(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5936561号
(P5936561)
(45)発行日　平成28年6月22日(2016.6.22)　(24)登録日　平成28年5月20日(2016.5.20)

(51)Int.Cl.　F I
*G06T 7/00 (2006.01)*　G06T 7/00 350Z

請求項の数 9　外国語出願　(全 6 頁)

(21)出願番号　特願2013-25318(P2013-25318)
(22)出願日　平成25年2月13日(2013.2.13)
(65)公開番号　特開2013-206458(P2013-206458A)
(43)公開日　平成25年10月7日(2013.10.7)
審査請求日　平成27年12月2日(2015.12.2)
(31)優先権主張番号　13/432,998
(32)優先日　平成24年3月28日(2012.3.28)
(33)優先権主張国　米国(US)

早期審査対象出願

(73)特許権者　000006013
三菱電機株式会社
東京都千代田区丸の内二丁目７番３号
(74)代理人　100110423
弁理士　曾我　道治
(74)代理人　100094695
弁理士　鈴木　憲七
(74)代理人　100111648
弁理士　梶並　順
(74)代理人　100122437
弁理士　大宅　一宏
(74)代理人　100147566
弁理士　上田　俊一
(74)代理人　100161171
弁理士　吉田　潤一郎

最終頁に続く

(54)【発明の名称】画像における外観及びコンテキストに基づく物体分類

(57)【特許請求の範囲】
【請求項１】
画像内の物体を分類するための方法であって、
外観分類器を前記画像に適用するステップであって、前記物体の候補と、該候補と関連付けられた統計とを求め、前記外観分類器は一組の窓を用い、前記候補は選択された窓内にあるものと、
コンテキスト分類器を、前記画像の前記選択された窓に適用して、前記画像内の前記物体の識別情報及びロケーションを求めるものと、
を含み、前記ステップはプロセッサにおいて実行され、
前記コンテキスト分類器は、物体コンテキスト特徴及びグローバルコンテキスト特徴を用い、
前記物体コンテキスト特徴は、前記候補のサイズを全ての候補のサイズの平均で除算したものと、前記候補の信頼スコアを平均信頼スコアで除算したものと、該信頼スコアを最大信頼スコアで除算したものと、前記候補のロケーションを全ての候補の平均ロケーションで除算したものと、全ての他の候補からの平均距離を、一対の前記候補間の距離である候補ペアワイズ距離の平均で除算したものとを含む、前記候補間の相対情報を符号化する、画像内の物体を分類するための方法。
【請求項２】
前記統計は、前記候補と関連付けられたロケーション、サイズ、信頼スコアを含む、請求項１に記載の方法。

【請求項３】
前記外観分類器は、前記分類の誤検出誤り率を低減させる情報を出力する、請求項１に記載の方法。
【請求項４】
前記外観分類器及び前記コンテキスト分類器は、前記物体の様々なクラスについてトレーニングされる、請求項１に記載の方法。
【請求項５】
前記一組の窓は、複数のスケール、回転、及び平行移動用のものである、請求項１に記載の方法。
【請求項６】
前記グローバルコンテキスト特徴は、前記候補を含むシーン内の構造に関する情報を符号化する、請求項１に記載の方法。
【請求項７】
前記外観分類器及び前記コンテキスト分類器はサポートベクターマシンを用いる、請求項１に記載の方法。
【請求項８】
前記外観分類器は勾配のヒストグラムを用いる、請求項１に記載の方法。
【請求項９】
前記外観分類器はハールライク特徴を用いる、請求項１に記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
この発明は包括的にはコンピュータービジョンに関し、より詳細には、センサーによって捕捉された画像内の物体を分類することに関する。
【背景技術】
【０００２】
通常、カメラは、目視し解析することができる画像及びビデオよりも多くの画像及びビデオを捕捉する。したがって、画像及びビデオの解析を自動化する、コンピューターに基づく方法及びシステムがますます必要とされている。画像及びビデオの自動化された解析における基本的な作業は、カメラによって捕捉されたシーン内の様々なクラスの物体の識別及びロケーション特定である。
【０００３】
物体分類のための最も一般的な手法は走査窓を用いる。この手法では、窓が画像上で走査されるとともに、窓内のピクセルに分類器が適用される。通常、窓は長方形であり固定サイズである。分類器は、物体が窓内にあるか否かを示す。必要な場合、画像は物体に適合するようにサイズ変更することができる。サイズ変更は、サイズ変更された画像が窓と一致するか又は窓よりも小さくなるまで繰り返し行われる。この力まかせ探索は、全てのロケーション及びサイズにわたって繰り返される。この方法は、物体の各クラスについて繰り返すことができる。
【０００４】
これらの方法は事実上、窓内のピクセルから利用可能な外観情報しか用いない。しかしながら、そのような手法は、シーン内の様々な物体の、互いに対する相対的な外観及びレイアウトの双方に基づく構造情報、又は全体的な物体構造若しくはシーン構造に基づく事前値(priors)を利用することをしていない。
【０００５】
幾つかの方法はコンテキスト情報に基づく。これらの方法は、３Ｄシーン構造を用いて、物体が画像内に位置する尤度を推測する。例えば、車両上に配置されたカメラを用いて人物を検出するために、画像平面内の道路、歩道、及び建物のロケーションの知識を用いて、ロケーションごとの物体尤度を生成することができる。同様に、カメラの幾何学的形状及び画像の手がかりを用いて、画像からシーンの粗い３Ｄ構造を推測することができ、

これを用いて物体尤度を生成することができる。コンテキストによる物体尤度は、次に分類器スコアと組み合わせて、検出性能を改善することができる。しかしながら、シーン構造の取得は困難な作業であり、これらの方法の有効性を制限する。

【０００６】

代替的な方法は、画像内の物体の一部分を探索し、これらの部分を組み合わせて物体を検出及びロケーション特定する。これらの方法は、物体の一部分の外観に基づく分類と、幾何学的関係とを組み合わせて単一の分類器にする。同時分類の計算複雑度が高いことに起因して、物体の部分対間の距離等の単純な幾何学的関係しか並列に用いることができない。

【０００７】

別の方法は、「上方」、「下方」、「隣」、「上」等の、様々な物体の単純な空間関係に基づく特徴を用いる。これらの関係は、マルチクラス分類器内の外観特徴と組み合わされる。

【発明の概要】

【発明が解決しようとする課題】

【０００８】

より効率的に画像内の物体を分類する方法が、求められている。

【課題を解決するための手段】

【０００９】

外観分類器を画像に適用し、物体の候補と、この候補と関連付けられた統計とを求めることによって、画像内の物体が分類される。外観分類器は一組の窓を用い、候補は選択された窓内にある。

【００１０】

次に、コンテキスト分類器が、画像の選択された窓にのみ適用され、画像内の物体の識別情報及びロケーションが求められる。

【発明の効果】

【００１１】

この発明は、より効率的に画像内の物体を分類する方法を提供する。

【図面の簡単な説明】

【００１２】

【図１】この発明の実施の形態による分類方法の流れ図である。

【発明を実施するための形態】

【００１３】

図１に示すように、この発明の実施の形態は、センサー、例えばカメラによって捕捉された画像１０１内の１つ又は複数の異なる物体を分類するための方法を提供する。この明細書において用いられるとき、分類は、物体を識別及びロケーション特定することを意味する。この方法への入力は、他の撮像様式を用いた光学画像と、距離画像と、医学画像及び物理学画像とを含む任意の画像とすることができる。この方法は、外観分類器１１０及びコンテキスト分類器１３０を連続的に用いる。この方法のステップは、既知のメモリ及び入出力インターフェースに接続されたプロセッサ１００において実行することができる。

【００１４】

オフラインステップにおいて、外観分類器及びコンテキスト分類器をトレーニングする。外観分類器は、物体の特定のクラスを表す画像内のピクセルの「外観」を候補１１１として用いる。分類器は、物体の各クラスについてトレーニングすることができる。外観分類器の出力は、候補と関連付けられた様々なクラス、ロケーション、サイズ、スコアの統計１１２を含む。この分類器は、誤検出誤り率(a false positive error rate)を低減する他の情報も出力することができる。

【００１５】

コンテキスト分類器は、候補及び関連付けられた統計を用いる。この分類器は、物体の

各クラスについてトレーニングすることもできる。

【００１６】

外観分類器は、複数のスケール、回転、平行移動について一組の窓すなわち選択された窓に適用される。各窓について外観特徴が求められる。次に、選択された窓は、この窓が候補１１１を含む場合、「１」(正)として分類される。画像内の窓のロケーション、向き、及びサイズも与えられる。換言すれば、候補は、分類することができる物体を含む可能性が高い選択された窓である。

【００１７】

各候補について、物体コンテキスト特徴及びグローバルコンテキスト特徴１２１を抽出する(１２０)。物体コンテキスト特徴は、候補間の相対情報を符号化する。例えば、候補のサイズが全ての候補のサイズの平均によって除算される。グローバルコンテキスト特徴は、シーン内の構造又は画像構造に対する候補情報、例えば画像平面内の候補のロケーションを符号化する。

【００１８】

コンテキスト分類器は、コンテキスト特徴及び候補に適用される。コンテキスト分類器によって「１」(正)と分類された候補は、分類済み物体１３１としてラベル付けされ、例えばその物体の識別情報及びロケーションが知られる。

【００１９】

トレーニング

人又は自動車等の特定の物体クラスのための分類器は、任意の従来の技法を用いてトレーニングすることができる。本発明者らは、外観分類器及びコンテキスト分類器のためにサポートベクターマシン(ＳＶＭ)を用いる。分類器は、一組の正及び負のトレーニング例を必要とする二次計画法ソルバーを用いてトレーニングされる。一組の正及び負のトレーニング例は、データベース内のトレーニング画像の集合から選択された物体窓及び非物体窓である。トレーニングされた分類器は、実数値の信頼スコアを出力する。この信頼スコアをしきい値処理して二値出力(０／１)を得ることができる。

【００２０】

外観特徴

窓ごとに様々な外観特徴ベクトルを求めることができる。本発明者らは、勾配ヒストグラム(ＨｏＧ)特徴と、ハールライク特徴との組み合わせを用いた。ヒストグラムは候補の外観統計を表す。

【００２１】

ＨｏＧは画像の上に重なる長方形であり、この長方形内の各ピクセルについて勾配の大きさ及び向きが求められる。０度～３６０度の範囲において向きが離散した一組のビンが定義される。各ビンは連続した範囲の向き(角度)をカバーする。各ピクセルの勾配の大きさが、勾配の向きに対応する適切なビンに加えられる。これによって、特徴ベクトルとして用いられるヒストグラムが得られる。通常、検出窓内で、そのような長方形エリアの重なり合わない組が定義され、各長方形エリアのＨｏＧ特徴ベクトルを連結したものは、分類のための特徴ベクトルとして用いられる。

【００２２】

ハールライク特徴は、画像の上に重なる２つ以上の連続した長方形であり、この画像について、各長方形内のピクセルの値が合算され、定数(多くの場合、１又は－１)を乗算され、次に組み合わされる(加算される)。ハールライク特徴は、積分画像法として知られる画像表現を用いて非常に高速に計算される。

【００２３】

コンテキスト特徴

物体コンテキスト特徴

物体コンテキスト特徴は、候補間の相対関係を定量化する任意の尺度である。本発明者らは、サイズを平均候補物体サイズで除算したものと、スコアを平均スコアで除算したものと、スコアを最大スコアで除算したものと、ロケーションを候補の平均ロケーションで

除算したものと、全ての他の候補からの平均距離を、候補ペアワイズ距離の平均で除算したものとを用いる。

【００２４】

グローバルコンテキスト特徴

候補と全体としてのシーン又は画像との間の関係を定量化する任意の尺度を、グローバルコンテキスト特徴として用いることができる。例えば、画像平面内のロケーション、ピクセルのサイズを用いる。

【００２５】

発明の効果

この発明は、画像内の１つ又は複数の異なる物体を分類するための方法を提供する。分類は、外観分類器及びコンテキスト分類器によって逐次行われる。外観分類器は候補物体を選択し、それによってコンテキスト分類器の計算が大幅に低減される。

【００２６】

外観分類器はＨｏＧ特徴又はハールライク(Haar-like)特徴を用いる。これらは求めるのが比較的容易である。

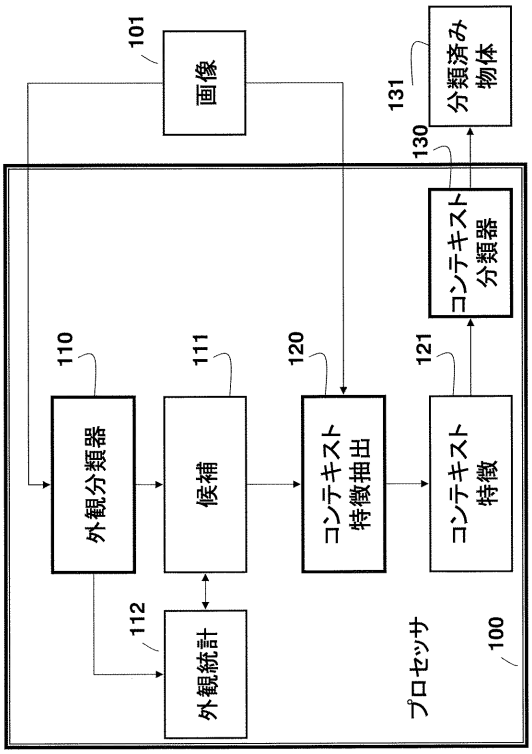
【００２７】

双方の分類器が、外観特徴、及び相対特徴、及びグローバルコンテキスト特徴を用いて、誤検出誤り率を一桁低減する。

【００２８】

本発明者らのコンテキスト分類器は、外観分類器によって分類された小さな候補組にのみ適用され、特徴をＯ(ｎ)時間(ここで、ｎは候補数である)で求めることができるので、非常に効率的である。

【図１】



フロントページの続き

(74)代理人 100161115
弁理士 飯野 智史
(72)発明者 オンセル・チュゼル
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ヒューロン・アヴェニュー ７００、アパートメント １７エム
(72)発明者 ジェイシャンカー・ケイ・ピライ
アメリカ合衆国、メリーランド州、カレッジ・パーク、ロウォルト・ドライブ ４３１４、アパートメント ２０２

審査官 佐田 宏史

(56)参考文献 特開２００７－１９３４０４（ＪＰ，Ａ）
特開２００３－０５７１９３（ＪＰ，Ａ）
特開２００４－０６２７１９（ＪＰ，Ａ）
C. Galleguillos et al., "Object Categorization using Co-Occurrence, Location and Appearance", IEEE Conference on Computer Vision and Pattern Recognition, 2008. , 米国, IEEE , ２００８年 6月２３日, pp.1-8
C. Galleguillos et al., "Context based object categorization: A critical survey", Computer Vision and Image Understanding, NL, Elsevier Inc. , ２０１０年 3月 1日, Vol.114, No.6, pp.712-722, ＵＲＬ, http://www.cs.princeton.edu/courses/archive/spr11/cos598A/pdfs/Galleguillos10.pdf
P. Viola et al., "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001., 米国, IEEE, ２００１年１２月 8日, Vol.1, pp.511-518
Xinxiao Wu et al., "Action Recognition using Context and Appearance Distribution Features", 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, 米国, IEEE, ２０１１年 6月２０日, pp.489-496

(58)調査した分野(Int.Cl., ＤＢ名)
Ｇ０６Ｔ ７／００－７／６０