索のサーバ装置、および、クライアント装置として機能するものとして説明するが、サーバ機能、クライアント機能はそれぞれ独立した装置で構成しても構わない。また、サーバ装置は、単独の装置ではなく、複数のコンピュータ装置で処理を分散して実現しても構わない。複数のコンピュータ装置で構成される場合は、互いに通信可能なように Local Area Network (LAN) などで接続されている。コンピュータ装置は、パーソナルコンピュータ (PC) やワークステーション (WS) 等の情報処理装置によって実現することができる。

【0014】

図 1 において、CPU 101 は画像照合装置 100 全体を制御する Central Processing Unit である。ROM 102 は変更を必要としないプログラム (BIOS 等) やパラメータを格納する Read Only Memory である。RAM 103 は外部装置などから供給されるプログラムやデータを一時記憶する Random Access Memory である。外部記憶装置 104 はコンピュータ装置 100 に固定して設置されたハードディスク等の大容量の記憶装置である。この外部記憶装置 104 は、本装置のオペレーティングシステム (OS)、画像照合にかかるアプリケーション、照合する際に参照する照合データベース (以下、照合 DB という) が格納されている。また、後述するが、ビデオカメラで撮像した映像データも、この外部記憶装置 104 に格納されている。

【0015】

入力デバイスインターフェイス 105 はユーザの操作を受け、データを入力するポインティングデバイスやキーボードなどの入力デバイス 109 とのインターフェイスである。出力デバイスインターフェイス 106 はコンピュータ装置 100 の保持するデータや供給されたデータを表示するためのモニタ 110 とのインターフェイスである。通信インターフェイス 107 はインターネットなどのネットワーク回線 111 に接続するための通信インターフェイスである。ネットワーク回線 111 上には、ネットワークカメラ 112a 乃至 112c が接続されている。なお、ネットワークカメラ 112a 乃至 112c には、それぞれカメラ ID が割り当てられている。ここでは、ネットワークカメラ 112a、112b、112c にカメラ ID 1, 2, 3 が割り当てられているものとして説明する。また、接続されているネットワークカメラは 1 以上であればよく、その数は問わない。システムバス 108 は 101 ~ 107 の各ユニットを通信可能に接続する伝送路である。

【0016】

上記構成において、電源が ON になると、CPU 101 は ROM 102 に格納された BIOS プログラムを実行し、外部記憶装置 104 から OS (オペレーティングシステム) を RAM 103 上にロードし、OS を実行する。この結果、本装置が情報処理装置として機能することになる。そして、OS の制御下において、外部記憶装置 104 から、画像照合にかかるサーバプログラム、および、クライアントプログラムを実行することで、本装置が画像照合装置として機能することになる。

10

20

30

40

50

【 0 0 1 7 】

図 2 は、実施形態における映像照合装置の機能ブロック構成図である。同図は、検索サーバ、クライアントの両機能のプログラムを、CPU 101 が実行した場合の機能構成図でもある。

【 0 0 1 8 】

制御処理部 211 は、符号 201 乃至 210 で示される各処理部を制御するためのメイン処理を行う。

【 0 0 1 9 】

映像入力部 201 は、通信インターフェース 107 を介して、ネットワークカメラ 112 a 乃至 112 c それぞれからの映像フレームを受信し、外部記憶装置 104 にそれぞれ格納していく。入力するフレームレートは例えば 30 フレーム / 秒とする。上記の結果、外部記憶装置 104 は、ネットワークカメラ 112 a ~ 112 c それぞれの動画ファイルが作成されることになる。また、映像入力部 201 は、受信した映像フレームを顔検出部 205 に供給する。

【 0 0 2 0 】

顔検出部 205 は、顔向き検出部 207、顔画像特徴抽出部 209 を利用して、顔検出を行う。顔画像特徴抽出部 209 は、映像フレーム中の顔領域を検出し、その特徴量を抽出し、顔検出部 205 に返す。本実施形態では、顔画像特徴量として顔器官点の SIFT (Scale Invariant Feature Transform) 特徴量を求めるものとする。また、顔画像特徴抽出部 209 は、顔特徴量を抽出する際に、その顔領域の外接矩形の左上隅の座標、並びに、その外接矩形のサイズ (幅 W と高さ H) も顔検出部 205 に返す。なお、サイズの代わりに外接矩形の右下隅の座標でも構わない。

【 0 0 2 1 】

顔向き検出部 207 は、顔検出部 205 で検出された顔の向きを検出し、その向きを示す情報 (顔向きパターン番号) を、顔検出部 205 に返す。なお、画像中の人物の顔の向きを検出する技術は、公知であるのでその詳述は省略する (例えば非特許文献 1)。

【 0 0 2 2 】

実施形態における顔向き検出部 207 は、着目している顔の向きが、図 3 (a) に示す 5 × 5 個のパターンのいずれかに該当するかを判定する。なお、図示では 5 × 5 個のパターンで示しているが、この数は特に問わない。ネットワークカメラの解像度が十分に高ければ、特徴量や向きの精度が高くなるので、パターン数を更に増やしても構わない。図示のパターンでは、正面を向いている顔の向きを中心に位置し、上下左右方向へのずれ角が大きくなるほど、中心位置から外れるようにした。換言すれば、5 × 5 個のパターンの中心に近いほど、顔向きが正面に向くように顔向きのパターンを配置した。顔照合では、正面を向いている顔を用いた場合にその精度が高くなる。つまり 5 × 5 個の顔向きの中心位置が最も優先順位が高く、その中心位置から外れるにしたがって優先順位が低いものとした。図 3 (b) はその顔向きパターン番号の一例である。図示の如く、顔の向きが正面を向いている場合の顔向きパターン番号は “1” である。そして、その周りに向き顔向きパターン番号 2 ~ 9、さらにその外側に顔向きパターン番号 10 乃至 25 を割り当てた。なお、顔向きパターン番号 2 ~ 9 は、必ずしも図示の通りでなくても構わない。顔向きパターン番号 10 乃至 25 も同様である。

【 0 0 2 3 】

以上の説明の通り、顔検出部 205 は、ネットワークカメラ 112 a 乃至 112 c それぞれより受信した各映像フレームから、顔向き検出部 207、顔画像特徴抽出部 209 を利用して、人物の顔の位置、サイズ、特徴量、顔向きパターン番号を得る。

【 0 0 2 4 】

追尾処理部 205 は、制御処理部 211 からの追尾開始指示情報を受けると、現在のフレーム内の指示された顔を追尾開始対象の顔とし、それ以降のフレーム画像の該当する顔の追尾処理を行う。ここで追尾開始指示情報には、いずれのネットワークカメラからのフレーム画像かを示す情報 (カメラ ID)、その追尾対象の顔の初期位置を示す情報、追尾

10

20

30

40

50

する顔を特定するためのオブジェクトIDが含まれる。オブジェクトIDは、初期値は1であり、追尾すべき顔が発見されるたびに1増加されるものである。なお、画像中の人物を追尾する処理は、例えば、特許文献2に開示された技術を用いればよい。

【0025】

追尾処理部205は、上記の追尾開始指示情報を受けると、図4に示すようなオブジェクト追尾情報をRAM103上に生成する。1つのオブジェクト追尾情報は、カメラIDとオブジェクトIDで特定される。格納される情報には、本装置がその日の撮影記録を開始してから、追尾開始を指示された時のフレーム番号（時刻でも構わない）が含まれる。そして、さらには、該当する映像フレーム内の追尾対象の顔の外接矩形の左上隅の位置とサイズ（幅、高さ）、顔向きパターン番号、および、その顔の特徴量が含まれる。これらの情報は、顔向き検出部207、顔画像特徴抽出部209からの情報を利用する。

10

【0026】

そして、追尾処理部205は、映像フレームが順次入力され、該当する顔の追尾が成功しつつある限り、該当するオブジェクト追尾情報に、上記で説明したフレーム番号、位置、サイズ、顔向きパターン番号、特徴量を追記していく。そして、追尾不可となった場合、追尾処理部205は、制御処理部211に対して該当するオブジェクトIDに対する追尾が終了した旨を通知する。なお、追尾不可となるケースは、追尾中の人物（の顔）がカメラの視野範囲外に移動した場合、人物が後ろを向いたなど、顔として認識できなくなった場合がこれに当たる。

【0027】

20

制御処理部211は、追尾終了の通知を受けると、該当するオブジェクト追尾情報を、画像特徴蓄積部202に渡し、照合DBへの蓄積処理（登録処理）を行わせる。この画像特徴蓄積部202の詳細は後述するが、登録処理を終えると、制御処理部211は、登録済みのオブジェクト追尾情報をRAM103から消去する。

【0028】

なお、カメラの視野範囲外に移動した人物が、再度、カメラの視野範囲内に移動することもあり得る。実施形態の装置は、カメラの視野外になった人物と、再度視野内に移動した人物が同一人物か否かは判断しない。それ故、再度、カメラの視野範囲内に移動してきた人物の顔に対しては、新規なオブジェクトIDを発行し、追尾を開始することになる。従って、追尾処理部205は、結局のところ、カメラの撮像視野範囲にて、顔と認識されたオブジェクトの全てについて追尾処理を行う。従って、図4に示すオブジェクト追尾情報は現に追尾中の顔の数だけ生成されることになる。1台のネットワークカメラの視野内に、顔と認識できる程度にサイズの顔領域はせいぜい4,50個程度である。ネットワークカメラが3台であることを考慮しても、最大でも150個の顔が追尾対象となるに留まるので、現在のコンピュータでは十分に対処できる数と言える。

30

【0029】

因に、ネットワークカメラからの映像フレームを受信していて、図6における時刻t0にて人物Aの顔が初めて検出された場合、人物Aに対してオブジェクトIDが発行され、人物Aのオブジェクト追尾情報が生成される。これ以降、人物Aの追尾できなくなる時刻t3まで、人物Aの追尾処理が行われる。この間、時刻t1にて、別の人物Bの顔が、同じ撮影視野内に初めて検出された場合には、人物Bに対してオブジェクトIDが発行され、人物Bのオブジェクト追尾情報が生成される。そして人物Bの追尾できなくなる時刻t2まで、人物Bの追尾処理が行われることになる。

40

【0030】

次に、画像特徴蓄積部202の蓄積処理を図5を参照して説明する。この蓄積処理は、既に説明したように或る人物の顔を追尾処理が終了した場合に行われる。

【0031】

図5における符号500はオブジェクト追尾情報に格納された顔向きパターン番号の時間推移を表している。なお、図示では、オブジェクト追尾情報に格納された顔向きパターン番号は全部で6つの例を示しているが、これは説明を簡単にするためである。

50

【 0 0 3 2 】

まず、画像特徴蓄積部 2 0 2 は、オブジェクト追尾情報に格納された顔向きパターン番号の出現数の集計処理を行う。図 5 の符号 5 0 1 が集計結果を示し、顔パターンの右下の数字が、その出現数（累積数）を表している。

【 0 0 3 3 】

顔向きパターン番号は図 3（b）に示す如く 1 乃至 2 5 であるので、顔向き I D の集計結果を変数 F（顔向きパターン番号）として表す。ここで、実施形態においては、追尾対象の顔に対して、最大 4 つの顔向きの特徴量を照合 D B に登録するものとする。また、1 つの顔向きについて、最大 2 個の特徴量を照合 D B に登録する。もちろん、この数は一例であって、それ以上であっても構わない。

10

【 0 0 3 4 】

実施形態の場合、顔向きパターン番号は、その番号順に優先順位を設定されている（図 3（b）参照）。画像特徴蓄積部 2 0 2 は、集計した変数 F（1）乃至 F（2 5）をこの順番に調べ、出現数が非ゼロ（1 以上）である最初の 4 つを見つける。図 5 の場合、参照符号 5 0 2 で示すように、斜線で示した F（1）、F（2）、F（3）、F（6）がこれに当たる。つまり、オブジェクト追尾情報中の顔向きパターン番号 = 1, 2, 3, 6 となった特徴量が照合 D B に登録する候補として決定する。

【 0 0 3 5 】

ここで、顔向きパターン番号“1”の出現数 F（1）は“1”である。つまり、オブジェクト追尾情報に顔向きパターン番号“1”の特徴量は 1 つしか存在しなかったことになる。それ故、画像特徴蓄積部 2 0 2 は、オブジェクト追尾情報内の、顔向きパターン番号“1”となっている特徴量を照合 D B に登録する。

20

【 0 0 3 6 】

また、顔向きパターン番号“2”の出現数 F（2）は“4”であるので、そのうちの 2 つを登録するものとして決定する。実施形態では、4 つのうち、顔のサイズが大きい方から特徴量を登録するものとする。顔の大きい方が、その人物の顔を撮影した際に、カメラとの距離が短いことになり、高い精度が期待できるからである。また、顔向きパターン番号 = 3、6 の場合も同様である。

【 0 0 3 7 】

以上の結果、照合 D B に登録するのは、1 つのオブジェクト I D につき、4 つの顔向きの特徴量となり、また、1 つの顔向きについては最大で 2 個の特徴量となる。尚且つ、その顔の特徴量は、顔の向きが重複せず、かつ、正面を向いている程度が高いほど登録される可能性が高くできる。しかも、登録する特徴量は、顔のサイズが大きいほど優先されることになり、精度の高い整合処理が期待できる。

30

【 0 0 3 8 】

図 7 は実施形態における照合 D B の構造例を示している。この照合 D B は、外部記憶装置 1 0 4 に確保されるものである。同図に示すように、照合 D B の 1 つのレコードは、オブジェクト I D、特徴量（最大 4 つ）、カメラ I D、登録した顔の特徴量の顔を検出した期間の最初のフレーム番号（時刻）、座標、顔のサイズを格納するフィールドで構成される。このうち、特徴量が検索の際にキーとして利用されることになる。

40

【 0 0 3 9 】

次に、上記説明を踏まえ、実施形態における映像記録のメイン処理を図 8 のフローチャートに従って説明する。この処理制御部 2 1 1 が行うものでもある。

【 0 0 4 0 】

制御処理部 2 1 1 は、ステップ S 8 1 にて、通信インターフェース 1 1 1 を介してネットワークカメラ 1 1 2 a 乃至 1 1 2 c それぞれから映像フレームを受信し、外部記憶装置 1 0 4 に、被照合対象の動画画像ファイルとしてそれぞれ記録していく。この際に作成される 3 つの動画画像ファイルのファイル名は、例えばカメラ I D を付けて作成する。

【 0 0 4 1 】

次いで、制御処理部 2 1 1 は、現映像フレームを顔認識部 2 0 6 に渡し、顔認識処理を

50

実行させる。この結果、映像フレーム内に存在する各顔の認識結果を得ることができる。制御処理部 211 は、現映像フレームの認識結果と、1 つ前の映像フレームの認識結果から、現映像フレーム中に未追尾の顔が存在するか否かを判定する。それが存在した場合、処理はステップ S 84 に進む。このステップ S 84 にて、制御処理部 211 は、追尾処理部 205 に、その新規追尾する顔が存在する位置とサイズ、新規オブジェクト ID、カメラ ID、フレーム番号を引数にした追尾開始指示情報を与える。この結果、追尾処理部 205 は図 4 に示すようなオブジェクト追尾情報を作成し、その顔の追尾処理を行う。

一方、制御処理部 211 は、ステップ S 85 にて、追尾処理部 205 から追尾終了通知を受信したか否かを判定する。追尾終了通知を受信しなかった場合には、ステップ S 81 以降の処理を繰り返す。そして、追尾終了通知を受信した場合、制御処理部 211 はステップ S 86 にて、追尾終了したオブジェクト追尾情報を画像特徴蓄積部 202 に渡し、照合 DB への登録処理を行わせる。この登録処理を行った後、照合 DB に登録に使用したオブジェクト追尾情報を削除し、ステップ S 81 以降の処理を行う。

【0042】

次に、上記のステップ S 86 の DB 登録処理（画像特徴蓄積部 202 の処理）を図 9 のフローチャートに従って説明する。

【0043】

画像特徴蓄積部 202 は、ステップ S 91 にて、制御処理部 211 からオブジェクト追尾情報を受信する。そして、画像特徴蓄積部 202 は、受信したオブジェクト追尾情報を解析し、顔向きパターン番号の集計処理を行う。つまり、顔向きパターン番号 1 乃至 25 の出現回数 $F(1)$ 乃至 $F(25)$ を求める。

【0044】

次いで、画像特徴蓄積部 202 は、ステップ S 93 にて変数 i を 1、変数 j を 0 に初期化する。変数 i は出現回数 $F()$ を特定するための変数であり、変数 j は登録する特徴量の個数をカウントするための変数である。

【0045】

ステップ S 94 では、画像特徴蓄積部 202 は変数 i と 25 とを比較する。変数 i が 25 以下である場合、全出現回数 $F()$ をチェックしていないことになる。それ故、画像特徴蓄積部 202 は処理をステップ S 95 に進め、変数 j と照合 DB に登録するために予め設定された特徴量の上限数 N （実施形態では $N = 4$ ）とを比較する。 $j < N$ の場合、照合 DB に登録する特徴量の数が上限数に達していないので、画像特徴蓄積部 202 は処理をステップ S 96 に進める。このステップ S 96 では、顔向きパターン番号が i の出現回数 $F(i)$ が非ゼロ（1 以上）か否かを判定する。非ゼロの場合、画像特徴蓄積部 202 はステップ S 97 にて、顔向きパターン番号 = i の中で顔サイズが大きい方から所定数以下の特徴量（本実施形態の場合は、2 以下）を、照合 DB への登録対象として決定する。そして、登録対象が 1 つ決定したので、ステップ S 98 にて、画像特徴蓄積部 202 は、変数 j を 1 だけ増加させる処理を行う。そして、ステップ S 100 にて、画像特徴蓄積部 202 は、変数 i を 1 だけ増加させる処理を行い、ステップ S 94 に処理を戻す。

【0046】

さて、ステップ S 94、S 95 のいずれかの判定結果が No を示す場合、画像特徴蓄積部 202 はステップ S 99 に処理を進める。このステップ S 99 には、画像特徴蓄積部 202 は、登録対象として決定した特徴量を、オブジェクト ID、フレーム数、カメラ ID、位置、サイズ情報とともに、照合 DB に登録処理を行い、本処理を終える。なお、変数 j が N に到達する以前に、変数 i が 25 となることも起こり得る。この場合、登録する特徴量の個数は 4 未満の個数となる。

【0047】

以上が本実施形態における映像データ & 特徴量の登録処理である。次に、本実施形態における照合処理を説明する。照合処理は、クエリ映像入力部 203、顔認識部 206、顔画像特徴照合部 210、顔照合結果表示部 204、および、それらを制御する制御処理部 211 により実現される。以下、図 10 のフローチャートに従い、照合処理を説明する。

10

20

30

40

50

まずステップS601にて、制御処理部211はクエリ映像入力部203に対してクエリ映像の入力を指示する。クエリ映像の入力源は特に問わない。例えば、検索依頼人が所有する携帯端末を本装置と接続し、探そうとしている人物の顔が映った画像をクエリ映像として本装置に転送しても構わない。なお、本装置がインターネットに接続されている場合には、メールとして受信しても構わない。いずれにせよ、クエリ映像を入力した場合、クエリ人物の顔を特定する処理を行う。例えば、顔検出部207にてクエリ映像中に存在する全顔を検出し、検出した顔に矩形枠を付加して表示装置に表示し、検索依頼人にその中の1つを検索対象の顔画像（クエリ顔画像）と選択してもらう。なお、検出した顔が1つだけである場合には、無条件にその顔が選択されたものとして扱っても構わない。クエリ顔画像が決定されると、そのクエリ顔画像を検出の際に抽出した特徴量を検出キーとして決定する。

10

【0048】

次に、ステップS602にて、制御処理部211は、照合DBに登録されているオブジェクトIDの総個数M（レコード数）を得る。制御処理部211は、取得したM、クエリ顔画像の特徴量を引数にして、顔画像特徴照合部210に照合処理を要求する。

【0049】

顔画像特徴照合部210は、ステップS603にて、変数*i*をゼロに初期化する。そして、ステップS604にて、変数*i*がM以下であるか否かを判定する。*i* < Mの場合、全レコードとの照合が済んでいないことになるので、処理はステップS606に進む。このステップS606では、顔画像特徴照合部210は、照合DBの*i*番目のレコードを読み込み、その中に格納されている*Ni*個の特徴量を読み込む。実施形態では、最大で4つの顔向きについて、それぞれ最大で2個登録されているため、*Ni*個は最大で8である。そして、クエリ顔画像の特徴量と、*Ni*の特徴量それぞれとの距離を計算し、その中の最少距離を、*i*番目のレコードのオブジェクトIDの人物顔とクエリ顔画像との距離として決定する。顔画像特徴照合部210は、オブジェクトIDと決定した距離とをペアにして、RAM103の所定エリアに記憶する。この後、次のレコードとの照合処理を行うため、変数*i*を“1”だけ増加させ、ステップS604に処理を戻す。

20

【0050】

以上の処理を繰り返していき、全レコードとの比較処理を終えたとき（変数*i* = Mと判断されたとき）、RAM103には、全ての登録済みのオブジェクトIDに対する距離が格納されていることになる。そこで、顔画像特徴照合部210は、ステップS605にて、RAM103に記憶された距離を小さい順にソートし、本処理を終える。

30

【0051】

この後の処理としては、制御処理部211が照合結果を顔照合結果表示部204に渡すことで、照合結果の表示処理を行う。表示の仕方としては、例えば、距離が小さい順の予め定められた個数の顔画像一覧を表示させ、その中の1つが選択された場合には、その顔に対応するフレーム番号とカメラIDとから、動画画像の該当する位置から再生を開始するなどを行う。

【0052】

なお、上記実施形態では、1つのオブジェクト追尾情報中に、同じ顔向きパターン番号が複数存在したとき、その中の最大顔サイズの特徴量を登録対象として決定したが、登録対象の特徴量としてサイズ以外の条件にしてもよい。例えば、映像フレームのブレの度合いが小さいものを優先したり、目つぶりや口開き顔を登録されにくくすることが考えられる。

40

【0053】

また、動画を撮影するカメラでも静止画のカメラと同様に、その場所の明るさに従いシャッター速度が変わる場合がある。従って、暗い場所や被写体の動き速度により、顔画像のブレが生じる事があり、これは直接的に画像特徴量や属性情報の劣化の原因となる。ブレの推定に関しては、顔画像領域の周波数成分を求め、低周波成分と高周波成分との比率

50

を求め、これが低周波成分の比率が所定の値を超えた時にブレを生じていると判断する事が可能となる。更に、眼つぶりや口あき等がると、器官の画像特徴量が変質し、場合によっては属性情報にも誤りを生じる可能性が有る。

【 0 0 5 4 】

更には、上記のサイズや目つぶり、ブレなどの複数のパラメータを用いて登録対象の特徴量を決定する場合には、それらパラメータの種類の座標軸を持つ空間を想定し、その座標空間での距離を求めればよい。

【 0 0 5 5 】

また、顔向きパターン番号が 2 乃至 9 は、中心位置を取り囲むように配置されるので、それらは同一レベルと見ることもできる。従って、顔向きパターン番号 2 乃至 9 については、その中に集計結果が非ゼロとなっているものが複数存在する場合には、その中の最大サイズの特徴量を決定することを、登録数（実施形態では 4）に到達するまで繰り返せばよい。ただし、1つの顔向きパターン番号につき、登録する特徴量は最大でも 2 つとするのは同じである。上記は顔向きパターン番号 10 乃至 25 についても同様である。

【 0 0 5 6 】

また、上記実施形態では、所定数を 2 とし、顔向きパターン毎に所定数以下の顔を選択してその特徴量を登録するようにしている。しかし、この所定数は顔向きパターンによって異なる値としてもよい。例えば、顔向きパターン番号 1 は所定数を 4、顔向きパターン番号 2 ~ 9 は所定数を 3、顔向きパターン番号 10 ~ 25 は所定数を 2 というように、正面顔に近いパターンほど所定数を大きくする等、顔向きパターンによって異なる数の顔を選択するようにしてもよい。

【 0 0 5 7 】

以上説明したように、本実施形態によれば、同じ顔向きの特徴量は所定の数の特徴量のみ登録されることになるので、照合 DB が無駄に肥大化することを抑制できる。しかも、実施形態では、顔向きが正面であるほど優先順位を高くし、その優先順位の上記の所定数の特徴量を照合 DB に登録する。この結果、照合 DB の肥大化を更に抑制でき、しかも、精度の高い特徴量を登録することが可能となる。そして、同じ顔向きであっても、顔サイズなどの条件の高い精度が期待できる特徴量を登録するので、照合精度も高くすることができる。

【 0 0 5 8 】

なお、上記実施形態では、顔向きで分類して登録 DB に登録する例を示したが、例えば、表情（笑顔、泣き顔など）に分類して登録するものであってもよい。すなわち、本実施形態は、入力された映像から抽出された顔を、予め定められたグループ（顔向きや表情など）毎に分類して、登録する構成に広く適用できるものである。

【 0 0 5 9 】

なお、上記実施形態では、追尾処理部 205 が、追尾ターゲットの顔を追尾している際に、その顔の特徴量、向きを、オブジェクト追尾情報に追記していくものとして説明した。しかし、顔認識では、人間の顔であるか否かが判定できればよいので簡易な顔認識を行い、追尾終了した後に、照合 DB に登録するための、より精度の高い特徴量や顔向きを算出しても構わない。

【 0 0 6 0 】

また、上記実施形態では、追尾処理部 205 の追尾機能を利用して入力映像の各フレームから同一人物（オブジェクト ID が同一の人物）を検出できるようにしている。しかし、例えば、顔画像特徴抽出部 209 が抽出する顔画像特徴量に基づいて、入力された映像の各フレームに含まれる人物を識別することにより、入力映像から同一の人物を検出できるようにしてもよい。すなわち、上述の実施形態において、追尾処理部 205 は必須の構成ではなく、本実施形態はこれに限定されない。

【 0 0 6 1 】

[第 2 の実施形態]

上記第 1 の実施形態では、照合処理を行う際に入力するクエリ映像として 1 枚の画像を

10

20

30

40

50

利用する例であった。本第2の実施形態では、動画像をクエリ映像とする例を説明する。なお、装置構成は第1の実施形態と同じとし、その説明は省略する。

【0062】

以下、図11のフローチャートを参照し、第2の実施形態における照合処理を説明する。

【0063】

制御処理部211は、ステップS701にて、クエリ映像入力部293に対してクエリ動画映像の入力を指示する。クエリ動画映像の入力源は特に問わない。例えば、検索依頼人が所有する携帯端末を本装置と接続し、探そうとしている人物の顔が映った動画像をクエリ動画像として本装置に転送しても構わない。なお、本装置がインターネットに接続されている場合には、メールとして受信しても構わない。また、場合によっては、本システムに蓄積保存された動画像中から、検索する人物を目視で探し出し、それを指定しても構わない。この場合、その人物の他の場所（時間）での行動を検索することと等価となる。

10

いずれにせよ、クエリ動画像を入力した場合、クエリ人物の顔を特定する処理を行う。例えば、顔検出部207にてクエリ映像中に存在する全顔を検出し、検出した顔に矩形枠を付加して表示装置に表示し、検索依頼人にその中の1つを検索対象の顔画像（クエリ顔画像）と選択してもらう。なお、検出した顔が1つだけである場合には、無条件にその顔が選択されたものとして扱っても構わない。クエリ顔画像が決定されると、そのクエリ顔画像を追跡し、顔向きパターン番号に分類し、各顔向きパターン番号の若い順番から1つずつ特徴量を決定することで計 N_q 個の特徴量を決定する。なおクエリ特徴量の個数 N_q は、多いほどよいが、多くなるほど照合に時間がかかる。ここでは、照合DBへの登録と同様に $N_q = 4$ とする。

20

【0064】

次に、ステップS702にて、制御処理部211は、照合DBに登録されているオブジェクトIDの総個数 M （レコード数）を得る。制御処理部211は、取得した M 、クエリ顔画像の N_q 個の特徴量を引数にして、顔画像特徴照合部210に照合処理を要求する。

顔画像特徴照合部210は、ステップS703にて、変数 i をゼロに初期化する。そして、ステップS704にて、変数 i が M 以下であるか否かを判定する。 $i < M$ の場合、全レコードとの照合が済んでいないことになるので、処理はステップS706に進む。このステップS706では、顔画像特徴照合部210は、照合DBの i 番目のレコードを読み込み、その中に格納されている N_i 個の特徴量を読み込む。実施形態では、 N_i 個は最大で4であるのは既に説明した。そして、ステップS707にて、 N_q 個のクエリ特徴量と、 N_i 個の特徴量それぞれとの距離を総当たりで計算する。そして、ステップS708にて、その中の最少距離を、 i 番目のレコードのオブジェクトIDの人物顔とクエリ像顔画像との距離として決定する。顔画像特徴照合部210は、オブジェクトIDと決定した距離とをペアにして、RAM103の所定エリアに記憶する。この後、次のレコードとの照合処理を行うため、変数 i を“1”だけ増加させ、ステップS704に処理を戻す。

30

【0065】

以上の処理を繰り返していき、全レコードとの比較処理を終えたとき（変数 $i = M$ と判断されたとき）、RAM103には、全ての登録済みのオブジェクトIDに対する距離が格納されていることになる。そこで、顔画像特徴照合部210は、ステップS705にて、RAM103に記憶された距離を小さい順にソートし、本処理を終える。

40

【0066】

以上の結果、本第2の実施形態によれば、精度の劣化を抑えながら登録する情報量の削減が可能となっただけでなく、検索時も登録時と同様、複数の代表顔を用いる事で、クエリと登録顔で同じ顔向きで照合する確率を増やす事が可能となる。

【0067】

[第3の実施形態]

50

本第3の実施形態では、クエリとして、複数のカメラにより同一人物を撮影する例である。人物検索の精度は顔向きに依存する事から、異なる画角でカメラを複数台設置し、複数の顔向き画像を得る。

【0068】

この場合、クエリ人物の顔を、複数の異なる方向から撮影するので、一度の撮影で複数の顔向きパターンを得ることが約束されることになる。本処理では、撮影した個々の画像中のクエリ人物とされる同一人物を指定する構成を有し、それぞれの顔の特徴量と顔の向きを得ればよく、これまでの説明から明らかであるので、説明は省略する。なお、撮影する人物が一人であることが約束されているのであれば、人物を指定する構成を省略してもよい。

10

【0069】

本第3の実施形態によれば、精度の劣化を抑えながら登録する情報量の削減が可能となった。更に、検索時に、複数の画角の異なるカメラによりクエリ人物を撮影する事により、複数の代表顔を用いる事で、クエリと登録顔で同じ顔向きで照合する確率を増やす事が可能となる。

【0070】

[第4の実施形態]

本第4の実施形態では、クエリ人物の顔として、複数のカメラにより同一人物を動画撮影する例である。

【0071】

20

人物検索の精度は顔向きに依存する事から、異なる画角でカメラを複数台設置し、複数の顔向きを含んだ動画を得る。それ故、カメラの個数をM個とするなら、M個の動画画像が得られる。第2の実施形態では、1個の動画画像からクエリ人物の顔の特徴量を得るものであったが、本第4の実施形態では、M個の動画画像を入力し、そのM個の動画画像について顔向きパターン番号の集計を行ってNq個のクエリ特徴量を求めればよい。つまり、図9のS91をM回行えばよいので、その説明は省略する。

【0072】

以上の結果、精度の劣化を抑えながら登録する情報量の削減が可能となる。更に、検索時に、複数の画角の異なるカメラによりクエリ人物を動画撮影する事により、非常に多くの代表顔を用いる事で、クエリと登録顔で同じ顔向きで照合する確率を増やす事が可能となる。

30

【0073】

(その他の実施例)

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア(プログラム)を、ネットワーク又は各種記憶媒体を介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ(またはCPUやMPU等)がプログラムを読み出して実行する処理である。

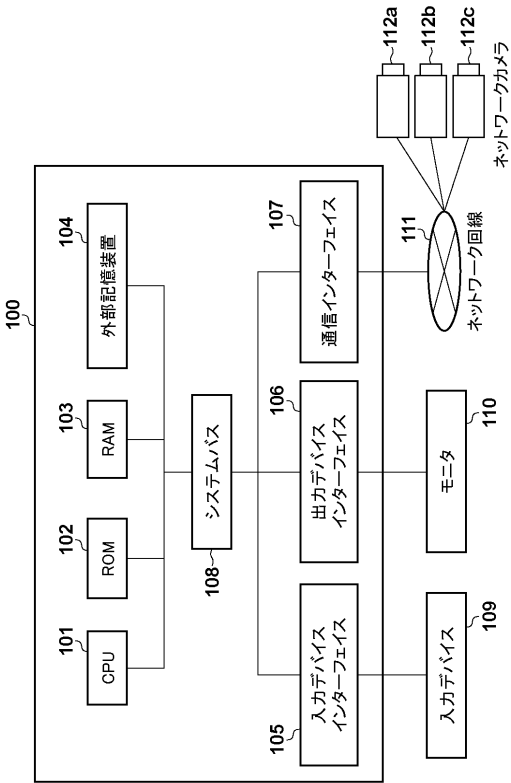
【符号の説明】

【0074】

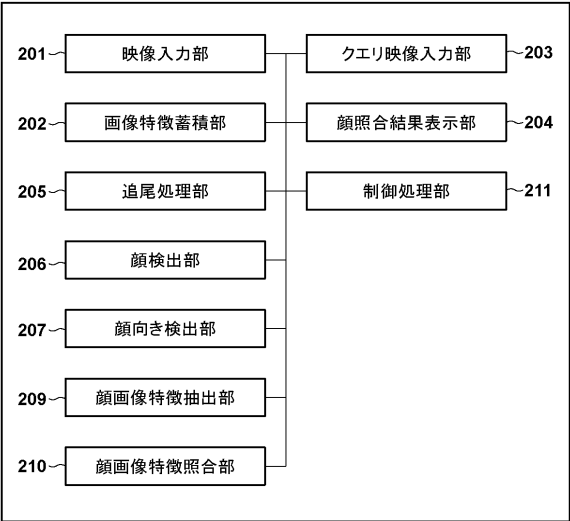
40

201...映像入力部、202...画像特徴蓄積部、203...クエリ映像入力部、204...顔照合結果表示部、205...追尾処理部、206...顔検出部、207...顔向き検出部、209...顔画像特徴抽出部、210...顔画像特徴照合部、211...制御処理部

【 図 1 】



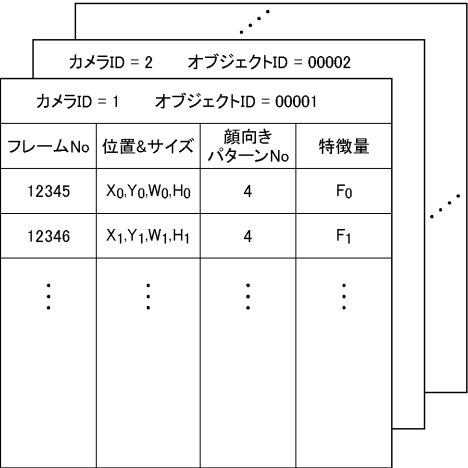
【 図 2 】



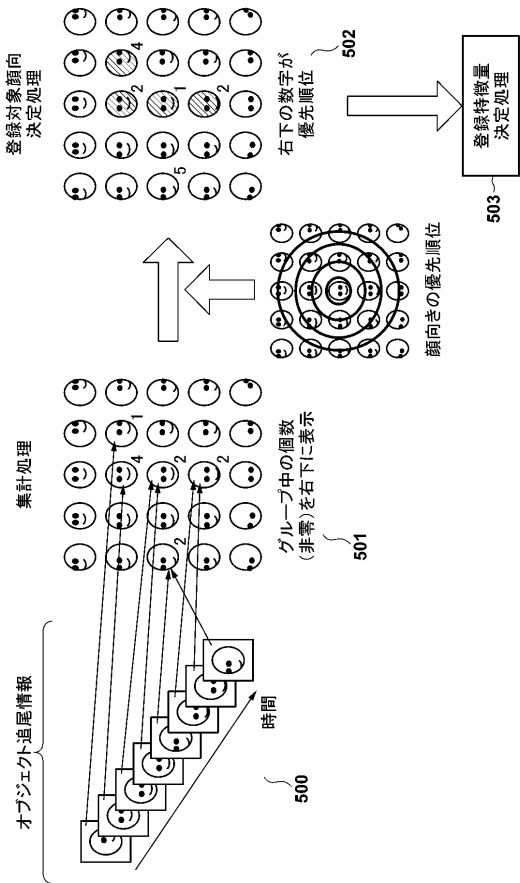
【 図 3 】



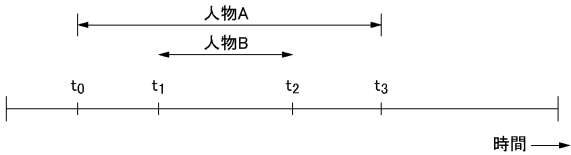
【 図 4 】



【図5】



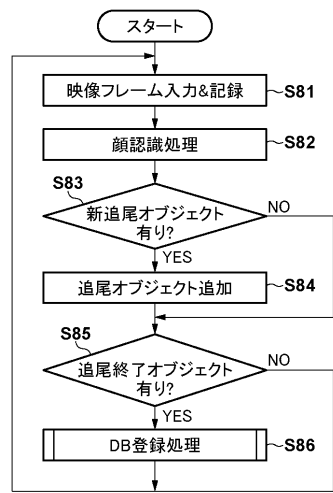
【図6】



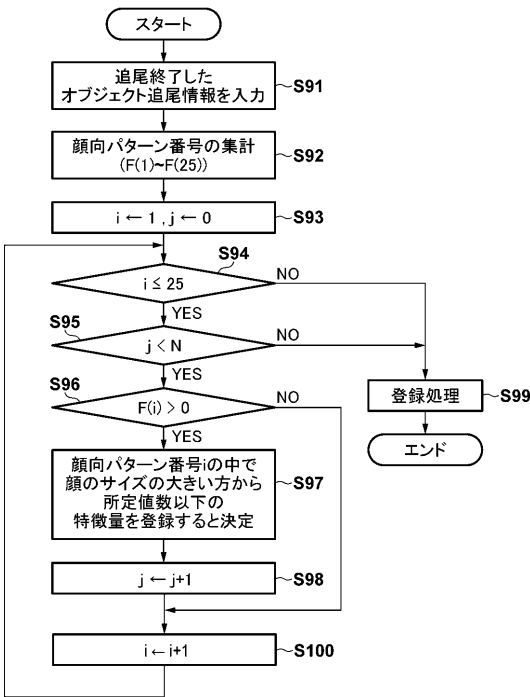
【図7】

オブジェクトID	特徴量	カメラID	フレームNo.位置,サイズ
00001	XXXXXXXX	1	XXXXXXXXXX
00002	XXXXXXXX	1	XXXXXXXXXX
00003	XXXXXXXX	2	XXXXXXXXXX
⋮	⋮	⋮	⋮

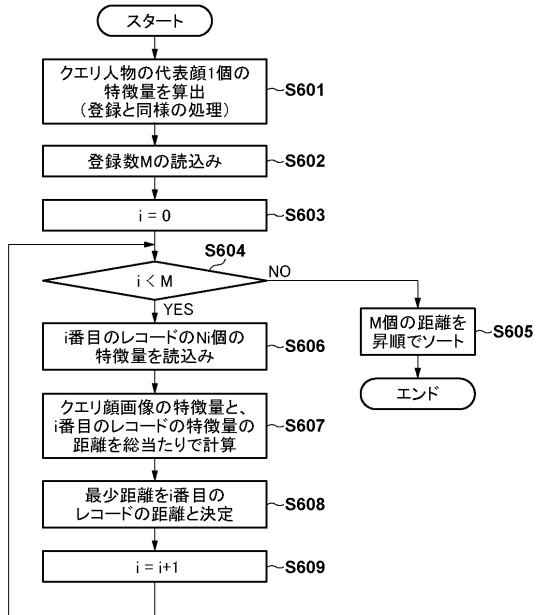
【図8】



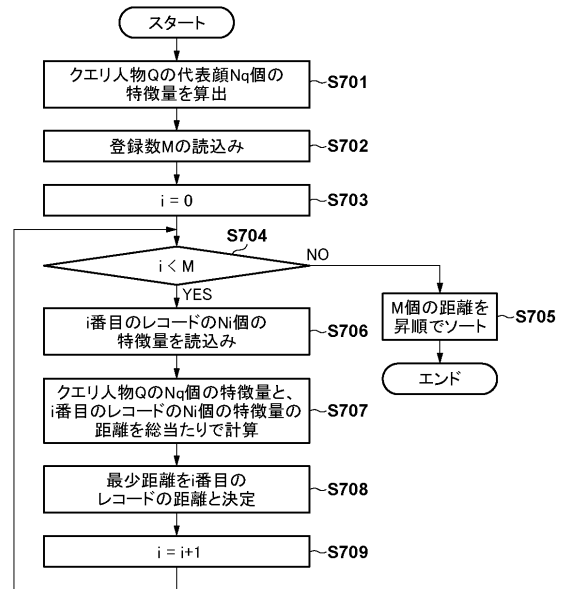
【図9】



【図 10】



【図 11】



フロントページの続き

(72)発明者 椎山 弘隆

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

(72)発明者 松下 昌弘

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) 5B043 AA09 BA04 CA09 DA05 EA05 EA08 FA02 FA07 GA02 HA02
HA08
5L096 CA04 CA05 DA01 FA18 FA59 FA67 HA05 HA07 JA11 KA15