

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号  
特許第5588395号  
(P5588395)

(45) 発行日 平成26年9月10日 (2014. 9. 10)

(24) 登録日 平成26年8月1日 (2014. 8. 1)

(51) Int. Cl.

F I

G O 6 T 1/00 (2006. 01)

G O 6 T 1/00 2 0 0 E

請求項の数 8 (全 18 頁)

(21) 出願番号	特願2011-106382 (P2011-106382)	(73) 特許権者	502096543
(22) 出願日	平成23年5月11日 (2011. 5. 11)		パロ・アルト・リサーチ・センター・イン
(65) 公開番号	特開2011-248876 (P2011-248876A)		コーポレーテッド
(43) 公開日	平成23年12月8日 (2011. 12. 8)		P a l o A l t o R e s e a r c h
審査請求日	平成26年5月8日 (2014. 5. 8)		C e n t e r I n c o r p o r a t e d
(31) 優先権主張番号	12/788, 852		アメリカ合衆国、カリフォルニア州 94
(32) 優先日	平成22年5月27日 (2010. 5. 27)		304、パロ・アルト、コヨーテ・ヒル・
(33) 優先権主張国	米国 (US)		ロード 3333
早期審査対象出願		(74) 代理人	110001210
			特許業務法人 Y K I 国際特許事務所
		(72) 発明者	ブラティーク・シャーカー
			アメリカ合衆国 カリフォルニア州 94
			087 サニーヴェイル カードレーン・
			ウェイ 992
			最終頁に続く

(54) 【発明の名称】 画像をオブジェクト及びそのパーツに関して効率的に解釈するためのシステムと方法

(57) 【特許請求の範囲】

【請求項 1】

画像を解釈する方法であって、

因子分解可能な上限を有する目的関数を最適化する、複数変数の構成を発見することを含み、因子分解可能とは、前記上限が、非相互作用パーツの和または非負の相互作用パーツの積の形式に操作処理される、または操作処理され得ることを意味するものであり、

前記発見が、構成候補として既知の要素の効率的な動的順序づけに依存する反復アルゴリズムを、優先待ち行列における上限スコアの降順で適用することによって行われ、

前記反復アルゴリズムは、

a . 優先待ち行列 (優先 Q ) を確立することであって、候補の優先順位は上限 (候補) と共に上がるものと、

b . カレント候補値を確立し、かつこれを ( 0 , 0 . . . , 0 ) に設定することと、

c . 最良候補値を確立し、かつこれを ( 0 , 0 . . . , 0 ) に設定することと、

d . 優先 Q に後続候補 (カレント) の全要素を挿入することであって、初期要素値 ( 0 , 0 . . . , 0 ) に後続候補関数を適用して後続候補を得るとともに、当該初期要素値の後続候補に再び前記後続候補関数を適用し、さらにこれにより得られた後続候補にも前記後続候補関数を適用して、全ての関連データを生成するような、後続候補関数が定められ、

e . 最良候補値が満足解であるか否かについて問合せを行うことと、

ステップ e が正しければ、最良候補を返すことと、  
ステップ e が正しくなければ、( i ) 優先 Q の先頭における候補をカレント候補値へ割り付けかつこれを優先 Q から外し、( i i ) 優先 Q に後続候補 ( カレント ) の全エレメントを挿入し、かつこの後、( i i i ) 真のスコア ( 最良候補 ) > 真のスコア ( カレント ) であるかどうかを決定することであって、真のスコアは前記目的関数の実際の値であって、次に、

ステップ e の ( i i i ) が真であれば、ステップ e の開始へ戻ることと、  
ステップ e の ( i i i ) が真でなければ、カレント候補値における値を包含するように最良候補を更新し、ステップ e の開始へ戻ることを含み、

当該方法は、少なくとも 1 つのプロセッサを用いたシステムにより行われる方法。

10

#### 【請求項 2】

前記反復アルゴリズムは、実行中に、その時点で入手可能な最良の完全解を得るために中断されてもよい任意時間アルゴリズムである、請求項 1 に記載の方法。

#### 【請求項 3】

前記反復アルゴリズムは、適切な基準に基づいて完了前に停止され、当該適切な基準は、限定されないが、カレントスコア及び前記上限の最大値としきい値との比較を含む、請求項 1 に記載の方法。

#### 【請求項 4】

前記目的関数は第一項と第二項とを含み、前記第一項はオブジェクトまたはシーンの個々のパーツの特性を記述し、かつ前記第二項はこれらのパーツ間の関係及び / または制約を記述するものである、請求項 1 に記載の方法。

20

#### 【請求項 5】

前記目的関数の第一項は、前記オブジェクトまたはシーンの個々のアピランモデルを明確に記述するものである、請求項 4 に記載の方法。

#### 【請求項 6】

前記目的関数の第 2 項は、前記オブジェクトまたはシーンのパーツ間の空間的關係モデルを記述するものである、請求項 4 に記載の方法。

#### 【請求項 7】

前記目的関数は星座モデルより与えられる、請求項 1 に記載の方法。

#### 【請求項 8】

30

画像を解釈する方法であって、

因子分解可能な上限を有する目的関数を最適化する、複変数の構成を発見することを含み、因子分解可能とは、前記上限が、非相互作用パーツの和または非負の相互作用パーツの積の形式へと操作処理される、または操作処理され得るものを意味するものであり、前記目的関数は、第一項及び第二項を含み、前記第一項はオブジェクトまたはシーンの個々のパーツの特性を記述し、かつ前記第二項はこれらのパーツ間の関係及び / または制約を記述するものであり且つオブジェクトまたはシーンの個々のパーツの空間的關係モデルを記述するものであり、前記空間的關係モデルは、2 つのパート間の距離を含み、

前記発見が、構成候補として知られるエレメントの効率的な動的順序づけに依存する反復アルゴリズムを、優先待ち行列における上限スコアの降順で適用することによって行われ、

40

既に与えられたカテゴリのオブジェクトまたはそのようなオブジェクトの一部が画像中に存在するか否かを決定する目的で画像解釈を呼び出し、

当該方法が、少なくとも 1 つのプロセッサを用いたシステムにより行われる方法。

#### 【発明の詳細な説明】

#### 【技術分野】

#### 【0001】

本出願は画像の解釈に関し、より具体的には下記、即ち、  
 局在化：画像内の特異的な既知のオブジェクトまたは特異的な既知のカテゴリの任意のオ

50

プロジェクトを発見し、かつそのオブジェクトの個々のパーツを位置決めすること。ここで、オブジェクトは形状（企業ロゴ等）であっても、車または顔面等の自然物であってもよい、

検出：画像が特異的な既知のオブジェクトまたは特異的な既知のカテゴリの任意のオブジェクトを含んでいるかどうかを検査すること、

カテゴリ分類：画像に、そのコンテンツを基礎としてカテゴリラベルを割り付けること、に関する画像解釈の問題点を解決することに関する。

#### 【背景技術】

##### 【0002】

オブジェクトまたはシーンカテゴリのパーツベースモデルは、技術上周知である。しかしながら、関連の「bag-of-features」モデル等のモデルが画像のタグ付け及びカテゴリ分類にうまく使用されてきているものの、オブジェクト表現の重大部分は、オブジェクト特徴間の空間的関係にあるという長い間の認識がある。空間的関係のモデリングは、オブジェクトの認識、カテゴリ分類及び局在化のパフォーマンスを向上させる。空間的関係を用いるモデルは、オブジェクト及びそのパーツに関する自動画像解釈に使用されることも可能である。この後者の課題、即ちオブジェクトパーツの効率的な局在化は、それ自体で関心対象である。これは、目等のオブジェクトパーツの中には、あまりに重要であるために内在的な関心対象オブジェクトであるとされ得るものがあることに起因する。

##### 【0003】

空間的関係のモデリングは、モデル変数間に依存性をもたらす。これは、このようなモデルの使用を計算量的に困難なものにするが、その理由は、複数の変数に対して目的関数を同時に最大化しなければならないことにある。例えば、パーツのアピアランス及び空間的関係の双方を組み込む周知のモデルファミリである星座モデルについて考察されたい。簡単に言えば、このモデルは、特徴的なアピアランスを有するK個のパーツ集合から成る。例えば、顔面のカテゴリの場合、パーツは目、鼻及び口である可能性がある。よって完全な顔面は、個々のパーツを局在化することによって検出される。

##### 【0004】

オブジェクトのパーツが独立していて、各パーツが画像内にN個のロケーション候補を有していれば、N個の構成候補の「優良度」を評価することによって各パーツの最良ロケーションを発見することができ、合計KN個の評価を行ってオブジェクトが検出される。しかしながら、パーツを独立したものとしてモデリングする結果、局在化の質が悪くなる。

#### 【発明の概要】

#### 【発明が解決しようとする課題】

##### 【0005】

個々のパーツの局在化を向上させるためには、その典型的な空間的配置に関する情報をモデリングすることが有益である。このような幾何学的情報のモデリングは、局在化パフォーマンスを大幅に向上させる。しかしながら、これもまた、パーツ間の依存性を招来する。例えば、顔の場合、モデリングされ得る情報は顔面の対称性または典型的なパーツ間距離である。しかしながらこのモデルでは、目のパーツとして所定のロケーションを仮定すると、鼻が発見されると思われるロケーション集合に影響が出る。その結果、もはやパーツを独立的に画像へ嵌め込むことはできない。代わりに、全パーツの接合ロケーションとしての $N^K$ 個の可能性について最適構成が探索される必要がある。これはKの指数関数であることから、この探索は、パーツ数の増加に伴ってすぐに処理しにくいものとなる。具体的には、7個のパーツ間の空間的関係のモデリングが事実上の限界とされている場合が多い。従って、パーツ間に依存性があるモデルを最適化する新しい方法を発見することが望ましい。

#### 【課題を解決するための手段】

##### 【0006】

本出願は、構成候補の効率的な動的順序づけに依存する反復アルゴリズムを優先待ち行列における上限スコアの降順で適用することにより、因子分解可能な上限を有する目的関数を最適化する複数変数の構成を発見して画像を解釈する方法及びシステムである。一例として、オブジェクトの星座モデルを考察されたい。これは、オブジェクトの個々のパーツ、並びにこれらのパーツ間の空間的関係のアピアランスモデルを明示する。これらは、その値が画像内のオブジェクトの類似を表す単一関数へと結合される。オブジェクトが存在する構成を画像内に発見するために、我々はこの関数を全ての構成候補に渡って最大化する。上述の反復アルゴリズムの目的は、このような最適構成を効率的に発見することにある。

【図面の簡単な説明】

10

【0007】

【図1】本発明の画像解釈システム及び方法のコンピュータ実装を示す概観である。

【図2】星座モデルを用いた出力の一例である。

【図3】本出願によるプロセスのフロー図である。

【図4】局在化の概観である。

【図5】顔認識の詳細である。

【図6】車両認識の詳細である。

【図7】使用したパーツ数を基礎とするパフォーマンスの向上を示す。

【図8】本発明の概念の使用により達成される速度の改善を示す。

【図9】複数の反復の効果を示す。

20

【発明を実施するための形態】

【0008】

図1は、本出願のコンセプトが実装されている環境100を示す。環境100は、ネットワーク102及び通信チャネル104を介して複数のコンポーネント間の通信を提供する。環境100内のコンポーネントには、カテゴリ分類システム106と、局在化システム108と、オブジェクト認識システム110と、全関数評価システム112と、パーツ間関係システム114と、パーツアピアランスシステム116と、因子分解可能上限システム118とが含まれる。これらのシステムは各々、別々のプロセッサ及びコンピュータメモリを有しかつ適切なソフトウェアを実行する別々のコンピュータ操作可能サーバ上に存在してもよい。またこれらは、クラウドコンピューティング環境へ1つのエレメントとして付着されてもよく、またはサーバ120等の同じサーバ上に位置決めされてもよい。各サーバは、通信リンク104を介して、ウェブ対応コンピュータ122、(画像生成能力を有する)電話124または他のウェブ対応携帯デバイス126等の、但しこれらに限定されない、画像を送受信することができる様々なデバイスの任意のものへ接続されてもよい。またサーバ120は、電子的に格納された画像128を含むコンピュータデータベースとして編成されてもよい、コンピュータメモリで構成される情報レポジトリとして考慮される場合もある。或いは、本出願の全体システムは、スタンドアロンコンピューティングデバイス上に位置決めされてもよい。デバイス122-126は、後述する画像解釈オペレーションを達成するような方法で、システム106-118との間で画像を送受信してもよい。但し、これらのオペレーションの様々なステップはシステム106-118によって達成される。

30

40

【0009】

次に、図2を参照すると、星座モデルを顔200、オートバイ210及び自動車220へ適用したプロセスの結果が示されている。星座モデルは、コンピュータビジョンの分野では周知であり、コンピュータビジョンにおけるカテゴリレベルのオブジェクト認識のための確率モデルである。パーツベースの他のモデルと同様に、星座モデルは、オブジェクトのクラスを相互的な幾何学的制約下にあるK個のパーツの集合によって表現しようとする。星座モデルは、異なるパーツ間の幾何学的関係を考慮することから、ある特定のオブジェクトカテゴリについて、これらのパーツの相対的ロケーション、相対的スケール及びアピアランスを明示的にモデリングする。

50

## 【0010】

星座モデルは2つのコンポーネント、即ちアピアランスモデルと、空間的關係モデルとを含む。本出願の実施形態の9例は、3つの異なるアピアランスモデル及び3つの異なる空間的な先験的確率にスパンする。

## 【0011】

星座モデルが強調されているが、より一般的には、因子分解可能な上限を有する限り、星座モデルのうちの1つ以外の目的関数を使用される可能性もある。次の論考においてさらに詳述するが、因子分解可能とは、上限が何らかの算術演算または他の演算によって非相互作用パーツの合計または非負の相互作用パーツの積の形式へと操作される、または操作され得ることを意味する。

10

## 【0012】

従って、コンピュータビジョンの文献ではオブジェクトまたはシーンカテゴリのパーツベースモデルが一般的であり、かつ関連の「bag-of-features」モデルが画像のタグ付け及びカテゴリ分類にうまく使用されてきているが、オブジェクト表現の重大部分は、オブジェクト特徴間の空間的關係にあるという長い間の認識がある。空間的關係のモデリングは、オブジェクトの認識、カテゴリ分類及び局在化のパフォーマンスを向上させる。空間的關係を用いるモデルは、オブジェクト及びそのパーツに関する自動画像解釈を呼び出すために使用されることも可能である。この後者の課題、即ちオブジェクトパーツの効率的な局在化は、それ自体で関心対象である。これは、図2から分かるように、目201、202、鼻204、206及び口203、205等のオブジェクトパーツの中には、あまりに重要であるために内在的な関心対象オブジェクトであるとされ得るものがあることに起因する。

20

## 【0013】

一例として、多くの異なる形状、サイズ及びデザインを有するオートバイ（具体的なオートバイは、オートバイ210である）を考察されたい。画像内でのオートバイの認識は困難である。しかしながら、オートバイのパーツは、車輪211、212及びヘッドライト213等のより一貫したアピアランスを有する傾向がある。従って、オートバイの認識に取り組む一方法は、車輪211、212、ヘッドライト213及び予測される相対的な空間的關係性にあるパーツを探すことである。同様の例は、車輪221、222、ドア223及びヘッドライト224を有する自動車220であろう。この認識は、星座モデルが実行しようとするものである。この場合、車輪及びヘッドライトといえども変化に富んでいることから、その検出も完全ではない。典型的には、あらゆるロケーションが車輪の中心であるという幾分かの類似性を有するが、大部分のロケーションでこの類似の値は小さい、という複数の仮説が検討されなければならない。各パーツの複数の仮説は、最も可能性の高い構成に対する組合せ探索に問題を生じさせる。

30

## 【0014】

従って、本出願は、星座モデルの目的関数を最適化するという上述の問題点に適用される、複数の離散変数の関数を最適化するための一般的アルゴリズムを提供する。目的関数の因子分解可能な上限を用いることにより、本出願のアルゴリズムは、優先待ち行列を動的に成長させて $N^K$ 個の候補を効率的に暗に順序づけする。ソリューション候補は、優先待ち行列から1つずつ検索され、その候補の真の目的関数が計算され、次いで待ち行列からまだ抽出されていない全ての候補の上限と比較される。これは、後に説明するように、因子分解可能性を活用することによって効率的に実行されることが可能である。

40

## 【0015】

未探索候補の全ての上限を超える真のスコアを有する候補が識別されると、これは、これまでで最良の候補が最適ソリューションであることを保証する。我々は、実際には、近似または制限モデルによらない場合でも、本方法が代替の最適化スキームより数倍も効率的であることを実証している。

## 【0016】

コンピュータビジョン及びパターン認識における他の多くの問題点は、離散変数に渡る

50

関数の最大化としての枠内に入れることができる。K個の変数が存在し、その各々がN個の値のうちの1つをとるとすれば、このような最大化は、 $N^K$ 個の可能な組合せに渡る探索になる。本アルゴリズムは、このような多くの問題点に対する確実な最適ソリューションを効率的に発見するために適用されてもよい。一般的な要件は、目的関数上で因子分解可能な上限が発見されることである。

【0017】

次に、以下の論考では、複数のパーツ及びこれらのパーツ間の空間的關係によってモデリングされる最適化アルゴリズム及びオブジェクトロケーションへのその適用について述べる。この場合、最適化アルゴリズムはパーツのアピランス及びパーツ間距離（より一般的には、距離に加えて、または距離の代わりに角度、共直線関係、他等のものを含んでもよいこれらのパーツの空間的構造）を受信することに言及しておく。以下の説明は、専ら星座モデルについて行うが、記述されるべき最適化アルゴリズムは一般的であり、僅かな修正によって追加の問題点に適用され得ることは認識されるべきである。

10

【0018】

核心的な最適化アルゴリズムは、K個の離散変数、 $l_1, \dots, l_2, \dots, l_K$ に渡って最適化を考慮する。

【数1】

$$l_1^*, l_2^*, \dots, l_K^* = \arg \max_{l_1, \dots, l_K} f(l_1, \dots, l_K)$$

20

但し、 $l_i$ は離散変数をとる。各 $l_i$ は、異なる値集合をとることができる。単純化を期して、各々が0からN-1までのN個の値をとるものと想定する。これは、目的関数 $f()$ が実行可能領域より外部であるようにして定義されるという理解によって最適化と呼ばれる。可能なソリューションの数は、 $N^K$ である。

【0019】

関数 $g(l_1, \dots, l_K)$ は、目的関数 $f()$ 上の上限であり、即ち、 $(l_1, \dots, l_K) \cdot g(l_1, \dots, l_K) = f(l_1, \dots, l_K)$ である。可能な $N^K$ 個のソリューションが全て、優先待ち行列においてその上限スコアの降順で配列されることを想像されたい。最も高い上限を有するソリューション候補（優先待ち行列の先頭または先頭エレメント）が選択され、その真のスコア $f()$ が評価される。この真のスコアが待ち行列における次のエレメントの上限 $g()$ を超えていれば、これは、あらゆる未探索ソリューション候補の上限及び真のスコアも超える。従って、最適ソリューションが発見される。そうでなければ、システムは最良候補発見の追跡を継続し、プロセスが反復される。ループ終了時点で探索された最良のソリューションが、我々の最適ソリューションであるものとして保証される。

30

【0020】

優先待ち行列には、 $N^K$ 個の候補が配列されたままであり、よってシステムは依然として高い計算複雑度を負っていることに留意されたい。この負担は、上限が因子分解可能であれば軽減されることが可能である。

40

【0021】

関数 $g(l_1, \dots, l_K)$ は、 $g()$ を非相互作用パーツの和（または積）として表すことができれば、 $f()$ の因子分解可能な上限である。

【数2】

$$g(l_1, \dots, l_K) = \sum_{i=1}^K h_i(l_i) \geq f(l_1, \dots, l_K) \quad (1)$$

50

すると、

【数 3】

$$h_i(l_i = a) \geq h_i(l_i = b) \\ \Rightarrow g(\dots, l_i = a, \dots) \geq g(\dots, l_i = b, \dots)$$

になる。

10

【0022】

因数分解可能という言い回しは、本出願の概念が主として倍数因子を成分として有する確率関数に対処するものであるために使用されている。本明細書において記述している概念は、 $i$  個に渡る和が正項の積に置換されても同じく効力を発する。

【0023】

これは、上限によってソートされる優先待ち行列を動的に編成しかつこれを必要に応じで成長させるための我々のアルゴリズムに繋がる。単純化を期して、かつ一般性を失うことなく、 $h_i(a)$  は、 $a$  が 0 から  $N - 1$  まで増大するにつれて低減するものと仮定する。(この仮定が成り立たなければ、我々は各  $i$  における候補を  $h_i$  でソートし、逆写像のためのソート - 指数アレイを保持する。我々は次に、このソーティングにより提供された候補の新しい順序づけを使用する。この新しい順序づけでは、上述の仮定の成立が保証される。) 明らかに、 $(0, 0, \dots, 0)$  が最も高い上限を有する候補であり、後続候補  $(1, 0, \dots, 0)$ ,  $(0, 1, \dots, 0)$ ,  $\dots$ ,  $(0, 0, \dots, 1)$  のうちの 1 つがこれに続く。後続候補の集合は、カレント候補を取り上げ、各位置指数を一度に 1 だけ増分することによって取得されることが可能である。従って、我々は各位置における指数を次に最良である可能性のある値で置換する。例えば、 $K = 4$ 、及び考慮対象の候補を  $(2, 3, 0, 1)$  であるとすると、第 1 のパーツは 3 番目により選択肢であり、第 2 のパーツは 4 番目により選択肢であり、第 3 のパーツは最良の選択肢であり、かつ第 4 のパーツは 2 番目により選択肢である。この配列下では、この候補の後続候補は  $(3, 3, 0, 1)$ 、 $(2, 4, 0, 1)$ 、 $(2, 3, 1, 1)$  及び  $(2, 3, 0, 2)$  である。優先待ち行列へ追加されるのがこれらの後続候補のみであれば、それが既に待ち行列内に存在していて未だ考慮されていないものであるか、これらの 4 つの直近後続候補のうちの 1 つであるかのいずれかで、次に高い上限を有する候補が既に待ち行列内に存在することが確信される。これにより、待ち行列を動的に成長させることができるようになり、かつ全ての候補を前もってソートする必要がない。候補は、常時、その後続候補より高い上限を有することが保証される。

20

30

【0024】

この観測は、次の擬似コードをもたらす。

```
priority Q = { } ;
current = ( 0 , 0 , ... , 0 ) ;
best = current ;
priority Q . add ( successors ( current ) ) ;
while not ( f ( best ) > g ( head ( priority Q ) ) ) {
    current = head ( priority Q ) ;
    remove_head ( priority Q ) ;
    priority Q . add ( successors ( current ) ) ;
    if not ( f ( best ) > f ( current ) )
        best = current ;
}
```

40

50

return best ;

【 0 0 2 5 】

このように、優先待ち行列は、探索が進むにつれて動的に成長される。アルゴリズムがソリューションを返すと、これは最適なソリューションを提供する。最も高度な探索アルゴリズムの場合のように、このアルゴリズムの最悪の複雑度は  $N^K$  であり、実際に、待ち行列を保持しかつ上限を計算するオーバーヘッドに起因して単純な全数検索よりも悪い。しかしながら、実際には、空間の大部分が明確に探索されることはなく、このアプリケーションにおけるアルゴリズムは極めて効率的になる。この経験的効率は、上限に依存する。不良ソリューションに対して上限が厳密であるほど、実際に探索される空間は少なくなる。

10

【 0 0 2 6 】

注目すべき詳細は、もう1点ある。ソリューション候補  $(2, 3, 0, 1)$  (この場合、 $K = 4$ ) について考察されたい。これは、全てより高い上限を有することが保証される3つの先行候補、 $(1, 3, 0, 1)$ 、 $(2, 2, 0, 1)$  及び  $(2, 3, 0, 0)$  を有する。全ての先行候補が調査されれば、 $(2, 3, 0, 1)$  は待ち行列に三度追加される可能性がある。僅かなブックキーピングの追加は、この不必要なオーバーロードを回避する。後続候補リストは、一度に1位置を、 $i = 1$  から開始して最初の非ゼロ位置まで増分することによって計算される。従って、後続候補  $(0, 0, 2, 1) = \{(1, 0, 2, 1), (0, 1, 2, 1), (0, 0, 3, 1)\}$  になる。これは、各候補が一意的な先行候補を有することを保証し、かつアルゴリズムの実行時間効率を向上させる。理想的には、ある候補は、その最も低い上限の先行候補を除外する場合にのみ待ち行列に入れられるべきであるが、これを実施する計算オーバーヘッドは任意の潜在的節減分を超過する場合がある。

20

【 0 0 2 7 】

この反復アルゴリズムのアプリケーション例として、何らかのオブジェクトカテゴリの星座モデルについて再度考察されたい。例えば、このカテゴリは人間の顔であって、目、鼻及び口等のパーツを使用することができる。概して、特徴的なアピアランスを有する  $K$  個のパーツが存在する。これらのパーツの幾何学的構成は、ベクトル  $L = [l_1, \dots, l_k]$  によって表される。但し、 $l_i = (x_i, y_i)$  はパーツ  $i$  のロケーションである。

30

【 0 0 2 8 】

画像  $I$  を所与とすると、 $I$  におけるパーツの構成を記述するためにロケーションベクトル  $L$  が決定されてもよい。これを行うために、システムは、画像内のパーツ構成  $L$  の確率、 $p(L | I)$  を最大化する。これは、

【数 4】

$$p(L|I) \propto p(L)p(I|L). \quad (2)$$

40

と表すことができる。

【 0 0 2 9 】

第一項  $p(L)$  は、「先行空間」と呼ばれる。第二項 (「アピアランス項」と呼ばれる) は、下記のようにモデリングされることが多い。

【数 5】

$$p(I | L) = \prod_i p(I[l_i] | l_i, A_i) \quad (3)$$

50



## 【 0 0 3 0 】

ここで、 $I[l_i]$  は、画像  $I$  における  $l_i$  を中心とした画像パッチのアピアランスを表す。これは、そのロケーションから抽出された生のピクセル値である可能性もあれば、スケール不変特徴変換 (SIFT) 記述子によって与えられるもの等のアピアランスの他の何らかの表現である可能性もある。 $A_i$  は、 $i$  \* 番目のパーツのアピアランスを符号化するモデルパラメータの集合である。例えば、 $I[l_i]$  が画像パッチの SIFT 記述子であれば、 $A_i$  はトレーニング記述子の平均及び共分散である可能性がある。

## 【 0 0 3 1 】

この因子分解は、画像内にロケーションを与えられている各パーツのアピアランスが他の全てのパーツのロケーション及びアピアランスとは独立しているものと仮定している。この仮定は、必ずしも厳密には成り立たない。例えば、顔の特徴のアピアランスは人の性別によって影響され、よって、1つのパーツの女性的アピアランスは、別のパーツが男性的に見える確率に影響を与える。この仮定は、パーツ同士が重なる場合にも破られる(但し、何れにしても、多くの先行空間はこのような構成を考慮から外す)。しかしながら、概して、この近似は実際には広く使用され、かつ D. Crandall、P. Felzenszwalb 及び D. Huttenlocher 共著「統計モデルを用いるパーツベース認識のための先行空間」CVPR、2005年、及び T. K. Leung、M. C. Burl 及び P. Perona 共著「ランダムラベルグラフマッチングを用いるクラッシーン内の顔の発見」ICCV、1995年、に示されているように有益な結果を与えることができる。上記文献は共に、本参照によりその全体が開示に含まれる。

## 【 0 0 3 2 】

従って、完全なモデルは下記のようなになる。

## 【数 6】

$$p(L|I) \propto p(L) \prod_i p(I[l_i]|l_i, A_i) \quad (4)$$

## 【 0 0 3 3 】

この場合もやはり、局在化のタスクは、この式を最大化するパーツ構成  $L = |l_1, \dots, l_k|$  を発見することである。

## 【 0 0 3 4 】

上述の星座モデルの最適化は、目的関数が(対数形式で)、

## 【数 7】

$$f(l_1 \dots l_K) = \log p(l_1 \dots l_K) + \sum_i \log p(I[l_i]|l_i, A_i)$$

である本出願のアルゴリズム要件と容易に結びつく。

## 【 0 0 3 5 】

本来の因子分解は、

## 【数 8】

$$h_i(l_i) = \log p(I[l_i]|l_i, A_i)$$

である。

$0 \leq \log p(l_1, \dots, l_k)$  であることに留意されたい。従って、上限はアピアラ

ンスモデルによって与えられ、かつ完全なスコアまたは真のスコアは、これに先行空間の対数を加算することによって得られる。

#### 【 0 0 3 6 】

この定式化の下では、本アルゴリズムは良き直観的説明を有する。探索は、パーツモデルに対する局所応答に従って順序づけされる。画像内の各パーツに対する最良の応答は、より強い仮定は先行空間の結合時に拒絶されていることからより弱い適合を考慮してまず検査される仮説を構成する。星座モデルは確率的であるが、本アルゴリズムは、確率ではなくエネルギー関数を基礎とする類似モデルに等しく適用可能である。

#### 【 0 0 3 7 】

図 3 を参照すると、先に述べた、かつ先に開示した擬似コードのオペレーションを反映する最適化アルゴリズムのフロー図 3 0 0 が提示されている。諸ステップは、複数のパーツによってモデリングされたオブジェクトを位置決めしかつこれらのオブジェクト間の空間的関係を識別するためのアルゴリズムの適用方法を明示している。

#### 【 0 0 3 8 】

図 3 のフロー図 3 0 0 は、ステップ 3 0 5 において、候補の優先順位が上限（候補）に伴って上がる優先待ち行列（優先 Q）を確立する。上限（）関数は、[ 0 0 2 1 ] における  $g()$  と同じであり、式 1 のように計算される。ステップ 3 1 0 において、プロセスは、 $(0, 0, \dots, 0)$  に設定されるカレント候補値も確立し、かつステップ 3 2 0 において、 $(0, 0, \dots, 0)$  に設定される最良候補値も確立する。次にステップ 3 3 0 において、プロセスは優先 Q に後続候補（カレント）の全エレメントを挿入する。任意の候補の後続は、[ 0 0 2 4 ] 及び [ 0 0 2 7 ] において説明したように計算される。この時点で、ステップ 3 4 0 において、真のスコア（最良候補） $>$  上限（先頭（優先 Q））であるかどうかを決定するための検査によって反復ループが始まる。我々の表記法では、真のスコア（）は  $f()$  と同じである。ステップ 3 4 0 における検査が成功すれば（YES、であれば）、ステップ 3 5 0 においてプロセスは最良候補を返す。しかしながら、ステップ 3 4 0 における検査が失敗すれば（NO、であれば）、ステップ 3 6 0 においてプロセスは優先 Q の先頭における候補をカレントに割り付け、これを優先 Q から外す。次に、ステップ 3 7 0 において、プロセスの反復ループは優先 Q 内に後続候補（カレント）の全エレメントを挿入し、この後、ステップ 3 8 0 において真のスコア（最良候補） $>$  真のスコア（カレント）であるかどうかを決定する別の検査へ移動する。ステップ 3 8 0 における検査が成功すれば（YES、であれば）、プロセスはステップ 3 4 0 に戻って継続される。ステップ 3 8 0 における検査が失敗であれば（NO、であれば）、ステップ 3 9 0 においてプロセスはカレントの値を包含して最良候補を更新し、次いでプロセスはステップ 3 4 0 に戻ってさらなる処理を行う。

#### 【 0 0 3 9 】

図 4 は、4 つの星座モデル（顔モデルが 2 つ、オートバイモデルが 1 つ及び車モデルが 1 つ）の最適構成（即ち、パーツロケーション）を発見するための提案アルゴリズムの適用を実証する一連の検査結果を提示している。モデルパラメータは、カテゴリラベル及びパーツロケーションを用いて手動で注釈された画像（即ち、画像テンプレート）の監督された学習を介してトレーニングされることが可能である。このようなモデルを自動的に学習する代替方法は、例えば F e r g u s 他著「非監督スケール不変学習によるオブジェクトクラスの認識」C V P R、2 0 0 3 年 6 月、W e b e r 他著「オブジェクトカテゴリの自動的発見へ向けて」C V P R、1 0 1 - 1 0 9 ページ、2 0 0 0 年、及び W e b e r 他著「認識用モデルの非監督学習」E C C V、1 8 - 3 2 ページ、2 0 0 0 年、において論じられているように容易に利用可能である。上述の文献は各々、その全体が開示に含まれる。

#### 【 0 0 4 0 】

星座モデルのセットアップに注目すると、先に述べたように、星座モデルは 2 つの要素、即ちパーツアピアランスモデル及び空間的關係モデルから成る。異なる種類のアピアランスモデル及び先行空間がトレーニングされ、これらは混合されかつ合わされて幾つかの

10

20

30

40

50

星座モデルが作られる。本アルゴリズムは、これらのモデルによって評価される。これを行うに当たっては、アルゴリズム擬似コードの探索（図3のフローチャートも参照されたい）のために、使用されるモデルの種類によって変更を行う必要はないことが留意される。実際、その優位点は探索が最適ソリューションを与えることにあり、パフォーマンスが不良であれば、変更を要するのはモデルの方である。

#### 【0041】

本明細書に記述している検査では、3つのアピアランスモデルが使用される。そのうちの1つでは、各画像ロケーションがスケール不変特徴変換（SIFT）記述子（Lowe著「スケール不変キーポイントからの弁別的画像特徴」、IJCV、60（2）：91-110、2004年、において論じられているもの等。上記文献は、本参照によりその全体が開示に含まれる）によって、ロケーション周囲の $41 \times 41$ ピクセル近傍内に表現される。各パーツの尤度 $p(I[l_i] \mid l_i, A_i)$ は、モデルのトレーニングに使用される画像から推定されるパラメータ $\mu_i$ 、 $\sigma_i$ を用いてガウス分布としてモデリングされる。

10

#### 【0042】

別のアピアランスモデルでは、画像ロケーションのアピアランスが生みのピクセル値によって $41 \times 41$ ピクセル近傍内に与えられる。パーツ $A_i$ は、マーキングされたパーツロケーション付近の近傍パッチであり、かつアピアランススコア $p(I[l_i] \mid l_i, A_i)$ は、正規化された相互相関 $NCC(A_i, I[l_i])$ として計算される。この定式化では、 $p(I[l_i] \mid l_i, A_i)$ は適切な確率分布ではなく、表記法 $h(I[l_i] \mid l_i, A_i)$ がより適切である。しかしながら、確率分布だけでなく任意関数に適用可能であることから、最適化アルゴリズムは変わらない。

20

#### 【0043】

第3のアピアランスモデルでは、ある画像ロケーションにおけるアピアランスが90個の異なるハールフィルタの応答として $41 \times 41$ ピクセル近傍内に符号化される。アピアランススコア $h(I[l_i] \mid l_i, A_i)$ は、ランダムな画像ロケーションから真のパーツロケーションのアピアランスを見分けるようにトレーニングされたアダプスト分類器の応答として計算される。ハール-アダプストモデルの場合、画像全体に渡る均一な格子点がパーツロケーション候補（ほぼ全ての画像において $N > 1000$ 個のロケーション）として使用されるが、ハリス-ラプラス関心点（Mikolajczyk他著「アフィン領域検出器の比較」IJCV、65（1/2）：43-72、において論じられているようなもの。上記文献は、本参照によりその全体が開示に含まれる）は、最初の2個（ $N > 100$ ）におけるロケーション候補として処理される。

30

#### 【0044】

上述の実験では、 $41 \times 41$ ピクセルの近傍が使用されたが、これはモデル要件ではなく、任意サイズの近傍（非正方形及び非長方形の近傍を含む）を代わりに使用できることに留意されたい。

#### 【0045】

異なる先行空間について、追加実験を行った。1つの先行空間 $p(L)$ は、パーツ間の対間距離を用いてモデリングされた。K個のパーツでは、このような距離が $K(K-1)/2$ 個存在する。ロケーション $l_i$ 及び $l_j$ 間の距離を $d_{i,j}$ で表し、かつ全距離 $[d_{1,2}, \dots, d_{K-1,K}]$ のベクトルを $D$ で表して、このモデルでは、 $p(L)$ を $p(D(L))$ に比例するものと定義する。実験において、 $p(D(L))$ は、対角共分散行列を有するガウス分布であった。この分布のパラメータは、トレーニング例から学習された。

40

#### 【0046】

別の先行空間は上述のものに類似するものであったが、全距離が $l_1$ 及び $l_2$ 間の距離でスケールされる点異なる。その目的は、スケール不変を達成することにある。

#### 【0047】

第3のタイプの先行空間は顔画面のみに対処し、対称先行候補と呼ばれる。これは、顔

50

の軸を中心にして互いの鏡となっていなければ、左右のパーツロケーションにペナルティを与える。本明細書において、これは、我々の対称先行候補を任意に複雑にすることが可能であって、何れにしても因子分解可能である必要はないことの例証として包含されている。

#### 【0048】

実験の一部として、図4に示すような3種のデータセット、即ち、顔のデータセット410、420、オートバイのデータセット430及び車のデータセット440が使用された。顔画像のバイオIDデータセット(410, 420)は、Jesorsky他著「ハウズドルフ距離を用いるロバストな顔検出」Audio and Video Based Person Authentication、2001年、において論じられているように、23個のサブジェクトの1521個の顔画像から成る。上記文献は、本参照によりその全体が開示に含まれる。各画像は、パーツ名及びロケーションで完全に注釈を付けられる。顔の表情及び照度には、著しい変形が存在する。オリジナル画像にもスケール変化があり、よって全ての画像は、顔がほぼ同じサイズを有するようにスケールし直されている。

10

#### 【0049】

さらに、Caltech101データセットからのオートバイ及び車側面のビューが検査されている。Caltech101データセットは、Fei-Fei他著「少数訓練例からの生成的ビジュアルモデルの学習：101オブジェクトカテゴリ上でテスト済みのインクリメンタルなBayesアプローチ」CVPR 2004 Workshop on Generative-Model Based Vision、2004年、において論じられている。上記文献は、本参照によりその全体が開示に含まれる。この場合やはり、各カテゴリ内の画像は、特定のオブジェクトが全画像を通じてほぼ同じサイズになるようにスケールし直された。バイオIDの顔モデルのトレーニングには、約100個の画像が使用され、Caltechのオートバイ及び車については各々30個のトレーニング画像が使用された。

20

#### 【0050】

「妥当な」モデルが使用されていることを確立するために、2つの標的課題、即ちオブジェクト検出及びオブジェクト局在化に定性的なパフォーマンス測度が設けられる。これらのタスクは、段落[0001]において定義されている。

30

#### 【0051】

第1のタスクに関しては、受信者動作特性(ROC)曲線を用いてパフォーマンスが測定される。カテゴリ毎に、我々のアルゴリズムを用いて星座モデルの最良一致スコアが発見される。これは、対応するカテゴリ内の各検査画像について、さらには、「背景」画像(Caltech101「背景」集合からの50画像)についても行われる。「背景」画像は、負の手本として使用される。ここから、受容しきい値を変えることによってROCプロットが計算される。

#### 【0052】

受信者動作特性(ROC)曲線または単にROC曲線は、2値分類器システムのその弁別しきい値の変更に伴う感度対(1 - 特異度)をグラフにプロットするための信号検出理論における概念である。ROCは、真陽性の割合(TPR = 真陽性率)対偽陽性の割合(FPR = 偽陽性率)をプロットすることによって等しく表現されることが可能である。これは、基準の変更に伴う2つの動作特性(TPR及びFPR)の比較であることから、相対動作特性曲線としても知られる。ROC分析は、コストコンテキストまたはクラス分布から独立して(またはこれらを明示する前に)最適モデルである可能性のあるものを選択しかつ最適以下のものを放棄するためのツールを提供する。

40

#### 【0053】

第2のタスクは、パーツの局在化である。この場合、モデルパーツの最良ロケーションは、我々のアルゴリズムによって発見される最適なモデル構成に対応するものとして識別される。これらのロケーションは、全てのカテゴリ内検査画像について記録され、評価の

50

ために、手動で注釈を付けられた「グラウンドトゥース」ロケーションと比較される。一例として、推測されたロケーションと手動で注釈を付けられた「グラウンドトゥース」ロケーションとの間の平均ユークリッド距離を計算して、パーツの局在化の精度を定量化することができる。

【 0 0 5 4 】

図 4 には、幾つかの実験結果が提示されている。ヒストグラム 4 5 0 には、幾つかのモデルの局在化誤差がプロットされている。また、検出パフォーマンスを明示した R O C 曲線 4 6 0 も提示されている。

【 0 0 5 5 】

図 5 は、各々 6 個、1 0 個及び 1 4 個のパーツでモデルを構築するバイオ I D の顔データセット 5 0 0 による実験結果を詳述している。6 パーツモデル 5 1 0 では、目の瞳 5 1 1、5 1 2、2 つの鼻孔 5 1 3、5 1 4 及び左右の口角 5 1 5、5 1 6 が使用されている。1 0 パーツモデル 5 2 0 では、眉毛の 4 つの端点 5 2 1、5 2 2、5 2 3、5 2 4 が追加されている。1 4 パーツのモデルでは、システムは 4 つの目の端を使用してもよい。

【 0 0 5 6 】

図 6 は、2 つの車輪の中心 6 3 1、6 3 2、フロントライト及びテールライト 6 3 4 及び座席の前側 6 3 5 及び後側 6 3 6 の 6 パーツを使用するオートバイのデータセット 6 3 0 を用いた結果を提示している。

【 0 0 5 7 】

また、2 つの車輪の中心 6 4 1、6 4 2、タイヤガードの内角 6 4 3、6 4 4、フロントライト及びテールライト 6 4 5、6 4 6、フロントガラス輪郭の最低ポイント 6 4 7 及びルーフの 2 つの末端 6 4 8、6 4 9 の 9 パーツを使用する車のデータセット 6 4 0 を用いた結果が提示されている。

【 0 0 5 8 】

これらの例は、各カテゴリ内の個々のオブジェクトのアピアランス及びジオメトリの変化に関わらず、パーツが如何に高精度で局在化されるかを実証している。

【 0 0 5 9 】

図 7 は、各々 6 パーツ、1 0 パーツ及び 1 4 パーツによる 3 つの顔モデル 7 0 0、7 3 0、7 2 0 の顔検出パフォーマンスを明示している。全てのモデルがバイオ I D データセットからの画像上でトレーニングされて検査され、かつ対間距離ガウス先行候補を使用している。R O C 曲線は、誤警報率 ( X 軸 ) 対 ヒット率 ( Y 軸 ) をプロットしている。

【 0 0 6 0 】

図 8 は、6 パーツ ( 8 1 0 )、1 0 パーツ ( 8 2 0 ) 及び 1 4 パーツ ( 8 3 0 ) を基礎とする最適ソリューションを提示している。主たる結果は、最適性を保証しながらの提案している探索アルゴリズムの効率である。本アルゴリズム 8 4 0 が探索する平均構成数は、s t a r モデル 8 5 0、2 - f a n モデル 8 6 0 及び全数検索 8 7 0 が探索する構成数と比較される。2 - f a n がより限定的なモデルを要するとしても、アルゴリズム 8 4 0 の効率は経験的に 2 - f a n をも凌ぐことに留意されたい。結果は、使用されたモデル及び検査データセット毎に編成されている。

【 0 0 6 1 】

学習されたモデル毎に、大域的最適化を発見するためにアルゴリズムが行った平均反復回数が示されている。各反復は、1 つのパーツ構成の評価に対応する。比較のために、s t a r モデル 8 5 0 及び 2 - f a n モデル 8 6 0 が実行する評価回数が提示されている。s t a r モデル 8 5 0 及び 2 - f a n モデル 8 6 0 は限定空間を探索し、よってその表現力は完全な先行空間による星座モデルよりも弱いことに留意されたい。2 - f a n モデルによる最適構成は、完全な星座モデルでは必ずしも最適とはならない。これに対して、提案している最適化アルゴリズムは、完全な星座モデルの大域的最適化を確実に発見する。

【 0 0 6 2 】

また、可能性のある全ての構成数も記されているが、3 または 4 以上のパーツに関してこの数字が実際には実行不可能であることは知られている。図から分かるように、本アル

10

20

30

40

50

ゴリズム／擬似コード（図3のフローチャートも参照されたい）が行う評価は、より限定的なモデルである2-fanモデルが行う回数よりさらに少ない。これは、計算効率も向上させながら、確率分布 $p(L|I)$ をより精確にモデリングすることを可能にする。

#### 【0063】

図9は、反復回数が画像内のパーツロケーション候補の数にどのように依存するかを明示している。また、X軸910上のパーツロケーション候補（例えば、キーポイント）の数、Y軸920上のパーツロケーション候補の数の関数としての我々のアルゴリズムによる反復（探索される構成）の数の散布図900、3-fanの反復数に関する対数0及び6パーツによるフル探索もプロットされている930。最悪のケースで我々のアルゴリズムにより探索された構成の数は、フル探索及び3-fanより何倍も望ましいものである。散布図を調べると、本アルゴリズムに関して幾つかの洞察が得られる。画像が、そのモデルのカテゴリのオブジェクトを含んでいれば、アルゴリズムは極めて少ない反復の後に終了することが多い。これらは、パーツアピランスモデルが正しい候補に高得点をつける状況を反映している。画像が背景カテゴリに属していれば、パーツモデルは偽のロケーションでマッチングし、明確な勝者はほとんどいない。従って、勝利する構成を宣言するまでには、さらに多くの構成を探索する必要がある。

#### 【0064】

当然ながら、実際には、上限が受容しきい値を下回ると、探索を停止することによって背景画像は迅速に拒絶される。最終的には、「カテゴリ内」の画像に関して探索が百万回以上もの反復を実行する幾つかのケースも存在する。通常、これらの状況では最適ソリューションがごく早期に発見されるが、アルゴリズムは、これが実際に最適であることを確認できるまで多くの代替ソリューションを調査しなければならない。その最適構成が正しいものであれば、これは、先行空間モデルが正しい構成に過度に不利益を与えていることを含意する。最適構成が間違っただけであれば、これは、アピランスモデルに問題があることを指す場合が多い。このパターンは、星座モデルに関して、我々のアルゴリズムがモデルの目的関数の分析及びデバッグに有益である可能性を示唆している。

#### 【0065】

我々のアルゴリズムの計算効率により有効化される当然の功績として、星座モデル内のパーツ数は増大される。文献では、適用される最大の星座モデルは7パーツから成る。我々は、10パーツ及び14パーツ（バイオIDデータセット上）を有する星座モデルを適用している。

#### 【0066】

大部分の先行研究における代替先行空間は、ガウス分布によってモデリングされている。これは、探索プロセスの幾つかの最適化を許容するものであった。例えば、探索時間を評価数 $N^{k+1}$ から $N^k$ へ短縮するために、ガウス分布の簡易な関数形式が使用された。また、分布の特性は、優先待ち行列の探索を高速化するためにも使用された。

#### 【0067】

本出願のアルゴリズムは、先行空間に凸状、まして連続性等の制約を与えない。本方法の多用性を例証するために、本明細書ではこれが、顔パーツの空間的關係が対称性を用いてモデリングされるモデルに適用される。モデルには、顔の右側からの各パーツに対して、左側の対応するパーツも包含される。対称性の先行候補は、次のように計算される。まず、各反復において（部分的ソリューションではなく）全てのパーツロケーションが発生することから、最良の顔軸 - 両目及び両唇端の中点を結ぶ線、が計算される。次に、各パーツペア（例えば、左鼻孔、右鼻孔）を結ぶ各線分について、顔軸がこれを二等分していなければ、ペナルティが累算される。

#### 【0068】

本アルゴリズムは、この新しい星座モデルを効率的に最適化することができる。本アルゴリズムは、先行分布が容易な分析形式を有していなくても、モデルを首尾良く最適化することができる。

#### 【0069】

本方法は、パーツ間の空間的関係を有するカテゴリモデルの最適化を効率的に実行する。本方法は、星座モデルの周知のファミリに適用され、これらのモデルの計算効率を数倍も向上させることが示された。またこれは、より多くのパーツによるモデルの使用によってカテゴリ分類パフォーマンスの向上を可能にすることから、この高速化が実際に有益であることも実証している。

#### 【 0 0 7 0 】

本アルゴリズムは、許容的ヒューリスティックを有する  $A^*$  探索の概念に幾分関連する探索を実行する。しかしながら、典型的な  $A^*$  探索では、探索ノードは部分解（子孫葉ノードの全集合）を表す。従って、各ノードで計算されるべきヒューリスティック関数は全ての派生解に対する上限でなければならず、有界関数の全体階層が利用可能であることを要する。本アルゴリズムでは、各反復において、解（葉）ノードが考慮される。従って、目的関数は任意に複雑である可能性があり、必要なものは1つの因子分解可能な上限が全てである。星座局在化実験において、本出願のこの「反復アルゴリズム」は、多様で複雑な先行空間にも変わらず機能する。これは、探索の考慮事項は文献において考察されるモデルを限定する傾向があるのに対して、本アルゴリズムの著しい優位点である。

10

#### 【 0 0 7 1 】

本アルゴリズムは、任意時間アルゴリズムである。これは、その実行中のいつでも、部分解ではなく、その時点で入手可能な最良の完全解を得るために中断されることが可能である。

#### 【 0 0 7 2 】

20

提案しているアルゴリズムは、 $k$ -fanモデルにおける探索効率を向上させるためにも使用されることが可能である。本アルゴリズムは、 $k$ 個のアンカパーツ構成の評価をスケジュールするために使用される。これらの $k$ 個のアンカパーツによる所定の構成に関して、「真のスコア」は、 $k$ 個のアンカパーツ並びに他のパーツの最適配列を含む合計モデルスコアになる。

#### 【 0 0 7 3 】

本明細書に記述している概念は、最適化の速度を数倍も向上させる利点を提供する。これは、局在化をさらに大幅に効率的にし、よって本方法は潜在的に幾つかの現実的なタスクに適用され得る。さらに、速度の向上に起因して、より多くのオブジェクトパーツをモデリングすることができる。その結果、局在化及びカテゴリ分類のパフォーマンスは向上する。

30

#### 【 0 0 7 4 】

画像のカテゴリ分類に関しては、個々のカテゴリモデルを基礎とする多クラス分類のための幾つかの周知技術のうちの1つを使用することができる。我々は、幾つかのカテゴリモデルに関して、我々のアルゴリズムを用いて最適一致を発見する。最も高い一致スコアを有するカテゴリは、その画像のカテゴリとして割り付けられる。当然ながら、取り上げるカテゴリは1つだけでなくてもよい。画像の検出検査に合格する全てのカテゴリは、有効なカテゴリラベルとして受容される可能性もある。他に、割り付けられる1つまたは複数のカテゴリを解く前に様々なカテゴリスコアが正規化される多くの方法を使用することができる。

40

#### 【 0 0 7 5 】

これまでに開示した、及び他の特徴及び機能またはこれらの代替物のうちの様々なものが、望ましくは他の多くの異なるシステムまたはアプリケーションへと結合されてもよいことは認識されるであろう。また、同じく以下の請求の範囲に包含されることが意図される様々な、現時点では不測である、または予期されていない代替物、修正、変形またはこれらにおける改良が、当業者によって続いて実行される場合もある。

【図 1】

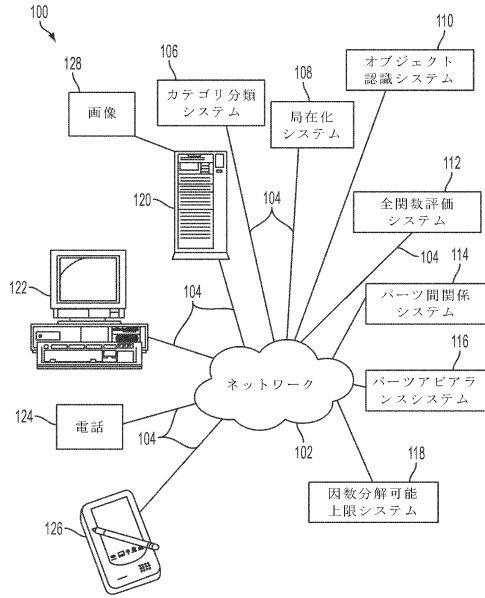


図 1

【図 2】

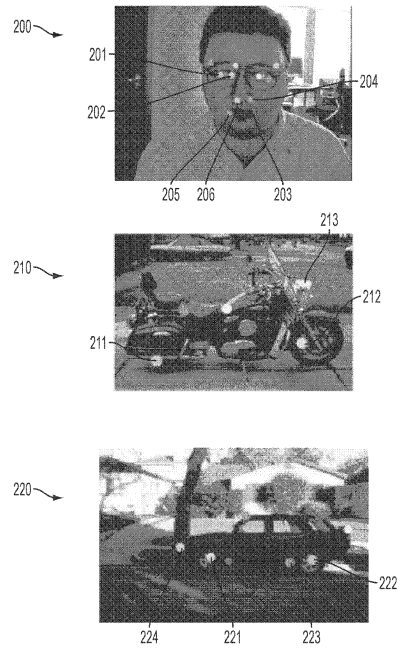


図 2

【図 3】

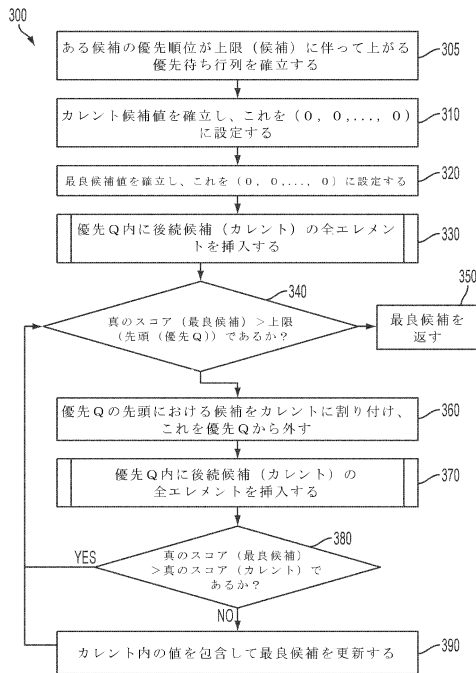


図 3

【図 4】

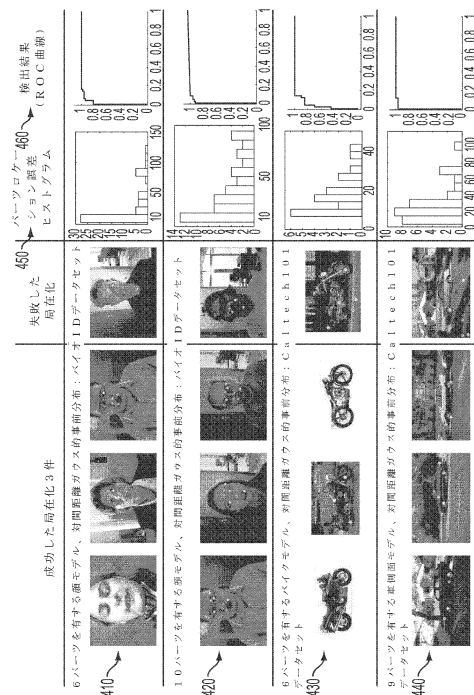


図 4



【図 5】

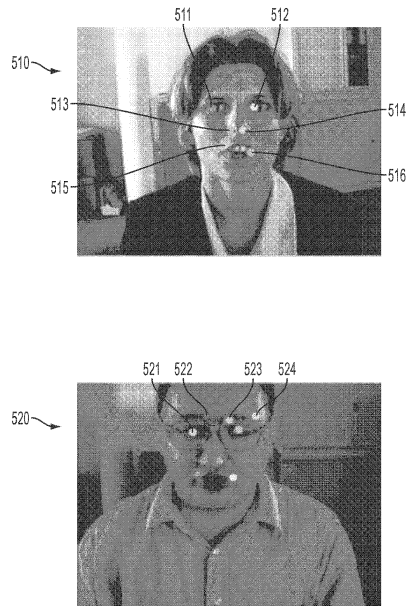


図 5

【図 6】

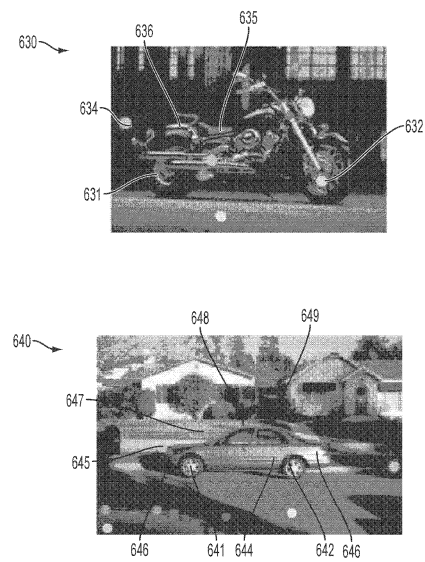


図 6

【図 7】

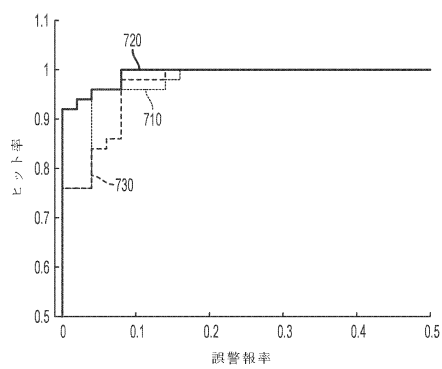


図 7

【図 9】

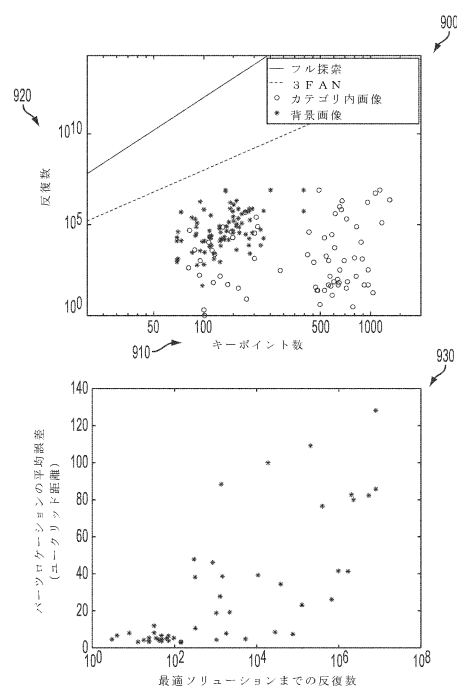


図 9

【図 8】

	840	850	860	870
データセット、モデル	我々の方法	STAR	2-FAN	フル探索
810 バイオID 顔、6 パーツ	6.0e+5	5.0e+5	4.2e+8	3.8e+17
820 バイオID 顔、10 パーツ	2.6e+6	5.0e+5	4.2e+8	5.8e+29
830 バイオID 顔、14 パーツ	3.3e+6	5.0e+5	4.2e+8	4.4e+41
バイオID 顔、6 パーツ (同時実行最適)	1.4e+6	5.0e+5	4.2e+8	3.8e+17
オートバイ、6 パーツ	5.3e+4	1.8e+4	2.8e+6	1.5e+13
車側面、6 パーツ	4.4	6.0e+4	1.5e+7	5.8e+21

図 8

## フロントページの続き

(72)発明者 エフゲニ・バルト

アメリカ合衆国 カリフォルニア州 94086 サニーヴェイル ベルフラワー・アヴェニュー  
678 アpartment ナンバー21

審査官 千葉 久博

(56)参考文献 特開2002-015284(JP, A)

神谷保徳, 外5名, "一般物体認識におけるタイプの異なる局所特徴の統合利用", 画像ラボ,  
日本, 日本工業出版株式会社, 2010年 1月10日, 第21巻, 第1号, p.1-5神谷保徳, 外3名, "一般物体認識のためのマルチモーダル星座モデル", 電子情報通信学会論  
文誌, 日本, 社団法人電子情報通信学会, 2009年 8月 1日, 第J92-D巻, 第8号,  
p.1104-1114山田貢己, 外2名, "因子分解法と部分空間法による顔向き推定", 電子情報通信学会技術研究  
報告, 日本, 社団法人電子情報通信学会, 2002年 1月10日, 第101巻, 第568号,  
p.1-8

(58)調査した分野(Int.Cl., DB名)

G06T 1/00

G06K 9/00 - 9/03, 9/46 - 9/52,  
9/62 - 9/82

G06T 7/00, 7/20 - 7/60