

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6299427号
(P6299427)

(45) 発行日 平成30年3月28日 (2018.3.28)

(24) 登録日 平成30年3月9日 (2018.3.9)

(51) Int. Cl.

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 6 5 0 Z

G 0 6 T 7/215 (2017.01)

G 0 6 T 7/215

請求項の数 25 (全 42 頁)

(21) 出願番号 特願2014-107335 (P2014-107335)
 (22) 出願日 平成26年5月23日 (2014.5.23)
 (65) 公開番号 特開2014-235737 (P2014-235737A)
 (43) 公開日 平成26年12月15日 (2014.12.15)
 審査請求日 平成28年8月2日 (2016.8.2)
 (31) 優先権主張番号 61/829, 980
 (32) 優先日 平成25年5月31日 (2013.5.31)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 14/171, 677
 (32) 優先日 平成26年2月3日 (2014.2.3)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 000003207
 トヨタ自動車株式会社
 愛知県豊田市トヨタ町1番地
 (74) 代理人 100100549
 弁理士 川口 嘉之
 (74) 代理人 100085006
 弁理士 世良 和信
 (74) 代理人 100113608
 弁理士 平川 明
 (74) 代理人 100123319
 弁理士 関根 武彦
 (74) 代理人 100123098
 弁理士 今堀 克彦
 (74) 代理人 100143797
 弁理士 宮下 文徳

最終頁に続く

(54) 【発明の名称】 シーン推定方法およびシーン推定装置

(57) 【特許請求の範囲】

【請求項 1】

コンピュータが行うシーン推定方法であって、
移動プラットフォームによって取得された画像を分割するための領域である分割領域のレイアウトを決定する決定ステップと、

前記移動プラットフォームの動きを取得する属性取得ステップと、
前記移動プラットフォームの動きに基づいて、前記レイアウトの中心点の位置を調整することで、前記分割領域のレイアウトを動的に変更する調整ステップと、

前記分割領域によって前記画像を分割する分割ステップと、
前記画像に含まれる画素のそれぞれを画素タイプに分類する画素タイプ分類ステップと

10

前記分割領域に含まれる複数の画素に対応する複数の画素タイプの統計を表すデータである画素特徴データを、前記分割領域ごとに生成する特徴生成ステップと、

前記各分割領域に対応する画素特徴データに基づいて、画像のシーンを分類するシーン分類ステップと、

を含む、シーン推定方法。

【請求項 2】

前記移動プラットフォームの動きは、前記画像から独立した一つ以上の要素に基づいて決定される、

請求項 1 に記載のシーン推定方法。

20

【請求項 3】

前記決定ステップでは、前記分割領域のレイアウトを、前記移動プラットフォームの動きに少なくとも基づいて決定する、

請求項 2 に記載のシーン推定方法。

【請求項 4】

前記画素特徴データは、各分割領域に含まれる画素タイプのヒストグラムである、

請求項 1 から 3 のいずれかに記載のシーン推定方法。

【請求項 5】

少なくとも二つ以上の前記分割領域が、画像内に設定された前記中心点に向かって収束するように配置される、

請求項 1 から 4 のいずれかに記載のシーン推定方法。

【請求項 6】

前記中心点は、シーンの分類の際に使用される一つまたは複数の態様を判断する期待値を設定する基準点であり、前記各分割領域の形状は、前記中心点に基づいて決定される、

請求項 5 に記載のシーン推定方法。

【請求項 7】

前記中心点は、予測されたエピ極である、

請求項 6 に記載のシーン推定方法。

【請求項 8】

前記各分割領域は、前記中心点から画像の端に向けて延出する、

請求項 5 から 7 のいずれかに記載のシーン推定方法。

【請求項 9】

道路上を移動する移動プラットフォームから、前記道路を含んだ画像データを取得する画像取得ステップをさらに含む、

請求項 1 から 8 のいずれかに記載のシーン推定方法。

【請求項 10】

前記画像に対応する前記分割領域は 5 つ以上である、

請求項 1 から 9 のいずれかに記載のシーン推定方法。

【請求項 11】

前記シーン分類ステップでは、画像によって示されたコンテンツ全体の印象を決定し、前記コンテンツ全体の印象をさらに用いて、前記画像を一つ以上のシーンカテゴリに分類する、

請求項 1 から 10 のいずれかに記載のシーン推定方法。

【請求項 12】

前記シーン分類ステップは、

各分割領域に含まれる画素特徴データの集合に基づいて、一つ以上のシーン変数に関連付いた分布を決定するステップと、

前記一つ以上のシーン変数に関連付いた分布に基づいて、分類先シーンを決定するステップと、を含む、

請求項 1 から 11 のいずれかに記載のシーン推定方法。

【請求項 13】

請求項 1 から 12 のいずれかに記載のシーン推定方法の各ステップをコンピュータに実行させるためのプログラム。

【請求項 14】

移動プラットフォームによって取得された画像を分割するための領域である分割領域のレイアウトを決定する決定手段と、

前記移動プラットフォームの動きを取得する属性取得手段と、

前記移動プラットフォームの動きに基づいて、前記レイアウトの中心点の位置を調整することで、前記分割領域のレイアウトを動的に変更する調整手段と、

前記分割領域によって前記画像を分割する分割手段と、

10

20

30

40

50

前記画像に含まれる画素のそれぞれを画素タイプに分類する画素タイプ分類手段と、
前記分割領域に含まれる複数の画素に対応する複数の画素タイプの統計を表すデータである画素特徴データを、前記分割領域ごとに生成する特徴生成手段と、
前記各分割領域に対応する画素特徴データに基づいて、画像のシーンを分類するシーン分類手段と、
を有する、シーン推定装置。

【請求項 15】

前記移動プラットフォームの動きは、前記画像から独立した一つ以上の要素に基づいて決定される、

請求項 14 に記載のシーン推定装置。

10

【請求項 16】

前記決定手段は、前記分割領域のレイアウトを、前記移動プラットフォームの動きに少なくとも基づいて決定する、

請求項 15 に記載のシーン推定装置。

【請求項 17】

前記画素特徴データは、各分割領域に含まれる画素タイプのヒストグラムである、

請求項 14 から 16 のいずれかに記載のシーン推定装置。

【請求項 18】

少なくとも二つ以上の前記分割領域が、画像内に設定された前記中心点に向かって収束するように配置される、

20

請求項 14 から 17 のいずれかに記載のシーン推定装置。

【請求項 19】

前記中心点は、シーンの分類の際に使用される一つまたは複数の態様を判断する期待値を設定する基準点であり、前記各分割領域の形状は、前記中心点に基づいて決定される、

請求項 18 に記載のシーン推定装置。

【請求項 20】

前記中心点は、予測されたエピ極である、

請求項 19 に記載のシーン推定装置。

【請求項 21】

前記各分割領域は、前記中心点から画像の端に向けて延出する、

30

請求項 18 から 20 のいずれかに記載のシーン推定装置。

【請求項 22】

道路上を移動する移動プラットフォームから、前記道路を含んだ画像データを取得する画像取得手段をさらに有する、

請求項 14 から 21 のいずれかに記載のシーン推定装置。

【請求項 23】

前記画像に対応する前記分割領域は 5 つ以上である、

請求項 14 から 22 のいずれかに記載のシーン推定装置。

【請求項 24】

前記シーン分類手段は、画像によって示されたコンテンツ全体の印象を決定し、前記コンテンツ全体の印象をさらに用いて、前記画像を一つ以上のシーンカテゴリに分類する、

40

請求項 14 から 23 のいずれかに記載のシーン推定装置。

【請求項 25】

前記シーン分類手段は、

各分割領域に含まれる画素特徴データの集合に基づいて、一つ以上のシーン変数に関連付いた分布を決定し、前記一つ以上のシーン変数に関連付いた分布に基づいて、分類先シーンを決定する、

請求項 14 から 24 のいずれかに記載のシーン推定装置。

【発明の詳細な説明】

【技術分野】

50

【 0 0 0 1 】

本発明は画像処理に関し、特に、画像からシーンを推定する技術に関する。

【 背景技術 】

【 0 0 0 2 】

(関連出願への相互参照)

本出願は、"Driver Assistance by a Lightweight Scene Classification System"と題され、2013年5月31日に出願された、米国仮特許出願第61/829,980号の優先権を主張する。該出願は、その全体が本明細書中に参照として組み込まれる。

また、本出願は、2014年2月3日に出願された、米国特許出願第14/171,677号の優先権を主張する。該出願は、その全体が本明細書中に参照として組み込まれる。

10

【 0 0 0 3 】

画像解析のために用いられる既存の技術は、例えば、オブジェクト認識方法、画素レベルでのセグメンテーション、探索窓のスキャニングなど、計算コストが高い方法をしばしば用いている。

これらの既存の技術では、オブジェクト認識において、画像をどのようにセグメンテーションするかを決定するために、画像に含まれる画素データを分析する。この処理は長い処理時間を必要とするため、ユーザに対する待ち時間や遅延をもたらす。

すなわち、モバイルコンピューティングにおいて、リアルタイムでビデオストリームを取得して分析することは非実用的であると言える。

20

【 0 0 0 4 】

また、現状、車両のリアビューカメラのような、多くの既存の車両用ビデオシステムは、取得したイメージについての分析はほとんど行っていない。これらのシステムは、様々な情報（例えば、駐車する際の車両位置）をハイライトするため、道路に図形をオーバーレイ表示することはあっても、画像が示すシーンを解釈することはない。

また、あるシステムは、画像を圧縮ないし縮小し、保存したり他の車両と共有したりするが、同様に、画像を解析したり解釈したりする機能は通常有していない。

また、他の車両用ビデオシステムは、画像中から、例えば歩道や横断歩道といったような、運転に関連した特徴物を識別する機能を有するが、画像全体が表すシーンを解析する機能は有していない。

30

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 5 】

本発明は、上記の問題点を考慮してなされたものであり、画像に含まれるシーンを低い計算コストで分類する方法を提供することを目的とする。

【 課題を解決するための手段 】

【 0 0 0 6 】

上記課題を解決するための、本発明の一形態に係る方法は、

コンピュータが行うシーン推定方法であって、画像を取得する画像取得ステップと、前記画像を分割するための領域である分割領域を決定し、前記画像を分割する分割ステップと、前記画像に含まれる画素を、一つ以上の画素レベルの特徴に基づいて、一つ以上の画素タイプに分類する画素タイプ分類ステップと、前記分割領域に含まれる画素の画素タイプに基づいて、前記分割領域ごとに、前記画素タイプの分布を表すデータである画素特徴データの集合を決定する特徴決定ステップと、前記各分割領域に含まれる画素特徴データの集合に基づいて、画像のシーンを分類するシーン分類ステップと、を含むことを特徴とする。

40

【 0 0 0 7 】

また、本発明に係るシーン推定装置は、

画像を取得する画像取得手段と、前記画像を分割するための領域である分割領域を決定し、前記画像を分割する分割手段と、前記画像に含まれる画素を、一つ以上の画素レベル

50

の特徴に基づいて、一つ以上の画素タイプに分類する画素タイプ分類手段と、前記分割領域に含まれる画素の画素タイプに基づいて、前記分割領域ごとに、前記画素タイプの分布を表すデータである画素特徴データの集合を決定する特徴決定手段と、前記各分割領域に含まれる画素特徴データの集合に基づいて、画像のシーンを分類するシーン分類手段と、を有することを特徴とする。

【0008】

また、これらの形態の一つ以上の他の実装として、方法、システム、装置、コンピュータプログラム、コンピュータ記憶装置上に符号化された方法の実行などが含まれる。

【0009】

これらの実装例は、以下の特徴のうち一つ以上を任意に含んでもよい。

10

例えば、前記分割ステップでは、前記画像から独立した一つ以上の要素に基づいて前記分割領域の形状を決定してもよい。

また、前記一つ以上の要素は、移動プラットフォームの動きを含み、前記分割ステップでは、前記各分割領域の形状を、前記移動プラットフォームの動きに少なくとも基づいて決定してもよい。

また、前記分割領域における画素タイプの分布は、各分割領域における画素タイプのヒストグラムであってもよい。

また、少なくとも二つ以上の前記分割領域が、画像内に設定された中心点に向かって収束するように配置されてもよい。

また、前記中心点は、シーンの分類の際に使用される一つまたは複数の態様を判断する期待値を設定する基準点であり、前記各分割領域の形状は、前記中心点に基づいて決定されてもよい。

20

また、前記中心点は、エッジを予測するものであってもよい。

また、前記各分割領域は、前記中心点から画像の端に向けて延出するものであってもよい。

また、前記画像取得ステップでは、道路上を移動する移動プラットフォームから、前記道路を含んだ画像データを取得してもよい。

また、前記画像に対応する前記分割領域は5つ以上であることを特徴としてもよい。

また、前記シーン分類ステップでは、画像によって示されたコンテンツ全体の印象を決定し、前記コンテンツ全体の印象をさらに用いて、前記画像を一つ以上のシーンカテゴリに分類してもよい。

30

また、前記シーン分類ステップは、各分割領域に含まれる画素特徴データの集合に基づいて、一つ以上のシーン変数に関連付いた分布を決定するステップと、前記一つ以上のシーン変数に関連付いた分布に基づいて、分類先シーンを決定するステップと、を含んでもよい。

【0010】

本明細書によって開示される発明は、特にいくつかの面において有利である。

例えば、本発明に係るシーン分類技術は、移動プラットフォームからの画像ストリームを分析することによって、複数かつ同時にシーンタイプを分類することができる。

本シーン分類は、複数の（例えば、二つ以上）範囲について、現在の景色の要旨を取得しうる。

40

本シーン分類技術はまた、場面の特徴を予測する新しいシステムと方法を含む。

以下に詳細に説明するように、分類されたシーンの出力には、ドライバーへのアドバイスやアシストにおいて、多数の有益な用途がある。

【0011】

本発明に係るシーン分類技術は、現在の視覚状態の分析に基づいて視像を既定のカテゴリ群（予測されたシーン特性など）と関連付けることにより、視像からシーンを分類するための軽量の計算手法の実行可能性を示すことができる。本技術は、一連の経験的な機械学習タスクの問題を軽減することにより、対象物認識方法の複雑性の多くを避けるという利点を提供する。

50

【 0 0 1 2 】

本明細書に記載の技術は、計算コストの高い画像処理ステップを避けるために、全体画像から情報を好都合に抽出できる、新規なシーン分類アルゴリズムを含む。

また、本発明に係る技術は、少なくとも二つのカスケード状の分類器を含むシステムであってもよい。当該分類器は、原画像を画素タイプのセットに変換する第一の画素レベル分類器と、画素タイプの統計情報を入力データとして使用し、各種シーン変数と関連付けられたラベルのセットを推定し出力する第二の分類器を含む。

例えば、第二の分類器は、分類された画素タイプを分割領域に割り当てることによって第一の分類器の出力を受信および要約し、その後、各分割領域から取得したその画素タイプコンテンツの統計情報を、確率モデルへの入力データとして使用することができる。

10

この手法によるシーン分類技術では、100万画素のオーダーから少数の別個のシーン変数を好都合に類推することができる。さらに、第一の分類器によって画像の複雑性を軽減することにより、以降のステップのデータ処理負荷が大幅に低減される。また、第一の分類器の効率により、画像のサイクルタイムのほぼ全体が判断され得る。

【 0 0 1 3 】

各分類器は、例えば、画素タイプおよびシーン特性を手作業でラベル付けした画像のセットによって学習することができる。第一の分類器は、この画像を、第二の分類器への入力となる豊富な特徴を持つ、扱いやすい少量のデータへと変換する。また、第二の分類器は、シーン変数と関連付けられたラベルの最有力の組合せを用いてシーンの特徴付けることにより、シーン変数群に対応するいくつかの従属次元でシーンを分類することができる。

20

このシステムは、ビデオシーケンスを分析する際に、ラベルのリスト（各シーン変数カテゴリに属する一ラベルなど）を出力として生成する。例えば、第二の分類器は、ラベルと関連付けられたシーン変数ノードのセットを出力ノードとして指定できる、確率モデルとして実装してもよい。また、ラベルのリストは、画像に添付されたラベルの確実性または信頼性を表す値であってもよい。

【 0 0 1 4 】

なお、上記の利点はあくまで例であり、システムは、他の多数の利点および利益を有し得ることが理解されるべきである。

【 発明の効果 】

30

【 0 0 1 5 】

本発明によると、画像に含まれるシーンを低い計算コストで分類する方法を提供することができる。

【 図面の簡単な説明 】

【 0 0 1 6 】

【 図 1 】 画像からシーンを推定するシステムを例示するブロック図である。

【 図 2 】 推定アプリケーションとシーン利用アプリケーションを備えるコンピュータを例示するブロック図である。

【 図 3 】 画像からシーンを推定する方法のフローチャート例である。

【 図 4 A 】 画像からシーンを推定する更なる方法のフローチャート例である。

40

【 図 4 B 】 画像からシーンを推定する更なる方法のフローチャート例である。

【 図 5 】 さまざまなアプリケーションにおいて分類されたシーンを利用する方法のフローチャート例である。

【 図 6 A 】 分割レイアウトを例示する図である。

【 図 6 B 】 分割レイアウトを例示する図である。

【 図 6 C 】 分割レイアウトを例示する図である。

【 図 6 D 】 分割レイアウトを例示する図である。

【 図 6 E 】 分割レイアウトを例示する図である。

【 図 6 F 】 車両データにおける中心点の配置を例示する図である。

【 図 6 G 】 車両データにおける中心点の配置を例示する図である。

50

【図 6 H】車両データにおける中心点の配置を例示する図である。

【図 6 I】車両データにおける中心点の配置を例示する図である。

【図 6 J】画素レベルの分類の例を示す図である。

【図 6 K】分割レイアウトを例示する図である。

【図 6 L】車両データにおける中心点の配置を例示する図である。

【図 6 M】車両データにおける中心点の配置を例示する図である。

【図 6 N】車両データにおける中心点の配置を例示する図である。

【図 6 O】異なる道路シーンにおける分割レイアウトの例を示す図である。

【図 6 P】異なる道路シーンにおける分割レイアウトの例を示す図である。

【図 6 Q】異なる道路シーンにおける分割レイアウトの例を示す図である。

10

【図 6 R】所定のイメージに適用できる 2 つの分割レイアウトを全体的に表す図である。

【図 6 S】所定のイメージに適用できる 2 つの分割レイアウトを全体的に表す図である。

【図 7 A】混同行列の例を示す図である。

【図 7 B】シーン分類における評価の結果を例示する図である。

【図 8 A】確率モデルを例示する図である。

【図 8 B】確率モデルを例示する図である。

【図 9】画像からシーンを推定するためのプロセスを例示する図である。

【図 10】画素レベル分類の更なる例を例示する図である。

【発明を実施するための形態】

【0017】

20

シーンを推定を行うシステムおよび方法について以下に記載する。なお、本開示は、例として示され、何らかの限定をするためのものではない。また、添付の図面においては、参照符号が同様の要素を参照するために使用されている。

【0018】

(システム概要)

図 1 は、ある実施形態において、画像に基づいてシーンを推定するシステム 100 のブロック図を例示したものである。

システム 100 は、サーバ 101、クライアント装置 115 および移動プラットフォーム 135 を含む。システム 100 の構成要素は、ネットワーク 105 によって通信可能に結合している。

30

他の実施形態において、システム 100 は、例えば、地図情報を提供するための地図サーバ、交通情報を提供するための交通情報サーバなど、図 1 に示されない他の構成要素を含んでいてもよい。

【0019】

ネットワーク 105 は、既知の有線ネットワークであっても無線ネットワークであってもよい。また、その構成は、スター型、トークンリング型やその他の構成を取ることができる。さらにまた、ネットワーク 105 は、ローカルエリアネットワーク (LAN)、ワイドエリアネットワーク (WAN) (例えばインターネット)、公衆ネットワーク、プライベートネットワーク、仮想ネットワーク、または、複数の装置が通信に用いる相互接続されたその他の任意のデータバスであってもよい。さらに別の形態では、ネットワーク 105 はピアツーピアネットワークであってもよい。ネットワーク 105 は、複数の異なる通信プロトコルでデータ送信するための通信ネットワークと接続されたり、このような通信ネットワークを含んだりしても良い。

40

ある実施形態では、ネットワーク 105 は、Bluetooth (登録商標) 通信ネットワークや携帯電話通信ネットワークを含み、SMS (ショートメッセージサービス)、MMS (マルチメディアメッセージサービス)、HTTP (ハイパーテキスト転送プロトコル)、直接データ接続、WAP、電子メールなどのデータを送受信する。

なお、図 1 では 1 つのネットワーク 105 のみが、クライアント装置 115 および移動プラットフォーム 135 に接続されているが、複数のネットワーク 105 がこれらの構成要素に接続されていてもよい。

50

【 0 0 2 0 】

一実施形態において、推定アプリケーション 1 0 9 がサーバ 1 0 1 において利用できる。推定アプリケーション 1 0 9 は、信号線 1 0 4 を介してネットワーク 1 0 5 に接続される。

サーバ 1 0 1 は、プロセッサおよびメモリを備え、ネットワークによる通信を行うことができる、ハードウェアサーバや仮想サーバ等のサーバである。一実施形態において、サーバ 1 0 1 は、クライアント装置 1 1 5 や移動プラットフォーム 1 3 5 に対してデータを送受信することができる。また、一実施形態において、サーバ 1 0 1 は、画像に基づいて分類されたシーンを利用するアプリケーションである、シーン利用アプリケーション 1 1 7 を含んでいてもよい。

10

なお、図 1 では、1 台のサーバ 1 0 1、移動プラットフォーム 1 3 5、クライアント装置 1 1 5 を例示するが、システム 1 0 0 は直接的に、またはネットワーク 1 0 5 によって相互接続される多くの構成要素を含むことができる。

【 0 0 2 1 】

一実施形態において、推定アプリケーション 1 0 9 は、クライアント装置 1 1 5 で実行されてもよい。クライアント装置 1 1 5 は、信号線 1 0 8 を介してネットワーク 1 0 5 に接続している。

一実施形態において、クライアント装置 1 1 5 は、サーバ 1 0 1 や移動プラットフォーム 1 3 5 に対してデータを送受信することができる。

クライアント装置 1 1 5 は、メモリとプロセッサを含む、情報処理とネットワーク 1 0 5 へのアクセスが可能なコンピュータであり、例えば、ラップトップコンピュータ、デスクトップコンピュータ、タブレットコンピュータ、携帯電話、個人情報端末 (P D A)、モバイル電子メール装置、および、その他の電子機器などである。

20

一実施形態において、クライアント装置 1 1 5 は、画像に基づいて分類されたシーンを利用するアプリケーションである、シーン利用アプリケーション 1 1 7 を含んでいてもよい。

ユーザ 1 2 5 は、クライアント装置 1 1 5 を利用する利用者である。

なお、図 1 では、1 台のクライアント装置 1 1 5 を例示するが、システム 1 0 0 は、複数のクライアント装置 1 1 5 を含んでいてもよい。

【 0 0 2 2 】

30

一実施形態において、推定アプリケーション 1 0 9 は、シーン利用アプリケーション 1 1 7 に対して、画像に対応するシーンと、シーンの分類結果を表すデータを送信することができる。また、シーン利用アプリケーション 1 1 7 は、分類されたシーンを表すユーザインタフェースを提供するためのグラフィックを生成することができる。また、シーン利用アプリケーション 1 1 7 は、生成したグラフィックを、クライアント装置 1 1 5 に接続された表示装置 (図示せず) を用いて提示してもよい。

【 0 0 2 3 】

一実施形態において、推定アプリケーション 1 0 9 は、移動プラットフォーム 1 3 5 上で実行されてもよい。移動プラットフォーム 1 3 5 は、信号線 1 3 4 を通してネットワーク 1 0 5 に接続している。

40

一実施形態において、移動プラットフォーム 1 3 5 は、サーバ 1 0 1 やクライアント装置 1 1 5 に対してデータを送受信することができる。

移動プラットフォーム 1 3 5 は、メモリとプロセッサを有するコンピュータを含んでもよい。一実施形態において、移動プラットフォーム 1 3 5 は、道路上を移動可能なコンピュータである。例えば、移動プラットフォーム 1 3 5 は、車両、自動車、バス、生体に埋め込まれた装置、非一時的 (non-transitory) なコンピュータ機器 (例えば、プロセッサ、メモリ、あるいは非一時的なコンピュータ機器の組み合わせ) を有する他のモバイルシステムなどであってもよい。

ユーザ 1 2 5 は、移動プラットフォーム 1 3 5 の操作者であり、移動プラットフォーム 1 3 5 を利用する人間である。例えば、ユーザ 1 2 5 は、車両を運転するドライバーであ

50

ってもよい。

ー実施形態において、移動プラットフォーム 135 は、カメラ 107 と、画像に基づいて分類されたシーンを利用するアプリケーションである、シーン利用アプリケーション 117 を含んでいてもよい。

なお、図 1 では、1 台の移動プラットフォーム 135 を例示するが、システム 100 は、複数の移動プラットフォーム 135 を含んでいてもよい。

【0024】

推定アプリケーション 109 は、画像からシーンを推定するためのロジックを含んだアプリケーションである。

いくつかの実施例において、推定アプリケーション 109 は、FPGA (Field Programmable Gate Array) や ASIC (Application Specific Integrated Circuit) などのハードウェア、コンピュータのプロセッサで実行可能なソフトウェア、または、それらの組み合わせを用いて実装される。

ある実施形態では、推定アプリケーション 109 は、装置とサーバの両方に分散して格納されてもよいし、装置とサーバのいずれかのみに格納されてもよい。

推定アプリケーション 109 については、図 2 ~ 図 4 B を参照しながら後ほど詳しく説明する。

【0025】

カメラ 107 は、画像 (ビデオ画像および静止画像を含む) を記録する光学的装置である。カメラ 107 は、任意のフレームレートを使用して、ビデオストリームのフレームを取り込むことができ、また、任意の方法を用いて撮影した画像を符号化することができる。

ー実施形態において、カメラ 107 は、周辺環境を画像で記録することができる。

例えば、移動プラットフォームの場合、カメラ 107 は、道路、空、山、他の車両、オブジェクト (例えば、道路脇の構造物、建物、樹木など) を含んだ、移動プラットフォーム 135 の周囲の環境を取り込むことができる。

ー実施形態において、カメラ 107 は、移動プラットフォーム 135 の前面に、前方を向いてマウントされていてもよい。別の実施形態において、カメラ 107 は、移動プラットフォーム 135 の他の部分にマウントされていてもよい。

例えば、クライアント装置の場合、カメラ 107 は、フロントあるいはリアカメラであってもよく、セットトップカメラ等であってもよい。

カメラ 107 は、取得した画像を含む画像データを、推定アプリケーション 109 に送信することができる。

例えば、カメラ 107 は、道路上を移動している移動プラットフォーム 135 (例えば車両) から、道路を含んだシーンを表す画像データを取得することができる。

また、カメラ 107 は、画像データ (例えばリアルタイムビデオストリーム) を、本明細書の他の箇所でも更に説明するような処理およびクラス分類を行う推定アプリケーション 109 に送信することができる。

【0026】

シーン利用アプリケーション 117 は、様々なアプリケーションにおいてシーンを利用するためのロジックを含む。

例えば、シーン利用アプリケーション 117 は、推定アプリケーション 109 によって推定されたシーンを活用して、ユーザに高品質なナビゲーション機能を提供する地図アプリケーションであってもよい。

いくつかの実施形態において、シーン利用アプリケーション 117 は、FPGA (Field Programmable Gate Array) や ASIC (Application Specific Integrated Circuit) などのハードウェア、コンピュータのプロセッサで実行可能なソフトウェア、または、それらの組み合わせを用いて実装される。

ある実施形態では、シーン利用アプリケーション 117 は、装置とサーバの両方に分散して格納されてもよいし、装置とサーバのいずれかのみに格納されてもよい。

シーン利用アプリケーション 117 については、図 2 および図 5 を参照しながら後ほど詳しく説明する。

【0027】

位置決定システム 120 は、クライアント装置 115 や移動プラットフォーム 135 の地理的な位置を決定するためのシステムである。

一実施形態において、位置決定システム 120 は、信号を受信するように構成された電子装置（例えばクライアント装置 115 や移動プラットフォーム 135）に、位置信号を供給する。

位置決定システム 120 は、信号線 122 で示したように無線送信機と受信機を介して、または、信号線 124 で示したようにネットワーク 105 を介して、あるいは他の適当な伝達手段を介して位置信号を供給してもよい。

10

クライアント装置 115 および移動プラットフォーム 135 は、位置決定システム 120 から位置信号を受信する受信機（例えば GPS 受信機、ネットワークインタフェース等）を含んでいてもよい。また、クライアント装置 115 および移動プラットフォーム 135 は、位置信号を処理することで、自装置（自端末）の地理的な位置を決定してもよい。

一実施形態において、位置決定システム 120 は、GPS、ディファレンシャル GPS（DGPS）、補助 GPS（A-GPS）、多辺測量あるいは三角測量エンジンを含んだ電気通信システム、ネットワーク 105 におけるクライアント装置 115 や移動プラットフォーム 135 の IP アドレスを取得し、他の情報ソース（例えばインターネットプロバイダのデータベースやインターネットレジストリ）と IP アドレスをクロスリファレンスすることで、当該クライアント装置 115 や移動プラットフォーム 135 の大まかな位置を推定するジオロケーションシステム、あるいは他の位置情報システムや装置などを含んでいてもよい。

20

【0028】

様々な実施形態において、"hiball"のようなコンステレーションシステム、磁気追跡システム、光学追跡システム、慣性追跡システムなど、どのような位置追跡システムが用いられてもよいことを理解すべきである。

また、位置決定システム 120 は、単体の要素として示されているが、サーバ 101、ネットワーク 105、クライアント装置 115、移動プラットフォーム 135 を含むシステムの他の要素と組み合わせてもよい。

30

推定アプリケーション 109 やシーン利用アプリケーション 117 は、位置決定システム 120 と接続され、位置信号や位置データを受信するように構成されてもよい。例えば、API を介してこれらの情報にアクセスできるようにしてもよい。

【0029】

（推定アプリケーション 109 の例）

図 2 を参照して、推定アプリケーション 109 とシーン利用アプリケーション 117 を有するコンピュータの例について、詳細に説明する。

図 2 は、一実施形態における、推定アプリケーション 109、プロセッサ 235、メモリ 237、通信部 241、入出力部 243、記憶装置 245、シーン利用アプリケーション 117、およびカメラ 107 を含むコンピュータ 200 のブロック図である。コンピュータ 200 の構成要素は、バス 220 によって通信可能に接続される。

40

バス 220 は、コンピュータが有する構成要素間、または、コンピュータ間でデータを送受信することができれば、どのような種類の既知の通信バスであってもよい。

一実施形態において、コンピュータ 200 は、サーバ 101、クライアント装置 115、または移動プラットフォーム 135 のうちのいずれかである。

【0030】

プロセッサ 235 は、算術論理ユニット、マイクロプロセッサ、汎用コントローラ、または、計算を実行して表示装置に表示信号を提供する他のプロセッサアレイのような、一つ以上の処理ユニットを含む。

プロセッサ 235 は、データ信号を処理し、そのアーキテクチャは、CISC（Comple

50

x Instruction Set Computer)、RISC(Reduced Instruction Set Computer)、これら両方の命令セットの組合せとして実装されたアーキテクチャなど様々なアーキテクチャでありうる。なお、図2には一つのプロセッサ235だけが示されているが、複数のプロセッサ235が含まれていてもよい。上記以外のプロセッサ、オペレーティングシステム、センサ、表示装置、あるいは物理的構成も採用可能である。

【0031】

メモリ237は、プロセッサ235が実行可能な命令やデータを格納する手段であり、一つ以上のコンピュータによって読み取り可能な非一時的記憶媒体を含む。メモリ237に格納される命令やデータは、本明細書に示す技術を実行するためのコードを含んでもよい。一実施形態において、メモリ237は、推定アプリケーション109やシーン利用アプリケーション117を格納する。

10

メモリ237は、DRAM、SRAM、組み込み型メモリ、フラッシュメモリやその他の既存のメモリ装置を含んでもよい。ある実施形態においては、メモリ237は、ハードディスクドライブ、フロッピーディスク(登録商標)ドライブ、CD-ROM装置、DVD-ROM装置、DVD-RAM装置、DVD-RW装置、フラッシュメモリ装置や情報記憶の分野において既知のその他の大容量記憶装置などの、非一時的(non-volatile)メモリや永久記憶装置を含んでもよい。

【0032】

通信部241は、サーバ101、移動プラットフォーム135、クライアント装置115のうちの一つ以上と通信を行う手段である。その通信先は、推定アプリケーション109が格納されている場所によって異なる。

20

通信部241は、ネットワーク105を形成する他の要素と通信を行うため、ネットワーク105に有線または無線等によって接続される。

いくつかの実施形態において、通信部241は、ネットワーク105や他の通信チャネルへの直接的な物理的接続のためのポートやネットワークインタフェースを含む。例えば、通信部241は、USB、SD、CAT-5など、クライアント装置115と有線通信するためのポートを含む。

ある実施形態では、通信部241は、IEEE 802.11、IEEE 802.16、Bluetooth(登録商標)、DSRC(Dedicated Short-Range Communication)または他の適当な無線通信方式を含む一つまたは複数の無線通信方式を用いて、クライアント装置115や他の通信チャネルとデータを交換するための無線送受信部を含む。

30

一実施形態において、無線送受信部は、GPS信号を含む位置データや、コンピュータ200の位置を表す他の位置データを受信するように構成された受信機を含んでもよい。

なお、図2では、単体の通信部241を例示したが、通信部241は、実際には、一つ以上の異なる通信装置であってもよいし、統合通信装置であってもよい。

【0033】

ある実施形態では、通信部241は、SMS(ショートメッセージサービス)、MMS(マルチメディアメッセージサービス)、HTTP(ハイパーテキスト転送プロトコル)、直接データ接続、WAP、電子メールやその他の適切な電子通信方式により、携帯電話通信網上でデータを送受信するためのセルラ通信トランシーバを含む。ある実施形態では、通信部241は、有線ポートと無線送受信機を含む。通信部241は、TCP/IP、HTTP、HTTPS、SMTPなどの標準的なネットワークプロトコルを用いてファイルやメディアオブジェクトを配布するために、ネットワーク105への従来型の接続を提供する。

40

【0034】

記憶装置245は、本明細書に記載される構造、処理、または機能を提供するためのデータを格納する非一時的(non-transitory)な記憶媒体である。一実施形態において、記憶装置245は、データを格納し、データへのアクセスを提供するデータベース管理システム(DBMS)を含んでもよい。

また、記憶装置245は、半導体素子メモリ(例:フラッシュメモリやRAM等)、ハ

50

ードディスクドライブ、磁気テープ装置、フロッピーディスク（登録商標）ドライブ、CD-ROM装置、DVD-ROM装置、DVD-RAM装置、DVD-RW装置や、情報記憶の分野において既知のその他の大容量記憶装置を含んでも良い。

【0035】

記憶装置245は、本明細書で後ほど詳細に説明するように、推定アプリケーション109が、カメラ107が撮像した画像を一つ以上の分割領域に分割するために用いる、所定の分割レイアウトのセットを格納することができる。

また、記憶装置245は、カメラ107によって取得された画像（動画または静止画等）、画素タイプ、分割領域における画素タイプの分布（例えば画素タイプのヒストグラム）、分割領域に対応する画素特徴データ（例えば分割領域に対応する尤度ベクトル）、画像からシーンを推測するために用いられる確率モデル247、確率モデルの学習における学習データセット、シーン変数、一つ以上のシーン変数に関連付いた分布、および、シーン分類結果（例えばシーンラベルのセット）等を格納することができる。

これらの、記憶装置245に格納されるデータについては、後ほど詳しく説明する。

一実施形態において、記憶装置245は、本明細書に記載される構造、処理、または機能を提供するための他のデータを格納してもよい。

【0036】

移動プラットフォーム135が、コンピュータ200を含むという実施形態においては、コンピュータが、外部コンポーネント250と接続される他のコンピュータを含んでもよい。外部コンポーネント250は、プラットフォームバス（例：CAN等）、一つ以上のセンサ（例えば、ECU、ECM、PCM等の制御ユニット、車両センサ、音響センサ、化学的センサ、生体センサ、位置センサ（例：GPS、コンパス、加速度計、ジャイロ等）、スイッチ、コントローラ、カメラ等）、エンジン、駆動手段、サスペンション手段、計装、温度制御、または、他の電氣的、機械的、構造的コンポーネント等であるが、これに限られない。

【0037】

図2に示したように、推定アプリケーション109は、画像プロセッサ202、分割モジュール204、画素分類モジュール206、エビデンス生成モジュール208、学習モジュール210、シーン分類モジュール212を含むことができる。各構成要素は、バス220を介して、プロセッサ235やコンピュータ200の他の構成要素と通信可能に接続される。

【0038】

画像プロセッサ202は、分類対象の画像のデータ量を減らすため、画像の前処理を行うルーチンを含むソフトウェアである。一実施形態において、画像プロセッサ202は、プロセッサ235で実行可能な命令の集合を含む。また、一実施形態において、画像プロセッサ202は、コンピュータ200のメモリ237に格納され、プロセッサ235によってアクセスされ、実行されてもよい。

画像プロセッサ202は、プロセッサ235およびコンピュータ200の他の構成要素と協働して通信するように最適化されてもよい。

【0039】

画像プロセッサ202は、カメラ107、記憶装置245、推定アプリケーション109のいずれかから、画像を取得することができる。

画像プロセッサ202は、画素分類モジュール206や、推定アプリケーション109内の他の構成要素によって処理されるデータの量を減らすために、画像を前処理する手段である。これにより、画像の分類における速度や効率を改善することができる。

いくつかの実施形態において、画像プロセッサ202は、画像を粗いパッチ配列にダウンコンバートすることができる。例えば、画像プロセッサ202は、画像内の画素を変換して、低解像度のパッチ配列を生成する。これにより、カメラ107によってキャプチャされた生の画像内のすべての画素を分類する必要性をなくすることができる。

本実施形態では、例えば、画像プロセッサ202が、所定のグリッドに対応する画素を

10

20

30

40

50

選択することにより、ダウンサンプリング画像を生成する。このダウンサンプリング画像は、図10に示すとおり、近隣画素が互いに類似する傾向があることから、場合によっては、原画像に対してカラー漫画調になることもある。

また、元の画素セットを低解像度の画素アレイ（パッチとも呼ばれる）に変換するため、他のダウンサンプリング、平均化、または圧縮技術を利用してもよい。

画像プロセッサ202は、分割モジュール204が画像を分割する前に、あるいは後に、画像を処理することができる。また、画像プロセッサ202は、処理後の画像を、分割モジュール204や画素分類モジュール206に送信してもよい。

【0040】

分割モジュール204は、画像を分割するレイアウトを決定するためのルーチンを含むソフトウェアである。

一実施形態において、分割モジュール204は、プロセッサ235で実行可能な命令の集合を含む。また、一実施形態において、分割モジュール204は、コンピュータ200のメモリ237に格納され、プロセッサ235によってアクセスされ、実行されてもよい。

分割モジュール204は、プロセッサ235およびコンピュータ200の他の構成要素と協働して通信するように最適化されてもよい。

【0041】

分割レイアウトを決定するために、分割モジュール204は、移動プラットフォーム135のカメラ107によって撮影された画像が示すであろうシーンを、画像を参照せずに予測し、当該予測に基づいて、対応する分割レイアウトを選択する。

シーンは、走行している移動プラットフォーム135に近づいてくる道路環境が反映される。一実施形態において、カメラ107は、移動プラットフォーム135に取り付けることができ、移動プラットフォーム135の動きに従って前方のシーンを取り込むように前を向いて取り付けられる。

【0042】

分割モジュール204は、画像データとは無関係に、分割レイアウトを決定することができる。例えば、分割モジュール204は、画像データを解析することなく、画像によって示されるシーンに対応する分割レイアウトを特定することができる。従って、カメラ107は、分割レイアウトが決定される前、後、決定中のいずれのタイミングであっても、画像を取得することができる。

【0043】

分割レイアウトは、一つ以上の分割領域を含む。この領域は、所定のタイプのシーンにおける、視覚的に区別可能な、または別個の要素に対応した領域として構成される。いくつかの実施形態では、シーンが、ルート（道路など）、物理的な周辺環境、運転状況（道路など）、動的環境（天候など）、障害物などの要素を含み得るが、これらに限定されない。いくつかの実施形態において、分割モジュール204によって予測的に判断された分割レイアウトは、シーンに含まれる主要素に概ね対応する形状を有する分割領域を含んでいてもよい。

【0044】

分割レイアウトは、事前に定義されたものを使用してもよいし、処理を行いながら定義してもよい。また、実行時に分割モジュール204によって調整されてもよい。また、これらの組み合わせを用いてもよい。

分割レイアウトは、分割レイアウトの向き、分割レイアウトの寸法、分割レイアウトに含まれる分割領域の数、対応する分割領域のセットの数、分割領域のジオメトリ（例えば分割領域のサイズと形状、分割領域の向きなど）等によって定義される。

例えば、サイズ、形状、分割レイアウトにおける各分割領域の向きは、本明細書の他の箇所の説明するように、デフォルト値として事前に定義され、分割モジュール204によって、一つ以上のコンテキスト属性に基づいて実行時に調整されてもよい。

【0045】

分割領域は、様々な形状やサイズをとることができる。例えば、分割領域は、くさび形、三角形、円形、扇形、または任意の他の幾何学的形状であってもよい。

また、分割レイアウトに含まれる分割領域は、互いに均一、不均一、類似、非類似、対象、非対称などであってもよいし、これらの組み合わせであってもよい。

各分割領域は、所定の分割レイアウトの中で同心状に配置されてもよい。また、分割領域として、異なる形状の組み合わせが用いられてもよいし、同一または類似の形状が用いられてもよい。また、分割領域の一部または全部の寸法は、調整されていてもよいし、調整されていなくてもよい。また、各分割領域は、互いに異なる寸法であってもよいし、分割領域の一部または全部の寸法が同じであってもよい。

【0046】

分割レイアウトの例を図6A～図6Eに示す。分割レイアウト600、605、610、620、625はそれぞれ例示であるが、これに限定されない。

図示した通り、これらのレイアウトに含まれる分割領域は、三角形、台形、半円形であってもよいし、他の形状であってもよい。

各分割領域は線で囲まれ、中心点626で交差しないし収束する。

各分割領域は可変サイズであってもよく、直線以外で囲まれてもよい。また、図6Dのように、中心点626を囲うように配置されてもよい。

【0047】

一実施形態において、分割モジュール204は、一つ以上のコンテキスト属性に基づいて分割レイアウトを決定する。コンテキスト属性とは、移動プラットフォーム135の状況や、その移動環境を表すデータである。コンテキスト属性は、コンテキストデータに基づいて推定、または決定されてもよい。

一実施形態において、コンテキストデータは、例えば、対応する画像が取得された（または前後の）時刻に対応する、移動プラットフォーム135の移動に関するデータ、車道と移動プラットフォーム135との地理的な位置関係を表す位置データ、地理的な位置の景観を表す周辺環境データ、移動プラットフォーム135の地理的な位置に対応する運転状況データ、移動プラットフォーム135の地理的な位置に対応するイベントデータ、現在の日時などのうちの一つ以上である。

コンテキスト属性は、カメラ107が取得した画像とは独立したデータである。

【0048】

例えば、図6Oに示したシーンのように、推定アプリケーション109によって分類されるシーンが、空が上にあり、道路が下にあり、特徴物が左側と右側にある、屋外の道路シーンであると仮定する。

この場合、分割モジュール204は、一つ以上のコンテキスト属性に基づいて、画像に対応するシーンが農村における道路シーンであると推定し、画像を分割領域のセットによって分割するための分割レイアウトを選択する。

例えば、図6Oに示したように、分割レイアウト675は、12個のくさび型の形状を持つ分割領域を含んでおり、各分割領域は、中心点（収束点）から伸びる線によって囲われている。この例に示されるように、道路シーン内のオブジェクトは、これらを囲むくさび型の分割領域に大略沿う傾向がある。例えば、地平線は一つ以上の分割領域に収まるかもしれないし、路面は一つ以上の分割領域に収まるかもしれない。

このような方法は、推定アプリケーション109が、CPUに負荷をかけ、待ち時間を発生させる画像の分析を行うことなく、画像に関連したパーツ（すなわち分割領域）に分割することができるため、有利である。

また、分類の品質を向上させることができ、マルチコアまたは並列プロセッサであるプロセッサ235のより効率的な使用を可能にすることができる。

図6Pおよび図6Qは、シーンを表した、分割レイアウト680および685を、対応する画像にオーバーレイ表示させた、さらなる例である。

【0049】

分割モジュール204は、外部コンポーネント250、記憶装置245のようなデータ

10

20

30

40

50

ストレージ、または他の適切なコンポーネントの一つ以上から、移動プラットフォーム 135 の移動に関するデータを受信してもよい。

移動プラットフォーム 135 の移動に関するデータとは、例えば、加速度、速度、ブレーキ力、加わった力、運転プロファイル、温度、湿度、振動、圧力、音量、燃費、日時、端末から外部にあるオブジェクト（例えば、道路脇の構造物、信号機、他の車両、歩行者、自転車等）までの距離、システム安全上のパラメータ（例えばエアバッグ、ABS 等）、走行方向、燃料残量、バッテリーレベル、または他の物理的または動的な状況を表すデータなどである。

【0050】

分割モジュール 204 は、通信部 241、一つ以上のセンサ 250、記憶装置 245 のようなデータストレージ、ナビゲーションユニット（例えば GPS）を含むシーン利用アプリケーション 117、ネットワーク 105 経由で（例えば API やブッシュ、プル等によって）情報を提供する地図・交通・天気サーバ、または他の適切なコンポーネントの一つ以上から、位置データや運転状況データを取得してもよい。

10

位置データは、移動プラットフォーム 135 の位置（例えば GPS 座標）、移動履歴、現在の日時を表すタイムスタンプ、移動プラットフォーム 135 の速度を表すデータを含んでいてもよい。

運転状況データは、移動プラットフォーム 135 の位置に関連付いた交通情報やそのアップデート、事故情報、予測遅延時間などを含んでいてもよい。

また、分割モジュール 204 は、コンピュータ 200 のオペレーティングシステムや、ネットワーク 105 に接続された時計サーバ（不図示）、あるいは他の類似するシステム 100 内の構成要素から受信したタイムスタンプに基づいて、現在の日時を決定してもよい。

20

【0051】

分割モジュール 204 は、不図示のイベントエンジンや、記憶装置 245 のようなデータストレージ、または他の適切なコンポーネントから、システム 100（例えば、マッピング、検索、あるいは情報サーバ）内のコンピュータ 200 によって処理可能な周辺環境データを受信してもよい。

周辺環境データとは、地理的に異なる場所に関連付いた土地の特徴（例えば、山、丘、水辺、草木、建物、構造物など）を表すデータである。

30

一実施形態において、周辺環境データは、指定された場所の特徴を示す地図（例えば衛星地図）から、推定アプリケーション 109 によって抽出されてもよい。

【0052】

分割モジュール 204 は、システム 100 内のコンピュータによって処理可能な不図示のイベントエンジンや、記憶装置 245 のようなデータストレージ、または他の適切なコンポーネントからイベントデータを受信してもよい。

例えば、カレンダーリング、スケジューリング、ソーシャルネットワーキングサービス、電子メール、現在のイベント、エンドユーザに対する他のソフトウェアサービス等を提供する、不図示のサーバシステムがネットワーク 105 に接続されており、分割モジュール 204 が、当該システムからネットワーク 105 経由でイベントデータを取得するようにしてもよい。

40

イベントデータは、現在の場所、または車両のルート（例えば位置データに基づいて決定されたルート）に沿って接近中の場所にて発生するイベント、発生中のイベント、または発生するであろうイベントを表すものであってもよい。

【0053】

下記は、一つ以上のコンテキスト属性に基づいて、分割レイアウトを決定する非限定的な追加例である。

コンテキスト属性は、「平日のラッシュ帯」のような時間帯を含んでいてもよく、分割モジュール 204 は、渋滞に対応する分割レイアウト（例えば、隣接車線、出口、合流車線、路肩、バリア、信号機などに対応するレイアウト）を決定するようにしてもよい。

50

また、コンテキスト属性は、交通量を含んでいてもよく、分割モジュール204は、交通量に対応した分割領域のセットを含む分割レイアウトを決定してもよい。

例えば、画像に含まれる交通量が少ない場合、分割レイアウトは、交通量が多い場合と比較して簡素なものであり、より少ない分割領域を含むものであってもよい。なぜならば、このような場合、交通量が多い場合と比較して、道路におけるオブジェクト、状況、関心エリア等が少ないためである。

また、コンテキスト属性は、（例えば車両のGPS装置によって受信される）事故情報を含んでいてもよく、分割モジュール204は、当該事故情報の一つ以上の属性に適合する分割レイアウト（例えば路肩で起きた事故に対応する領域を一つ以上含むようなレイアウト）を選択してもよい。

10

【0054】

また、コンテキスト属性は、運転中の車両の車種を表すデータであってもよく、分割モジュール204は、当該車種（例えばサイズや積載量等）に適した分割レイアウトを選択してもよい。

また、コンテキスト属性は、車両が走行している道路の種類を表すものであってもよく、分割モジュール204は、当該道路の種類（例えば、二車線の高速道路、四車線の高速道路、橋梁、トンネル、都市の通り、一方通行、未舗装道路など）に対応する分割レイアウト（例えば一致する領域を持つもの）を選択してもよい。

また、コンテキスト属性は、都市、田園地域、郊外地域、商業地域など、車両が走行しているエリアを含んでもよく、分割モジュール204は、当該エリアの共通属性（例えば地平線、平原、太陽、建物、歩道、駐車区画など）に適した領域を含む分割レイアウトを選択してもよい。

20

また、コンテキスト属性は、運転者の年齢や性別など、運転者の属性を含んでもよく、分割モジュール204は、若い運転者用の分割レイアウト、老年運転者用の分割レイアウトなど、当該運転者の属性に対応する分割レイアウトを選択してもよい。

また、コンテキスト属性は、場所（例えば、多くの車両と人が想定されるコンサート会場）にて発生するイベントタイプを含んでいてもよく、分割モジュール204は、当該シーンの種類に対応する分割レイアウトを選択してもよい。

【0055】

また、コンテキスト属性は、他の車両で利用された分割レイアウトを表すデータを含んでいてもよく、分割モジュール204は、対応する分割レイアウトを選択してもよい。

30

当該他の車両は、以前、対応する道路に沿って走行したことがあり、当該車両の学習モジュール210は、シーンを分類するための最も効果的な分割レイアウトを学習していたものであってもよい。

これは、分割レイアウト、特に、これらの車両の学習モジュール210によって学習され、改善された分割レイアウトを、システム100内の車両に共有させることができるという点において有利である。

さらなる例として、システム100内の移動プラットフォーム135に含まれる推定アプリケーション109は、分割レイアウトを互いに送信することで、互いに共有する構成であってもよい。送信は、例えば、直接行ってもよいし、サーバ101と、当該サーバで

40

使用可能な推定アプリケーション109のインスタンスを経由して行ってもよい。

例えば、通信可能な車両が、他の車両および、当該車両に含まれる推定アプリケーション109のインスタンスによって道路画像を分類するための中央サーバ（例えば、サーバ101）と、好ましい分割レイアウトデータと使用基準を共有してもよい。

一実施形態において、サーバ101は、移動プラットフォーム135に含まれる推定アプリケーション109のインスタンスによって決定された、より好ましく、新しく、改善されるように連続的にアップデートされる分割レイアウトを保持する、動的な集中リポジトリを含んでもよい。

また、移動プラットフォーム135の推定アプリケーション109のインスタンスは、集中リポジトリに対して、より好ましく、新しく、改善された分割レイアウトを取得する

50

ために問い合わせ、同期してもよく、分割モジュール 204 は、分割レイアウトを決定する際にこれらを利用してもよい。

【0056】

このように、分割モジュール 204 は、分割レイアウトを決定するために任意のコンテキスト属性を利用することができる。

これは、任意のタイプの道路シーンについての適切な分割レイアウトを、当該画像に含まれる画素を解析することなく決定できるため、推定アプリケーション 109 の処理速度と効率を向上できるという点において有利である。

また、道路シーンを一つのコンテキスト属性から識別することが困難である場合であっても、分割モジュール 204 は、道路シーンに対応するより正確な分割レイアウトを推定することができる。

10

なお、コンテキスト属性は走行中に変化するため、一実施形態において、選択された分割レイアウトも変更され、あるいは異なる分割レイアウトに選択されなおしてもよい。

例えば、分割レイアウトの決定（再決定）は、タイマによってトリガされてもよいし、カメラ 107 のフレームレートと同期していてもよいし、他の規則的、または不規則的なトリガイイベントなどに基づいて行われてもよい。

【0057】

一実施形態において、分割レイアウトの集合は予め定義されたもの（例えば、機械学習や手動設定により、またはこれらの組み合わせにより）であってもよく、当該予め定義された分割レイアウトを表すレイアウトデータは、メモリ 237、記憶装置 245、システム 100 における他の記憶装置などのストレージデバイスに格納されていてもよい。

20

一実施形態において、本明細書の他の箇所で説明したように、分割モジュール 204 は、一つ以上のコンテキスト属性に基づいて、記憶装置から分割レイアウトを選択することで、画像によって表されたシーンに対応する分割レイアウトを少なくとも部分的に決定できる。また、いくつかのケースにおいては、一つ以上のコンテキスト属性に基づいて分割レイアウトを変更することができる。

一実施形態において、分割レイアウトを構成するデータ（例えば変数、データ構造、オブジェクト等）は、当該分割レイアウトを構成する分割領域の形と数の特徴づけることができる。

【0058】

30

また、学習モジュール 210 は、分割モジュール 204 による分割レイアウトの選択を最適化するために、対応するコンテキスト属性（例えば車両の方向や位置データ）を持つテスト用画像のセットを用いて、特定のシーンタイプ（例えば道路環境）ごとに分割モジュール 204 がどのような分割レイアウトを選択すべきかについて学習を行うようにしてもよい。

これは、画素分類モジュール 206 やシーン分類モジュール 212 によって決定される所望の分類パラメータを最大化できるという点において有利である。

いくつかのケースでは、学習のために、分割レイアウトを、少なくとも部分的に手動で構成し、テスト画像に関連付けてもよい。

学習結果を反映した学習データは、記憶装置 245 内の学習モジュール 210 によって保持されてもよく、分割モジュール 204 は、分割レイアウトを決定する際に、当該決定を最適化するために、当該データに対して照会を行ってもよい。

40

【0059】

分割レイアウトは、一つ以上の分割領域のセットを含んでいてもよい。

また、各セットに含まれる分割領域は、当該レイアウト内のエリアにある特定の点に収束してもよい。

当該特定の点は、分割領域のセットにおける収束点または中心点とも呼ばれる。

最終的に、分割レイアウトが画像に適用されるとき、分割レイアウトに含まれる中心点は、対応する画像の画像領域の中のどこに位置してもよい。

中心点は、シーンの分類時に使用された一つ以上の態様を判断する期待値を設定できる

50

基準点となり得る。期待値とは、分析用の画像を分割するのに最適な方法に関する予想値や推定値のことである。場合によっては、特定の対象や関心領域が画像中に現れると予測される場合があり、このような場合、分割モジュール204は、当該画像に適用された当該画像領域に対応する箇所に中心点を配置することができる。

例えば、車両が交差点に接近していると判定された場合、交通信号機が表示される場所の近傍に中心点を置く分割レイアウトを選択または適応させることができる。この例では、車両が交差点に接近しているため、「交通信号機を含むシーンが現れる」という予測を行うことができる。

別の例では、分割モジュール204は、画像におけるいくつかのシーンの特徴が、例えば大きく開けた空のように情報が少ないものであり、一つの領域に空が大きく含まれるように中心点を調整可能（例えば分割レイアウトを適応させたり、対応する分割レイアウトを選択したり、あるいはその組み合わせにより）であることを、例えば位置データに基づいて知っていてもよい。

10

【0060】

いくつかの実施形態において、中心点および当該中心点に関連付けられた分割領域は、重要な分類可能な運転状況、道路上の障害物、道路の特性、周辺環境など、重要なデータ分類目標と関連付けられ得た、予想画像領域に対応するものであってもよい。一実施形態において、中心点は、画像の中心、すなわち、画像内の各種要素が収束するように見える予測された場所（視覚的重心）、予測消失点あるいはエピ極、その画像内の各種特徴の予測連結点あるいは収束点などに位置するように定義または適合されてもよい。エピ極は、画像内の他の点の起点になり得る点である。

20

【0061】

図6Rおよび図6Sは、一つの画像（例えばビデオ画像フレーム）に対する二種類の分割レイアウトの決定法を全体的に示した図である。

分割レイアウト690は、画像右側のターゲットを分類するために適応されたレイアウトであり、分割レイアウト695は、画像左側のターゲットを分類するために適応されたレイアウトである。（分類は、例えば画素分類モジュール206によって行われる）

分割モジュール204は、一つの画像（例えばビデオフレーム）に対して、各々が異なる中心点を有する複数の分割レイアウトをオーバーレイしてもよい。

これは、画素分類モジュール206が、一つの画像に対して複数の画像分類セット（例えば左側に重点を置いたもの、右側に重点を置いたもの）を確立してもよいことを意味する。

30

【0062】

一実施形態において、中心点は、分割レイアウト内の対応する分割領域の組の中心を表すことができる。例えば、分割モジュール204によって決定された分割レイアウトが、対応する画像に適用されたとき、分割領域は、図6Oおよび6Qに示したように、中心点から画像の端辺に対応する辺まで外方へ延びる形状となる。

図示したように、中心点は、車道の消失点に対応するように構成されてもよい。また、分割領域は、当該道路の消失点から放射状に二つ以上の分割領域に画像を分割するものであってもよい。

40

【0063】

予め定義された分割レイアウトは、当該分割レイアウトに対応するデフォルトの中心点の位置を含んでいてもよく、当該中心点は、実行時に、一つ以上の学習データやコンテキストデータに基づいて、分割モジュール204によって変更されてもよい。

いくつかの実施形態において、分割モジュール204は、移動プラットフォーム135の方位、運動、動き（例えば加速、速度、コーナリング、減速、位置の変動、上昇、下降、方向等）、移動プラットフォームの位置（例えば地理的なロケーション、経路予想、範囲データ等）などの一つ以上のコンテキスト属性に基づいて、中心点の位置を調整してもよい。

例えば、中心点に関連付いた二つ以上のパーティションが、車両の動きにあわせて移動

50

してもよい。

【 0 0 6 4 】

図 6 A および図 6 F ~ 図 6 I は、移動プラットフォーム 1 3 5 の動的な動きや場所といったような、一つ以上のコンテキスト属性に基づいて、分割モジュール 2 0 4 が、異なる時点における車両コンテキストを用いて、中心点 6 3 1 を動的に調整する例を説明する図である。

ここで、一つ以上のコンテキスト属性から決定される分割レイアウトが、通常の状態において図 6 A のようなものであったとする。

ここで、右方向を向いた場合、分割レイアウト 2 0 4 が有する分割領域は、車両が右を向いたことを示す一つ以上のコンテキスト属性（例えばセンサデータ、地図データ等）に

10

基づいて、図 6 F のように、中心点 6 3 1 を左方向に移動させることで形成されなおされてもよい。

また、左方向を向いた場合、分割レイアウト 2 0 4 が有する分割領域は、車両が左を向いたことを示す一つ以上のコンテキスト属性（例えばセンサデータ、地図データ等）に基づいて、図 6 G のように、中心点 6 3 1 を右方向に移動させることで形成されなおしてもよい。

また、坂を上りながら右方向を向いた場合、分割レイアウト 2 0 4 が有する分割領域は、車両が上りながら右を向いたことを示す一つ以上のコンテキスト属性（例えばセンサデータ、地図データ等）に基づいて、図 6 H のように、中心点 6 3 1 を左上方向に移動させることで形成されなおしてもよい。

20

また、坂を下りながら右方向を向いた場合、分割レイアウト 2 0 4 が有する分割領域は、車両が下りながら右を向いたことを示す一つ以上のコンテキスト属性（例えばセンサデータ、地図データ等）に基づいて、図 6 I のように、中心点 6 3 1 を左下方向に移動させることで形成されなおしてもよい。

【 0 0 6 5 】

図 6 L ~ 図 6 M は、車両コンテキストによって中心点を動的に調整するさらなる例を示す図である。

これらの図では、シーンの態様を曖昧にしないように、分割領域が省略されている。

図 6 L は、シーン内の道路の消失点に中心点 6 6 1 を持つ直線道路を示している。また、図 6 M は、右方向へのカーブに基づいて、中心点 6 6 1 を左方向に移動させた例である。

30

また、図 6 N は、左方向へのカーブに基づいて、中心点 6 6 1 を右方向に移動させた例である。

【 0 0 6 6 】

いくつかの実施形態において、中心点は、一つ以上のコンテキスト属性（例えば車両の方角、速度、予測経路、地図データ）に応じて、画像の予想領域に対する分割領域の形状のどこをどのように修正するかを計算するために、分割モジュール 2 0 4 によって使用される基準点であってもよい。

【 0 0 6 7 】

分割モジュール 2 0 4 は、決定または処理された分割レイアウトを表すデータを、画素分類モジュール 2 0 6、エビデンス生成モジュール 2 0 8 のうちの一つ以上に送信してもよく、当該分割レイアウトを表すデータを、記憶装置 2 4 5 やメモリ 2 3 7 に記憶させてもよい。

40

【 0 0 6 8 】

画素分類モジュール 2 0 6 は、画像に含まれる画素を一つ以上の画素タイプに分類するためのルーチンを含むソフトウェアである。

一実施形態において、画素分類モジュール 2 0 6 は、画像に含まれる画素を分類するために、以下に示す構造、処理、または機能を提供するための、プロセッサ 2 3 5 で実行可能な命令の集合を含む。

また、一実施形態において、画素分類モジュール 2 0 6 は、コンピュータ 2 0 0 のメモリ 2 3 7 に格納され、プロセッサ 2 3 5 によってアクセスされ、実行されてもよい。

50

また、画素分類モジュール 206 は、プロセッサ 235 およびコンピュータ 200 の他の構成要素と協働して通信するように最適化されてもよい。

【0069】

画素分類モジュール 206 は、分割モジュール 204 が決定した一つ以上の分割レイアウトを表すデータを、分割モジュール 204 や記憶装置 245 から取得することができる。

画素分類モジュール 206 は、画像を分割領域によって分割するために、一つ以上の分割レイアウトを当該画像に適用（例えば、オーバーレイ、マッチ等）することができる。

一実施形態において、画像内の画素は、各画素をある領域、すなわち $r_m = \{P_{i,j}\}_m$ ($m = 1, 2, \dots, M$) に割り当てて、連続する画素群を形成することにより、分割領域によって（例えば M 個の分割領域に）分割できる。

10

【0070】

画素分類モジュール 206 は、画素レベルの特徴に基づいて、各分割領域内の画素を、複数の画素タイプのうちの一つの画素タイプに分類することができる。

一実施形態において、複数の画素タイプは、画素分類モジュール 206 によって予め定められる。

【0071】

画素タイプは、画素の特徴を表すデータであり、当該画素を生成する面のタイプを表す。

画素タイプの例として、「空」、「樹木」、「建築構造物」、「路面」、「車線」、「バリアつきの歩道」、「車両」、「歩行者」等が挙げられるが、これらに限定されるものではない。

20

また、画素に含まれる画素レベルの特徴とは、画素の色やテクスチャなどであるが、これらに限定されるものではない。

例えば、画素分類モジュール 206 は、画素の色とテクスチャを判定し、当該色とテクスチャに基づいて、当該画素のタイプを決定することができる。

更なる例として、画素の色が緑であるならば、画素分類モジュール 206 は、当該画素のタイプを「樹木」と決定することができる。

一実施形態において、画素は、カメラ 107 で取得され、または処理されたオリジナルの画像に含まれるオリジナルの画素であり、そして、当該画素の特徴は、当該オリジナルの画素のサイズ、テクスチャ、色であってもよい。

30

一実施形態において、カメラ 107 で取得されたオリジナルの画像は、他のフォーマットに（例えば Raw データを JPEG や PNG 等に）に変換されてもよく、当該画素は、他のフォーマットに変換された画素であってもよい。画素の特徴は、当該変換された画素のサイズ、テクスチャ、色を含む。

さらに別の実施形態において、カメラ 107 で取得されたオリジナルの画像は、画像プロセッサ 202 でダウンサンプリングされたものであってもよく、画素は、ダウンサンプリングされた画像に含まれる、ダウンサンプリング後の画素であってもよい。画素の特徴は、当該ダウンサンプリングされた画素のサイズ、テクスチャ、色を含む。

【0072】

40

画素分類モジュール 206 は、分割領域ごとに、画素レベルの分類結果を得ることができ、画素レベルの分類結果によって、分割領域内の各画素（例えば、ダウンサンプリング後の画素やパッチなど）を一つ以上の画素タイプに分類することができる。

一例において、二つ以上の画素は、所定の画素タイプのセットから、同じ画素タイプに分類される場合がある。さらなる例において、二つ以上の画素は、所定の画素タイプのセットから、異なる画素タイプに分類される場合がある。

【0073】

例えば、各画素が特徴ベクトル $f_{i,j}$ によって表される、 $i \times j$ 個の画素 $P_{i,j}$ からなる画像を考える。当該特徴は、座標 (i, j) を中心とする色およびテクスチャのフィルタ群によって導出することができる。画素レベルの分類器は、ドメイン f から、 n 個の別

50

個の画素タイプ群のうちの一つまでの関数、すなわち $C: f \{c_1, c_2 \cdots c_n\}$ である。また、画素レベルの分類結果は、分類された画素タイプの配列である。

【0074】

画素分類モジュール206は、カメラ107やそのコントローラ、メモリ237、記憶装置245、画像プロセッサ202またはシステム100の他の構成要素から、画像（例えば、ビデオフレーム、写真など）を取得することができる。

一実施形態において、画素分類モジュール206は、画像プロセッサ202、メモリ237、記憶装置245またはシステム100の他の構成要素から、ダウンサンプリングされた画像を取得することができる。

画素分類モジュール206は、前述したものと同一または類似の処理を実行することによって、ダウンサンプリングされた画像の画素を一つ以上の画素タイプに分類することができる。

10

ダウンサンプリング画像では、局所的な依存関係が無視され得るものの、ダウンサンプリング画像の画素レベルの分類結果は、近隣画素が同じ画素タイプに分類される傾向があることから、画素レベルの分類結果は、イラスト調になることがある。

画素レベルの分類結果の例を、図10に示す。

【0075】

一実施形態において、画素分類モジュール206は、画像中の分割領域ごとの画素レベルの分類結果を、エビデンス生成モジュール208に送信することができる。

他の実施形態において、画素分類モジュール206は、画像中の分割領域ごとの画素レベルの分類結果を、記憶装置245またはメモリ237に格納することができる。

20

【0076】

エビデンス生成モジュール208は、画素分類モジュール206が生成した画素分類データに基づいて、画像の分割領域と関連した画素特徴データを生成するためのルーチンを含むソフトウェアである。

一実施形態において、エビデンス生成モジュール208は、画像の分割領域に関連付いた画素特徴データを生成するために、以下に記す構造、処理、または機能を提供するための、プロセッサ235で実行可能な命令の集合を含む。

また、一実施形態において、エビデンス生成モジュール208は、コンピュータ200のメモリ237に格納され、プロセッサ235によってアクセスされ、実行されてもよい。

30

エビデンス生成モジュール208は、プロセッサ235およびコンピュータ200の他の構成要素と協働して通信するように最適化されてもよい。

【0077】

エビデンス生成モジュール208は、画素分類モジュール206や記憶装置245から、画像の分割領域ごとに、画素レベルの分類結果を取得することができる。画素レベルの分類結果は、当該分割領域における各画素に対応する画素タイプを含む。

また、エビデンス生成モジュール208は、分割領域ごとに、当該分割領域に関連付いた画素レベルの分類結果に基づいて、画素特徴データのセットを決定することができる。

例えば、エビデンス生成モジュール208は、分割領域の画素の画素タイプに基づいて、当該分割領域における、画素特徴データのセットを決定することができる。

40

【0078】

画像の分割領域に関連付いた画素特徴データは、当該分割領域に含まれる一部の画像を表す。

例えば、画像の所定の分割領域において、画素特徴データのセットは、当該分割領域にあるオブジェクトの一つ以上の特徴を特徴づける。

一例において、画素特徴データのセットは、分割領域の画素の内容を示す統計的なデータである。例えば、画素特徴データのセットは、画像の分割領域内の画素タイプの分布を表す尤度ベクトルである。

例えば、尤度ベクトルは、分割領域における画素タイプの57%が路面であり、23%

50

がバリアであり、15%が樹木であり、5%が建築構造物であることを示す。もちろん、他の尤度ベクトルの例も採用可能である。

【0079】

一実施形態において、分割領域に関連付いた画素特徴データは、当該分割領域の画素を形容したもの、またはサマリである。例えば、画素特徴データは、当該分割領域における画素の統計のセットや、ヒストグラムであってもよい。

【0080】

さらなる例として、エビデンス生成モジュール208は、分割領域に含まれる画素タイプの集合に対応するヒストグラムを生成することで、画素タイプの集合に関連付いた、画素タイプの分布を決定することができる。当該画素タイプの分布は、分割領域に含まれる部分画像の画素タイプのセットの確率分布であってもよい。

例えば、各領域は、当該領域内の画素タイプのヒストグラム ($H_m = [|c_1|_m, \dots, |c_n|_m]$) によって表すことができる。なお、 $|c_i|_m$ は、領域 m における各画素タイプを表し、 $i = 1, 2, \dots, n$ である。

エビデンス生成モジュール208は、分割領域の画素特徴データのセットを、ヒストグラムを基とした尤度ベクトルとして生成することができる。

例えば、エビデンス生成モジュール208は、分割領域に対して、「当該分割領域の画素の80%が「空」という画素タイプに分類され、15%が「樹木」という画素タイプに分類され、5%が「路面」という画素タイプに分類された」という情報を持つ尤度ベクトルを生成する。

【0081】

画素特徴データのセットは、分割領域に対応するエビデンスデータのセットとして位置付けることができ、シーン分類モジュール212は、以下に説明するように、全分割領域における全画素特徴データを、画像をシーンに分類するためのエビデンスとして用いることができる。

例えば、各分割領域における画素特徴データは、確率モデルへの入力データとすることができ、画像のシーンを決定するための確率モデルへのエビデンスとして用いることができる。確率モデルについては、以下に詳述する。

【0082】

一実施形態において、エビデンス生成モジュール208は、シーン分類モジュール212に、画像の各分割領域に対応する画素特徴データを送信することができる。

他の実施形態において、エビデンス生成モジュール208は、各分割領域に対応する画素特徴データを、記憶装置245またはメモリ237に格納することができる。

【0083】

学習モジュール210は、確率モデルを学習するためのルーチンを含むソフトウェアである。一実施形態において、学習モジュール210は、プロセッサ235で実行可能な命令の集合を含む。また、一実施形態において、学習モジュール210は、コンピュータ200のメモリ237に格納され、プロセッサ235によってアクセスされ、実行されてもよい。

学習モジュール210は、プロセッサ235およびコンピュータ200の他の構成要素と協働して通信するように最適化されてもよい。

【0084】

実施形態における確率モデルとは、画像のシーンを分類するためのモデルである。

一実施形態において、確率モデルは、ベイズネットワークのような、確率的グラフィカルモデルである。確率モデルの例を、図8Aおよび図8Bに示す。

一実施形態において、学習モジュール210は、画像の分割領域に関連付いた画素特徴データ（例えば仮想的なエビデンスデータ）を含む学習データセットを取得する。学習データセットにおける画像のシーンは、予め分類され、複数のラベルによってラベリングされたものである。学習モジュール210は、下記に示すように、学習データセットを用いた確率モデルの学習を実行することができる。

例えば、学習モジュール 210 は、学習データセットを用いて、確率モデルの構造、入力ノード、出力ノードを学習することができる。

【0085】

ベイズネットワークは、同時確率分布を局所確率モデルに因数分解したものである。各々が、ネットワーク内の一つのノードに対応しており、ノード間の有向アークは、別のノードの確率モデルにおける、あるノードの確率モデルの条件を表す。推論（分類など）は、当該アークの因果方向に対向する方向に働く。推論は、ネットワーク内の低次のエビデンスから、ベイズネットワークの最上位に位置するクラスノード（シーン変数ノード）へと流れる。クラスノードでは、ベイズネットワークが、クラス変数（ここではシーン変数）に関する事後分布を生成することができる。

10

【0086】

ベイズネットワークは、異なる分割入力、異なる出力カテゴリ、異なる構造制約の下で、簡単に再学習することができる。これらの出力カテゴリは、ユーザによって提供されてもよく、場合によっては、現在のベイズネットワークにないカテゴリを含むことができる。

例えば、アプリケーションは、ユーザによって導入され、ベイズネットワークに存在しないシーンカテゴリを考慮し、シーンに対応する詳細を入力するようにユーザに促してもよい。

ベイズネットワークは、これらの例を用いて、新しい場面カテゴリを取り込むために再学習することができる。異なるエビデンスのタイプを入力とし、あるいは、異なる種類として定義されたノードを出力としてテストするためのモデルを簡単に変更できることは、

20

本明細書に記載のシーン識別技術の利点である。

このモデルの構造は、ドメインの特性を表すモデル変数間の依存関係を発見することができる。

【0087】

ベイズネットワークのような確率モデルの学習は、

(1) 入力変数および出力変数の構造を学習する

(2) 可変的な条件付き確率テーブルのパラメータを学習する

という二つの態様を含む。構造とパラメータの学習は、同時に行われてもよい。また、入次数およびノード順を制限するために制約を用いることができる。

30

また、ベイズネットワークの学習は、インタラクティブに実行されることができる。

【0088】

図 8 B に示したように、ベイズネットワークのような確率モデルは、複数のレベル（例えば最上位レベルと下部のレベル）で構築されることができる。図 8 B の例では、出力が最上位レベル、入力下部レベルである。この複数のレベル構造は、ベイズネットワークを使用して分類を行うための標準的な構造を提供することができる。本例の場合、確率モデルは複数の出力ノードを持っていてもよい。

ベイズネットワークの入力ノードは、仮想エビデンスノードと呼ばれる場合がある。仮想エビデンスノードは、出力を生成するためのエビデンスをベイズネットワークに提供することができる。各々の入力ノードは、画像に含まれる分割領域の一つに対応し、当該分割領域に関連付いた画素特徴データを取得し、確率モデルへの入力データとすることができる。例えば、一つの入力ノードが一つの分割領域に対応してもよい。この場合、ベイズネットワークの入力ノードの数は、画像の分割領域の数と同一となる。

40

【0089】

また、ベイズネットワークにおいて、入次数とノード順を制限するために、更なる制約が用いられてもよい。例えば、エビデンスノードの入次数は、二つに限定することができる。

一実施形態において、ノードの順序は、構造上の制約として課すことができ、その結果、条件付けアークが低次から高次へと移動できなくなる。

なお、シーン変数の間で課される更なる因果制約があってもよい。例えば、「周辺環境

50

」というシーン変数は車道のタイプに影響するが、逆は成り立たない。別の例では、「周辺環境」というシーン変数は、「運転状況」という変数に影響するが、逆は成り立たない。

結果的に、確率モデルは、ほぼ純粋なベイズ構造に従うが、学習時に行われたモデル選択の結果である追加アークを伴う。学習過程で作成されたアークは、入力変数と出力変数との間で発見された依存関係を表し得る。

【 0 0 9 0 】

一実施形態において、学習モジュール 2 1 0 は、ベイズネットワークを学習させるための仮想的なエビデンスとして、分割領域の画素特徴データを用いることができる。

一例において、仮想的なエビデンスを含むサンプルからベイズネットワークの近似学習を行うための、少なくとも三つの方法が存在する。

【 0 0 9 1 】

第一の方法は、仮想エビデンスの状態ごとの尤度比に準じて各エビデンス行の倍数を生成することにより、ほぼ同等の観察エビデンスデータセットに変換する方法である。

複数の仮想エビデンスノードが存在する場合、仮想エビデンスノード間の依存関係を捕捉する操作により、行セットの組合せ爆発を引き起こしうる。すなわち、仮想エビデンスノード状態の組合せごとの倍数が、状態組合せの尤度に比例して多重化しうる。その複雑度は、すべての仮想エビデンスノードを結合してサンプリング用の一ノードにすることに匹敵する。

【 0 0 9 2 】

同様に、学習モジュール 2 1 0 は、すべての仮想エビデンスノードの組合せからサンプリングを行い、サンプル内の項目に基づいてサンプル行を生成することができる。これは、仮想状態をサンプリングするロジックのようなものである。両手法とも、学習重みを等しくするための方法として、学習セットに複数の行コピーを作成する。あるいは、学習モジュール 2 1 0 が、その尤度に応じて、サンプリングされたトレーニングセットの各行に重みを付けても良い。

【 0 0 9 3 】

第二の方法は、学習された決定論的エビデンスモデルの混合である「マルチネット」を考慮した方法である。これらのモデルは、同じ構造を有し得るため、結果は、尤度によって重み付けされた条件付確率表 (CPT) の混合となる。この手法は、混合構成要素の組合せ爆発を被ることがあり、サンプリングによってセットが減少しやすい場合がある。

【 0 0 9 4 】

第三の方法は、仮想エビデンスノードの子として追加された仮想ノードによる、仮想エビデンスを考慮する方法である。当該仮想エビデンスノードは、その親に同等の尤度ベクトルを送るようにインスタンスが生成される。多くの場合、各ケースにおいて、ネットワークに追加された仮想ノード群が生じ、管理できない可能性のある方法を再び生成する。一実施形態において、学習モジュール 2 1 0 は、一つのノード群を有するネットワークを構築し、一つの学習ステップを実行し、その後ノードを次の群で置き換え、学習ステップを繰り返すことができる。なお、その他の方法も採用可能でもある。

【 0 0 9 5 】

本明細書に記載の確率モデルを学習するためのサンプルデータセットが、上記第一の方法におけるさらなる推定となり得る。そこでは、各仮想エビデンスノードを独立してサンプリングすることにより、この問題を、サンプリングされたデータを有する同等の問題に変換することができる。そのため、各ヒストグラムは、当該分布を用いてサンプリングすることにより、そのヒストグラムを推定した従来のエビデンスサンプル群を生成することができる。結果は、行数に、ヒストグラムデータセット内の行ごとのサンプルサイズを掛け合わせた拡張データセットとなる。

結果として生成される学習データセットの詳細は、例えば以下になる。

1) 5つのシーンラベルによってラベル付けされた、画像の 1 2 の領域ヒストグラムの 1 2 2 の行が、元のデータセットとして存在する

- 2) 各領域ヒストグラムが10回サンプリングされ、1220行が生成される
3) 1220行によって、5つのラベルと12個の特徴が表された最終データセットが得られる

学習データセットを用いて評価した結果の例を、図7Bに示す。

【0096】

シーン分類モジュール212は、画像で表現されたシーンを分類するためのルーチンを含むソフトウェアである。

一実施形態において、シーン分類モジュール212、画像で表現されたシーンを分類するために、以下に記す構造、処理、または機能を提供する、プロセッサ235で実行可能な命令の集合を含む。また、一実施形態において、シーン分類モジュール212は、コンピュータ200のメモリ237に格納され、プロセッサ235によってアクセスされ、実行されてもよい。

シーン分類モジュール212は、プロセッサ235およびコンピュータ200の他の構成要素と協働して通信するように最適化されてもよい。

【0097】

シーン分類モジュール212は、エビデンス生成モジュール208や記憶装置245から、画像内の複数の分割領域に関連する画素特徴データの複数のセットを取得することができる。

また、シーン分類モジュール212は、記憶装置245や学習モジュール210から、確率モデルを表すデータを取得することができる。

シーン分類モジュール212は、以下に詳述するように、複数の画素特徴データと確率モデルのセットに基づいて、画像のシーンを分類することができる。

【0098】

一実施形態において、複数の画素特徴データのセットは、それぞれが分割領域に対応する複数の尤度ベクトルである。

シーン分類モジュール212は、複数の尤度ベクトルと確率モデルに基づいて、一つ以上のシーン変数に関連づいた分布を決定することができる。

例えば、シーン分類モジュール212は、確率モデルへの入力データとして複数の尤度ベクトルを用いることができ、一つ以上のシーン変数に関連付いた分布を表す確率モデルの出力を生成することができる。

一つ以上のシーン変数に関連付いた分布は、一つ以上のシーン変数の同時分布、一つ以上のシーン変数の条件付き分布、シーン変数ごとの個別分布、またはこれらの組合せを含む。

一実施形態において各シーン変数の分布は、これらが画素タイプの分布に依存するように、他のシーン変数に依存する。これは、同時分布（例えば全てのシーン変数の分布）を形成するシーン変数による。

【0099】

シーン変数は、シーンのカテゴリを表す変数である。

シーン変数の例として、例えば、周辺環境、道路のタイプ、運転状況、動的環境、道路障害物状況などが挙げられるが、これに限られない。また、各シーン変数は、ラベルのセットを有していてもよい。シーン変数のラベルは、シーン変数に対する確率値を表す値である。

例えば、「道路」というシーン変数は、「カーブと勾配」「自動車専用道」「狭隘」「路肩なし」「駐車車両あり」等といったラベルを持つ。

また、「周辺環境」というシーン変数は、「ゾーニング」「開発区域」「商業地区」「山岳」「田園」「住宅地」「都市部」等といったラベルを持つ。

また、「運転状況」というシーン変数は、「自転車と歩行者」「交通と混雑度」「障害物なし」等といったラベルを持つ。

また、「動的環境」というシーン変数は、照明や天候といった視程に関するラベル（例えば、逆光による視界不良」「曇り」「晴れ」「天候による視界不良」など）等といった

10

20

30

40

50

ラベルを持つ。

また、「道路障害物状況」というシーン変数は、「障害物なし」、「工事中」、「合流または交差点」、「木の幹や柱」等といったラベルを持つ。

【0100】

シーン分類モジュール212は、画像におけるシーンの分類を、一つ以上のシーン変数に関連付いた分布に基づいて決定する。

例えば、シーン分類モジュール212は、一つ以上のシーン変数に関連付いた分布に基づいて、各シーン変数の一つのラベルを、画像に割り当てる。

更なる例として、シーン変数に関連する分布が、シーン変数の同時確率分布であると仮定する。当該同時確率分布は、以下のような場合に、高い確率値を有する。

(1) シーン変数「道路」が、第一のラベル「カーブと勾配」であり、

(2) シーン変数「環境」が、第二のラベル「山岳」であり、

(3) シーン変数「運転状況」が、第三のラベル「障害物なし」であり、

(4) シーン変数「動的環境」が、第四のラベル「曇り」であり、

(5) シーン変数「障害物」が、第五のラベル「合流または交差点」であった場合。

この結果、シーン分類モジュール212は、第一～第五のラベルを画像に割り当て、分類された画像のシーンは、前記第一～第五のラベルによって表されたものとなる。

【0101】

別の例において、シーン変数に関連する分布が、各シーン変数の個々の確率分布を含むと仮定する。この場合、

(1) シーン変数「道路」の確率分布は、当該シーン変数の第一のラベルが「カーブおよび勾配」である場合に高い確率値を有する。

(2) シーン変数「環境」の確率分布は、当該シーン変数の第二のラベルが「山岳」である場合に高い確率値を有する。

(3) シーン変数「運転状況」の確率分布は、当該シーン変数の第三のラベルが「山岳」である場合に高い確率値を有する。

(4) シーン変数「動的環境」の確率分布は、当該シーン変数の第四のラベルが「曇り」である場合に高い確率値を有する。

(5) シーン変数「障害物」の確率分布は、当該シーン変数の第五のラベルが「合流または交差点」である場合に高い確率値を有する。

この結果、シーン分類モジュール212は、第一～第五のラベルを画像に割り当て、分類された画像のシーンは、前記第一～第五のラベルによって表されたものとなる。

【0102】

さらなる例として、画像のM個の領域に対応する仮想エビデンスノードを有する確率モデルを用いてシーン分類を行ってもよい。各エビデンスノードは、画素特徴データ（仮想エビデンスなど）を、ラムダメッセージ λ_m という形態で受け取ることができ、状態尤度は、 H_m が提供する比で表される。確率モデルは、ノードのサブセット $S = \{S_1, \dots, S_V\}$ を有する。このサブセットは、シーン変数ノード群を表し、各シーン変数ノードは別個の状態空間を有する（例えば、各シーン変数は別個のラベル群を有する）。

シーン分類は、

(1) $P(S | \lambda_1 \dots \lambda_m)$ 、すなわちM個の分割領域からの λ_m メッセージが適用された場合のSの同時分布

(2) Sに関する最大事後確率(MAP)構成による同時分布の特徴、または

(3) Sの事後周辺分布

によって表すことができる。

【0103】

一実施形態において、シーン分類モジュール212は、画像全体のシーンを分類することができる。

例えば、シーン分類モジュール212は、画像によって示されるコンテンツの全体の全体的な印象を決定し、画像を一つ以上のシーンカテゴリのラベルに分類するため、全体的

10

20

30

40

50

な印象を使用することができる。

ここに記載された、この全体的なシーン分類は、シーンの一部のみに適用したり、シーンの一部のオブジェクトなどの、シーンを構成する一部のみを識別した結果に依存する他の分類とは異なる。

【0104】

一実施形態において、シーン分類モジュール212は、シーン分類の性能を評価することができる。例えば、仮のデータセットについてのテストは、シーン変数の分類のうち、良好な正確度 (Accuracy) (例えば70%を超える正確度) を示す。

評価結果の例を、図7Bに示す。

【0105】

シーン利用アプリケーション117は、ユーザに対してアドバイスまたはアシストするための様々なアプリケーションにおいて、シーンを利用するためのルーチンを含むソフトウェアである。

例えば、シーン利用アプリケーション117は、アプリケーションマーケットからクライアント装置115や移動プラットフォーム135にダウンロードされた「アプリ」であってもよく、車両における、インフォマティクス・ソフトウェアシステムであってもよい。

また、クライアント装置115等で動作するネイティブアプリケーションであってもよい。

また、ブラウザからアクセスされるウェブベースのアプリケーション等であってもよい。

シーン利用アプリケーション117は、ナビゲーション、マッピング、ソーシャルネットワーク、コミュニケーション、おすす、メディアなどを含む、様々な異なるアプリケーションを含み、または、表したものであってもよい。

一実施形態において、シーン利用アプリケーション117は、ユーザに対してアドバイスまたはアシストするための様々なアプリケーションにおいて、シーンを利用するため、以下に記す構造、処理、または機能を提供する、プロセッサ235で実行可能な命令の集合を含む。また、一実施形態において、シーン利用アプリケーション117は、コンピュータ200のメモリ237に格納され、プロセッサ235によってアクセスされ、実行されてもよい。

シーン利用アプリケーション117は、プロセッサ235およびコンピュータ200の他の構成要素と協働して通信するように最適化されてもよい。

【0106】

一実施形態において、シーン利用アプリケーション117は、シーン分類モジュール212や記憶装置245から分類された画像のシーンを取得し、ユーザにシーン分類結果を提示するユーザインタフェースを提供するために、グラフィックデータを生成することができる。シーン分類結果は、画像の分類されたシーンを表す。

シーン利用アプリケーション117は、クライアント装置115や移動プラットフォーム135にグラフィックデータを送信し、ユーザインタフェースをユーザに提供させることができる。

また、シーン利用アプリケーション117は、他のユーザインタフェースをユーザに提供するためのグラフィックデータを生成してもよい。

【0107】

シーン分類の目的のひとつに、シーン変数に割り当てられたラベルから、現在のビューの概略を取得することがある。

例えば、シーン分類は、「あなたが見ているものに基づいて、場所をどのように説明しますか?」といった質問に対する答えを提供することができる。

一実施形態では、シーン利用アプリケーション117は、シーンに特有なおすめをユーザに提供することができる。

例えば、分類されたシーンは、環境におけるコンテキストを提供することができ、シー

10

20

30

40

50

ン利用アプリケーション 117 は、ユーザに対する適切なおすすめを決定するために分類されたシーンを利用することができる。

例えば、個別化されたおすすめを作成する際に、分類されたシーンにおけるコンテキストを用いて、「プリトーを食べたくありませんか？」や「ここは、それを見つけられる場所ではありません」などのリクエストの実際性を判断することができる。

【0108】

一実施形態において、シーン利用アプリケーション 117 は、「シーン拡張検索」をユーザに提供することができる。

例えば、シーン利用アプリケーション 117 は、「市街地でワイナリーを見つける」「僻地でレストランを見つける」「行ったことのない住宅地にある、混んでいない公園を見つける」等といったような、ローカルな環境の特徴を用いて検索を補足することができる。

10

【0109】

一実施形態において、シーン利用アプリケーション 117 は、現在の環境における好ましさについてのスコアを決定することができる。

経路を選択する際にドライバーは、しばしば当該経路における快適さを考慮に入れる。

例えば、ドライバーは、「この場所からの眺めはどうか？」「この地域の日当たりや日陰はどの程度か？」「環境における自然と人造物の割合はどの程度か？」等を考慮する場合がある。

好ましい場所についての個々のドライバーの評価があれば、シーン利用アプリケーション 117 は、他の望ましい経路をドライバーに提案してもよい。当該望ましい経路は、最も短い、または、最も速いルートとは異なる場合がある。

20

【0110】

一実施形態において、シーン利用アプリケーション 117 は、クラウドをソースとしたシーン特徴の集約結果をユーザに提供することができる。

シーン利用アプリケーション 117 は、近隣箇所に関するクラウドソーシングされた情報を収集して、ある地域の全体像を形成することができる。

本例は、上述した例の拡張である。

クラウドソースのデータを用いることで、ドライバーは、例えば「道路上の人の群れがどこまで延びていますか？（例えばロードレースやパレードなどの場合）」という質問に対する回答といったような、集約された状況についての情報を近隣の車両から得ることができる。

30

【0111】

ここに記述されるシーン分類技術は、オブジェクト認識とは多くの点において異なる。オブジェクト認識技術は、画像内の物体の存在、識別情報、姿勢などを取得するタスクであるが、本明細書に記載のシーン分類は、画像全体に対して推論を行うタスクである。一方、オブジェクト認識技術は、オブジェクトと、オブジェクト以外（背景）を区別することにフォーカスを置いている。

オブジェクト認識技術と比較して、シーン分類技術は、画像を全体的に処理することによって、計算コストの削減を達成することができる。

40

例えば、シーン分類技術は、あらかじめ定義されたカテゴリのセットから、画像を「屋内」、「屋外」、「都市部」、「地方部」などの種類に割り当てることができる。

シーン分類は、「カテゴリのラベルに終端部を割り当てること」とみなすことができる。カテゴリでは、画像に複数のラベル群にわたる値を割り当てることができ、終端部はいくつかのシーン変数にわたる同時分布となり得る。

【0112】

(処理フローチャート)

図 3 は、画像からシーンを推定する方法 300 のフローチャートである。

まず、画素分類モジュール 206 が、カメラ 107、記憶装置 245、画像プロセッサ 202 またはシステム 100 のもう一つの構成要素から、画像を表すデータを取得する（

50

ステップ 3 0 2)。

また、分割モジュール 2 0 4 が、シーンに対応する分割領域を決定する (ステップ 3 0 4)。分割領域の形状は、前述したように、一つ以上のコンテキスト属性に基づいて決定される。分割レイアウトは複数の分割領域から構成され、各分割領域は、シーンごとに対応するものが用いられる。

画像はシーンを表すものであり、その取得タイミングは、分割領域を決定する前であってもよいし、決定中、または決定後であってもよい。

【 0 1 1 3 】

次に、画素分類モジュール 2 0 6 が、各分割領域の画素を一つ以上の画素タイプに分類する (ステップ 3 0 6)。例えば、画素分類モジュール 2 0 6 は、分割領域を画像に適用し (例えば整列させる)、どの画素がどの分割領域に属するか決定したうえで、画像の各分割領域に含まれる画素を、一つ以上の画素タイプに分類する。

例えば、画素分類モジュール 2 0 6 は、予め定められた 8 つの画素タイプから、分割領域の各画素を、一つの画素タイプに分類することができる。

【 0 1 1 4 】

次に、各分割領域について、エビデンス生成モジュール 2 0 8 が、当該分割領域内の画素の画素タイプに基づいて、分割領域に含まれる画像の一部分に対応する画素特徴データのセットを決定する (ステップ 3 0 8)。

次に、シーン識別モジュール 2 1 2 が、分割領域に対応する各々の画素特徴データのセットに基づいて、画像のシーンを分類する (ステップ 3 1 0)。

例えば、シーン識別モジュール 2 1 2 は、複数の分割領域に関連付いた画素特徴データの複数のセットを、確率モデルへの入力として使用し、当該確率モデルの出力に基づいて画像のシーンを分類することができる。

【 0 1 1 5 】

図 4 A および図 4 B は、画像からシーンを推定するもう一つの方法 4 0 0 のフローチャートである。

図 4 A を参照して説明する。

まず、学習モジュール 2 1 0 が、学習データセットに基づいて、入力ノードおよび出力ノードを含む確率モデルの学習を行う (ステップ 4 0 2)。

入力ノードは、画素特徴データを確率モデルに入力するための仮想エビデンスノードである。また、出力ノードは、シーン変数に関連付いた分布を出力するシーン変数ノードである。

次に、カメラ 1 0 7 が、移動プラットフォーム 1 3 5 の周辺環境 (例えばシーン) を表す画像 (例えばビデオフレームや静止画像) を取得する (ステップ 4 0 4)。

次に、画像プロセッサ 2 0 2 が、カメラ 1 0 7、記憶装置 2 4 5 やシステム 1 0 0 の他の構成要素から、画像を表すデータを取得する (ステップ 4 0 5)。

なお、画像プロセッサ 2 0 2 が、画像を処理する前に、画像に含まれる画素を減らしてもよい (ステップ 4 0 6)。例えば、画像プロセッサ 2 0 2 は、取得した画像をダウンサンプルして、ダウンサンプル画像を生成することができる。

次に、分割モジュール 2 0 4 が、画像に対応する分割レイアウトを、移動プラットフォーム 1 3 5 (例えば車両) の一つ以上のコンテキスト属性や、移動プラットフォーム 1 3 5 の移動環境に基づいて決定する (ステップ 4 0 8)。

次に、画素分類モジュール 2 0 6 が、画像の各分割領域に含まれる各画素を、画素タイプのセットの中からいずれかの画素タイプに分類する (ステップ 4 1 0)。

【 0 1 1 6 】

次に、エビデンス生成モジュール 2 0 8 が、分割領域ごとに、分割領域で分割された画像の一部を特徴付けるデータである、画素タイプのヒストグラムを生成することにより、画素タイプのセットと関連付いた画素タイプの分布を生成する (ステップ 4 1 2)。

【 0 1 1 7 】

図 4 B を参照して説明する。

10

20

30

40

50

次に、エビデンス生成モジュール 208 が、分割領域ごとに、生成したヒストグラムに基づいて尤度ベクトルを生成する（ステップ 414）。

次に、シーン分類モジュール 212 が、分割領域に関連付いた尤度ベクトルと、確率モデルに基づいて、一つ以上の、シーン変数に関連付いた分布を決定する（ステップ 416）。

次に、シーン分類モジュール 212 が、当該分布を用いて、画像のシーンを決定する（ステップ 418）。

また、シーン利用アプリケーション 117 が、様々なシーンアプリケーションにおいて、分類されたシーンを適用する（ステップ 420）。

【0118】

図 5 を参照して、ステップ 420 についてより詳しく説明する。

図 5 は、様々なアプリケーションにおいて、分類されたシーンを利用する方法 420 のフローチャートである。

まず、シーン利用アプリケーション 117 が、記憶装置 245 やシーン分類モジュール 212 から、分類されたシーンを表すデータを取得する（ステップ 502）。

次に、シーン利用アプリケーション 117 が、ユーザに おすすめを提供するか否かを決定する（ステップ 504）。ここで、おすすめをユーザに提供する場合、方法 420 はステップ 506 に遷移し、それ以外の場合、方法 420 はステップ 508 に遷移する。

ステップ 506 では、シーン利用アプリケーション 117 が、ユーザに対するおすすめを行うためのコンテキストを提供するために、分類されたシーンを適用する。その後、処理はステップ 508 に遷移する。

【0119】

次に、ステップ 508 で、シーン利用アプリケーション 117 が、ユーザのために検索を実行するか否かを決定する。ここで、ユーザのために検索を実行する場合、方法 420 はステップ 510 へ遷移し、それ以外の場合、方法 420 はステップ 512 へ遷移する。

ステップ 510 では、シーン利用アプリケーション 117 が、ローカル環境を用いた検索を補完するために、分類されたシーンを利用する。その後、方法 420 はステップ 512 へ遷移する。

【0120】

次に、ステップ 512 で、シーン利用アプリケーション 117 が、シーンを評価するか否かを決定する。ここで、シーンを評価する場合、方法 420 はステップ 514 へ遷移する。それ以外の場合、方法 420 はステップ 516 へ遷移する。

ステップ 514 では、シーン利用アプリケーション 117 が、分類されたシーンに関連した場所における、現在の状況の評価をする。その後、方法 420 はステップ 516 へ遷移する。例えば、シーン利用アプリケーション 117 は、シーンにおける環境の快適さを評価することができる。

【0121】

次に、ステップ 516 では、シーン利用アプリケーション 117 は、他のユーザから得たクラウドソースの情報の有無を判断する。ここで、クラウドソースの情報を利用する場合、方法 420 はステップ 518 へ遷移する。それ以外の場合、方法 420 は終了、または、上述の処理を繰り返す。

ステップ 518 では、シーン利用アプリケーション 117 が、地域の全体像を形成するために、クラウドソースの情報を適用する。その後、方法 420 は終了、または、上述の処理を繰り返す。

【0122】

（グラフィック表現）

図 6 J は、画像における画素レベルの分類を例示するグラフィック表現 600 である。本例では、画像中の画素 1 と画素 2 が、「空」という画素タイプに分類され、画素 3 と画素 5 は、「樹木」という画素タイプに分類され、画素 4 と画素 6 は、「建物」という画素タイプに分類される。

10

20

30

40

50

図 6 K は、分割レイアウト 6 5 5 を例示するグラフィック表現である。例示されたように、分割レイアウトは、中心点 6 5 6 で収束する 1 2 個の領域を含む。1 2 個の領域の各々は、中心点 6 5 6 から、画像の端に対応する端辺に向かって延びている。

図 6 J から図 6 K における、画素番号 1 ~ 6 は、いくつかの例において一致する。

【 0 1 2 3 】

図 8 B は、確率モデルを例示するグラフィック表現 8 2 0 である。

確率モデルは、複数の入力ノード 8 2 2 と、複数の出力ノード 8 2 4 a ... 8 2 4 n を含む。各々の入力ノード 8 2 2 は、画像における分割領域にそれぞれ対応しており、確率モデルへの入力データとして、当該分割領域に関連した画素特徴データのセットを用いる。

各々の出力ノード 8 2 4 は、シーン変数を表し、確率モデルは、シーン変数と関連した一つ以上の分布を出力することができる。

出力される分布は、シーン変数の同時分布、各シーン変数の個々の分布、シーン変数の条件つき分布、またはその組合せなどである。

2 つのノードをつないでいるラインまたはアーク（例えば、ノード 8 2 4 a とノード 8 2 4 n を繋ぐライン 8 2 6、ノード 8 2 4 a とノード 8 2 9 を繋ぐライン 8 2 8）は、二つのノード間の依存または関連を示す。

ここに示した確率モデルの構造は一例である。

さらなる実施形態において、確率モデルは、異なる入力ノード、異なる出力ノード、および、ノード間を繋ぐ異なるラインを持ってもよい。

【 0 1 2 4 】

図 7 A は、混同行列を例示するグラフィック表現 7 0 0 である。

確率モデルの学習過程において、シーン分類モジュール 2 1 2 は、シーン変数ごとに混同行列を生成し、全てのシーン変数に対応する混同行列を、シーン分類のパフォーマンスを評価するために使用することができる。

図 7 A に示したように、シーン変数に関連付いた混同行列は、予測されたラベル 7 0 2 と真のラベル 7 0 4 の数を表す行列である。

混同行列の行は、真のラベル 7 0 4 であり、混同行列の列は、予測されたラベル 7 0 2 である。

予測されたラベルは、画像における、推定されたラベルである。また、真のラベルは、画像における、実際のラベルである。

予測されたラベルが真のラベルと同一であれば、シーン変数についての画像の分類は正しいことになる。さもなければ、シーン変数についての画像の分類は誤っているということになる。

例えば、もし、「車道」というシーン変数についての、予測されたラベルが、真のラベル「カーブと勾配」と同じものであった場合、「カーブと勾配」を予測した画像の分類は正しいことになる。これに対し、もし、「車道」というシーン変数についての、予測されたラベルが、真のラベル「カーブと勾配」とは異なる「狭隘」というラベルであった場合、「狭隘」を予測した画像の分類は誤っていることになる。

【 0 1 2 5 】

混同行列の対角要素は、予測したラベルが正しかった数を意味する。

例えば、「5 0」という値を有する第一の対角要素は、シーン分類モジュール 2 1 2 が、ラベルを「1」と予測し、実際のラベルが「1」であったことが 5 0 回あったことを表す。

非対角要素は、分類エラーを表す。例えば、「2」という値を有する非対角要素は、シーン分類モジュール 2 1 2 が、ラベルを「2」と予測し、実際のラベルが「1」であったことが 2 回あったことを表す。

また、「4」という値を有する非対角要素は、シーン分類モジュール 2 1 2 が、ラベルを「1」と予測し、実際のラベルが「2」であったことが 4 回あったことを表す。

【 0 1 2 6 】

各列の合計は、対応するラベルが、画像に対応するラベルとして予測されたカウント数

を表す。

例えば、列 7 0 8 は、ラベル 1 として予測された回数が 5 9 回あり、そのうち、真のラベルがラベル 1 であった回数が 5 0 回あり、ラベル 2 であった回数が 4 回あり、ラベル 3 であった回数が 5 回あったことを表す。

各行の合計は、対応するラベルが、画像の真のラベルであった総カウント数を表す。

例えば、行 7 0 6 は、真のラベルがラベル 1 であった回数が 5 5 回あり、そのうち、予測されたラベルがラベル 1 であった回数が 5 0 回あり、ラベル 2 であった回数が 2 回あり、ラベル 3 であった回数が 3 回あったことを表す。

【 0 1 2 7 】

ラベルの再現率 (Recall) やカバレッジは、正しいラベルに分類されたケースと、全ラベルの分類ケースとの比率である。

ラベルに対応する再現率は、以下の比率として計算できる。

(1) ラベルに関連付いた対角要素

(2) ラベルに関連付いた行の合計

例えば、ラベル 1 に対応する再現率は、ラベル 1 に関連した対角エントリ「 5 0 」と、例えば行 7 0 6 の合計との間の比率 (例えば、 $50 / (50 + 2 + 3) = 0.91$) として計算できる。

ラベルに対応する再現率が高い値であることは、ラベルの分類がより正確なことを示す。

【 0 1 2 8 】

ラベルに対応する適合率 (Precision) は、以下の比率として計算できる。

(1) ラベルに関連付いた対角要素

(2) ラベルに関連付いた列の合計

例えば、ラベル 1 に対応する適合率は、ラベル 1 に関連した対角要素「 5 0 」と、例えば列 7 0 8 の合計の間の比率 (例えば、 $50 / (50 + 4 + 5) = 0.85$) として計算できる。

ラベルに対応する適合率が高い値であることは、ラベルの分類がより正確なことを示す。

【 0 1 2 9 】

図 7 B は、シーン分類の評価結果を例示するグラフィック表現 7 5 0 である。

表 1 は、「周辺環境」というシーン変数に関連付いた異なるラベル (例えば、山岳、田園、住宅地、都市部) に対応する再現率と適合率を例示する。

表 2 は、「道路」というシーン変数に関連付いた異なるラベル (例えばカーブと勾配、自動車専用道、狭隘、路肩なし、駐車車両あり) に対応する再現率と適合率を例示する。

表 3 は、「運転状況」というシーン変数に関連付いた異なるラベル (例えば自転車と歩行者、交通と混雑、障害物なし) に対応する再現率と適合率を例示する。

表 4 は、「道路障害物状況」というシーン変数に関連付いた異なるラベル (例えば障害物なし、工事中、合流と交差点、分岐) に対応する再現率と適合率を例示する。

【 0 1 3 0 】

表 1 と表 4 と比較すると、「周辺環境」というシーン変数のラベルの再現率と適合率の平均が、「道路障害物状況」というシーン変数のラベルの再現率と適合率の平均よりも優れていることがわかる。これは、シーン変数「道路障害物状況」の分類が、単に走行車線における細部を考慮するだけなのに対し、シーン変数「周辺環境」の分類は、画像全体を考慮するためである。表 3 において、自転車と歩行者は、画像中の小さな領域に現れるだけであるため、ラベル「自転車と歩行者」に対応する再現率は低い。

シーン分類モジュール 2 1 2 は、周辺予測の精度を評価することに加え、図 8 A に描かれているベイズネットワークによって学習された構造について観察することもできる。

【 0 1 3 1 】

図 8 A は、ベイズネットワークのような確率モデルを例示するグラフィック表現 8 0 0 である。この例では、シンボル S __ 1 ~ S __ 8 を含む、8 0 2 , 8 0 6 , 8 0 8 といった

10

20

30

40

50

各ブロックが、確率モデルの入力ノードを表し、分割領域ごとに割り当てられた画素特徴データ（例えば仮想的なエビデンス）である。

図中の入力ノードは、画像中の対応する分割領域のおおよその位置に配置されている。

また、シンボル $S_1 \sim S_8$ は、例えば、木、建物、空、葉、建築物、路面、車線、歩道、車両、歩行者などの異なる画素タイプを表す。

各々のブロックは、それぞれの分割領域における画素タイプの割合を表す値を含む。例えば、ブロック 802 では、画素タイプ「 S_1 」は 3.4% という値を持ち、画素タイプ「 S_2 」は 2.8% という値を持つ。

各々のブロックのタイトルにある数値は、中心点からの分割線の角度を表す。

例えば、「 $N_BL_60_80$ 」は、左側の 60 度から 80 度の間にある分割領域を表す。また、「 $N_BR_60_80$ 」は、右側の 60 度と 80 度の間にある分割領域を表す。

10

【0132】

灰色で塗られた、ブロック 804 のようなブロックは、確率モデルの出力ノードであり、「道路」「周辺環境」「運転状況」「動的環境」「道路障害物状況」のようなシーン変数に対応している。

各々のシーン変数は、ラベルのセットと、当該ラベルに対応する値を持つ。

例えば、ブロック 804 のシーン変数「動的環境」は、「逆光による視界不良」が 3.7% という確率値を持ち、「曇り空」が 77.8% という確率値を持ち、「晴れ」が 17.3% という確率値を持ち、「天候による視界不良」が 1.2% の確率値を持つことを表す。

20

【0133】

確率モデルを連結するアークまたはラインは、どの入力ノードがどのシーン変数に関連するかを示す。これらのアークは、エビデンスノードの入次数が 2 つに制限されているため、比較的疎である。

学習モジュール 210 で選択されるアークは、分割領域の位置と、異なるシーン変数との間に強い関連を示す。例えば、「運転状況」シーン変数が、画像の基部にある分割領域に繋がっており、「環境」シーン変数が、画像の周辺部にある分割領域に繋がっている場合、強い関連が示される。

図の下部にある二つの分割領域（ブロック 806、808 など）は、各領域の入来アークが他の分割領域のものであることから、その関連性は限られており、各領域のエビデンスは完全に隣接領域によってサポートされていることを示している。これらのアークまたはラインに沿って進んでいくと、区画依存という点から、分割領域から学習されるアークは一つだけであり、そのシーンに寄与したエビデンスは、この事例を除くすべての事例において、条件付きで独立していることを表している。

30

シーン変数の下位ネットワークはさらに接続されており、シーン変数間の強い依存関係を表している。例えば、「道路」というシーン変数における「カーブと勾配」は、「周辺環境」というシーン変数における「山」に対して強い相関を示す。

【0134】

図 9 は、画像からシーンを推定する処理を例示するグラフィック表現 900 である。

40

画像によってどのようなシーンが表されているか不明な状態で、分割モジュール 204 は、一つ以上のコンテキスト属性（例えば、車両、車両の移動環境など）に基づいて、画像の分割領域を形成する分割レイアウト 902 を決定する。

画素分類モジュール 206 は、画像に分割領域を適用したうえで、ブロック 904 に示すように、一つ以上の画素タイプに各分割領域内の画素进行分类する。

また、エビデンス生成モジュール 208 は、ブロック 906 に示すように、分割領域ごとに異なる、画素タイプのヒストグラムを生成する。

また、シーン分類モジュール 212 は、ブロック 908 に示すように、各々の分割領域のヒストグラムに基づいてシーン分類を実行する。例えば、シーン分類モジュール 212 は、所定のシーンのタイプのヒストグラムとの比較に基づいてシーン分類を行い、「山と

50

乾燥地形」というシーンを決定する。

【0135】

図10は、画素レベルの分類のもう一つの例を示すグラフィック表現1000である。

画像プロセッサ202は、オリジナル画像1002を、96ピクセル×54ピクセルの解像度にダウンサンプルすることができる。

画像内の車線に隣接するバリア1006aおよび1006bは、システムを学習させるための重要な特徴である。

画像1004は、ダウンサンプルされた画像に関連付いた画素タイプを示す。画像1004の中心から発散している線は、画像に適用されるくさび形の分割領域を表す。

また、画像1004の色（濃度）は、画素タイプを示す。例えば、領域1008は葉を表し、領域1010aおよび1010bはバリアを表す。

【0136】

以上の説明では、本発明を十分に理解できるように、多くの詳細について説明した。しかしながら、各実施形態はこれらの具体的な詳細無しでも良いことは当業者にとって明らかであろう。また、説明が不明瞭になることを避けるために、構造や装置をブロック図の形式で表すこともある。例えば、実施形態は、ユーザインタフェースおよび特定のハードウェアとともに説明される。しかし、ここでの説明は、データおよびコマンドを受信する任意のタイプの計算装置および任意の周辺機器について適用できる。

【0137】

本明細書における「1つの実施形態」または「別の実施形態」等という用語は、その実施形態と関連づけて説明される特定の特徴・構造・性質が少なくとも本発明の一つの実施形態に含まれることを意味する。「1つの実施形態における」等という用語は本明細書内で複数用いられるが、これらは必ずしも同一の実施形態を示すものとは限らない。

【0138】

本明細書の詳細な説明の一部は、非一時的（non-transitory）なコンピュータ可読記憶媒体に記憶されたデータビットに対する動作のアルゴリズムおよび記号的表現として提供される。これらのアルゴリズム的な説明および表現は、データ処理技術分野の当業者によって、他の当業者に対して自らの成果の本質を最も効果的に説明するために用いられるものである。なお、本明細書において（また一般に）アルゴリズムとは、所望の結果を得るための論理的な手順を意味する。処理のステップは、物理量を物理的に操作するものである。必ずしも必須ではないが、通常は、これらの量は記憶・伝送・結合・比較およびその他の処理が可能な電氣的または磁氣的信号の形式を取る。通例にしたがって、これらの信号をビット・値・要素・エレメント・シンボル・キャラクタ・項・数値などとして称することが簡便である。

【0139】

なお、これらの用語および類似する用語はいずれも、適切な物理量と関連付いているものであり、これら物理量に対する簡易的なラベルに過ぎないということに留意する必要がある。以下の説明から明らかなように、特に断らない限りは、本明細書において「処理」「計算」「コンピュータ計算（処理）」「判断」「表示」等の用語を用いた説明は、コンピュータシステムや類似の電子的計算装置の動作および処理であって、コンピュータシステムのレジスタやメモリ内の物理的（電子的）量を、他のメモリやレジスタまたは同様の情報ストレージや通信装置、表示装置内の物理量として表される他のデータへ操作および変形する動作および処理を意味する。

【0140】

本発明は、本明細書で説明される動作を実行する装置にも関する。この装置は要求される目的のために特別に製造されるものであっても良いし、汎用コンピュータを用いて構成しコンピュータ内に格納されるプログラムによって選択的に実行されたり再構成されたりするものであっても良い。このようなコンピュータプログラムは、コンピュータのシステムバスに接続可能な、例えばフロッピー（登録商標）ディスク・光ディスク・CD-ROM・MOディスク・磁気ディスクなど任意のタイプのディスク、読み込み専用メモリ（R

10

20

30

40

50

OM)、ランダムアクセスメモリ(RAM)、EPROM、EEPROM、磁気カード、フラッシュメモリ、光学式カード、電子的命令を格納するために適した任意のタイプの媒体などの、非一時的なコンピュータ可読記憶媒体に記憶される。

【0141】

発明の具体的な実施形態は、完全にハードウェアによって実現されるものでも良いし、完全にソフトウェアによって実現されるものでも良いし、ハードウェアとソフトウェアの両方によって実現されるものでも良い。好ましい実施形態は、ソフトウェアによって実現される。ここでソフトウェアとは、ファームウェア、常駐ソフトウェア、マイクロコードやその他のソフトウェアを含むものである。

【0142】

さらに、ある実施形態は、コンピュータが利用あるいは読み込み可能な記憶媒体からアクセス可能なコンピュータプログラムプロダクトの形態を取る。この記憶媒体は、コンピュータや任意の命令実行システムによってあるいはそれらと共に利用されるプログラムコードを提供する。コンピュータが利用あるいは読み込み可能な記憶媒体とは、命令実行システムや装置によってあるいはそれらと共に利用されるプログラムを、保持、格納、通信、伝搬および転送可能な任意の装置を指す。

【0143】

プログラムコードを格納・実行するために適したデータ処理システムは、システムバスを介して記憶素子に直接または間接的に接続された少なくとも1つのプロセッサを有する。記憶素子は、プログラムコードの実際の実行に際して使われるローカルメモリや、大容量記憶装置や、実行中に大容量記憶装置からデータを取得する回数を減らすためにいくつかのプログラムコードを一時的に記憶するキャッシュメモリなどを含む。

【0144】

入力/出力(I/O)装置は、例えばキーボード、ディスプレイ、ポインティング装置などであるが、これらはI/Oコントローラを介して直接あるいは間接的にシステムに接続される。

【0145】

システムにはネットワークアダプタも接続されており、これにより、私的ネットワークや公共ネットワークを介して他のデータ処理システムやリモートにあるプリンタや記憶装置に接続される。モデム、ケーブルモデム、イーサネット(登録商標)は、現在利用可能なネットワークアダプタのほんの一例である。

【0146】

最後に、本明細書において提示されるアルゴリズムおよび表示は特定のコンピュータや他の装置と本来的に関連するものではない。本明細書における説明にしたがったプログラムを有する種々の汎用システムを用いることができるし、また要求された処理ステップを実行するための特定用途の装置を製作することが適した場合もある。これら種々のシステムに要求される構成は、以下の説明において明らかにされる。さらに、本発明は、特定のプログラミング言語と関連づけられるものではない。本明細書で説明される本発明の内容を実装するために種々のプログラミング言語を利用できることは明らかであろう。

【0147】

実施形態の前述の説明は、例示と説明を目的として行われたものである。したがって、開示された実施形態が本発明の全てではないし、本発明を上記の実施形態に限定するものでもない。本発明は、上記の開示にしたがって、種々の変形が可能である。本発明の範囲は上述の実施形態に限定解釈されるべきではなく、特許請求の範囲にしたがって解釈されるべきである。本発明の技術に詳しい者であれば、本発明はその思想や本質的特徴から離れることなくその他の種々の形態で実現できることを理解できるであろう。同様に、モジュール・処理・特徴・属性・方法およびその他の本発明の態様に関する名前付けや分割方法は必須なものでもないし重要でもない。また、本発明やその特徴を実装する機構は異なる名前や分割方法や構成を備えていても構わない。さらに、当業者であれば、モジュール・処理・特徴・属性・方法およびその他の本発明の態様は、ソフトウェア、ハードウェア

10

20

30

40

50

ア、ファームウェアもしくはこれらの組合せとして実装できることを理解できるであろう。また、本発明をソフトウェアとして実装する場合には、モジュールなどの各要素は、どのような様式で実装されても良い。例えば、スタンドアローンのプログラム、大きなプログラムの一部、異なる複数のプログラム、静的あるいは動的なリンクライブラリー、カーネルロードダブルモジュール、デバイスドライバ、その他コンピュータプログラミングの当業者にとって既知な方式として実装することができる。さらに、本発明の実装は特定のプログラミング言語に限定されるものではないし、特定のオペレーティングシステムや環境に限定されるものでもない。以上のように、上記の本発明の説明は限定的なものではなく例示的なものであり、本発明の範囲は添付の特許請求の範囲にしたがって定められる。

【符号の説明】

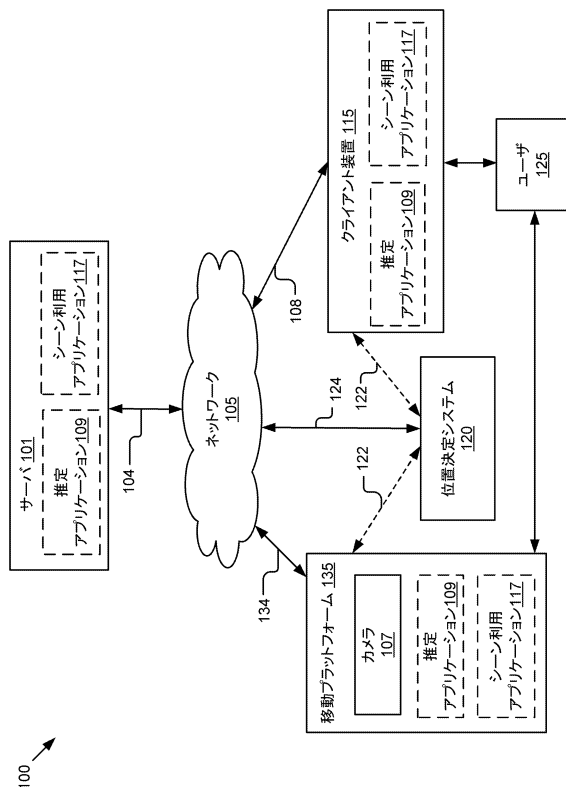
【 0 1 4 8 】

- 1 0 1 サーバ
- 1 0 7 カメラ
- 1 0 9 推定アプリケーション
- 1 1 5 クライアント装置
- 1 1 7 シーン利用アプリケーション
- 1 3 5 移動プラットフォーム
- 2 0 2 画像プロセッサ
- 2 0 4 分割モジュール
- 2 0 6 画素分類モジュール
- 2 0 8 エビデンス生成モジュール
- 2 1 0 学習モジュール
- 2 1 2 シーン分類モジュール

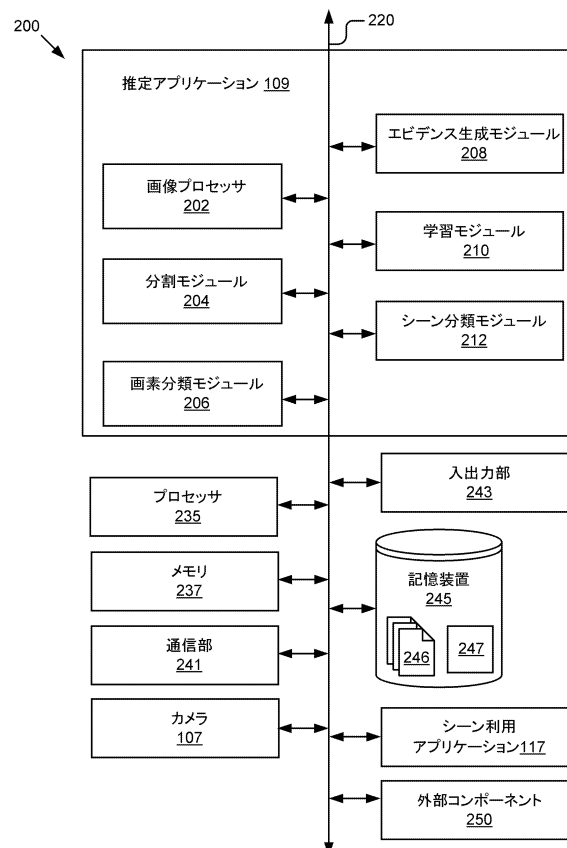
10

20

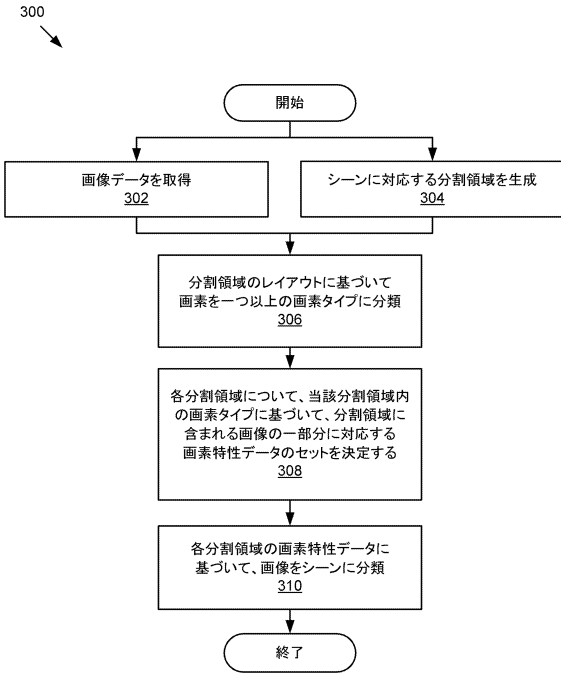
【 図 1 】



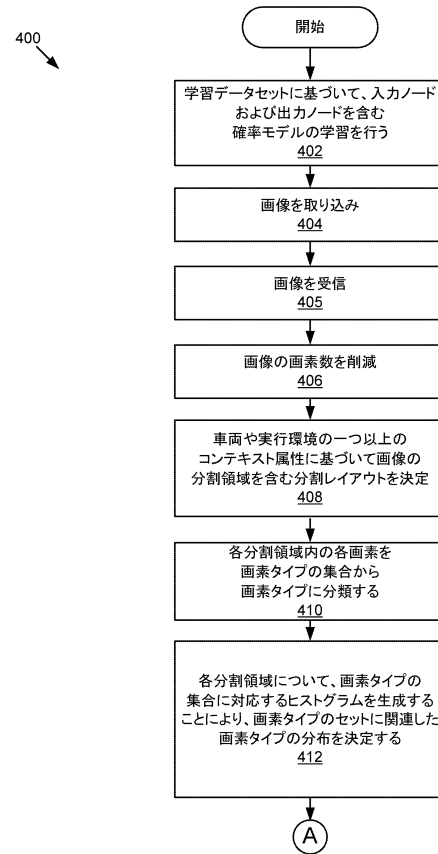
【 図 2 】



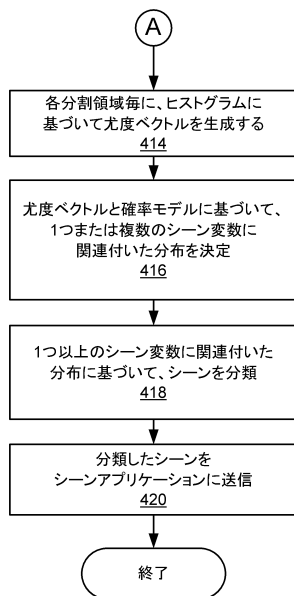
【図 3】



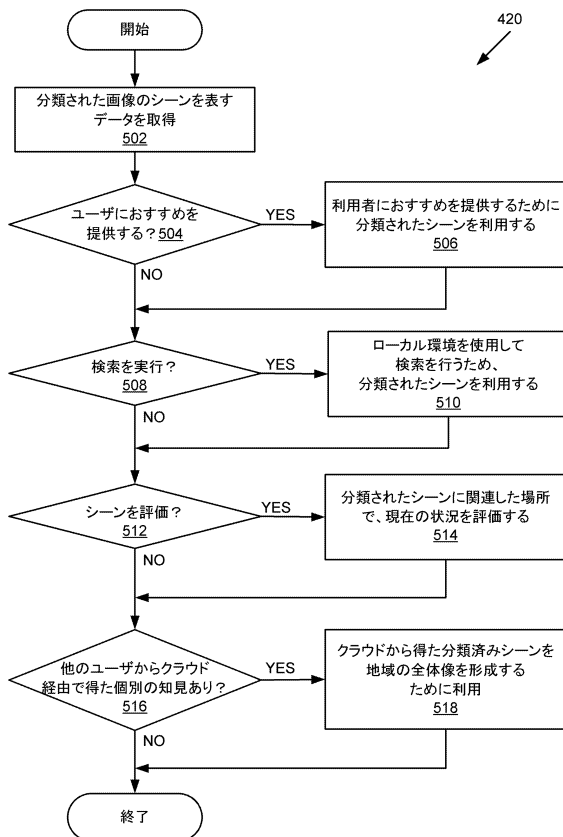
【図 4 A】



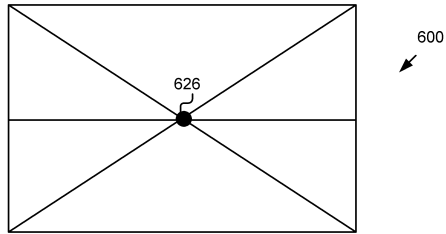
【図 4 B】



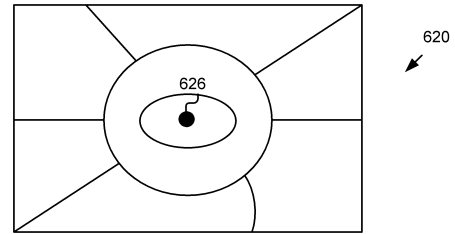
【図 5】



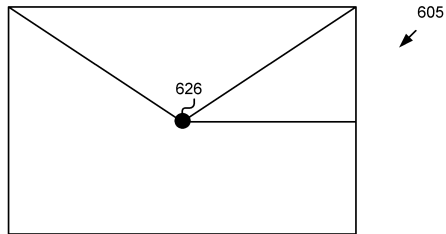
【図 6 A】



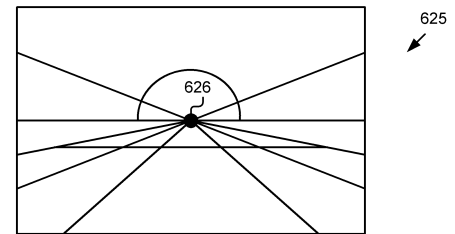
【図 6 D】



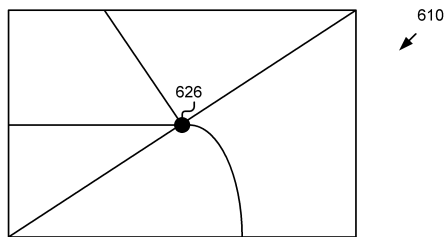
【図 6 B】



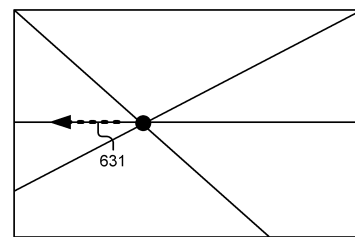
【図 6 E】



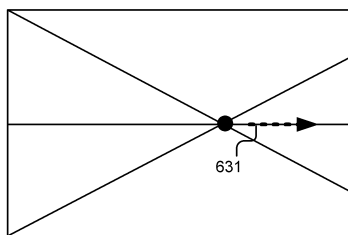
【図 6 C】



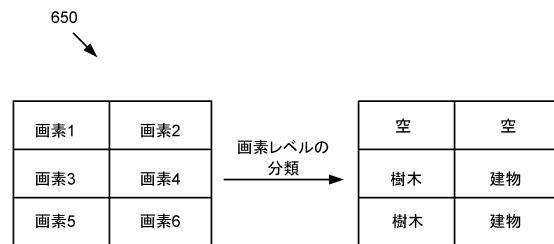
【図 6 F】



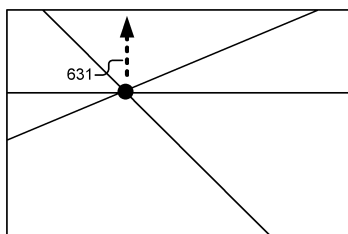
【図 6 G】



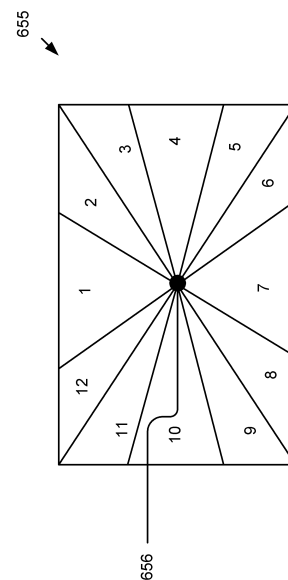
【図 6 J】



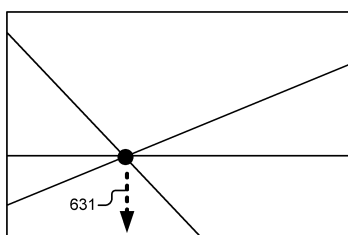
【図 6 H】



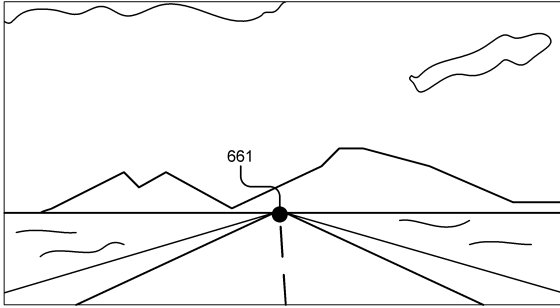
【図 6 K】



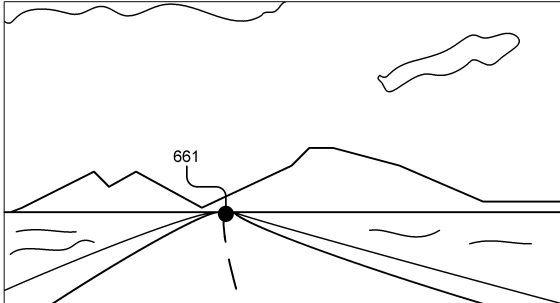
【図 6 I】



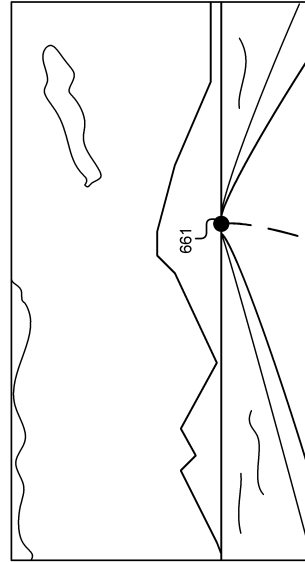
【図 6 L】



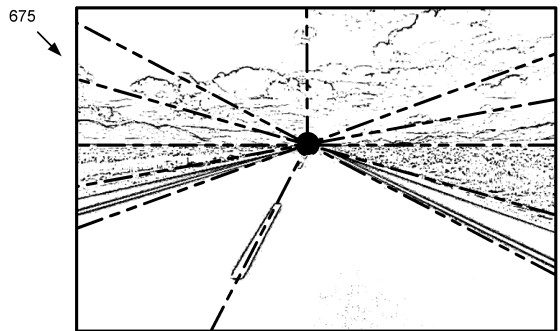
【図 6 M】



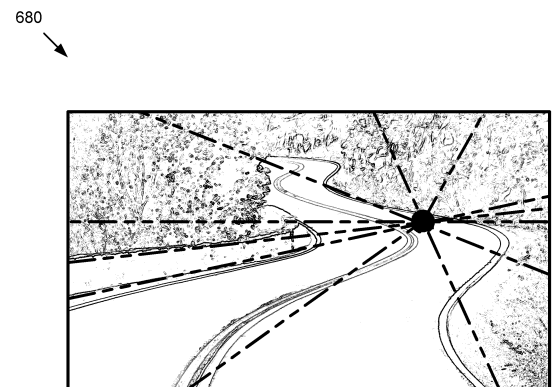
【図 6 N】



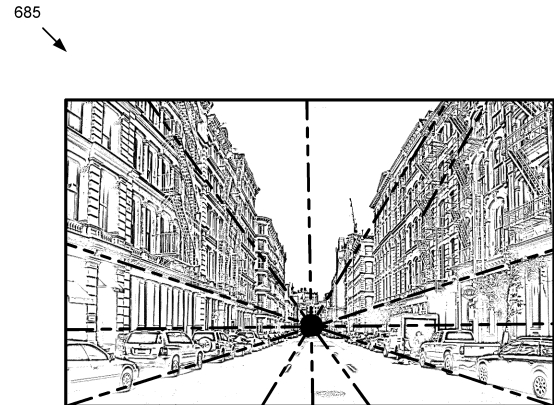
【図 6 O】



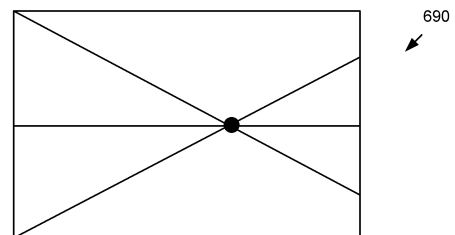
【図 6 P】



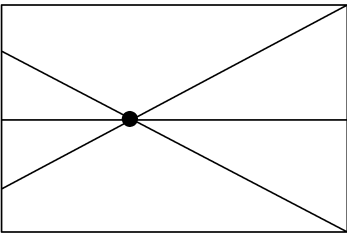
【図 6 Q】



【図 6 R】



【図 6 S】



695

【図 7 B】



	山	田園	住宅地	都市部
再現率	1.0	0.9	0.794	0.45
適合率	0.642	1.0	0.964	1.0
正確度				0.784

表 1 : 周辺環境

【図 7 A】



		予測ラベル 702		
		ラベル1	ラベル2	ラベル3
真ラベル 704	ラベル1	50	2	3
	ラベル2	4	60	7
	ラベル3	5	6	70

706

708

	カーブと 勾配	自動車 専用道	狭隘	路肩なし	駐車車両あり
再現率	1.0	0.75	0.0	0.85	0.56
適合率	0.61	1.0	NaN	1.0	1.0
正確度					0.770

表 2 : 道路

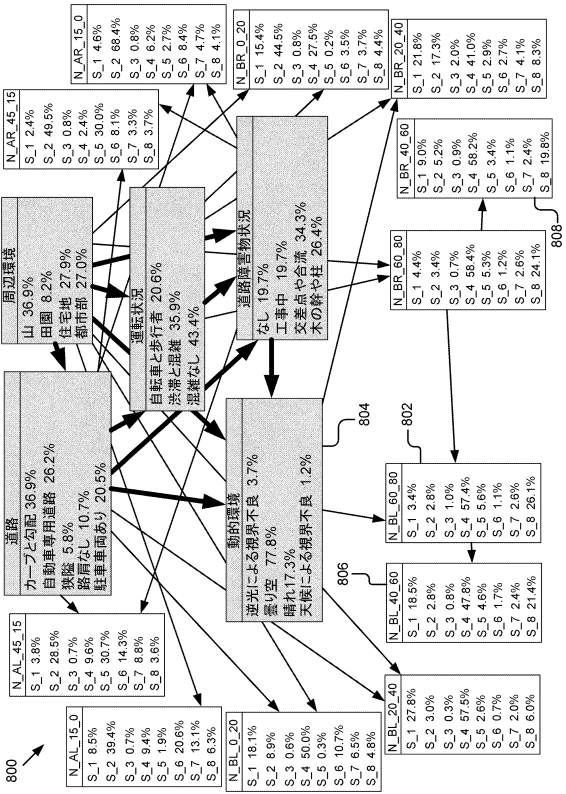
	自転車と 歩行者	渋滞と混雑	混雑なし
再現率	0.56	0.6	0.98
適合率	0.875	0.93	0.67
正確度			0.754

表 3 : 運転状況

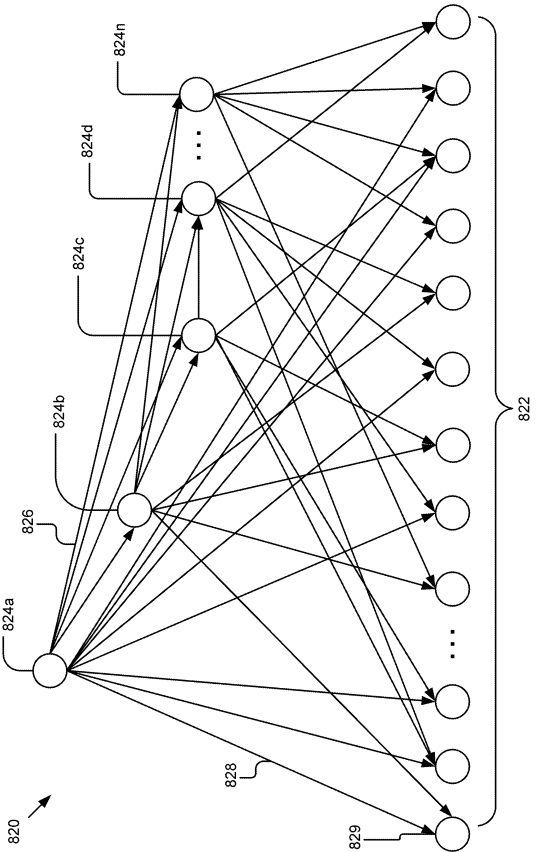
	なし	工事中	交差点や合流	木の幹や柱
再現率	0.42	0.96	0.6	1.0
適合率	1.0	0.92	0.86	0.55
正確度				0.738

表 4 : 道路障害物状況

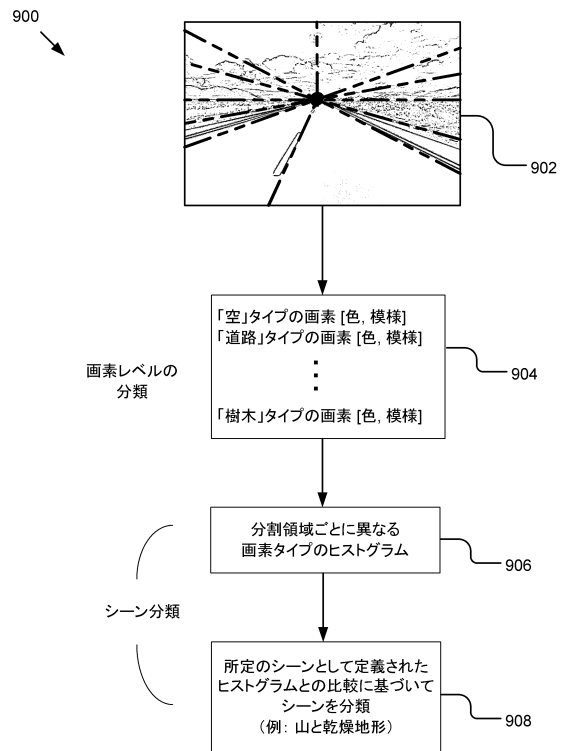
【図 8 A】



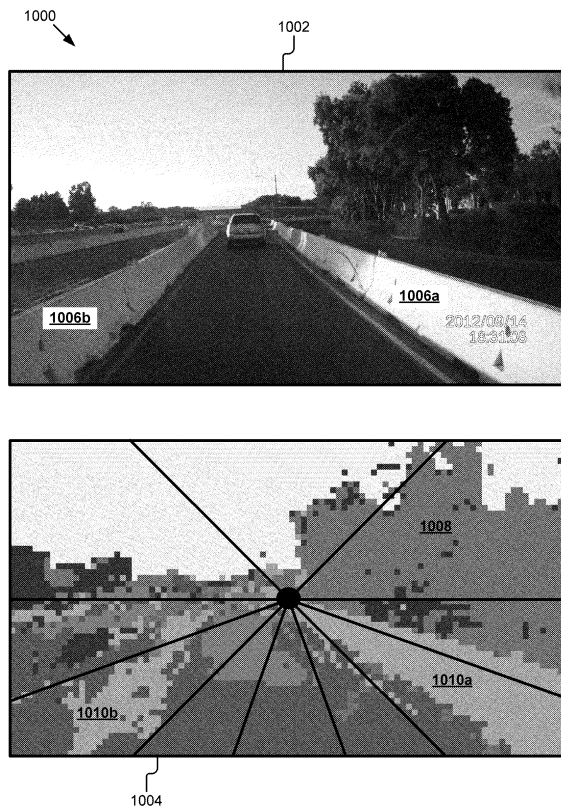
【図 8 B】



【図 9】



【図 10】



フロントページの続き

(74)代理人 100138357

弁理士 矢澤 広伸

(72)発明者 アゴスタ ジョン マーク

アメリカ合衆国 カリフォルニア州 9 4 3 0 6 パロアルト S . カリフォルニアアベニュー
1 6 4 8

(72)発明者 ピライ プリーティ

アメリカ合衆国 カリフォルニア州 9 4 0 8 6 サニーベール サウスベルナルドアベニュー
2 5 5 # 8

(72)発明者 尾口 健太郎

アメリカ合衆国 カリフォルニア州 9 4 0 2 5 メンロパーク シャロンパークドライブ 3 5
0 # I - 2 1

(72)発明者 ヤラ ガネッシュ

アメリカ合衆国 カリフォルニア州 9 4 0 4 3 マウンテンビュー ベルナルドアベニュー 4
6 5 トヨタ インフォテクノロジー センター , ユー . エス . エー . , インコーポレイテッド
内

審査官 佐藤 実

(56)参考文献 特開 2 0 0 7 - 0 7 2 9 8 7 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G 0 6 T 7 / 0 0 - 7 / 9 0