

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5675214号
(P5675214)

(45) 発行日 平成27年2月25日(2015.2.25)

(24) 登録日 平成27年1月9日(2015.1.9)

(51) Int.Cl.

G06T 7/00 (2006.01)

F 1

G06T 7/00 350B

請求項の数 14 (全 18 頁)

(21) 出願番号 特願2010-183391 (P2010-183391)
 (22) 出願日 平成22年8月18日 (2010.8.18)
 (65) 公開番号 特開2012-43156 (P2012-43156A)
 (43) 公開日 平成24年3月1日 (2012.3.1)
 審査請求日 平成25年8月13日 (2013.8.13)

(73) 特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 100076428
 弁理士 大塚 康徳
 (74) 代理人 100112508
 弁理士 高柳 司郎
 (74) 代理人 100115071
 弁理士 大塚 康弘
 (74) 代理人 100116894
 弁理士 木村 秀二
 (74) 代理人 100130409
 弁理士 下山 治
 (74) 代理人 100134175
 弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】情報処理装置、情報処理方法およびプログラム

(57) 【特許請求の範囲】

【請求項 1】

複数の判別器による結果を統合することによりパターンを判別するアンサンブル判別器を構成するそれぞれの判別器を生成するための情報処理装置であって、

入力手段を介して入力された複数のパターン間における特徴量の差異の分布を算出する算出手段と、

前記算出手段により算出された前記差異の分布を正規化することにより確率分布を生成する手段と、

前記生成された確率分布が示す確率に基づいて前記複数のパターンを判別するために参考するべきパターンにおける複数の画素を決定する決定手段と、

前記決定手段により決定された前記複数の画素の特徴量を用いて前記入力された複数のパターンを分類することにより、パターン判別に用いる判別器を複数生成する生成手段と、

を備えることを特徴とする情報処理装置。

【請求項 2】

前記複数のパターンを二つのクラスターに分割する分割手段と、

前記分割手段により分割された前記二つのクラスターからそれぞれ一つのパターンを選択して、選択されたパターン間の特徴量の差分を算出する手段と、

を更に備え、

前記決定手段は、前記差分の正の成分を正規化して求められた分布と、前記差分の負の

10

20

成分を正規化して求められた分布と、を用いて、前記複数のパターンを判別するために参考するべきパターン内の画素を決定し、

前記生成手段は、前記決定手段により決定された前記画素の特徴量に基づいて前記複数のパターンの判別を行うことにより前記複数の判別器を生成することを特徴とする請求項1に記載の情報処理装置。

【請求項3】

前記生成される判別器は決定木のノードを構成することを特徴とする請求項1または2に記載の情報処理装置。

【請求項4】

前記複数のパターンそれぞれの画素の特徴量に基づいて、前記複数のパターンの集合を10
2つのパターン群に分割する分割手段と、

前記分割手段により分割されたそれぞれのパターン群について特徴量のばらつきを算出し、前記算出された特徴量のばらつきの平均を算出する手段と、

を更に備え、

前記決定手段は、算出された前記特徴量のばらつきの平均を正規化して求められた確率分布を用いて、前記複数のパターンを判別するために参考するべきパターン内の画素を決定し、

前記生成手段は、前記決定手段により決定された前記画素の特徴量に基づいて前記複数のパターンの判別を行うことにより前記複数の判別器を生成することを特徴とする請求項1に記載の情報処理装置。20

【請求項5】

前記判別器は回帰木のノードを構成することを特徴とする請求項1または4に記載の情報処理装置。

【請求項6】

請求項1乃至5のいずれか1項に記載の情報処理装置の各手段としてコンピュータを機能させるためのプログラム。

【請求項7】

複数の判別器による結果を統合することによりパターンを判別するアンサンブル判別器を構成するそれぞれの判別器を生成するための情報処理装置における情報処理方法であつて、
30

前記情報処理装置の算出手段が、入力手段を介して入力された複数のパターン間における特徴量の差異の分布を算出する算出工程と、

前記情報処理装置の確率分布を生成する手段が、前記算出工程で算出された前記差異の分布を正規化することにより確率分布を生成する工程と、

前記情報処理装置の決定手段が、前記生成された確率分布が示す確率に基づいて前記複数のパターンを判別するために参考するべきパターンにおける複数の画素を決定する決定工程と、

前記情報処理装置の生成手段が、前記決定工程で決定された前記複数の画素の特徴量を用いて前記入力された複数のパターンを分類することにより、パターン判別に用いる判別器を複数生成する生成工程と、
40

を有することを特徴とする情報処理方法。

【請求項8】

前記算出手段は、対象物の姿勢を示す複数のパターン間における特徴量の差異の分布を算出することを特徴とする請求項1乃至5のいずれか1項に記載の情報処理装置。

【請求項9】

前記生成手段は、前記決定手段により決定された前記画素の特徴量と、予め設定された閾値と、の比較の結果を用いて、前記対象物の姿勢を求めるための判別を行う判別器を生成することを特徴とする請求項8に記載の情報処理装置。

【請求項10】

複数の判別器による結果を統合することによりパターンを判別するアンサンブル判別器
50

を構成するそれぞれの判別器を生成するための情報処理装置であって、

入力手段を介して入力された複数のパターン間における特徴量の差異の分布を算出する算出手段と、

前記算出手段により算出された前記差異の分布を正規化することにより確率分布を生成する手段と、

前記生成された確率分布が示す確率に基づいて前記複数のパターンを判別するために参考するべき前記入力された複数のパターンにおける特徴量の基底を決定する決定手段と、

前記決定手段により決定された前記特徴量の基底を用いて前記入力された複数のパターンを分類することにより、パターン判別に用いる判別器を複数生成する生成手段と、

を備えることを特徴とする情報処理装置。

10

【請求項 1 1】

前記基底は、特徴量の次元であることを特徴とする請求項 1 0 に記載の情報処理装置。

【請求項 1 2】

前記基底は、基底関数であることを特徴とする請求項 1 0 に記載の情報処理装置。

【請求項 1 3】

複数の判別器による結果を統合することによりパターンを判別するアンサンブル判別器を構成するそれぞれの判別器を生成するための情報処理装置における情報処理方法であつて、

前記情報処理装置の算出手段が、入力手段を介して入力された複数のパターン間における特徴量の差異の分布を算出する算出工程と、

20

前記情報処理装置の確率分布を生成する手段が、前記算出工程で算出された前記差異の分布を正規化することにより確率分布を生成する工程と、

前記情報処理装置の決定手段が、前記生成された確率分布が示す確率に基づいて前記複数のパターンを判別するために参考するべき前記入力された複数のパターンにおける特徴量の基底を決定する決定工程と、

前記情報処理装置の生成手段が、前記決定工程で決定された前記特徴量の基底を用いて前記入力された複数のパターンを分類することにより、パターン判別に用いる判別器を複数生成する生成工程と、

を有することを特徴とする情報処理方法。

【請求項 1 4】

30

請求項 1 0 乃至 1 2 のいずれか 1 項に記載の情報処理装置の各手段としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は情報処理装置、情報処理方法およびプログラムに関するものである。

【背景技術】

【0 0 0 2】

従来、パターン認識のための機械学習の技術として、複数の識別器で識別を行い、識別結果を統合することによって高精度な識別を行うアンサンブル学習法と呼ばれる手法がある（例えば、特許文献 1 を参照）。アンサンブル学習法は、推定の分散値の大きな識別器（弱識別器）でも、複数の識別器（弱識別器）を集めて多数決による識別を行うと推定値の分散が小さくなることを識別の原理としている。アンサンブル学習手法には、例えば、バギング、ブースティング、ランダムフォレストなどの手法がある。

40

【先行技術文献】

【特許文献】

【0 0 0 3】

【特許文献 1】米国特許第6,009,199号公報

【非特許文献】

【0 0 0 4】

50

【非特許文献 1】P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proc. of CVPR, 2001.

【非特許文献 2】A. Torralba, K. Murphy, W. Freeman, Sharing visual features for multiclass and multiview object detection, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 29, no. 5, pp. 854-869, 2007

【非特許文献 3】N. Dalal and B. Triggs, Histograms of Oriented Gradients for Human Detection, IEEE Computer Vision and Pattern Recognition, 886-893, 2005

【非特許文献 4】Choe, S. B. Faraway, J. J., Modelling Head and Hand Orientation During Motion Using Quaternions, SAE Transactions, VOL 113, PART 1, p.p. 186-192, 2004

【非特許文献 5】Qianjun Xu, Marie desJardins and Kiri L. Wagstaff. Active Constrained Clustering by Examining Spectral Eigenvectors. Proc of Discovery Science 2005.

【発明の概要】

【発明が解決しようとする課題】

【0005】

アンサンブル学習手法による画像認識では、弱識別器への入力として、例えば、Harr 基底関数（非特許文献 1）や局所パッチとのマッチング（非特許文献 2）のように、ある種の特徴量変換の基底関数が用いられることが多い。多くのアンサンブル学習手法では、弱識別器ごとに異なった基底関数を選択し、基底関数によって張られる空間上で学習データを判別学習している。その際、画像の一部分しか参照しないようなスパースな基底関数を用いるか、もしくはいわゆる integral image（非特許文献 1）等を用いることが多い。integral image を用いることによりデータの参照回数を抑えることができるので、高速な判別が実現できる。

【0006】

なお、特徴量の変換関数が小さな有限集合であれば弱識別器に与える最適な基底関数を総当たりで探索することができる。しかし、一般には変換関数の元は非可算無限であるため、総当たりで試行することは通常できない。そのような時には適当な範囲でランダムに変換関数を選んで基底関数を決定することが行われる。具体的にはランダムに十分な数の変換関数の候補が選ばれた基底のプールを作成し、プールの中から最適な基底関数を選択する方法等で近似される。

【0007】

ところで、パターン認識の応用課題の中には、対象物体の姿勢や形状の僅かな違いを識別するようなタスクがある。例えば、工業製品の組み立て自動化の際には、小さな切欠の有無で部品を識別したいという要望が生じることがある。また、マニピュレーターによる部品のピッキングために、部分的に似ても姿勢が全く異なるような部品を正しく見分けたいといった要望が存在することがある。このように空間上に疎にしか差異のないパターンを識別するタスクのことを、ここでは「似て非なるクラス」の識別問題と呼ぶ（ここで言う空間とは画像の二次元空間のみならず何らかの特徴空間も含む）。

【0008】

また明示的に「似て非なるクラス」ではない認識の課題においても、もし精度の向上を追求するのであれば、多かれ少なかれこのクラスの対象を適切に識別する必要が生じる。

【0009】

似て非なるクラスの識別をアンサンブル学習で行う場合、総当たりで最適な基底関数を選択する方法では、実際の差異ではなくノイズの方に過学習してしまって真に有効な基底を選べないことがありえる。また基底をランダムに選ぶ方法では、有効な基底を漏らしてしまうか、不十分な数しか選択できない恐れがある。アンサンブル学習法は手法の特性上、有効な弱識別器（判別正解率の期待値が 0.5 より大きい弱識別器）のバリエーションが十分多くないと汎化誤差が小さくならない。このため、できるだけ多数の基底の候補を含むプールと多数の弱識別器を用いることが望ましい。しかしそれでは学習および検出時

10

20

30

40

50

の計算量とデータ量が増大するという問題が生じる。

【課題を解決するための手段】

【0010】

上記の課題に鑑み、本発明は、特徴量空間の無数の基底関数の選択肢の中からパターン認識に有効な基底関数を優先的に選ぶことが可能な情報処理技術の提供を目的とする。

【0011】

上記の目的を達成するべく、本発明にかかる情報処理装置は、複数の判別器による結果を統合することによりパターンを判別するアンサンブル判別器を構成するそれぞれの判別器を生成するための情報処理装置であって、

入力手段を介して入力された複数のパターン間における特徴量の差異の分布を算出する算出手段と、

前記算出手段により算出された前記差異の分布を正規化することにより確率分布を生成する手段と、

前記生成された確率分布が示す確率に基づいて前記複数のパターンを判別するために参考するべきパターンにおける複数の画素を決定する決定手段と、

前記決定手段により決定された前記複数の画素の特徴量を用いて前記入力された複数のパターンを分類することにより、パターン判別に用いる判別器を複数生成する生成手段と、を備えることを特徴とする。

【発明の効果】

【0012】

本発明によれば、特徴量空間の無数の基底関数の選択肢の中からパターン認識に有効な基底関数を優先的に選ぶことが可能な情報処理技術の提供が可能になる。

【図面の簡単な説明】

【0013】

【図1】(a) 本発明の実施形態にかかるパターン認識方法を実行することが可能な情報処理装置の構成を説明する図、(b) パターン認識方法の処理の流れを概略的に示す模式図。

【図2】(a) 本実施形態のパターン認識方法の基本的な処理手順を説明する図、(b) クラス間の差異を模式的に示した図。

【図3】(a) 確率分布に従って画素を決定する処理を例示的に説明する図、(b) パターン認識過程における決定木の生成を例示的に説明する図。

【図4】(a) 第1実施形態のパターン認識処理の流れを説明する図、(b) 第1実施形態のパターン認識過程における決定木の生成を例示的に説明する図。

【図5】第1実施形態の変形例を説明する図。

【図6】第2実施形態のパターン認識処理の流れを説明する図。

【図7】第2実施形態の姿勢データのばらつきを算出する処理を模式的に説明する図。

【図8】第3実施形態のパターン認識処理の流れを説明する図。

【図9】第4実施形態における弱識別器の接続構成、およびパターン認識過程の結果の一例を表す図。

【図10】第4実施形態のパターン認識処理の流れを説明する図。

【図11】第2実施形態の学習過程における回帰木の生成を例示的に説明する図。

【発明を実施するための形態】

【0014】

(第1実施形態)

図1(a)の参照により、本発明の実施形態にかかるパターン認識方法を実行することが可能な情報処理装置の構成を説明する。ここでいう「パターン」とは例えば、静止画、動画、音声信号、文書データなどを認識の対象とすることが可能であるが、本発明の趣旨はこの例に限定されるものではない。具体例として、対象物を撮影した静止画を認識対象のパターンとし、対象物の名称、カテゴリ、3次元上の位置、姿勢等の情報を推定する例が挙げられる。ここでいう対象物とは、任意の物体を含み、例えば、人物、人物の顔、

10

20

30

40

50

臓器、工業製品、自動車、半導体基板などが含まれる。

【0015】

情報処理装置は、外部記憶装置101、中央処理装置(CPU)102、メモリ103、入出力装置104を備え、中央処理装置(CPU)102の全体的な制御の下にパターン認識方法を実行することが可能である。外部記憶装置101は、パターン認識方法の各処理ステップをコンピュータに実行させるためのパターン認識プログラムや、パターン認識方法の実行に用いる学習パターン、学習パターンを用いて作成された辞書等を保持する。また、外部記憶装置101は、パターン認識方法によって導かれた新規パターンの認識結果を保持することも可能である。中央処理装置(CPU)102はパターン認識プログラムの実行を司ることが可能である。

10

【0016】

また、中央処理装置(CPU)102は、外部記憶装置101、メモリ103、入出力装置104の動作を制御することが可能である。メモリ103は、中央処理装置(CPU)102によってパターン認識プログラムのサブルーチン、またはパターン認識プログラムの実行時に参照するべきデータを一時的に格納することが可能である。また、メモリ103はパターン認識方法によって導かれた新規パターンの認識結果を一時的に保持することも可能である。入出力装置104は、ユーザからの指示を受け付け、認識対象となる新規パターンを取り込む入力部と、パターン認識結果を他の情報処理装置や、例えば、ロボットを制御する制御装置などへ出力する出力部とを有する。認識対象のパターンとして2次元の画像を用いる場合、入出力装置104の入力部はカメラとして機能して、対象物体を撮影することも可能である。

20

【0017】

ユーザはパターン認識プログラムの実行開始のトリガーとして、入出力装置104を介して、実行を指示することも可能である。また、入出力装置104は、パターン認識結果を表示するための表示部を有しており、ユーザは、表示部に表示されたパターン認識結果を確認することも可能である。また、入出力装置104は、キーボードを有しており、ユーザはキーボードを介してパターン認識プログラムのパラメータを設定することが可能である。

【0018】

図1(b)は本発明の第1実施形態にかかるパターン認識方法の処理の流れを概略的に示す模式図である。まず、認識対象の形態を示すパターン(学習パターン)151において、学習パターン151間の特徴量の差異を基に確率分布152を作成する。そして確率分布152に従って、基底の選択を行い、特徴量の基底153a、b、c、…を弱識別器154a、b、c…ごとに選択してアンサンブル学習を行う。

30

【0019】

次に、図2(a)の参照により、本実施形態のパターン認識方法の基本的な処理手順を説明する。本実施形態では対象物体(部品)の姿勢推定を例とした分類問題を解くものとする。部品の画像が学習画像のパターンとして与えられたとき、その部品の姿勢をパターン認識方法の適用により識別する。ここでは数種類の姿勢のクラスを識別するタスクを想定する。学習パターンとしては、照明条件の違いや、大きさや傾きなど僅かな見え方の違いなど、一つの姿勢のクラスについて複数のバリエーションの画像が用意されているものとする。

40

【0020】

本発明は多数の識別器を用いるアンサンブル学習方式に広く適応が可能であるが、本実施形態ではその中の一つとしてランダムフォレスト方式に適応させることとする(特許文献1を参照)。ランダムフォレストは複数の決定木を用いてパターン認識を行う手法である。決定木の各ノードにおいて何らかのランダムな基準によって生成される情報の問い合わせを行い、問い合わせの結果に応じて学習パターンを次々と分割していく。葉ノードに達した時に残っている学習パターンをその葉ノードの分類結果として記憶する。判別時には、入力パターンに対して学習時と同じ情報の問い合わせを行ってノードをルートノード

50

から葉ノードへとたどる。葉ノードに達したら記憶されているパターンをその決定木の判別結果とする。全ての決定木の判別結果を投票等により統合して最終的な判別結果とする。

【0021】

学習は決定木ごとに行われる。まずステップS201で、全ての学習画像のパターン（パターン画像）を特徴量に変換する。本実施形態では照明変動に対して頑健な識別を行うことを目的としてラプラシアン・オブ・ガウシアン（Laplacian of Gaussian：以降「LoG」と表記する）の処理を施し、エッジの強度画像を得る。次に、ステップS202からS207までの処理を各決定木の各弱識別器に対して個々に行う。

【0022】

ステップS203では、パターン内において、クラス間の特徴量の差異の分布を計算する。ここで具体的に特徴量はLoGのエッジ強度であり、差異の値は姿勢のクラス内とクラス間の分散の比Fである。Fを以下の数式のように画素ごとに求める。

【0023】

【数1】

$$F(x, y) = \frac{\sum_{c} n_c \cdot (m(x, y) - m_c(x, y))^2}{\sum_{c} \sum_{I \in C} (m_c(x, y) - I(x, y))^2}$$

10

20

【0024】

但し、 $I(x, y)$ は座標 x, y の画素のエッジ強度、 $m_c(x, y)$ は姿勢クラスCのエッジ強度の平均値、 $m(x, y)$ は全パターンのエッジ強度の平均値、 n_c は各姿勢クラスの画像パターンの個数である。

【0025】

ここで、図2(b)の参照番号220は異なる姿勢クラス間の差異を模式的に示した図である。姿勢クラス1の画像221と姿勢クラス2の画像222との分散比を求めて二つのクラスの差異の分布223を求める。確率的選択の結果を示す画像224は差異の分布を基に確率的に選択された画素の位置を複数示している。ここでは図示を簡単にするために二つの姿勢のクラスのみを示しているが、本実施形態では実際には判別対象のパターンの全姿勢のクラスについて分散比を算出する。なお、ルートノードなどで対象となる学習パターンの数が多くて処理時間がかかるようなときは、学習パターンの集合からいくつかのパターンをランダムサンプリングして小さなサンプルで計算を行うようにしてもよい。

30

【0026】

次にステップS204において、クラス間の差異の分布Fを基にして確率分布Pを作成する。例えば、ここでは以下のようないふれ式を用いればよい。

【0027】

【数2】

$$P(x, y) = \frac{1}{Z} \exp[-\frac{\rho}{F(x, y)}]$$

40

【0028】

ただし、Zは $P(x, y) = 1$ となるための正規化パラメータ、 ρ は分散比の高低を確率分布にどの程度反映するかを決定するチューニングパラメータである。 $\rho = 0$ ならば確率は一様分布に近づく。すなわち分散比の高低を考慮しない分布になる。

【0029】

次にステップS205において、確率分布Pを用いて各ノード（各弱識別器）が判別のために参照する基底を決定する。通常のランダムフォレストにおいて、基底とは、いずれ

50

かの画素の値（特徴量）であり、ランダムに参照する画素を確率分布 P に従って決定する。確率分布による決定の方法として、具体的な例を図 3 (a) に図示する。まず確率分布 301 の画素ごとの確率 $P(x, y)$ の値をラスター方向に加算し、累積値のテーブル 302 を生成しておく。0 ~ 1 までの乱数を発生させて、乱数の値に対応する画素を選択する。図 3 (a) の場合、乱数 0.8123 に最も近い値 0.86 を有する累積テーブルの画素 I(4, 2) が選択される。

【0030】

次に、ステップ 206において、識別器の学習を行う。ここでは画像の一点のエッジ強度の値（特徴量）と、あらかじめランダムに設定した閾値とを比較する。比較の結果、学習画像のパターンを二つに分類して二つの新しいノードを生成する。これが各識別器における学習となる。そして分割された各パターンについて以降のノードでも、先に説明したステップ S203 から S206 を繰り返す。このようにして決定木が所定の深さに達するか、分割されたパターンの数が所定の数以下になればそのノードを葉ノードとする。葉ノードには残っている学習パターンの姿勢クラスの頻度を記憶する。

10

【0031】

本実施形態の動作の例を図 3 (b) に示す。ここでは簡単のために深さが 4 の小さなツリーを示す。なお、無意味な比較が行われる回数を減らすために、閾値は学習データの分布を参考にして適切に設定しておいてもよい。決定木 1 の各ノードにおいて、それぞれ画素 A, B, C, D, E と閾値 $1, 2, 3, 4, 5$ とを比較してその判定結果により分岐し、最後に葉ノードに達して姿勢の判別結果を得ることができる。決定木 2, 3 についても同様にして、姿勢の判別結果を得ることができる。

20

【0032】

次に図 4 (a) の動作フローチャートを用いて本実施形態の判別処理について詳述する。判別処理の動作と結果の一例を図 4 (b) に示す。判別処理は入力画像を各決定木で独立に判別させ、その結果を統合することで実行される。まずステップ S408 では学習時に行った前処理と同様の手法（ここでは LOG）を用いて入力画像を特徴量に変換する。次に決定木のルートノードを初期位置とする。そして各ノードでステップ S409 からステップ S412 までを繰り返す。

【0033】

ステップ S410 では、そのノードについて学習時と同じ画素を選択して値を観測する。そしてステップ S411 では、学習時にランダムに決定した閾値との比較を行う。図 4 (b) に示した動作例では決定木 1 の各ノードにおいて、それぞれ画素 A, B, E と閾値 $1, 2, 5$ とを比較してその判定結果により左、右、左と分岐し、最後に葉ノードに達して姿勢 1 という判別結果を得ている。

30

【0034】

本実施形態で用いる決定木の場合、各ノードで分類対象となるパターンの集合はパターンの分割によって刻々と変化するため、パターンの差異の値や差異に基づいて作成される確率分布が各ノードで変化していくことに注意が必要である。ルートノードから葉ノードに近づくにつれ、パターンは分岐によって細分化されていき、パターンの差異に特化した詳細な比較が行われる。これにより効率的にパターンの分類を行うことができる。

40

【0035】

また、汎化誤差を避けるために、適宜、チューニングパラメータを調整するような工夫があつてもよい。例えば、ノードの深さに応じての値を変えてもよい。具体的には判別の初期のノードでは $= 0$ として、あまりパターン間の差異を考慮しないようにし、徐々にの値を大きくしていくような方法が考えられる。また別の工夫として、例えば半分のノードは $= 0$ として確率分布を考慮せずに基底を選び、半分のノードは $= 1$ として確率分布を考慮して基底を選ぶ。これを交互に繰り返すというような工夫も考えられる。

【0036】

なお、本実施形態では特徴量に変換する前処理として LOG を使ったが、本発明の趣旨はこれに限るものではない。パターンのクラスを識別するために有効な特徴量の変換であ

50

ればどのような特徴量変換も候補となりうる。認識対象のパターンや条件によって適切なものを選べばよい。例えば、HOG (Histogram of Oriented Gradient) (非特許文献3を参照)などを選ぶことも可能である。本実施形態の变形例について、以下に簡単に説明する。例えば、図5に模式図を示すように部品の学習画像のパターンセット501をそれぞれ 7×7 のセルからなるHOG特徴量に変換する。各セルが8方位の輝度勾配の値を持つので一つの学習画像のパターンは392次元の特徴ベクトル502に変換される。この392次元の特徴量について次元ごとに全パターンの分散比を求めて分散比の分布503を作成し、先述と同様の方法で確率分布504を作成する。各識別器では確率分布504によってHOG特徴量の次元の一つを基底として選び、ランダムな値との比較を行つて学習データを分割していくべき。

10

【0037】

なお、本実施形態では確率分布の元となるパターンのクラス間の差異を表す指標として分散比を用いたが、これは他の指標であってもよい。具体的には、特徴量の値XとクラスのラベルYの2つを変量とする相互情報量I(X;Y)を利用することも考えられる。また、1クラスに1パターンしかデータがなければ、特徴量のクラス間の標準偏差を利用することも考えられる。正規化を適切に行うことで、様々な指標に対して本発明の適用が可能である。実際には用いる特徴量や認識タスクなどを勘案して有効な指標を選択することになる。

【0038】

(第2実施形態)

20

第1実施形態では姿勢識別の問題をアンサンブルの分類木を用いることで分類の問題として解いた。しかしこのためにはあらかじめ姿勢を有限個のクラスターに分割しておく必要があり、クラスタリングの方法の妥当性や最適な角度の分解能が不明といった問題が生じる。第2実施形態では姿勢推定問題をクラスの分類問題としては解かず、回帰木を用いて姿勢の数値を直接推定する。

【0039】

回帰木は関数の近似に用いられる手法である。構造は分類木とほぼ同じであるが、葉ノードにはパターンのクラスではなく、入力に対して関数が出力すべき数値を対応づける点が異なる。本実施形態では入力を部品の画像とし、出力を物体の姿勢の値である極角、偏角とする。ここでは説明の簡単のために姿勢の面内回転の違いは存在しないと仮定して二自由度の姿勢の推定を行うタスクを考える。

30

【0040】

通常の回帰木では、特徴量の候補を総当たりで試してデータを分割してみて、最適な特徴量を選んで決定木の分岐を行う。具体的には推定したい変量(ここでは姿勢の値)の分割後の二群内のばらつきが最も小さくなるような特徴量を選択するのが一般的である。そして各ノードでこれを繰り返して分岐とデータの分割を行っていく。

【0041】

本実施形態のアンサンブル学習手法では、最適な特徴量を基底として選択するのではなく、分割後のばらつきの値を評価して有効な特徴量が優先して選ばれるように確率分布を作成する。そして確率分布を基に特徴量を選択して基底とする。本実施形態の処理の例を図11に示す。ここでは簡単のために深さが4の小さなツリーを示す。なお、無意味な比較が行われる回数を減らすために、閾値は学習データの分布を参考にして適切に設定しておいてもよい。決定木1の各ノードにおいて、それぞれ画素A, B, C, D, Eと閾値
1, 2, 3, 4, 5とを比較してその判定結果により分岐し、最後に葉ノードに達して姿勢の値の判別結果を得ることができる。決定木2、3についても同様にして、姿勢の値の判別結果を得ることができる。

40

【0042】

図6を参照して第2実施形態のパターン認識処理の流れを説明する。まず、ステップS601では第1実施形態と同様にLOGによって前処理を行う。そしてノードごとに、どの画素の値に着目して判別対象のパターンを分割したら、どの程度有効な分割が行われる

50

のかを評価する。具体的にはステップ S 604 で、適当な閾値である画素の値（これを基底 f_i ）との比較により、特徴量である画素の値が大きいパターン群と特徴量である画素の値の小さいパターン群の二つの領域に分割する（第1のデータセット、第2のデータセット）。次にステップ S 605 で分割後の二つのパターン群において、画素の値（特徴量）として、姿勢の方位のばらつきの指標 V を算出する。ばらつきの指標 V については後述する。ばらつきの指標 V をその画素の基底としての有効度として画素ごとに記憶する。なお、上記の分割の閾値として、例えば、二分割されたデータの個数がなるべく等分になる値を用いる。または、指標 V が最大になる値を用いてもよい。

【0043】

ばらつきの指標 V の算出の方法を以下に説明する。ここでは図7(a)のように姿勢の空間として2次元球面 S^2 とその上に分布している姿勢のデータ群を考える。姿勢は極角 θ 、偏角 ϕ で表される。極角 θ 、偏角 ϕ から座標 x, y, z への変換は、以下の数3の式により求めることができる。

【0044】

【数3】

$$\begin{aligned}x &= r \sin \theta \cos \phi \\y &= r \sin \theta \sin \phi \\z &= r \cos \theta\end{aligned}$$

10

20

【0045】

この座標の値を用いて n 個の姿勢データから姿勢の平均ベクトルを合成すると、長さは、以下の数4の式により求められる。

【0046】

【数4】

$$L = \frac{1}{n} \left(\left(\sum_i x_i \right)^2 + \left(\sum_i y_i \right)^2 + \left(\sum_i z_i \right)^2 \right)^{\frac{1}{2}}$$

30

【0047】

ここで、以下の数5の式で定義される値は n 個の姿勢データのばらつきを反映した値である。

【0048】

【数5】

$$v = 1 - L$$

【0049】

$v = 0$ で姿勢の方位のばらつきが最小であり、 $v = 1$ でばらつきが最大となる。データを n_1 個と n_2 個の二つに分割したときの二つの群（第1のデータセット、第2のデータセット）のばらつきの平均は、以下の数6式により求めることができる。平均のばらつき V を各基底の判別の有効度の変量とする（図7(b)）。平均のばらつき V は、基底 f_i の有効度を示す。

40

【0050】

【数6】

$$V = \frac{(n_1 - 1)v_1 + (n_2 - 1)v_2}{n_1 + n_2 - 2}$$

【0051】

50

以降の処理は、第1実施形態と多くの部分について共通である。ステップS607において、変量V(x,y)から確率分布P(x,y)を作成する。ステップS608において、確率的に画素を選択して基底とし、ステップS609において、回帰木の分岐とデータの分割を行う。葉ノードには、残っている姿勢のデータを平均して平均の方位を求めてこれを記憶する。姿勢の平均の方法は、例えば、先述のように平均ベクトルを合成し、その極角と偏角を求めて姿勢の値とすればよい。

【0052】

姿勢の推定時には、入力画像を与えて各回帰木を辿り、各回帰木について姿勢の値を得る。これを平均して最終的な姿勢の推定結果とする。ここで、出力結果に外れ値が含まれていても頑健に結果が得られるように、M推定等の一般的なロバスト推定を用いてよい。

10

【0053】

ここでは面内回転を考慮せずに二自由度の姿勢推定について説明したが、別の実施形態として、面内回転を含む三自由度の自由姿勢を推定する形態が考えられる。基本的な形態は二自由度のときと同じであるが、三自由度の場合は四元数を使って姿勢の方位を表す必要がある。分割後の姿勢のばらつきを評価するためには姿勢の方位間の差分と平均を定義する必要がある。

【0054】

これは以下のようにして定義する。まず二つの姿勢の差分である角度差は姿勢の方位を長さ1の四元数をqで表し、別の姿勢pに対して以下の数7式として定義する。

20

【0055】

【数7】

$$\text{dist}(q, p) = \arccos(q \cdot p)$$

【0056】

次に四元数による姿勢の平均の計算方法は近似的な方法を含めていくつか提案されている（例えば、非特許文献4）。これらを用いて先述の方法と同様の計算を行えば分割後の姿勢のばらつきを評価することができる。このようにして三自由度姿勢についても本実施形態の適用が可能である。

【0057】

30

（第3実施形態）

第3実施形態では本発明の適用可能な決定木を用いて3つ以上の物体の種類の識別について説明する。物体としては人工物から一般物体まで様々な対象が考えられるが、ここでは具体的に工業用部品同士を例にして識別することを考える。条件としては多数の部品の種類が存在し、それぞれについて学習画像が一枚ずつ与えられているとする。すなわちここでは1クラス1パターンの多クラスのパターン識別を解くことになる。学習データには図2(b)の部品クラス211と部品クラス212のように、部分的には似通っていて、切欠の有無などで適切に見分ける必要があるようなものが存在すると仮定する。

【0058】

本実施形態では第1実施形態と同様に複数の決定木を分類木のアンサンブルとして使う。本実施形態では2点の画素の値を比較することで、決定木を分岐させていくことにする。

40

【0059】

簡単な例としてもし2クラスのパターンを見分けるのであれば、クラス間でパターンの差分を求め、正の値の領域のどこか1点の画素と、負の値の領域のどこか1点の画素を比較すれば、比較的信頼性の高い判別を行うことが期待できる。但し本実施形態では1クラス1パターンの多クラスを扱うため、後述するような工夫を用いる。

【0060】

図8の参照により第3実施形態のパターン認識処理の流れを説明する。本実施形態における学習時の処理の流れを説明する。本実施形態では、まずステップS801において、

50

入力画像に対して L o G 等の前処理を行う。ステップ S 8 0 2 から S 8 1 1 までの処理を繰り返す。次に前処理で得られた特徴量の値を基にステップ S 8 0 3 において、k - 平均クラスタリングを用いて学習パターンを便宜的に二つのクラスターに分割する。次にステップ S 8 0 4 では、先のステップ S 8 0 3 で分割された二つのクラスターからランダムに各一つずつパターンを選択してペアを作成し、ステップ S 8 0 5 において、二つのパターンの画素ごとの差分値を計算する。ステップ S 8 0 6 において、正の成分と負の成分とにそれぞれ分けて個別にその絶対値を加算して、加算結果を保存する。ステップ S 8 0 4 ~ ステップ S 8 0 7 を所定の数だけ繰り返して、正の値の成分と負の値の成分の分布の二枚の画像を得る。これを第 1 実施形態で行ったのと同様の手法によりステップ S 8 0 8 で正規化して二つの確率分布を作成する。

10

【 0 0 6 1 】

そして、ステップ S 8 0 9 ではこの二つの確率分布に従って 2 点を確率的に決定する。確率分布に基づく累積マップからの確率的な選択の方法も第 1 実施形態と同様に行えばよい。そして、ステップ S 8 1 0 では 2 点の特徴量の比較を行って、対象物の認識を行うため、比較の結果に従って学習パターンの分割とツリーの分岐を行う。このときには先ほどのクラスターの分割結果は考慮せずに新たな分割を行う。以上のように分割を進め、一つのパターンのみがノードに残るまでツリーを分岐する処理を繰り返す。一つのパターンのみのノードに到達したらパターンの種類を記憶して葉ノードとする。なお、途中でどちらかのノードのパターンが空になるような問い合わせは無効として選択をやり直す。

【 0 0 6 2 】

20

なお、クラスタリングの際には計算量低減のため、全データを使わずに部分的にランダムサンプリングしたセットでクラスタリングを行ってもよい。また、クラスタリングの方法は k - 平均クラスタリングに限らず、例えばスペクトラルクラスタリング手法（非特許文献 5 を参照）等を使ってもよい。データの性質と計算量とのトレードオフを考慮しながら選択すればよい。また、識別器による 2 点の比較の際には、第 1 実施形態で閾値を設けたようにどちらかにランダムな閾値を足してから比較するような形態にしてもよい。以上が本実施形態のパターン認識手法の学習の動作フローである。判別時には、こうして形成された複数の決定木をそれぞれ独立に辿り、葉ノードに達したときのパターンをその決定木の判別結果とする。全ての決定木で多数決をとって出現頻度の最大のパターンを判別結果とする。本実施形態の変形として、例えば、第 1 実施形態と同様に各ノードで 1 点の画素情報をランダムな値と比較するような形態でもよい。画像のクラス間の差異としては、例えば、パターンの各画素の標準偏差を利用して確率分布を作成することも可能である。

30

【 0 0 6 3 】

（第 4 実施形態）

第 4 実施形態としては、本発明を人物の顔の検出タスクに適用した例を説明する。ここでの問題設定の枠組みおよび識別器の構成は基本的には非特許文献 1 の Violaらの形態に則るものである。Violaらは Adaboost 手法（非特許文献 1 参照）で弱識別器を学習し、図 9 (a) に示すいわゆるカスケード構造で多段の判定を行った。これにより顔と非顔の 2 クラスのパターンの識別を行っている。また、判別の閾値を大きく下回るような、明らかに顔でないと判定できるパターンに関しては各段階において棄却を行うため、高速な判定が可能である。発明者はこれに本発明を適用して改良を行う。

40

【 0 0 6 4 】

Violaらは弱識別器に与える特徴量として Haar 基底関数を用いている。24 × 24 の解像度の画像で最大 4 のサブドメインを持つ Haar 基底関数の場合、4 万個程度のバリエーションが存在する。Violaらはこの全バリエーションを総当たりで試し、その時点でもっとも誤判別率が低くなる Haar 基底を選択している。もし画像の解像度を 24 × 24 よりも高くするか、より複雑なパターンの Haar 基底を用いようとする、バリエーションの数が膨大になって現実的な時間では総当たりで試すことはできなくなる。

【 0 0 6 5 】

50

そこで本実施形態ではパターン間の特徴量の差異を元に確率分布を作成して有効な基底が優先的に選択されるようにする。図10の参照により、第4実施形態のパターン認識処理の流れを説明する。また学習時のある一つのノードの動作の例を図9(c)に模式図として示す。以下では特に非特許文献1との相違の部分に絞って説明する。

【0066】

まず、クラス間の差異の分布を求める。本実施形態ではステップS1003で顔クラスの学習画像902(正事例学習画像)と非顔クラスの学習画像903(負事例学習画像)のデータに対して線形判別分析を行う。線形判別分析の結果、2クラスの分散比を最大化する一つの基底が得られる。この基底はn個の画素に対応するn次元の空間を1次元に圧縮する変換基底である。変換基底の係数の結果例を904に示す。係数904は実数値であり、正の値と負の値があることに留意されたい。10

【0067】

この係数とよく一致するようなHaar基底を優先的に選択すれば、2つのクラスを効率的に判別できることが期待できる。これについては様々な実現の方法が考えられるが、本実施形態では例えば以下のようない操作で実現する。ステップS1004でまず、n個のHaar基底をランダムに生成する。1つのHaar基底の例を905に示す。

【0068】

次に、ステップS1005においてこのn個のHaar基底と、線形判別分析の基底の係数904との一致度のスコアを測る。具体的にはスコアとしては例えば正規化相関(相互相関)の値を求めればよい。そして正規化相関の絶対値の大きな基底が有効な基底と考え、これらの基底がn個の中から優先されて選択されるようにする。ステップS1006では第1実施形態で行ったのと同様の方法を用いて一致度のスコアの値を正規化して確率分布に変換する。そしてS1007では第1実施形態で行ったのと同様にして累積分布のテーブルを作成して、確率的に一つのHaar基底を選択して判別(S1008)のための基底906とする。このようにした後、棄却閾値を決定し(S1009)、棄却された学習データを除き(S1010)、学習データの重み付けを更新する(S1011)。学習は、非特許文献1と同一の方法でもってAdaboostにより行う。誤判別データの重み付の仕方や判別閾値907の設定の仕方については非特許文献1に十分に開示されているのでここでは説明を省く。このようにして識別器を学習させる。20

【0069】

検出時には、非特許文献1と同様にいわゆるスライディングウインドウによる走査処理を行って画像の全領域について顔か否かの判定を行う。また異なった見えのサイズに対応するために、複数の解像度の階層画像を用いて複数回の判定を行う。30

【0070】

なお、本実施形態では基底関数を一つ選んで1次元空間上で判別と閾値の決定を行ったが、本発明の適用可能な実施形態においては弱識別器が判別を行う空間は1次元に限定されない。例えば、二つ以上の基底関数を選んで多次元の空間を張り、データの判別を行うような形態も考えられる。これにより、より複雑なHaar基底の組み合わせの表現を実現することが可能である。弱識別器の判別の手法としては、例えば、ソフトマージンのサポートベクターマシンなどを用いることができる。なお、組み合わせの効果を得るために、Haar基底関数で変換した後にさらに2乗するなど、非線形な変換を含む方が好ましい。本実施形態ではViolaの手法に順じて、図9(a)のように弱識別器のカスケードを一つのみ用いる実現形態を説明したが、図9(b)のようなカスケードのアンサンブルを用いて判別結果を多数決で統合するような形態も考えられる。この形態では基底選択のランダム性をより有效地に用いてアンサンブル学習が行われることが期待できる。40

【0071】

上述の実施形態によれば、特徴量空間の無数の基底関数の選択肢の中からパターン認識に有効な基底関数を優先的に選ぶことが可能な技術の提供が可能になる。

【0072】

また、基底関数の選択の仕方が決定的でないために弱識別器の多様なバリエーションが50

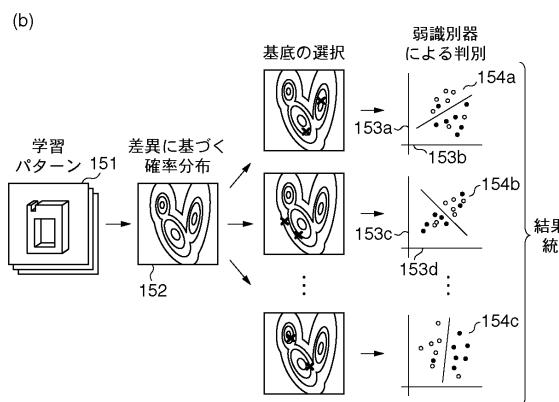
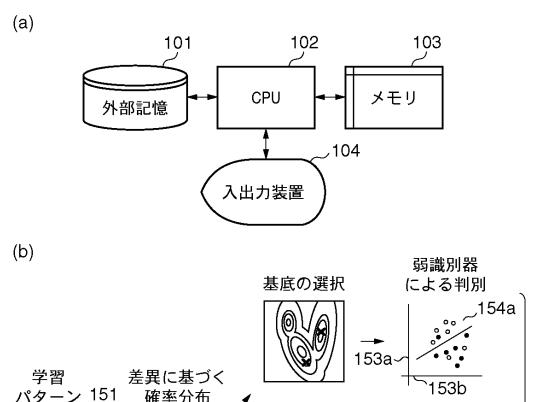
保持される。このようにして結果的に従来のアンサンブル学習手法に比べて、識別の精度を高めることができる。また、学習時および検出時の計算量やデータ量をより小さくすることができる。

【0073】

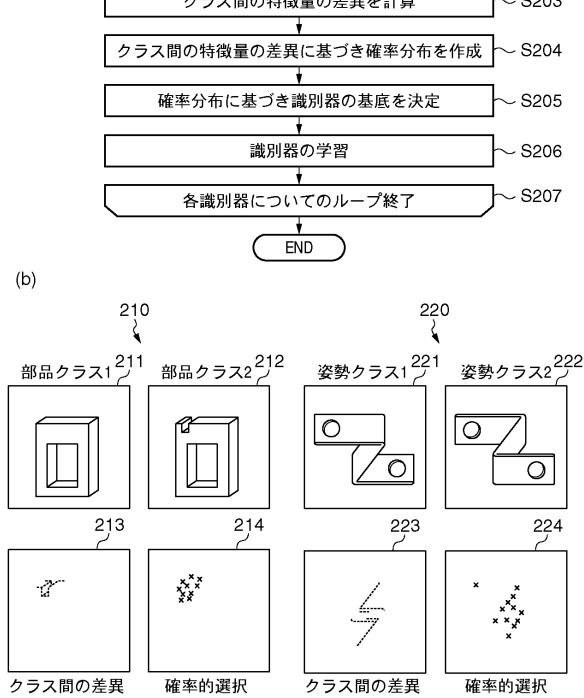
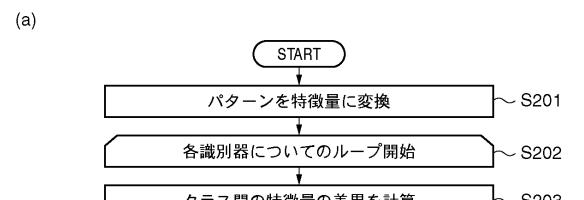
(その他の実施例)

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア（プログラム）を、ネットワーク又は各種記憶媒体を介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ（またはCPUやMPU等）がプログラムを読み出して実行する処理である。

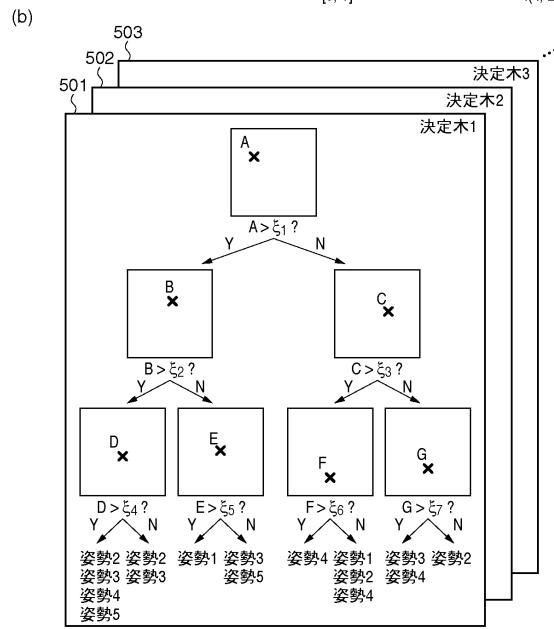
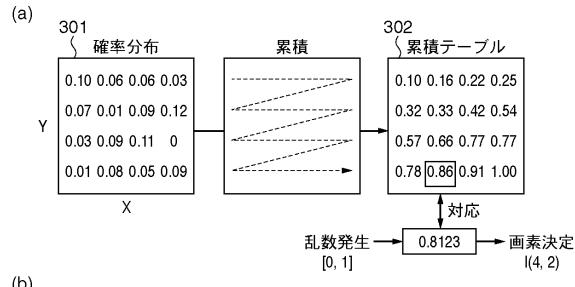
【図1】



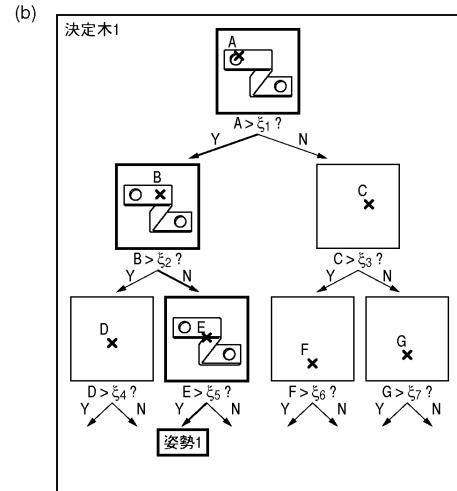
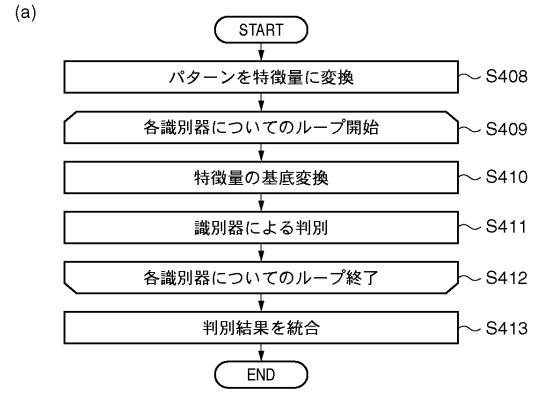
【図2】



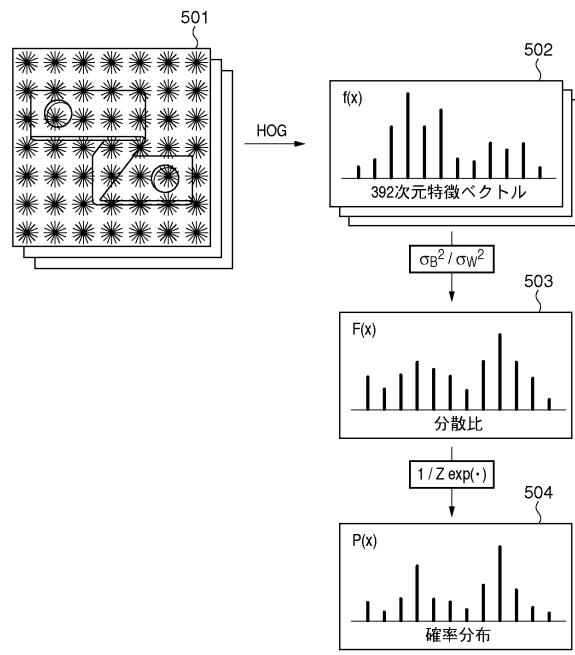
【図3】



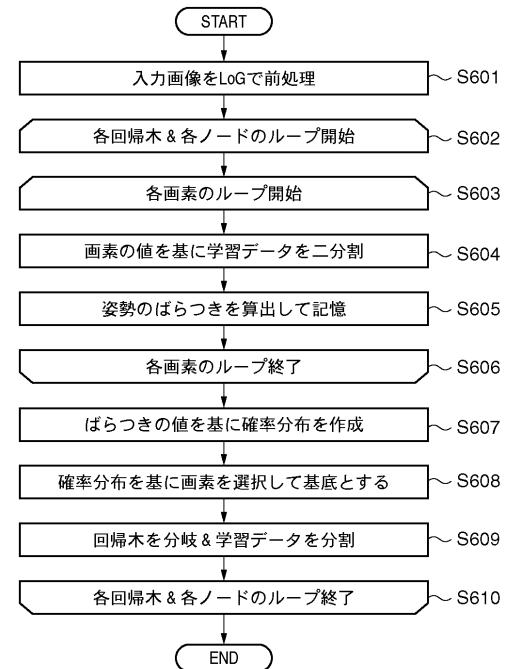
【図4】



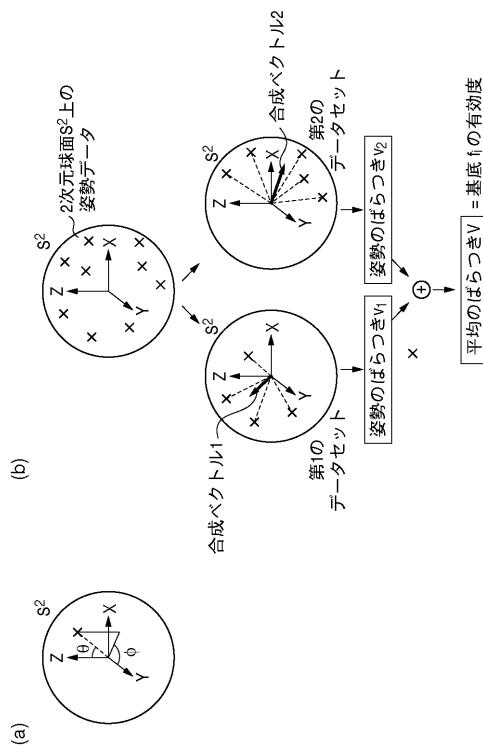
【図5】



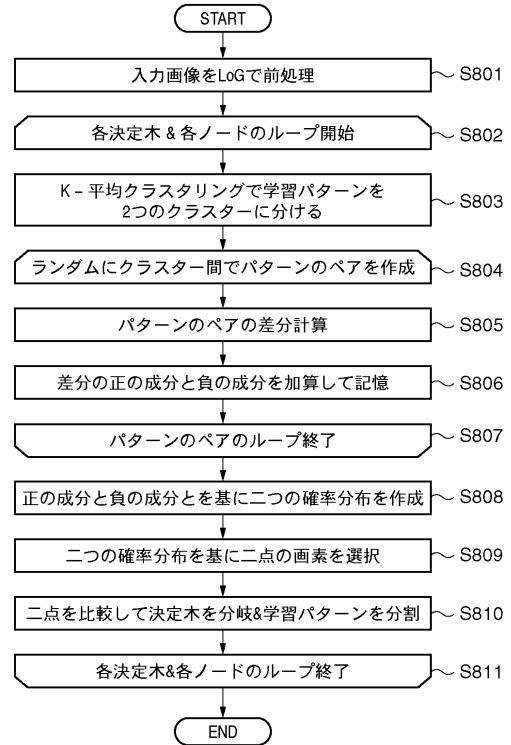
【図6】



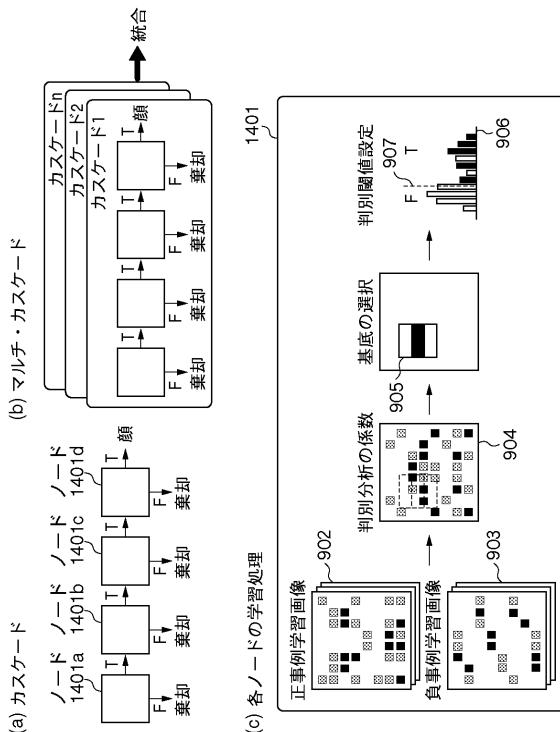
【図7】



【図8】



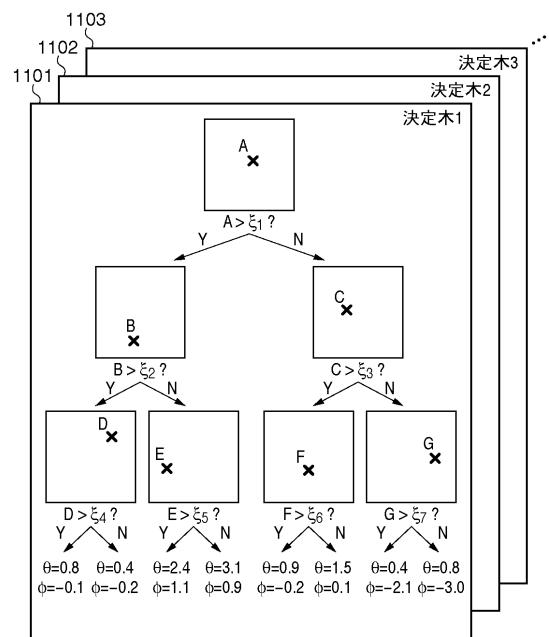
【図9】



【図10】



【図 1 1】



フロントページの続き

(72)発明者 館 俊太
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 御手洗 裕輔
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 吉井 裕人
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 奥野 泰弘
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 森 克彦
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 真継 優和
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 佐藤 実

(56)参考文献 増田 真平, 投機的変数選択法によるオンラインクラス分類問題の高速学習, 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, 2007年 3月 8日, 第106巻 第589号
, p.129-134
新村 秀一, ORと統計の融合 数理計画法を用いた最適線形判別関数(5), 経営の科学 オ
ペレーションズ・リサーチ, 日本, 社団法人日本オペレーションズ・リサーチ学会, 2002年
5月 1日, 第47巻 第5号, p.315-321

(58)調査した分野(Int.Cl., DB名)

G 06 T 7 / 00