

(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2014-93023

(P2014-93023A)

(43) 公開日 平成26年5月19日(2014.5.19)

(51) Int.Cl. F1 テーマコード(参考)
 G06T 7/60 (2006.01) G06T 7/60 180D 5L096

審査請求 未請求 請求項の数 16 OL (全 25 頁)

<p>(21) 出願番号 特願2012-244382 (P2012-244382) (22) 出願日 平成24年11月6日 (2012.11.6)</p>	<p>(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 (74) 代理人 100090273 弁理士 園分 孝悦 (72) 発明者 冨手 要 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 鳥居 寛 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) 5L096 FA02 FA18 FA59 FA62 FA64 FA67 FA69 FA77 FA78</p>
--	---

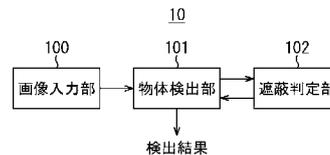
(54) 【発明の名称】 物体検出装置、物体検出方法及びプログラム

(57) 【要約】

【課題】 遮蔽が生じている多くの状況において精度良く物体を検出できるようにする。

【解決手段】 物体検出部101は、検出対象となっている人体の全身検出結果および頭部検出結果(検出スコア)を得る。遮蔽判定部102は、物体検出部101において得られた結果を用いて、検出対象に遮蔽が存在するか否かを判定し、遮蔽状態に応じて検出スコアを補正する。このとき、前景の人体の頭部検出結果と、後景の人体の全身検出結果とで重複領域が存在するか否かを判定し、重複領域がある場合には、遮蔽領域を算出し、算出した遮蔽領域に応じて検出スコアを補正する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

入力画像から前景に位置する第 1 の物体及び後景に位置する第 2 の物体を検出する物体検出装置であって、

前記第 1 及び第 2 の物体の部分領域を検出する第 1 の検出手段と、

前記第 1 及び第 2 の物体の姿勢を示す情報を検出する第 2 の検出手段と、

前記第 1 の検出手段によって検出された部分領域と前記第 2 の検出手段によって検出された情報とに基づいて、前記第 1 の物体による前記第 2 の物体の遮蔽状態を判定する判定手段と、

前記判定手段によって判定された遮蔽状態に応じて、前記第 2 の物体の検出結果を補正する補正手段と、

を備えることを特徴とする物体検出装置。

10

【請求項 2】

前記第 1 及び第 2 の検出手段は、前記物体を複数の部分に分けて位置及び大きさを検出することを特徴とする請求項 1 に記載の物体検出装置。

【請求項 3】

前記第 1 の検出手段は、人体の頭部の位置及び大きさを検出することを特徴とする請求項 1 又は 2 に記載の物体検出装置。

【請求項 4】

前記第 2 の検出手段は、人体の全身の位置及び大きさを検出することを特徴とする請求項 1 乃至 3 の何れか 1 項に記載の物体検出装置。

20

【請求項 5】

前記判定手段は、前記第 1 の検出手段によって検出された前記第 1 の物体の部分領域が、前記第 2 の検出手段によって検出された前記第 2 の物体の領域と交差するか否かによって遮蔽状態を判定することを特徴とする請求項 4 に記載の物体検出装置。

【請求項 6】

前記判定手段は、前記第 1 の検出手段によって検出された前記第 1 の物体の部分領域と前記第 2 の検出手段によって検出された前記第 2 の物体の領域との包含関係により遮蔽状態を判定し、遮蔽領域を計算することを特徴とする請求項 4 に記載の物体検出装置。

【請求項 7】

前記補正手段は、前記遮蔽領域の面積に応じて前記第 2 の物体の検出結果を補正することを特徴とする請求項 6 に記載の物体検出装置。

30

【請求項 8】

前記第 2 の検出手段は、さらに前記人体の体軸を検出することを特徴とする請求項 4 に記載の物体検出装置。

【請求項 9】

前記判定手段は、前記第 2 の検出手段によって検出された体軸の角度に応じて遮蔽状態の判定方法を切り替えることを特徴とする請求項 8 に記載の物体検出装置。

【請求項 10】

前記判定手段は、前記第 2 の検出手段によって検出された前記第 1 の物体の体軸と、前記第 2 の物体の全身を表す矩形とが交差した場合に生じる領域を遮蔽領域として計算することを特徴とする請求項 8 又は 9 に記載の物体検出装置。

40

【請求項 11】

前記判定手段は、前記第 2 の検出手段によって検出された前記第 1 及び第 2 の物体の体軸を延長した直線が、前記第 1 の検出手段によって検出された前記第 1 の物体の部分領域において互いに交差した場合に、前記第 1 及び第 2 の物体の部分領域が重複していると判定することを特徴とする請求項 8 に記載の物体検出装置。

【請求項 12】

前記判定手段により前記第 1 及び第 2 の物体の部分領域が重複していると判定された場合に、前記補正手段は、前記第 2 の物体の検出結果を補正しないようにすることを特徴と

50

する請求項 1 1 に記載の物体検出装置。

【請求項 1 3】

前記判定手段は、前記第 1 及び第 2 の物体の部分領域が重複していると判定した場合に、前記第 1 の検出手段による検出結果を、前記体軸を基準とした検出結果に変更することを特徴とする請求項 1 1 又は 1 2 に記載の物体検出装置。

【請求項 1 4】

前記判定手段は、前記第 1 及び第 2 の物体の部分領域が重複していると判定した場合に、互いの体軸がなす角度と、検出結果の候補とを決定することを特徴とする請求項 1 3 に記載の物体検出装置。

【請求項 1 5】

入力画像から前景に位置する第 1 の物体及び後景に位置する第 2 の物体を検出する物体検出方法であって、

前記第 1 及び第 2 の物体の部分領域を検出する第 1 の検出工程と、

前記第 1 及び第 2 の物体の姿勢を示す情報を検出する第 2 の検出工程と、

前記第 1 の検出工程において検出された部分領域と前記第 2 の検出工程において検出された情報とに基づいて、前記第 1 の物体による前記第 2 の物体の遮蔽状態を判定する判定工程と、

前記判定工程において判定された遮蔽状態に応じて、前記第 2 の物体の検出結果を補正する補正工程と、

を備えることを特徴とする物体検出方法。

【請求項 1 6】

入力画像から前景に位置する第 1 の物体及び後景に位置する第 2 の物体を検出する物体検出装置を制御するためのプログラムであって、

前記第 1 及び第 2 の物体の部分領域を検出する第 1 の検出工程と、

前記第 1 及び第 2 の物体の姿勢を示す情報を検出する第 2 の検出工程と、

前記第 1 の検出工程において検出された部分領域と前記第 2 の検出工程において検出された情報とに基づいて、前記第 1 の物体による前記第 2 の物体の遮蔽状態を判定する判定工程と、

前記判定工程において判定された遮蔽状態に応じて、前記第 2 の物体の検出結果を補正する補正工程と、

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。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、特に、遮蔽が生じている場合に用いて好適な物体検出装置、物体検出方法及びプログラムに関する。

【背景技術】

【0002】

1 枚の画像から部分的に遮蔽が生じている人物であっても、精度良く、かつ頑健に検出できるようにする技術は、動作解析などに応用が可能であり、近年盛んに研究されている。このような技術は、特にセキュリティシステム、安全運転支援、医療福祉といった分野などにおいて、その応用が検討されている。このように監視カメラや車載カメラ等の分野においても、画像中の人体に遮蔽が生じている場合に頑健に人体を検出する手法が知られている。

【0003】

例えば特許文献 1 に開示されている方法では、まず、入力画像全体から予め設定した人体の検出条件よりも検出スコアが大きくなる領域を検出する。この領域が最も前景に存在する人体を検出する領域になっている。次に、この領域を基準に遮蔽が生じる可能性がある周辺探索領域を計算する。そして、周辺探索領域内で前記条件よりも判定条件を緩くした条件で検出処理を行うことにより、前景の人体によって遮蔽が生じた後景の人物を検出

10

20

30

40

50

することができる。

【0004】

また、特許文献2に開示されている方法では、外光やガードレール等の高輝度物体により人体が部分的に遮蔽されている領域を、画像のコントラスト情報を利用して検出する。そして、検出した遮蔽領域は、遮蔽領域である旨を加味したスコアを計算することによりロバストに人体を検出している。

【0005】

また、非特許文献1に開示されている方法では、人体を構成するパーツの1つに遮蔽物が存在するか否かを示すoccluderを明示的に組み込むことにより、遮蔽物が存在しても頑健に人体を検出している。このoccluderを組み込むことによって、机やテーブルなどで人体が隠された場合に生じる直線状の強いエッジを検出して遮蔽状態を判定している。

10

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2010-49435号公報

【特許文献2】特開2011-165170号公報

【非特許文献】

【0007】

【非特許文献1】Proceedings of the Neural Information Processing Systems (NIPS) 2011. 「Object Detection with Grammar Models」

20

【非特許文献2】P. Viola and M. Jones (2001). "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Conference on Computer Vision and Pattern Recognition.

【非特許文献3】Dalal, N., & Triggs, B. (2005). "Histograms of oriented gradients for human detection.", IEEE CVPR

【非特許文献4】Platt, J. C. (1999). Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. Advances in Large Margin Classifiers.

【非特許文献5】Zadrozny, B., & Elkan, C. (2002). Transforming classifier scores into accurate multiclass probability estimates. Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

30

【非特許文献6】P. Felzenszwalb, D. McAllester, D. Ramanan (2008) "A Discriminatively Trained, Multiscale, Deformable Part Model", IEEE Conference on Computer Vision and Pattern Recognition.

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかし、特許文献1では、明示的に遮蔽判定をしているわけではなく、検出された人体の近くでかつ遮蔽が存在しそうな人体については、検出閾値を下げて検出しやすくするという方法をとっている。そのため、誤検出が多くなり検出精度は下がる。

40

【0009】

また、特許文献2では、人体を遮蔽する物体が高輝度物体であるという特性を利用し、コントラストの情報のみで遮蔽領域を検出しているため、必ずしも高コントラストになるとは限らない人体同士の重なりによる遮蔽などには対応することができない。

【0010】

非特許文献1では、机やテーブルといった直線的な強いエッジを対象としているため、机やテーブルによる人体の遮蔽には対応ができるが、形状が自由に變形する遮蔽（例えば、検出対象の前に他の人物が立っているなど）には対応できない。

【0011】

本発明は前述の問題点に鑑み、遮蔽が生じている多くの状況において精度良く物体を検

50

出できるようにすることを目的としている。

【課題を解決するための手段】

【0012】

本発明の物体検出装置は、入力画像から前景に位置する第1の物体及び背景に位置する第2の物体を検出する物体検出装置であって、前記第1及び第2の物体の部分領域を検出する第1の検出手段と、前記第1及び第2の物体の姿勢を示す情報を検出する第2の検出手段と、前記第1の検出手段によって検出された部分領域と前記第2の検出手段によって検出された情報とに基づいて、前記第1の物体による前記第2の物体の遮蔽状態を判定する判定手段と、前記判定手段によって判定された遮蔽状態に応じて、前記第2の物体の検出結果を補正する補正手段と、を備えることを特徴とする。

10

【発明の効果】

【0013】

本発明によれば、遮蔽が生じている多くの状況において精度良く物体を検出することができる。

【図面の簡単な説明】

【0014】

【図1】実施形態に係る物体検出装置の簡単な構成例を示すブロック図である。

【図2】実施形態に係る物体検出装置の詳細な構成例を示すブロック図である。

【図3】実施形態に係る物体検出装置の検出処理部が行う処理手順の一例を示すフローチャートである。

20

【図4】各検出器の結果から頭部の位置を推定する概要を説明するための図である。

【図5】全身検出器の結果と頭部との位置関係の定義例を示す図である。

【図6】頭部の正解基準を用いて頭部位置推定結果を評価する処理を説明する図である。

【図7】実施形態において、統合結果出力部が行う統合処理手順の一例を示すフローチャートである。

【図8】統合結果出力部から出力される処理結果の具体例を説明する図である。

【図9】前景となる人体により後景の人体が部分的に遮蔽されている様子を説明する図である。

【図10】図9における遮蔽状態の判定方法と遮蔽領域の計算方法とを説明する図である。

30

【図11】2人以上の人物が重なって遮蔽が生じている状態の一例を示す図である。

【図12】第1の実施形態における遮蔽判定部による処理手順の一例を示すフローチャートである。

【図13】パーツベースの検出手法を用いた全身検出器を説明する図である。

【図14】パーツベースの検出器と人体の体軸との関係を説明する図である。

【図15】複数のパーツ検出器による頭部の推定結果と体軸の推定結果とから、遮蔽判定を行う事例を説明する図である。

【図16】遮蔽領域の面積と位置とを算出する方法を説明する図である。

【図17】第2の実施形態における遮蔽判定部による処理手順の一例を示すフローチャートである。

40

【図18】前景の人体によって後景の人体が遮蔽されており、かつ両人体の頭部が重なっている場合の検出処理を説明する図である。

【図19】第3の実施形態における遮蔽判定部による処理手順の一例を示すフローチャートである。

【図20】統合処理前の頭部および体軸の検出結果の一例を示す図である。

【発明を実施するための形態】

【0015】

(第1の実施形態)

以下、本実施形態では、対象物を複数の部分領域に分割して、各部分領域の対象物らしさを部分領域スコアとして算出する。そして、一部または全ての部分領域スコアを統合し

50

て統合スコアを算出し、統合スコアを元に対象物であるかを判定することによって対象物検出処理を実施する。さらに、検出処理の過程で算出した部分領域のスコアと統合スコアとから、複数の対象物の位置姿勢情報を算出し、遮蔽判定を行うことにより頑健な検出処理を実現する。本実施形態で検出する対象物は特に限定されないが、検出対象物を人物とした場合について説明する。

【0016】

図1及び図2は、本実施形態に係る物体検出装置10の構成例を示すブロック図である。以下、図1及び図2を参照しながら本実施形態の構成について説明する。なお、本実施形態の物体検出装置10は、ネットワークまたは各種記録媒体を介して取得したソフトウェア(プログラム)を、CPU、メモリ、ストレージデバイス、入出力装置、バス、表示装置などにより構成される制御装置にて実行することで実現できる。また、不図示の制御装置については、汎用の制御装置を用いてもよいし、本発明のソフトウェアに最適に設計されたハードウェアを用いてもよい。

10

【0017】

< 検出処理に関わる構成 >

図1に示すように、本実施形態の物体検出装置10は、検出処理において、画像入力部100、物体検出部101、及び遮蔽判定部102を介して、対象物の検出結果を出力する。以下、これらの構成ブロックについて説明する。

【0018】

画像入力部100は、画像処理装置に画像を入力する部分である。画像入力部100に入力される画像は、カメラなどから得た動画像の1フレームの画像でもよく、ハードディスクドライブなどのストレージデバイスに保存された画像でもよい。

20

【0019】

以下、入力画像の注目領域に対する処理について説明する。注目領域は、入力画像の一部の領域であり、後述する対象物検出器のサイズと同じ画像サイズである。入力画像中の注目領域を、画像内でスライドさせながら順次設定することにより、画像全体から対象物を検出することができる。また、さらに入力画像を拡大あるいは縮小することにより、様々なサイズに写った対象物を検出することができる。

【0020】

物体検出部101は、検出対象となっている対象物の推定位置およびスコア(尤度)を算出する。以下、物体検出部101の詳細について、図2を参照しながら説明する。

30

【0021】

物体検出部101は、図2に示すように、画像入力部100、第1の検出処理部211~第nの検出処理部21n、及び各検出処理部に対応した第1の共通部位推定部221~第nの共通部位推定部22nを備えている。さらに、第1のスコア補正辞書231~第nのスコア補正辞書23n、第1のスコア補正部241~第nのスコア補正部24n、及び統合結果出力部250を備えている。以下、画像入力部100に入力された1枚の画像に対する処理を例にして、各構成要素について説明する。

【0022】

第1の検出処理部211~第nの検出処理部21nには、対象物の異なる部位や状態を検出する検出器が、予め格納されている。検出対象物を人物とした場合、各検出処理部の異なる検出器には、顔検出器、頭部検出器、上半身検出器、全身検出器などの人物の異なる部位を対象とした検出器を用いることができる。人物の異なる部位を検出する検出器を用いることにより、人物の一部が他の物体から遮蔽されている場合や、画像中から人物の一部がはみ出している場合にも人物を検出できるようになる。

40

【0023】

これらの複数の検出器は、それぞれ相互に補間し合うような検出器を準備することが望ましい。検出器を相互に補間し合う例として、例えば、頭部検出器と全身検出器との組み合わせが考えられる。まず、頭部検出器は、胴体以下が他の物体から遮蔽されていても人物を検出することが可能であり、また、身体部分の姿勢変動に影響を受けずに人物を検出

50

できるという利点がある。しかし、頭部は特徴的な形状が少ないため、検出性能は全身検出器に比べて劣る傾向にあることが欠点である。一方、全身検出器は、対象としている部位が大きいと、人物の特徴を捉えやすく検出性能が比較的高いという利点があるが、遮蔽に弱いという欠点がある。そこで、頭部検出器と全身検出器とを同時に利用することにより、相互の欠点を補うことができるようになり、人物検出の精度向上が期待できる。

【0024】

一方、顔検出器では、非特許文献2に開示されている方法のように、学習画像の顔範囲のHaar-Like特徴量を収集し、AdaBoostにより統計的に顔らしい特徴を識別できるようにして顔検出器を学習する。また、頭部や上半身、全身などその他の人物部位を学習する場合は、画像特徴としては非特許文献3に記載されているHOG特徴量を用いる。頭部検出器や上半身検出器、全身検出器を準備する場合には、各部位の学習画像を準備し、それぞれのHOG特徴量を取得して、SVM(サポートベクタマシン)やAdaBoost等の識別器により各部位の検出器を学習する。学習の結果(例えば、AdaBoostの弱識別器など)は、検出器辞書として保存し、検出時に利用する。

10

【0025】

また、各検出器は、人物検出の尤もらしさを検出スコアとして算出する。例えばAdaBoostでは、各弱識別器の出力の重み付き和を検出スコアとして出力する。またSVMでは、識別超平面との距離を検出器スコアとして算出する。上記以外の判別処理でも、尤度など対象物らしさを表すスコアを出力する方法であれば、どのような方法を用いてもよい。検出スコアは、対象物を示す確率値などに変換するなどして、各検出器の検出スコアを比較可能な値にしておくことが好ましい。なお、以下では、検出器スコアが高いほど、各検出器が対象としている人物部位、または人物状態らしさが高い出力が得られているものとする。

20

【0026】

以下、本実施形態では、複数の検出器として、顔検出器、頭部検出器、全身検出器の3つを用いる場合について説明するが、用いる検出器の構成はこの限りではない。

【0027】

次に、第1の検出処理部211、第2の検出処理部212、第nの検出処理部21nによる処理について説明する。図3は、第1の検出処理部211～第nの検出処理部21nが行う処理手順の一例を示すフローチャートである。以下、第1の検出処理部211が全身検出器の検出処理を行う例について説明する。

30

【0028】

まず、ステップS301において、入力画像の画像特徴量を算出する。この処理では、第1の検出処理部211の検出器は全身検出器であるため、入力画像からHOG特徴量を算出する。次に、ステップS302において、検出処理を行う画像の特定位置の画像特徴量を取得する。そして、ステップS303において、検出器辞書を用いて、処理対象の画像特徴量の対象物らしさを判別し、検出スコアを算出する。

【0029】

次に、ステップS304において、入力画像全体において検出スコアを算出したか否かを判定する。画像全体を探索するため、各画像中の位置で検出スコアを算出するよう判別位置を変えながら画像全体に対して行う。この判定の結果、入力画像全体において検出スコアを算出していない場合は、ステップS302に戻る。なお、判別位置を変更する際に、画像サイズも変更することにより、画像中で異なるサイズで写っている人物も検出できるようになる。

40

【0030】

一方、ステップS304の判定の結果、入力画像全体において検出スコアを算出した場合は、画像中の各位置での検出スコアが得られる。ここで、この結果を全て第1の共通部位推定部221に送るようにしてもよいが、明らかに人物ではないと判断できる低い検出スコアの検出結果については、これ以降の処理を省略して全体の処理負荷を低減することができる。したがって、次のステップS305においては、所定のスコア以上の結果を残

50

す閾値処理を行い、無駄な検出結果を削除する。ステップ S 3 0 5 の処理の結果、画像中で検出スコアが所定値よりも高い位置の位置情報とその検出スコアを共通部位推定部 2 2 1 に出力する。

【 0 0 3 1 】

以上、1つの検出処理部の処理結果を説明したが、物体検出装置 1 0 全体としては、この検出処理部の処理を検出処理部の数だけ繰り返す。

【 0 0 3 2 】

次に、第 1 の共通部位推定部 2 2 1 ~ 第 n の共通部位推定部 2 2 n について説明する。第 1 の共通部位推定部 2 2 1 ~ 第 n の共通部位推定部 2 2 n では、対象物の共通部位の位置を各検出器の結果から推定する。本実施形態では、異なる部位を検出する検出器の結果を統合するために、各検出器から対象物の共通な部位の位置または範囲を推定し、推定した部位の位置関係を元に検出結果を統合する。

10

【 0 0 3 3 】

例えば、各検出処理部において人体の腕や足、胴体が検出され、共通部位として頭部が推定される。複数の検出処理部で検出される結果より共通部位を推定して、後段の処理で検出候補を絞り込むことも可能になる。以下、本実施形態では、人物の頭部を共通部位として定義し、第 1 の共通部位推定部 2 2 1 が各検出器の検出結果から頭部位置を推定する手順について説明する。ただし、本実施形態では推定する共通部位を頭部にしたが、各検出器で共通に推定可能な部位であれば、特に限定するものではない。例えば、共通に推定可能な部位として人体の喉元から腰の中心を結ぶ体軸などが挙げられる。なお、検出対象物が人物である場合、人物の頭部は比較的遮蔽されにくい部位であるため共通部位として適している。

20

【 0 0 3 4 】

図 4 は、各検出器の結果から頭部の位置を推定する概要を説明するための図である。各検出器の検出処理の結果、検出対象の位置・範囲の情報が得られており、本実施形態では、検出結果の位置・範囲は、検出対象を囲む矩形枠で得られるものとする。図 4 に示す例では、検出結果を矩形枠で示しており、それぞれ顔検出結果枠 4 0 1、頭部検出結果枠 4 0 2、全身検出結果枠 4 0 3 の情報が得られている。矩形枠の座標 X は、画像座標の 2 点を用いて、以下の式 (1) により表される。

【 0 0 3 5 】

【 数 1 】

$$\mathbf{X} = (x_1, y_1, x_2, y_2)^T \quad \dots (1)$$

30

【 0 0 3 6 】

ここで、 x_1 、 y_1 は矩形の左上点の画像座標であり、 x_2 、 y_2 は矩形の右下点の画像座標である。第 1 の共通部位推定部 2 2 1 では、この矩形枠から共通部位として頭部位置・範囲を推定する。図 4 に示す例の場合は、顔検出結果枠 4 0 1 から推定した頭部位置・範囲を矩形枠 4 1 1 で表し、全身検出結果枠 4 0 3 から推定した頭部位置・範囲を矩形枠 4 1 3 で表している。検出結果枠から頭部の位置を推定する場合には、予め検出結果枠と頭部との位置関係を定義しておき、検出結果枠から頭部の位置に変換することによって頭部の位置を推定する。

40

【 0 0 3 7 】

図 5 は、全身検出器の結果と頭部との位置関係の定義例を示す図である。図 5 に示す例では、全身検出器に対して頭部の位置は、全身検出器の高さ h_B の 1 5 % を頭部高さ h_H とし、全身検出器の幅 w_B の 5 0 % を頭部幅 w_H と定義している。また、x 軸方向には $0.25 w_B$ のオフセットが定義されている。全身検出器から頭部の位置を推定する際には、全身結果の座標 X から、図 5 に示す定義に従って頭部座標 X_H として求める。頭部座標 X_H は以下の式 (2) により表される。

【 0 0 3 8 】

【数 2】

$$\mathbf{X}_h = (x_{h1}, y_{h1}, x_{h2}, y_{h2})^T \quad \dots (2)$$

【0039】

ここで、 x_{h1} 、 y_{h1} は推定した頭部範囲の左上点の座標であり、 x_{h2} 、 y_{h2} は推定した頭部範囲の右下点の座標である。第 1 の共通部位推定部 2 2 1 では、第 1 の検出処理部 2 1 1 の処理の結果として得られた各検出結果について、式 (1) に示した座標 X から頭部推定座標 X_h を算出する。

【0040】

なお、頭部範囲の定義は、各数値を予め人が入力・設計してもよいし、実際の全身検出結果から得られる頭部位置の平均から設計するようにしてもよい。頭部の位置の平均を取得する場合には、複数のサンプル画像に全身検出器による検出処理を行い、検出結果内の頭部位置の平均値を算出することによって求めることができる。

【0041】

以上の説明では、全身検出器から頭部位置を推定する方法を例にして第 1 の共通部位推定部 2 2 1 の動作について説明した。他の検出器の検出結果から頭部の位置を推定する場合についても、全身検出器の場合と同様に、各検出結果と頭部との位置関係を定義しておき、検出結果から頭部の位置を推定する。全身検出器では、検出結果の内部の頭部の位置を推定したが、推定する位置は検出結果の内部である必要はない。

【0042】

例えば、図 4 に示す顔検出結果枠 4 0 1 から推定した頭部の位置を示す矩形枠 4 1 1 は、顔検出結果枠 4 0 1 の外側にある。また、頭部そのものを検出する頭部検出器の頭部検出結果枠 4 0 2 では、共通部位推定部の処理を省略して、頭部検出結果そのものを共通部位として推定したという結果として出力するようにしてもよい。

【0043】

次に、第 1 のスコア補正辞書 2 3 1 ~ 第 n のスコア補正辞書 2 3 n と第 1 のスコア補正部 2 4 1 ~ 第 n のスコア補正部 2 4 n について説明する。本実施形態では、各検出結果から推定した共通部位の位置と、各検出スコアとを用いて複数の異なる検出結果を統合する。ここで、共通部位の位置は検出結果から推定した結果であり、その推定精度は検出器によって異なる。本実施形態では、頭部の位置を共通部位として推定しているが、頭部の位置の推定性能は、頭部の位置に近い、あるいは頭部と関係が深い検出器の方が良いと考えられる。そこで、共通部位の推定性能の差を考慮した統合を行うために、第 1 のスコア補正部 2 4 1 ~ 第 n のスコア補正部 2 4 n は、それぞれ第 1 のスコア補正辞書 2 3 1 ~ 第 n のスコア補正辞書 2 3 n を用いて共通部位の推定性能差に基づいた検出スコアの補正を行う。そして、補正した検出スコアを用いて周囲の検出結果を統合することにより、対象物の検出結果の位置精度が向上することが期待できる。

【0044】

第 1 のスコア補正部 2 4 1 ~ 第 n のスコア補正部 2 4 n では、それぞれの検出器の検出スコアを、それぞれ第 1 のスコア補正辞書 2 3 1 ~ 第 n のスコア補正辞書 2 3 n に記録された情報を用いて変換する。第 1 のスコア補正辞書 2 3 1 ~ 第 n のスコア補正辞書 2 3 n には、各検出器が共通部位を推定する信頼度に基づいて検出スコアを補正するための情報が格納されている。

【0045】

スコア補正では、検出器ごとに補正係数を各スコア補正辞書に保存し、スコア補正時には係数を検出スコアに乗じて補正スコアを算出すればよい。補正係数の例としては、頭部検出器の補正係数を 1 とし、顔検出器の補正係数を 0.8、全身検出器の補正係数を 0.5 などとする。このように、頭部に近い検出器（頭部位置の推定性能が高い検出器）では大きい補正係数を設定し、頭部から遠い検出器（頭部位置の推定性能が低い検出器）では低い補正係数を設定する。この補正係数は、姿勢・撮影条件・遮蔽の発生部位などにより

10

20

30

40

50

、変わり得るものであり、それらの状態を判定して適応的に設定してもよい。本実施形態では、検出器性能に関する事前確率を予め求めておき、その値を用いる。この補正係数を検出スコアに乗じて補正スコアを得ることにより、検出器の検出結果と共通部位の推定の性能とを考慮した補正スコアを得ることができる。補正スコアは、対象物らしさを示す検出スコアに、共通部位の位置推定の確からしさによって重み付けしたスコアとなっており、対象物らしさと位置の確からしさを合わせて示すことになる。

【 0 0 4 6 】

なお、補正係数をユーザが入力して設定してもよいが、補正係数は各検出器で推定する頭部位置の正解確率によって設定することが好適である。そこで、各検出器の頭部の位置の推定に係る正解確率については、事前に求めておく必要がある。以下、図 6 を参照しながら頭部の位置の推定に係る正解確率の求め方と各スコア補正辞書に保存する補正係数とについて説明する。

10

【 0 0 4 7 】

まず、頭部の位置が既知な画像サンプル群を準備する。図 6 (A) は、画像 6 0 0 の人物の頭部位置が既知である画像の例を示しており、頭部範囲の座標が頭部正解 6 0 1 として記録されている。ここで、画像 6 0 0 は、人物が一人しか写っていない、または、一人の人物範囲に切り出された画像であることが望ましい。このように頭部の位置が既知である画像を大量に準備する。

【 0 0 4 8 】

次に、図 6 (B) は、図 6 (A) の画像に対して顔検出を実施した結果を示している。顔検出の結果、検出処理で説明した処理と同様に、画像 6 0 0 全体に顔検出器の検出処理が逐次行われる。ここでは、画像 6 0 0 の中で顔検出の検出スコアが最も高い検出結果 6 1 1 に着目する。画像 6 0 0 には人物が一人しか写っていないため、最も高いスコアを示す検出結果 6 1 1 が顔であると考えられる。

20

【 0 0 4 9 】

次に、この顔検出結果から頭部の位置を推定した推定結果 6 1 2 を算出する。この頭部の位置の推定結果 6 1 2 と頭部正解 6 0 1 とを比較して、頭部の推定が正しく行われたか否かを評価する。頭部正解 6 0 1 と頭部の位置の推定結果 6 1 2 とを比較する際には、例えば、各位置の中心間距離が所定範囲内であれば、推定結果が正解であるものとする。また、他の基準としては、矩形形状の頭部正解 6 0 1 と頭部の位置の推定結果 6 1 2 との重複率を算出し、所定の重複率以上を示す場合に推定結果が正解であるものとしてもよい。矩形の重複率の算出方法としては、例えば、以下の式 (3) によって算出できる。

30

【 0 0 5 0 】

【 数 3 】

$$\alpha = \frac{S_{bc}}{S_h + S_c - S_{bc}} = \frac{\text{重複領域面積}}{\text{全領域面積}} \quad \dots (3)$$

【 0 0 5 1 】

ここで、 S_b は頭部正解の面積、 S_c は推定した頭部範囲の面積、 S_{bc} は頭部正解と推定した頭部範囲の重複した領域の面積である。以上の正解判定を、準備した全ての画像サンプル群に対し実行し、頭部推定が正解となった確率を求めることができ、その確率を補正係数とする。なお、画像サンプルに対して検出結果自体が得られない場合には、頭部推定は不正解として判定する。

40

【 0 0 5 2 】

他の検出器についても同様に、頭部推定の正解確率を各検出器について求め、それぞれの正解確率を各検出器の補正係数として利用すればよい。図 6 (D) に示す例では、全身検出器の検出結果 6 3 0 から推定した頭部位置 6 3 1 と頭部正解 6 0 1 との位置関係を評価する。図 6 (D) に示す例の場合、頭部正解 6 0 1 から頭部位置 6 3 1 が大きくずれているため、全身検出器からの頭部の位置の推定は不正解となる。

50

【 0 0 5 3 】

また、図 6 (C) には、頭部検出器の検出結果の正解判定の例を示している。頭部検出器の結果についても、他と同様に頭部正解との評価を行い、頭部位置を示す性能を評価して補正係数を算出してもよい。頭部検出器では、頭部の位置の推定を必ずしも行う必要がないので、その場合は検出結果そのものの位置と頭部正解との評価を行う。

【 0 0 5 4 】

また、上記の重複率 を用いて補正係数を算出する場合には、各画像サンプルで正解・不正解の 2 値判定により正解確率を算出している。そこで、この情報を利用して非特許文献 4 に開示されている Platt scaling や非特許文献 5 に開示されている Isotonic Regression を行って、スコア補正を行ってもよい。また、スコア補正を全く行わなくてもよい。

10

【 0 0 5 5 】

以上の処理により補正スコアが算出されると、統合結果出力部 2 5 0 は、これら検出器の結果を統合し、同じ人物に対して複数の検出器から出力される情報を 1 つにまとめる。なお、本実施形態では、同じ人物周辺において同じ検出器から重複して出力された検出結果をまとめることを目的としているわけではない。

【 0 0 5 6 】

以下、同じ人物に対して複数の検出器から出力される情報を 1 つにまとめる処理について説明する。図 7 は、本実施形態において、統合結果出力部 2 5 0 が行う統合処理手順の一例を示すフローチャートである。図 7 に示す処理では、ステップ S 7 0 1 ~ ステップ S 7 0 4 の処理を全身検出器の個々の出力結果に対してループしながら行う例について説明する。以下、注目している全身検出器の出力結果の番号を i ($i = 1, \dots, L$) とし、その番号の頭部推定座標を $X_{hB,i}$ とする。

20

【 0 0 5 7 】

まず、ステップ S 7 0 1 において、全身検出器の検出結果が残っているか否かを判定する。この結果、残っている場合にはステップ S 7 0 2 に進み、そうでない場合は処理を終了する。次に、ステップ S 7 0 2 において、頭部の位置の推定結果の中で座標 $X_{hB,i}$ が示す領域との重複率が最も高いものを選択する。このとき、選択した推定結果の番号を j ($j = 1, \dots, M$) とし、その番号の頭部の推定座標を $X_{hH,j}$ とする。ここで座標 $X_{hB,i}$ が示す領域と座標 $X_{hH,j}$ が示す領域との重複率 $A_0(X_{hB,i}, X_{hH,j})$ は、以下の式 (4) より求める。

30

【 0 0 5 8 】

【 数 4 】

$$A_0(X, Y) = \frac{P(X, Y)}{S(X) + S(Y) - P(X, Y)} = \frac{\text{重複領域面積}}{\text{全領域面積}} \quad \dots (4)$$

【 0 0 5 9 】

ここで、 $P(X, Y)$ は矩形 X と矩形 Y とで重複した領域の面積である。また、 $S(X)$ 、 $S(Y)$ は、それぞれ矩形 X 、矩形 Y の面積である。

【 0 0 6 0 】

次に、ステップ S 7 0 3 において、顔の位置の推定結果の中で、座標 $X_{hB,i}$ が示す領域との重複率が最も高いものを選択する。このとき、選択した推定結果の番号を k ($k = 1, \dots, N$) とする。そして、ステップ S 7 0 4 において、各全身検出器の検出結果 i について、以下の式 (5) に示すベクトル R_i を出力する。

40

【 0 0 6 1 】

【 数 5 】

$$R_i = (X_{hH,j}, S_{B,i} + S_{H,j} + S_{F,k}) \quad \dots (5)$$

【 0 0 6 2 】

ここで、 $S_{B,i}$ 、 $S_{H,j}$ 、 $S_{F,k}$ はそれぞれ、 i 番目の全身検出器の補正スコア、 j 番目

50

の頭部検出器の補正スコア、k番目の顔検出器の補正スコアであり、それらの和が統合スコアとして統合結果出力部250から出力される。本実施形態では、各検出器のスコアを補正して単純な和をとって統合スコアとする。なお、各検出器の種類によっては補正が必要のない場合があり、このことは検出精度を比較することによって要否を判断できる。また、スコアを補正する場合でもしない場合でも各検出器のスコアの線形和を取って統合スコアとすることもできる。この場合の線形係数は、各検出器のスコアを入力ベクトルとしたSVMなどの学習によって得ることができる。

【0063】

図8は、統合結果出力部250から出力される処理結果の具体例を説明する図である。図8(A)は、統合結果出力部250に入力された時点での検出結果を示しており、人物周辺に複数の検出結果が得られている状態である。なお、図8(A)に示す例では、説明を簡略化するために顔検出器の結果については省略しており、頭部検出器の検出結果及び全身検出器の検出結果のみを図示している。

10

【0064】

波線の矩形801は全身検出器の検出結果であり、波線の矩形804は全身検出器から推定された頭部位置の領域である。図8(A)に示す例では、1つの全身検出器の検出結果と、その頭部の推定結果とを示している。また、実線の矩形802、803は頭部検出器の2つの検出結果を示している。これらは、頭部を検出する処理で画像中の探索位置を変えながら検出処理を行った結果、人物の頭部の周辺に複数の検出結果が得られた結果である。統合結果出力部250では、共通部位である頭部位置と推定情報を用いて、これらの検出結果をまとめる。

20

【0065】

図8(B)は、図8(A)に示す検出結果を統合結果出力部250で処理した結果を示しており、全身検出器に基づいた矩形804に示す頭部の推定位置と最も重複度の高い矩形802に示す頭部の検出結果が選択され、統合結果として残されている。逆に、頭部検出器の誤検出と思われる矩形803に示す頭部の検出結果は、対応する全身検出の結果がないため、削除されている。

【0066】

< 遮蔽判定処理 >

遮蔽判定部102は、物体検出部101において統合された結果を用いて、検出対象に遮蔽が存在するか否かを判定し、遮蔽状態に応じて検出スコアを補正する。以下、物体検出部101の出力結果である統合スコアを検出スコアと記載して説明する。具体的には、遮蔽判定部102は、検出スコアの高い頭部の検出結果に対応する全身推定結果と、その周辺に存在する頭部の検出結果とを順次参照して遮蔽を判定する。なお、本実施形態では、検出する人体は常に直立した状態であることを想定しているものとする。

30

【0067】

図9は、第1の物体である前景となる人体900により第2の物体である後景の人体901が部分的に遮蔽されている様子を説明する図である。この時、物体検出部101の処理結果として、実線の矩形910、911に示す頭部の検出結果が得られており、波線の矩形920、921は、全身の検出結果を示している。

40

【0068】

図10は、図9における遮蔽状態の判定方法と遮蔽領域の計算方法とを説明する図である。まず、遮蔽状態を判定するために、前景の人体900の頭部の検出結果を表す矩形910に着目する。ここで、カメラで撮影した被写体は、透視投影の影響により、手前の物ほど大きく映り、奥にある物ほど小さく映る。そして、手前の物体ほど遮蔽が生じる可能性は小さい。この原理を踏まえ、物体検出部101から出力された全ての検出結果を検出スコアと検出枠のサイズとで大きい順にソートする。この処理を施すことにより、遮蔽が生じている可能性が低く、かつ遮蔽の原因になる手前の人体を特定することができる。そして、遮蔽の原因になりそうな人体から順に、他の検出結果との遮蔽判定を行う。ただし、近接する人体の場合は、必ずしも透視投影による顕著なサイズの違いは生じない。この

50

ような場合は、テクスチャの連続性などを考慮してどちらの人体が前景に存在する人体かを判定するなどしてもよい。

【0069】

そして、人体900の頭部の検出結果を表す矩形910と人体901の全身の検出結果を表す矩形921との交点、もしくは重複領域1002を算出する。ここで、一般的に最も前景と想定される人体の頭部と他の人体の全身の検出結果とが重複していた場合で、かつその全身の検出結果に対応する頭部の検出スコアが閾値よりも高い場合、その人体は高い確率で遮蔽状態にあると判定できる。そして、交差もしくは重複する全身の検出結果に対応する頭部の検出スコアの最も低いものから順に人体の一部が遮蔽状態にあると判定して、検出スコアを補正する。

10

【0070】

以下、遮蔽状態を判定して検出結果を補正する処理について説明する。

図12は、遮蔽判定部102による処理手順の一例を示すフローチャートである。図12に示す処理では、ステップS1201～ステップS1207の処理を物体検出部101で検出された個々の出力結果に対してループしながら行う。ここで、注目している全身検出器の出力結果の番号を i ($i = 1, \dots, L$)とする。

【0071】

まず、ステップS1201において、検出スコアの高い順であって、かつ頭部のサイズの大きい順に検出結果をソートする。この処理は、前述したように検出スコアが高く、かつ頭部のサイズが大きい人体ほど前景に存在する可能性が高いという前提知識に基づくものである。

20

【0072】

また、カメラから奥側に物体が存在する場合、もしくは子供の頭が検出された場合などは、検出スコアは高いが頭部のサイズは小さくなる傾向にある。この場合、検出スコアの高い順にソートした後、構築するシステムを実際に稼働させる状況で検出サイズに一定の閾値などを設けて、検出され得る最小の頭部のサイズを決定しておく。これにより、検出候補の尤度を信頼性の高いものにできる。

【0073】

さらに、被写体が柄の長い帽子を被っていた場合などは、検出スコアが低く頭部のサイズが大きくなることがある。このような場合には、第1の検出処理部211～第 n の検出処理部21 n に、柄の長い帽子を被っている頭部に対して高い検出結果を出力する検出器を備えることにより、検出スコアの補正を行って再度ソートすることが望ましい。

30

【0074】

次に、ステップS1202において、前景と推定される人体の全身及び頭部の検出結果を参照し、前景と推定される人体を選択する。以下、前景と推定される人体の全身の検出結果を A_B とし、その人体の頭部の検出結果を A_H とする。この処理では、ステップS1201でソートが完了しているので、検出スコアの高いものから選択することとなる。

【0075】

次に、ステップS1203において、後景の人体の全身及び頭部の検出結果を参照し、ステップS1202で選択した人体の検出スコアよりも低く、かつ頭部の検出結果に対応する全身の検出結果が算出されている人体を選択する。以下、この処理で選択される後景の人体の全身の検出結果を B_B とし、その人体の頭部の検出結果を B_H とする。

40

【0076】

次に、ステップS1204においては、ステップS1202で選択した前景の人体の頭部の検出結果(A_H)と、後景の人体の全身の検出結果(B_B)とで交点が2点以上存在するか、もしくは重複領域($A_H \cap B_B$)が存在するか否かを判定する。この判定の結果、交点もしくは重複領域が存在しない場合はステップS1207に進み、存在する場合はステップS1205に進む。

【0077】

ステップS1205においては、前景の人体の全身の検出結果(A_B)を利用して、後

50

景の人体における遮蔽領域 (A_B B_B) を算出する。図 10 に示す例の場合、各検出結果の左上及び右下の座標値を利用して遮蔽領域 1001 を算出する。まず、人体 901 の全身の検出結果の左上点を (x_j, y_j) として、同様に人体 900 の全身の検出結果の左上点を (x_i, y_i) とすると、幅 w_o および高さ h_o ($w_{o,i}, h_{o,i}$) は、以下の式 (6) で表される。

【0078】

【数6】

$$(w_{o,i}, h_{o,i}) = (x_i - x_j, y_i - y_j) \quad \dots (6)$$

【0079】

そして、人体 901 の全身を検出する際に、幅 w_B および高さ h_B は既に算出されているため、遮蔽領域 1001 の面積 S_o は、以下の式 (7) により求めることができる。

【0080】

【数7】

$$S_{o,i} = \frac{w_{B,i} h_{B,i}}{(w_{B,i} - w_{o,i})(h_{B,i} - h_{o,i})} \quad \dots (7)$$

ただし、 $(w_{B,i} - w_{o,i})(h_{B,i} - h_{o,i}) \neq 0$

【0081】

次に、ステップ S1206 において、ステップ S1205 で算出した遮蔽領域の面積から検出スコアの補正値を決定する。本実施形態で示した例のように、検出する人体が直立していることを仮定した場合は、遮蔽領域が大きければ大きいほど検出スコアは低下するため、遮蔽領域の面積と検出スコアの間には負の相関があると考えられる。したがって、遮蔽が存在する場合は、検出スコアに対して以下の式 (8) により補正を施す。

【0082】

【数8】

$$R_{o,i} = R_i \delta(S_{o,i}) \quad \dots (8)$$

【0083】

ここで、 $(S_{o,i})$ は、値域が 0 から 1 の単調増加関数である。仮に遮蔽が無ければ補正係数は 1 となり、遮蔽の面積が大きくなればなるほど、大きな係数がかかるように設計されている。このような補正をかけることによって、遮蔽のために生じた検出スコアの低下を取り消すことができる。

【0084】

また、前述した頭部の位置を推定する際の正解確率の求め方と同様に、遮蔽が生じるパターン毎に、遮蔽パターンと各検出器の検出スコアとの相関関係を求めておくことにより、遮蔽時の補正係数をスコア補正辞書として記録することも可能である。

【0085】

遮蔽時の補正係数を予め求めておくためには、まず、頭部位置と全身位置が既知の画像サンプル群を準備する。なお、準備する画像は、人体が遮蔽されている状態の画像と遮蔽されていない状態の画像とをそれぞれ準備することが望ましい。さらに、遮蔽されている画像とそうでない画像とで、人体の姿勢およびサイズが同じ状態で写っていることも望まれる。これは、遮蔽状態の有無により検出スコアにどのような変化が生じるかを計測するためである。

【0086】

まず、物体検出部 101 により、遮蔽状態がない人体の検出スコア $S_{c,i}$ を算出する。次に、遮蔽パターン毎に人体の検出スコア $S_{B,i}$ を算出する。そして、遮蔽パターン毎に遮蔽面積と検出スコアとの相関を最小二乗法などにより計算する。遮蔽面積と検出スコア

10

20

30

40

50

との相関は、遮蔽パターン毎に異なるため、起こりうる遮蔽状態のパターンを予め想定して遮蔽パターン数のスコア補正辞書を保持しておくのが望ましい。

【0087】

最後に、遮蔽の状態に応じて予め算出しておいた遮蔽パターン p における補正係数 $S_{O,p}$ を検出スコア $S_{B,i}$ に乗算することにより、以下の式(9)に従って検出スコアを計算する。本実施形態では、人体の頭部や顔以外の全身領域において遮蔽が生じることを想定しているため、全身検出器から出力される検出スコア $S_{B,i}$ にのみ乗算した。一方、構築するシステムによっては、頭部検出器から出力される検出スコア $S_{F,k}$ や顔検出器から出力される補正スコア $S_{H,j}$ に、遮蔽状態に応じた補正スコアを乗算してもよい。

【0088】

【数9】

$$R_i = (X_{HH,j} \cdot S_{B,i} \cdot S_{O,p} + S_{H,j} + S_{F,k}) \quad \dots (9)$$

【0089】

以上のように、遮蔽パターン毎にスコア補正辞書を用意し、検出スコアの補正をすることにより、遮蔽状態であっても精度良く人体検出を行うことができる。

【0090】

次に、ステップ S1207 において、検出スコアが閾値以上のすべての検出結果に対して遮蔽判定を行ったか否かを判定する。この判定の結果、まだ遮蔽判定を行っていない検出結果が残っている場合は、ステップ S1202 に戻り、処理を繰り返す。処理を繰り返す場合は、ステップ S1202 に戻って前景となる検出結果を選択し直し、検出スコアが閾値未満の検出結果しかなかった場合は処理を終了する。

【0091】

図11は、2人以上の人物が重なって遮蔽が生じている状態の一例を示す図である。図11に示す例では、一番手前の人体1100により後方の人体1101に遮蔽が生じており、さらに後方の人体1102は、2つの人体1100、1101に遮蔽されている。遮蔽領域1103は、2つの人体1100、1101の影響により人体1102に生じる遮蔽領域を示している。

【0092】

前述したように、透視投影の原理を踏まえ、ステップ S1201 では、物体検出部 101 で出力された全ての検出結果を検出スコアと検出枠のサイズとで大きい順にソートする。この処理により、遮蔽が生じている可能性が低く、かつ遮蔽の原因になる手前の人体1100を特定することができる。そして、遮蔽の原因になりそうな人体から順に、他の検出結果との遮蔽判定を行う。

【0093】

また、人体1102は人体1101の近隣にあり、さらに頭部検出器は人体1102の頭部に対して高い検出スコアを出力する。そのため、人体1102の頭部の検出スコアが人体1101の頭部の検出スコアよりも大きい場合には、人体1102の頭部と人体1101の全身とが融合したような領域が、人体1101の全身の検出結果として出力される。また、人体1102は人体1101の一部として検出されるため、正しく検出されない。このような誤検出が生じた場合には、本実施形態で説明した遮蔽状態判定方法を用い、人体1102に遮蔽状態が生じていると判定する。そして、人体1102の頭部のサイズのスケールに合う全身矩形を選択し、かつ人体1101の頭部とその全身矩形を検出することによって誤検出を防ぐようにする。

【0094】

以上のように本実施形態によれば、従来の手法に比べ、検出対象物に遮蔽が存在した場合でも検出スコアを落とさずに検出することができるようになる。また、包含関係により求めた遮蔽領域の面積や遮蔽パターン毎に検出スコアを補正しているため、最終的に出力する検出結果は、従来よりも精度が向上する。さらに、物体検出過程で算出される頭部検出器と全身検出器との出力結果をそのまま利用し、高度で複雑な計算をせずに検出スコア

10

20

30

40

50

を補正することができる。

【0095】

(第2の実施形態)

本実施形態では、対象物を複数の移動可能なパーツに分割して検出する検出処理部を用いた場合の遮蔽状態の判定方法、並びに検出スコアの補正方法について説明する。本実施形態においても、検出対象物は人物とし、共通部位は人物の頭部とした例について説明する。ただし、第1の実施形態では、人物が常に直立した状態であることを想定していたが、本実施形態では、前屈みになった前傾姿勢や、しゃがみといった姿勢変化にも対応できる。また、第1の実施形態で説明した内容と同じ構成及び処理については説明を省略する。

10

【0096】

なお、本実施形態に係る物体検出装置10の全体構成は、基本的には第1の実施形態で説明した図1及び図2と同様である。但し、第1の検出処理部211～第nの検出処理部21nの検出対象が異なっており、統合結果出力部250の処理内容も異なっている。また、本実施形態で使用する検出器としては、頭部検出器と全身検出器とを使用する例について説明する。対象物の小さな姿勢変化などに対応した検出を行うために、例えば、非特許文献6に記載されているようなパーツベースの検出手法が知られている。

【0097】

図13は、パーツベースの検出手法を用いた全身検出器を説明する図である。図13の点線の矩形1302は、全身検出器の一つのパーツであり、図13に示す例では、全身検出器は8個のパーツで構成されている。また、実線の矩形1301は、パーツベースの検出の結果得られる全身の検出結果である。

20

【0098】

図13の(A)に示す例と図13(B)に示す例とでは、人物の姿勢が異なるため、検出の結果得られる各パーツの位置もそれぞれ異なる。パーツベースの検出結果では、各パーツの検出スコアと位置関係とに基づいて算出される全体としての検出スコアが得られるとともに、図13に示す実線や破線で表される対象物や各パーツの位置・範囲の情報が得られる。

【0099】

以下、このようなパーツベースの検出器を用いた場合に、検出結果から頭部の位置を推定(共通部位を推定)する例について説明する。まず、パーツベースの検出器の結果から、第1の共通部位推定部221～第nの共通部位推定部22nで頭部の位置を推定する処理について説明する。簡単な場合としては、頭部を検出対象とするパーツが含まれる場合には、その頭部のパーツ位置を頭部の位置の推定結果とすればよい。また、頭部のパーツが推定する頭部範囲と一致しない場合(例えば、頭部から肩までを検出対象としたパーツがある場合)は、第1の実施形態で説明したように、検出結果の頭部パーツから頭部位置を推定すればよい。

30

【0100】

一方、図13に示すように、頭部を明確に示さないパーツ群で検出器が構成されている場合には、複数のパーツの位置情報を用いて頭部の位置を推定することができる。複数のパーツの位置情報から頭部の位置を推定する場合には、各パーツの座標情報を並べたベクトルから、推定する頭部の位置を線形変換により求める。8個のパーツから頭部位置の左上のx座標 x_{h1} を推定する線形変換式としては、例えば以下の式(10)に示すものを用いて頭部位置を推定する。

40

【0101】

【数 1 0】

$$x_{h1} = \mathbf{X}_p^T \mathbf{B}_{h1} = \left(1, \frac{x_{p1}}{w}, \frac{y_{p1}}{h}, \frac{x_{p2}}{w}, \frac{y_{p2}}{h} \dots \frac{x_{pn}}{w}, \frac{y_{pn}}{h} \right) (b_0, b_1, b_2, b_3, b_4 \dots b_{2n-1}, b_{2n})^T \dots (10)$$

【0 1 0 2】

ここで、 X_p はパーツ座標のベクトルに定数 1 を追加したものであり、 B_{h1} は変換係数ベクトルである。また、 x_{pn} 、 y_{pn} はそれぞれ n 番目のパーツの中心座標であり、 b は x_{h1} 座標を求めるためのそれぞれの項の変換係数であり、定数項 b_0 を含む。また、 w 、 h はそれぞれ対象物領域（図 1 3 に示す実線の矩形 1 3 0 1）の幅、高さである。なお、頭部の推定位置 X_h を求めるためには、異なる変換係数を用いて同様に y_{h1} 、 x_{h2} 、 y_{h2} を求めればよい。

【0 1 0 3】

前述した例では、各パーツの中心座標のみから頭部の位置を推定しているが、検出の結果得られる対象物領域（図 1 3 に示す実線の矩形 1 3 0 1）の座標情報をパーツ座標ベクトルに加えてもよい。変換係数ベクトル B は、頭部の正解基準を与えた画像サンプル群と、その画像サンプル群へのパーツベース検出器の検出結果から最小二乗法により求めることができる。頭部位置を推定する方法は、最小二乗法に限定することなく、頭部位置を目的変数、複数のパーツ位置を説明変数として他の回帰分析で求めることもできる。

【0 1 0 4】

また、式（1 0）の変換係数ベクトル B を計算する際に、頭部の正解基準を与える代わりに、人体の体軸（喉元から腰の中心を結んだ直線）を正解基準とする画像サンプル群を与えることにより、複数のパーツ群から体軸を推定することが可能になる。例えば、図 1 4 に示すように、人体を直立した状態から少し前傾にした際の姿勢や、更に屈んだしゃがみの姿勢なども検出することが可能である。この場合、体軸 1 4 0 1（喉元と腰の中心とを結ぶ直線）の各点の座標セット X_U は、以下の式（1 1）により求めることができる。

【0 1 0 5】

【数 1 1】

$$\mathbf{X}_U = (x_{U1}, y_{U1}, x_{U2}, y_{U2})^T \dots (11) \quad 30$$

【0 1 0 6】

体軸 1 4 0 1 を推定する際には、図 1 4（A）の波線の矩形 1 4 0 2 に示す 8 個のパーツから頭部の位置を推定する式（1 0）に示した線形変換式を用いる。頭部の正解基準の代わりに人体の体軸 1 4 0 1（喉元から腰の中心まで）を正解基準として与えた画像サンプル群と、その画像サンプル群へのパーツベース検出器の検出結果とから、頭部の推定と同様に最小二乗法により体軸 1 4 0 1 を推定することが可能である。体軸を推定する方法は、前述と同様に最小二乗法に限定するものではなく、構築するシステムにおいて好適な手段を用いればよい。更に、頭部や体軸以外にも頭部と体軸とを含む上半身矩形を同様の処理で算出することも可能である。

【0 1 0 7】

ここで、8 個のパーツ群のそれぞれの検出器は、図 1 4（B）の波線の矩形 1 4 0 4 に示すように、人体の傾きによってその位置は大きく変動する。そのため、体軸 1 4 0 3 を精度良く算出するには、人体の様々な傾き（姿勢）毎にカテゴリ分けして、変換係数ベクトル B を計算しておくことよい。ただし、姿勢のカテゴリを多くすると、姿勢の識別精度は高まる一方で、計算量も増加するため、姿勢のカテゴリ数は計算量とのバランスに応じて設定する。したがって、構築するシステムによって識別したい姿勢のカテゴリ数を決める必要がある。

【0 1 0 8】

< 遮蔽判定方法とスコア補正方法 >

次に、本実施形態における遮蔽判定方法と検出スコアの補正方法とについて説明する。

図15は、本実施形態における複数のパーツ検出器による頭部の推定結果と体軸の推定結果とから、遮蔽判定を行う事例を説明する図である。人体1501、1509は、同一の姿勢で同じ人体である。図15(A)に示す人体1501は、前景の人体1500により部分的に遮蔽されており、図15(B)に示す人体1509は、前屈みになった前景の人体1508により部分的に遮蔽されている。

【0109】

まず、図15(A)において、第1の実施形態で説明した方法により、人体の頭部検出枠(矩形1502、1505)、及び全身検出枠(矩形1503、1506)を算出する。そして、本実施形態で前述した方法により、体軸1504、1507を算出する。ここで、人体1501は部分的に遮蔽されているため、検出スコアが低下し、この段階では本来の検出スコアで検出することはできない。

10

【0110】

図16は、遮蔽領域1603の面積と位置とを算出する方法を説明する図である。遮蔽領域1603の面積および位置は、図14に示した8個のパーツ検出器(図16の点線1600~1602)と矩形1502~1507が示す各種検出結果とから算出する。

【0111】

図16(A)において、体軸1504、1507は、鉛直上向きを基準としたときの角度が10度未満である。そのため、前景の人体および後景の人体の姿勢は「どちらも立位である」と想定することができる。この場合は、第1の実施形態で説明した方法により遮蔽領域1603の面積および位置を計算する。そして、算出した遮蔽領域と重畳するパーツ検出器(図16(A)の点線1600~1602)を8個の中から特定し、人体1501の検出スコアを補正する。

20

【0112】

具体的な補正方法は、第1の実施形態で説明した方法と同様に、遮蔽が存在しない画像サンプル群と遮蔽が存在する画像サンプル群とをそれぞれ入力し、どのパーツ検出器が遮蔽されているかで場合分けを行う。そして、どのパーツ検出器が遮蔽された場合にどの程度統合スコアが低下するのかの相関を計算しておく。そして、遮蔽状態に応じた補正係数をかけない場合は、遮蔽領域1603によって遮蔽されたパーツ検出器を特定し、そのパーツ検出器が遮蔽されていない場合の出力値を代わりに、パーツ検出器の出力値とすることにより、統合スコアの低下を防ぐ。また、遮蔽が無い場合のパーツの検出スコアを予め算出していなくても、統合スコアが下がらないように適当な定数を代入するなどしてもよい。

30

【0113】

一方、図15(B)に示す例では、前傾になっている人体1508が直立した人体1509を部分的に遮蔽している。図15(B)に示す例の場合も、図15(A)と同様に頭部および全身の検出結果が出力され、体軸1510、1511が算出されている。

【0114】

図16(B)に示すように、体軸1510は、鉛直上向きを基準とした場合に、角度が10度以上あるため、第1の実施形態で用いた計算方法をそのまま利用することはできない。そこで、角度が10度以上の場合、体軸1510の直線を延長し、後景の人体の全身検出枠との交点(x_{U3} , y_{U3})を算出する。そして、体軸1510を延長した直線と後景の全身検出枠で囲まれる領域とを遮蔽領域1604として算出する。

40

【0115】

検出スコアの補正方法については、図16(A)に示す補正方法と同様に遮蔽領域1604と重畳する領域を有するパーツ検出器を特定し、予め計算した遮蔽が無い場合のパーツ検出器の検出スコアを代わりに採用することによって補正を行う。

【0116】

本実施形態におけるパーツ検出器は、人体の姿勢が変動することにより、個々の位置及

50

び検出スコアが大きく変化する。そのため、上記のスコア補正を行う際に、姿勢カテゴリ毎にパーツ検出器の移動量と検出スコアとの相関関係を非特許文献6に記載された方法を用いて予め計算しておく必要がある。

【0117】

また、体軸1510は、人体1508の内部を示す直線であるため、実際の遮蔽領域を示すものではない。したがって、厳密に遮蔽領域を推定する必要がある場合は、体軸1510の代わりに人体1508のエッジ情報を抽出して遮蔽領域を求めてもよい。

【0118】

図17は、本実施形態における遮蔽判定部102による処理手順の一例を示すフローチャートである。

まず、ステップS1201の処理は、第1の実施形態で説明した図12のステップS1201の処理とほぼ同一の処理である。

【0119】

次に、ステップS1701において、前景と推定される人体の全身及び頭部の検出結果と、体軸の検出結果とを参照し、前景と推定される人体を選択する。以下、体軸の検出結果を A_U とする。

【0120】

次に、ステップS1702において、後景の人体の全身及び頭部の検出結果と、体軸の検出結果とを参照し、ステップS1701で選択した人体の検出スコアよりも低く、かつ頭部の検出結果に対応する全身の検出結果が算出されている人体を選択する。以下、体軸の検出結果を B_U とする。このように、本実施形態では、頭部の検出結果及び全身の検出結果の他に、体軸の検出結果も参照している。

【0121】

次に、ステップS1703において、ステップS1701で参照した前景に存在するであろう人体Aの体軸 A_U の直線、あるいは体軸 A_U を延長した直線と後景に存在するであろう人体Bの全身の検出結果 B_B との交差判定を行う。この判定の結果、交点がない場合は、遮蔽している物体がないため、ステップS1701に戻る。

【0122】

一方、ステップS1703の判定の結果、交点がある場合は、ステップS1704において、鉛直上向きを基準とした体軸 A_U の傾き(角度)を算出する。そして、ステップS1705において、体軸 A_U の角度が10度未満であるか否かを判定する。この判定の結果、体軸 A_U の角度が10度未満である場合は、前景の人体が"立位"の姿勢であると判定できるので、ステップS1706の処理に移る。一方、体軸 A_U の角度が10度以上である場合は、前景の人体が"前傾"あるいは"しゃがみ"の姿勢であると判定し、ステップS1707の処理に移る。ここでは説明を簡単にするために体軸の角度10度を基準に処理を変えているが、構築するシステムによっては、立位を示す体軸の角度を任意に設定してもよい。

【0123】

ステップS1706においては、前景の人体が立位であると判定し、第1の実施形態で説明した遮蔽領域の計算方法を用いて遮蔽領域の位置と面積を算出する。一方、ステップS1707において、体軸 A_U あるいは体軸 A_U を延長した直線と全身検出結果 B_B とで囲まれる領域の位置および面積を算出する。

【0124】

次に、ステップS1708において、8個のパーツ検出器の中でどのパーツ検出器が遮蔽されていたかを特定し、遮蔽が無い場合のパーツ検出器の検出スコアに置き換えて統合スコアを補正する。なお、パーツ検出器の検出スコアは予め設定した定数で置き換えるなどしてもよい。ステップS1207については、図12のステップS1207と同様である。

【0125】

本実施形態では、後景の人物が立位の場合を例に説明したが、立位以外の前傾やしゃが

10

20

30

40

50

みといった姿勢でも同様の処理で遮蔽を検出することができる。以上、全身検出器から推定される頭部位置の推定枠を統合する処理については第1の実施形態と同様である。本実施形態では、立位以外の姿勢で遮蔽された場合にも遮蔽の姿勢に応じた検出スコアの補正を行うことができるため、精度良く検出できるようになる。

【0126】

(第3の実施形態)

第1及び第2の実施形態では、後景の頭部以外の部位に遮蔽が生じていた場合について説明した。本実施形態では、前景及び後景の人物の頭部、或いは共通部位が重なった場合に、遮蔽を検出する方法について説明する。本実施形態でも、検出対象物は人物とし、推定する共通部位は人物の頭部とする。また、以下の説明では、前景及び後景の人物の頭部が重なっている場合は、遮蔽フラグがONであるものと定義して説明する。また、第1及び第2の実施形態で説明した内容と同じ構成及び処理については説明を省略する。

10

【0127】

図18は、前景の人体1800によって後景の人体1801が遮蔽されており、かつ両人体の頭部が重なっている場合の検出処理を説明する図である。

図18に示すような場面において、第1または第2の実施形態で説明した物体検出部101を用いて検出処理を行うと、人体が2体あるにも関わらず頭部の検出結果1802の1体分しか検出されない。これは、物体検出部101の統合処理の結果、検出スコアの低い体軸及び全身の検出結果は、検出スコアの高い頭部と全身及び体軸の検出結果に統合されてしまうためである。

20

【0128】

そこで本実施形態では、統合処理を行う前に体軸の検出結果を利用して頭部同士の重なりの有無を検出し、頭部同士の重なりが在った場合には、物体検出部101の統合処理の処理内容を頭部基準から体軸基準に変更する。これは、頭部基準で統合処理すると得られない正しい出力結果を、体軸基準の統合処理に処理内容を変更することによって得られるようにするためである。

【0129】

なお、本実施形態のように頭部が重なり合う場合、それぞれの統合スコアはそれほど低下しない。これは、本実施形態ではパーツベースの検出器で物体検出するため、物体の不連続性の影響を特に受けないためである。したがって、本実施形態では頭部の重なりによる遮蔽が生じていても検出スコアの低下は少ないため、検出スコアの補正処理は行わないようにする。ただし、共通部位を頭部以外に設定した場合は、構築するシステムの特性に応じて遮蔽の有無により検出スコアの補正を行ってもよい。

30

【0130】

図19は、本実施形態における遮蔽判定部102による遮蔽判定の処理手順の一例を示すフローチャートである。まず、第2の実施形態で説明した物体検出部101の処理により、人体1800、1801の頭部の検出結果1802、全身の検出結果1803、1804、及び体軸1805、1806がそれぞれ仮に算出されているものとする。また、ステップS1201、S1207、S1701及びS1702の処理は、第2の実施形態で説明した図17の処理と同一の処理であるため、説明は省略する。

40

【0131】

次に、ステップS1901において、ステップS1701及びS1702で参照した人体Aの体軸 A_U を延長した直線と、人体Bの体軸 B_U を延長した直線との交点を算出する。そして、頭部の検出結果1802の矩形内部に交点が存在するかどうかを判定する。この判定の結果、交点が矩形内部にない場合は頭部検出の結果で統合処理ができるため、遮蔽フラグはOFFのまま、ステップS1902に進む。一方、交点が矩形内部にある場合は、各体軸と組になる頭部が重なり合っていると判定し、遮蔽フラグをONにして、ステップS1903に進む。

【0132】

ステップS1902においては、ステップS1901で人体Aおよび人体Bの頭部が重

50

なり合っていないと判断できるため、第1及び第2の実施形態で説明した頭部検出結果を含めた統合処理により検出スコアを算出する。

【0133】

そして、ステップS1903においては、人体Aもしくは人体Bの頭部によって、遮蔽状態が生じていると判断できるため、統合処理において頭部検出結果は利用せずに、第2の実施形態で説明した体軸の角度を基準とした統合処理を行う。

【0134】

図20は、統合処理前の頭部および体軸の検出結果の一例を示す図である。図20において、頭部の検出結果の候補2001は、検出結果の候補群を示しており、それぞれ体軸1806、2002と対応している。

10

【0135】

第1の実施形態で説明した統合処理を実施すると、頭部の検出結果の候補2001は、頭部の検出結果1802に統合されてしまう。なお、第1の実施形態の統合処理では、検出スコアと頭部の重複度を基準に処理を行うため、頭部が遮蔽された人体1801の頭部の検出結果は統合処理によって、人体1800の頭部の検出結果1802に吸収されてしまう。つまり、遮蔽フラグが立つ場合に、頭部を基準にして統合処理を行うと、遮蔽された人体1801の頭部、体軸、全身のそれぞれの検出結果が遮蔽している人体1800の各検出結果に吸収されてしまう。

【0136】

そこで本実施形態では、遮蔽フラグがONである場合は、頭部基準から体軸基準に統合処理の基準を変更する。図20に示すように、体軸2002も頭部の検出結果1901と同様に複数検出されるため、最も高い検出スコアの体軸と各体軸の検出候補の直線とのなす角度を基準に検出結果を統合する。例えば、体軸1805と他の体軸の検出結果とを角度基準（例えば閾値を30度）で比べると、体軸2002はいずれもなす角が30度未満なので体軸1805に統合される。なお、体軸1806は30度以上なす角に違いがあるため統合されない。また、角度の代わりに体軸の検出結果の端点（喉元および腰の中心座標）を利用し、2直線の距離を基準に統合処理を行ってもよい。

20

【0137】

なお、本実施形態では、遮蔽が生じた場合に検出結果の統合処理によって本来検出される検出結果の候補が、他の結果に吸収されることを防止することを目的とした例外処理である。検出結果によっては遮蔽フラグがONとなる検出候補が多くなる場合もあり、遮蔽フラグがONとなる検出候補の検出スコアの値でソート処理を施し、検出スコアの高い順に予め設定した数（例えば1、2など）の検出結果を検出候補として選択する。検出候補に選択されなかった検出結果は統合処理により削除される。検出候補として残す数は、構築するシステムの環境により設定することが望ましい。

30

【0138】

以上のように本実施形態の統合方法では、前後で違う姿勢を取っている複数の人体に対して体軸の検出結果を活用して頭部の検出結果が重なっていても、その後の統合処理で削除されることなく頑健に人体検出を行うことができる。

【0139】

（その他の実施形態）

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。

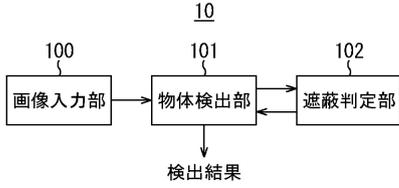
40

【符号の説明】

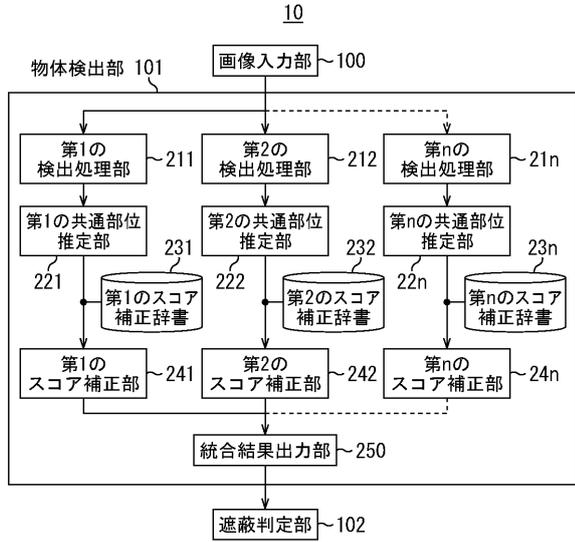
【0140】

- 101 物体検出部
- 102 遮蔽判定部

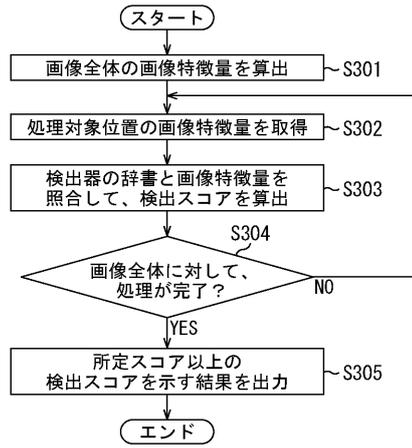
【 図 1 】



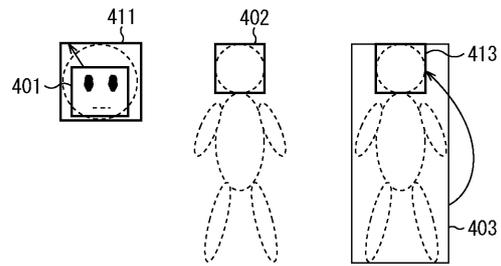
【 図 2 】



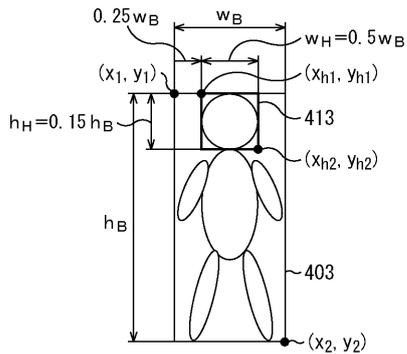
【 図 3 】



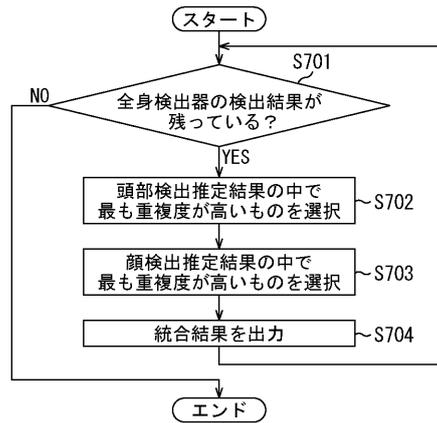
【 図 4 】



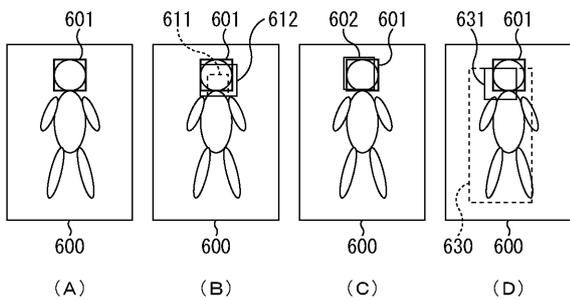
【 図 5 】



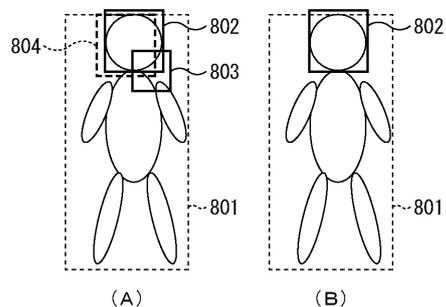
【 図 7 】



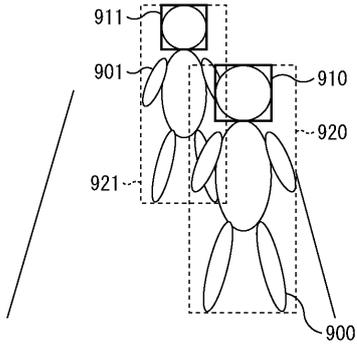
【 図 6 】



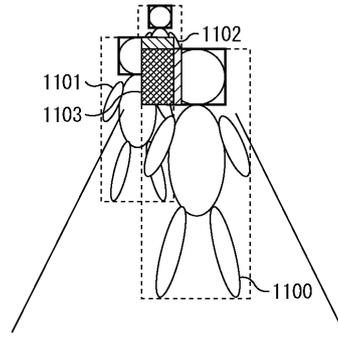
【 図 8 】



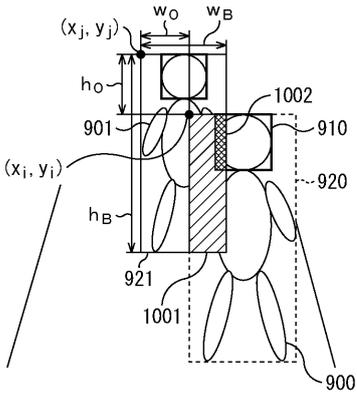
【 図 9 】



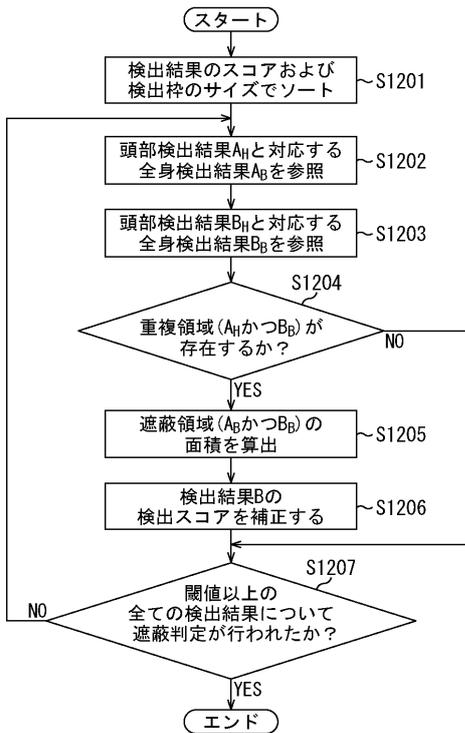
【 図 1 1 】



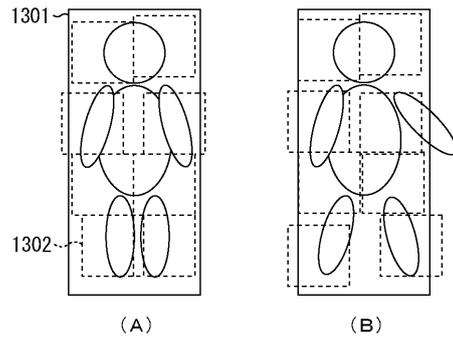
【 図 1 0 】



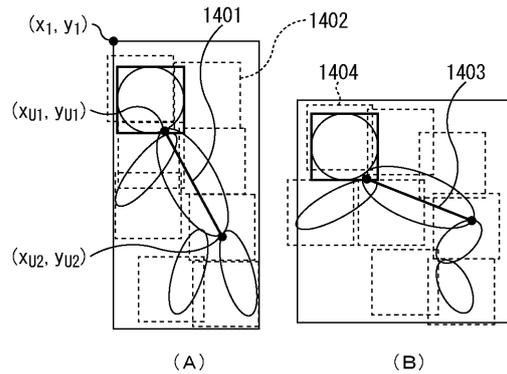
【 図 1 2 】



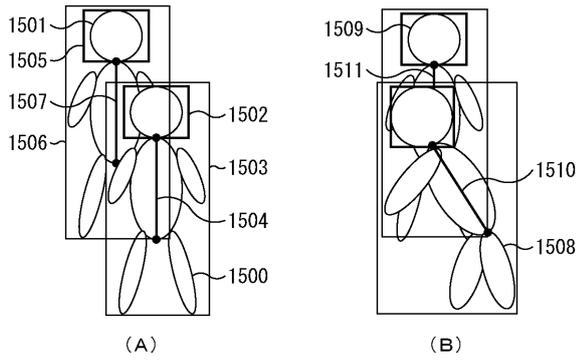
【 図 1 3 】



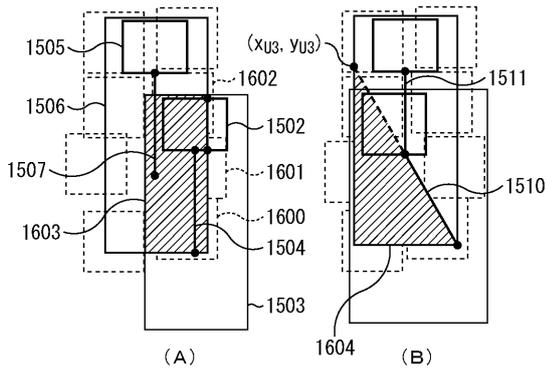
【 図 1 4 】



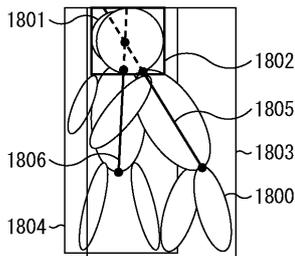
【図15】



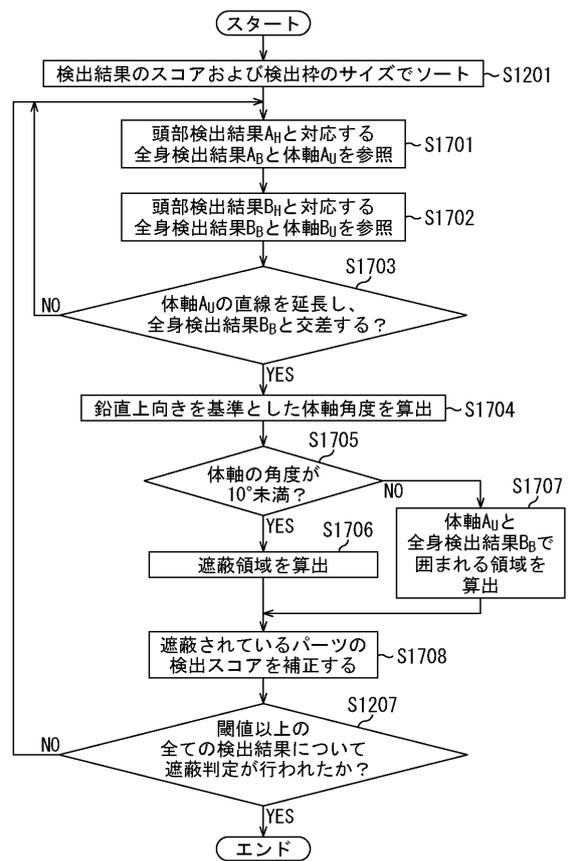
【図16】



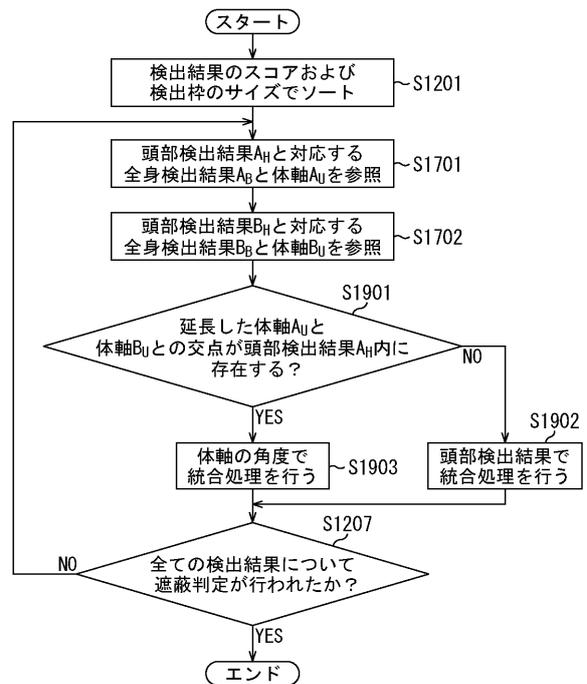
【図18】



【図17】



【図19】



【 図 2 0 】

