

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-97607
(P2008-97607A)

(43) 公開日 平成20年4月24日(2008.4.24)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 200D	5B050
G06F 17/30 (2006.01)	G06F 17/30 170B	5B075
G06T 7/00 (2006.01)	G06F 17/30 210D	5L096
	G06T 7/00 250	

審査請求 未請求 請求項の数 4 O L (全 18 頁)

(21) 出願番号 特願2007-262997 (P2007-262997)
 (22) 出願日 平成19年10月9日 (2007.10.9)
 (31) 優先権主張番号 11/548377
 (32) 優先日 平成18年10月11日 (2006.10.11)
 (33) 優先権主張国 米国 (US)

(71) 出願人 000002369
 セイコーエプソン株式会社
 東京都新宿区西新宿2丁目4番1号
 (74) 代理人 100095728
 弁理士 上柳 雅誉
 (74) 代理人 100107261
 弁理士 須澤 修
 (74) 代理人 100127661
 弁理士 宮坂 一彦
 (72) 発明者 ファイ ゾウ
 カナダ オンタリオ トロント スプリング
 デール ブルバード 136

最終頁に続く

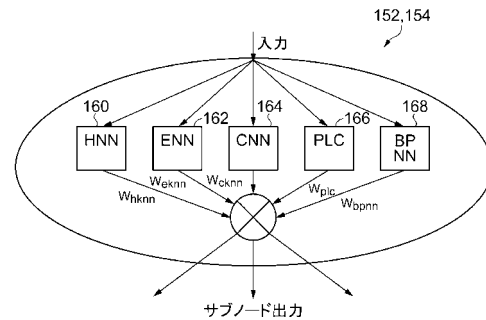
(54) 【発明の名称】 入力イメージを自動的に分類する方法

(57) 【要約】

【課題】 イメージの分類を自動的に行うための方法および装置。

【解決手段】 入力イメージを自動的に分類化する方法は、入力イメージの特徴を抽出するステップおよび入力イメージを表している特徴ベクトルを生成するステップを含む。特徴ベクトルは複数の異なる分類器を用いてプロセスされ、その複数の異なる分類器の組み合わせ出力に基づいて入力イメージを分類する。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

入力イメージを自動的に分類する方法であって、
前記入力イメージの特徴を抽出する工程と、
前記抽出した特徴に基づいて前記入力イメージを表す特徴ベクトルを生成する工程と、
分類器を用いて前記特徴ベクトルを分類する工程と、
複数の前記分類器の出力の組み合わせに基づいて前記入力イメージを分類する工程と、を
有し、
前記複数の分類器は、K平均最近傍分類器、知覚分類器、およびバックプロパゲーションニューラルネットワーク分類器のうち少なくとも二つを含むイメージ自動分類方法。

10

【請求項 2】

前記抽出した特徴は、前記入力イメージのカラーの同調性、エッジ方向の同調性、およびテキストの同時生起を表す、請求項 1 に記載の方法。

【請求項 3】

前記特徴抽出工程より前に、前記入力イメージをプリプロセッシングする工程をさらに含み、
該プリプロセッシング工程は、ノイズのフィルタリングおよび前記入力イメージの正規化のうち少なくとも一つを含む、請求項 2 に記載の方法。

【請求項 4】

前記分類工程は、前記複数の分類器を用いて重み付け出力を生成し、該重み付き出力を評価して前記入力イメージを分類する、請求項 3 に記載の方法。

20

【発明の詳細な説明】**【技術分野】****【0001】**

本発明はイメージ処理に関し、より具体的にはイメージを自動的に分類するための方法、装置およびそのためのコンピュータプログラムを実装したコンピュータ可読媒体に関する発明である。

【背景技術】**【0002】**

大型の媒体管理システムにおいては、格納されているイメージを効率よく且つ有効に取り出せるように、大まかなセマンティックの類似性を有するイメージを分類するのが望ましい。イメージを手作業で分類するのは時間がかかり現実的でない。特に、多数のイメージを分類している際にはそうである。そのためにイメージを自動的に分類する技術が望まれる。

30

【0003】

イメージを自動的に分類するための技術が考慮されてきている。例えば、Normile et al に付与された米国特許第 5 8 7 2 8 6 5 号明細書はイメージおよびビデオシーケンスを自動的に分類するためのシステムを公開している。このシステムは、あるカテゴリーを定義する固有値および固有ベクトルを決めるために初期カテゴリー集合を使った分類アプリケーションを実行する。次に、イメージの属性、画素ドメインでの直交分解、およびニューラルネットベースの分類を用いる直交分解のうちの一つを使って入力ビデオシーケンスの分類が行われる。平均のピンカラーヒストグラム、輝度 (intensity) に対する平均のルミナンス、平均の動きベクトルおよびテキストパラメータに基づく基本属性 (primitive attributes) の集合がビデオシーケンスのフレーム毎に生成される。ビデオシーケンスのフレームは、固定ベクトルによって定義された標準空間 (canonical space) に変換されて、カテゴリーを定義する固定値および固有ベクトルを基本属性と比較できるようになり、それによってフレームの分類が可能になる。

40

【0004】

Kimmel に付与された米国特許第 6 0 3 1 9 3 5 号明細書は、変形可能な輪郭 (deformable contours) を用いてイメージをセグメント化するための方法および装置を開示してい

50

る。セグメント化の対象となるターゲットオブジェクト（例えば、ボーダー）に関する事前情報（a priori information）が入力される。ターゲットオブジェクトを訓練用イメージ（training images）でトレースすることによって手作業でセグメント化して装置を訓練する。次に検索イメージを選択し、最も近い近傍の訓練用イメージを選択する。検索輪郭を形成するために訓練用イメージにおいてトレースした輪郭が次に検索イメージに転送される。事前情報および訓練情報に基づいて高度に類似していると考えられるターゲットオブジェクトの領域にくっつくように検索輪郭の形が変わる。そうして検索輪郭の最終のセグメント化が完了する。

【0005】

Burmanに付与された米国特許第6075891号明細書は、ハイバースペクトル画像を利用するためのリテラルでないパターン認識の方法およびシステムを開示している。空間でないスペクトル情報および電磁気スペクトルバンドデータを含むオブジェクトのオペティカル特性を定義するイメージの集合を生成するためにオブジェクトをスキャンする。次に、イメージ集合の中の単一の画素からスペクトルシグネチャが抽出される。そして、感知したデータに関するカテゴリーを識別するためにスペクトルシグネチャがフィルタされ、正規化され、マテリアル分類システム（material categorization system）に送られる。入力スペクトルシグネチャをなす構成マテリアルの量を検出および推定するために制約付きのミキシング方程式を解く遺伝的アルゴリズム（genetic algorithm）を採用する。

【0006】

Krumm et al.に付与された米国特許第6477272号明細書は、検索イメージの中のモデル化したオブジェクトの場所を識別するためのシステムおよびプロセスを開示している。検索イメージの中で場所を識別しようとしているオブジェクトのモデルイメージがキャプチャされる。一連の画素カラー範囲の同じ組み合わせ内に入るカラーを示すとともに同じ一連の距離範囲内に入る距離で分離された全ての画素対の数を生成することによって各モデルイメージを計算する。次に、モデルイメージ毎に同時発生ヒストグラム（co-occurrence histogram）が計算される。検索イメージのオーバーラップする部分から一連の検索ウィンドウが生成される。モデルイメージ用に確立された画素のカラー範囲および距離範囲を用いて検索ウィンドウ毎に同時発生ヒストグラムも計算される。類似性を評価（assess）するために各モデルイメージと各検索ウィンドウとの間の比較が行われる。そして、モデルイメージおよび検索イメージウィンドウからのそれぞれの同時発生ヒストグラムが類似性値を生み出すために比較される。類似性値がしきい値を超えていれば、オブジェクトは検索ウィンドウ内にあると考える。

【0007】

Krummに付与された米国特許第6611622号明細書は、シーンのイメージの中に描かれた人物およびオブジェクトを識別するオブジェクト認識システムおよびプロセスを開示している。イメージの中で識別の対象となる人物およびオブジェクトのモデルヒストグラムが作られる。識別している人物およびオブジェクトに一致しそうな領域を抽出するためにイメージがセグメント化される。抽出した領域毎にヒストグラムを計算し、抽出した各領域のヒストグラムとモデルヒストグラムそれぞれとの間の類似性の程度を評価する。モデルヒストグラムのうちのひとつの類似性の程度を示す、つまり規定のしきい値を超えるヒストグラムをもつ抽出領域がモデルヒストグラムと対応付いている人物またはオブジェクトに一致するとして指定される。

【0008】

Minamiに付与された米国特許第6668084号明細書は、ターゲットオブジェクトの形状およびルミナンス分布を識別する検索モデルを作るイメージ認識方法を開示している。オブジェクトの相関性を示すフィット具合（goodness-of-fit）を検索モデルのそれぞれ毎に計算し、計算したフィット具合に基づいて検索モデルが並べ替えられる。オブジェクトの形状をポリゴンとしてモデル化し、ルミナンス値をポリゴンの内側境界と考える。

【0009】

Andersonの米国特許出願公開第2001/0012062号明細書は、イメージを分析

10

20

30

40

50

および分類するためのシステムおよび方法を開示している。分析モジュールは選択した判定基準に対してキャプチャしたイメージファイルを審査して、望ましいカテゴリーのイメージを自動的にアクセスできるようにしかるべきカテゴリータグを生成してそれらをイメージと一緒に格納する。カテゴリータグを生成するために一つの分析モジュールは、赤、緑、青 (RGB) の遷移点でのイメージの最後のラインを分析する。もう一つの分析モジュールがイメージデータを YCC フォーマットに変換するためにガンマ補正およびカラー空間変換を行ってから、カテゴリータグを生成するためにイメージデータの最後のラインを Y Y C 遷移点で分析する。

【 0 0 1 0 】

Luo et al. の米国特許出願公開第 2 0 0 3 / 0 0 5 3 6 8 6 号明細書は、カラーイメージの中の内容領域 (subject matter region) を検出するための方法を開示している。イメージの中の各画素にはカラーおよびテクスチャに基づいて内容領域に属するとして確信値 (belief value) が指定される。確信値のしきい値化によって空間的に連続した候補の内容領域が形成される。そして、領域が望ましい内容に属する確率を判定するためにこの形成された空間に連続した内容領域が分析される。検出した内容領域および対応付けられた確率のマップが生成される。

10

【 0 0 1 1 】

Luo et al. の米国特許出願公開第 2 0 0 2 / 0 1 3 1 6 4 1 号明細書は、イメージの類似性を決めるためのシステムおよび方法を開示している。対象となるイメージの主体 (main subject) またはバックグラウンドの知覚的に有意義な特徴を決定する。こうした特徴にはカラーテクスチャおよび/または形状が含まれる。主体は連続して値をつけられた確信マップ (belief map) で示される。次に、対象となるイメージが格納されているイメージのどれかに類似しているかどうか決定するために、決定された知覚的に有意義な特徴をデータベースに格納されたイメージの知覚的に有意義な特徴と比較される。

20

【 0 0 1 2 】

Deng et al. の米国特許出願公開第 2 0 0 2 / 0 1 8 3 9 8 4 号明細書は、デジタルイメージを分類するためのシステム並びに方法を開示している。各イメージを、シーケンシャルに進行して一連の分類タスクにかけることで、キャプチャしたイメージを選択されたクラスに基づいて分類する。分類タスクは、イメージをクラスを指定するべきかどうか決めるためのアルゴリズムを要するノードである。イメージ内で、ある特定のクラスを識別できるかどうかを決めるためにノード毎にコントラストベースの分析および/またはメタデータの分析を採用する。

30

【 0 0 1 3 】

Schneiderman の米国特許出願公開第 2 0 0 4 / 0 0 6 6 9 6 6 号明細書は、オブジェクト検出プログラムにおいてサブ分類器 (sub-classifiers) の集合を決めるためのシステムおよび方法を開示している。候補としての係数サブセットを作成するモジュールは複数の係数候補サブセットを作成する。係数は 2 次元のデジタル化されたイメージに対して行われた変換操作 (transform operation) の結果であり、空間、周波数およびオリエンテーションのローカライズがなされたデジタル化されたイメージに基づいた対応するビジュアル情報を表している。訓練用のモジュールは複数の係数候補サブセットそれぞれのサブ分類器を訓練する。サブ分類器選択モジュールはいくつかのサブ分類器を選択する。選択されたサブ分類器は、オブジェクトがイメージのウィンドウ内に入っているかどうか決定するために、各入力イメージを調べる。オブジェクトの見た目が変わることを考慮に入れるために統計モデル化を使う。

40

【 0 0 1 4 】

Crandall et al. の米国特許出願公開第 2 0 0 4 / 0 1 7 0 3 1 8 号明細書は、デジタルイメージの中のカラーイメージを検出するための方法を開示している。ターゲットオブジェクトを含んだモデルイメージおよび潜在的にターゲットオブジェクトを含む検索イメージに対してカラー量子化を行う。複数の検索ウィンドウが生成され、各モデルイメージおよび検索イメージ毎の空間カラージョイント確率関数が計算される。空間カラージョイ

50

ント確率関数としてカラー同時発生エッジヒストグラムを選ぶ。ターゲットオブジェクトを含んだ検索ウィンドウを指定できるようにモデルイメージに対する各検索ウィンドウの類似性が評価される。

【0015】

【特許文献1】米国特許第5872865号明細書

【発明の開示】

【発明が解決しようとする課題】

【0016】

上記の引例はイメージを分類するためのテクニックを開示しているけれども、改善が望まれる。そこで、本発明の目的はイメージを自動的に分類するための新規の方法、装置およびそのためのコンピュータプログラムを実装したコンピュータ可読媒体を提供することである。

【課題を解決するための手段】

【0017】

従って、一つの態様においては、入力イメージを自動的に分類する方法を提供しており、この方法は次の工程を含んでいる：

前記入力イメージの特徴を抽出して、前記抽出された特徴に基づいて前記入力イメージを表す特徴ベクトルを生成し、

分類器を用いて特徴ベクトルを分類して前記分類器の出力の組み合わせに基づいて前記入力イメージを分類する。

一つの実施例において、分類工程には複数の分類器を用いて重み付き出力を生成して、入力イメージを分類するためにその重み付き出力を評価 (evaluate) することが含まれる。抽出した特徴は、入力イメージにおけるカラーの同調性 (coherence)、エッジのオリエンテーション、テキストの同時生起を表している。複数の分類器は、K中間近接傍分類器 (K-mean-nearest neighbor classifiers) と、知覚分類器、およびバックプロパゲーションニューラルネットワーク分類器のうち少なくとも二つを含む。

【0018】

一つの実施例においては、特徴抽出工程の前に、入力イメージをプリプロセッシングするステップをさらに含むものでもよい。このプリプロセッシング工程はノイズの除去 (フィルタリング)、入力イメージの正規化のうち少なくとも一つを含むものでよい。

【0019】

こうした方法、装置およびイメージを自動的に分類するためのコンピュータプログラムを実装したコンピュータ可読媒体はフレキシブルであるとともにロバストで、しかも公知のイメージ分類テクニックと比べて精度が向上している。

【発明を実施するための最良の形態】

【0020】

以下、本発明の実施形態を図面に基づいて説明する。

【0021】

以下の説明において、入力イメージを自動的に分類するための方法およびシステムの一つの実施例を提供している。イメージの分類時に、入力イメージの特徴が抽出され、その抽出した特徴に基づいて入力イメージを表す特徴ベクトルが生成される。(特徴ベクトルはシグネチャベクトルまたはフューチャベクトルと呼ばれることがある。)この特徴ベクトルは分類器(分類子ともいう)を用いて処理され、その分類器の組み合わせ出力に基づいて入力イメージにはクラスが割り当てられる。

【0022】

図1は、イメージを自動的に分類するための分類化システム140の概略図である。表示されているように、分類化システム140はノイズを取り除き、分類の対象となる各入力イメージを正規化するためのプリプロセッサ142を含んでいる。特徴ベクトルジェネレータ144はプリプロセッサ142によってフィルタおよび正規化されたイメージ出力それぞれを受け取り、イメージを分類するために使用するイメージの特徴nを表す特徴ベ

10

20

30

40

50

クトルを生成する。木のような階層構造 146 に配列された一連の分類化ノード（入力イメージをクラスおよびサブクラスに分類することを担当している）は特徴ベクトルジェネレータ 144 と通信する。階層構造 146 のトップとしてのノード 152 は入力イメージ毎に生成された特徴ベクトルを受け取り、入力イメージが分類されることになるクラスに基づいて基礎行（underlying row）のノード 154 に特徴ベクトル出力を提供する。

【0023】

この実施例において、ノード 152、154 はそれぞれ入力は一つで出力は複数である。各出力はそのノードで分類されたクラスまたはサブクラスを表す。例えば、図に示すように、ノード 152 は各入力イメージを一つの「ランドスケープ」、「ビルディング」および「人物」のうちの一つのクラスに分類する。基礎行の左側のノード 154 はノード 152 で「ランドスケープ」クラスが割り当てられた各入力イメージの特徴ベクトルを受け取り、その入力イメージをさらに「山」、「野原」、「砂漠」といったサブクラスに分類する。基礎行の真ん中にあるノード 154 はノード 152 で「ビルディング」のクラスが割り当てられた各入力イメージの特徴ベクトルを受け取り、その入力イメージをさらに「教会」、「家」、「塔」といったサブクラスに分類する。基礎行の右側のノード 154 はノード 152 で「人物」クラスが割り当てられた各入力イメージの特徴ベクトルを受け取り、その入力イメージをさらに「男性」および「女性」といったサブクラスに分類する。表示している分類化システム 140 は 3 つのノードからなる一列の基礎行のノード 154 しか含んでいないけれども、当業者ならばこれは例証を簡単にするためであることを理解するだろう。普通、多数の基礎行のノードが提供され、それらの各基礎行は多数のノード 154 をもち、入力イメージをうまく定義した詳細なサブクラスに分類できるようになっている。

10

20

【0024】

図 2 は、図 1 のイメージ分類システムの一部を形成する分類化ノードの概略図である。分類化するノード 152、154 はそれぞれ、3 つの K 平均最近傍分類器 160、162、164 と、1 セットの N 個のバイナリ知覚リニア分類器 166 と、一つの N クラスのバックプロパゲーションニューラルネットワーク分類器（以降、BPNN 分類器という）168 とを含んで構成され、ここで N は図 2 に表示したノードで分類されているクラスまたはサブクラスの数である。わかるように、各分類化ノードは多数の複数の異なる分類器を含んでおり、各複数の異なる分類器は違う範囲に強い（area of strength）。

30

【0025】

K 平均最近傍分類器 160 は、入力イメージのために生成された特徴ベクトルの色相成分に基づいて各入力イメージを分類することを担当している。K 平均最近傍分類器 162 は、特徴ベクトルのエッジオリエンテーション成分に基づいて各入力イメージを分類することを担当している。K 平均最近傍分類器 164 は、特徴ベクトルのテキスチャ同時発生成成分に基づいて各入力イメージを分類することを担当している。

【0026】

図 3 は、図 2 の分類化ノードの一部を形成する知覚リニア分類器の概略図である。図 3 に表示しているように、各バイナリ知覚リニア分類器 166 は、特徴ベクトル入力 170 と、ウェイト 172、加算関数 174、しきい値素子 176、および出力 178 を含んでいる。しきい値素子 176 は、加算関数 174 によって生成された総和に基づいて出力を管理する設定を有する。総和がしきい値素子 176 の設定以下の場合には、バイナリ知覚リニア分類器 166 はマイナス 1（-1）値を出力する。総和がしきい値素子 176 の設定以上の場合には、バイナリ知覚リニア分類器 166 は 1 値を出力する。

40

【0027】

図 4 は、図 2 の分類化ノードの一部を形成するバックプロパゲーション人工ニューラルネットワーク分類器の概略図である。BPNN 分類器 168 は、444 個の入力ノード 182 と、ウェイトを介して入力ノード 182 に連結した 20 個の隠れたノード 184 からなる隠れたレイヤー、および隠れたノード 184 に連結した N 個の出力ノード 186 より少なくともなる出力レイヤーを含んだ入力レイヤーを含み、ここで N は図 3 に示した分類

50

化ノードによって処理されるサブクラス数である。

【0028】

よく知られているように、BPNN分類器168はバイナリ知覚リニア分類器166よりもパワフルである。入力ノード182でのアクティビティいかににより、隠れたノード184のレイヤー全体でのエラーにかかる量でウェイトは変わる。入力特徴ベクトルを分類している最中に、隠れたノード184は出力し、隠れたノード184でのエラーが計算される。訓練を実施している最中に、隠れたノード184のレイヤー全体でのエラーが隠れたレイヤーを通してバックプロパゲートされ、隠れたノード184にかかるウェイトが調整される。

【0029】

図5は、自動イメージ分類時に分類化システム140が行う大まかなステップを表示するフローチャートである。最初、入力イメージを自動的に分類しようとするとき、ノイズを取り除いてからそのイメージを正規化するためにプリプロセッサ142(図1参照)が入力イメージをフィルタする(ステップS202)。入力イメージが一旦フィルタされ正規化されたら、特徴ベクトルジェネレータ144(図1参照)が入力イメージの特徴抽出を行って入力イメージを分類するために使用しようとしている入力イメージの特徴nを表す特徴ベクトルを生み出す(ステップS204)。この実施例では、特徴ベクトルは444個のピンを有し、カラーの同調性と、エッジオリエンテーションの同調性、およびテクスチャの同時発生に基づいている。特徴ベクトルは次に階層構造146(図1参照)に供給され、そこで特徴ベクトルは複数の異なる分類器、具体的にはK平均最近傍分類器160~164と、バイナリ知覚リニア分類器166、およびBPNN分類器168によって処理されることでイメージを分類する(ステップS206)。

【0030】

ステップS202で入力イメージがプリプロセッサ142(図1参照)によって処理されているときに、入力イメージはノイズを取り除くために3x3のボックスフィルタを通され、640x480の色相、彩度、輝度のイメージ(「HSI」)を形成するために正規化される。

【0031】

ステップS204での特徴抽出時に、入力HSIイメージのための444ピン特徴ベクトルが構築される。この特徴ベクトルは入力HSIイメージから抽出した特徴数に基づいている。最初に、HSIイメージにおける各画素のカラーの同調性/非同調性を調べることによって、図6に示すHSIカラー空間における48x2のローカライズされたカラー同調性ベクトル(LCCV)マトリックスがポピュレートされる。マトリックスの各行は、色相(H)に指定された32のピンと、彩度(S)に指定された8のピン、および輝度(I)に指定された8のピンを含んでいる。マトリックスの1行はイメージの画素のカラーの同調性を表し、もう一方の行はイメージの画素のカラーの非同調性を表している。

【0032】

この実施例では、良く知られたカラー同調性ベクトル(CCV)ではなく、各画素のローカライズされたカラーの同調性ベクトル(LCCV)を使う。ローカライズされたカラー同調性ベクトル(LCCV)は計算がもっと効率的である。カラーの同調性/非同調性を決定する際に、問題の画素上でセンターにある指定したウィンドウ内の画素のうちの一つがまったく同じカラーであるとき且つそのときに限って画素はローカルなカラーの同調性であると考えられる。この実施例では、イメージの画素がカラーの同調性を有するかあるいは非同調であるかどうかを決定するのに5x5のカラーの同調性ウィンドウを使う。図7は対象となる画素P上でセンターにある5x5のカラー同調性ウィンドウを示している。対象となる画素Pを取り囲む数字は、カラーの同調性ウィンドウ内の画素と対応付けられたHSIカラー情報を表現している。シェーディングして示した取り囲んでいる画素は対象となる画素Pとまったく同一のカラー情報を有する。従って、図7の例では、対象となる画素Pは8個のローカルな同調性をもつ近傍画素と16個のローカルなカラー非同調の近傍画素とを有する。

10

20

30

40

50

【 0 0 3 3 】

H S I イメージの各画素のローカルエッジのオリエンテーションを調べることによって、図 8 に示す 46×2 のローカライズされたエッジのオリエンテーションの同調性ベクトルのマトリックスもポピュレートされる。ローカライズされたエッジのオリエンテーションの同調性ベクトルマトリックスの各行はエッジ方向の 45 のピンと非エッジのピン 1 個を含む。各エッジ方向のピンは 8 度を含む (encompass 8 degrees)。マトリックスの 1 行はイメージ画素のローカルエッジのオリエンテーションの同調性を表す一方、別の行はイメージ画素のローカルエッジのオリエンテーションの非同調性を表している。

【 0 0 3 4 】

ローカライズされたエッジのオリエンテーションの同調性ベクトルマトリックスのポピュレートする時に、問題の画素でセンターにある特定のウィンドウ内の画素のうちの一つが同じエッジオリエンテーションを有しているとき且つそのときに限って画素はローカル的にエッジのオリエンテーションが同調性を有していると考える。この実施例では、 5×5 のエッジオリエンテーション同調性ウィンドウを使う。図 9 は、対象となる画素 P でセンターにあるエッジオリエンテーション同調性ウィンドウの一例を示す。ウィンドウの中の数字はそのエッジオリエンテーション同調性ウィンドウ内の画素と対応付けられたエッジオリエンテーションを表す。従って、図 9 の例において、対象となる画素 P は 8 のローカルエッジのオリエンテーションの同調性を有する画素をもつとともに 16 のローカルエッジのオリエンテーションの非同調性を有する画素を有する。

【 0 0 3 5 】

H S I イメージの画素に基づいて 16×16 のテクスチャ同時発生マトリックス (GLCM) もポピュレートされる。GLCM の定義については、Michael S. Lew (Ed) の著作「Principals of Visual Information Retrieval」(Springer-Verlag, 2001, ISBN: 1-85233-381-2) に説明されている。この出版物の内容は参照することにより本明細書に組み込まれているものとする。知られているように、GLCM はテクスチャを特徴化し、H S I イメージにおいて隣接画素の輝度レベルの異なる組み合わせがどれだけ頻繁に発生するかを表にする。GLCM の各行は可能性のある画素レベルの輝度を表す一方、各列は隣接画素の可能性のある輝度レベルを表している。GLCM の各セル内の値は、ある特定の画素の輝度レベル対が H S I イメージにおいて発生する回数を識別するものである。

【 0 0 3 6 】

上記の特徴抽出の最終結果は、図 10 に示す 11 の成分特徴ベクトルである。特徴ベクトルが生成された後、入力イメージデータが表すことに基づいてその特徴ベクトルの成分を正規化する。この実施例では、色相成分が $3 \times 640 \times 480$ により正規化され、ここで 3 はカラーのチャンネル数で、 640×480 はステップ S 202 で正規化されたイメージの解像度である。エッジオリエンテーションの成分はエッジ画素の数により正規化され、テクスチャの同時発生成分は 640×480 で正規化される。

【 0 0 3 7 】

特徴ベクトルが正規化された後、特徴ベクトルは階層構造 146 のトップとしてのノード 152 に適用され、その中にある分類器の各々に供給される。つまり、K 平均最近傍分類器 160 は、特徴ベクトルの色相成分を受け取り、K 平均最近傍分類器 160 が入力イメージはノード 152 で分類された各クラスを表していると思う度を表す重み付き出力 W_{hnn} を生成する。K 平均最近傍分類器 162 は、特徴ベクトルのエッジオリエンテーション成分を受け取り、K 平均最近傍分類器 162 が入力イメージはノード 152 で分類された各クラスを表していると思う度を表す重み付き出力 W_{enn} を生成する。K 平均最近傍分類器 164 は、特徴ベクトルのテクスチャ同時発生成分を受け取り、K 平均最近傍分類器 164 が入力イメージはノード 152 で分類された各クラスを表していると思う度を表す重み付き出力 W_{cnn} を生成する。

【 0 0 3 8 】

つまり、K 平均最近傍分類器 160 ~ 164 による特徴ベクトル成分の処理時に、特徴ベクトル成分は以下に示すベクトル交点方程式 (intersection equation) (1) を用い

10

20

30

40

50

て、ノードで評価されたクラスの中にある平均的な表現ベクトル成分 (average representation vector components) と比較される。

【 0 0 3 9 】

【 数 1 】

$$I(v1,v2)=\left(\frac{\text{elements}}{\sum_{i=1} \min(v1[i],v2[i])}\right) / \max(W1,W2) \text{-----式(1)}$$

ここで、I は 2 つのベクトル成分のベクトル交点、

1 は入力特徴ベクトル成分、

2 は平均的ベクトル成分、

1 [i] および 2 [i] はベクトル成分 1 およびベクトル成分 2 それぞれの i 番目の元の値で、

元はベクトル成分に入っているピン総数で、

W 1 と W 2 はベクトル成分 1 およびベクトル成分 2 それぞれにおけるピンの値の累積総和である。

【 0 0 4 0 】

上記のことを達成するために、訓練のプロセス時に、クラス毎に K の数学的平均表現ベクトル (K mathematical mean representation vectors) が決定され、そのクラス内に入るであろうイメージのジェネリックな表現としての機能をする。これらのベクトルは M / K の訓練用イメージベクトルの平均値をとることによって決定され、ここで M は訓練用イメージベクトルの総数で、K は平均的ベクトルの数で、ここで K は分類化システム 1 4 0 の要件に基づいて選択される。

【 0 0 4 1 】

K のベクトル交点を決めてから、それを用いて入力イメージが属するクラスが決められる。例えば、最近傍の平均的な表現ベクトルの大半が「クラス 1」に属する場合には、特徴ベクトルは「クラス 1」に属すると宣言される。

【 0 0 4 2 】

バイナリ知覚リニア分類器 1 6 6 は特徴ベクトル全体を受け取り、バイナリ知覚リニア分類器 1 6 6 が入力イメージはノード 1 5 2 によって分類された各クラスを表していると思う度を表す出力 W_{PLC} を生成する。つまり、バイナリ知覚リニア分類器 1 6 6 は、一致するクラスと属する特徴ベクトルとそうでないクラスとを分けるハイパープレーン (hyper-plane) を得る。バイナリ知覚リニア分類器 1 6 6 の結果を得て、最高の分類器の結果値を最終のクラスと宣言する。シグモイド関数を起動関数 (activation function) として使う。

【 0 0 4 3 】

B P N N 分類器 1 6 8 の各入力ノードは特徴ベクトルの対応付けられたピンを受け取る。B P N N 分類器 1 6 8 は、B P N N 分類器 1 6 8 がノード 1 5 2 によって分類された各クラスを表していると思う度を表す重み付き出力 W_{bpnn} を生成する。

【 0 0 4 4 】

K 平均最近傍分類器 1 6 0 ~ 1 6 4、バイナリ知覚リニア分類器 1 6 6、B P N N 分類器 1 6 8 の重み付き出力は次にピンソートが行われる。最高の累積和をもつピンと対応付けられたクラスを次に決定し、そのクラスが入力イメージがノード 1 5 2 によって分類されたクラスであると考えられる。

【 0 0 4 5 】

一旦、ノード 1 5 2 が入力イメージをクラスに分類したら、上に説明したと同じようなやり方でさらにサブクラスの分類を行うために、そのクラスと対応付けられたノード 1 5 4 に特徴ベクトルを適用する。

【 0 0 4 6 】

例えば、ノード 1 5 2 のクラス出力が「ビルディング」であると決定されたら、次にそ

10

20

30

40

50

の「ビルディング」に対応するノード154は更なるサブクラス分類のために特徴ベクトルを受け取る。ノード154が特徴ベクトルを受け取ると、そこに入っている分類器は、入力イメージをさらに分類するために、上に説明したと同じようなやり方で特徴ベクトルに対して働く。

【0047】

小さな集合の訓練用イメージしか使えない時には分類化を強調するために、ベクトルの密度に基づいて縮小された訓練用イメージから抽出したベクトルの集合を用いて分類しているノード152, 154のBPNN分類器168を訓練する。図11は、訓練用イメージとしての集合Sからベクトルの集合を生成するために行われるステップを表示している。最初に、訓練用イメージが属するクラスを統計的に表す訓練用イメージの特徴ベクトルを生成するために、集合Sにおける訓練用イメージに対して特徴抽出を行う(ステップS302)。次に、制御されたランダム性を加えるためにスキュー特徴ベクトルを生成する(ステップS304)。次に、分類化ノードの分類器を訓練するために使用する訓練用イメージの特徴を表す特徴ベクトルの集合を生み出すベクトルの密度に基づいて、訓練用イメージ特徴ベクトルおよびスキュー特徴ベクトルの集合を縮小する(ステップS306)。

10

【0048】

ステップS302で、最初に、訓練用イメージとしての集合Sの中の各イメージを、先に説明したと同じようなやり方で、特徴ベクトルVを表現する。次に、訓練用イメージの特徴ベクトルVをベクトル集合Hにコピーする。訓練用イメージとしての集合Sの中の各カテゴリはm個のサブ集合に分割される。そこでmは事前に定義された値である。この実施例では、mは $n_s / 7$ で、 n_s は訓練用イメージから抽出する特徴の集合である。次に訓練用の特徴 n_{sub} のサブ集合S'を訓練用イメージとしての集合Sからランダムに選ぶ。次に、訓練用の特徴 n_{sub} のサブ集合S'の数学的平均ベクトル V_{mean} を以下の方程式(2)の従って計算する。

20

【0049】

【数2】

$$V_{mean}(i) = (1/n) \sum_{S^l} G_j(i) \text{ -----式(2)}$$

30

ここで、 $G_j(i)$ はサブ集合 S^l における j^{th} のベクトルの i^{th} の元で、 $V_{mean}(i)$ は数学的平均ベクトル V_{mean} の i^{th} の元である。

【0050】

計算した数学的平均ベクトル V_{mean} をベクトル集合Hに加算する。選んだ訓練用としての特徴 n_{sub} のルート平均平方サンプルベクトル(以降、RMSサンプルベクトルという) V_{rms} も以下の方程式(3)に従って計算する。

【0051】

【数3】

$$V_{rms}(i) = \sqrt{(1/n) \sum_{S^l} (G_j(i))^2} \text{ -----式(3)}$$

40

ここで、

$V_{rms}(i)$ はRMSベクトル V_{rms} の i^{th} の元である。

【0052】

次に、RMSサンプルベクトル V_{rms} をベクトル集合Hに加算する。次に、選んだ訓練用の特徴 n_{sub} の最小の訓練用イメージの特徴ベクトル V_{min} を以下の方程式(4)に従って計算する。

50

【 0 0 5 3 】

【 数 4 】

$$V_{min}(i) = \min(G_1(i), G_2(i), \dots, G_n(i)) \text{ -----式(4)}$$

ここで、

$V_{min}(i)$ は最小の訓練用イメージの特徴ベクトル V_{min} の i^{th} の元である。

【 0 0 5 4 】

次に、最小の訓練用イメージの特徴ベクトル V_{min} をベクトル集合 H に加算する。ステップ $S204$ で、選んだ訓練用の特徴 n_{sub} におけるサンプルベクトルの元 G ごとに、スキューベクトル V_{skewj} を以下の方程式 (5) に従って計算する。

10

【 0 0 5 5 】

【 数 5 】

$$V_{skewj}(i) = G_j(i) + P^{\mp} \left(\text{rand}(V_{min}(i), V_{mean}(i), V_{rms}(i)) - G_j(i) \right) \text{ -----式(5)}$$

ここで、

20

j は選んだ訓練用の特徴 n_{sub} におけるサンプルベクトル (0, 1, ..., n) の指数 (index) で、

p はランダムに選んだ統計的ベクトルのパーセンテージ (0~1) で、

$\text{rand}(a, b, c)$ は a, b または c からランダムに選んだ一つを表す。

【 0 0 5 6 】

計算したスキューベクトル $V_{skew0}, V_{skew1}, \dots, V_{skewN}$ もベクトル集合 H に加算する。それから、訓練用イメージとしての集合 S におけるイメージがすべてプロセスされるまで、上記のステップを繰り返す。これにより ($2n_s + 3m = n_e$) ベクトルを含んだベクトル集合 H が生み出される。

【 0 0 5 7 】

30

分かるように、余分の特徴ベクトル (本件の場合、ベクトル数学的平均ベクトル V_{mean} 、特徴ベクトル V_{min} および RMS ベクトル V_{rms}) を挿入することで特徴ベクトルの集合を拡大する。訓練用のイメージ集合における特徴毎に歪曲したベクトル (distorted vector) V_{skew} も挿入する。訓練用のイメージ集合が小さいときには、訓練時こうしたベクトルを加えることが役に立つ。訓練用のイメージ集合が大きいたときにはこうしたベクトルを追加することが訓練に役立つ程度が小さくなる。

【 0 0 5 8 】

上記のプロセスは、各訓練用のイメージが属するクラスを統計的に表現するベクトルを作成するために、訓練用のイメージ集合に含まれている情報を用いる。既存の訓練用のイメージ集合に表現ベクトルを付加することで、一様でない初期の訓練用のイメージデータの分布から生じるクラス情報のギャップを埋めることによって個々のクラスは平滑な推定データの分布でよりうまく定義できるようになる。さらに、このプロセスは、これから説明するように、クラスの定義を表現しているとは考えない訓練用のイメージ集合におけるベクトルがステップ $S206$ 時に消去される確率が高くなる。

40

【 0 0 5 9 】

分類化システム 140 が訓練できる類似したデータのより連続してうまく定義された範囲を提供するためにスキューしたベクトルを生成する。訓練用のイメージ集合にある程度の制御された分散を提供することによって、スキューしたベクトルによって、精度を損なうことなく、より広義な意味での訓練が可能になる。わずかに制御されたランダム性は、訓練用のイメージが完璧に理に適ったデータに優れた分類を与えないときに、訓練用のラ

50

ーニングシステムによって、クラスにフィットする可能なデータの範囲に一様に沿って分布されていない訓練用のイメージが高度に集中していることが原因で、頻繁に経験するよくある疎なデータの問題を減らすのに役立つ。

【 0 0 6 0 】

一旦、ベクトルの集合としてのベクトル集合Hを生成したら、ベクトル集合Hを縮小する(ステップS306)。このプロセス時に、 $\{h_k\}$ で表記したベクトル集合Hの中のベクトルごとに、サブ集合 $C_i - \{h_k\}$ 上のベクトルの内部クラスの密度 $DI\{h_k\}$ を計算する。ここで、 C_i は $\{h_k\}$ が属するクラスの全ての元を含んだベクトル集合Hのサブ集合である。内部クラスの密度 $DI\{h_k\}$ は以下の方程式(6)で定義される。

【 0 0 6 1 】

【数6】

$$DI(h_k) = \left(\sum_{C_i - \{h_k\}} I(\{h_k\}, \{h_j\}) \right) / (N_{C_i} - 1) \quad \text{-----式(6)}$$

ここで、

【 0 0 6 2 】

【数7】

$$(\{h_1\}, \{h_2\}) = \left(\sum_{i=1}^e \min(h_1(i), h_2(i)) \right) / \max(A_{h_1}, A_{h_2}) \quad \text{-----式(7)}$$

式(7)、 N_{C_i} は C_i におけるサンプル数で、

e はサンプルベクトルにおける元の総数で、

A_{h_1} および A_{h_2} はベクトル $\{h_1\}$ およびベクトル $\{h_2\}$ それぞれにおける元の値の累積総和である。

【 0 0 6 3 】

ベクトル集合Hにおけるベクトル毎に、サブ集合 C_o 上のベクトル $\{h_k\}$ の外部クラスの密度 $DO(h_k)$ を計算する。ここで、 C_o は $\{h_k\}$ が属しないクラスの全ての元を含んだベクトル集合Hのサブ集合である。外部クラスの密度 $DO(h_k)$ は以下の方程式(8)で定義される。

【 0 0 6 4 】

【数8】

$$DO(h_k) = \left(\sum_{C_o} I(\{h_k\}, \{h_j\}) \right) / (N_{C_o}) \quad \text{-----式(8)}$$

ここで、

N_{C_o} はサブ集合 C_o におけるサンプル数である。

【 0 0 6 5 】

ベクトル集合Hにおけるベクトル毎に、ベクトル $\{h_k\}$ の同調性の密度 $DC(h_k)$ を計算する。同調性の密度 $DC(h_k)$ は以下の方程式(9)に従って定義される。

【 0 0 6 6 】

【数9】

$$DC(h_k) = DI(h_k) - \alpha \cdot DO(h_k) \quad \text{-----式(9)}$$

10

20

30

40

50

ここで、 α は事前に定義した係数である。

【0067】

最も高い同調性の密度をもつベクトルがベクトル集合 R に格納され、そのベクトルは分類ノードの分類器を訓練するのに使用する拡張した訓練用のイメージ集合からの特徴を表している。

【0068】

上記のプロセスは訓練用のイメージ集合の密度データに基づいてベクトルの集合を小さくする。つまり、ベクトル集合 H においてとっておくべきベクトルはどれが決定するために、訓練用のイメージ集合の密度分布を活用する。ベクトル縮小の段階があることにより、データ集合に完全に依存する場合よりも、最終的な訓練用のイメージ集合のサイズがもっと正確になる。これにより、特定の問題点や性能上の拘束のニーズに合うようにユーザが訓練用のイメージ集合を微調整できるようになる。また、ベクトル縮小の段階は、どんな形でもラーニングを必要としないので、性能が高速になる。また、異なった訓練用のイメージ集合をプロセスする場合、結果の一貫性が一層高くなる。ベクトル縮小の段階はモデルに依存しておらず、ある特定の分類システム専用の設計にはなっていない。ベクトル縮小の段階はジェネリックな性質なので、異なったサブラーニングシステムを含んでいることのあるハイブリッドなシステムに効果的に使用できる。

10

【0069】

これまでに説明してきた実施例は特定の分類器に関係している。当業者ならば、様々な分類器が複数の異なる場合には、代替えの分類器を使用できることが分かるだろう。

20

【0070】

ベクトル集合の拡張および縮小は、パーソナルコンピュータやその他のコンピューティングシステム環境などプロセッシングユニットによって実行されるコンピュータが実行可能な命令を含んだソフトウェアアプリケーションにより実行可能である。ソフトウェアアプリケーションはスタンドアロン版のイメージ分類化ツールとしてランできるし、あるいはメディア管理システムに組み込んでこうしたメディア管理システムに向上した機能性を提供することもできる。ソフトウェアアプリケーションは、ルーチン、プログラム、オブジェクトコンポーネント、データストラクチャなどのプログラムモジュールを含んでいることがあり、コンピュータが判読可能な媒体に格納したコンピュータ可読プログラムコードとして実施されることがある。コンピュータ可読媒体は、後からコンピュータシステムで読める、データを格納できるどんなデータストレージデバイスでも構わない。コンピュータ可読媒体の例としては、例えば、リードオンリーメモリ、ランダムアクセスメモリ、CD-ROM、磁気テープ、オプティカルデータストレージデバイスなどを含む。コンピュータが判読可能なプログラムコードは、コンピュータ可読プログラムコードを分散して格納および実行できるように、結合型のコンピュータシステムなどネットワークで分散することができる。

30

【0071】

ノードベースの分類化ストラクチャにより、モジュール性および編成が向上するので、装置が新たなイメージカテゴリーを含むべく拡張することが可能になる。数学的平均の代表的なベクトルを組み込むことで、大型の事前に存在しているイメージコレクションがなくても、優れたインスタントベースの分類を実行できる。多様なタイプが異なった分類器を含む重み付きのコミティベースの分類システムを用いることによって、各分類器がもつ異なった分野の強みを利用することでより優れた精度を達成する。特徴ベクトルのコンポーネントとして LCCV および LEOCV を用いると、正確であるけれど、保持しようとしているイメージの非常に凝縮された表現が可能になるので、訓練および分類がより高速になる。

40

【0072】

実施例を説明してきたが、当業者ならば、添付の請求項で定義した精神ならびに範囲から逸脱することなく、変形および変更ができることを理解するだろう。

【図面の簡単な説明】

50

【 0 0 7 3 】

【 図 1 】 イメージ分類システムの概略図。

【 図 2 】 図 1 のイメージ分類システムの一部を形成する分類化ノードの概略図。

【 図 3 】 図 2 の分類化ノードの一部を形成する知覚リニア分類器の概略図。

【 図 4 】 図 2 の分類化ノードの一部を形成するバックプロパゲーション人工ニューラルネットワーク分類器の概略図。

【 図 5 】 自動イメージ分類時に行われる大まかなステップを示すフローチャート。

【 図 6 】 ローカライズしたカラーの同調性ベクトルマトリックスを示す図。

【 図 7 】 図 6 のローカライズしたカラーの同調性マトリックスを生成するために使用するカラーの同調性ウィンドウを示す図。

10

【 図 8 】 ローカライズしたエッジオリエンテーションの同調性ベクトルマトリックスを示す図。

【 図 9 】 図 8 のローカライズしたエッジオリエンテーションの同調性ベクトルマトリックスを生成するために使用するエッジオリエンテーションの同調性ウィンドウを示す図。

【 図 1 0 】 特徴ベクトルを示す図。

【 図 1 1 】 イメージセット強調 (image set enhancement) の訓練時に行われるステップを示すフローチャート。

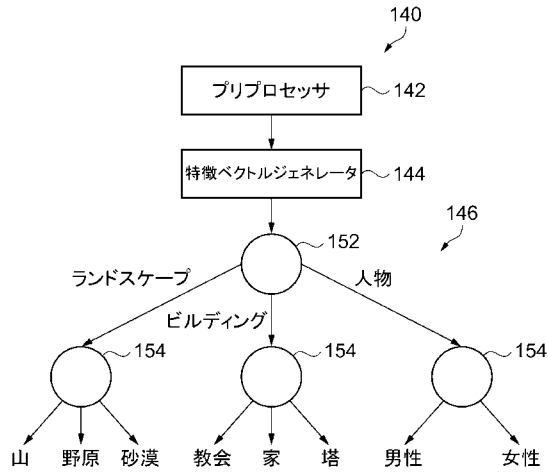
【 符号の説明 】

【 0 0 7 4 】

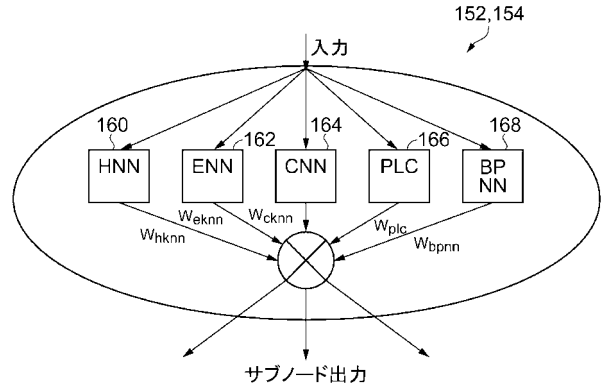
1 4 0 ... 分類化システム、 1 4 2 ... プリプロセッサ、 1 4 4 ... 特徴ベクトルジェネレータ、 1 4 6 ... 階層構造、 1 5 2 , 1 5 4 ... ノード、 1 6 0 , 1 6 2 , 1 6 4 ... K 平均最近傍分類器、 1 6 6 ... バイナリ知覚リニア分類器、 1 6 8 ... バックプロパゲーションニューラルネットワーク (B P N N) 分類器、 H ... ベクトル集合、 S ... 集合、 S 2 0 2 ... ステップ；入力イメージをプリプロセス、 S 2 0 4 ... ステップ；特徴抽出に基づいて特徴ベクトルを生成、 2 0 6 ... ステップ；複数の異なる分類器を用いて特徴ベクトルをプロセスしてイメージを分類、 3 0 2 ... ステップ；訓練用のイメージを拡張してサンプルベクトルを生成、 3 0 4 ... ステップ；スキューベクトルを生成、 3 0 6 ... ステップ；サンプルおよびスキューベクトルの集合を縮小。

20

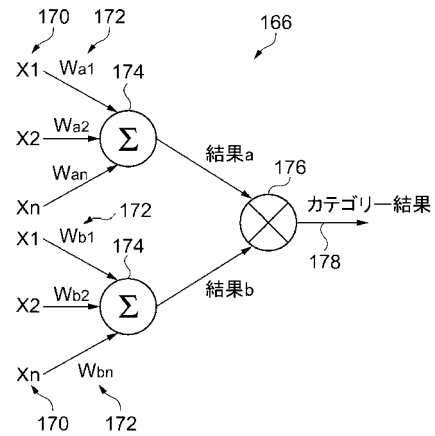
【 図 1 】



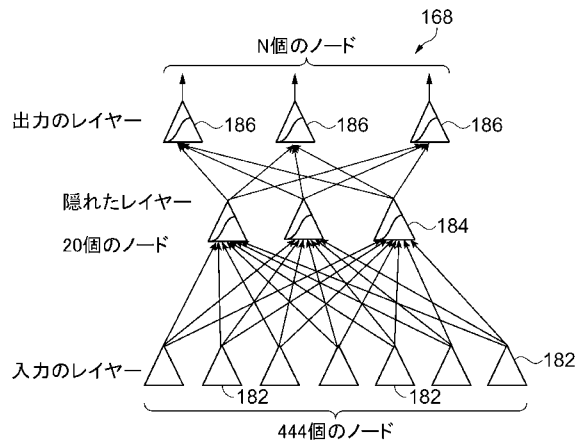
【 図 2 】



【 図 3 】



【 図 4 】

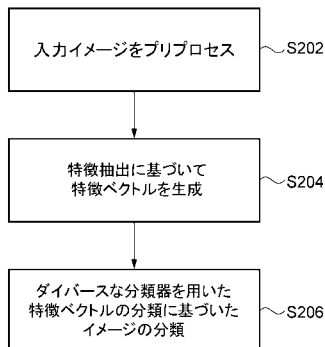


【 図 6 】

HI	SI	II
HC	SC	IC

32個のビン(H) 8個のビン(S) 8個のビン(I)

【 図 5 】



【 図 8 】

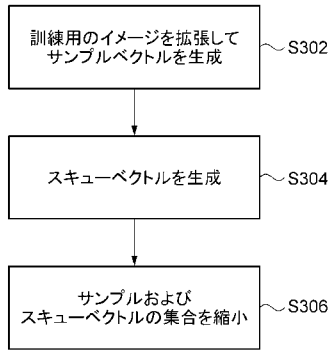
DI	NI
DC	NC

45個のビン (エッジ方向) 1個のビン (エッジでない部分)

【 図 10 】

HI	SI	II	DI	NI	HC	SC	IC	DC	NC	CC
----	----	----	----	----	----	----	----	----	----	----

【 図 1 1 】



【 図 7 】

P

1	1	5	7	6
3	3	2	2	4
4	6	6	6	8
5	3	6	6	6
5	5	6	6	6

【 図 9 】

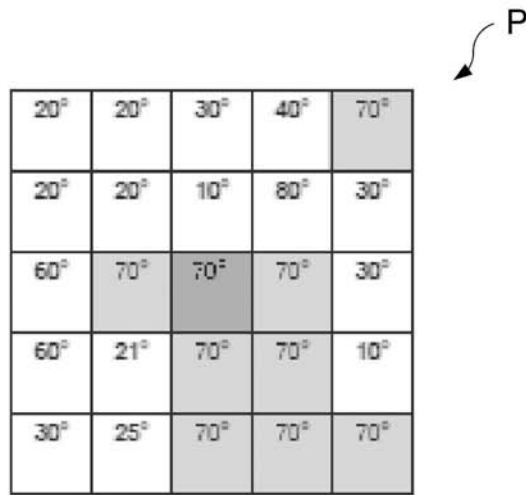


Diagram 9 shows a 5x5 grid of numerical values. The values are as follows:

20°	20°	30°	40°	70°
20°	20°	10°	80°	30°
60°	70°	70°	70°	30°
60°	21°	70°	70°	10°
30°	25°	70°	70°	70°

A force vector P is shown pointing towards the top-right corner of the grid.

フロントページの続き

(72)発明者 アレキサンダー シェン ライ ウォン

カナダ オンタリオ スカーボロー アレクスミア ブルバード 201 ユニット 72

Fターム(参考) 5B050 BA10 BA15 EA04 EA18 GA08

5B075 ND08 NR12

5L096 AA02 EA39 FA06 FA46 FA67 FA74 HA11 JA22 KA04 MA07