

明 細 書

発明の名称： 情報処理装置、制御方法、及びプログラム

技術分野

[0001] 本発明は情報処理装置、制御方法、及びプログラムに関する。

背景技術

[0002] 監視カメラの映像を解析することで不審な物体を検出する技術が開発されている。例えば特許文献1や特許文献2は、静止している状態が継続する物体を置き去られた不審物として検出し、その物体を置き去った人物を推定する技術を開示している。

先行技術文献

特許文献

[0003] 特許文献1：特開2011-049646号公報

特許文献2：特開2012-235300号公報

発明の概要

発明が解決しようとする課題

[0004] 前述した特許文献1や特許文献2では、継続して静止している物体が、置き去られた物体として検出される。しかしながら、静止している物体が必ずしも置き去られた物体であるとは限らない。例えば空港のロビーにおいて目的の飛行機を待っている客が、周辺に荷物を置いた状態で椅子に座り続けているといったケースがある。静止している物体を置き去られた物体として検出してしまうと、このような荷物も置き去られた物体として誤検出されてしまう。

[0005] なお、特許文献1には、移動物体が近傍に存在する物体については不審物として検出しないようにすることで、単に足下に置かれただけの物体を不審物として検出しないようにするという技術も開示されている。しかしながらこの方法では、上述のように荷物を足下に置いた状態で椅子に座り続けるケースのように、荷物の持ち主が移動しないケースでは、その荷物が置き去ら

れた物体として誤検出されてしまう。

[0006] また、どちらの特許文献にも、アラートの確認方法については明確に記載されていない。置き去られた物体だけを矩形で囲んで表示しても、誰がどのように置き去ったのかといった情報を容易に把握することは難しい。

[0007] 本発明は、上記の課題に鑑みてなされたものである。本発明の目的の一つは、置き去られた物体を検知した後に、その物体が置き去られた状況を容易に把握することを可能とする技術を提供することである。

課題を解決するための手段

[0008] 本発明の情報処理装置は、１）動画データから置き去り物体を検出する置き去り物体検出手段と、２）前記置き去り物体が置かれた時刻である置き去り時刻を推定する置き去り時刻推定手段と、３）前記推定された置き去り時刻付近の動画データを取得し、前記推定された置き去り時刻に基づいて前記取得した動画データの提示方法を制御して、前記取得した動画データの少なくとも一部を提示する提示制御手段と、を有する。

[0009] 本発明の制御方法は、コンピュータによって実行される。当該制御方法は、１）動画データから置き去り物体を検出する置き去り物体検出ステップと、２）前記置き去り物体が置かれた時刻である置き去り時刻を推定する置き去り時刻推定ステップと、３）前記推定された置き去り時刻付近の動画データを取得し、前記推定された置き去り時刻に基づいて前記取得した動画データの提示方法を制御して、前記取得した動画データの少なくとも一部を提示する提示制御ステップと、を有する。

[0010] 本発明のプログラムは、本発明の制御方法が有する各ステップをコンピュータに実行させる。

発明の効果

[0011] 本発明によれば、だれがどのように物体を置き去ったのかという、物体の置き去り状況を把握しやすい映像の提示方法が提供される。

図面の簡単な説明

[0012] 上述した目的、およびその他の目的、特徴および利点は、以下に述べる好

適な実施の形態、およびそれに付随する以下の図面によってさらに明らかになる。

[0013] [図1]本実施形態の情報処理装置の動作の概要を例示する図である。

[図2]実施形態1の情報処理装置の構成を例示する図である。

[図3]情報処理装置を実現するための計算機を例示する図である。

[図4]実施形態1の情報処理装置によって実行される処理の流れを例示するフローチャートである。

[図5]動画データから静止物体を検出する方法を概念的に示す図である。

[図6]実施形態2の置き去り物体検出部の構成を例示する図である。

[図7]実施形態2の置き去り物体検出部によって実行される処理の流れを例示するフローチャートである。

[図8]動画フレームの一部の画像領域として定められる周辺領域を例示する図である。

[図9]置き去り判定部によって実行される置き去り判定処理の流れを例示する図である。

発明を実施するための形態

[0014] 以下、本発明の実施の形態について、図面を用いて説明する。尚、すべての図面において、同様な構成要素には同様の符号を付し、適宜説明を省略する。また、特に説明する場合を除き、各ブロック図において、各ブロックは、ハードウェア単位の構成ではなく、機能単位の構成を表している。

[0015] [実施形態1]

<概要>

置き去り物体を検出する方法として、動画データから静止物体を検出し、静止物体を置き去り物体として扱う方法がある。そして、静止物体を検出する方法として、動画データを構成する各動画フレームから物体を検出し、その物体が静止している場合に、その物体を静止物体として検出する方法がある。ここで、「物体が静止している」とは、物体の位置の変化が小さい（所定の閾値以下の値）状態が継続することを意味する。よって、物体の位置の

変化が小さい状態が継続した場合に、その物体が静止物体として検出される。

[0016] しかし、置き去り物体が検知された時刻は、そのものが置き去られた時刻よりもある一定時間経過しており、置き去った人物は同じフレーム内に映っていない。このため、どのような人物がどのように物体を置き去ったのかをすぐに把握することが困難である。

[0017] そこで本実施形態の情報処理装置は、動画データから置き去り物体が検出された場合に、さらなる処理を行う。図1は、本実施形態の情報処理装置（後述する図2における情報処理装置2100）の動作の概要を例示する図である。図1は、情報処理装置2100の理解を容易にするための例示であり、情報処理装置2100の動作を限定するものではない。

[0018] 情報処理装置2100は、動画データ12から置き去り物体を検出する。さらに情報処理装置2100は、その静止物体が置き去られた時刻を推定する。以下、推定された置き去り時刻を置き去り推定時刻と呼ぶ。そして、情報処理装置2100は、動画データ12のうち、置き去り推定時刻付近の画像（静止画像又は動画画像）を提示する。ここで、「置き去り推定時刻付近」とは、置き去り推定時刻に近いことを意味する。例えば、「置き去り推定時刻付近」は、置き去り推定時刻よりも所定時間前の時刻以降であって、なおかつ置き去り推定時刻よりも所定時間後の時刻以前の期間に含まれることを意味する。

[0019] ここで、情報処理装置2100は、画像の提示方法を、置き去り推定時刻に基づいて制御する。後述するように、例えば画像提示を制御する処理には、置き去り推定時刻付近の動画データ12をその他の期間の動画データ12よりも遅い再生速度で再生するように、速度を変えて動画データ12を再生する処理や、置き去り推定時刻付近の動画データ12を静止画に展開して提示し、その際に、置き去り推定時刻付近はそれ以外の期間よりも密に静止画を抽出して表示する処理などが含まれる。

[0020] 本実施形態の情報処理装置2100によれば、置き去り推定時刻付近の画

像が、置き去り推定時刻に基づいて制御された態様で提示される。このように提示される画像を見ることで、置き去り物体が置き去られた状況を容易に把握することができる。

[0021] 以下、本実施形態の情報処理装置 2100 についてさらに詳細に説明する。

[0022] <情報処理装置 2100 の機能構成の例>

図 2 は、実施形態 1 の情報処理装置 2100 の構成を例示する図である。情報処理装置 2100 は、置き去り物体検出部 2120、置き去り時刻推定部 2140、及び提示制御部 2160 を有する。置き去り物体検出部 2120 は、動画データ 12 から置き去り物体を検出する。置き去り時刻推定部 2140 は、置き去り物体として検出された物体（対象物体）が置き去られた時刻である置き去り時刻を推定する。提示制御部 2160 は、推定された置き去り時刻（すなわち、置き去り推定時刻）付近の動画データ 12 を取得し、取得した動画データ 12 の少なくとも一部を提示する。この提示は、置き去り推定時刻に基づいて制御される。

[0023] <情報処理装置 2100 のハードウェア構成>

情報処理装置 2100 の各機能構成部は、各機能構成部を実現するハードウェア（例：ハードワイヤードされた電子回路など）で実現されてもよいし、ハードウェアとソフトウェアとの組み合わせ（例：電子回路とそれを制御するプログラムの組み合わせなど）で実現されてもよい。以下、情報処理装置 2100 の各機能構成部がハードウェアとソフトウェアとの組み合わせで実現される場合について、さらに説明する。

[0024] 図 3 は、情報処理装置 2100 を実現するための計算機 1000 を例示する図である。計算機 1000 は任意の計算機である。例えば計算機 1000 は、Personal Computer (PC)、サーバマシン、タブレット端末、又はスマートフォンなどである。計算機 1000 は、情報処理装置 2100 を実現するために設計された専用の計算機であってもよいし、汎用の計算機であってもよい。

- [0025] 計算機１０００は、バス１０２０、プロセッサ１０４０、メモリ１０６０、ストレージデバイス１０８０、入出力インタフェース１１００、及びネットワークインタフェース１１２０を有する。バス１０２０は、プロセッサ１０４０、メモリ１０６０、ストレージデバイス１０８０、入出力インタフェース１１００、及びネットワークインタフェース１１２０が、相互にデータを送受信するためのデータ伝送路である。ただし、プロセッサ１０４０などを互いに接続する方法は、バス接続に限定されない。プロセッサ１０４０は、CPU (Central Processing Unit) や GPU (Graphics Processing Unit) などの演算装置である。メモリ１０６０は、RAM (Random Access Memory) などを用いて実現される主記憶装置である。ストレージデバイス１０８０は、ハードディスク、SSD (Solid State Drive)、メモリカード、又は ROM (Read Only Memory) などを用いて実現される補助記憶装置である。ただし、ストレージデバイス１０８０は、RAM など、主記憶装置を構成するハードウェアと同様のハードウェアで構成されてもよい。
- [0026] 入出力インタフェース１１００は、計算機１０００と入出力デバイスとを接続するためのインタフェースである。ネットワークインタフェース１１２０は、計算機１０００を通信網に接続するためのインタフェースである。この通信網は、例えば LAN (Local Area Network) や WAN (Wide Area Network) である。ネットワークインタフェース１１２０が通信網に接続する方法は、無線接続であってもよいし、有線接続であってもよい。
- [0027] 例えば計算機１０００は、ネットワークを介してカメラ１０やビデオ管理サーバ１０１０と通信可能に接続されている。ただし、計算機１０００をカメラ１０やビデオ管理サーバ１０１０と通信可能に接続する方法は、ネットワークを介した接続に限定されない。また、計算機１０００は、カメラ１０と通信可能に接続されていなくてもよい。
- [0028] ストレージデバイス１０８０は、情報処理装置２１００の各機能構成部（置き去り物体検出部２１２０、置き去り時刻推定部２１４０、及び提示制御部２１６０）を実現するプログラムモジュールを記憶している。プロセッサ

１０４０は、これら各プログラムモジュールをメモリ１０６０に読み出して実行することで、各プログラムモジュールに対応する機能を実現する。

[0029] なお、計算機１０００は、複数の計算機を利用して実現されてもよい。例えば置き去り物体検出部２１２０、置き去り時刻推定部２１４０、及び提示制御部２１６０をそれぞれ異なる計算機で実現することができる。この場合、各計算機のストレージデバイスに記憶されるプログラムモジュールは、その計算機で実現される機能構成部に対応するプログラムモジュールだけでもよい。

[0030] ＜カメラ１０について＞

カメラ１０は、繰り返し撮像を行って動画データ１２を生成できる任意のカメラである。例えばカメラ１０は、特定の施設や道路などを監視するために設けられている監視カメラである。

[0031] 情報処理装置２１００の機能の一部又は全部は、カメラ１０によって実現されてもよい。すなわち、情報処理装置２１００を実現するための計算機１０００として、カメラ１０を利用してもよい。また、ビデオ管理サーバ１０１０の機能もこの中に含まれていてもよい。この場合、カメラ１０は、自身で生成した動画データ１２を蓄積、処理する。このように情報処理装置２１００を実現するために用いられるカメラ１０には、例えば、インテリジェントカメラ、ネットワークカメラ、又は IP (Internet Protocol) カメラなどと呼ばれるカメラを利用できる。

[0032] ＜ビデオ管理サーバ１０１０について＞

ビデオ管理サーバ１０１０は、カメラ１０、計算機１０００と通信可能に接続されており、カメラ１０で生成された動画データ１２を蓄積し、要求に応じて動画データ１２を配信する機能を有する。計算機１０００から動画データ１２の配信要求があった場合、カメラ１０で生成された動画データ１２のうち、指定された時刻の動画データ１２を計算機１０００へ配信する。この際、計算機１０００は、カメラ１０から直接動画データ１２を取得して処理するようになっていてもよいし、あるいは、カメラ１０で取得した動画デ

ータ１２はビデオ管理サーバ１０１０に蓄積され、計算機１０００は、ビデオ管理サーバにアクセスして動画データ１２を取得し、処理するようになっていてもよい。

[0033] ビデオ管理サーバ１０１０は、情報処理装置２１００と共に計算機１０００を用いて実現されていてもよい。この場合、情報処理装置２１００とビデオ管理サーバ１０１０とは、互いにバス１０２０を介してデータをやり取りする。

[0034] ＜処理の流れ＞

図４は、実施形態１の情報処理装置２１００によって実行される処理の流れを例示するフローチャートである。置き去り物体検出部２１２０は、動画データ１２から置き去られた物体を検出する（Ｓ１１０）。置き去り時刻推定部２１４０は、物体を検出した時刻に基づいて、その物体が置き去られた置き去り時刻の推定を行う（Ｓ１１２）。提示制御部２１６０は、置き去り推定時刻付近の動画データ１２を取得し、置き去り推定時刻に基づいて、動画データ１２の制御及び提示を行う（Ｓ１１４）。

[0035] ＜動画データ１２の取得方法＞

情報処理装置２１００は、処理対象とする動画データ１２を取得する。情報処理装置２１００が動画データ１２を取得する方法は様々である。例えば情報処理装置２１００は、カメラ１０から送信される動画データ１２を受信する。また例えば、情報処理装置２１００は、カメラ１０にアクセスし、カメラ１０に記憶されている動画データ１２を取得する。

[0036] また、カメラ１０によって生成された動画データ１２は、ビデオ管理サーバ１０１０にも送られ、ここで動画データ１２が蓄積される。

[0037] なお、カメラ１０によって生成された動画データ１２をまずビデオ管理サーバ１０１０に動画データ１２を蓄積し、その後、蓄積された動画データ１２を計算機１０００から読み出して処理を行うようにしてもよい。

[0038] 情報処理装置２１００の一部又は全部の機能がカメラ１０で実現される場合、情報処理装置２１００は、情報処理装置２１００自身によって生成され

た動画データ 12 を取得する。この場合、動画データ 12 は、例えば情報処理装置 2100 の内部にある記憶装置（例えばストレージデバイス 1080）に記憶されている。そこで情報処理装置 2100 は、これらの記憶装置から動画データ 12 を取得する。

[0039] 情報処理装置 2100 が動画データ 12 を取得するタイミングは任意である。例えば情報処理装置 2100 は、カメラ 10 によって動画データ 12 を構成する新たな動画フレーム 14 が生成される度に、その新たに生成された動画フレーム 14 を取得することで、動画データ 12 を取得する。その他にも例えば、情報処理装置 2100 は、定期的に未取得の動画フレーム 14 を取得してもよい。例えば情報処理装置 2100 が 1 秒間に 1 回動画フレーム 14 を取得する場合、情報処理装置 2100 は、1 秒間に生成される複数の動画フレーム 14（例えば動画データ 12 のフレームレートが 30fps(frames/second)であれば、30枚の動画フレーム 14）をまとめて取得する。

[0040] 置き去り物体検出部 2120 は、動画データ 12 を構成する全ての動画フレーム 14 を取得してもよいし、一部の動画フレーム 14 のみを取得してもよい。後者の場合、例えば置き去り物体検出部 2120 は、カメラ 10 によって生成される動画フレーム 14 を、所定数に 1 つの割合で取得する。

[0041] <置き去り物体の検出：S110>

置き去り物体検出部 2120 は、動画データ 12 から静止物体を検出し、それを置き去り物体とする（S110）。動画データから静止物体を検出する技術には、既存の技術を利用することができる。なお、ここでは静止物体を置き去り物体とみなす例について述べたが、置き去り物体の検知はそれに限らない。置き去り物体の検知をより高精度に行う方式については後述する。以下、動画データ 12 から静止物体を検出する方法の一例を説明する。

[0042] 置き去り物体検出部 2120 は、動画データ 12 を構成する各動画フレーム 14 から物体を検出する。置き去り物体検出部 2120 は、同じ物体が含まれる複数の動画フレーム 14 それぞれについて、その物体の位置の変化量を算出する。そして置き去り物体検出部 2120 は、その物体の位置の変化

量が所定量以下である状態が所定長の期間継続した場合に、その物体を静止物体として検出する。ここで、置き去り物体検出部 2120 によって或る物体が静止物体として検出された場合において、その物体の位置の変化量が所定量以下である状態が継続した上記所定長の期間を、静止判定期間と呼ぶ。

[0043] 図 5 は、動画データ 12 から静止物体を検出する方法を概念的に示す図である。この例において、静止物体検出部 2020 は、物体の位置の変化量が所定値以下である期間の長さが p 以上である場合に、その物体を静止物体として検出する。

[0044] 図 5 の動画データ 12 では、時点 t 以降の各動画フレーム 14 から物体 20 が検出される。静止物体検出部 2020 は、時系列で隣接する 2 つの動画フレーム 14 の各組み合わせについて物体 20 の位置の差分を算出することで、物体 20 の位置の変化量を算出する。

[0045] まず時点 t において、物体 20 を持った人物 30 が現れる。人物 30 は、時点 $t+a$ まで、物体 20 を持って移動している。そのため、時点 t から $t+a$ までの期間について算出される物体 20 の位置の変化量は、いずれも所定量より大きい値となる。そのため、この期間に生成された動画フレーム 14 を用いた判定では、物体 20 は、静止物体として検出されない。

[0046] その後、時点 $t+a$ 以降、物体 20 は地面に置かれている。そのため、時点 $t+a$ 以降の各動画フレーム 14 から検出される物体 20 の位置の変化量は、いずれも所定量以下となる。したがって、置き去り物体検出部 2120 は、時点 $t+a$ から時点 $t+a+p$ の期間に生成された動画フレーム 14 を用いた判定により、物体 20 を静止物体として検出する。

[0047] その他にも例えば、置き去り物体検出部 2120 は、動画フレーム 14 を用いて背景画像（背景のみからなる画像）を生成し、背景差分によって静止物体を検出してもよい。まず、置き去り物体検出部 2120 は、時系列の動画フレーム 14 から背景画像を生成する。背景画像の生成には、既知の様々な手法を用いることができる。次に、置き去り物体検出部 2120 は、新たに取得する動画フレーム 14 と背景画像との差分を算出する。そして、置き

去り物体検出部 2120 は、差分が大きくなる領域を抽出する。置き去り物体検出部 2120 は、動画フレーム 14 を取得するたびにこの処理を実行し、得られた抽出結果をフレーム間で比較し、物体の位置の変化量が所定量以内に収まっているかどうかを判定する。

[0048] 図 5 のケースにおいて、上述の背景差分を利用するとする。この場合、時点 t から $t+a$ までの期間について算出される物体 20 の位置の変化量は所定量より大きい値となる。一方、時点 $t+a$ 以降では、物体 20 の位置の変化量が所定量以下となる。よって、背景差分を利用する場合においても、時点 $t+a$ から時点 $t+a+p$ の期間に生成された動画フレーム 14 を用いた判定により、物体 20 が静止物体として検出される。なお、背景差分で用いる背景画像は複数あってもよい。例えば、情報処理装置 2100 は、特許文献 W02014/192441 に記載されているような、複数の時間周期で作成した背景画像を用いて背景差分を計算し、静止物体を判定するように構成されていてもよい。

[0049] なお、物体の移動を判定する方法は、上記の方法に限定されない。例えば、「物体中の特徴点を抽出し、特徴点を隣接フレーム間で対応付けることによって、物体の移動量を求める」という方式など、様々な方式を用いることができる。

[0050] <置き去り時刻の推定：S112>

次に、置き去り時刻推定部 2140 が、物体が置き去られた時刻の推定を行う (S112)。上述のように、静止物を置き去り物として検知する場合、物体が静止してから一定時間経過後、その物体が静止物として検知される。よって、基本的には、検知時刻 $t+a+p$ から静止判定期間 p を引いた時刻 $t+a$ が置き去り推定時刻となる。ただし、時刻 $t+a$ が必ずしも正確な置き去り時刻ではない場合もある。人が往来する場所などがカメラ 10 によって撮像される場合、対象物体が静止していると判定され始める時点は、対象物体が実際に置かれた時点と一致するとは限らないためである。よって、置き去り時刻推定部 2140 は、その前後の時刻の動画データ 12 を探索し、置き去り推定時刻を求めるようにしてもよい。

[0051] 例えば、置き去り時刻推定部 2140 は、 $t+a+p$ の時刻に検知された静止物領域と同じ領域の特徴量を時刻 $t+a$ のフレームから抽出する。また、置き去り時刻推定部 2140 は、背景画像においても、同じ領域の特徴量を抽出する。時刻 $t+a$ のフレームから抽出した特徴量が、背景のものと同一とみなせる場合（例えば、特徴量間の類似性が一定値以上であるか、あるいは特徴量間の距離が一定値以下の場合）には、その時点ではまだ物体が置かれていなかった可能性が高いため、置き去り時刻推定部 2140 は、それより後の時刻のフレームを順次探索し、背景と同じでない特徴量が抽出されるフレームを見つけた時点で、その時刻を置き去り推定時刻としてもよい。逆に、時刻 $t+a$ のときに同じ領域の特徴量を抽出したときに、背景と異なる場合には、それ以前に物体が置かれた可能性が高いため、置き去り時刻推定部 2140 は、それより前の時刻のフレームを順次探索し、背景と同じ特徴量が抽出されるフレームを見つけた時点で、その時刻を置き去り推定時刻としてもよい。

[0052] あるいは、置き去り時刻推定部 2140 は、 $t+a+p$ の時刻に検知された静止物領域の特徴量を抽出し、時刻 $t+a$ のときに同じ領域の特徴量を抽出し、これらを比較してもよい。この場合は、時刻 $t+a+p$ のときの特徴量と時刻 $t+a$ のときの特徴量が同一と見なせる場合（例えば、特徴量間の類似性が一定値以上であるか、あるいは特徴量間の距離が一定値以下の場合）には、その時点ですでに物体が置かれている可能性が高いため、置き去り時刻推定部 2140 は、それより前の時刻のフレームを順次探索し、時刻 $t+a+p$ のときと同じでない特徴量が抽出されるフレームを見つけた時点で、その時刻を置き去り推定時刻としてもよい。逆に、時刻 $t+a+p$ のときの特徴量と時刻 $t+a$ のときの特徴量が異なる場合には、それ以降に物体が置かれた可能性が高いため、置き去り時刻推定部 2140 は、それより後の時刻のフレームを順次探索し、背景と同じでない特徴量が抽出されるフレームを見つけた時点で、その時刻を置き去り推定時刻としてもよい。

[0053] <画像の提示 : S114>

次に、提示制御部 2160 が、置き去り推定時刻付近の動画データ 12 をビデオ管理サーバから取得し、取得した動画データ 12 を、ユーザ（オペレータ）に対して状況を把握しやすい形に制御して提示する（S114）。「置き去り推定時刻付近」の定め方は任意である。例えば前述したように、置き去り推定時刻付近は、置き去り推定時刻よりも所定時間前の時刻以降であって、なおかつ置き去り推定時刻よりも所定時間後の時刻以前である期間として定められる。この場合、提示制御部 2160 は、動画データ 12 を構成する動画フレームのうち、置き去り推定時刻よりも所定時間前の時刻に生成された動画フレームから、置き去り推定時刻よりも所定時間後の時刻に生成された動画フレームまでの各動画フレームを取得する。

[0054] 提示制御部 2160 は、置き去り推定時刻に基づいて画像（取得した動画データ 12 の少なくとも一部）の提示方法を制御する。以下では、その制御方法の具体例について述べる。

[0055] <<（１）置き去り推定時刻付近の動画を再生>>

提示制御部 2160 は、置き去り推定時刻付近の動画データ 12 に関する情報をビデオ管理サーバ 1010 から取得する。例えば提示制御部 2160 は、置き去り推定時刻より所定時間前から、置き去り物体が検知された時刻までの動画データ 12 を取得する。次に、提示制御部 2160 は、この動画データ 12 を再生して表示する。例えば再生表示は、ボタンによって制御される。この場合、提示制御部 2160 は、ユーザによって所定のボタンが押されたことを検知したことに応じて、動画データ 12 の再生を開始する。その他にも例えば、提示制御部 2160 は、置き去り物体が検出されたことに応じて、自動的に動画データ 12 を取得して再生してもよい。

[0056] この際、提示制御部 2160 は、動画データ 12 の再生速度を、再生している動画データ 12 の時刻と置き去り推定時刻の時間間隔に応じて制御することが好適である。置き去り推定時刻付近には、実際に置き去った人物が映っている可能性が高い。一方、置き去り推定時刻から検知時刻までは、検知された物体自体は静止したままであるため、早送りで再生しても、物体の視

認性は落ちにくい。そこで提示制御部 2160 は、置き去り推定時刻に近い時刻（例えば、置き去り推定時刻との差が所定値以下の時刻）付近の動画データ 12 の再生速度を、それ以外の期間の動画データ 12 の再生速度よりも遅くする。例えば、前者の再生速度は通常の再生速度や通常よりも遅い（例えば 0.5 倍）再生速度とし、後者の再生速度は通常よりも速い（例えば通常の 2 倍）の再生速度とする。これにより、どのような人物がどのように物を置き去ったかという重要な部分を重点的に確認できるようになり、置き去り時点の状況把握がしやすくなる。

[0057] また、この再生方法は、時間の順方向ではなく、時間の逆方向に再生するようになっていてもよい。すなわち、提示制御部 2160 は、置き去り物体が検知された時刻付近から置き去り推定時刻の方に向けて動画データ 12 を再生してもよい（いわゆる逆再生）。これにより、各動画フレームは、その生成時点が遅い順に表示される。この場合も、置き去り推定時刻付近の動画データ 12 はゆっくり逆再生され、それ以外の時刻の動画データ 12 は早送り逆再生されるようにすることが好適である。これにより、検出された置き去り物体を中心に動画データ 12 を確認することができ、それに注視して動画データ 12 を確認すればよくなる。

[0058] 時間の順方向に再生する場合は、最初の時点では、まだ置き去り物体が画面に表れていないため、画面中に複数の人物が映っていた場合、どの人物に着目して動画データ 12 を確認すればよいかが特定できない。このため、複数の人物を同時に確認し続ける必要が生じ、状況の把握がしにくくなる。

[0059] これに対し、逆方向に再生する場合には、すでに置き去られた物体が画面中に映っており、そこに着目してその物体に逆方向に近づいてくる人物のみに着目すればよくなるため、注視すべき人物を絞り込みやすくなる。よって、状況の確認がしやすくなり、短時間で効率的に状況を把握し、置き去り人物を特定できるようになる。

[0060] <<（２）置き去り推定時刻付近の動画を静止画に展開して表示>>

別の画像提示方法として、動画データ 12 を静止画のフレーム（動画フレ

ーム)に展開し、画面上に並べて表示することが考えられる。ここでいう「展開」とは、動画データから、その動画データを構成する複数の動画フレームの内の1つ以上を抽出することを意味する。この際、展開する動画フレームの密度を、置き去り推定時刻付近については密にし(すなわち、展開する動画フレームの時間間隔を短くする)、それ以外の期間では疎にする(展開する動画フレームの時間間隔を長くする)ようにする。言い換えれば、置き去り推定時刻付近の動画データ12については、単位時間当たりの動画データ12から抽出される動画フレームの枚数を多くし、それ以外の期間の動画データ12については、単位時間当たりの動画データ12から抽出される動画フレームの枚数を少なくする。例えば、置き去り推定時刻付近の動画データ12については、1秒間の動画データ12ごとに5枚の動画フレームを抽出して画面に表示し、それ以外の期間の動画データ12については、1秒間の動画データ12ごとに1枚の動画フレームを抽出して画面に表示する。このように展開する動画フレームの密度を置き去り推定時刻に基づいて制御することにより、置き去った人が映っていると考えられる部分を重点的に静止画で提示できるようになり、誰がどのように物体を置き去ったかが把握しやすくなる。

[0061] また、動画データ12を動画フレームに展開して表示する際、提示制御部2160は、置き去り推定時刻に基づいて、表示する動画フレームの大きさを制御してもよい。例えば提示制御部2160は、置き去り推定時刻付近の動画フレームを、他の期間の動画フレームよりも大きく表示する。これにより、置き去った時刻付近の人物の挙動が把握しやすくなる。なお、動画フレームの枚数の制御と大きさの制御は、いずれか一方のみが行われてもよい、双方が行われてもよい。

[0062] <<(3) 展開した動画を一定時間ごとに切り替えて紙芝居風に提示>>

提示制御部2160は、上述の(2)の方法によって動画データ12から展開された動画フレームを、画面に並べる代わりに、時間方向に一定間隔で切りかえながら表示させてもよい。(2)で述べたように、置き去り推定時

刻付近の静止画像の密度が高くなっていれば、等時間間隔で静止画を切りかえて表示した際、推定時刻付近は、ゆっくり再生していることと等価になり、（１）と同様の効果が得られるようになる。

[0063] 現場にいる警備員が利用する端末は、携帯端末や AR (Augmented Reality) グラスなどのように、画面の大きさがそれほど大きくない端末であることもある。この場合、展開された動画フレームを（３）の方法で制御して提示することで、効率的に現場の警備員に状況を伝えることができるようになる。また、もともとの動画データに比べると、このように飛び飛びに選んだ静止画列の方がデータ容量は一般的に小さくなる。そのため、通信帯域幅が狭い通信環境下においてもデータ伝送が可能となる。なお、（１）で述べたように、この場合も、時間の逆方向に向かって画像を切りかえる（時間的に新しい方から古い方に向かって画像を順次提示する）ようにすることで、（１）の場合と同様に、より状況把握をしやすくなる。

[0064] 以上、３通りの画像提示制御方法について述べたが、いずれの場合も、提示制御部 2160 は、状況を視認しやすくするために、他の情報を重畳したり、強調表示したりしてもよい。例えば、画像を提示する際、対象となる物体を枠で囲って表示することが考えられる。その際、提示制御部 2160 は、枠を点滅表示させるようにしてもよい。また、提示制御部 2160 は、置き去り検知時刻付近では、枠の色を変えたり、枠を太くしたり、点滅表示させたり、あるいは音を鳴らしたりして、ユーザの注意を喚起するようにしてもよい。

[0065] あるいは、提示制御部 2160 は、対象物体に関する情報（以下、物体情報）を併せて表示してもよい。例えば物体情報には、対象物体を切り出して拡大表示した画像、対象物体の色や形に関する情報、対象物体が撮像された時点や期間、対象物体が含まれる動画フレーム 14 の識別子（フレーム番号）、及び対象物体を撮像したカメラ 10 の識別子などが含まれる。また、対象物体の大きさの情報等も付加して提示するようにしてもよい。物体の大きさについては、カメラのキャリブレーション情報を用いてその物体の画像領

域を実空間上の領域に変換することにより、推定することができる。対象物体の大きさを示す情報は、対象物体が爆弾等の危険物である場合に、その危険性を判定する上で有用である。

[0066] なお、上述の画像提示方法によって制御された置き去り推定時刻付近の画像の提示先は任意である。例えば画像の提示先は、警備員室でカメラ10の動画データ12を監視している監視員が使用している端末のディスプレイ装置である。その他にも例えば、置き去り推定時刻付近の画像の提示先は、現場で警備を行っている警備員が使用している端末のディスプレイ装置である。その他にも例えば、置き去り推定時刻付近の画像の提示先は、警備会社や警察などの所定の機関で使用されている端末のディスプレイ装置であってもよい。

[0067] [実施形態2]

<概要>

置き去り物体を検出する方法として、実施形態1で述べたように、動画データから静止物体を検出し、静止物体を置き去り物体として扱う方法がある。しかし、静止物体は必ずしも置き去り物体とは限らない。そこで、動画データから静止物体が検出された場合に、さらに人物検出を行い、検出された静止物体が置き去り物体かどうかを判定する。

[0068] 例えば、静止物体として検出された対象物体が、置き去られた物体ではなく、持ち主のそばに置かれている物体であるとする。このケースでは、対象物体の周辺に持ち主がいる状態が継続する。一方で、対象物体が置き去られた物体であるケースでは、対象物体の周辺から持ち主がいなくなる。そこで置き去り物体検出部2120は、これらのケースを区別するために、それぞれ異なる時点における対象物体が含まれる動画フレーム14について、対象物体の周辺から人物を検出する人物検出処理を行ってその結果を比較する。こうすることで、静止物体が検出されたそれを置き去り物体として検出する方法と比較し、置き去り物体を精度良く検出することができる。

[0069] 図6は、本実施形態の置き去り物体検出部2120の構成の概要を例示す

る図である。置き去り物体検出部 2120 は、静止物体検出部 2020、人物検出部 2040、置き去り判定部 2050 を有する。静止物体検出部 2020 は、実施形態 1 で述べたのと同様に、動画データ 12 から静止物体を検出する。人物検出部 2040 は、静止物体として検出された物体（対象物体）が含まれる複数の動画フレーム 14 それぞれを対象に、対象物体の周辺から人物を検出する人物検出処理を実行する。置き去り判定部 2050 は、静止物検出処理の結果と人物検出処理の結果から、検出された静止物が置き去り物体であるかどうかを判定する、置き去り判定処理を実行する。

[0070] <処理の流れ>

図 7 は、置き去り物体検出部 2120 の処理の流れを例示するフローチャートである。静止物体検出部 2020 は、動画データ 12 から静止物体を検出する（S102）。人物検出部 2040 は、対象物体が含まれる複数の動画フレーム 14 それぞれを対象に、対象物体の周辺から人物を検出する人物検出処理を実行する（S104）。置き去り判定部 2050 は、複数の動画フレームそれぞれを対象とした人物検出処理の結果を比較することで、検知された物体が置き去り物体であるかどうかを判定する（S106）。

[0071] <静止物体の検出：S102>

静止物体検出部 2020 は、動画データ 12 から静止物体を検出する（S102）。動画データ 12 から静止物体を検出する技術には、既存の技術を利用することができる。例えば、実施形態 1 で述べた置き去り物体検出（S110）の技術を用いることができる。

[0072] <人物検出処理：S104>

人物検出部 2040 は、対象物体（静止物体として検出された物体）が含まれる複数の動画フレーム 14 それぞれを対象に、対象物体の周辺から人物を検出する人物検出処理を行う（S104）。以下、動画フレーム 14 において、対象物体の周辺の画像領域を、周辺領域と呼ぶ。「対象物体が含まれる複数の動画フレーム 14」は、例えば前述した図 5 の例においては、時点 t 以降に生成された各動画フレーム 14 である。これらの動画フレーム 14

には、静止物体として検出された物体である物体 20 が含まれている。

[0073] 動画フレーム（すなわち画像データ）から人物を検出する方法には、既存の技術（例えば特徴量マッチングやテンプレートマッチングなど）を利用できる。例えば特徴量マッチングを利用する場合、人物検出部 2040 は、人間の外見的特徴を表す特徴量を持つ画像領域を動画フレーム 14 から検出することにより、人物検出を行う。

[0074] 人物検出処理は、人物全体を表す画像領域を検出する処理であってもよいし、人物の一部（例えば頭部）を表す画像領域を検出する処理であってもよい。ここで、カメラ 10 の撮像範囲に多くの人物や物体が含まれる場合、人物の一部が他の人物や物体によって隠されてしまう蓋然性が高い。この点、頭部は下半身などと比べて他の人物などに隠されてしまう蓋然性が低い部位（カメラ 10 によって撮像される蓋然性が高い部位）であると言える。また、頭部は個々の人の特徴を良く表す部位であると言える。そのため、人物検出処理として頭部を検出する処理を行うことには、1) 動画フレーム 14 から人物を検出できる蓋然性が高い、2) 他の人物との識別が容易な態様で人物を検出できるなどといった利点がある。

[0075] 周辺領域は、動画フレーム 14 全体の画像領域であってもよいし、動画フレーム 14 の一部の画像領域であってもよい。動画フレーム 14 の一部の画像領域を周辺領域とする場合、例えば対象物体の位置を基準とした所定形状の画像領域を周辺領域とする。ここで、物体の位置は、その物体を表す画像領域に含まれる任意の位置（例えば中心位置など）とすることができる。また、所定形状は、例えば所定のサイズを持つ円や矩形などとすることができる。周辺領域には、対象物体が含まれていてもよいし、含まれていなくてもよい。所定形状を定義する情報は、予め人物検出部 2040 に設定されていてもよいし、人物検出部 2040 からアクセス可能な記憶装置に記憶されていてもよい。

[0076] 図 8 は、動画フレーム 14 の一部の画像領域として定められる周辺領域を例示する図である。図 8 において、対象物体は物体 20 である。図 8 (a)

において、所定形状は、半径 d の円形である。また、周辺領域 40 の中心位置は、物体 20 の中心位置である。一方、図 8 (b) において、所定形状は、長辺が $d1$ であり、なおかつ短辺が $d2$ の矩形である。また、周辺領域 40 の中心位置は、物体 20 の中心位置から上方に所定距離 e 離れた位置である。

[0077] なお、周辺領域 40 の画像上の位置に応じ、周辺領域 40 を定める所定形状を変化させてもよい。例えば、カメラ 10 に近い場所では所定形状のサイズを大きくし、カメラ 10 から遠い場所ではカメラ 10 のサイズを小さくするように定義する。その他にも例えば、周辺領域 40 の実空間での大きさが一定となるように、所定形状を定める。ここで、周辺領域 40 の実空間での大きさは、カメラ 10 のキャリブレーション情報を用いて推定することができる。なお、カメラ 10 のキャリブレーション情報は、カメラ 10 上の座標を実空間上の座標に変換するために必要な種々のパラメータ（カメラの位置や姿勢、レンズの歪など）の情報を含む。キャリブレーション情報の生成には、既知の手法を利用することができる。

[0078] <<人物検出処理の対象とする動画フレーム 14>>

人物検出部 2040 は、対象物体を含む複数の動画フレーム 14 のうち、少なくとも 2 つの動画フレーム 14 を人物検出処理の対象とする。以下、この 2 つの動画フレーム 14 を第 1 動画フレームと第 2 動画フレームと呼ぶ。第 2 動画フレームは、第 1 動画フレームよりも後に生成された動画フレーム 14 である。以下、第 1 動画フレームと第 2 動画フレームとして扱う動画フレーム 14 を定める方法について説明する。

[0079] 人が物体を或る場所に置き去る場合、少なくともその物体がその場所に置かれるまでは、その人物がその物体の周辺に存在すると言える。一方、その物体が置かれてからしばらく後には、その人物がその物体の周辺に存在しない蓋然性が高い。よって、対象物体が置かれた時点付近又はその時点よりも前の時点に対象物体の周辺に存在する人物が、対象物体が置かれてからしばらく後に対象物体の周辺に存在していなければ、対象物体が置き去られてい

ると推測できる。

[0080] そこで、例えば第1動画フレームには、前述した静止判定期間の開始時点又はそれに近い時点（所定時間前又は後の時点）に生成された動画フレーム14を採用できる。例えば図5の例において、静止判定期間の開始時点は時点 $t+a$ である。よって、静止判定期間の開始時点に生成された動画フレーム14を第1動画フレームとすると、動画フレーム14-2が第1動画フレームとなる。

[0081] その他にも例えば、第1動画フレームは、静止判定期間の終了時点 $t+a+p$ に抽出された対象物体が実際に置かれた時点（以下、置き去り時点と呼ぶ）を基準として決定されてもよい。人が往来する場所などがカメラ10によって撮像される場合、対象物体が静止していると判定され始める時点（静止判定期間の開始時点）は、対象物体が実際に置かれた時点と一致するとは限らないためである。

[0082] そのために、人物検出部2040は、対象物体の置き去り時点を推定する。例えば人物検出部2040は、静止判定期間の終了時点 $t+a+p$ に生成された動画フレーム14から対象物体の画像特徴量を抽出し、その特徴量が時点 $t+a+p$ よりも前に生成された各動画フレーム14において同じ位置で検出されるかどうかを遡って調べる。そして、人物検出部2040は、例えば対象物体の画像特徴量が検出されなくなる時点が、置き去り推定時刻であると推定する。置き去り時点を基準として決定される第1動画フレームは、例えば、置き去り時点に生成された動画フレーム14や、置き去り時点よりも所定時間前又は後に生成された動画フレーム14である。

[0083] その他にも例えば、第1動画フレームには、対象物体が含まれる動画フレーム14のうち、最も生成時点が早い動画フレーム14（対象物体が出現した動画フレーム14）を採用できる。例えば図5の例において、対象物体が出現した動画フレーム14は、動画フレーム14-1である。

[0084] その他にも例えば、第1動画フレームには、第2動画フレームの生成時点よりも所定の長さ前の時点に生成された動画フレーム14を採用してもよい

。

[0085] 一方、第2動画フレームには、例えば前述した静止判定期間の終了時点又はそれに近い時点（所定時間前又は後の時点）に生成された動画フレーム14を採用できる。図5の例において、静止判定期間の終了時点は $t+a+p$ である。よって、静止判定期間の終了時点に生成された動画フレーム14を第2動画フレームとすると、動画フレーム14-3が第2動画フレームとなる。その他にも例えば、第2動画フレームには、第1動画フレームの生成時点よりも所定の長さ後の時点に生成された動画フレーム14を採用してもよい。

[0086] 人物検出部2040は、前述した第1動画フレームと第2動画フレームに加え、他の動画フレーム14をさらに人物検出処理の対象としてもよい。例えば第1動画フレームに含まれる人物が屈んでいる場合、第1動画フレームに人物が含まれているにもかかわらず、姿勢変動によりその人物を検出できない可能性がある。そこで例えば、人物検出部2040は、第1動画フレームの生成時点の前後所定時間に生成された複数の動画フレーム14も、人物検出処理の対象とする。同様に人物検出部2040は、第2動画フレームの生成時点の前後所定時間に生成された複数の動画フレーム14も、人物検出処理の対象とする。

[0087] <置き去り判定処理の実行：S106>

置き去り判定部2050は、対象物体が含まれる複数の動画フレーム14それぞれを対象とした人物検出処理の結果を比較することで、検出された静止物体が置き去り物体であるかどうかを判定する（S106）。

[0088] 図9は、置き去り判定部2050によって実行される置き去り判定処理の流れを例示する図である。この例において、置き去り判定部2050は、第1動画フレームに対する人物検出結果と第2動画フレームに対する人物検出結果を比較する。

[0089] まず置き去り判定部2050は、第1動画フレームを対象とする人物検出処理によって人物が検出されたか否かを判定する（S202）。人物が検出されなかった場合、図9の処理は終了する（置き去りとは判定されない）。

第1動画フレームにおいて対象物体の周辺から人物が検出されないケースは、例えば、対象物体がカメラ10の撮像範囲外から移動してきたのではなく、カメラ10の撮像範囲内に突然出現したケースである。例えば、対象物体が、カメラ10の撮像範囲内の或る場所に固定で設置されている物体（看板や銅像など）であるとする。この場合、対象物体が他の物体（例えば車などの機械）に隠されていると、対象物体がカメラ10によって撮像されない。しかし、対象物体を隠している物体が移動すると、対象物体がカメラ10によって撮像されるようになる。つまり、カメラ10の視点からは、対象物体が突然出現したことになる。その他にも例えば、環境光の変化などによって背景画像との差分が生じることで、対象物体が動画データ12に突然現れたかのように検出される場合がある。これらのケースでは、対象物体は置き去られていない蓋然性が高いと言える。そのため、置き去り判定部2050は対象物体を置き去り物体と判定しない。

[0090] 第1動画フレームを対象とする人物検出処理によって人物が検出された場合（S202：YES）、人物が検出された領域（人物領域）から人物の照合に必要な特徴量（顔や頭部などの身体の一部の特徴量や、服若しくは持ち物などの特徴量）を抽出した上で、図9の処理はS204に進む。S204において、置き去り判定部2050は、第2動画フレームを対象とする人物検出処理によって、第1動画フレームを対象とする人物検出処理によって検出された人物と同じ人物が検出されたか否かを判定する。この際、第1動画フレームから人物が検出された場合には、その人物領域から、人物の照合に必要な特徴量を抽出する。そして、置き去り判定部2050は、第1動画フレームから抽出された特徴量と第2動画フレームから抽出された特徴量の類似度がある所定の閾値よりも高い場合に、これらの動画フレームから同一の人物が検出されたと判定する。

[0091] 同じ人物が検出された場合（S204：YES）、図9の処理は終了する（対象物体を置き去り物体と判定しない）。このケースでは、第1動画フレームの生成時点と第2動画フレームの生成時点において、同一の人物が対象

物体の周辺に存在している。すなわちこの状況は、対象物体を置いた人物がその周辺に留り続けている状況であり、対象物体は置き去られていない蓋然性が高いと言える。よって、置き去り判定部2050は対象物体を置き去り物体と判定しない。

[0092] これに対し、第2動画フレームを対象とする人物検出処理によって、第1動画フレームを対象とする人物検出処理によって検出された人物と同じ人物が検出されなかった場合（S204：NO）、置き去り判定部2050は対象物体を置き去り物体と判定する（S206）。このケースは、1）第2動画フレームを対象とする人物検出処理によって人物が検出されないか、又は2）第2動画フレームを対象とする人物検出処理によって検出された人物が、第1動画フレームを対象とする人物検出処理によって検出された人物とは異なるケースである。いずれのケースであっても、対象物体を置いた人物がその対象物体の周辺に存在しなくなっており、対象物体が置き去られている蓋然性が高いと言える。そこで置き去り判定部2050は対象物体を置き去り物体と判定する。

[0093] なお、第1動画フレームと第2動画フレームから複数の人物が検出されることが考えられる。この場合、例えば、第1動画フレームから検出された複数の人物のいずれかが、第2動画フレームから検出された複数の人物のいずれかと同じである場合、置き去り判定部2050は、「第2動画フレームを対象とする人物検出処理によって、第1動画フレームを対象とする人物検出処理によって検出された人物と同じ人物が検出された」と判定する。一方、第1動画フレームから検出された全ての人物が、第2動画フレームから検出された複数の人物のいずれとも異なる場合、置き去り判定部2050は、「第2動画フレームを対象とする人物検出処理によって、第1動画フレームを対象とする人物検出処理によって検出された人物と同じ人物が検出されなかった」と判定する。

[0094] その他にも例えば、置き去り判定部2050は、第1動画フレームから検出される複数の人物それぞれについて、対象物体を置き去った人物（置き去

り主)である蓋然性の高さを判定し、その判定結果を用いて人物同一判定を行ってもよい。例えば置き去り判定部2050は、第1動画フレームから検出される各人物について、その検出位置と対象物体との距離に応じて、置き去り主である尤度を算出する。すなわち、人物の検出位置と対象物体との距離が小さいほど、その人物の尤度を大きくする。この際、人物の検出位置と対象物体との距離は、動画フレーム14上の距離でもよいし、実空間上の距離でもよい。実空間上の距離は、カメラ10のキャリブレーション情報を用いて推定することができる。そして、例えば置き去り判定部2050は、第1動画フレームから検出される人物のうち、所定値以上の尤度を持つ人物のいずれかが、第2動画フレームからも検出された場合に、「第2動画フレームを対象とする人物検出処理によって、第1動画フレームを対象とする人物検出処理によって検出された人物と同じ人物が検出された」と判定する。

[0095] また、上述の例では第1動画フレームを対象とする人物検出結果と第2動画フレームを対象とする人物検出結果を比較しているが、置き去り判定部2050は必ずしもこれらを比較する必要はない。例えば、第1動画フレームを対象とする人物検出処理の結果の代わりに、第1動画フレームの生成時点の前後所定時間に生成されたいずれかの動画フレーム14を対象とする人物検出処理の結果を利用してもよい。例えば、第1動画フレーム及び第1動画フレームの生成時点の前後所定時間に生成された複数の動画フレーム14それぞれを対象として人物検出処理を行った結果、いずれか複数の動画フレーム14から同じ人物が検出されたとする。この場合、置き去り判定部2050は、最も鮮明にその人物が検出された人物検出処理の結果を利用する。第2動画フレームについても同様である。

[0096] また、置き去り判定部2050は、3つ以上の動画フレーム14を比較することで、対象物体を置き去り物体と判定してもよい。例えば人物検出部2040は、第1動画フレームと第2動画フレームとの間に生成される1つ以上の動画フレーム14についても、人物検出処理を行う。以下、第1動画フレームと第2動画フレームとの間に生成される動画フレーム14を、中間動

画フレームと呼ぶ。そして置き去り判定部2050は、第1動画フレーム、第2動画フレーム、及び1つ以上の中間動画フレームそれぞれに対する人物検出処理の結果に基づいて、対象物体を置き去り物体と判定する。こうすることで、対象物体を置いた人物がその周辺に留まり続けているかどうかを、より高い精度で判定することができる。

[0097] 例えば置き去り判定部2050は、上述のS204における判定の代わりに、第1動画フレームから検出された人物と同じ人物が、第2動画フレームに加え、1つ以上の中間動画フレームからも検出されるか否かを判定する。そして、例えば置き去り判定部2050は、第2動画フレーム及び中間動画フレームの中に、第1動画フレームから検出された人物と同じ人物が検出されない動画フレーム14がある場合には、対象物体を置き去り物体と判定する。一方、第2動画フレーム及び全ての中間動画フレームから、第1動画フレームから検出された人物と同じ人物が検出される場合、置き去り判定部2050は対象物体を置き去り物体と判定しない。なお、第1動画フレームから検出された人物は、必ずしも全ての中間動画フレームから検出されなくてもよい。例えば置き去り判定部2050は、所定割合以上の中間動画フレームから、第1動画フレームから検出された人物と同じ人物が検出された場合には、対象物体を置き去り物体と判定しないようにしてもよい。

[0098] その他にも例えば、置き去り判定部2050は、まずは第1動画フレームから検出された人物と第2動画フレームから検出された人物とが同一であるか否かを判定し、その判定の精度が十分でない場合のみ、中間動画フレームを利用するようにしてもよい。例えば置き去り判定部2050が、第1動画フレームから検出された人物の特徴量と、第2動画フレームから検出された人物の特徴量との類似度に基づいて、これらの人物が同一であると判定するとする。この場合、例えば置き去り判定部2050は、1) 類似度が第1所定値以上であれば、検出された人物が同一であると判定し、2) 類似度が第2所定値(第1所定値未満の値)未満であれば、検出された人物が同一でないと判定し、3) 類似度が第2所定値以上であって第1所定値未満であれば

、判定精度が不十分であると判定する。そして３）の場合、置き去り判定部２０５０は、第１動画フレームから検出された人物が中間動画フレームから検出されるかどうかをさらに判定する。

[0099] その他にも例えば、置き去り判定部２０５０は、置き去られた物体の種別を推定するようにしてもよい。置き去り物体が何であるかによって、置き去り物体が検出された際の対処などが異なるためである。例えば、置き去り物体が、荷物を運ぶためのカートや、床に置く注意を促す掲示ボード（Cautionと書かれたボードなど）などである場合は、その存在によって大きな問題は生じないため、置き去り物体として検知する必要はない。このような場合には、置き去り判定部２０５０は、置き去り物体として判定しないようにしてもよい。

[0100] 以上、置き去り物体検出部２１２０の動作について述べたが、実施形態２における、置き去り時刻推定部２１４０の動作は、実施形態１の時と同様である。また、提示制御部２１６０についても、基本的には実施形態１の時と同様である。ただし、状況をより視認しやすくするため、提示制御部２１６０は、人物検出結果に基づいて、対象物体を置き去ったと推測される人物に関する情報（以下、人物情報）も合わせて画面に提示してもよい。なお、対象物体を置き去ったと推測される人物は、第１動画フレームにおいて対象物体の周辺から検出され、なおかつ第２動画フレームにおいて対象物体の周辺から検出されなかった人物である。

[0101] 例えば提示制御部２１６０は、画像を提示する際、対象物体の持ち主と考えられる人物の領域を枠で囲って表示してもよい。すなわち提示制御部２１６０は、置き去り判定部２０５０で対象物体の持ち主である尤度が高いと判定された人物の領域を枠で囲うようにして提示してもよい。この際、最も尤度が高い人物のみを枠で囲って表示するようにしてもよいし、ある一定以上の尤度を持つ人物を枠で囲うようにしてもよい。複数の人物を枠で囲って提示する際には、尤度に応じて、枠の色を変えたり、太さを変えたり、あるいは、尤度の値を直接枠と一緒に表示するようにして、尤度の違いがわかるよ

うに提示してもよい。なお、後述の追跡処理により、第1動画フレームの前後も対象人物の位置が求まっている場合には、求まっているフレームにおいても、人物の枠を提示するようになっていてもよい。

[0102] また、提示制御部2160は、人物情報として、上述した情報以外の他の情報を提示してもよい。例えば人物情報は、対象物体を置き去ったと推測される人物の画像やその特徴量、その人物がカメラ10の撮像範囲に入った時点（到着時点）、その人物がカメラ10の撮像範囲から出た時点（立ち去り時点）などを含む。ここで、到着時点や立ち去り時点の推定は、例えば、対象物体を置き去ったと推定される人物が検出された動画フレーム14の前後に生成された複数の動画フレーム14を用いてその人物の追跡処理を行い、対象物体が置かれた場所からその人物が動き出した時点を求めることによって推定できる。その他にも例えば、カメラ10の撮像範囲からその人物が消失した時点に基づいて、到着時点や立ち去り時点を推定するようにしてもよい。なお、人物の追跡処理には、既知の方式を用いることができる。なお、追跡処理によってその人物の移動速度を求め、その移動速度も人物に含めるようにしてもよい。対象物体を置き去ったと推定される人物の移動速度は、周辺の他のカメラにおける、当該人物の出現時点の予測などに用いることができる。

[0103] また、対象物体を置き去ったと推定される人物が、ブラックリストなど、人物に関する他の情報に含まれる場合、人物情報は、この「他の情報」に示されるその人物の情報を含んでもよい。

[0104] なお、以上の説明では、置き去りを検知する場合について述べてきたが、置き去り検知の技術は、落書き検知や持ち去り検知にも適用できる。よって、落書きをしている人物や、持ち去りをした人物の状況を動画データで提示する際にも、上述の技術を用いることができる。

[0105] 以上、図面を参照して本発明の実施形態について述べたが、これらは本発明の例示であり、上記各実施形態を組み合わせた構成や、上記以外の様々な構成を採用することもできる。

[0106] 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。

1. 動画データから置き去り物体を検出する置き去り物体検出手段と、
前記置き去り物体が置かれた時刻である置き去り時刻を推定する置き去り時刻推定手段と、

前記推定された置き去り時刻付近の動画データを取得し、前記推定された置き去り時刻に基づいて前記取得した動画データの提示方法を制御して、前記取得した動画データの少なくとも一部を提示する提示制御手段と、を有する情報処理装置。

2. 前記置き去り物体検出手段は、

静止物体として検出された物体である対象物体が含まれる第1動画フレーム、および前記対象物体が含まれ、なおかつ前記第1動画フレームよりも後に生成された第2動画フレームそれぞれを対象に、前記対象物体の周辺から人物を検出する人物検出処理を実行する人物検出手段と、

前記第1動画フレームおよび前記第2動画フレームそれぞれを対象とした前記人物検出処理の結果を比較することで置き去り状態を判定する置き去り判定手段と、を有し、

前記置き去り判定手段で置き去り状態と判定された場合に、置き去り物体を検出したと判定する、1.に記載の情報処理装置。

3. 前記提示制御手段は、前記取得した動画データに含まれる複数の部分それぞれの再生速度を、前記推定された置き去り時刻に基づいて決定する、1.または2.に記載の情報処理装置。

4. 前記提示制御手段は、前記取得した動画データを逆方向に再生し、前記取得した動画データのうち、前記推定された置き去り時刻付近の部分の再生速度を、それ以外の部分の再生速度よりも遅くする、3.に記載の情報処理装置。

5. 前記提示制御手段は、前記取得した動画データに含まれる1つ以上の動画フレームを抽出し、前記抽出した動画フレームを並べて表示し、

前記取得した動画データの複数の部分それぞれから抽出される前記動画フレームの枚数は、前記推定された置き去り時刻に基づいて決定される、1. または2. に記載の情報処理装置。

6. 前記提示制御手段は、前記取得した動画データに含まれる1つ以上の動画フレームを抽出し、前記抽出した動画フレームを順次切り替えて表示し、

前記取得した動画データの複数の部分それぞれから単位時間ごとに抽出される前記動画フレームの枚数は、前記推定された置き去り時刻に基づいて決定される、1. または2. に記載の情報処理装置。

7. 前記提示制御手段は、前記抽出した動画フレームを、生成時点が遅い順に表示する、6. に記載の情報処理装置。

8. 前記提示制御手段は、前記推定された置き去り時刻付近の動画データから単位時間ごとに抽出する動画フレームの枚数を、それ以外の期間の動画データから単位時間ごとに抽出する動画フレームの枚数よりも多くする、5. から7. いずれか一つに記載の情報処理装置。

9. 前記提示制御手段は、前記取得した動画データの少なくとも一部と共に、前記検出された置き去り物体に関する情報を提示する、1. から8いづれか一つに記載の情報処理装置。

[0107] 10. コンピュータによって実行される制御方法であって、

動画データから置き去り物体を検出する置き去り物体検出ステップと、
前記置き去り物体が置かれた時刻である置き去り時刻を推定する置き去り時刻推定ステップと、

前記推定された置き去り時刻付近の動画データを取得し、前記推定された置き去り時刻に基づいて前記取得した動画データの提示方法を制御して、前記取得した動画データの少なくとも一部を提示する提示制御ステップと、を有する制御方法。

11. 前記置き去り物体検出ステップは、

静止物体として検出された物体である対象物体が含まれる第1動画フレー

ム、および前記対象物体が含まれ、なおかつ前記第 1 動画フレームよりも後に生成された第 2 動画フレームそれぞれを対象に、前記対象物体の周辺から人物を検出する人物検出処理を実行する人物検出ステップと、

前記第 1 動画フレームおよび前記第 2 動画フレームそれぞれを対象とした前記人物検出処理の結果を比較することで置き去り状態を判定する置き去り判定ステップと、を有し、

前記置き去り判定ステップで置き去り状態と判定された場合に、置き去り物体を検出したと判定する、10. に記載の制御方法。

12. 前記提示制御ステップにおいて、前記取得した動画データに含まれる複数の部分それぞれの再生速度を、前記推定された置き去り時刻に基づいて決定する、10. または 11. に記載の制御方法。

13. 前記提示制御ステップにおいて、前記取得した動画データを逆方向に再生し、前記取得した動画データのうち、前記推定された置き去り時刻付近の部分の再生速度を、それ以外の部分の再生速度よりも遅くする、12. に記載の制御方法。

14. 前記提示制御ステップにおいて、前記取得した動画データに含まれる 1 つ以上の動画フレームを抽出し、前記抽出した動画フレームを並べて表示し、

前記取得した動画データの複数の部分それぞれから抽出される前記動画フレームの枚数は、前記推定された置き去り時刻に基づいて決定される、10. または 11. に記載の制御方法。

15. 前記提示制御ステップにおいて、前記取得した動画データに含まれる 1 つ以上の動画フレームを抽出し、前記抽出した動画フレームを順次切り替えて表示し、

前記取得した動画データの複数の部分それぞれから単位時間ごとに抽出される前記動画フレームの枚数は、前記推定された置き去り時刻に基づいて決定される、10. または 11. に記載の制御方法。

16. 前記提示制御ステップにおいて、前記抽出した動画フレームを、生

成時点が遅い順に表示する、１５．に記載の制御方法。

１７． 前記提示制御ステップにおいて、前記推定された置き去り時刻付近の動画データから単位時間ごとに抽出する動画フレームの枚数を、それ以外の期間の動画データから単位時間ごとに抽出する動画フレームの枚数よりも多くする、１４．から１６いずれか一つに記載の制御方法。

１８． 前記提示制御ステップにおいて、前記取得した動画データの少なくとも一部と共に、前記検出された置き去り物体に関する情報を提示する、１０．から１７．いずれか一つに記載の制御方法。

[0108] １９． １０．から１８．いずれか一つに記載の制御方法の各ステップをコンピュータに実行させるプログラム。

[0109] この出願は、２０１７年７月２８日に出願された日本出願特願２０１７－１４６６３６号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。