

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7540974号
(P7540974)

(45)発行日 令和6年8月27日(2024.8.27)

(24)登録日 令和6年8月19日(2024.8.19)

(51)国際特許分類 F I
G 0 6 T 7/292(2017.01) G 0 6 T 7/292

請求項の数 12 (全16頁)

(21)出願番号	特願2021-83598(P2021-83598)	(73)特許権者	000208891 K D D I 株式会社 東京都新宿区西新宿二丁目 3 番 2 号
(22)出願日	令和3年5月18日(2021.5.18)	(74)代理人	100092772 弁理士 阪本 清孝
(65)公開番号	特開2022-177391(P2022-177391 A)	(74)代理人	100119688 弁理士 田邊 壽二
(43)公開日	令和4年12月1日(2022.12.1)	(72)発明者	小林 達也 埼玉県ふじみ野市大原二丁目 1 番 1 5 号 株式会社 K D D I 総合研究所内
審査請求日	令和5年8月7日(2023.8.7)	審査官	藤原 敬利

最終頁に続く

(54)【発明の名称】 複数物体追跡装置、方法及びプログラム

(57)【特許請求の範囲】

【請求項 1】

撮影範囲がオーバーラップする複数のカメラで撮影した画像に基づいて複数の物体を追跡する複数物体追跡装置において、

カメラ画像ごとに物体を検出する手段と、

各物体の検出結果と各物体の予測位置とを多対一で対応付ける手段と、

前記検出した物体のカメラ画像上での大きさを計算する手段と、

前記検出した物体のカメラ画像上での遮蔽度合いを計算する手段と、

前記検出した物体のカメラ画像上での大きさの降順で、前記遮蔽度合いが所定の閾値以下の上位少なくとも一つのカメラを、その検出結果を追跡結果の更新に用いるカメラとして選択する手段と、

前記選択する手段が選択したカメラの検出結果及び前記多対一の対応付けに基づいて追跡結果を更新する手段とを具備し、

前記追跡結果の更新過程で予測した各物体の予測位置を前記対応付ける手段にフィードバックすることを特徴とする複数物体追跡装置。

【請求項 2】

相互に対応付けられなかった検出結果及び予測位置に基づいて、追跡を開始する物体及び追跡を終了する物体の少なくとも一方を処理する手段を更に具備したことを特徴とする請求項 1 に記載の複数物体追跡装置。

【請求項 3】

10

20

前記カメラを選択する手段は、遮蔽度合いが所定の閾値以下のカメラが存在しないと、カメラ画像上での大きさが所定の閾値を超えるカメラから遮蔽度合いが最低のカメラを選択することを特徴とする請求項 1 または 2 に記載の複数物体追跡装置。

【請求項 4】

相互に遮蔽関係にある複数の物体を対象に、一の物体の画像特徴量と他の物体の画像特徴量との類似度を計算する手段を具備し、

前記カメラを選択する手段は、カメラ画像上での物体の大きさ、遮蔽度合い及び画像特徴量の類似度に基づいてカメラを選択することを特徴とする請求項 1 ないし 3 のいずれかに記載の複数物体追跡装置。

【請求項 5】

前記カメラを選択する手段は、カメラ画像上での物体の大きさ、遮蔽度合い及び画像特徴量の類似度の重み付き平均に基づいてカメラを選択することを特徴とする請求項 4 に記載の複数物体追跡装置。

【請求項 6】

前記カメラを選択する手段は、遮蔽度合いよりも画像特徴量の類似度を優先してカメラを選択することを特徴とする請求項 5 に記載の複数物体追跡装置。

【請求項 7】

前記遮蔽度合いを計算する手段は、カメラごとに一の物体の検出結果と他の物体の予測位置の当該カメラへの投影領域との遮蔽度合いを当該一の物体と他の物体との遮蔽度合いとみなすことを特徴とする請求項 1 ないし 6 のいずれかに記載の複数物体追跡装置。

【請求項 8】

前記遮蔽度合いを計算する手段が、
物体の外接矩形をベースに遮蔽度合いを計算する手段と、
物体の回転矩形をベースに遮蔽度合いを計算する手段とを具備し、
前記外接矩形をベースに計算した遮蔽度合いの最小値が前記所定の閾値以上であると前記回転矩形をベースに遮蔽度合いを計算することを特徴とする請求項 1 ないし 7 のいずれかに記載の複数物体追跡装置。

【請求項 9】

物体の3Dモデルをカメラ画像へ投影して得られる輪郭をベースに遮蔽度合いを計算する手段を更に具備し、

前記回転矩形をベースに計算した遮蔽度合いの最小値が前記所定の閾値以上であると前記輪郭をベースに遮蔽度合いを計算することを特徴とする請求項 8 に記載の複数物体追跡装置。

【請求項 10】

前記輪郭をベースに遮蔽度合いを計算する手段は、Instance Segmentationによって検出した領域に基づいて遮蔽度合いを計算することを特徴とする請求項 9 に記載の複数物体追跡装置。

【請求項 11】

コンピュータが、撮影範囲がオーバーラップする複数のカメラで撮影した画像に基づいて複数の物体を追跡する複数物体追跡方法において、

カメラ画像ごとに物体を検出し、
各物体の検出結果と各物体の予測位置とを多対一で対応付け、
前記検出した物体のカメラ画像上での大きさを計算し、
前記検出した物体のカメラ画像上での遮蔽度合いを計算し、
前記検出した物体のカメラ画像上での大きさの降順で、前記遮蔽度合いが所定の閾値以下の上位少なくとも一つのカメラを、その検出結果を追跡結果の更新に用いるカメラとして選択し、

前記更新に用いるカメラとして選択したカメラの検出結果及び前記多対一の対応付けに基づいて追跡結果を更新し、

前記追跡結果の更新過程で予測した各物体の位置を前記各物体の検出結果と対応付ける

10

20

30

40

50

ためにフィードバックすることを特徴とする複数物体追跡方法。

【請求項 1 2】

撮影範囲がオーバーラップする複数のカメラで撮影した画像に基づいて複数の物体を追跡する複数物体追跡プログラムにおいて、

カメラ画像ごとに物体を検出する手順と、

各物体の検出結果と各物体の予測位置とを多対一で対応付ける手順と、

前記検出した物体のカメラ画像上での大きさを計算する手順と、

前記検出した物体のカメラ画像上での遮蔽度合いを計算する手順と、

前記検出した物体のカメラ画像上での大きさの降順で、前記遮蔽度合いが所定の閾値以下の上位少なくとも一つのカメラを、その検出結果を追跡結果の更新に用いるカメラとして選択する手順と、

10

前記選択する手順で選択したカメラの検出結果及び前記多対一の対応付けに基づいて追跡結果を更新する手段とを手順と、

前記追跡結果の更新過程で予測した各物体の位置を前記各物体の検出結果と対応付けるためにフィードバックする手順と、をコンピュータに実行させることを特徴とする複数物体追跡プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、撮影範囲がオーバーラップする複数のカメラで撮影した複数の物体を高速かつオクルージョンに対して頑健に追跡する複数物体追跡装置、方法及びプログラムに関する。

20

【背景技術】

【0002】

カメラ画像中の複数物体を追跡する技術は監視カメラ、移動ロボットあるいは自動運転車等の様々な利用シーンでの活用が期待されている。しかしながら、カメラ画像中の複数物体を高精度に検出し、カメラ画像中で複数物体が重複するオクルージョンが頻発する状況下において長期に渡り軌跡を高精度に推定し続けることは依然として困難である。

【0003】

オクルージョン発生時における高精度な物体追跡を実現するための様々な技術が提案されている。特許文献1には、オクルージョン発生時に画像中で複数物体の重なる領域に対して個々の物体の特徴に基づき個々の物体の存在位置を特定する技術が開示されている。特許文献2には、オクルージョンの発生を検知した際にオクルージョンが及んでいない部分領域に特化した追跡を行うことで個々の物体の追跡を維持する技術が開示されている。

30

【0004】

オクルージョンに対する頑健性を向上させるためのアプローチとしては、撮影領域がオーバーラップする複数のカメラを用い、複数視点の情報に基づいて個々の物体を追跡することも有効である。

【0005】

そのためには、カメラの配置に関する情報を用いて、個々のカメラの画像と共通の世界座標系との対応付けを行い、個々のカメラの画像中の物体位置を共通の世界座標系の物体位置に変換した上で、同一物体をカメラ間で正しく対応付ける必要がある。

40

【0006】

特許文献3には、個々のカメラ中の追跡物体を複数カメラ間で正確に対応付ける技術が開示されている。

【0007】

非特許文献1には、個々のカメラで検出・追跡した複数物体の軌跡 (tracklet) の類似度を世界座標系上で評価し、複数カメラ間で軌跡を統合する技術が開示されている。

【0008】

非特許文献2には、個々のカメラで検出した結果を追跡中の物体と多対一で対応付け、

50

複数の検出結果と追跡結果の対応に基づき、複数カメラで協調を取りながら世界座標系における個々の物体追跡を行う技術が開示されている。

【0009】

特許文献4には、個々のカメラの物体検出結果を一つの統合追跡手段に入力し、統合追跡手段において物体と各カメラとの距離を考慮して、世界座標系における個々の物体追跡を行う技術が開示されている。

【0010】

特許文献5には、カメラ画像中の物体間の重なりを考慮して、オクルージョンが最も生じづらいカメラを使って物体追跡を行う技術が開示されている。

【先行技術文献】

【特許文献】

【0011】

【文献】特開2009-15827号公報

【文献】特開2016-126624号公報

【文献】特開2019-96062号公報

【文献】特開2019-114280号公報

【文献】特開2001-94975号公報

【非特許文献】

【0012】

【文献】Yuhang He et al. "Multi-Target Multi-Camera Tracking by Tracklet-to-Target Assignment," in IEEE Trans. on Image Processing, 2020.

【文献】Li He et al. "Efficient Multi-View Multi-Target Tracking Using a Distributed Camera Network," in IEEE Sensors Journal, 2020.

【発明の概要】

【発明が解決しようとする課題】

【0013】

上記のいずれの方法でも、撮影範囲がオーバーラップする多数のカメラが存在する環境下において高速かつオクルージョンに対して頑健に複数物体を追跡することは依然として困難であった。

【0014】

すなわち、特許文献1, 2が開示する、オクルージョン発生中の物体を頑健に追跡する手法は、物体が部分的に遮蔽される場合には対応可能であるが、完全に遮蔽される場合の追跡が困難である。

【0015】

特許文献3が開示するように、複数カメラ間で追跡物体を対応付ける手法は、撮影範囲がオーバーラップするカメラ台数が増えた際に対応付けの組み合わせがカメラ台数 N_c の2乗のオーダーで増えるため、高速な処理が困難になる。

【0016】

例えば、天井に多数（例えば20台）の全方位カメラを設置するようなケースでは、特定の物体が20台のカメラで同時に撮影され得るため、カメラ同士の組み合わせが190通り存在し、対応付けの処理負荷が増大する。

【0017】

非特許文献1が開示するように、軌跡の類似度を評価する手法も同様の理由で、カメラごとに追跡した軌跡同士の類似度を評価するための処理負荷の増大が問題になる。

【0018】

非特許文献2や特許文献4が開示するように、個々のカメラで検出した結果を追跡中の物体と多対一で対応付け、複数の検出と追跡の対応に基づいて世界座標系における個々の物体追跡を行う手法は、撮影範囲がオーバーラップするカメラ台数が増加しても対応付けの回数が線形のオーダーで増加するため、比較的処理負荷に優れる。しかしながら、一部のカメラでオクルージョンが発生した際の人物位置を精度良く追跡することが困難である。

10

20

30

40

50

【0019】

特許文献5が開示するように、カメラごとにオクルージョンの度合いを判定し、最もオクルージョンの発生しづらいカメラにおいて物体を追跡する手法は、特定のシーンにおいてはオクルージョンに頑健な物体追跡を行うことが期待される。

【0020】

しかしながら、俯瞰視点における複数物体とカメラとを結ぶ直線の角度に基づいて重なり度合いを評価する手法では、天井カメラ等、俯瞰に近い角度で撮影されるケースでは、正確にオクルージョンの度合いを推定することができない。

【0021】

また、カメラと追跡物体の幾何的な情報のみに基づいて担当カメラを判定するため、障害物の影響で特定の条件で物体が検出できないような場合には対応ができない。

10

【0022】

本発明の目的は、上記の技術課題を解決し、個々のカメラで検出した結果を追跡中の物体と多対一で対応付け、複数の検出と追跡との対応関係に基づいて世界座標系における個々の物体追跡を行う際に、各カメラ中の個々の物体のオクルージョン状態を考慮し、物体の位置更新に使用するカメラを逐次的に切り替えて物体を追跡することで、高速かつオクルージョンに頑健な追跡を実現する複数物体追跡装置、方法及びプログラムを提供することにある。

【課題を解決するための手段】

【0023】

上記の目的を達成するために、本発明は、撮影範囲がオーバーラップする複数のカメラで撮影した画像に基づいて複数の物体を追跡する複数物体追跡装置において、以下の構成を具備した点に特徴がある。

20

【0024】

(1) カメラ画像ごとに物体を検出する手段と、各物体の検出結果と各物体の予測位置とを多対一で対応付ける手段と、その検出結果を追跡結果の更新に用いるカメラを前記物体の遮蔽度合いを考慮して選択する手段と、前記選択したカメラの検出結果及び前記多対一の対応付けに基づいて追跡結果を更新する手段とを具備し、前記追跡結果の更新過程で予測した各物体の予測位置を前記対応付ける手段にフィードバックするようにした。

【0025】

(2) 検出した物体のカメラ画像上での大きさを計算する手段と、検出した物体のカメラ画像上での遮蔽度合いを計算する手段とを具備し、カメラを選択する手段は、前記物体の大きさ及び遮蔽度合いに基づいてカメラを選択するようにした。

30

【0026】

(3) 相互に遮蔽関係にある複数の物体を対象に、一の物体の画像特徴量と他の物体の画像特徴量との類似度を計算する手段を具備し、カメラを選択する手段は、カメラ画像上での物体の大きさ、遮蔽度合い及び画像特徴量の類似度に基づいてカメラを選択するようにした。

【0027】

(4) 遮蔽度合いを計算する手段は、カメラごとに一の物体の検出結果と他の物体の予測位置を当該カメラへ投影した投影領域との遮蔽度合いを当該一の物体と他の一の物体との遮蔽度合いとみなすようにした。

40

【発明の効果】

【0028】

(1) その検出結果を追跡結果の更新に用いるカメラを物体の遮蔽度合いを考慮して選択し、選択したカメラの検出結果に基づいて追跡結果を更新するようにしたので、物体の識別性が高いカメラを逐次的に選択することが可能となり、高精度な追跡を実現できるようになる。

【0029】

(2) 物体のカメラ画像上での大きさや遮蔽度合いは物体の識別性に影響するので、これ

50

らを指標にカメラを選択することで高精度な追跡を実現できるようになる。

【0030】

(3) 各物体から抽出した画像特徴量の類似度は各物体の同一性を評価する指標となるので、物体の大きさ、遮蔽度合い及び画像特徴量の類似度に基づいてカメラを選択すれば、物体の識別性や選択性の高いカメラを効率的に選択できるようになる。

【0031】

(4) 物体同士の遮蔽度合いを一方の物体の検出結果と他方の物体の追跡結果の投影領域との重なりに基づいて評価できるので、物体同士の重なり量が大きいために一方の物体しか検出できない場合でも当該物体同士の遮蔽度合いを定量的に評価できるようになる。

【図面の簡単な説明】

【0032】

【図1】本発明の一実施形態に係る複数物体追跡装置1の構成を示した機能ブロック図である。

【図2】物体追跡部の機能を説明するための図である。

【図3】カメラ選択の条件を説明するための図である。

【図4】外形矩形ベースで遮蔽度合いを計算する例を示した図である。

【図5】回転矩形ベースで遮蔽度合いを計算する例を示した図である。

【図6】輪郭ベースで遮蔽度合いを計算する例を示した図である。

【図7】一の物体の検出結果と他の一の物体の追跡結果とに基づいて遮蔽度合いを計算する例を示した図である。

【発明を実施するための形態】

【0033】

以下、図面を参照して本発明の実施の形態について詳細に説明する。図1は、本発明の一実施形態に係る複数物体追跡装置1の構成を示した機能ブロック図であり、物体検出部10、対応付け処理部20、追跡判定部30及び物体追跡部40を主要な構成としている。

【0034】

このような複数物体追跡装置1は、CPU、ROM、RAM、バス、インタフェース等を備えた少なくとも一台の汎用のコンピュータやサーバに各機能を実現するアプリケーション（プログラム）を実装することで構成できる。あるいはアプリケーションの一部をハードウェア化またはソフトウェア化した専用機や単能機としても構成できる。

【0035】

本実施形態では、各カメラは他のカメラと撮影範囲がオーバーラップしており、撮影範囲が孤立したカメラは存在しない。各カメラの内部パラメータ及び共通の世界座標系における外部パラメータは校正済であり、各カメラの画像はほぼ時刻同期している。また、連続するフレーム間で追跡対象の物体が極端に大きく移動しない程度のフレームレートで連続画像を撮像できている。

【0036】

撮影範囲のオーバーラップの程度は任意であるが、オーバーラップ領域においてオクルージョンへの頑健性を向上させるため、基本的には解析エリア内の大部分の領域において追跡対象の物体が2台以上のカメラで撮像される環境を想定している。

【0037】

物体検出部10は、複数のカメラからカメラ画像を取得して各カメラ画像に映る物体を検出し、検出結果にカメラ識別子*i*及び物体識別子*j*を付して後段の対応付け処理部20へ出力する。物体検出のアルゴリズムとしては、コンピュータビジョンの領域で一般に利用されているR-CNNやYOLO、SSD等の任意のアルゴリズムを利用できる。

【0038】

本実施形態では、物体検出の結果をD ($D=[D_1, D_2 \dots D_{Nc}]$, Nc はカメラ台数) で表現し、カメラ*i*における物体*j*の検出結果を D_i ($D_i=[D_{i,1}, D_{i,2} \dots D_{i,j} \dots D_{i,Ndi}]$, Ndi は*i*番目のカメラにおける物体検出数) で表現する。物体検出の結果Dは対応付け処理部20に出力される。本実施形態では、各物体検出の結果に各物体の種類、位置、大きさ及び当該検出の

10

20

30

40

50

信頼度が含まれる。

【0039】

本実施形態では、複数物体追跡装置1が追跡対象とする物体の種類（クラス）を特に制限しないが、例えば人物、移動ロボット、犬、猫、車等の各種の移動体が考えられる。なお、単一種類の物体（例えば人物）追跡に特化しても良く、その場合は物体の種類が一種類となるので物体の種類は検出結果に含まれない。

【0040】

カメラ画像上での物体の位置及び大きさを表す方法には様々なものがあり、本実施形態では物体の外接矩形の左上座標（ x, y ）並びにその幅 w 及び高さ h の計4パラメータで表す場合について説明するが、本発明はこれのみに限定されるものではない。

10

【0041】

対応付け処理部20は、物体検出部10からカメラ i ごとに検出結果 D_i を取得し、後述する物体追跡部40から追跡中の物体の現フレームでの予測位置 $T=[t^{\wedge}_1, t^{\wedge}_2 \dots t^{\wedge}_k \dots t^{\wedge}_{Nt}]$ （ t^{\wedge}_k は前フレームで追跡されていた物体のグローバル座標系における予測位置の情報を含む。 Nt は前フレーム時点での追跡中の物体数）を取得し、各カメラ i の検出結果 D_i と追跡中の各物体の予測位置 T とをフレームごとに物体単位で多対一に対応付ける。

【0042】

具体的には、各カメラ i において独立に予測位置 t^{\wedge}_k と検出結果 $d_{i,j}$ との間のコスト関数 $c(t^{\wedge}_k, d_{i,j})$ を評価し、次式(1)のコスト行列 C_i を計算する。

【0043】

20

【数1】

$$C_i = \begin{bmatrix} c(\hat{t}_1, d_{i,1}) & \dots & c(\hat{t}_1, d_{i,Nd_i}) \\ \vdots & \ddots & \vdots \\ c(\hat{t}_{Nt}, d_{i,1}) & \dots & c(\hat{t}_{Nt}, d_{i,Nd_i}) \end{bmatrix} \quad (1)$$

【0044】

コスト関数 C_i は、予測位置 t^{\wedge}_k と検出結果 $d_{i,j}$ とが同一物体の場合に小さい値を取るよう設計された関数であり、一般的に画像上の対応付けに用いられる、グローバル座標系における予測位置 t^{\wedge}_k と検出結果 $d_{i,j}$ との距離、カメラ画像上の予測位置 t^{\wedge}_k の投影領域の中心と検出結果 $d_{i,j}$ の中心との距離、あるいはカメラ画像上の予測位置 t^{\wedge}_k の投影領域と検出結果 $d_{i,j}$ との間のIoU（Intersection over Union）の逆数等を用いることが可能である。

30

【0045】

前記対応付け処理部20は、対応付けの方法として貪欲法やハンガリアン法等の一般的なマッチング手法を用いることができるが、ここでは貪欲法による対応付けを例にして説明する。

【0046】

まず、コスト行列 C_i の各要素から最小値 $c(t^{\wedge}_k, d_{i,j})$ を選択する。最小値が所定の閾値を上回れば対応付けは終了する。最小値が閾値を下回れば当該要素の予測位置 t^{\wedge}_k と検出結果 $d_{i,j}$ とを対応付け、予測位置 t^{\wedge}_k 及び検出結果 $d_{i,j}$ が重複して対応付けられることのないように、対応付けられた予測位置 t^{\wedge}_k 及び検出結果 $d_{i,j}$ を以降の対応付け対象から除外して、最小値が閾値を上回るまで上述の対応付けを繰り返す。以上の処理によって、カメラ i における検出結果 $d_{i,n}$ と予測位置 t^{\wedge}_m との対応付け結果 P_i [次式(2)]及び対応付けられずに残った検出結果 D_i^r [次式(3)]が得られる。

40

【0047】

【数2】

$$P_i = [p_1^i, \dots, p_{Np}^i] (p_i^i = [\hat{t}_m, d_{i,n}]) \quad (2)$$

【0048】

50

【数3】

$$D_i^r = \left[d_{i,r_1}, \dots, d_{i,r_{N_{D_r}^i}} \right] \quad (3)$$

【0049】

対応付け処理部20は、上記の各処理を全てのカメラにおいて繰り返し、対応付け結果Piを合算することで、次式(4)で与えられる全カメラ合算の対応付け結果Pを取得する。

【0050】

【数4】

$$P = \left[p_1, \dots, p_{N_p} \right] \left(p_l = \left[\widehat{t}_m, d_{i'_1, n_1}, \dots, d_{i'_{N_{p_l}}, n_{N_{p_l}}} \right] \right) \quad (4)$$

10

【0051】

対応付け結果Pでは、各カメラの検出結果と予測位置とが多対一で対応付けられることとなる。ここで、Npは対応付けられたペアの数で、その最大値は追跡中の物体数Ntとなる。Nplはl番目のペアplにおいて予測位置と対応付けられた検出結果の数を表し、その最大値はカメラの台数Ncとなる。

【0052】

最終的に、全カメラ合算の対応付け結果P、どのカメラの検出結果とも対応付けられずに残った予測位置[次式(5)：未対応付け予測位置]及び各カメラにおいて対応付けられずに残った検出結果[次式(6)：未対応付け検出結果]が得られる。

20

【0053】

【数5】

$$\widehat{T}^r = \left[\widehat{t}_1^r, \dots, \widehat{t}_{N_{T_r}}^r \right] \quad (5)$$

【0054】

【数6】

$$D^r = \left[D_1^r, \dots, D_{N_c}^r \right] \quad (6)$$

30

【0055】

対応付け処理部20は、未対応付け予測位置 \widehat{T}^r 及び未対応付け検出結果 D^r を追跡判定部30へ出力する一方、対応付け結果Pを物体追跡部40へ出力する。

【0056】

追跡判定部30は、物体追跡の新規開始判定部31及び終了判定部32を具備し、前記未対応付け予測位置 \widehat{T}^r 及び未対応付け検出結果 D^r に基づいて物体追跡の新規開始及び終了の判定処理を行う。

【0057】

新規開始判定部31は、各カメラにおいて対応付けられずに残った検出結果 D^r を複数フレームに渡って評価し、所定のフレーム数に渡って連続して同一位置で検出された検出結果を追跡対象の物体が存在する(新規に追跡を開始する)領域とみなして、物体追跡部40に当該領域のグローバル座標系における位置の情報を出力する。

40

【0058】

ここで、誤った追跡開始(False Positive)を抑制するため、非特許文献2と同様に各検出結果の位置をグローバル座標系に投影し、グローバル座標系で複数のカメラから同時に検出された検出位置のみを新規に追跡を開始する領域として物体追跡部40へ出力するようにしても良い。

【0059】

また、物体が移動している場合は、上記の連続して同一位置で検出する条件を満たすこ

50

となく、追跡対象の未検出 (False Negative) が発生することが想定される。そこで、新規開始判定部31は各カメラにおいて対応付けられずに残った未対応付け検出結果D^rを記録しておき、次のフレームで入力されるD^rとの間のコスト行列を計算し、マッチングを行っても良い。そして、連続してマッチングが行われた検出結果を、追跡対象の物体が存在する (新規に追跡を開始する) 領域として、物体追跡部40に当該領域のグローバル座標系における位置の情報を出力するようにしても良い。

【0060】

終了判定部32は、どのカメラの検出結果とも対応付けられずに残った未対応付け予測位置T^rを複数フレームに渡って評価し、所定のフレーム数、連続して対応付けられなかった予測位置T^rを、存在しない (追跡を終了する) 物体として物体追跡部40へ出力する。

10

【0061】

物体追跡部40は、カメラ選択部41、更新部42、フィードバック部43及び追加削除処理部44を含み、図2に示すように、物体同士のオクルージョンが発生しているカメラ[同図(a)]や、物体から遠いカメラ[同図(b)]を排除し、オクルージョンが発生しておらず、かつ物体から十分に近いカメラ[同図(c)]を逐次的に選択し、当該カメラの物体検出結果に基づいて各物体の追跡位置を更新する。

【0062】

カメラ選択部41は、カメラ画像上で検出された物体の大きさや遮蔽度合いに基づいて、その検出結果を追跡中の物体の位置更新に用いるカメラを選択する。

【0063】

更新部42は、対応付け処理部20から取得した全カメラ合算の対応付け結果P及び前記カメラ選択の結果に基づいて追跡中の各物体 $T=[t_1, t_2 \dots t_k \dots t_{Nt}]$ の位置を更新し、更新結果を複数物体の追跡結果として外部へ出力する。

20

【0064】

フィードバック部43は、各物体の次フレームにおける位置 $T^{\wedge}=[t^{\wedge}_1, t^{\wedge}_2 \dots t^{\wedge}_k \dots t^{\wedge}_{Nt}]$ を予測し、各予測位置 t^{\wedge}_k を前記対応付け処理部20へフィードバックする。

【0065】

追加削除処理部44は、前記追跡判定部30から取得した新規追跡開始及び追跡終了の各情報に基づいて、それぞれ追跡物体の新規追加及び削除に係る処理を行う。

【0066】

図3は、前記カメラ選択部41の動作を説明するための図であり、ここでは説明の簡略化のためにカメラ数 $Nc=3$ 、追跡中の物体数 $Nt=2$ とし、物体 $T=[t_1, t_2]$ を追跡中であり、3つのカメラ $i (i=1, 2, 3)$ の各検出結果 $d_{1,1}, d_{2,1}, d_{3,1}$ と予測位置 t^{\wedge}_1 とが多対一で対応付けられ、3つのカメラ $i (i=1, 2, 3)$ の各検出結果 $d_{1,2}, d_{2,2}, d_{3,2}$ と予測位置 t^{\wedge}_2 とが多対一で対応付けられているものとする。

30

【0067】

一般に、カメラから遠い物体はカメラ画像中に小さく映り (例えば、カメラ $i=2$)、カメラに近い物体は大きく映る (例えば、カメラ $i=3$) 傾向がある。また、複数の物体を追跡する状況下では撮影角度によってカメラ画像中の物体間の遮蔽 (オクルージョン) の度合いが異なる。

40

【0068】

そして、追跡精度は一般に、カメラ画像中に大きく映る物体は小さく映る物体よりも高くなり、他の物体によって小さく遮蔽される (あるいは全く遮蔽されない) 物体は大きく遮蔽される物体よりも高くなる傾向にある。カメラ選択部41はカメラ画像中の物体の大きさ及び遮蔽度合いに応じて、その物体検出の結果を追跡位置の更新に利用するカメラを逐次的に選択する。

【0069】

本実施形態では、追跡位置 t_1 を更新する際、初めに対応付けられた3つの検出結果 $d_{1,1}, d_{2,1}, d_{3,1}$ をカメラ画像における物体領域の大きさの降順でソートする。次いで、上位の検出結果から順に遮蔽度合い O を評価していく。

50

【 0 0 7 0 】

図4は、物体間の遮蔽度合いを外接矩形（AABB：Axis-aligned bounding box）ベースで計算する例を示した図であり、物体領域に正立姿勢で外接する外接矩形 $R=[x, y, w, h]$ を定義し、その位置、形状及び大きさで各物体の領域を代表させる。検出結果 $d_{i, j}$ の外接矩形を $R_{d_{i, j}}$ で表すと、検出結果 $d_{1, 1}$ の遮蔽度合い $O_{d_{1, 1}}$ は外接矩形同士の間重なりとして次式(7)で表すことができる。

【 0 0 7 1 】

【数7】

$$O_{d_{1,1}} = \frac{\text{Area}\left(\left(R_{d_{1,1}} \cap R_{d_{1,2}}\right) \cup \left(R_{d_{1,1}} \cap R_{d_{1,3}}\right)\right)}{\text{Area}\left(R_{d_{1,1}}\right)} \quad (7)$$

10

【 0 0 7 2 】

カメラ選択部41は、外接矩形の遮蔽度をその大きさ（本実施形態では、面積）の降順で評価し、最初に遮蔽度が閾値以下となったカメラを、その検出結果を追跡位置の更新に利用するカメラとして選択する。閾値以下となるカメラが存在しなければ遮蔽度合いが最小となるカメラを選択する。閾値以下となるカメラが複数存在する場合は複数のカメラを同時に選択しても良い。

【 0 0 7 3 】

図3の例では、予測位置 t^{\wedge}_1 と対応付けられた面積上位の検出結果 $d_{3, 1}$ は遮蔽されておらず、その遮蔽度合いが0なので、検出結果 $d_{3, 1}$ を出力するカメラ $i=3$ が選択される。検出結果 $d_{1, 1}$ 、 $d_{2, 1}$ は他の検出結果に遮蔽されており、選択されるか否かは閾値の設定次第であるが、遮蔽度が閾値以上であれば検出結果 $d_{3, 1}$ のみが選択されることになる。

20

【 0 0 7 4 】

図5は、物体間の遮蔽度合いを回転矩形（Rotated Rect）ベースで計算する例を示した図であり、原則として外形矩形ベースの計算方法で遮蔽度合いの最小値が閾値以下となるカメラを選択できない場合に、次善の計算方法として用いられる。

【 0 0 7 5 】

回転矩形ベースの計算方法では、グローバル座標系における地面から垂直方向のベクトルを画像平面上に投影し、投影されたベクトルを回転軸とした回転矩形 R' で物体領域の位置、形状及び大きさを代表する。

30

【 0 0 7 6 】

物体領域を回転矩形 R' で代表すると、外接矩形 R で代表する場合よりも物体領域によりフィットすることが期待できるため、遮蔽度合いをより正確に計算することが可能になる。

【 0 0 7 7 】

なお、外形矩形ベース及び回転矩形ベースのいずれによっても追跡位置の更新に利用するカメラを選択できなければ、物体の3Dモデルが定義されていることを前提に、当該3Dモデルのカメラ画像中への投影像のシルエット（輪郭）を計算し、輪郭（多面体）ベースで遮蔽度合いを計算するようにしても良い。

【 0 0 7 8 】

物体の3Dモデルが定義されていない場合でも、図6に示すように、Instance Segmentationの手法を用いて検出した物体領域 IS に次式(8)を適用し、外接矩形領域 R や回転矩形領域 R' との重複領域の面積によって遮蔽度を計算するようにしても良い。Instance Segmentationは物体の領域をオブジェクト単位でセグメンテーション可能な手法であり、Mask R-CNN等の手法が一般に知られている。

40

【 0 0 7 9 】

【数8】

$$O_{d_{1,1}} = \frac{\text{Area}\left(\left(IS_{d_{1,1}} \cap R_{d_{1,2}}\right) \cup \left(IS_{d_{1,1}} \cap R_{d_{1,3}}\right)\right)}{\text{Area}\left(IS_{d_{1,1}}\right)} \quad (8)$$

50

【 0 0 8 0 】

なお、外接矩形ベースの計算手法，回転矩形ベースの計算手法及び輪郭ベースの計算手法は、上記のように追跡位置の更新に利用するカメラを選択できるまで当該順序で適用することに限定されず、適宜の順序及び組み合わせで用いることができる。

【 0 0 8 1 】

具体的には、外接矩形ベースで計算した遮蔽度合いの最小値が閾値以上の場合に輪郭ベースで遮蔽度合いを計算したり、回転矩形ベースで計算した遮蔽度合いの最小値が閾値以上の場合に輪郭ベースで遮蔽度合いを計算したりしても良い。

【 0 0 8 2 】

一般に遮蔽度合いの大きさは、[外接矩形ベースで計算した遮蔽度合い] [回転矩形ベースで計算した遮蔽度合い] [輪郭ベースで計算した遮蔽度合い]の傾向があり、かつ計算コストの大きさは[外接矩形ベースでの計算コスト] [回転矩形ベースでの計算コスト] [輪郭ベースでの計算コスト]の傾向がある。したがって、外接矩形ベースで計算する遮蔽度は厳密性に欠け、実際の遮蔽度より大きい値を返す可能性がある一方、計算コストは低くなる。

10

【 0 0 8 3 】

そこで、上記の順序で3つの計算方法を併用すれば、輪郭ベースのみで遮蔽度合いを計算する場合と較べて、少ない処理負荷で効率的に遮蔽度合いを計算することが可能になる。

【 0 0 8 4 】

前記更新部42は、各物体の追跡結果としての位置（追跡位置）を、前記選択したカメラの検出結果に基づいて更新する。位置の更新にはカルマンフィルタやパーティクルフィルタ等、一般的な状態空間モデル推定手法を用いることが可能であり、本実施形態では状態空間モデル推定手法の種類を限定しない。ここでは、線形カルマンフィルタを用いた場合について説明する。

20

【 0 0 8 5 】

線形カルマンフィルタを用いる場合、時刻 f における追跡結果 t_k のグローバル座標系における位置及び移動速度を表す状態ベクトル X_f は、状態遷移行列 A 及びプロセスノイズベクトル w_f を用いて次式(9)で表され、選択された検出結果 d_s の位置（観測値）ベクトル Z_f は、観測行列 H 及び観測ノイズベクトル v_f を用いて次式(10)で表される。

【 0 0 8 6 】

30

【数 9】

$$X_f = AX_{f-1} + w_f \quad (9)$$

【 0 0 8 7 】

【数 10】

$$Z_f = HX_f + v_f \quad (10)$$

【 0 0 8 8 】

状態ベクトル X_f は位置ベクトル Z_f を用いて次式(11)で更新される。

【 0 0 8 9 】

40

【数 11】

$$X_f = \widehat{X}_f + K_f(Z_f - H\widehat{X}_f) \quad (11)$$

【 0 0 9 0 】

$X_f = AX_{f-1}$ は前フレームの状態ベクトル X_{f-1} から予測される現フレームの状態ベクトルの予測値であり、 K_f は次式(12)で表され、予測値と検出のどちらを重視するかを表すカルマンゲインである。 U_f は推定誤差の共分散行列、 R は観測ノイズの共分散行列を表す。

【 0 0 9 1 】

【数 12】

50

$$K_f = U_f H^T (H U_f H^T + R)^{-1} \quad (12)$$

【 0 0 9 2 】

つまり、更新された状態ベクトル X_f のうち、位置の成分が追跡結果 t_k のグローバル座標系における推定位置であり、物体追跡部40はこれを複数物体追跡装置1の出力結果として外部に出力する。

【 0 0 9 3 】

また、次フレームにおける X_f の予測値 $X_{f+1}^{\wedge} = A X_f$ のうち、位置の成分が追跡結果の予測位置 t_k^{\wedge} であり、前記フィードバック部43は当該予測位置 t_k^{\wedge} を対応付け処理部20へ出力する。カメラ選択部41が複数のカメラを選択していれば、複数の検出結果のグローバル座標系における重心位置が検出結果として扱われる。

10

【 0 0 9 4 】

追加削除処理部44は、前記追跡判定部30から取得した追跡の新規開始又は終了の情報に基づいて追跡結果を追加または削除する。

【 0 0 9 5 】

なお、上記の実施形態では物体の遮蔽度合いを計算する際、カメラごとに各物体の遮蔽度合いを計算していたが、本発明はこれのみに限定されるものではなく、一の物体の検出結果 d と他の物体の追跡結果 t を当該カメラに投影した投影領域との重なりに基づいて遮蔽度合いを計算するようにしても良い。

20

【 0 0 9 6 】

このようにすれば、図7に一例を示すように、カメラ $i=1$ において2つの物体の遮蔽度合いが大きいために一方の物体のみが検出($d_{1,1}$)されて他方の物体が検出($d_{1,2}$)されず、その結果、一方の物体の遮蔽割合が0と誤認されてしまう場合でも、一方の物体の検出結果 $d_{1,2}$ と他方の物体の追跡結果 t_2 との遮蔽度合いを計算し、これを物体同士の遮蔽度合いとみなすことで、一方の物体の遮蔽度合いを計算できるようになる。

【 0 0 9 7 】

また、上記の実施形態ではカメラ画像上での物体の大きさ及び遮蔽度合いに基づいて、追跡位置の更新に利用するカメラ(検出結果)を選択するものとして説明した。しかしながら、本発明はこれのみに限定されるものではなく、各物体から抽出した画像特徴量と当該物体とオーバーラップする他の物体から抽出した画像特徴量との類似度を比較し、類似度のより低いカメラが優先的に選択されるようにしても良い。類似度が低いほど各物体の識別性や選択性が高くなり、高精度な追跡が期待できるからである。

30

【 0 0 9 8 】

更に、上記のように物体の遮蔽度合いに加えて画像特徴量の類似度を参照し、2つの指標値に基づいて追跡位置の更新に利用するカメラを選択するのであれば、これらの指標値に重み付けを行い、その重み付き平均値に基づいてカメラを選択するようにしても良い。

【 0 0 9 9 】

例えば、検出結果 $d_{i,j}$ の遮蔽度合い $O_{d_{i,j}}$ と類似度 $S_{d_{i,j}}$ との重み付き平均 $OS_{d_{i,j}}$ を次式(13)で計算し、カメラ選択部41は、重み付き平均 $OS_{d_{i,j}}$ を当該検出領域の大きさの降順で評価し、最初に重み付き平均 OS が閾値以下となったカメラを、その検出結果を追跡位置の更新に利用するカメラとして選択するようにしても良い。

40

【 0 1 0 0 】

【 数 1 3 】

$$OS_{d_{i,j}} = \alpha O_{d_{i,j}} + (1 - \alpha) S_{d_{i,j}} \quad (13)$$

【 0 1 0 1 】

ここで、 $S_{d_{ij}}$ は検出領域 d_{ij} と同一カメラにおいて相互に遮蔽関係にある他の検出領域との間の類似度を表す(例えば、検出結果 d_{11} と d_{12} とが遮蔽関係にあり、その類似度が $S_{11,12}$ であれば $S_{d_{11}} = S_{11,12}$ となる)。仮に複数の検出領域と遮蔽関係にある場合は、その最

50

大値としても良い。その場合、例えば、検出結果 d_{11} と d_{12} 、および d_{11} と d_{13} がそれぞれ遮蔽関係にあり、その類似度がそれぞれ $S_{11,12}$ と $S_{11,13}$ であれば類似度 $S_{d11} = \max(S_{11,12}, S_{11,13})$ となる。

【0102】

なお、 α は遮蔽度合い $O_{di,j}$ および類似度 $S_{di,j}$ の重視の度合いを意味するパラメータであり、 $\alpha = 0.5$ の場合には両者を均等に重視することを意味する。ここで、 α の値は複数の値を仮で設定して類似する（あるいは同一の）映像に対して複数物体追跡を行うことで、追跡精度に優れる値を経験的に設定することが可能になる。

【0103】

これにより、カメラ台数が比較的少数あるいは追跡中の物体数が非常に多い環境において、いずれのカメラにおいても遮蔽度合いが大きい物体が存在するようなケースにおいて、追跡の頑健性を維持することが期待できる。この場合、前記対応付け処理部20は、コスト関数 $c(t_k, d_{i,j})$ として、距離やIoUの逆数に加えて画像特徴量の類似度の逆数を用いて対応付けを行うことが望ましい。

10

【0104】

そして、上記の実施形態によれば複数のカメラで捉えた複数の物体を正確に追跡できるようになるので、地理的あるいは経済的な格差を超えて多くの人々に多様なサービスやエンターテインメントを提供できるようになる。その結果、国連が主導する持続可能な開発目標(SDGs)の目標9「レジリエントなインフラを整備し、包括的で持続可能な産業化を推進する」や目標11「都市を包摂的、安全、レジリエントかつ持続可能にする」に貢献することが可能となる。

20

【符号の説明】

【0105】

10...物体検出部, 20...対応付け処理部, 30...追跡判定部, 31...新規開始判定部, 32...終了判定部, 40...物体追跡部, 41...カメラ選択部, 42...更新部, 43...フィードバック部, 44...追加削除処理部

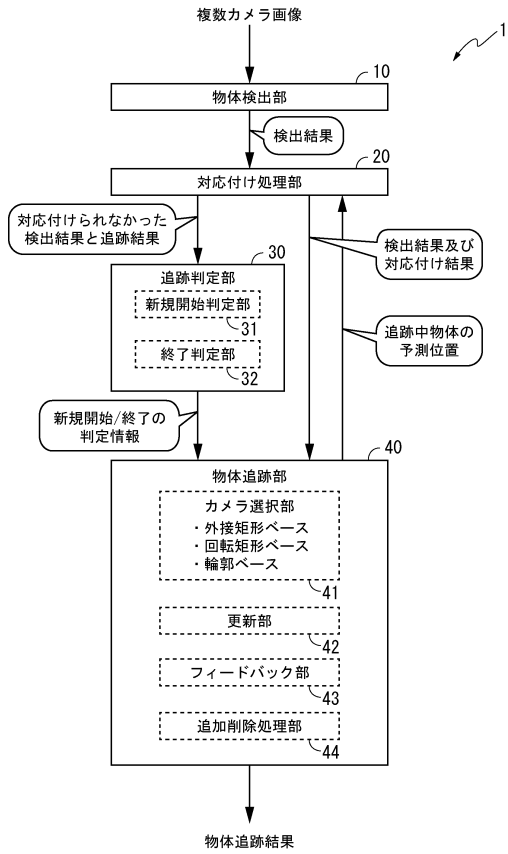
30

40

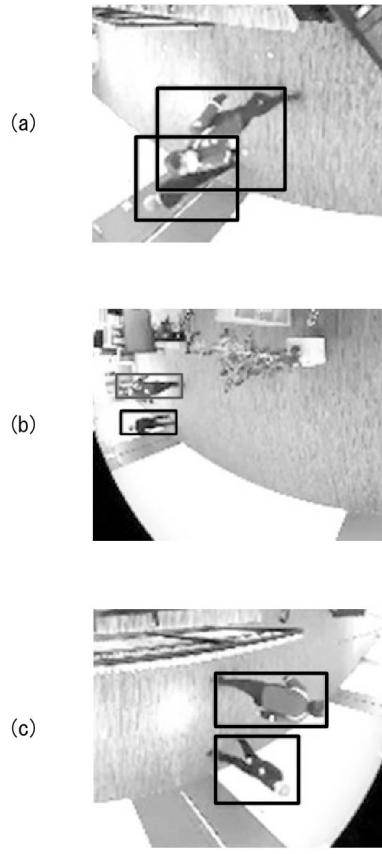
50

【 図 面 】

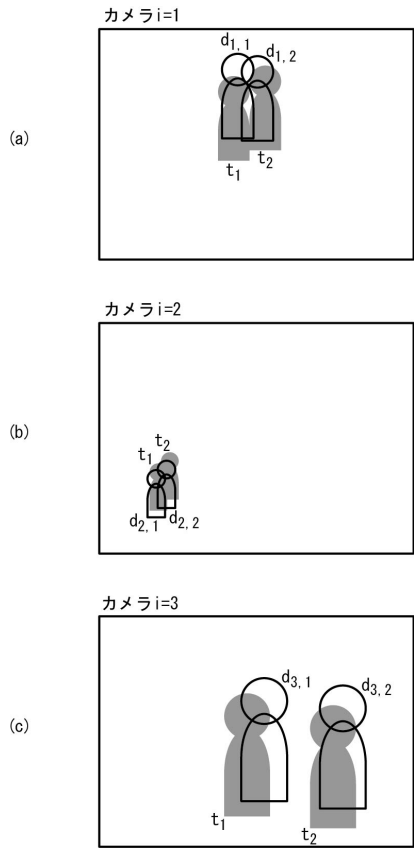
【 図 1 】



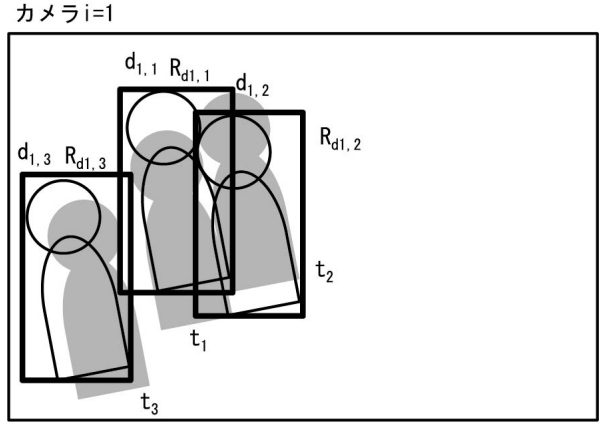
【 図 2 】



【 図 3 】



【 図 4 】



$$O_{d1,1} = \frac{Area((R_{d1,1} \cap R_{d1,2}) \cup (R_{d1,1} \cap R_{d1,3}))}{Area(R_{d1,1})}$$

10

20

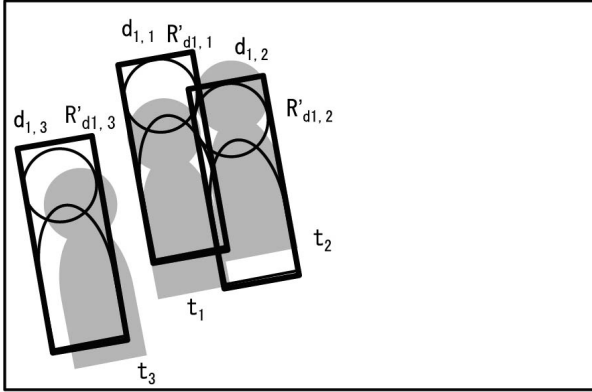
30

40

50

【 図 5 】

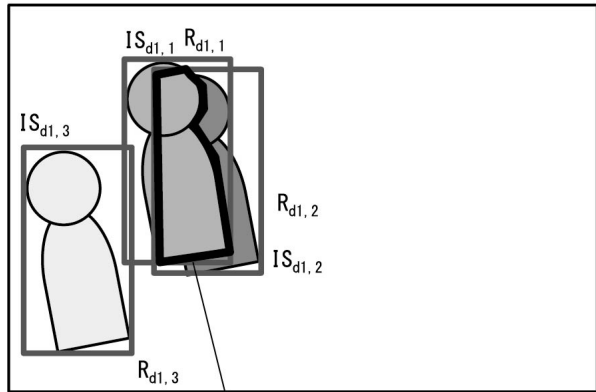
カメラ i=1



$$O_{d_{1,1}} = \frac{\text{Area}((R'_{d_{1,1}} \cap R'_{d_{1,2}}) \cup (R'_{d_{1,1}} \cap R'_{d_{1,3}}))}{\text{Area}(R'_{d_{1,1}})}$$

【 図 6 】

カメラ i=1



$$O_{d_{1,1}} = \frac{\text{Area}(\underbrace{IS_{d_{1,1}} \cap R_{d_{1,2}}}_{0} \cup \underbrace{IS_{d_{1,1}} \cap R_{d_{1,3}}}_{0})}{\text{Area}(IS_{d_{1,1}})}$$

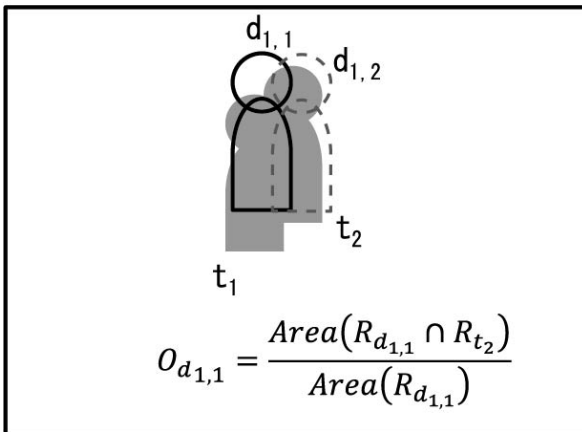
R ... Object Detectionの検出領域
IS ... Instance Segmentationの検出領域

10

20

【 図 7 】

カメラ i=1



$$O_{d_{1,1}} = \frac{\text{Area}(R_{d_{1,1}} \cap R_{t_2})}{\text{Area}(R_{d_{1,1}})}$$

30

40

50

フロントページの続き

- (56)参考文献 特開2002-008040(JP,A)
特開2019-121019(JP,A)
特開2016-126624(JP,A)
- (58)調査した分野 (Int.Cl., DB名)
- G06T 1/00 - 7/90
G06V 10/00 - 20/90
H04N 7/18
H04N 23/00 - 23/959