

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6800901号
(P6800901)

(45) 発行日 令和2年12月16日(2020.12.16)

(24) 登録日 令和2年11月27日(2020.11.27)

(51) Int.Cl.		F I			
G06T	7/11	(2017.01)	G06T	7/11	
G06T	7/00	(2017.01)	G06T	7/00	350C
G08G	1/16	(2006.01)	G08G	1/16	C

請求項の数 10 (全 18 頁)

<p>(21) 出願番号 特願2018-39979 (P2018-39979)</p> <p>(22) 出願日 平成30年3月6日(2018.3.6)</p> <p>(65) 公開番号 特開2019-153235 (P2019-153235A)</p> <p>(43) 公開日 令和1年9月12日(2019.9.12)</p> <p>審査請求日 令和1年8月16日(2019.8.16)</p>	<p>(73) 特許権者 000003078 株式会社東芝 東京都港区芝浦一丁目1番1号</p> <p>(74) 代理人 110002147 特許業務法人酒井国際特許事務所</p> <p>(72) 発明者 関 晃仁 東京都港区芝浦一丁目1番1号 株式会社東芝内</p> <p>審査官 山田 辰美</p>
---	---

最終頁に続く

(54) 【発明の名称】 物体領域識別装置、物体領域識別方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

画像の各画素が予め定めた複数の物体クラスのいずれに属するかを識別して画素毎に物体種別のラベリングを行う物体領域識別装置であって、

前記各画素において前記複数の物体クラス毎のベースコストを算出するベースコスト算出部と、

前記画像上の隣接する画素間で物体クラスが遷移する際に要する遷移コストを推定する遷移コスト推定部と、

前記画像上に設定した走査方向に沿って、前記ベースコストと前記遷移コストを前記複数の物体クラス毎に累積して、前記各画素における前記複数の物体クラス毎の累積コストを算出する累積コスト算出部と、

前記累積コストに基づいて、前記各画素の物体クラスを決定するクラス決定部と、
を備える物体領域識別装置。

【請求項2】

前記累積コスト算出部は、前記画像上に設定した複数の走査方向毎に前記複数の物体クラス毎の前記遷移コストを推定し、

前記クラス決定部は、前記各画素において、前記複数の物体クラスのうち、前記複数の走査方向毎に算出した前記累積コストを加算した統合累積コストの値が最も小さい物体クラスを、当該画素の物体クラスとして決定する

請求項1に記載の物体領域識別装置。

【請求項 3】

前記累積コスト算出部による前記累積コストの算出および前記クラス決定部による物体クラスの決定は、動的計画法を用いて行われる

請求項 1 または 2 に記載の物体領域識別装置。

【請求項 4】

前記遷移コスト推定部は、前記各画素における前記累積コストが真の物体クラスで最小となるような前記遷移コストを推定するように、学習用データを用いて学習される

請求項 1 乃至 3 のいずれか一項に記載の物体領域識別装置。

【請求項 5】

前記遷移コスト推定部は、ニューラルネットワークにより構成される

10

請求項 4 に記載の物体領域識別装置。

【請求項 6】

前記遷移コスト推定部は、前記画像の注目画素とその周辺の画像小領域を入力とし、前記注目画素における前記遷移コストを推定して出力する

請求項 1 乃至 5 のいずれか一項に記載の物体領域識別装置。

【請求項 7】

前記遷移コスト推定部は、別のニューラルネットワークによって算出された特徴量を入力とし、前記画像の注目画素における前記遷移コストを推定して出力する

請求項 1 乃至 5 のいずれか一項に記載の物体領域識別装置。

【請求項 8】

20

前記遷移コスト推定部は、前記画像上で前記注目画素を特定する位置情報をさらに入力とする

請求項 6 または 7 に記載の物体領域識別装置。

【請求項 9】

画像の各画素が予め定めた複数の物体クラスのいずれに属するかを識別して画素毎に物体種別のラベリングを行う物体領域識別方法であって、

前記各画素において前記複数の物体クラス毎のベースコストを算出し、

前記画像上の隣接する画素間で物体クラスが遷移する際に要する遷移コストを推定し、

前記画像上に設定した走査方向に沿って、前記ベースコストと前記遷移コストを前記複数の物体クラス毎に累積して、前記各画素における前記複数の物体クラス毎の累積コストを算出し、

30

前記累積コストに基づいて、前記各画素の物体クラスを決定する

物体領域識別方法。

【請求項 10】

コンピュータを、画像の各画素が予め定めた複数の物体クラスのいずれに属するかを識別して画素毎に物体種別のラベリングを行う物体領域識別装置として機能させるためのプログラムであって、

前記コンピュータに、

前記各画素において前記複数の物体クラス毎のベースコストを算出する機能と、

前記画像上の隣接する画素間で物体クラスが遷移する際に要する遷移コストを推定する機能と、

40

前記画像上に設定した走査方向に沿って、前記ベースコストと前記遷移コストを前記複数の物体クラス毎に累積して、前記各画素における前記複数の物体クラス毎の累積コストを算出する機能と、

前記累積コストに基づいて、前記各画素の物体クラスを決定する機能と、

を実現させるためのプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明の実施形態は、物体領域識別装置、物体領域識別方法およびプログラムに関する

50

。【背景技術】

【0002】

画像の画素毎に物体種別をラベリングすることは、セマンティックセグメンテーションと呼ばれ、幅広い応用がある。例えば、自動車の自動運転では、車載カメラの映像から自動車、道路、歩行者などの領域を識別することで自動車の走路を決定できる。また、監視カメラの映像から人物の領域を識別することで、不審者の検出や人数カウント等に利用できる。セマンティックセグメンテーションによるラベリングが高精度なほど、その後の処理の信頼性も上がる。

【0003】

セマンティックセグメンテーションの手法としては、画素毎に独立して物体ラベルを求める方法が広く知られている。しかし、物体ラベルは画素毎に大きく変化しないため、画像の大域的な情報を用いることで、ラベリング精度の向上が期待できる。そこで、大域的な情報を用いて画素毎の物体ラベルを決定する方法も提案されている。しかし、大域的な情報を用いて画素毎の物体ラベルを決定する従来の方法は、各画素の物体ラベルを周辺画素との関係から試行錯誤的に求めるものであるため、高精度のラベリングを行うためには計算を繰り返し行う必要がある。このため、特に解像度の高い画像に対しては計算時間が長くなってしまうという問題があり、大域的な情報を用いた高精度なラベリングをより少ない計算量で実現できるようにすることが求められている。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2008-217706号公報

【非特許文献】

【0005】

【非特許文献1】Vijay Badrinarayanan et al., "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.

【非特許文献2】Shuai Zheng et al., "Conditional Random Field as Recurrent Neural Networks", International Conference on Computer Vision, 2015.

【非特許文献3】H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008.

【非特許文献4】Akihito Seki and Marc Pollefeys, "SGM-Nets: Semi-global matching with neural networks", IEEE CVPR, July, 2017.

【発明の概要】

【発明が解決しようとする課題】

【0006】

本発明が解決しようとする課題は、大域的な情報を用いた高精度なラベリングをより少ない計算量で実現できる物体領域識別装置、物体領域識別方法およびプログラムを提供することである。

【課題を解決するための手段】

【0007】

実施形態の物体領域識別装置は、画像の各画素が予め定めた複数の物体クラスのいずれに属するかを識別して画素毎に物体種別のラベリングを行う物体領域識別装置であって、ベースコスト算出部と、遷移コスト推定部と、累積コスト算出部と、クラス決定部と、を備える。ベースコスト算出部は、前記各画素において前記複数の物体クラス毎のベースコストを算出する。遷移コスト推定部は、前記画像上の隣接する画素間で物体クラスが遷移する際に要する遷移コストを推定する。累積コスト算出部は、前記画像上に設定した走査方向に沿って、前記ベースコストと前記遷移コストを前記複数の物体クラス毎に累積して

10

20

30

40

50

、前記各画素における前記複数の物体クラス毎の累積コストを算出する。クラス決定部は、前記累積コストに基づいて、前記各画素の物体クラスを決定する。

【図面の簡単な説明】

【0008】

【図1】実施形態に係る物体領域識別装置の機能的な構成例を示すブロック図。

【図2】遷移コストの概要を説明する図。

【図3-1】画像上に4つの走査方向を設定する例を説明する図。

【図3-2】画像上に8つの走査方向を設定する例を説明する図。

【図4】累積コストを計算する過程を説明する図。

【図5】実施形態に係る物体領域識別装置の処理手順を示すフローチャート。

10

【図6】遷移コスト推定部を学習する際の構成例を示すブロック図。

【図7】学習用データの一例を示す図。

【図8】遷移コスト推定部を学習する際の処理手順を示すフローチャート。

【図9】第1ロスを説明する図。

【図10】第2ロスを説明する図。

【図11】遷移コスト推定部の具体的な構造の一例を示す図。

【図12】走査方向と遷移コストとの関係を説明する図。

【図13】遷移コスト推定部の具体的な構造の他の例を示す図。

【図14】図13に示す構造の遷移コスト推定部を用いた場合の物体領域識別装置の構成例を示す図。

20

【図15】実施形態に係る物体領域識別装置のハードウェア構成例を示すブロック図。

【発明を実施するための形態】

【0009】

以下、添付図面を参照しながら、実施形態の物体領域識別装置、物体領域識別方法およびプログラムについて詳細に説明する。本実施形態は、画像の各画素が予め定めた複数の物体クラスのいずれに属するかを識別して画素毎に物体種別のラベリングを行うものであり、特に、大域的な情報を用いた高精度なラベリングをより少ない計算量で実現する、新規なセマンティックセグメンテーションの手法を提案するものである。

【0010】

大域的な情報を用いたセマンティックセグメンテーションのエネルギー関数は、下記式(1)のように表すことができる。

30

【数1】

$$E(\hat{C}) = \min_C \sum_x \left(U(x, c^x) + \sum_{y \in N_x} p(c^x, c^y) \right) \quad \dots(1)$$

ここで、 x は画像における画素位置、 y は x に隣接する画素位置、 c^x 、 c^y はそれぞれ画素位置 x 、 y における物体クラス c である。 U は画素位置 x 毎・物体クラス c 毎のコストであり、 p は画素位置 x と画素位置 y との間で物体クラス c が変化(遷移)する際に要するペナルティである。 C は画像全体の物体クラスを表し、大域的な情報を用いたセマンティックセグメンテーションは、上記式(1)により最もエネルギーの小さい C^* を求める問題である。

40

【0011】

上記式(1)の右辺第2項のペナルティ p (ペアワイズ項と呼ばれることもある)が0の場合、右辺第1項のコスト U が最小となる物体クラス c が画素位置 x 毎に決定される。これは、画素毎に独立して物体ラベルを求めるセマンティックセグメンテーションの手法(例えば、非特許文献1に開示される手法)と同様である。上記式(1)では、右辺第1項のコスト U に対して、右辺第2項のペナルティ p が加算されることで、画素毎の急激なラベルの変化が抑制される。

【0012】

大域的な情報を用いたセマンティックセグメンテーションの従来手法(例えば、特許文

50

献 1 に開示される手法や非特許文献 2 に開示される手法)では、上記式(1)を解いて高精度なラベリングを行うためには、 $c \times$ (画素位置 x における物体クラス c) を更新しながら安定するまで繰り返し計算を行う必要があり、計算時間が長くなる。これに対し本実施形態では、後述の「累積コスト」の概念を導入し、動的計画法を用いて上記式(1)を近似的に解くことで、1回の操作で画像の各画素の物体クラスを決定する。これにより、繰り返しの計算を行う従来手法と比較して計算時間を大幅に短縮することができる。

【0013】

図 1 は、本実施形態に係る物体領域識別装置 10 の機能的な構成例を示すブロック図である。この物体領域識別装置 10 は、図 1 に示すように、ベースコスト算出部 11 と、遷移コスト推定部 12 と、累積コスト算出部 13 と、クラス決定部 14 と、を備える。

10

【0014】

ベースコスト算出部 11 は、処理対象となる画像を入力とし、この画像の各画素において複数の物体クラス毎のベースコストを算出する。ここで、複数の物体クラスは、用途に応じて予め定められる。上述の自動運转向けの用途では、一例として、「自動車」、「道路」、「歩行者」、「その他」の 4 つの物体クラスが予め定められる。以下では、この事例を適宜想定して説明するが、この例に限定されるものではない。

【0015】

ベースコスト算出部 11 により算出されるベースコストは、上記式(1)のコスト U に相当する。ベースコスト算出部 11 は、例えば、非特許文献 1 に開示されるニューラルネットワーク (SegNet) などを利用して実現できる。ただし、非特許文献 1 のニューラルネットワークは、物体クラス毎に計算した値が最大値を取る物体クラスをその画素の物体クラスとして決定するため、非特許文献 1 のニューラルネットワークにより計算される値 s を、下記式(2)により、ベースコスト U に変換する。

20

【数 2】

$$U = 1 - s \quad \dots(2)$$

【0016】

ここでは、非特許文献 1 のニューラルネットワークで計算される値 s がシグモイド関数などを施すことで 0 ~ 1 に正規化されていることを想定し、上記式(2)によりベースコスト U に変換している。ただし、ベースコスト算出部 11 は、真の物体クラスで値が小さくなるようなベースコスト U を算出できればよく、その実現方法は上述の例に限らない。なお、非特許文献 1 のニューラルネットワークは、事前に大量の画像を使って学習(ネットワークパラメータが更新)されるが、その学習方法については説明を省略する。

30

【0017】

遷移コスト推定部 12 は、画像上の隣接する画素間で物体クラスが遷移する際に要する遷移コストを推定する。遷移コストは、同一の物体クラスでは 0、異なる物体クラス間では非負となるようにし、遷移が容易な物体クラスほど 0 に近い値を持つようにする。遷移コスト推定部 12 により推定される遷移コストは、上記式(1)のペナルティ p に相当する。

【0018】

遷移コストの概要を図 2 を使って説明する。ここでは、自動車の自動運转向けの用途を想定し、物体クラスとして、「自動車」、「道路」、「歩行者」、「その他」の 4 つの物体クラスが予め定められているものとする。図 2 に示す画像例は、自動車 110 が道路 120 上を走行している様子を撮影した画像 100 であり、自動車 110 に属する画素 101 と道路 120 に属する画素 102 とが図 2 に示す位置で隣接している。なお、図 2 では説明のため、画素 101 と画素 102 を実際の画素の大きさよりも大きく描いている。

40

【0019】

図 2 に示す画像 100 において、画素 101 と画素 102 の位置では、画素 101 から画素 102 への方角で真の物体クラスが「自動車」から「道路」に変化している。よって、画素 101 から画素 102 への方角では、物体クラスが「自動車」から「道路」に遷移する遷移コストは 0 に近くなり、その他の物体クラスの遷移、例えば「自動車」「歩行

50

者」や、「その他」「自動車」などの遷移コストは大きな値が推定されることが望ましい。また、図2の画像100では、画素103と画素104がともに道路120に属するため、これら画素103と画素104の位置では、物体クラスが「道路」から他の物体クラスに遷移する遷移コストをはじめ、あらゆる物体クラスの遷移に対する遷移コストは大きな値が推定されることで、画素103と画素104の間で物体クラスが変わってしまうのを防ぐ。

【0020】

遷移コスト推定部12は、処理対象となる画像の位置と方向に応じて上述の遷移コストを推定する。この遷移コスト推定部12は、処理対象となる画像の各画素における後述の累積コストが真の物体クラスで最小となるような遷移コストを推定するように、学習用データをを用いて事前に学習される。遷移コスト推定部12は、例えば、ニューラルネットワークとして構築することができる。この場合、事前の学習によりニューラルネットワークのパラメータ（各層の重みやバイアス）が更新される。なお、遷移コスト推定部12の具体的な学習方法については、詳細を後述する。

10

【0021】

累積コスト算出部13は、処理対象となる画像上に設定した走査方向に沿って、上述のベースコストと遷移コストを複数の物体クラス毎に累積して、処理対象となる画像の各画素における複数の物体クラス毎の累積コストを算出する。クラス決定部14は、累積コスト算出部13により算出された累積コストに基づいて、処理対象となる画像の各画素の物体クラスを決定する。

20

【0022】

本実施形態に係る物体領域識別装置10は、上述のように、動的計画法を用いて上記式(1)を近似的に解くことで、処理対象となる画像の各画素の物体クラスを決定する。つまり、累積コスト算出部13による累積コストの算出およびクラス決定部14による物体クラスの決定が、動的計画法を用いて行われる。

【0023】

動的計画法を用いた累積コストの算出および物体クラスの決定の概要は、以下の通りである。すなわち、画像上の任意の画素位置Aに対し、画像上に設定した走査方向に対して1つ戻った画素位置を画素位置Bとする。このとき、画素位置Aの物体クラス毎のベースコストに、画素位置Bにおける遷移コストと累積コストの最小の組み合わせを加算したものを、画素位置Aの累積コストとして算出する。この操作を、画素位置Aを走査方向に沿って1画素ずつずらしながら行って、画像の全画素について物体クラス毎の累積コストを算出する。画像に複数の走査方向を設定した場合は、上述の操作を走査方向毎に行い、走査方向毎に算出した累積コストを物体カテゴリ毎に加算して統合累積コストとする。そして、画像の各画素において、累積コスト（走査方向が複数の場合は統合累積コスト）が最も小さい値をとる物体クラスを、その画素の物体クラスとして決定する。

30

【0024】

動的計画法を用いた画像処理の技術としては、非特許文献3に開示される技術が知られている。上述の動的計画法を用いた累積コストの算出および物体クラスの決定は、例えば、この非特許文献3に開示される技術を応用することで実現できる。ただし、非特許文献3に開示される技術はステレオ画像を利用した視差画像の生成に関するものであり、これをそのまま適用することはできない。このため、新たに定式化を行う。以下では、非特許文献3に開示される技術を応用した累積コストの算出および物体クラスの決定の具体例について詳しく説明する。

40

【0025】

まず、処理対象となる画像上に走査方向を設定する。走査方向は複数設定することができる。走査方向の数が多ければ、セマンティックセグメンテーションのラベリング精度の向上が期待できる反面、計算量が多くなる。このため、用途に応じて要求される精度と装置の演算処理能力とに応じて、最適な数の走査方向を設定することが望ましい。図3-1は、画像200上に4つの走査方向を設定した例を示しており、画素位置 x_0 に向かって

50

、水平に左から右への走査方向 201 と、水平に右から左への走査方向 202 と、垂直に上から下への走査方向 203 と、垂直に下から上への走査方向 204 とが設定されている。また、図 3 - 2 は、画像 200 上に 8 つの走査方向を設定した例を示しており、図 3 - 1 で示した 4 つの走査方向に加えて、画素位置 x_0 に向かって、左下から右上へ 45 度の走査方向 205 と、右上から左下へ 45 度の走査方向 206 と、左上から右下へ 45 度の走査方向 207 と、右下から左上へ 45 度の走査方向 208 とが設定されている。

【0026】

次に、処理対象となる画像上に設定した走査方向毎に、画像の各画素における累積コストを算出する。走査方向毎に行う処理は同一であるため、以下では 1 つの走査方向に着目して説明する。

10

【0027】

図 4 は、画素位置 x_0 に向かって累積コストを計算する過程を説明する図である。図 4 の横軸は画素位置、縦軸は物体クラスをそれぞれ示している。画素位置 x_n は、画素位置 x_0 から走査方向に対して n 画素分戻った画素位置を示している。物体クラスの数 は固定であり、ここでは 4 つの物体クラス $c_1 \sim c_4$ (自動車・道路・歩行者・その他) があるものとしている。図 4 の例では、画素位置 x_3 と画素位置 x_2 の間と、画素位置 x_1 と画素位置 x_0 との間で物体クラスが変化しており、それぞれ遷移コスト p' , p'' が生じている。図中の EV は、画素位置 x_1 と画素位置 x_0 との間を拡大したものである。 P_{2_1} は、物体クラス c_2 から物体クラス c_1 への遷移コストである。

【0028】

20

累積コスト算出部 13 は、下記式 (3) に従って、走査方向 r 毎に、画像の各画素における物体クラス c 毎の累積コスト L_r を算出する。下記式 (3) は、画素位置 x_0 における物体クラス c_k の累積コスト $L_r(x_0, c_k)$ を求める式である。

【数 3】

$$L_r(x_0, c_k) = U(x_0, c_k) + \min \left(L_r(x_1, c_k), \min_{i \in c} (L_r(x_1, c_i) + P_{i \rightarrow k}(x_1)) \right) \quad \dots (3)$$

ただし、 $U(x_0, c_k)$ は画素位置 x_0 における物体クラス c_k のベースコスト、 $L_r(x_1, c_k)$ は画素位置 x_1 における物体クラス c_k の累積コスト、 $L_r(x_1, c_i)$ は画素位置 x_1 における物体クラス c_i の累積コスト、 $P_{i \rightarrow k}(x_1)$ は画素位置 x_1 と画素位置 x_0 との間で物体クラスが c_i から c_k に遷移する際に要する遷移コストである。

30

【0029】

上記式 (3) に示すように、画像の各画素における走査方向毎の累積コストは、走査方向に対して 1 つ戻った画素位置における累積コストと遷移コストを比較して最小になる物体クラスの累積コストを当該画素のベースコストに加算することで求められる。累積コスト算出部 13 は、画像上に設定した走査方向に沿って各画素の累積コストをベースコストに順次加算していくことで、画像の各画素における物体クラス毎の累積コストを算出する。

【0030】

40

なお、この計算方法では、走査方向に沿って隣接画素間で累積コストを加算していくため、装置の演算処理能力によっては値がオーバーフローする懸念もある。このような場合には、下記式 (4) に示すように、走査方向に対して 1 つ戻った画素位置における累積コストの最小値を減算した L_r' を L_r の代わりに用いてもよい。

【数 4】

$$L_r'(x_0, c_k) = L_r(x_0, c_k) - \min_{i \in c} L_r(x_1, c_i) \quad \dots (4)$$

【0031】

走査方向毎に累積コストを求める上述の処理は他の走査方向との相互作用がないため、複数方向で並列して計算することが可能である。さらに、走査方向と平行な画素位置では

50

相互作用がないため、位置に関しても並列に計算可能である。例えば、図3-1に示した例のように4つの走査方向を設定した場合には、走査方向201~204で並列に計算でき、かつ、例えば走査方向201については、画像200の縦方向の画素分だけ並列に計算可能である。

【0032】

次に、累積コスト算出部13は、以上のように走査方向毎に算出した累積コスト L_r を下記式(5)に従って加算し、画像の各画素における物体クラス毎の統合累積コスト L を求める。下記式(5)は、画素位置 x_0 における物体クラス c_k の統合累積コスト $L(x_0, c_k)$ を求める式である。なお、画像上に設定された走査方向が1つのみであれば、統合累積コスト $L =$ 累積コスト L_r である。

10

【数5】

$$L_r(x_0, c_k) = \sum_{r \in R} L_r(x_0, c_k) \quad \dots(5)$$

【0033】

累積コスト算出部13により物体クラス毎の統合累積コスト L が求められると、クラス決定部14が、画像の各画素において、下記式(6)に従って、統合累積コスト L の値が最も小さい物体クラスを選択し、これを当該画素の物体クラスとして決定する。下記式(6)は、画素位置 x_0 の物体クラス c^{x_0} を決定する式である。

【数6】

$$c^{x_0} = \operatorname{argmin}_{k \in c} L(x_0, c_k) \quad \dots(6)$$

20

【0034】

以上のようにして、処理対象となる画像の各画素の物体クラスが決定されることにより、処理対象となる画像は画素毎に物体種別のラベリングが行われる。例えば、上述の自動運転の事例では、処理対象となる画像の各画素が「自動車」、「道路」、「歩行者」、「その他」のいずれかの物体ラベルでラベリングされ、画像における物体領域が識別される。

【0035】

図5は、本実施形態に係る物体領域識別装置10の処理手順を示すフローチャートである。なお、図5の各ステップの具体的な内容は上述した通りであるので、詳細な説明は適宜省略する。

30

【0036】

処理対象となる画像が入力されると(ステップS101)、まず、ベースコスト算出部11が、ステップS101で入力された画像の各画素において予め定めた複数の物体クラス毎のベースコストを算出する(ステップS102)。

【0037】

次に、遷移コスト推定部12が、ステップS101で入力された画像上の隣接する画素間で物体クラスが遷移する際に要する遷移コストを推定する(ステップS103)。

【0038】

40

次に、累積コスト算出部13が、ステップS101で入力された画像上に設定した走査方向毎に、各画素の物体クラス毎の累積コストを算出する(ステップS104)。そして、累積コスト算出部13は、ステップS104で走査方向毎に算出した累積コストを加算して、統合累積コストを算出する(ステップS105)。

【0039】

次に、クラス決定部14が、ステップS101で入力された画像の各画素について、複数の物体クラスのうち、ステップS105で算出された統合累積コストの値が最も小さい物体クラスを、その画素の物体クラスとして決定する(ステップS106)。

【0040】

以上説明したように、本実施形態に係る物体領域識別装置10は、大域的な情報を用い

50

たセマンティックセグメンテーションを実現するにあたり、処理対象となる画像上に設定した走査方向毎に上記式(3)に従って各画素の物体クラス毎の累積コストを算出する。そして、複数の走査方向を設定した場合は、走査方向毎の累積コストを加算して統合累積コストを算出する。そして、画像の各画素において統合累積コスト(走査方向が1つの場合は累積コスト)の値が最も小さい物体クラスを、その画素の物体クラスとして決定する。したがって、従来技術のように繰り返しの計算を行うことなく1回の操作で各画素の物体クラスを決定することができ、大域的な情報を用いた高精度なラベリングをより少ない計算量で実現することができる。

【0041】

次に、上述の遷移コストを推定する遷移コスト推定部12がニューラルネットワークにより構成されている場合を想定し、この遷移コスト推定部12の具体的な学習方法について説明する。遷移コスト推定部12の学習時には、図6に示すように、クラス決定部14の代わりに更新部15が設けられ、累積コスト算出部13の出力が更新部15に入力される。更新部15は、累積コスト算出部13の出力が教示データに近づくように、遷移コスト推定部12を学習(ニューラルネットワークのネットワークパラメータを更新)する。

10

【0042】

教示データとは、画像に対応する正しいラベリング結果を示すデータである。画像と教示データとの組み合わせが学習用データである。学習用データの一例を図7に示す。図7に例示する学習用データは、図2で例示した画像100と、この画像100に対応する教示データ300とを含む。教示データ300は、画像100の自動車110が映る領域310内の各画素が「自動車」の物体ラベルでラベリングされ、画像100の道路120が映る領域320内の各画素が「道路」の物体ラベルでラベリングされ、その他の領域330内の各画素が「その他」の物体ラベルでラベリングされていることを示している。更新部15が遷移コスト推定部12を適切に学習するには、多数の学習用データが必要となる。

20

【0043】

遷移コスト推定部12は、画素が隣接する方向に対して画像全体で固定値の遷移コストを推定するタイプと、画素位置の周辺のパターンに応じて変動する遷移コストを推定するタイプとの2つのタイプに大別される。前者のタイプの遷移コスト推定部12は、推定する遷移コストが固定値であるため、任意の正值を設定すればよい。遷移コストの決め方の1つの指針として、教示データに含まれる物体ラベル数に応じて決めればよい。すなわち、隣接する画素間で物体クラスが遷移する態様についてヒストグラムを作成し、出現頻度が大きい遷移に関しては相対的に小さい遷移コストを割り当て、出現頻度が小さい遷移に関しては相対的に大きい遷移コストを割り当てればよい。このため、前者のタイプの遷移コスト推定部12は、事前の学習が容易である。

30

【0044】

しかし、実際には、画像パターンによって異なる遷移コストを与えるべきであると考えられる。例えば図2に示した画像例の画素103, 104は典型的な道路に属する画像パターンとなっている。このような場合には、自動車から歩行者に遷移する際の遷移コストなど、最もらしからぬ遷移に対する遷移コストは大きくするべきである。後者のタイプの遷移コスト推定部12は、このような要求に応えるものであり、例えばニューラルネットワークを用いて実現することができる。

40

【0045】

以下では、後者のタイプの遷移コスト推定部12の具体的な学習方法について説明する。なお、遷移コスト推定部12を学習する際は、多数の学習用データを用いて、以下に示す遷移コスト推定部12の学習を繰り返し行う。

【0046】

遷移コスト推定部12の学習は、例えば、非特許文献4に開示される技術を応用して実施することができる。ただし、非特許文献4に開示される技術はステレオ画像を利用した視差画像の生成に関するものであり、これをそのまま適用することはできない。このため

50

、新たに定式化を行う。この例の場合、まず、遷移コストの更新を行う画素位置を複数個選出する。続いて、それらの画素位置を使って第1ロスおよび第2ロスを算出する。そして、これらの第1ロスおよび第2ロスに基づいて、遷移コスト推定部12を学習する。

【0047】

図8は、遷移コスト推定部12を学習する際の処理手順を示すフローチャートである。以下、このフローチャートに沿って、遷移コスト推定部12の具体的な学習方法について詳しく説明する。

【0048】

ステップS201では、遷移コストの更新を行う画素位置を画像から複数個選出する。これらの画素位置は、画像内のすべての画素位置を選択してもよいが、現実的な計算速度で遷移コスト推定部12の学習を行うには、画像内から離散的に一部の画素位置を選出せば十分である。

10

【0049】

ステップS202では、第1ロスを算出する。さらに、遷移コスト推定部12は、ニューラルネットワークで構築されるため、一般的な学習手法である誤差逆伝播法(Backpropagation)に必要な微分も求める。

【0050】

第1ロスは、下記式(7)に従って算出することができる。下記式(7)は、画素位置 x_0 における走査方向 r に対応する第1ロス $H_1(x_0, r)$ を求める式である。

【数7】

20

$$H_1(x_0, r) = \sum_{c_i^{x_0} \neq c_{gt}^{x_0}} \max(0, L_r(x_0, c_{gt}^{x_0}) - L_r(x_0, c_i^{x_0}) + m) \quad \dots(7)$$

【0051】

上記式(6)により、統合累積コストの値が最小となる物体クラスをその画素位置の物体クラスとして決定するため、真の物体クラス c_{gt} はそれ以外の物体クラス c_i よりマージン m 以上小さくなるように、ヒンジロスの形式で設計された上記式(7)に従って、第1ロスを算出する。なお、画素位置 x_0 はステップS202で選択された画素位置である。上記式(7)では、走査方向 r も考慮している。例えば、自動車はほとんどの場合、道路上に存在することから、例えば走査方向が下から上であるとき、物体クラスが自動車からその他に変わることはあるが、走査方向が上から下であるときに、物体クラスが自動車からその他になることはまれなはずである。走査方向 r を考慮することにより、こうした傾向を遷移コスト推定部12の学習に反映させることができる。

30

【0052】

続いて、上記式(7)から、遷移コストに対する微分を求める。物体クラスの遷移の一例を図9に示す。図9に示す例では、画素位置 x_0 で物体クラス c_1 が真の物体クラス c_{gt} となっている。ここで、画素位置 x_3 から画素位置 x_0 までの間で、物体クラス c_3 が物体クラス $c_1 (= c_{gt})$ に遷移するパス401と、物体クラス c_3 が物体クラス $c_4 (= c_i)$ に遷移するパス402とを考える。これらのパス401, 402の累積コストを上記式(3)に従って求める。すると、パス401の累積コストは下記式(8)、パス402の累積コストは下記式(9)でそれぞれ記述できる。

40

【数8】

$$L_r(x_0, c_{gt}^{x_0}) = U(x_0, c_{gt}^{x_0}) + U(x_1, c_1^{x_1}) + U(x_2, c_3^{x_2}) + U(x_3, c_3^{x_3}) + P_{r3 \rightarrow 1}(x_2) \quad \dots(8)$$

【数9】

$$L_r(x_0, c_4^{x_0}) = U(x_0, c_4^{x_0}) + U(x_1, c_4^{x_1}) + U(x_2, c_3^{x_2}) + U(x_3, c_3^{x_3}) + P_{r3 \rightarrow 4}(x_2) \quad \dots(9)$$

【0053】

上記式(8)および上記式(9)を上記式(7)を代入し、下記式(10)の条件を満たす場合、遷移コストに対する微分は下記式(11)のように求めることができる。下記

50

式(10)の条件を満たさない場合は、遷移コストに対する微分はすべて0である。

【数10】

$$L_r(x_0, c_{gt}^{x_0}) - L_r(x_0, c_4^{x_0}) + m > 0 \quad \dots(10)$$

【数11】

$$\frac{\partial C'}{\partial P_{3 \rightarrow 4}(x_2)} = -1, \frac{\partial C'}{\partial P_{3 \rightarrow 1}(x_2)} = 1, \frac{\partial C'}{\partial P_{2 \rightarrow 1}(x_1)} = 0, \frac{\partial C'}{\partial P_{3 \rightarrow 1}(x_1)} = 0 \quad \dots(11)$$

【0054】

よって、微分が非0となる画素位置およびその遷移を保持しておき、後述のステップS204において遷移コスト推定部12を誤差逆伝播法により学習する際に、微分の値を利用することができる。

10

【0055】

ステップS203では、第2ロスを算出する。第1ロスは累積コストを計算する際の途中の画素位置における真の物体クラスを考慮していない。よって、遷移コストによっては、例えば図9に示すパス403のように、パス401とは異なるパスを通して真の物体クラス c_{gt} に到達することもありえる。この場合、パス403の累積コストは下記式(12)のようになる。

【数12】

$$L_r(x_0, c_{gt}^{x_0}) = U(x_0, c_{gt}^{x_0}) + U(x_1, c_1^{x_1}) + U(x_2, c_1^{x_2}) + U(x_3, c_3^{x_3}) + P_{r3 \rightarrow 1}(x_3) \quad \dots(12)$$

20

【0056】

上記式(12)と上記式(8)とを比べると、右辺第3項と第5項が異なっていることが分かる。これによって正しく遷移コストが求められないため、本ステップで第2コストを算出する。第2ロスの算出には2つの条件が必要である。1つめは、画素位置 x_0 と画素位置 x_1 の両方で真の物体クラス c_{gt} が既知であること、2つめは、画素位置 x_1 で累積コストが最小値をとる物体クラスが真の物体クラス c_{gt} と一致していることである。これら2つの条件が満たされた画素位置に対して、下記式(13)により第2ロスを計算する。下記式(13)は、画素位置 x_0 における走査方向 r に対応する第2ロス $H_2(x_0, r)$ を求める式である。

30

【数13】

$$H_2(x_0, r) = \sum_{d_1^{x_0} \neq d_{gt}^{x_0}} \max(0, S(\cdot) - N(\cdot) + m) \quad \dots(13)$$

【0057】

上記式(13)において、 $S(\cdot)$ は、図10に示すパス501のように、画素位置 x_1 における真の物体クラス c_{gt} から画素位置 x_0 で真の物体クラス c_{gt} に到達するパスである。また、 $N(\cdot)$ はそれ以外のパスであり、図10に示すパス群502のように、画素位置 x_1 における真でない物体クラス c_i から画素位置 x_0 で真の物体クラス c_{gt} に到達するパス群と、図10に示すパス群503のように、画素位置 x_0 で真でない物体クラス c_i に到達するパス群である。図10に示す例では、 $S(\cdot)$ と $N(\cdot)$ はそれぞれ下記式(14)、下記式(15)のように記述できる。

40

【数14】

$$S(\cdot) = U(x_0, c_{gt}^{x_0}) + L_r(x_1, c_{gt}^{x_1}) + P_{r4 \rightarrow 1}(x_1) \quad \dots(14)$$

【数15】

$$N(\cdot) = U(x_0, c_{\#}^{x_0}) + L_r(x_1, c_{*}^{x_1}) + P_{r* \rightarrow \#}(x_1) \quad \dots(15)$$

上記式(15)の*とは画素位置 x_1 と画素位置 x_0 との物体クラスの組み合わせであり、両画素位置で真の物体クラス c_{gt} となる組み合わせは除く。

50

【 0 0 5 8 】

第2ロスを求める上記式(13)から、第1ロスと同様に、遷移コストに対する微分を求めることができる。

【 0 0 5 9 】

ステップS204では、ステップS202とステップS203で計算される第1ロスと第2ロスの双方を用いて、遷移コスト推定部12を学習する。例えば、第1ロスと第2ロスから、下記式(16)に示すようにロスHを作成する。そして、このロスHが小さくなるように、遷移コスト推定部12を学習する。

【 数 1 6 】

$$H = \sum_r \left\{ \sum_{x \in X_1} H_1(x, r) + \sum_{x \in X_2} H_2(x, r) \right\} \quad \dots (16)$$

10

【 0 0 6 0 】

上記式(16)で示されるロスHは、第1ロスと第2ロスの組み合わせである。したがって、上述の第1ロスに関する遷移コストに対する微分の値と、第2ロスに関する遷移コストに対する微分の値とを用いて、遷移コスト推定部12を構成するニューラルネットワークを誤差逆伝播法により適切に学習(ネットワークパラメータを更新)することができる。

【 0 0 6 1 】

遷移コスト推定部12の学習は、入力する学習用データを順次切り替えながら、累積コスト算出部13による累積コストの算出および図8のフローチャートに従った処理を繰り返す。学習の終了条件としては、例えば、所定回数以上の繰り返し回数を設定してもよいし、Validationデータを用いた検証によりセマンティックセグメンテーションのエラーが収束したと判断される場合に終了してもよい。

20

【 0 0 6 2 】

次に、遷移コスト推定部12の具体的な構造例について説明する。図11は、遷移コスト推定部12の具体的な構造の一例を示す図である。図11に示す遷移コスト推定部12は、画像パッチ601を入力とするニューラルネットワークである。画像パッチ601は、画素位置 x_0 (注目画素)とその周辺の画像小領域である。画像パッチ601はマルチスケール化したり、カラー画像を入力したりしてもよい。また、画像パッチ601に加えて、画像全体における画像パッチ601の位置(注目画素の位置)を特定する位置情報602を遷移コスト推定部12の入力として用いてもよい。

30

【 0 0 6 3 】

例えば車載カメラにより撮影された画像では、画像の下方は道路の領域が多く、自動車は画像の中央付近に映るなど、分布に偏りがある。この情報を有効に利用するために、位置情報602をニューラルネットワークの入力の1つとして用いることが望ましい。

【 0 0 6 4 】

ニューラルネットワークの最終段603では、出力604を正值とするため、ELU(Exponential Linear Unit)と定数を足している。出力604が正值となるレイヤであればELUに限定されない。出力604は、物体クラスの遷移コストである。例えばN個の物体クラスがあり、M個の方向で遷移コストを推定する場合(走査方向がM個ある場合)の出力数は、 $N \times (N - 1) \times M$ である。

40

【 0 0 6 5 】

なお、図12に示すように、物体クラス c_1 に属する画素701と物体クラス c_2 に属する画素702がある場合、走査方向703に沿って物体クラスが c_1 から c_2 に遷移する際の遷移コストと、走査方向703と逆方向である走査方向704に沿って物体クラスが c_2 から c_1 に遷移する際の遷移コストとが同じになるように学習することで、出力604の数を半分に減らしてもよい。出力604の数を減らすことにより、計算時間をさらに短縮したり、学習に用いる学習用データの数を削減したりすることが期待できる。

【 0 0 6 6 】

50

図13は、遷移コスト推定部12の具体的な構造の他の例を示す図である。図13に示す遷移コスト推定部12は、ニューラルネットワークとして構成されることは図11の例と同様であるが、図11に示した画像パッチ601の代わりに別のニューラルネットワーク801の特徴量を入力として用いる。ニューラルネットワーク801として、例えばベースコスト算出部11として非特許文献1のニューラルネットワークを用いる場合は、このニューラルネットワークの特徴量をエンコードする部分（解像度を下げて畳み込み演算をすることにより計算された特徴量マップ）を用いることができる。すなわち、ベースコスト算出部11の中間出力である特徴量を、遷移コスト推定部12の入力として用いることができる。遷移コスト推定部12の入力は、画素位置ごとに異なる特徴量となっていることが望ましい。非特許文献1のニューラルネットワークでは、低解像度から高解像度までマルチスケールの情報を用いることができる。

10

【0067】

また、図11の例と同様に、画像全体における画素位置を特定する位置情報802を特徴量と併せて遷移コスト推定部12に入力するようにしてもよい。さらに、これ以外にも遷移コスト推定部12の入力として、画像が撮影された場所、地域、時間、天候、カメラの種類などの情報を用いたり、RGB画像以外にも深度を加えたRGBDや深度画像のみや、赤外画像などを用いたりしてもよい。ニューラルネットワークの最終段803は、図11の例と同様であり、物体クラスの遷移コストを出力804とする。

【0068】

図14は、図13に示す構造の遷移コスト推定部12を用いた場合の物体領域識別装置10の構成例を示す図である。図14に示す物体領域識別装置10は、入力画像901の各画素における物体クラス毎のベースコストをベースコスト算出部11で算出する。また、大域的な情報を用いるため、ベースコスト算出部11の中間出力である特徴量を用いて、遷移コスト推定部12において物体クラス間の遷移コストを推定する。そして、ベースコスト算出部11の出力と遷移コスト推定部12の出力とを用いて、累積コスト算出部13において各画素における物体クラス毎の累積コストを算出する。走査方向が複数ある場合は、走査方向毎の累積コストを加算して統合累積コストを算出する。そして、累積コスト算出部13の出力を用いて、クラス決定部14において入力画像901の各画素の物体クラスを決定し、入力画像901に対応するラベリング結果（セマンティックセグメンテーションの結果）902を出力する。

20

30

【0069】

上述の特許文献1や非特許文献2で開示される従来手法により推定されるペナルティ（ペアワイズ項）は、ガウス関数にRGBで表現される画素値や画素位置の距離などを入力し、その標準偏差や複数のガウス関数の重みで調整した後、画像特徴に関係なく決まっている物体クラス間の重みであるCompatibilityを最後に掛け合わせ算出したものである。ところが物体クラス間の重みは画像特徴と相関があるため、従来手法により推定されるペナルティは十分な表現能力がない。これに対し、本実施形態に係る物体領域識別装置10の遷移コスト推定部12は、入力された画像パッチやニューラルネットワークの中間出力である特徴量に対して、特徴量を改めて自動で学習し、学習された特徴量から遷移コスト（ペナルティに相当）を直接推定するため、十分な表現能力を持つ。すなわち、本実施

40

【0070】

上述した本実施形態に係る物体領域識別装置10は、例えば、汎用のコンピュータ装置を基本ハードウェアとして用いることで実現することが可能である。すなわち、上述の物体領域識別装置10の各部の機能は、汎用のコンピュータ装置に搭載された1以上のプロセッサにプログラムを実行させることにより実現することができる。このとき、物体領域識別装置10は、上記のプログラムをコンピュータ装置にあらかじめインストールすることで実現してもよいし、CD-ROMなどの記憶媒体に記憶して、あるいはネットワーク

50

を介して上記のプログラムを配布して、このプログラムをコンピュータ装置に適宜インストールすることで実現してもよい。

【0071】

図15は、本実施形態に係る物体領域識別装置10のハードウェア構成例を示すブロック図である。物体領域識別装置10は、例えば図15に示すように、CPU (Central Processing Unit) などのプロセッサ21と、RAM (Random Access Memory) やROM (Read Only Memory) などのメモリ22と、HDD (Hard Disk Drive) やSSD (Solid State Drive) などのストレージデバイス23と、液晶パネルなどの表示装置26や画像を撮影するカメラ27といった機器を接続するための機器I/F24と、装置外部と通信を行う通信I/F25と、これら各部を接続するバス28とを備えた一般的なコンピュータとしてのハードウェア構成を有する。

10

【0072】

物体領域識別装置10が図15のようなハードウェア構成を有する場合、例えば、プロセッサ21がメモリ22を利用して、ストレージデバイス23などに格納されたプログラムを読み出して実行することにより、上述の各部(ベースコスト算出部11、遷移コスト推定部12、累積コスト算出部13、クラス決定部14(更新部15))の機能を実現することができる。

【0073】

なお、上述の物体領域識別装置10の各部の機能は、その一部または全部を、ASIC (Application Specific Integrated Circuit) やFPGA (Field-Programmable Gate Array) などの専用のハードウェア(汎用のプロセッサではなく専用のプロセッサ)により実現することもできる。また、複数のプロセッサを用いて上述した各部の機能を実現する構成であってもよい。

20

【0074】

また、本実施形態に係る物体領域識別装置10は、複数台のコンピュータ装置を用い、上述の各部の機能を複数台のコンピュータ装置に分散して実現した物体領域識別システムとして構成してもよい。また、本実施形態に係る物体領域識別装置10は、クラウドシステム上で動作する仮想マシンであってもよい。

【0075】

以上、本発明の実施形態を説明したが、ここで説明した実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。ここで説明した新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。ここで説明した実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。

30

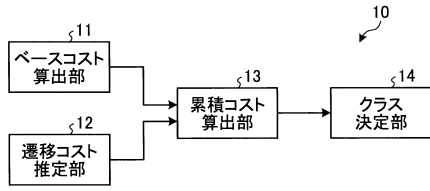
【符号の説明】

【0076】

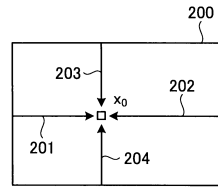
- 10 物体領域識別装置
- 11 ベースコスト算出部
- 12 遷移コスト推定部
- 13 累積コスト算出部
- 14 クラス決定部
- 15 更新部

40

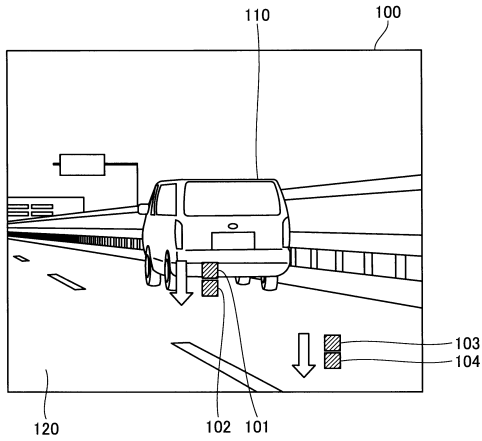
【図1】



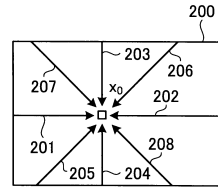
【図3-1】



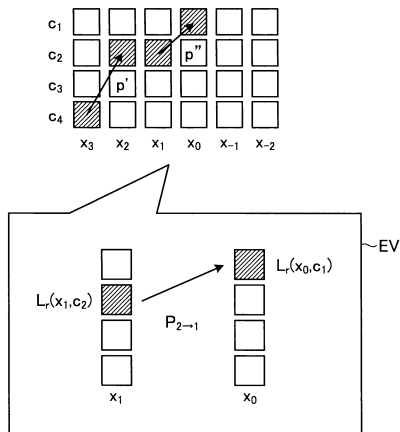
【図2】



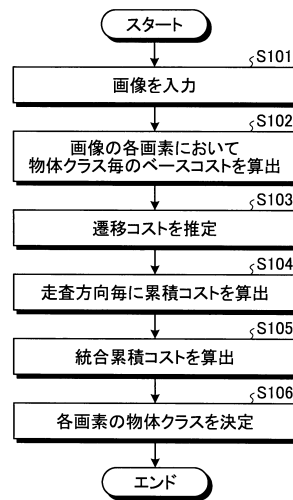
【図3-2】



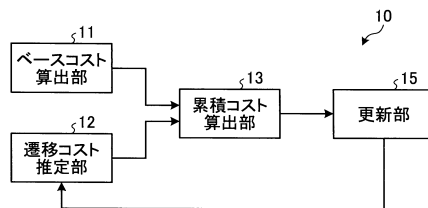
【図4】



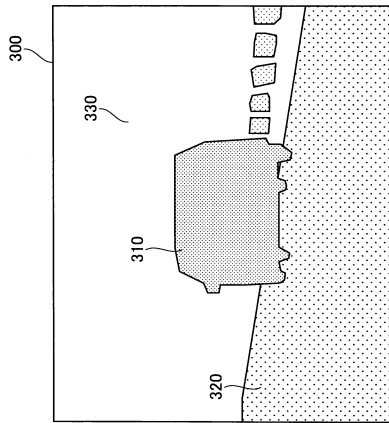
【図5】



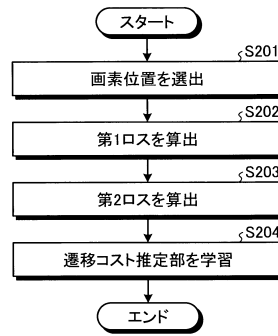
【図6】



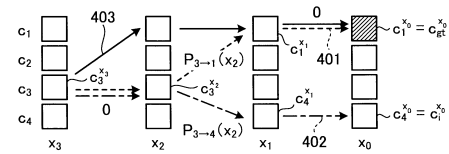
【図7】



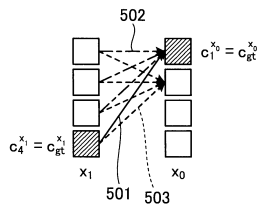
【図8】



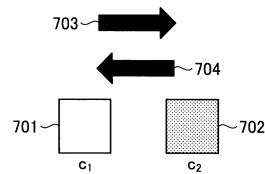
【図9】



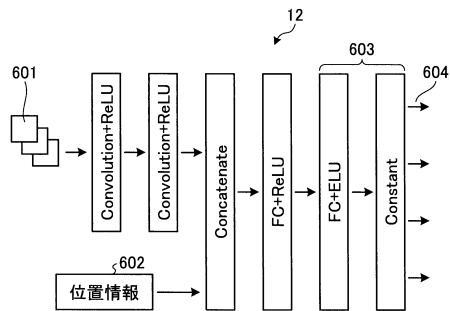
【図10】



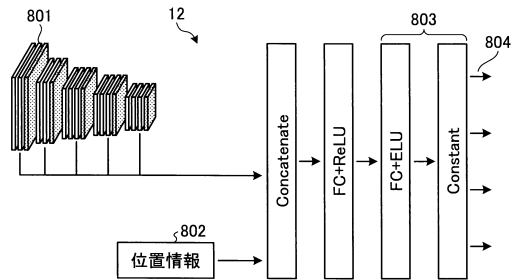
【図12】



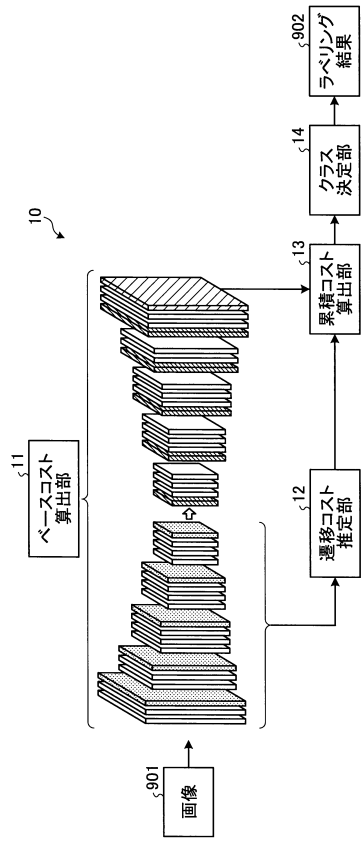
【図11】



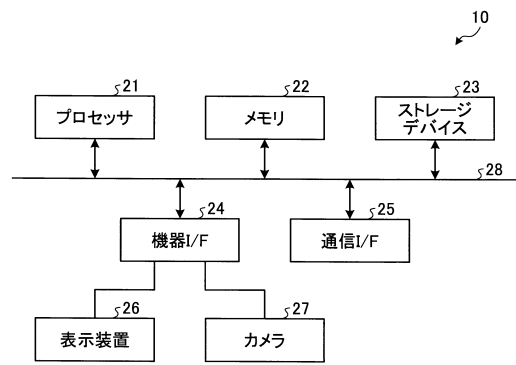
【図13】



【図14】



【図15】



フロントページの続き

- (56)参考文献 特開2018-014041(JP,A)
国際公開第2017/163759(WO,A1)
国際公開第2013/172413(WO,A1)

- (58)調査した分野(Int.Cl., DB名)
G06T 7/00-7/90
G08G 1/16
IEEE Xplore