

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6720263号
(P6720263)

(45) 発行日 令和2年7月8日 (2020.7.8)

(24) 登録日 令和2年6月19日 (2020.6.19)

(51) Int.Cl. F I
G 0 6 T 7 / 0 0 (2017.01) G 0 6 T 7 / 0 0 3 5 0 B

請求項の数 13 外国語出願 (全 21 頁)

(21) 出願番号	特願2018-188341 (P2018-188341)	(73) 特許権者	517038176
(22) 出願日	平成30年10月3日 (2018.10.3)		株式会社ストラロビジョン
(65) 公開番号	特開2019-67402 (P2019-67402A)		大韓民国 37668 キョンサンブク
(43) 公開日	平成31年4月25日 (2019.4.25)		ド ポハン-シ ナム-グ ジゴク-ロ
審査請求日	平成30年11月1日 (2018.11.1)		394 フィフス ペンチャー-ドン ス
(31) 優先権主張番号	15/723,538		イト 304-308
(32) 優先日	平成29年10月3日 (2017.10.3)	(74) 代理人	110001737
(33) 優先権主張国・地域又は機関	米国 (US)		特許業務法人スズエ国際特許事務所
		(72) 発明者	キム、ヨンジョン
			大韓民国、キョンサンブクド、ポハン-
			シ、ナム-グ、チュオナム-ロ、67、ナ
			ンバー4-4427

最終頁に続く

(54) 【発明の名称】 回帰分析により2Dバウンディングボックスからスード (PSEUDO) - 3Dボックスを獲得する方法及びこれを利用した学習装置及びテスト装置

(57) 【特許請求の範囲】

【請求項1】

トレーニングイメージ上の2Dバウンディングボックスからスード (p s e u d o) - 3Dボックスを獲得する方法において、

(a) 学習装置が、前記2Dバウンディングボックスによって囲まれた客体を含む前記トレーニングイメージを獲得するか獲得するように支援する段階；

(b) 前記学習装置が、(i) 各々のスード (p s e u d o) - 3D方向パターンに対応する確率情報を参照して前記客体のスード (p s e u d o) - 3D方向を分類するか分類するように支援し、(i i) 回帰分析を利用して前記スード (p s e u d o) - 3Dボックス頂点の2D座標を獲得するか獲得するように支援する段階；及び

(c) 前記学習装置が、(i) 前記スード (p s e u d o) - 3Dボックス頂点の2D座標と前記スード (p s e u d o) - 3Dボックスに対応されるGT (Ground Truth) の2D座標との間の差及び (i i) 分類された前記スード (p s e u d o) - 3D方向と前記スード (p s e u d o) - 3D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援する段階；

を含み、

前記 (b) 段階で、

前記学習装置は回帰分析を利用して前記2Dバウンディングボックス頂点からの各変位を計算する過程を遂行するか遂行するように支援し、

10

20

前記 (b) 段階のうち (i i) の過程で、
前記学習装置は、

(b-1) (i) 前記スード (pseudo) - 3Dボックスの第1面に対応される第1ターゲット2Dボックスの中心である第1ベクタの終点を定めるための前記2Dバウンディングボックスの中心から前記第1ベクタと (i i) 前記客体の前記スード (pseudo) - 3D方向を参照して前記スード (pseudo) - 3Dボックスの前記第1面と対蹠される第2面に対応される第2ターゲット2Dボックスの中心である第2ベクタの終点を定めるための前記2Dバウンディングボックスの中心からの前記第2ベクタ及び (b-2) (i) 前記第1ターゲット2Dボックス幅に対する前記2Dバウンディングボックス幅の比率、 (i i) 前記第2ターゲット2Dボックス幅に対する前記2Dバウンディングボックス幅の比率、 (i i i) 前記第1ターゲット2Dボックス高さに対する前記2Dバウンディングボックス高さの比率、及び (i v) 前記第2ターゲット2Dボックス高さに対する前記2Dバウンディングボックス高さの比率と (b-3) 前記変位と前記比率を参照して前記第1ターゲット2Dボックスの頂点であるP5、P6、P7、P8の座標と前記第2ターゲット2Dボックスの頂点であるP1、P2、P3、P4の座標を計算するか計算するように支援するが、

10

前記P8は前記2Dバウンディングボックスの頂点であり、前記P1は前記P8の対角線に位置した前記2Dバウンディングボックスの頂点であり、前記P4及び前記P5は前記2Dバウンディングボックスのラインセグメントに接しないことを特徴とする方法。

【請求項2】

20

前記 (b) 段階のうち (i i) の過程で、

前記学習装置は前記2Dバウンディングボックスの頂点V1とその対角線に位置した頂点V2を前記スード (pseudo) - 3Dボックスの二つの頂点P1とP8に決定するか決定するように支援し、前記客体の前記スード (pseudo) - 3D方向を参照して前記2Dバウンディングボックスのラインセグメントに接しない二つの頂点P4とP5の前記変位を計算するが、前記変位は前記2Dバウンディングボックスの前記ラインセグメントと平行をなすことを特徴とする請求項1に記載の方法。

【請求項3】

前記 (b) 段階のうち (i) の過程で、

スード (pseudo) - 3D方向の前記パターンはデータベースに保存され、前記確率の予測に使用されることを特徴とする請求項1に記載の方法。

30

【請求項4】

前記学習装置は、

CNN (convolutional neural network) を含むが、

(i) 前記トレーニングイメージを獲得して操作して少なくとも一つのコンボリューションレイヤを利用して少なくとも一つの特徴マップを作り、 (i i) 前記特徴マップ上の前記2Dバウンディングボックスに対応される領域を定め、少なくとも一つのプーリングレイヤを利用して前記領域から特徴ベクタを抽出し、 (i i i) 第1FC (fully connected) レイヤを利用して前記入力された特徴ベクタから前記客体のスード (pseudo) - 3D方向を定め、 (i v) 第2FC (fully connected) レイヤを利用して前記スード (pseudo) - 3Dボックスの頂点に対する前記2D座標を定めて前記スード (pseudo) - 3Dボックスの頂点の2D座標を獲得するか獲得を支援することを特徴とする請求項1に記載の方法。

40

【請求項5】

テストイメージ上の客体を囲んだスード (pseudo) - 3Dボックスを獲得する方法において、

(a) テスト装置は、 (i) 学習装置がトレーニング用2Dバウンディングボックスに囲まれたトレーニング用客体を含んだトレーニングイメージを獲得するか獲得するように支援し、 (i i) (i i-1) スード (pseudo) - 3D方向の各パターンに対応する確率情報を参照して前記トレーニング用客体のスード (pseudo) - 3D方向を分

50

類し、(i i - 2) 回帰分析を利用してトレーニング用前記スード (p s e u d o) - 3 Dボックス頂点の2 D座標を獲得するか獲得するように支援し、(i i i) (i i i - 1) 前記獲得したトレーニング用前記スード (p s e u d o) - 3 Dボックス頂点の2 D座標とトレーニング用前記スード (p s e u d o) - 3 Dボックスに対応されるGTの2 D座標との間の差と(i i i - 2) 分類された前記トレーニング用スード (p s e u d o) - 3 D方向と前記トレーニング用スード (p s e u d o) - 3 D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援した状態で、テスト用2 Dバウンディングボックスに囲まれたテスト用客体を含む前記テストイメージを獲得するか獲得するように支援する段階；及び

10

(b) 前記テスト装置が、(i) スード (p s e u d o) - 3 D方向の各パターンに対応される確率情報を参照して前記テスト用客体のスード (p s e u d o) - 3 D方向を分類し、(i i) 回帰分析を利用して前記テスト用スード (p s e u d o) - 3 Dボックス頂点の2 D座標を獲得するか獲得するように支援する段階；

を含み、

前記(b) 段階のうち(i i) の過程で、

前記テスト装置が、

(b - 1) (i) 前記テスト用2 Dバウンディングボックスの中心と第1ターゲット2 Dボックスの中心である第1点との間の変位を演算するが、前記第1ターゲット2 Dボックスは前記テスト用スード (p s e u d o) - 3 Dボックスの第1面に対応して(i i) 前記テスト用2 Dバウンディングボックスの前記中心と第2ターゲット2 Dボックスの中心である第2点との間の変位を演算するが、前記第2ターゲット2 Dボックスは前記テスト用スード (p s e u d o) - 3 Dボックスの第2面に対応し、前記第2面は前記テスト用スード (p s e u d o) - 3 Dボックスの前記第1面と対蹠される位置に存在し；

20

(b - 2) (i) 前記テスト用2 Dバウンディングボックスの幅に対する前記第1ターゲット2 Dボックスの幅、(i i) 前記テスト用2 Dバウンディングボックスの前記幅に対する前記第2ターゲット2 Dボックスの幅、(i i i) 前記テスト用2 Dバウンディングボックスの高さに対する前記第1ターゲット2 Dボックスの高さ、そして(i v) 前記テスト用2 Dバウンディングボックスの前記高さに対する前記第2ターゲット2 Dボックスの高さ、各々の比率を演算し；

30

(b - 3) 前記各変位と前記各比率を参照し、前記第1ターゲット2 Dボックスの頂点をなすP 5、P 6、P 7、及びP 8の座標と、前記第2ターゲット2 Dボックスの頂点をなすP 1、P 2、P 3、及びP 4の座標を演算するが、前記P 8は前記テスト用2 Dバウンディングボックスの頂点であり、前記P 1は前記テスト用2 Dバウンディングボックスの前記P 8に対して対角線上に位置した頂点であり、前記P 4及び前記P 5は前記テスト用2 Dバウンディングボックスの前記ラインセグメントと接触しないことを特徴とする方法。

【請求項6】

前記(b) 段階のうち(i i) の過程で、

前記テスト装置が前記テスト用2 Dバウンディングボックスの頂点V 1とその対角線上の位置にある頂点V 2を前記テスト用スード (p s e u d o) - 3 Dボックスの頂点P 1とP 8に決定し、前記テスト用客体の前記スード (p s e u d o) - 3 D方向を参照して前記テスト用2 Dバウンディングボックスのラインセグメントに接しない二つの頂点P 4とP 5の変位を演算するが、前記変位は前記テスト用2 Dバウンディングボックスの前記ラインセグメントと平行であることを特徴とする請求項5に記載の方法。

40

【請求項7】

トレーニングイメージ上の2 Dバウンディングボックスからスード (p s e u d o) - 3 Dボックスを獲得する学習装置において、

前記2 Dバウンディングボックスによって囲まれた客体を含む前記トレーニングイメージを獲得するか獲得するように支援する通信部；及び

50

(i) 各々のスード (pseudo) - 3D方向パターンに対応する確率情報を参照して前記客体のスード (pseudo) - 3D方向を分類するか分類するように支援し、(ii) 回帰分析を利用して前記スード (pseudo) - 3Dボックス頂点の2D座標を獲得するか獲得するように支援するプロセス；及び (I) 前記スード (pseudo) - 3Dボックス頂点の2D座標と前記スード (pseudo) - 3Dボックスに対応されるGT (Ground Truth) の2D座標との間の差及び (II) 分類された前記スード (pseudo) - 3D方向と前記スード (pseudo) - 3D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援するプロセスを遂行するプロセッサ；

10

を含むことを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得し、

前記 (ii) のプロセスで、

前記プロセッサは回帰分析を利用して前記2Dバウンディングボックス頂点からの各変位を計算する過程を遂行するか遂行するように支援することを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得し、

前記 (ii) のプロセスで、

前記プロセッサは、

(b-1) (i) 前記スード (pseudo) - 3Dボックスの第1面に対応される第1ターゲット2Dボックスの中心である第1ベクタの終点を定めるための前記2Dバウンディングボックスの中心から前記第1ベクタと (ii) 前記客体の前記スード (pseudo) - 3D方向を参照して前記スード (pseudo) - 3Dボックスの前記第1面と対蹠される第2面に対応される第2ターゲット2Dボックスの中心である第2ベクタの終点を定めるための前記2Dバウンディングボックスの中心からの前記第2ベクタ及び (b-2) (i) 前記第1ターゲット2Dボックス幅に対する前記2Dバウンディングボックス幅の比率、(ii) 前記第2ターゲット2Dボックス幅に対する前記2Dバウンディングボックス幅の比率、(iii) 前記第1ターゲット2Dボックス高さに対する前記2Dバウンディングボックス高さの比率、及び (iv) 前記第2ターゲット2Dボックス高さに対する前記2Dバウンディングボックス高さの比率と (b-3) 前記変位と前記比率を参照して前記第1ターゲット2Dボックスの頂点であるP5、P6、P7、P8の座標と前記第2ターゲット2Dボックスの頂点であるP1、P2、P3、P4の座標を計算するか計算するように支援するが、

20

30

前記P8は前記2Dバウンディングボックスの頂点であり、前記P1は前記P8の対角線に位置した前記2Dバウンディングボックスの頂点であり、前記P4及び前記P5は前記2Dバウンディングボックスのラインセグメントに接しないことを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得することを特徴とする学習装置。

【請求項8】

前記 (ii) のプロセスで、

前記プロセッサは前記2Dバウンディングボックスの頂点V1とその対角線に位置した頂点V2を前記スード (pseudo) - 3Dボックスの二つの頂点P1とP8に決定するか決定するように支援し、前記客体の前記スード (pseudo) - 3D方向を参照して前記2Dバウンディングボックスのラインセグメントに接しない二つの頂点P4とP5の前記変位を計算するが、前記変位は前記2Dバウンディングボックスの前記ラインセグメントと平行をなすことを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得することを特徴とする請求項7に記載の学習装置。

40

【請求項9】

前記 (i) のプロセスで、

50

スード (pseudo) - 3D方向の前記パターンはデータベースに保存され、前記確率の予測に使用されることを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得することを特徴とする請求項7に記載の学習装置。

【請求項10】

前記学習装置は、

CNN (convolutional neural network) を含むが、

(i) 前記トレーニングイメージを獲得して操作して少なくとも一つのコンボリューションレイヤを利用して少なくとも一つの特徴マップを作り、(ii) 前記特徴マップ上の前記2Dバウンディングボックスに対応される領域を定め、少なくとも一つのプーリングレイヤを利用して前記領域から特徴ベクタを抽出し、(iii) 第1FC (fully connected) レイヤを利用して前記入力された特徴ベクタから前記客体のスード (pseudo) - 3D方向を定め、(iv) 第2FC (fully connected) レイヤを利用して前記スード (pseudo) - 3Dボックスの頂点に対する前記2D座標を定めて前記スード (pseudo) - 3Dボックスの頂点の2D座標を獲得するか獲得するように支援することを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得することを特徴とする請求項7に記載の学習装置。

【請求項11】

前記客体のスード (pseudo) - 3D方向を分類するプロセスは、前記スード (pseudo) - 3Dボックス頂点の2D座標を獲得するプロセス以前に遂行されることを特徴とする請求項7に記載の学習装置。

【請求項12】

テストイメージ上の客体を囲んだスード (pseudo) - 3Dボックスを獲得するテスト装置において、

学習装置がトレーニング用2Dバウンディングボックスに囲まれたトレーニング用客体を含んだトレーニングイメージを獲得するか獲得するように支援し、(i) (i-1) スード (pseudo) - 3D方向の各パターンに対応する確率情報を参照して前記トレーニング用客体のスード (pseudo) - 3D方向を分類し、(i-2) 回帰分析を利用してトレーニング用前記スード (pseudo) - 3Dボックス頂点の2D座標を獲得するか獲得するように支援し、(ii) (ii-1) 前記獲得したトレーニング用前記スード (pseudo) - 3Dボックス頂点の2D座標とトレーニング用前記スード (pseudo) - 3Dボックスに対応されるGTの2D座標との間の差と(ii-2) 分類された前記トレーニング用スード (pseudo) - 3D方向と前記トレーニング用スード (pseudo) - 3D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援した状態で、テスト用2Dバウンディングボックスに囲まれたテスト用客体を含む前記テストイメージを獲得するか獲得するように支援する通信部；及び

(I) スード (pseudo) - 3D方向の各パターンに対応される確率情報を参照して前記テスト用客体のスード (pseudo) - 3D方向を分類し、(II) 回帰分析を利用して前記テスト用スード (pseudo) - 3Dボックス頂点の2D座標を獲得するか獲得するように支援するプロセスを遂行するプロセッサ；

を含むことを特徴とするテストイメージ上の客体を囲んだスード (pseudo) - 3Dボックスを獲得し、

前記(II)のプロセスで、

前記プロセッサが、

(b-1) (i) 前記テスト用2Dバウンディングボックスの中心と第1ターゲット2Dボックスの中心である第1点との間の変位を演算するが、前記第1ターゲット2Dボックスは前記テスト用スード (pseudo) - 3Dボックスの第1面に対応して(ii) 前記テスト用2Dバウンディングボックスの前記中心と第2ターゲット2Dボックスの中

心である第2点との間の変位を演算するが、前記第2ターゲット2Dボックスは前記テスト用スード（pseudo）-3Dボックスの第2面に対応し、前記第2面は前記テスト用スード（pseudo）-3Dボックスの前記第1面と対蹠される位置に存在し；

（b-2）（i）前記テスト用2Dバウンディングボックスの幅に対する前記第1ターゲット2Dボックスの幅、（ii）前記テスト用2Dバウンディングボックスの前記幅に対する前記第2ターゲット2Dボックスの幅、（iii）前記テスト用2Dバウンディングボックスの高さに対する前記第1ターゲット2Dボックスの高さ、そして（iv）前記テスト用2Dバウンディングボックスの前記高さに対する前記第2ターゲット2Dボックスの高さ、各々の比率を演算し；

（b-3）前記各変位と前記各比率を参照し、前記第1ターゲット2Dボックスの頂点をなすP5、P6、P7、及びP8の座標と、前記第2ターゲット2Dボックスの頂点をなすP1、P2、P3、及びP4の座標を演算するが、前記P8は前記テスト用2Dバウンディングボックスの頂点であり、前記P1は前記テスト用2Dバウンディングボックスの前記P8に対して対角線上に位置した頂点であり、前記P4及び前記P5は前記テスト用2Dバウンディングボックスの前記ラインセグメントと接触しないことを特徴とするテストイメージ上の客体を囲んだスード（pseudo）-3Dボックスを獲得することを特徴とするテスト装置。

【請求項13】

前記（II）のプロセスで、

前記プロセッサが前記テスト用2Dバウンディングボックスの頂点V1とその対角線上の位置にある頂点V2を前記テスト用スード（pseudo）-3Dボックスの頂点P1とP8に決定し、前記テスト用客体の前記スード（pseudo）-3D方向を参照して前記テスト用2Dバウンディングボックスのラインセグメントに接しない二つの頂点P4とP5の変位を演算するが、前記変位は前記テスト用2Dバウンディングボックスの前記ラインセグメントと平行であることを特徴とするテストイメージ上の客体を囲んだスード（pseudo）-3Dボックスを獲得することを特徴とする請求項12に記載のテスト装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はトレーニングイメージ上の2Dバウンディングボックスからスード（pseudo）-3Dボックスを獲得する方法及び装置に関し、より詳細には、（a）前記2Dバウンディングボックスによって囲まれた客体を含む前記トレーニングイメージを獲得する段階；（b）（i）各々のスード（pseudo）-3D方向パターンに対応する確率情報を参照して前記客体のスード（pseudo）-3D方向を分類し、（ii）回帰分析を利用して前記スード（pseudo）-3Dボックス頂点の2D座標を獲得する段階；及び（c）（i）前記スード（pseudo）-3Dボックス頂点の2D座標と前記スード（pseudo）-3Dボックスに対応されるGT（Ground Truth）の2D座標との間の差及び（ii）分類された前記スード（pseudo）-3D方向と前記スード（pseudo）-3D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整する段階；を含むことを特徴とする学習方法、テスト方法及びこれを利用する学習装置、テスト装置に関する。

【背景技術】

【0002】

自律走行車は周辺環境を感知して人の介入なしに走行することができる自動車である。自律走行車はレーダやレーザ、GPS、走行記録計及びコンピュータビジョンなど周辺状況を探知する多様な技術を利用する。

【0003】

自律走行と個人用ロボットなどの分野ではコンピュータビジョンを利用して2次元（2

10

20

30

40

50

D) 単眼イメージを通じて3次元(3D)情報を推定することは重要である。一般的に、イメージ上の客体を囲んだ2Dボックスが形成され、以後前記2Dボックスから3Dモデルが作られる。

【0004】

既存技術では前記客体を囲んだ2Dボックスを探すために一般的にテンプレート基盤方式を使用した。前記2Dボックスを形成する既存の方法の中にはスライディングウィンドウ方式がある。この方法は全体イメージ上でウィンドウのようなボックスを大きさを変化させて反復的に移動させ、前記ウィンドウのようなボックス内部にある各客体を感知する。即ち、前記イメージ内の前記客体は大きさや尺度が異なる場合があるので、前記イメージは数回スケールダウンされ、前記ウィンドウのようなボックスは再び前記イメージ上を移動しながら大きさが各々異なる前記客体を探す。

10

【0005】

既存方法の中でまた他の一つはアンカーボックス技術である。この方法によれば、多様なアンカーボックスが一定の位置中央にあり、多様なアンカーボックスのうち確率が最も高いアンカーボックス、即ち、GT (ground truth) 客体と重畳が最も多い領域を有するアンカーボックスが回帰分析を利用して決定される。

【0006】

以後、3Dバウンディングボックスは前記決定されたアンカーボックスから生成されるが制限が多い。第一に、前記3Dバウンディングボックスは6面であり、6面のうち3面は全数調査が要求され得る。第二に、前記3Dバウンディングボックスの3面を決定するのに単一テンプレートが使用される場合、前記客体の3D方向が変われば回帰分析の境界条件が異なり得るため正確性が低くなり得る。第三に、前記3Dバウンディングボックスを得るための既存の方法は多くの演算資源が要求される。例えば、3Dバウンディングボックスを探すために直方体テンプレートやボクセルをマッチングさせるのに多くの演算時間がかかる。

20

【0007】

ここに、本発明ではかかる重複的演算をなくし、感知正確性を改善するための新しい方法を提案する。

【発明の概要】

【発明が解決しようとする課題】

30

【0008】

本発明は前記のようなすべての問題点を解決することを目的とする。

【0009】

また、本発明はスード(pseudo)-3Dボックスを利用して3D客体と外接する3Dバウンディングボックスを単純化することを他の目的とする。

【0010】

本発明はスード(pseudo)-3Dボックス頂点の2D座標を利用して演算時間と資源を減らすことをまた他の目的とする。

【課題を解決するための手段】

【0011】

40

前記のような本発明の目的を達成し、後述する本発明の特徴的な効果を実現するための、本発明の特徴的な構成は下記のとおりである。

【0012】

本発明の一態様によれば、トレーニングイメージ上の2Dバウンディングボックスからスード(pseudo)-3Dボックスを獲得する方法において、(a)学習装置が、前記2Dバウンディングボックスによって囲まれた客体を含む前記トレーニングイメージを獲得するか獲得するように支援する段階；(b)前記学習装置が、(i)各々のスード(pseudo)-3D方向パターンに対応する確率情報を参照して前記客体のスード(pseudo)-3D方向を分類するか分類するように支援し、(ii)回帰分析を利用して前記スード(pseudo)-3Dボックス頂点の2D座標を獲得するか獲得するよう

50

に支援する段階；及び（c）前記学習装置が、（i）前記スード（pseudo）－3Dボックス頂点の2D座標と前記スード（pseudo）－3Dボックスに対応されるGT（Ground Truth）の2D座標との間の差及び（ii）分類された前記スード（pseudo）－3D方向と前記スード（pseudo）－3D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援する段階；を含むことを特徴とする方法が提供される。

【0013】

一例として、前記（b）段階で、前記学習装置は回帰分析を利用して前記2Dバウンディングボックス頂点からの各変位を計算する過程を遂行するか遂行するように支援することを特徴とする。

10

【0014】

一例として、前記（b）段階のうち（ii）の過程で、前記学習装置は前記2Dバウンディングボックスの頂点V1とその対角線に位置した頂点V2を前記スード（pseudo）－3Dボックスの二つの頂点P1とP8に決定するか決定するように支援し、前記客体の前記スード（pseudo）－3D方向を参照して前記2Dバウンディングボックスのラインセグメントに接しない二つの頂点P4とP5の前記変位を計算するが、前記変位は前記2Dバウンディングボックスの前記ラインセグメントと平行をなすことを特徴とする。

【0015】

20

一例として、前記（b）段階のうち（ii）の過程で、前記学習装置は（b-1）（i）前記スード（pseudo）－3Dボックスの第1面に対応される第1ターゲット2Dボックスの中心である第1ベクタの終点を定めるための前記2Dバウンディングボックスの中心から前記第1ベクタと（ii）前記客体の前記スード（pseudo）－3D方向を参照して前記スード（pseudo）－3Dボックスの前記第1面と対蹠される第2面に対応される第2ターゲット2Dボックスの中心である第2ベクタの終点を定めるための前記2Dバウンディングボックスの中心からの前記第2ベクタ及び（b-2）（i）前記第1ターゲット2Dボックス幅に対する前記2Dバウンディングボックス幅の比率、（ii）前記第2ターゲット2Dボックス幅に対する前記2Dバウンディングボックス幅の比率、（iii）前記第1ターゲット2Dボックス高さに対する前記2Dバウンディングボックス高さの比率、及び（iv）前記第2ターゲット2Dボックス高さに対する前記2Dバウンディングボックス高さの比率と（b-3）前記変位と前記比率を参照して前記第1ターゲット2Dボックスの頂点であるP5、P6、P7、P8の座標と前記第2ターゲット2Dボックスの頂点であるP1、P2、P3、P4の座標を計算するか計算するように支援するが、前記P8は前記2Dバウンディングボックスの頂点であり、前記P1は前記P8の対角線に位置した前記2Dバウンディングボックスの頂点であり、前記P4及び前記P5は前記2Dバウンディングボックスのラインセグメントに接しないことを特徴とする。

30

【0016】

一例として、前記（b）段階のうち（i）の過程で、スード（pseudo）－3D方向の前記パターンはデータベースに保存され、前記確率の予測に使用されることを特徴とする。

40

【0017】

一例として、前記学習装置は、CNN（convolutional neural network）を含むが、（i）前記トレーニングイメージを獲得して操作して少なくとも一つのコンボリューションレイヤを利用して少なくとも一つの特徴マップを作り、（ii）前記特徴マップ上の前記2Dバウンディングボックスに対応される領域を定め、少なくとも一つのプーリングレイヤを利用して前記領域から特徴ベクタを抽出し、（iii）第1FC（fully connected）レイヤを利用して前記入力された特徴ベクタから前記客体のスード（pseudo）－3D方向を定め、（iv）第2FC（fu

50

l l y c o n n e c t e d) レイヤを利用して前記スード (p s e u d o) - 3 D ボックスの頂点に対する前記 2 D 座標を定めて前記スード (p s e u d o) - 3 D ボックスの頂点の 2 D 座標を獲得するか獲得を支援することを特徴とする。

【0018】

本発明の他の態様によれば、テストイメージ上の客体を囲んだスード (p s e u d o) - 3 D ボックスを獲得する方法において、(a) テスト装置は、(i) 学習装置がトレーニング用 2 D バウンディングボックスに囲まれたトレーニング用客体を含んだトレーニングイメージを獲得するか獲得するように支援し、(i i) (i i - 1) スード (p s e u d o) - 3 D 方向の各パターンに対応する確率情報を参照して前記トレーニング用客体のスード (p s e u d o) - 3 D 方向を分類し、(i i - 2) 回帰分析を利用してトレーニング用前記スード (p s e u d o) - 3 D ボックス頂点の 2 D 座標を獲得するか獲得するように支援し、(i i i) (i i i - 1) 前記獲得したトレーニング用前記スード (p s e u d o) - 3 D ボックス頂点の 2 D 座標とトレーニング用前記スード (p s e u d o) - 3 D ボックスに対応される G T の 2 D 座標との間の差と (i i i - 2) 分類された前記トレーニング用スード (p s e u d o) - 3 D 方向と前記トレーニング用スード (p s e u d o) - 3 D 方向に対応される G T との間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援した状態で、テスト用 2 D バウンディングボックスに囲まれたテスト用客体を含む前記テストイメージを獲得するか獲得するように支援する段階；及び (b) 前記テスト装置が、(i) スード (p s e u d o) - 3 D 方向の各パターンに対応される確率情報を参照して前記テスト用客体のスード (p s e u d o) - 3 D 方向を分類し、(i i) 回帰分析を利用して前記テスト用スード (p s e u d o) - 3 D ボックス頂点の 2 D 座標を獲得するか獲得するように支援する段階；を含むことを特徴とする方法が提供される。

【0019】

一例として、前記 (b) 段階のうち (i i) の過程で、前記テスト装置が前記テスト用 2 D バウンディングボックスの頂点 V 1 とその対角線上の位置にある頂点 V 2 を前記テスト用スード (p s e u d o) - 3 D ボックスの頂点 P 1 と P 8 に決定し、前記テスト用客体の前記スード (p s e u d o) - 3 D 方向を参照して前記テスト用 2 D バウンディングボックスのラインセグメントに接しない二つの頂点 P 4 と P 5 の変位を演算するが、前記変位は前記テスト用 2 D バウンディングボックスの前記ラインセグメントと平行であることを特徴とする。

【0020】

一例として、前記 (b) 段階のうち (i i) の過程で、前記テスト装置が (b - 1) (i) 前記テスト用 2 D バウンディングボックスの中心と第 1 ターゲット 2 D ボックスの中心である第 1 点との間の変位を演算するが、前記第 1 ターゲット 2 D ボックスは前記テスト用スード (p s e u d o) - 3 D ボックスの第 1 面に対応して (i i) 前記テスト用 2 D バウンディングボックスの前記中心と第 2 ターゲット 2 D ボックスの中心である第 2 点との間の変位を演算するが、前記第 2 ターゲット 2 D ボックスは前記テスト用スード (p s e u d o) - 3 D ボックスの第 2 面に対応し、前記第 2 面は前記テスト用スード (p s e u d o) - 3 D ボックスの前記第 1 面と対蹠される位置に存在し；(b - 2) (i) 前記テスト用 2 D バウンディングボックスの幅に対する前記第 1 ターゲット 2 D ボックスの幅、(i i) 前記テスト用 2 D バウンディングボックスの前記幅に対する前記第 2 ターゲット 2 D ボックスの幅、(i i i) 前記テスト用 2 D バウンディングボックスの高さに対する前記第 1 ターゲット 2 D ボックスの高さ、そして (i v) 前記テスト用 2 D バウンディングボックスの前記高さに対する前記第 2 ターゲット 2 D ボックスの高さ、各々の比率を演算し；(b - 3) 前記各変位と前記各比率を参照し、前記第 1 ターゲット 2 D ボックスの頂点をなす P 5、P 6、P 7、及び P 8 の座標と、前記第 2 ターゲット 2 D ボックスの頂点をなす P 1、P 2、P 3、及び P 4 の座標を演算するが、前記 P 8 は前記テスト用 2 D バウンディングボックスの頂点であり、前記 P 1 は前記テスト用 2 D バウンディングボックスの前記 P 8 に対して対角線上に位置した頂点であり、前記 P 4 及び前記 P 5 は前

記テスト用 2 D バウンディングボックスの前記ラインセグメントと接触しないことを特徴とする。

【0021】

本発明のまた他の態様によれば、トレーニングイメージ上の 2 D バウンディングボックスからスード (pseudo) - 3 D ボックスを獲得する学習装置において、前記 2 D バウンディングボックスによって囲まれた客体を含む前記トレーニングイメージを獲得するか獲得するように支援する通信部；及び (i) 各々のスード (pseudo) - 3 D 方向パターンに対応する確率情報を参照して前記客体のスード (pseudo) - 3 D 方向を分類するか分類するように支援し、(ii) 回帰分析を利用して前記スード (pseudo) - 3 D ボックス頂点の 2 D 座標を獲得するか獲得するように支援するプロセス；及び (I) 前記スード (pseudo) - 3 D ボックス頂点の 2 D 座標と前記スード (pseudo) - 3 D ボックスに対応される GT (Ground Truth) の 2 D 座標との間の差及び (II) 分類された前記スード (pseudo) - 3 D 方向と前記スード (pseudo) - 3 D 方向に対応される GT との間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援するプロセスを遂行するプロセッサ；を含むことを特徴とするトレーニングイメージ上の 2 D バウンディングボックスからスード (pseudo) - 3 D ボックスを獲得することを特徴とする学習装置が提供される。

10

【0022】

一例として、前記 (ii) のプロセスで、前記プロセッサは回帰分析を利用して前記 2 D バウンディングボックス頂点からの各変位を計算する過程を遂行するか遂行するように支援することを特徴とするトレーニングイメージ上の 2 D バウンディングボックスからスード (pseudo) - 3 D ボックスを獲得することを特徴とする。

20

【0023】

一例として、前記 (ii) のプロセスで、前記プロセッサは前記 2 D バウンディングボックスの頂点 V1 とその対角線に位置した頂点 V2 を前記スード (pseudo) - 3 D ボックスの二つの頂点 P1 と P8 に決定するか決定するように支援し、前記客体の前記スード (pseudo) - 3 D 方向を参照して前記 2 D バウンディングボックスのラインセグメントに接しない二つの頂点 P4 と P5 の前記変位を計算するが、前記変位は前記 2 D バウンディングボックスの前記ラインセグメントと平行をなすことを特徴とするトレーニングイメージ上の 2 D バウンディングボックスからスード (pseudo) - 3 D ボックスを獲得することを特徴とする。

30

【0024】

一例として、前記 (ii) のプロセスで、前記プロセッサは (b-1) (i) 前記スード (pseudo) - 3 D ボックスの第 1 面に対応される第 1 ターゲット 2 D ボックスの中心である第 1 ベクタの終点を定めるための前記 2 D バウンディングボックスの中心から前記第 1 ベクタと (ii) 前記客体の前記スード (pseudo) - 3 D 方向を参照して前記スード (pseudo) - 3 D ボックスの前記第 1 面と対蹠される第 2 面に対応される第 2 ターゲット 2 D ボックスの中心である第 2 ベクタの終点を定めるための前記 2 D バウンディングボックスの中心からの前記第 2 ベクタ及び (b-2) (i) 前記第 1 ターゲット 2 D ボックス幅に対する前記 2 D バウンディングボックス幅の比率、(ii) 前記第 2 ターゲット 2 D ボックス幅に対する前記 2 D バウンディングボックス幅の比率、(iii) 前記第 1 ターゲット 2 D ボックス高さに対する前記 2 D バウンディングボックス高さの比率、及び (iv) 前記第 2 ターゲット 2 D ボックス高さに対する前記 2 D バウンディングボックス高さの比率と (b-3) 前記変位と前記比率を参照して前記第 1 ターゲット 2 D ボックスの頂点である P5、P6、P7、P8 の座標と前記第 2 ターゲット 2 D ボックスの頂点である P1、P2、P3、P4 の座標を計算するか計算するように支援するが、前記 P8 は前記 2 D バウンディングボックスの頂点であり、前記 P1 は前記 P8 の対角線に位置した前記 2 D バウンディングボックスの頂点であり、前記 P4 及び前記 P5 は前記 2 D バウンディングボックスのラインセグメントに接しないことを特徴とするトレーニング

40

50

ングイメージ上の2Dバウンディングボックスからスード(pseudo)-3Dボックスを獲得することを特徴とする。

【0025】

一例として、前記(i)のプロセスで、スード(pseudo)-3D方向の前記パターンはデータベースに保存され、前記確率の予測に使用されることを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード(pseudo)-3Dボックスを獲得することを特徴とする。

【0026】

一例として、前記学習装置は、CNN(convolutional neural network)を含むが、(i)前記トレーニングイメージを獲得して操作して少なくとも一つのコンボリューションレイヤを利用して少なくとも一つの特徴マップを作り、(ii)前記特徴マップ上の前記2Dバウンディングボックスに対応される領域を定め、少なくとも一つのプーリングレイヤを利用して前記領域から特徴ベクタを抽出し、(iii)第1FC(fully connected)レイヤを利用して前記入力された特徴ベクタから前記客体のスード(pseudo)-3D方向を定め、(iv)第2FC(fully connected)レイヤを利用して前記スード(pseudo)-3Dボックスの頂点に対する前記2D座標を定めて前記スード(pseudo)-3Dボックスの頂点の2D座標を獲得するか獲得するように支援することを特徴とするトレーニングイメージ上の2Dバウンディングボックスからスード(pseudo)-3Dボックスを獲得することを特徴とする。

【0027】

一例として、前記客体のスード(pseudo)-3D方向を分類するプロセスは、前記スード(pseudo)-3Dボックス頂点の2D座標を獲得するプロセス以前に遂行されることを特徴とする。

【0028】

本発明のまた他の態様によれば、テストイメージ上の客体を囲んだスード(pseudo)-3Dボックスを獲得するテスト装置において、学習装置がトレーニング用2Dバウンディングボックスに囲まれたトレーニング用客体を含んだトレーニングイメージを獲得するか獲得するように支援し、(i)(i-1)スード(pseudo)-3D方向の各パターンに対応する確率情報を参照して前記トレーニング用客体のスード(pseudo)-3D方向を分類し、(i-2)回帰分析を利用してトレーニング用前記スード(pseudo)-3Dボックス頂点の2D座標を獲得するか獲得するように支援し、(ii)(ii-1)前記獲得したトレーニング用前記スード(pseudo)-3Dボックス頂点の2D座標とトレーニング用前記スード(pseudo)-3Dボックスに対応されるGTの2D座標との間の差と(ii-2)分類された前記トレーニング用スード(pseudo)-3D方向と前記トレーニング用スード(pseudo)-3D方向に対応されるGTとの間の差のうち少なくとも一つを参照して決定されるロス情報をバックプロパゲーションして自身のパラメータを調整するか調整するように支援した状態で、テスト用2Dバウンディングボックスに囲まれたテスト用客体を含む前記テストイメージを獲得するか獲得するように支援する通信部；及び(I)スード(pseudo)-3D方向の各パターンに対応される確率情報を参照して前記テスト用客体のスード(pseudo)-3D方向を分類し、(II)回帰分析を利用して前記テスト用スード(pseudo)-3Dボックス頂点の2D座標を獲得するか獲得するように支援するプロセスを遂行するプロセッサ；を含むことを特徴とするテストイメージ上の客体を囲んだスード(pseudo)-3Dボックスを獲得することを特徴とするテスト装置が提供される。

【0029】

一例として、前記(II)のプロセスで、前記プロセッサが前記テスト用2Dバウンディングボックスの頂点V1とその対角線上の位置にある頂点V2を前記テスト用スード(pseudo)-3Dボックスの頂点P1とP8に決定し、前記テスト用客体の前記スード(pseudo)-3D方向を参照して前記テスト用2Dバウンディングボックスのラ

インセグメントに接しない二つの頂点 P 4 と P 5 の変位を演算するが、前記変位は前記テスト用 2 D バウンディングボックスの前記ラインセグメントと平行であることを特徴とするテストイメージ上の客体を囲んだスード (p s e u d o) - 3 D ボックスを獲得することを特徴とする。

【0030】

一例として、前記 (I I) のプロセスで、前記プロセッサが (b - 1) (i) 前記テスト用 2 D バウンディングボックスの中心と第 1 ターゲット 2 D ボックスの中心である第 1 点との間の変位を演算するが、前記第 1 ターゲット 2 D ボックスは前記テスト用スード (p s e u d o) - 3 D ボックスの第 1 面に対応して (i i) 前記テスト用 2 D バウンディングボックスの前記中心と第 2 ターゲット 2 D ボックスの中心である第 2 点との間の変位を演算するが、前記第 2 ターゲット 2 D ボックスは前記テスト用スード (p s e u d o) - 3 D ボックスの第 2 面に対応し、前記第 2 面は前記テスト用スード (p s e u d o) - 3 D ボックスの前記第 1 面と対蹠される位置に存在し； (b - 2) (i) 前記テスト用 2 D バウンディングボックスの幅に対する前記第 1 ターゲット 2 D ボックスの幅、 (i i) 前記テスト用 2 D バウンディングボックスの前記幅に対する前記第 2 ターゲット 2 D ボックスの幅、 (i i i) 前記テスト用 2 D バウンディングボックスの高さに対する前記第 1 ターゲット 2 D ボックスの高さ、そして (i v) 前記テスト用 2 D バウンディングボックスの前記高さに対する前記第 2 ターゲット 2 D ボックスの高さ、各々の比率を演算し； (b - 3) 前記各変位と前記各比率を参照し、前記第 1 ターゲット 2 D ボックスの頂点をなす P 5、P 6、P 7、及び P 8 の座標と、前記第 2 ターゲット 2 D ボックスの頂点をなす P 1、P 2、P 3、及び P 4 の座標を演算するが、前記 P 8 は前記テスト用 2 D バウンディングボックスの頂点であり、前記 P 1 は前記テスト用 2 D バウンディングボックスの前記 P 8 に対して対角線上に位置した頂点であり、前記 P 4 及び前記 P 5 は前記テスト用 2 D バウンディングボックスの前記ラインセグメントと接触しないことを特徴とするテストイメージ上の客体を囲んだスード (p s e u d o) - 3 D ボックスを獲得することを特徴とする。

【発明の効果】

【0031】

本発明によれば、次のような効果がある。

【0032】

本発明は 3 次元客体を表現する 3 次元モデルを簡便化することができるようになる。

【0033】

本発明はバウンディングボックスの頂点の座標決定に必要な演算を最小化することで演算時間と資源の消費を減少させることができるようになる。

【0034】

本発明は人の介入を最小化するための、3 次元ボックスの獲得を学習することができる方法を提供することができるようになる。

【図面の簡単な説明】

【0035】

本発明の実施例の説明に利用されるために添付された下記図面は本発明の実施例のうち単に一部に過ぎず、本発明が属する技術分野の通常の技術者においては発明的作業がなされることなく本図面に基づいて他の図面が得られることがある。

【0036】

【図 1】図 1 は、本発明の一実施例によってトレーニングイメージ内の 2 D バウンディングボックスからスード (p s e u d o) - 3 D ボックスを獲得するための学習装置の構成を示す図面である。

【図 2】図 2 は、本発明の一実施例によってトレーニングイメージからスード (p s e u d o) - 3 D ボックスを獲得することができる CNN (c o n v o l u t i o n a l n e u r a l n e t w o r k) を含んだ学習装置の構成を概略的に示した図面である。

【図 3】図 3 は、本発明の一実施例によってトレーニングイメージ内の 2 D バウンディン

グボックスからスード (pseudo) - 3Dボックスを獲得する学習装置の流れを簡略に示した図面である。

【図4】図4は、本発明の一実施例によって2Dバウンディングボックスからスード (pseudo) - 3Dボックス頂点座標を獲得する方法を簡略に示す図面である。

【図5】図5は、本発明の一実施例によって2Dバウンディングボックスからスード (pseudo) - 3Dボックス頂点座標を獲得するまた他の方法を簡略に示す図面である。

【図6】図6は、本発明のまた他の一実施例によってテストイメージ内の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得するためのテスト装置の構成を示す図面である。

【図7】図7は、本発明の一実施例によって客体のスード (pseudo) - 3D方向のパターンに対応される確率予測に使用される客体のスード (pseudo) - 3D方向のパターンを概略的に示した図面である。

【発明を実施するための形態】

【0037】

後述する本発明に対する詳細な説明は、本発明が実施され得る特定の実施例を例示として示す添付図面を参照する。これらの実施例は当業者が本発明を実施することができるように充分詳細に説明される。本発明の多様な実施例は相互異なるが、相互排他的である必要はないことを理解されたい。例えば、ここに記載されている特定の形状、構造及び特性は一実施例にかかる本発明の精神及び範囲を逸脱せずに他の実施例で具現され得る。また、各々の開示された実施例内の個別構成要素の位置または配置は本発明の精神及び範囲を逸脱せずに変更され得ることを理解されたい。従って、後述する詳細な説明は限定的な意味で捉えようとするものではなく、本発明の範囲は、適切に説明されると、その請求項が主張することと均等なすべての範囲と、併せて添付された請求項によってのみ限定される。図面で類似する参照符号はいくつかの側面にかけて同一か類似する機能を指称する。

【0038】

図1は本発明によって少なくとも一つの客体を含むトレーニングイメージ内の2Dバウンディングボックスからスード (pseudo) - 3Dボックスを獲得するための学習装置100の構成を示した図面である。ここに、前記スード (pseudo) - 3Dボックスと前記2Dバウンディングボックスは前記客体に正確に外接され得る。また、前記学習装置100は機能を遂行するいくつかの装置に分けられ得るが、本明細書では便宜上一つの装置として具現された前記学習装置100を記述する。本発明による前記学習装置100はコンボリューションレイヤ、RPN (region proposal network) 及びFC (fully connected) レイヤなどの一部が含まれ得る。

【0039】

図1を参照すれば、前記トレーニングイメージ内の前記2Dバウンディングボックスから前記スード (pseudo) - 3Dボックスを獲得するための前記学習装置100には通信部110とプロセッサ120が含まれる。また、図1ではデータベース130が学習装置100の外部にあると示されているが、前記データベース130は学習装置100に含まれ得る。

【0040】

前記通信部110は外部装置や内部構成要素間のデータをやり取りし得る。特に、前記通信部110は前記トレーニングイメージを獲得し得る。

【0041】

図2は2Dバウンディングボックス及びスード (pseudo) - 3Dボックスを獲得するための前記学習装置100に含まれたCNNの構成を示す。前記CNN内の第1部分201は前記トレーニングイメージ内の前記2Dバウンディングボックスを獲得するように構成され、前記CNN内の第2部分202は前記2Dバウンディングボックスから前記スード (pseudo) - 3Dボックスを獲得するために構成される。

【0042】

前記第1部分201にはコンボリューションレイヤ210とRPN220、プーリング

レイヤ230及び二つのFCレイヤ240、245が含まれ得る。ここに、前記コンボリ
ューションレイヤ210は前記トレーニングイメージから少なくとも一つの特徴マップを
生成し得る。また、前記FCレイヤ240、245は各々客体分類及びボックス回帰を遂
行し得る。

【0043】

前記第2部分202にはプーリングレイヤ250と他の二つのFCレイヤ260、26
5が含まれ得る。ここに、前記FCレイヤ260、265は各々方向分類及び変位回帰を
遂行し得る。

【0044】

また他の例示として、前記FCレイヤ240、245は単一FCレイヤとして具現され
てもよく、前記FCレイヤ260、265も単一FCレイヤとして具現されてもよい。前
記実線矢印は下向きの流れを示し、前記点線矢印はバックプロパゲーションの上向きの流
れを示し得る。

【0045】

図2を参照すれば、前記トレーニングイメージは前記コンボリューションレイヤ210
に入力され得て、以後前記コンボリューションレイヤ210から出力される前記特徴マッ
プは前記RPN220に入力され得る。以後、前記トレーニングイメージ上のいくつかの
ROI (region of interest) に対する情報は前記RPN220から
出力され得る。前記ROI情報は前記第1部分201内の前記プーリングレイヤ230と
前記第2部分202内の前記プーリングレイヤ250に入力され得る。

【0046】

第一に、前記プーリングレイヤ230はmax poolingであり得るプーリング
を遂行し得て、前記二つのFCレイヤ240、245に入力され得るプーリング特徴マッ
プを出力し得る。前記FCレイヤ240は前記客体分類に対する情報、例として前記客体
が自動車である情報を出力し得て、前記FCレイヤ245は前記2Dバウンディングボッ
クスと関連して前記ボックス回帰情報、例として、前記2Dバウンディングボックス頂点
の座標に対する情報を出力し得る。

【0047】

第二に、前記RPN220から前記プーリングレイヤ250に入るようにする前記入力
は前記FCレイヤ240の結果を参照してフィルタリングされ得て、以後前記プーリング
レイヤ250に入力され得る。即ち、前記RPN220から前記プーリングレイヤ250
に入り得る前記入力は前記ROI情報のうち特定ROI情報になり得る。前記特定ROI
情報は前記FCレイヤ240から出力されたクラス情報、例えば自動車、を参照して選択
される。

【0048】

前記プーリングレイヤ250はaverage poolingを遂行し得て、その結
果は前記二つのFCレイヤ260、265に入力され得る。

【0049】

一方、前記第1部分201にある前記FCレイヤ245から出力された前記ボックス回
帰情報は前記第2部分202内の前記FCレイヤ265に入力され得る。

【0050】

ここに、前記FCレイヤ260は前記方向分類情報、例えば前記自動車が向かっている
スード (pseudo) - 3D方向に対する情報、を出力し得て、これは前記プーリング
レイヤ250の出力情報を利用して演算される。また、前記FCレイヤ265は前記変位
回帰に対する情報、例えば前記スード (pseudo) - 3Dボックス頂点の2D座標情
報、を出力し得て、これは前記プーリングレイヤ250出力情報及び前記FCレイヤ24
5から獲得した前記2Dバウンディングボックスに対して前記ボックス回帰情報を利用し
て演算される。前記FCレイヤ260、265から出たかかる結果は前記スード (pse
udo) - 3Dボックスを構成するのに活用され得る。

【0051】

ここで、前記第1部分201は場合によって前記学習装置100に含まれないこともある。即ち、他の装置にある前記第1部分201が前記トレーニングイメージから出た前記2Dバウンディングボックス情報を獲得し得る。以後、前記学習装置100は前述した他の装置から出た前記2Dバウンディングボックス情報が既に含まれたトレーニングイメージ（即ち、いわゆる修正されたトレーニングイメージ）を受信し得る。

【0052】

この場合、前記第2部分202は前記プーリングレイヤ250前に前記プーリングレイヤ250に入力され得る特徴マップを獲得するコンボリユーションレイヤ（未図示）が含まれ得る。また、前記RPN220に入力される前記特徴マップは前記プーリングレイヤ250によって活用され得るが、これは必須条件ではない。

10

【0053】

図3は本発明の一実施例によって前記トレーニングイメージ内の前記2Dバウンディングボックスから前記スード（pseudo）-3Dボックスを獲得する前記学習装置100の流れを示すフローチャートである。

【0054】

第1プロセスとして、前記通信部110はS310段階で前記修正されたトレーニングイメージ、即ち前記2Dバウンディングボックスが囲んだ前記客体を含む前記トレーニングイメージを獲得するか獲得するように支援する。前記修正されたトレーニングイメージは前記客体のみならず背景がある全体イメージや前記2Dバウンディングボックス内の前記客体のみを含むように切った一部イメージを含み得る。

20

【0055】

第2プロセスとして、前記プロセッサ120は前記修正されたトレーニングイメージにコンボリユーション演算を適用するか適用するように支援して特徴マップを獲得し、以後前記特徴マップにプーリング演算を適用するか適用するように支援してプーリング特徴マップを獲得する。以後、前記プーリング特徴マップは上で言及したように前記二つのFCレイヤ260、265に入力される。参考までに、前記第2プロセスは図3に示されていない。

【0056】

第3プロセスとして、前記プロセッサ120は前記FCレイヤ260をもってS320段階の（A）で前記データベース130に保存された前記スード（pseudo）-3D方向の各パターンに対応される予測確率に対する情報を参照して前記客体のスード（pseudo）-3D方向を分類する過程を遂行するようにする。図7を参照すれば、かかるパターンは例えば自動車である前記客体の（a）は後面、（b）は右側後面、（c）は右側面、（d）は右側正面、（e）は正面、（f）は左側正面、（g）は左側面、（h）は左側後面を示し得る。例えば、図4にある前記自動車は右側正面に対応される方向パターンを有するものとして分類され得る。また、前記プロセッサ120は前記FCレイヤ265をもって前記S320段階の（B）で回帰分析を利用して前記スード（pseudo）-3Dボックス頂点の2D座標に対する情報を獲得するようにする。ここで、前記回帰分析は前記スード（pseudo）-3D方向のすべてのパターンに対して遂行される。また、前記スード（pseudo）-3D方向の分類は前記スード（pseudo）-3Dボックス頂点の2D座標を獲得するのに先立って遂行され得る。

30

40

【0057】

第4プロセスとして、前記プロセッサ120はS330段階で（i）獲得した前記スード（pseudo）-3Dボックス頂点の2D座標と前記スード（pseudo）-3Dボックスに対応されるGT（ground truth）の2D座標の差及び（ii）分類された前記スード（pseudo）-3D方向と前記スード（pseudo）-3D方向に対応されるGTの方向の差のうち少なくとも一つを参照して決定されたバックプロパゲーションロス情報で前記CNNのパラメータを調整するか調整するように支援する。ここに、前記スード（pseudo）-3Dボックスに対応されるGTの2D座標は前記客体に正確に外接する3Dボックス頂点の2D座標になり得る。また、前記バックプロパゲ

50

ートされた情報には (i) 回帰分析を利用して演算された、前記スード (p s e u d o) - 3 D 方向の各パターンとそれに対応する 2 D 座標との間の差及び (i i) 前記スード (p s e u d o) - 3 D 方向及び前記 G T の 2 D 座標との間の差、各々を示すロスが含まれ得る。

【0058】

図4は本発明の一実施例によって座標獲得方法を簡略に示す図面である。図4は前記客体に外接する点線からなる前記2Dバウンディングボックスと前記回帰によって獲得される座標を有した頂点がある実線からなる前記スード (p s e u d o) - 3 D ボックスを示す。

【0059】

図4を参照すれば、前記プロセッサ120は回帰分析を利用して前記2Dバウンディングボックス頂点から前記スード (p s e u d o) - 3 D ボックス頂点までの各変位を演算する過程を遂行するか遂行するように支援する。ここに、前記スード (p s e u d o) - 3 D ボックスの頂点と前記スード (p s e u d o) - 3 D ボックスに対応される前記 G T の差に対する情報は前記 C N N のパラメータを調整するためにバックプロパゲートされ得る。

【0060】

もう少し具体的には、前記プロセッサ120は前記2Dバウンディングボックスの頂点 V 1 及びその対角の頂点 V 2 を前記スード (p s e u d o) - 3 D ボックスの二つの頂点 P 1 及び P 8 に決定するか決定するように支援する。以後前記プロセッサ120は回帰分析を利用して P 1 及び P 8 から (前記2Dバウンディングボックスの点線で表示されたラインセグメントと接しない) 二つの頂点 P 4 及び P 5 の変位を演算するか演算するように支援する。ここで、前記変位は前記2Dバウンディングボックスのラインセグメントと平行をなし得る。前記座標 P 4 及び P 5 は各々 P 1 及び P 8 からの変位から直接演算されてもよく、残りの頂点 P 2、P 3、P 6、P 7 座標は前記 P 4 及び P 5 座標から推定され得る。

【0061】

図5は本発明の一実施例によってまた他の前記座標獲得方法を簡略に示す図面である。図5は前記2Dバウンディングボックスの中心と、前記スード (p s e u d o) - 3 D ボックスの第1面に対応される第1ターゲット2Dボックスの推定上の中心と、前記第1面の対蹠点である第2面に対応される第2ターゲット2Dボックスの推定上の中心を示す。

【0062】

図5を参照すれば、前記プロセッサ120は前記第1ターゲット2Dボックスの中心である第1ベクタの終点を決定するために前記2Dバウンディングボックス中心からの前記第1ベクタを、前記第2ターゲット2Dボックスの中心である第2ベクタの終点を決定するために前記2Dバウンディングボックス中心からの前記第2ベクタを演算する過程を遂行するか遂行するように支援する。二つのターゲット2Dボックスの二つの中心は前記回帰分析によって獲得される。前記第1ターゲット2Dボックスの頂点は P 5、P 6、P 7、P 8 であり、前記第2ターゲット2Dボックスの頂点は P 1、P 2、P 3、P 4 である。また、前記プロセッサ120は前記第1ターゲット2Dボックスの幅と前記2Dバウンディングボックスの幅の比率、前記第2ターゲット2Dボックスの幅と前記2Dバウンディングボックスの幅の比率、前記第1ターゲット2Dボックスの高さと前記2Dバウンディングボックスの高さの比率、前記第2ターゲット2Dボックスの高さと前記2Dバウンディングボックスの高さの比率を演算する過程を遂行するか遂行するように支援する。例えば、(i) P 7 から P 8 まで両方向矢印で表示される前記第1ターゲット2Dボックスの幅と (i i) また他のさらに長い両方向矢印で表示される前記2Dバウンディングボックスの幅の比率が獲得される。また、かかる比率は前記回帰分析で得られ得る。併せて、前記プロセッサ120は前記変位と前記比率を参照して P 1、P 2、P 3、P 4、P 5、P 6、P 7、P 8 の座標を演算する過程を遂行するか遂行を支援する。ここで、P 1 は前記2Dバウンディングボックスの頂点であり、P 8 は前記 P 1 の対角線に位置した前記2

10

20

30

40

50

Dバウンディングボックスの頂点である。また、P4及びP5は前記2Dバウンディングボックスのラインセグメントに接しない。即ち、P1、P4、P5、P8の座標がわかれば、残りの頂点の座標は自明である。ここで、前記スード（pseudo）-3Dボックスの頂点と前記スード（pseudo）-3Dボックスに対応される前記GTとの間の差に対する情報も前記CNNのパラメータを調整するためにバックプロパゲートされ得る。

【0063】

図6はテストイメージ内の2Dバウンディングボックスからスード（pseudo）-3Dボックスを獲得するためのテスト装置を示す。

【0064】

図6を参照すれば、前記テスト装置600は通信部610とプロセッサ620が含まれ得る。前記テスト装置600は、前記データベース130の場合と類似して、データベース630を含むこともある。

【0065】

前記テスト装置600はいくつかの装置に分けて各機能が遂行され得る。但し、本明細書では便宜上単一装置として具現されるテスト装置600を記述する。

【0066】

本発明による前記テスト装置600はコンボリューションレイヤ、RPN、FCレイヤなどの一部を含む。

【0067】

ここで、“トレーニング用”という形容詞句は前記学習装置100を説明する時に使用する概念を示すが、前記テスト装置600の説明に使用される“テスト用”に對置される用語として使用される。前記“トレーニング用”という形容詞句は便宜のために前記説明では省略した。

【0068】

詳細には、前記学習装置100が（i）前記トレーニング用2Dバウンディングボックスが囲んだ前記トレーニング用客体を含んだ前記トレーニングイメージを獲得するか獲得するように支援し、（ii）（ii-1）前記スード（pseudo）-3D方向の各パターンに対応される確率情報を参照して前記トレーニング用客体の前記スード（pseudo）-3D方向を分類する過程及び（ii-2）回帰分析を利用して前記トレーニング用スード（pseudo）-3Dボックス頂点の2D座標を獲得する過程を遂行するか遂行するように支援し、（iii）（iii-1）前記トレーニング用スード（pseudo）-3Dボックス頂点の獲得された2D座標と前記トレーニング用スード（pseudo）-3Dボックスに対応される前記GTの2D座標との間の差及び（iii-2）前記分類されたトレーニング用スード（pseudo）-3D方向と前記トレーニング用スード（pseudo）-3Dボックスに対応される前記GT方向の差を参照して決定された情報をバックプロパゲーションして前記パラメータを調整するか調整するように支援した状態で、前記通信部610はテスト用2Dバウンディングボックスに囲まれたテスト用客体を含んだ前記テストイメージを獲得するか獲得するように支援し得る。

【0069】

また、前記プロセッサ620は（i）スード（pseudo）-3D方向の各パターンに対応される確率情報を参照して前記テスト用客体のスード（pseudo）-3D方向を分類する過程及び（ii）回帰分析を利用して前記テスト用スード（pseudo）-3Dボックス頂点の2D座標を獲得する過程を遂行するか遂行するように支援し得る。ここで、前記テスト用スード（pseudo）-3Dボックス頂点の2D座標に対する前記回帰分析は前記スード（pseudo）-3D方向の一つのパターンに対してのみ遂行され得る。

【0070】

残りの過程は、バックプロパゲーションによる前記パラメータの調整と同じテスト用には不要な部分を除いては前記内容とほぼ同一であり省略する。

【0071】

一方、以上で説明された本発明にかかる実施例は多様なコンピュータ構成要素を通じて遂行され得るプログラム命令語の形態で具現されてコンピュータで判読可能な記録媒体に記録され得る。前記コンピュータで判読可能な記録媒体はプログラム命令語、データファイル、データ構造などを単独または組み合わせて含まれ得る。前記コンピュータで判読可能な記録媒体に記録されるプログラム命令語は本発明のために特別に設計されて構成されたものか、コンピュータソフトウェア分野の当業者に公知となって使用可能なものでもよい。コンピュータで判読可能な記録媒体の例には、ハードディスク、フロッピーディスク及び磁気テープのような磁気媒体、CD-ROM、DVDのような光記録媒体、フロプティカルディスク (f l o p t i c a l d i s k) のような磁気-光媒体 (m a g n e t o - o p t i c a l m e d i a) 、及びROM、RAM、フラッシュメモリなどのようなプログラム命令語を保存して遂行するように特別に構成されたハードウェア装置が含まれる。プログラム命令語の例には、コンパイラによって作られるもののような機械語コードだけではなく、インタプリタなどを用いてコンピュータによって実行され得る高級言語コードも含まれる。前記ハードウェア装置は本発明にかかる処理を遂行するために一つ以上のソフトウェアモジュールとして作動するように構成されることがあり、その逆も同様である。

10

【0072】

以上、本発明が具体的な構成要素などのような特定の事項と限定された実施例及び図面によって説明されたが、これは本発明のより全般的な理解を助けるために提供されたものであるに過ぎず、本発明が前記実施例に限定されるものではなく、本発明が属する技術分野において通常の知識を有する者であればかかる記載から多様な修正及び変形が行なわれ得る。

20

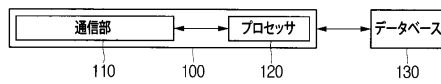
【0073】

従って、本発明の思想は前記説明された実施例に極限されて定められてはならず、後述する特許請求の範囲だけではなく、本特許請求の範囲と均等または等価的に変形されたすべてのものは本発明の思想の範疇に属するといえる。

【図 1】

図 1

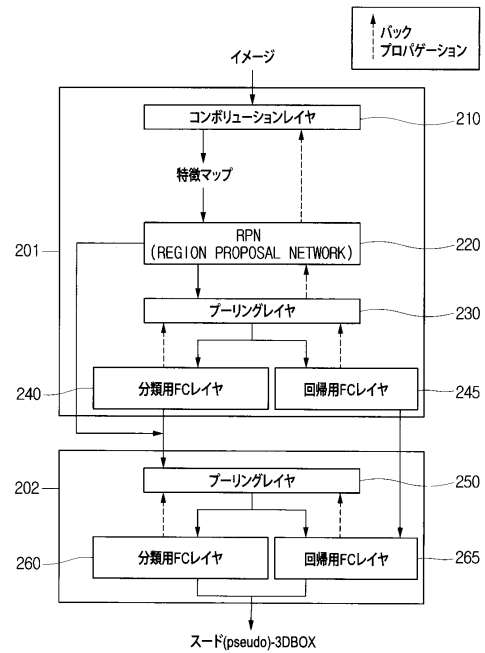
FIG. 1



【図 2】

図 2

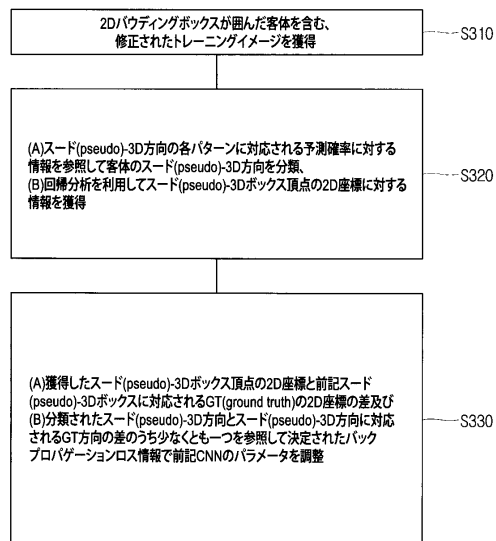
FIG. 2



【図 3】

図 3

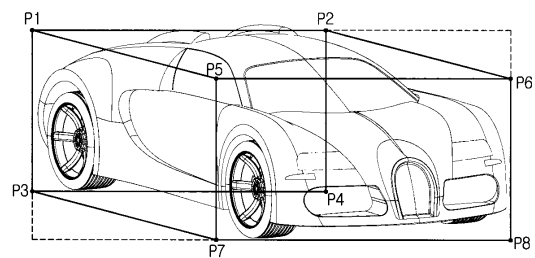
FIG. 3



【図 4】

図 4

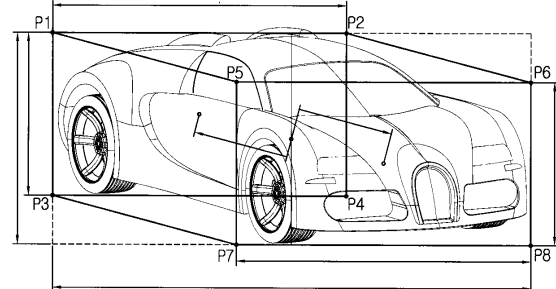
FIG. 4



【図 5】

図 5

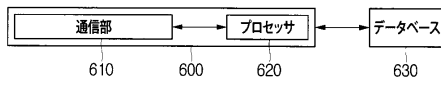
FIG. 5



【図 6】

図 6

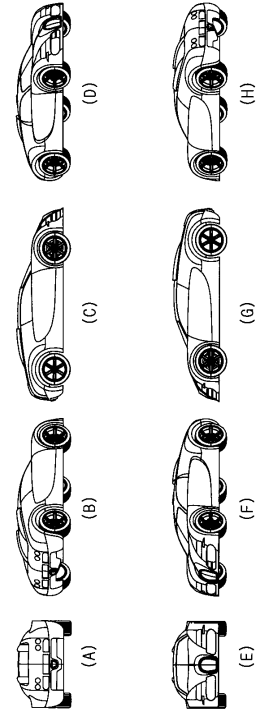
FIG. 6



【図 7】

図 7

FIG. 7



フロントページの続き

- (72)発明者 ナム、ウンユン
大韓民国、ギョンサンブクド、ポハンシーシ、ナムグ、エオニルーエウプ、ユアシーギル 9ベ
オシーギル、57、ナンバー203-803
- (72)発明者 ボ、シュクフン
大韓民国、ギョングード、アニャンシーシ、マナング、マナシーロ 55ベオシーギル、20-9
、ナンバービー02
- (72)発明者 シュン、ミュンチュル
大韓民国、ギョンサンブクド、ポハンシーシ、ブクグ、ジャンヤンロー 174ベオシーギル、
13
- (72)発明者 エオ、ドンフン
大韓民国、ギョンサンブクド、ポハンシーシ、ナムグ、チェオナムロー、67、ナンバー4-4
427
- (72)発明者 リュウ、ウージュ
大韓民国、ギョンサンブクド、ポハンシーシ、ナムグ、チェオナムロー、67、ナンバー4-4
427
- (72)発明者 ジャン、タエウォン
大韓民国、ソウル、ガンナムグ、エオンユーロ 113-ギル、18-5
- (72)発明者 ジョン、キュンジョン
大韓民国、ギョンサンブクド、ポハンシーシ、ナムグ、ジオクロー、294、ヒュンダイグリー
ン・アパートメント、ナンバー232-501
- (72)発明者 ジュ、ホンモ
大韓民国、ギョンサンブクド、ポハンシーシ、ナムグ、ジオクロー、20、セウニ・アパート
メント、ナンバー5-1805
- (72)発明者 チョ、ホジン
大韓民国、ギョンサンブクド、ポハンシーシ、ナムグ、チェオナムロー、67、ナンバー4-4
427

審査官 松浦 功

- (56)参考文献 特開2011-133273 (JP, A)
RAD, M. et al., BB8: A Scalable, Accurate, Robust to Partial Occlusion Method for Pred
icting the 3D Poses of Challenging Objects without Using Depth [online], 2017年
3月31日, [検索日 2019.09.03], インターネット, URL, <https://arxiv.org/pdf/1703.10896v1.pdf>
MOUSAVIAN, A. et al., 3D Bounding Box Estimation Using Deep Learning and Geometry, 201
7 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [online], IEEE, 2
017年 7月, pp. 5632-5640, [検索日 2019.09.03], インターネット, URL, <https://ieeexplore.ieee.org/document/8100080>
LOWE, D.G., Local feature view clustering for 3D object recognition, Proceedings of th
e 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVP
R 2001 [online], IEEE, 2001年12月, pp. I-682-I-688, [検索日 2019.09.03], イン
ターネット, URL, <https://ieeexplore.ieee.org/document/990541>

(58)調査した分野(Int.Cl., DB名)

G06T 7/00 - 7/90
G06T 1/00
IEEE Xplore