

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2009年11月5日(05.11.2009)

PCT

(10) 国際公開番号
WO 2009/133856 A1

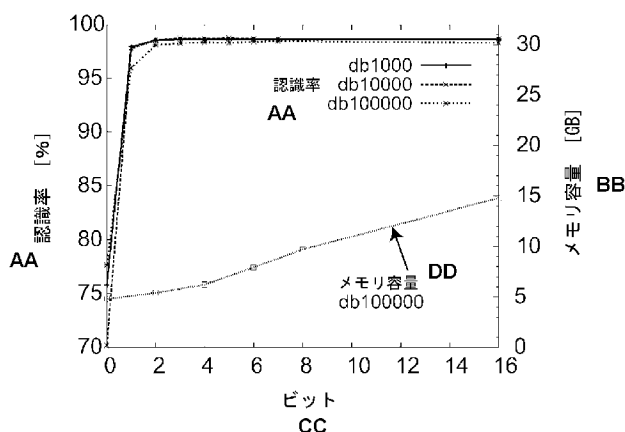
- (51) 国際特許分類:
G06T 7/00 (2006.01)
- (21) 国際出願番号: PCT/JP2009/058285
- (22) 国際出願日: 2009年4月27日(27.04.2009)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2008-117359 2008年4月28日(28.04.2008) JP
特願 2008-119853 2008年5月1日(01.05.2008) JP
- (71) 出願人 (米国を除く全ての指定国について): 公立大学法人大阪府立大学(OSAKA PREFECTURE UNIVERSITY PUBLIC CORPORATION) [JP/JP]; 〒5998231 大阪府堺市中区学園町1-1 Osaka (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 野口 和人 (NOGUCHI, Kazuhito) [JP/JP]; 〒5998231 大阪府堺市中区学園町1-1 公立大学法人大阪府立大学内 Osaka (JP). 黄瀬 浩一 (KISE, Kouichi) [JP/JP]; 〒5998231 大阪府堺市中区学園町1-1 公立大学法人大阪府立大学内 Osaka (JP). 岩村 雅一 (IWAMURA, Masakazu) [JP/JP]; 〒5998231 大
- (74) 代理人: 野河 信太郎 (NOGAWA, Shintaro); 〒5300047 大阪府大阪市北区西天満5丁目1-3 南森町パークビル 野河特許事務所 Osaka (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF,

[続葉有]

(54) Title: METHOD FOR CREATING IMAGE DATABASE FOR OBJECT RECOGNITION, PROCESSING DEVICE, AND PROCESSING PROGRAM

(54) 発明の名称: 物体認識用画像データベースの作成方法、処理装置および処理用プログラム

[図5]



AA RECOGNITION RATIO
 BB MEMORY CAPACITY
 CC BIT
 DD MEMORY CAPACITY

(57) Abstract: Provided is a method for creating an image database for object recognition, which comprises a feature extraction step of extracting local descriptors from object images which are to be registered in an image database, a scalar quantization step of quantizing a numeric value indicating each dimension of each of the local descriptors into a predetermined number of bit digits, and a registration step of systematizing each of the local descriptors after the quantization to be able to be searched for in the closest vicinity, giving to the local descriptor an identifier of the image from which the local descriptor has been extracted, and registering the local descriptor to which the identifiers are given in the image database. The registration step comprises extracting the local descriptors from the object images when a search query is given, scalar-quantizing each dimension, determining a local descriptor in the closest vicinity of each of the local descriptors from the image database, and performing registration so as to be able to identify one image by majority vote processing from the images including any determined local descriptor. The scalar quantization step comprises quantizing each dimension of each of the local descriptors into 8 bits or less. Also provided are a processing program for the method and a processing device for performing the processing.

processing device for performing the processing.

(57) 要約:

[続葉有]

WO 2009/133856 A1



CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, 添付公開書類:
TG).

— 国際調査報告 (条約第 21 条(3))

画像データベースに登録すべき物体の画像から局所記述子を抽出する特徴抽出工程と、各局所記述子の各次元を表す数値を所定のビット桁数に量子化するスカラー量子化工程と、量子化後の各局所記述子を最近傍探索可能に体系化しかつそれが抽出された画像の識別子を付与して画像データベースに登録する登録工程とを備え、前記登録工程は、検索質問が与えられたときそれから局所記述子を抽出して各次元をスカラー量子化し、各局所記述子に対し最近傍の局所記述子を前記画像データベース中から決定し、決定されたいずれかの局所記述子を含む画像の中から多数決処理によって一の画像を特定し得るよう登録を行い、前記スカラー量子化工程は、各局所記述子の各次元を 8 ビット以下に量子化する物体認識用画像データベースの作成方法、その方法の処理用プログラムおよび当該処理を行う処理装置。

明 細 書

発明の名称：

物体認識用画像データベースの作成方法、処理装置および処理プログラム

技術分野

[0001] この発明は、物体認識用画像データベースの作成方法、処理装置および処理プログラムに関する。より詳細には、画像に係るデータベースの中から、検索質問（クエリー）としての画像に表された対象物と一致する対象物を示す画像を検索できるような画像データベースの作成方法、処理装置および処理プログラムに関する。

背景技術

[0002] SIFT (Scale-Invariant Feature Transform) などの局所記述子を用いると、隠れや照明条件の変動に比較的頑健な物体認識が実現可能であるため、現在、注目を集めている（例えば、非特許文献 1、2 参照）。認識の基本は、“Bag of Words”あるいは“Bag of Features”と呼ばれるモデルであり、局所記述子の配置や共起を考慮せず、その頻度のみを手がかりとして物体を認識するものである。

[0003] ここで、局所記述子は、画像の局所的な特徴を表現したものであり、画像の変動（幾何学的変換や照明条件、解像度の差異）に対して頑健（ロバスト）な特性を有するよう、所定の手順で抽出される。また、局所記述子は、画像の局所的な領域から決定されるので、隠れに対してもロバストである。この明細書において局所記述子は特徴ベクトルともいう。ベクトル量で表現されるからである。

[0004] 一般に、画像から抽出される局所記述子の数は、数百から数千、場合によっては数万となるため、局所記述子の照合に必要な処理時間や、記憶に必要なメモリ容量は膨大となる。従って、認識精度を一定レベルに保ったまま、いかに処理時間やメモリ容量を削減するかが重要な研究課題となっている。

- [0005] 例えば、代表的な局所記述子のSIFTでは、各局所記述子は128次元のベクトルとして表される。また、SIFTに主成分分析を施すことにより次元数を削減したPCA-SIFTが知られている。それでも、実用的なPCA-SIFTの次元数は一例で36次元である。そして、各次元の数値を表すデータ型としては、一般的な数値表現に適用される32ビットのfloat型あるいはint型が一般的である。より高い精度が必要な場合は64ビットのdouble型が用いられる。一方、数値の範囲が限られる場合あるいは精度を犠牲にしてもメモリ容量を削減したい場合は特別に16ビットのshort int型が用いられることもある。データ容量の削減を優先して特別にshort int型を用いた36次元のPCA-SIFTの場合でも、各局所記述子は16ビット×36次元=512ビット（64バイト）のメモリを要する。
- [0006] 最近傍検索では、ベクトルの距離計算を行って最近傍の局所記述子を決定することが一般的である。各次元のデータの精度を落とせば正確な最近傍探索ができなくなり、従って、画像認識の精度（認識率）が低下すると考えるのが常識であった。
- [0007] そこで、従来、多くの手法では、モデル作成用の画像から得た局所記述子をベクトル量子化（類似する局所記述子を集めて所定数のグループに分類することにより、同一グループの各局所記述子とそのグループの代表値に置き換えて表現する手法、クラスタリング）して数千から数十万程度のvisual word（前記代表値に相当するもの）を定め、それを用いて画像を記述するというアプローチを採用している（例えば、非特許文献3参照）。未知の画像を認識する際には、その画像から得た局所記述子をvisual wordに変換し、頻度などを計測することになる。このようなアプローチでは、visual wordの数が十分少なければ高速な処理が期待できる。しかし一方で、visual wordの数が多くなれば、十分な認識率を達成できないことが指摘されている（例えば、非特許文献4参照）。visual wordの数が増えればそれだけ、ベクトル量子化に必要な計算時間を無視することができず、また、visual word自体の記録にもメモリ容量の面で問題が生じる。
- [0008] 以上の利点・問題点は、極端な場合、すなわち、モデル作成用の画像から

得た個々の局所記述子をそのままvisual wordとする場合に、最も顕著になる。例えば、VGAサイズの一般的な画像からは、2千程度の局所記述子が抽出される。従って、VGAサイズの10万画像をモデル作成に用いる場合、visual wordの数は2億となり、照合ならびに記憶に膨大な計算資源が必要となる。一方で、大量の局所記述子をモデルに用いることにより、高精度な認識を実現することが可能となる。

- [0009] 処理時間の問題に対する一つの解決策は、局所記述子の照合に「近似最近傍探索」を導入することである（例えば、非特許文献5および特許文献1参照）。これによって、例えば上記の規模の認識タスクの場合、単純な全数照合の場合と比べて、認識率を殆ど低下させることなく、処理速度を 10^{-6} 未満にできることが知られている。一方、メモリ容量の問題については、ベクトル量子化を粗くすることが解決策の一つであるが、これは認識率の低下を意味するため、必ずしも得策ではない。

先行技術文献

特許文献

- [0010] 特許文献1：国際公開第2008/026414号パンフレット

非特許文献

- [0011] 非特許文献1：D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004
- 非特許文献2：J. Ponce, M. Hebert, C. Schmid and A. Zisserman Eds., Toward Category-Level Object Recognition, Springer, 2006
- 非特許文献3：J. Sivic and A. Zisserman, Video google: A text retrieval approach to object matching in videos, Proc. ICCV2003, Vol. 2, pp. 1470-1477, 2003
- 非特許文献4：D. Nister and H. Stewenius, Scalable recognition with a vocabulary tree, Proc. CVPR2006, pp. 775-781, 2006
- 非特許文献5：野口和人、黄瀬浩一、岩村雅一、"近似最近傍探索の多段階化

による物体の高速認識、“画像の認識・理解シンポジウム(MIRU2007) 論文集、pp. 111-118, July, 2007

発明の概要

発明が解決しようとする課題

[0012] 以上のような事情のもとで、この発明は、従来とは異なる視点から解決策を提供するものである。即ち、局所記述子を用いた物体認識において、認識精度を一定レベルに保ったまま、いかに処理時間やメモリ容量を削減するかという視点で解決策を提供するものである。

[0013] 具体的には、発明者らは、認識率への影響が無視できる範囲内で局所記述子の各次元を表す数値の精度を極限まで縮小する可能性について検討した。これは、前記特許文献1のように多数の局所記述子から最近傍の局所記述子を短時間に探索できる手法が提供されており、メモリ容量の削減にある程度の目処がつけばvisual wordを用いない直接的な検索により高精度の物体認識が可能になると考えたからである。

特に、物体のクラスの認識のみならず、物体のインスタンスの認識を目的とする用途では、高精度、高速かつメモリ効率のよい手法が望まれている。visual wordは本質的にグルーピングを伴うため、クラスレベルの認識には有効であってもインスタンスレベルの認識とは相容れない側面がある。この発明の手法は、特にインスタンスレベルの認識に有効であると考えられる。ただし、クラスレベルの認識、および、visual wordを用いた手法への適用を除外するものではない。

ここで、クラスの認識とは、認識対象（検索質問）が与えられると例えば椅子、自動車などの物体のクラスを結果として返す（認識する）ものである。これに対し、インスタンスの認識とは、例えば自動車というクラスの中の特定のモデルなどの物体のインスタンスを認識するものである。

課題を解決するための手段

[0014] 発明者らは鋭意検討を行った結果、前述の技術常識に反し意外にも、実数型または整数型データより少ない構成ビット数で局所記述子の各次元を表現

しても、2ビットまでは認識率が殆ど低下しないことを見出した。特徴ベクトルの各次元は実数型または整数型データ、具体的には32ビットのデータを用いて表現するのがこの技術分野の常識である。ところが、それよりも少ない構成ビット数、例えば、8ビット以下の精度で局所記述子の各次元を表現しても、2ビットまでは認識率が殆ど低下しないことを見出した。また、1ビットであっても、0ビット、即ち、ベクトルの距離計算を行わない手法に比べると2ビットの認識率に近い良好な結果が得られることを見出した。さらに、この結果に多数決処理の寄与が大きいことを見出した。そして、これらの知見に基づき、この出願に至った。

[0015] 即ち、前述の課題を解決する手段として、この発明は、ベクトル量子化により得られるvisual wordを絞り込むというアプローチではなく、個々の局所記述子の記録に必要なメモリ量を削減するというアプローチをとる。具体的には、局所記述子を表現するベクトルについて、記録に必要な各次元のビット数を、実数型または整数型データ(32ビット)より少ない構成ビット数に削減する。これは、局所記述子にスカラー量子化を適用することと見なすことができる。

この発明は、画像データベースに登録されるべき画像であって物体を示す画像から、その画像の各所の局所的な特徴をそれぞれ表現するベクトルを複数の局所記述子として抽出する特徴抽出工程と、前記ベクトルの次元ごとにスカラー量子化を行うスカラー量子化工程と、前記画像とそれに対応する複数のベクトルを前記画像データベースに登録する工程であって、前記スカラー工程で量子化された各ベクトルを、近似最近傍探索を実行できるよう、かつ、各ベクトルが抽出された画像を特定する画像識別子がそれぞれ付加されるように体系化し登録する登録工程とを備え、各工程をコンピュータが実行し、前記登録工程は、複数の画像が前記画像データベースに登録され、問題とする物体を示す一つの画像がクエリーとして与えられたとき、コンピュータが、前記抽出工程と同様の手順で前記クエリーから複数のクエリー局所記述子を抽出し、前記スカラー量子化工程と同様の手順で各クエリー局所記述

子を量子化し、前記画像データベースに登録されたベクトルの中から各クエリー局所記述子の近傍ベクトルとしてのベクトルを前記近似最近傍探索のアルゴリズムを用いて検索し、前記近傍ベクトルに付加された画像識別子を取得し、取得された画像識別子に基づいて問題とする物体を示す少なくとも一つの画像を決定し得るように各ベクトルの登録を行い、前記スカラー量子化工程は、各ベクトルの次元を所定ビット数のスカラー値に量子化する物体認識用画像データベースの作成方法を提供する。

[0016] また、異なる観点から、この発明は、画像データベースに登録されるべき画像であって物体を示す画像から、その画像の各所の局所的な特徴をそれぞれ表現するベクトルを複数の局所記述子として抽出する特徴抽出部と、前記ベクトルの次元ごとにスカラー量子化を行うスカラー量子化部と、前記画像とそれに対応する複数のベクトルを前記画像データベースに登録する工程であって、前記スカラー工程で量子化された各ベクトルを、近似最近傍探索を実行できるよう、かつ、各ベクトルが抽出された画像を特定する画像識別子がそれぞれ付加されるように体系化し登録する登録部と、複数の画像が前記画像データベースに登録され、問題とする物体を示す一つの画像がクエリーとして与えられたとき、前記抽出部が登録されるべき画像と同様に前記クエリーから複数のクエリー局所記述子を抽出し、前記スカラー量子化部が登録されるべき画像と同様に各クエリー局所記述子を量子化した後、前記画像データベースに登録されたベクトルの中から各クエリー局所記述子の近傍ベクトルとしてのベクトルを前記近似最近傍探索のアルゴリズムを用いて検索し、前記近傍ベクトルに付加された画像識別子を取得し、取得された画像識別子に基づいて問題とする物体を示す少なくとも一つの画像を決定する検索部とを備え、前記スカラー量子化工程は、各ベクトルの次元を所定ビット数のスカラー値に量子化する物体認識用画像データベースの処理装置を提供する。

[0017] さらに、異なる観点から、この発明は、画像データベースに登録されるべき画像であって物体を示す画像から、その画像の各所の局所的な特徴をそれ

ぞれ表現するベクトルを複数の局所記述子として抽出する特徴抽出部と、前記ベクトルの次元ごとにスカラー量子化を行うスカラー量子化部と、前記画像とそれに対応する複数のベクトルを前記画像データベースに登録する工程であって、前記スカラー工程で量子化された各ベクトルを、近似最近傍探索を実行できるよう、かつ、各ベクトルが抽出された画像を特定する画像識別子がそれぞれ付加されるように体系化し登録する登録部と、複数の画像が前記画像データベースに登録され、問題とする物体を示す一つの画像がクエリーとして与えられたとき、前記抽出部が登録されるべき画像と同様に前記クエリーから複数のクエリー局所記述子を抽出し、前記スカラー量子化部が登録されるべき画像と同様に各クエリー局所記述子を量子化した後、前記画像データベースに登録されたベクトルの中から各クエリー局所記述子の近傍ベクトルとしてのベクトルを前記近似最近傍探索のアルゴリズムを用いて検索し、前記近傍ベクトルに付加された画像識別子を取得し、取得された画像識別子に基づいて問題とする物体を示す少なくとも一つの画像を決定する検索部、の各部としてコンピュータを機能させ、前記スカラー量子化工程は、各ベクトルの次元を所定ビット数のスカラー値に量子化する物体認識用画像データベースの処理プログラムを提供する。

発明の効果

- [0018] この発明による物体認識用画像データベースの作成方法において、前記スカラー量子化工程は、各局所記述子の各次元を実数型または整数型データより少ない構成ビット数に量子化するので、個々の局所記述子の表現に必要なメモリ容量を削減することができる。従って、多数の局所記述子(あるいはvisual word)を記録する場合であっても、スカラー量子化工程を経ない従来の手法に比べて小さな容量のメモリで画像データベースを作成することができる。
- [0019] 常識的には、特徴ベクトルを少ないビット数で表現すると、認識率の低下など悪影響が出ることが予測される。しかし、後述する実験結果でも示されるように、各次元を2ビットまで削減しても量子化を行わない場合の認識率と

比べて殆ど低下しないことを見出した。各次元が1ビットの場合であっても、0ビット、即ち、ベクトルの距離計算を行わない手法に比べて非常に高い認識率が得られる。

[0020] この発明において、局所記述子は、画像の局所的な特徴をベクトル量で表現するものである。その具体的な態様は、例えば、SIFTである。後述する実施形態においては、局所記述子の一例としてPCA-SIFTを適用している。

[0021] 各局所記述子を最近傍探索可能に体系化するには、画像データベースに登録すべき画像から局所記述子を抽出し、それらの各局所記述子を前記画像と対応付けてデータベースとして登録しておくことをいう。ある局所記述子を与えられたとき、前記データベースに登録された局所記述子の中から最近傍の局所記述子を決定する処理が最近傍探索である。最近傍か否かを評価する手法の一例は、2つのベクトルの距離を計算し、最短距離のものを最近傍とするものである。

[0022] 以下、この発明の好ましい態様について説明する。

好ましくは、前記スカラー量子化工程は、各局所記述子の各次元を8ビット以下に量子化する。より好ましくは、前記スカラー量子化工程は、各局所記述子の各次元を2ビットに量子化する。後述する実験結果にも示されるように、例えば、特徴ベクトルの各次元を記録するビット数を2ビットとした場合、量子化を行わない場合と比較して、所要メモリ容量を削減することができる（後述の実験例では1/3程度）。このとき、処理時間はビット演算のためにやや長くなる（後述の実験例では1.6倍程度）ものの、認識率は殆ど低下しない。また、各次元を8ビットにした場合でも、後述の実験例では所要メモリ容量が2/3程度に削減可能である。

[0023] また、前記登録工程は、量子化後の各局所記述子に所定のハッシュ関数を適用してハッシュ表のインデックス値を算出し、算出されたインデックス値に対応付けて量子化後の各局所記述子の各次元の値、登録すべき画像を識別する識別子および前記画像の参照先を登録するものであり、かつ、同一のインデックス値に対応付けて登録される局所記述子の数が所定の数を超える場

合、そのインデックス値についてはいずれの局所記述子も登録しないようにしてもよい。ハッシュ表を用いることにより高速の最近傍探索処理を実現することができる。さらに、同一のインデックス値に対応付けて登録される局所記述子の数が所定の数を超える場合、それらの局所記述子も登録しないようにすることによって、識別力の弱い局所記述子を最近傍探索処理の対象から除外することができる。従って、探索の精度を犠牲にすることなくより短時間で探索を行うことができる。

[0024] ハッシュ表で分類された一つの類（インデックス）に属する局所記述子の数が多い場合、それらの局所記述子は識別能力が低いといえる。即ち、入力画像の局所記述子からインデックス値を算出してハッシュ表を参照した場合、その類に属する候補が多数登録されているわけである。それらの局所記述子は、認識対象の絞込みにあまり貢献しておらず、識別能力が低いといえる。識別能力の低い局所記述子を探索対象から除外しておけば、識別能力の高い局所識別子だけを参照して、効率的な認識が行われる。

さらにまた、前記スカラー量子化工程は、検索質問から抽出された局所記述子を量子化するとき、量子化の閾値を跨ぐ複数の値についてインデックス値を算出し、いずれかのインデックス値に対応付けられて登録された局所記述子の中から最近傍の局所記述子を決定してもよい。このようにすれば、量子化に伴う誤差の範囲を含めてハッシュのインデックス値を算出するので、量子化による認識率の低下を抑制することができる。

[0025] 例えば、画像データベースに登録された画像が、検索質問に係る物体を異なる視点からみたものである場合、登録された画像と検索質問との間で対応関係にある特徴ベクトルの各次元の値が異なる場合がある。即ち、各特徴ベクトルは、視点の変化（幾何的変換）に対して頑健な特性を持つが、完全な不変量という訳ではなく、変動する場合がある。ハッシュ表は、各特徴ベクトルの各次元の値から所定の計算手順（ハッシュ関数の計算）により離散値であるインデックス値を算出する。ここで、対応する特徴ベクトルの量子化後の値が変動により異なると、異なるインデックス値が算出されてしまう可

能性が高い。検索対象の特徴ベクトルと検索質問の特徴ベクトルのインデックス値が異なると、その対の特徴ベクトルからは正しい探索結果が得られない。そこで、変動により各次元の値が量子化の閾値を超える場合を考慮し、量子化された各次元の値だけでなく、量子化の閾値を跨いだ値についてもインデックス値を計算し、いずれかインデックス値に対応付けて登録されている特徴ベクトルの中から最近傍の特徴ベクトルを決定する。このようにすれば、変動に伴う認識率の低下を抑制することができる。換言すれば、特徴ベクトルのある次元の値が量子化の閾値に近い場合に閾値を跨ぐ可能性も考慮してインデックス値を計算することにより、認識率を確保することができる。

前記特徴抽出工程は、各次元が32ビットで表されるベクトル値を各局所記述子として抽出てもよい。

また、検索質問の各局所記述子に対し最近傍の局所記述子を決定する処理は、局所記述子間の距離計算を行って最も近い距離のものを決定する処理であってもよい。

ここで示した種々の好ましい態様は、それら複数を組み合わせることでもできる。

図面の簡単な説明

- [0026] [図1] この発明に係るスカラー量子化の手法の一例を説明する説明図である。
- [図2] この実施形態の実験例で用いた登録画像の例を示す図である。
- [図3] この実施形態に係る画像データベースにおける特徴ベクトルの各次元の分布 $f(x)$ を示すグラフである。
- [図4] この実施形態の実験例で用いた検索質問の例を示す図である。
- [図5] この実施形態の実験例において、特徴ベクトルの容量と認識率の関係を示すグラフである。
- [図6] この実施形態の実験例において、登録画像数と認識率、処理時間の関係を示すグラフである。
- [図7] この実施形態の実験例において、パラメータ c と認識率、処理時間の関

係を示すグラフである。

[図8]この実施形態の実験例において、パラメータ b 、 e と認識率、処理時間の関係を示すグラフである。

[図9]この実施形態の実験例と従来手法の代表例であるANNとについて、認識率と処理時間の特性を示すグラフである。

[図10]この実施形態に係る物体認識用画像データベースへの画像の登録手順を示すフローチャートである。

[図11]図10の手順で画像が登録されてなる画像データベースを対象に、認識（検索）を行う際の手順を示すフローチャートである。

[図12]この発明による物体認識用画像データベースの処理装置について、その機能的構成を示すブロック図である。

発明を実施するための形態

[0027] <<近似最近傍探索の従来法>>

この発明の詳細について述べる前に、近似最近傍探索の従来法についてまとめておく。ここで取り上げる従来法は、後述の実験例に用いたものである。

[0028] 最近傍探索で最も時間がかかるのは距離計算である。最近傍探索の高速化には、個々の距離計算自体を高速に行う方法と、距離計算の対象を効率よく絞り込む方法の2通りがある。近似最近傍探索では、主に後者の絞り込みを大胆に行うことで、処理時間を削減する。ところがその代償として、最近傍が距離計算の対象から外れてしまい、求まらない可能性がある。どの程度の近似が適切であるのかは、対象とするタスクに依存するため、近似最近傍探索の手法では、近似の程度を調節するためのパラメータを設けている。以下では、近似最近傍探索の代表的な手法としてANNについて述べる。

[0029] ANN(Approximate Nearest Neighborの略。詳細は、S. Arya, D. M. Mount, R. Silverman and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching," Journal of the ACM, vol. 45, no. 6, pp. 891-923, 1998参照)は、2分木を用いて近似最近傍探索を高速に行う手法である。木のノ

ードは、特徴空間を分割したhyperrectangleに対応しており、葉ノードには単一の特徴ベクトルが対応付けられている。ANNでは、木構造の探索によって距離計算の対象となる特徴ベクトルを収集し、その中で距離が最短のものを近似最近傍探索の結果として出力する。ANNには近似の程度を表すパラメータとして許容誤差 ε がある。 ε が大きければ、より大幅な近似を行って対象となる特徴ベクトルを絞り込むため、処理時間が短縮できる。

[0030] <<実施形態>>

以下、図面を用いてこの発明をさらに詳述する。なお、以下の説明は、すべての点で例示であって、この発明を限定するものと解されるべきではない。

[0031] 1.1. 考え方

局所記述子を用いて認識を行う従来手法の問題点は、特徴ベクトルの数が多いため、大規模な認識を行う場合にメモリ使用量と処理時間が膨大になってしまうというものである。

[0032] 処理時間の問題に対して、野口らは近似最近傍探索による識別器を多段階に接続することで処理時間を削減する手法（非特許文献5および特許文献1参照）を提案している。しかし、この手法ではメモリ使用量の問題に触れていない。この手法のメモリ使用量の内訳をみると、約8割が特徴ベクトルを保持するために使われていることがわかる。そこで、この実施形態では野口らの手法を基本とし、特徴ベクトルを量子化して少ないビット数で表現することで、メモリ使用量を削減することを試みる。

[0033] ベクトルを量子化する方法には、代表的なものとしてベクトル量子化とスカラー量子化がある。ベクトル量子化は、あらかじめ、ある個数の代表ベクトルを求めておき、入力ベクトルを最近傍の代表ベクトルの符号に置き換えるものである。ベクトル量子化は、局所記述子を用いた認識手法で多く使われている（非特許文献3参照）。しかし、大規模なデータに対して量子化誤差の少ない代表ベクトル(前述のvisual word)を効率よく求めることは難しい。また、代表ベクトルの数が多くなると、入力ベクトルの最近傍の代表ベク

トルを探索する処理に時間がかかるという問題点がある。一方、スカラー量子化は、入力ベクトルの次元ごとに量子化を行うものである。同じ符号数で比較した場合はベクトル量子化と比べて量子化誤差が大きくなる。しかし、量子化の際に最近傍探索が必要ないため処理時間は少なく済むと考えられる。そこで、この発明では処理時間を重視してスカラー量子化を用いることとする。

[0034] 量子化には、メモリ使用量を削減できるという利点ばかりではなく、認識率が低下する可能性があるという問題点もある。これは、もとは別々の特徴ベクトルであったものが、量子化により同じ特徴ベクトルになり識別性が低下するからである。特徴ベクトルの識別性の低下が認識率に及ぼす影響は、この発明の認識結果が投票によって決定されるため、簡単にはわからない。これは、正解の得票数が逆転しなければ、誤った票が他の画像に入っても誤認識は生じないためである。そこで、量子化のビット数と認識率の関係については実験で検証する。

[0035] 1.2. スカラー量子化

まず、この発明に係るスカラー量子化について説明する。スカラー量子化では、次元ごとに量子化を行う。各次元2bitで量子化する場合を図1に示す。 $f(x)$ は特徴ベクトルのある次元の分布である。 t は量子化の閾値であり、例えば $-\infty$ から t_1 の範囲のものは0に符号化する。検索の際に距離計算を行う場合には v の値を用いて検索質問の特徴ベクトルとの距離を計算する。 t と v の値については、次式を満たすように次元ごとに定める。

[0036]

[数1]

$$\begin{aligned} \int_{-\infty}^{v_1} f(x)dx &= \int_{v_1}^{t_1} f(x)dx = \int_{t_1}^{v_2} f(x)dx \\ &= \dots = \int_{t_3}^{v_4} f(x)dx = \int_{v_4}^{\infty} f(x)dx \end{aligned} \quad (1)$$

[0037] 1.3. ハッシュ関数

野口らの手法と同様、この実施形態でも、ハッシュ関数を用いてデータをハッシュ表に登録し、検索に用いる。ハッシュ表にデータを登録、検索する場合には、ハッシュ関数により処理を行うインデックスを知る必要がある。ここではこの実施形態で用いるハッシュ関数について説明する。

[0038] PCA-SIFTによって得られた36次元の実数値ベクトル x は、主成分分析の結果であるので、前方の次元の方が固有値が大きいという性質がある。そこで、 x の第1次元から第 d 次元までをとり、

[0039] [数2]

$$\hat{\mathbf{x}} = (x_1, x_2, \dots, x_d)$$

とする。次に、

[0040] [数3]

$$u_j = \begin{cases} 1 & \text{if } x_j - \mu_j \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

を用いて次元ごとに二値化を行いビットベクトル

[0041] [数4]

$$\mathbf{u} = (u_1, \dots, u_d)$$

を作成する。ここで μ_j は、画像データベース中すべての x_j の平均値である。そして、

[0042] [数5]

$$H_{\text{index}} = \left(\sum_{i=1}^d u_i 2^i \right) \bmod H_{\text{size}} \quad (2)$$

によってハッシュ値を求める。ここで H_{index} は、ハッシュ表のサイズである。

[0043] 1.4. データ登録

特徴ベクトルをハッシュ表に登録するためには、前述のハッシュ関数によりハッシュ値を求める必要がある。ハッシュ表には、画像IDとともに量子化によってデータ量を削減した特徴ベクトルを登録する。登録時に衝突が生じた場合は、複数の特徴ベクトルをリストとして登録する。このとき、リストが長くなりすぎると、検索の際に距離計算のコストがかかりすぎるといった問題が生じる。そこでこの実施形態では、リスト長 n に対する閾値 c を設け、 $n > c$ を満たすとリスト全体をハッシュ表から削除する。同じハッシュ値を持つ特徴ベクトルが多いということは、その特徴ベクトルが画像の識別にあまり寄与しないことを意味する。従って、削除しても影響は比較的少ないと考えられる。

以上の処理を、データベースに登録するすべての特徴ベクトルに対して施すことにより、データの登録は完了する。

[0044] 図10は、この実施形態に係る物体認識用画像データベースへの画像の登録手順を示すフローチャートである。図10に示すように、処理を行うコンピュータは、登録すべき画像が入力されると、まず、その画像から局所記述子を抽出する（ステップS11）。続いて、各局所記述子を前述の手順でスカラ量子化し、各次元を8ビット以下、好ましくは2ビットに削減する（ステップS13）。そして、量子化された各局所記述子にハッシュ関数を適用し、インデックス値を算出する（ステップS15）。

[0045] 次に、前記コンピュータは、算出されたインデックス値に対応付けてその局所記述子の量子化されたベクトル値に画像IDを付与しハッシュ表に登録する（ステップS17）。即ち、画像データベースに登録する。ここで、同一のインデックス値に所定数を超える局所記述子が登録されたとき、そのインデックス値に対応付けられる局所記述子をハッシュ表から削除し登録しないようにする。これにより、識別力の低い局所記述子をデータベースに登録しないようにする。次に、抽出されたすべての局所記述子がハッシュ表に登録されたか否かを判断し（ステップS19）、残りの局所記述子が存在する場合（ステップS19のNo）は、残りの局所記述子についてもステップS13~17の処理を繰り返す。すべての局所記述子の登録が終了したら（ステップS19のYes）、その画像のデータベースへの登録は終了する。

[0046] 1.5. 検索

次に検索について述べる。この実施形態では、検索質問Qから得た各特徴ベクトルqに対して、ハッシュ関数を適用しハッシュ表から特徴ベクトルを検索する。得られた特徴ベクトルの集合をXとすると、次にXの中からqの最近傍となる特徴ベクトル x_* を求める。

[0047] そして、 x_* に対応する画像IDに投票する。もし、最近傍となる特徴ベクトルが複数ある場合には、それらすべてに対して投票処理を施す。検索質問のすべての特徴ベクトルに対してこの処理を行い、最終的に最も得票数の多いものを回答とする。ここで、「投票」とは、情報処理の分野で部分的に証拠を集計するために用いる処理であって、得られた証拠に基づいて選択肢の何れかに得点を与え、すべての証拠を集計した結果最高得点を得た選択肢を選択する処理をいう。一般的に各証拠は異なる点数の票を持つ。

[0048] この処理において、最も重要なステップは、いかにqに対する特徴ベクトルを検索するかにある。最も単純な手法は、登録時と同様にqに対してもビットベクトルを求め、ハッシュ関数によって同じハッシュ値を持つ特徴ベクトルを求めることである。ところが、このような処理では、距離の計算回数は十分削減できるものの、次の理由によって十分な認識率を得ることができない

。それは、特徴ベクトルの各次元の値が撮影条件によって変動するためである。もし、閾値を超えるような変動があると、ビットベクトルが異なるものとなり、もはや対応する特徴ベクトルを得ることができなくなる。

この実施形態では、値の変動幅 e をパラメータとして、変動への対処を施す。具体的には、

[0049] [数6]

$$\mathbf{q} = (q_1, \dots, q_d)$$

とするとき、

[0050] [数7]

$$|q_j - \mu_j| \leq e$$

を満たす次元 j に対しては、 u_j だけではなく

[0051] [数8]

$$u_j = (u_j + 1) \bmod 2 \quad (0 \text{ ならば } 1, 1 \text{ ならば } 0)$$

[0052] も用いて、特徴ベクトルを検索する。ただし、このような「両方試す」という処理を制限なく導入すると、膨大な計算時間が必要になってしまう。この処理では、処理の対象となる次元数を b とすると、 $2b$ 通りのビットベクトルを用いてハッシュ表にアクセスすることになる。そこでこの実施形態では、 b をあまり大きくない値に留めることとする。

[0053] [数9]

$$|q_j - \mu_j| \leq e$$

を満たす次元の数が b を上回るときには、次元のインデックスが大きいものから b 個を採用する。

[0054] 図11は、図10の手順で画像が登録されてなる画像データベースを対象に、認識（検索）を行う際の手順を示すフローチャートである。図11に示

すように、検索質問の画像が入力されると、処理を行うコンピュータは、まず、その画像から局所記述子を抽出する（ステップS21）。続いて、各局所記述子を前述の手順でスカラー量子化し、各次元を8ビット以下、好ましくは2ビットに削減する（ステップS23）。そして、量子化された各局所記述子にハッシュ関数を適用し、インデックス値を算出する（ステップS25）。ステップS21、S23およびS25の処理は、図10のステップS11、S13およびS15の処理とそれぞれ対応しており、同様の処理である。ただし、ステップS25では、画像の変動を考慮して、量子化の閾値を跨ぐ複数の値についてもインデックス値を計算する。即ち、前述の「両方試す」処理をさらに行う。

[0055] 次に、前記コンピュータは、算出されたインデックス値に対応付けてハッシュ表に登録されている局所記述子うちで最近傍の局所記述子を決定し、最近傍の局所記述子が含まれる画像IDに一票を投じる。次に、検索質問から抽出されたすべての局所記述子について最近傍探索とそれに基づく投票がなされたか否かを判断し（ステップS29）、残りの局所記述子が存在する場合（ステップS29のNo）は、残りの局所記述子についてもステップS23～27の処理を繰り返す。すべての局所記述子について投票が行われた（ステップS29のYes）、最多得票数を得た画像を認識結果として出力する（ステップS31）。そして、認識処理は終了する。

[0056] <<ブロック図>>

図12は、この発明による物体認識用画像データベースの処理装置について、その機能的構成を示すブロック図である。図12で、画像データベース25に画像21を登録する際は、登録すべき画像21が処理装置1の入力として与えられ、その出力として画像データベース25に画像21に係るデータが登録される。一方、画像データベース25に登録された画像を検索する際は、検索質問23が処理装置1の入力として与えられ、処理装置1は画像データベース25を参照し、出力として画像データベース25に登録された画像27あるいはそれを特定する情報を出力する。

[0057] 特徴抽出部11は、入力として与えられた画像21または23から特徴ベ

クトルを抽出する。スカラー量子化部 13 は、抽出された各特徴ベクトルをスカラー量子化する。登録部 15 は、スカラー量子化された各特徴ベクトルのデータにそれが抽出された画像を識別する識別子を付与し、画像データベース 25 に登録する。検索部 17 は、スカラー量子化された各特徴ベクトルについて、画像データベース 25 に登録された局所記述子のうちで最近傍ものを最近傍探索処理により決定する。そして、決定された局所記述子を含む画像に投票し、検索質問の各局所記述子に係る投票の結果に基づいて画像データベース 25 に登録された画像の中から一の画像を特定する。特定された画像 27 あるいはその識別情報を認識結果として出力する。

[0058] 図 12 の特徴抽出部 11、スカラー量子化部 13、登録部 15、検索部 17 の機能を実現するハードウェアは、主としてコンピュータとメモリである。即ち、コンピュータが、所定の物体認識用画像データベースの処理プログラムを実行することにより、各部の機能が実現される。

[0059] <<実験例>>

2.1. 実験条件

実験では、近似最近傍探索の従来法である ANN をこの実施形態と比較した。ここで、ANN としてはインターネット < URL: <http://www.cs.umd.edu/mount/ANN/> > で提供されているソースコードを用いた。

[0060] 実験には以下に述べる画像データベース、検索質問画像を用いた。局所記述子としては、PCA-SIFT を用いた。PCA-SIFT は、インターネット < URL: <http://www.cs.cmu.edu/yke/pcasift/> > で提供されている。PCA-SIFT では 36 次元の特徴ベクトルが得られる。ハッシュ表のサイズは

[0061] [数 10]

$$H_{\text{size}} = 2^d$$

とした。以下に示す処理時間は、検索質問の画像 1 枚あたりの認識に要した時間を表す。

ただし、特徴ベクトルの抽出に必要な時間は含めない。使用計算機は、CPU が

AMD Opteron 2.8GHz、メモリ32GBのものである。

[0062] 2.1.1. 画像データベース

実験に用いた画像について説明する。まず、収集方法の異なるA、B、Cの3種類のデータセットを準備した。Aは、Googleのイメージ検索を用いて収集した3,100枚の画像である。検索キーワードとしては、“ポスター”、“雑誌”、“表紙”などを用いた。図2(a)に例を示す。

[0063] BはPCA-SIFTのサイトで公開されている画像であり、画像数は18,500枚である。このデータは主に自然写真や人物の写真などで構成されている。図2(b)に例を示す。Cは、写真共有サイトのflickrにおいて“animal”、“birthday”、“food”、“japan”などのタグにより収集した78,400枚の画像からなる。主に図2(c)に示すような物体や自然の写真、人物の写真などを含む。なお、収集の際には、600×600 pixel以下のサイズの画像は除外し、画像の長辺が640 pixel以下になるように縮小した。また、特徴ベクトルが100個以下の画像も除外した。画像の一辺の長さの平均はA、B、Cそれぞれ498, 612, 554 pixelであった。

[0064] 次に、A、B、Cの画像を用いて、表1に示した画像数からなる4種類のデータベースを作成し実験に用いた。ここで、大きいデータベースは、小さいデータベースをその一部として含む。1万枚の画像データベースの特徴ベクトルの各次元の分布 $f(x)$ を図3に示す。図3(a)～(d)で、横軸は各次元の値であり、縦軸は頻度である。1次元目は双峰性の分布であり、2次元目以降は単峰性の分布を示す。また、次元が大きくなるにつれて分散が小さくなる。平均値はいずれも0の付近である。また、特徴ベクトルの個数は画像1枚あたり平均2千個であった。実験では、どの枚数のデータベースを用いる場合でも、1万枚の画像データベースから得られた $f(x)$ を量子化に用いた。

[0065]

[表1]

登録画像数		5,000	10,000	50,000	100,000
内訳	A	1,667	3,100	3,100	3,100
	B	1,667	3,450	18,500	18,500
	C	1,666	3,450	28,400	78,400

[0066] 2. 1. 2. 検索質問画像

検索質問としては、データセットA、B、Cのそれぞれから100、200、200枚の合計500枚を無作為に選択した。次に、これらをA4の用紙に印刷し、カメラを用いて撮影した。得られた画像の例を図4に示す。図4(a)～(c)に示すとおり、紙面全体が写る配置で、紙面に対するカメラの光軸の角度 θ を 90° 、 75° 、 60° に変化させた。また、角度を 90° として紙面の一部を撮影した(図4(d))。その結果、1枚の紙面に対して、合計4通りの画像を得た。さらに、撮影した画像を 512×341 pixelに縮小し、PCA-SIFTにより特徴ベクトルを求めた。その結果、画像1枚あたり平均612個の特徴ベクトルが得られた。

[0067] 2. 2. 特徴ベクトルのメモリ容量と認識率

まず、量子化によりデータ量を削減すると、どの程度認識率に影響がどうか調べた。結果を図5に示す。横軸は特徴ベクトルの各次元の量子化ビット数を表している。例えば、2bitの場合には、PCA-SIFTの特徴ベクトルが36次元なので、一ベクトルあたり $2\text{bit} \times 36 = 72\text{bit}$ 必要となる。ここで、16bitと0bitの場合には少し特殊であることを述べておく。16bitの場合には量子化を行わず生のデータをそのまま用いた(野口らの手法)。即ち、この実験では、メモリ容量削減を優先し、各次元を16bit表現した場合を比較の基準にしている。0bitの場合には、距離計算を行うための特徴ベクトルが存在しないため、ハッシュ表から得られた集合Xのすべてに対して投票を行った。登録画像数は1000、10000、100000枚、パラメータは16bitの場合により結果の得られた $b=10$ 、 $c=10$ 、 $d=28$ 、 $e=400$ について調べた。また、メモリ使用量は100000枚のデータベースについてのみ調べた。

[0068] グラフから、2bitまで削減しても、認識率に大きな影響はみられないことがわかる。1bitの場合には登録画像数が多いほど認識率の低下が著しいことがわかる。これは、登録されている特徴ベクトルの数が増えるほど、識別性の不足が問題となるためと思われる。0bitの場合には認識率が大きく低下することがわかる。このことは、各次元2bit、つまり特徴ベクトル1つあたり9バイトで表現できることを意味する。各次元を16bitで表現した場合と比べると、特徴ベクトル単体で1/8の容量であり、実際のメモリ使用量も約1/3と大幅な圧縮が可能であることがわかる。以降の実験では各次元2bitで量子化を行った場合の詳しい性質についてみていく。

[0069] 2.3. スケーラビリティ

次に、この実施形態の2bitと16bitの場合について登録画像数と認識率と処理時間の関係を比較した。

登録画像数を5000枚から10万枚まで変化させ、実験を行った。パラメータは先ほどと同様に $b=10$ 、 $c=10$ 、 $d=28$ 、 $e=400$ を用いた。登録画像数を変化させた結果を図6に示す。

[0070] グラフより、認識率は殆ど変わらないが、処理時間が増加していることがわかる。例えば10万画像を登録した場合の処理時間は、135.8msから210.9msと1.6倍となった。

処理時間の増加分は距離計算によるものと考えられる。それは、ハッシュ関数は特徴ベクトルの量子化と無関係なため、16bitの場合でも2bitの場合でも、距離計算の対象となるベクトルは変化しないからである。距離計算の処理時間が増加した理由は二つ考えられる。一つは、量子化された符号を中央値 v に変換する処理である。もう一つは、各次元の2bitのデータ取りだす際に起こるbit演算が原因であると考えられる。

[0071] 2.4. この実施形態のパラメータと認識率・処理時間

ここまでの実験で、2bitで量子化すれば認識率を落とさずにデータ量を削減できることがわかった。本節では、2bitで量子化した場合についてこの実施形態のパラメータと認識率処理時間の関係に大きな変化がみられないかを

確認する。この実施形態の主なパラメータである b 、 c 、 d 、 e について実験を行った。まず、衝突の閾値 c とハッシュ構築に使う次元数 d と認識率、処理時間の関係について述べる。このとき画像データベースは1万枚、ハッシュ表のサイズとしては $H_{size=2}$ とした。 $e=200$ 、 $b=10$ 、 $d=24$ 、 26 、 28 とし、 c を変化させた結果を図7に示す。 c が減少するにつれ、処理時間が減少していることがわかる。ただし、 c を小さくしすぎると、認識率が低下した。これは、認識に寄与していたものも削除してしまったためと考えられる。一方、 c を増加させた場合に、計算時間は増加するものの、認識率が減少することは殆どなかった。これは、最近傍にはなり得ない特徴ベクトルを検索したとしても、距離計算によって排除可能なためと考えられる。

[0072] 次に、両方試す次元数の閾値 b と認識率、処理時間の関係について述べる。 $d=28$ とした上で、 $e=200$ 、 400 、 600 、 $c=10$ とし、 b を変化させた結果を図8に示す。 b を増加させると、処理時間は増加するものの、認識率は改善する。両方試す処理の対象とする幅 e が小さいものほど、 b が小さい場合に認識率がよく上昇することがわかる。これは次のように説明できる。検索処理では、検索質問の特徴ベクトルの後ろの次元から $-e$ から $+e$ の範囲内のもの b 個を処理の対象とする。このとき、 e が大きければ、処理が必要な部分に到達する前に b 個の閾値にかかるため、認識率が低下する。また、ある程度 b を増やすと、 b の増加に対して認識率、処理時間も伸びが鈍くなる。これは各次元の値が $-e$ から $+e$ の範囲にあるベクトルがそれ以上存在せず、処理の対象とするインデックスが増加しないからである。

[0073] これらのことから、 e をあまり大きな値にしなければよい結果が得られることがわかる。

以上の実験の結果は、野口らの手法で得られたものと大きな差はみられなかった。そこから、量子化がパラメータに与える影響は軽微であるといえる。

[0074] 2.5. ANNとの比較

近似最近傍探索の従来法であるANNに対して、2bitで量子化を行った場合と

比較するため、1万枚の画像を登録しパラメータをさまざまに変え、横軸に認識率、縦軸に処理時間を描いたグラフを図9に示す。

[0075] ANNで許容誤差 ϵ を2から1000まで変化させたものを線で描き、評価の基準とした。右にプロットされているものほど認識率が高く、下にプロットされているものほど処理時間が短い。そのため、右下にプロットされているものほど優れているといえる。このグラフからわかるとおり、この実施形態はANNを用いた場合よりも優れているといえる。

[0076] [表2]

手法	パラメータ	60度		75度		90度		一部		平均	
		精度	時間	精度	時間	精度	時間	精度	時間	精度	時間
ANN 16bit	$\epsilon = 5$	98.2	232.3	99.0	243.6	99.2	244.1	99.0	742.8	98.9	365.7
	$\epsilon = 10$	97.4	48.4	98.8	53.1	98.6	53.1	98.6	153.2	98.4	76.9
	$\epsilon = 30$	92.8	10.0	97.4	11.7	97.4	11.9	97.2	32.1	96.2	16.4
ANN 2bit	$\epsilon = 3$	98.6	865.2	99.0	909.1	98.8	904.8	99.0	2871.9	98.9	1387.8
	$c = 5$	97.6	121.1	98.6	132.1	98.6	132.1	98.6	394.6	98.4	195.0
	$c = 10$	91.6	12.3	97.2	14.4	97.2	14.5	96.8	39.7	95.7	20.2
提案手法 16bit	$b = 14, c = 10, d = 28, e = 600$	98.4	240.9	98.8	296.2	99	299.8	99	815.8	98.8	413.2
	$b = 10, c = 10, d = 28, e = 400$	98	23.3	98.8	28.3	98.6	28.8	98.6	76.6	98.5	39.2
	$b = 4, c = 10, d = 28, e = 200$	95.2	0.8	97.6	1.1	97.6	1.2	98	4.1	97.1	1.8
提案手法 2bit	$b = 14, c = 10, d = 28, e = 600$	98.4	295.6	98.8	362.0	8.6	368.5	99.4	997.8	98.8	506.0
	$b = 10, c = 10, d = 28, e = 400$	98.0	28.8	98.8	35.0	98.6	35.7	98.8	94.7	98.6	48.6
	$b = 4, c = 10, d = 28, e = 200$	93.4	1.1	98.4	1.5	97.6	1.5	96.6	4.7	96.5	2.2

[0077] 次に、各種パラメータの代表的な値を用いた認識率と処理時間を表2に示す。まず、ANNについて認識率98.9%を得るために必要な処理時間を16bitと2bitで量子化した場合について比較すると、2bitで量子化した場合には、同じ認識率を得るために必要な処理時間が大幅に増加するものの、この実施形態についてみると、ANNと比べ処理時間の増加は穏やかである。以上から、量子化を行うと、同じ認識率を得るために必要な処理時間は増加することがわかる。しかし、量子化によってメモリ使用量を大幅に削減できるという利点があることがわかる。

[0078] 以上の実験例から、各次元2bitで量子化した場合に、量子化なしの場合と比べ処理時間を犠牲にすることで、メモリ使用量を1/3にすることができた。また、メモリ使用量、処理時間、認識率はそれぞれトレードオフの関係にあることがわかった。

[0079] 前述した実施の形態の他にも、この発明について種々の変形例があり得る。それらの変形例は、この発明の範囲に属しないと解されるべきものではない。この発明には、請求の範囲と均等の意味および前記範囲内のすべての変形とが含まれるべきである。

符号の説明

[0080] 1 : 画像データベース処理装置

1 1 : 特徴抽出部

1 3 : スカラー量子化部

1 5 : 登録部

1 7 : 検索部

2 1 : 登録画像

2 3 : 検索質問

2 5 : 画像データベース

2 7 : 認識結果

請求の範囲

- [請求項1] 画像データベースに登録されるべき画像であって物体を示す画像から、その画像の各所の局所的な特徴をそれぞれ表現するベクトルを複数の局所記述子として抽出する特徴抽出工程と、
前記ベクトルの次元ごとにスカラー量子化を行うスカラー量子化工程と、
前記画像とそれに対応する複数のベクトルを前記画像データベースに登録する工程であって、前記スカラー工程で量子化された各ベクトルを、近似最近傍探索を実行できるよう、かつ、各ベクトルが抽出された画像を特定する画像識別子がそれぞれ付加されるように体系化し登録する登録工程とを備え、
各工程をコンピュータが実行し、
前記登録工程は、複数の画像が前記画像データベースに登録され、問題とする物体を示す一つの画像がクエリーとして与えられたとき、コンピュータが、前記抽出工程と同様の手順で前記クエリーから複数のクエリー局所記述子を抽出し、前記スカラー量子化工程と同様の手順で各クエリー局所記述子を量子化し、前記画像データベースに登録されたベクトルの中から各クエリー局所記述子の近傍ベクトルとしてのベクトルを前記近似最近傍探索のアルゴリズムを用いて検索し、前記近傍ベクトルに付加された画像識別子を取得し、取得された画像識別子に基づいて問題とする物体を示す少なくとも一つの画像を決定し得るように各ベクトルの登録を行い、
前記スカラー量子化工程は、各ベクトルの次元を所定ビット数のスカラー値に量子化する物体認識用画像データベースの作成方法。
- [請求項2] 前記スカラー量子化工程は、各ベクトルの次元を2ビット以下のスカラー値に量子化する請求項1に記載の方法。
- [請求項3] 前記登録工程は、(1) 量子化された各ベクトルに所定のハッシュ関数を適用してハッシュ表のビン参照用のインデックス値を算出し、

(2) 量子化されたベクトルの各次元と付加された画像識別子とを、算出されたインデックス値で参照されるビンに一つのエントリーとして登録し、(3) 同一のビンに登録されたエントリーの数が閾値を超える場合、そのビンに登録されたすべてのエントリーを削除しかつそのビンにその後エントリーが登録されないように処理して各ベクトルを登録する請求項1または2に記載の方法。

[請求項4] 前記登録工程は、前記コンピュータが、前記近傍ベクトルを検索して前記画像を決定し得るように各ベクトルを登録し、前記コンピュータは、量子化されたベクトルの次元から前記インデックス値を算出し、さらに一以上の量子化されたベクトルの次元の値の隣の値から一以上のインデックス値を算出し、それらの算出されたインデックス値で参照される複数のビンに登録されたベクトルのうちから前記近傍ベクトルを検索する請求項3に記載の方法。

[請求項5] 前記近似最近傍探索のアルゴリズムは、各クエリー局所記述子と算出されたインデックス値で参照されるビンに登録されたベクトルとの距離を計算し、所定の距離内にある一以上のベクトルまたは最短距離にあるベクトルを特定する処理を含む請求項1～4のいずれか一つに記載の方法。

[請求項6] 画像データベースに登録されるべき画像であって物体を示す画像から、その画像の各所の局所的な特徴をそれぞれ表現するベクトルを複数の局所記述子として抽出する特徴抽出部と、前記ベクトルの次元ごとにスカラー量子化を行うスカラー量子化部と、前記画像とそれに対応する複数のベクトルを前記画像データベースに登録する工程であって、前記スカラー工程で量子化された各ベクトルを、近似最近傍探索を実行できるよう、かつ、各ベクトルが抽出された画像を特定する画像識別子がそれぞれ付加されるように体系化し登録する登録部と、

複数の画像が前記画像データベースに登録され、問題とする物体を示す一つの画像がクエリーとして与えられたとき、前記抽出部が登録されるべき画像と同様に前記クエリーから複数のクエリー局所記述子を抽出し、前記スカラー量子化部が登録されるべき画像と同様に各クエリー局所記述子を量子化した後、前記画像データベースに登録されたベクトルの中から各クエリー局所記述子の近傍ベクトルとしてのベクトルを前記近似最近傍探索のアルゴリズムを用いて検索し、前記近傍ベクトルに付加された画像識別子を取得し、取得された画像識別子に基づいて問題とする物体を示す少なくとも一つの画像を決定する検索部とを備え、

前記スカラー量子化工程は、各ベクトルの次元を所定ビット数のスカラー値に量子化する物体認識用画像データベースの処理装置。

[請求項7]

画像データベースに登録されるべき画像であって物体を示す画像から、その画像の各所の局所的な特徴をそれぞれ表現するベクトルを複数の局所記述子として抽出する特徴抽出部と、

前記ベクトルの次元ごとにスカラー量子化を行うスカラー量子化部と、

前記画像とそれに対応する複数のベクトルを前記画像データベースに登録する工程であって、前記スカラー工程で量子化された各ベクトルを、近似最近傍探索を実行できるように、かつ、各ベクトルが抽出された画像を特定する画像識別子がそれぞれ付加されるように体系化し登録する登録部と、

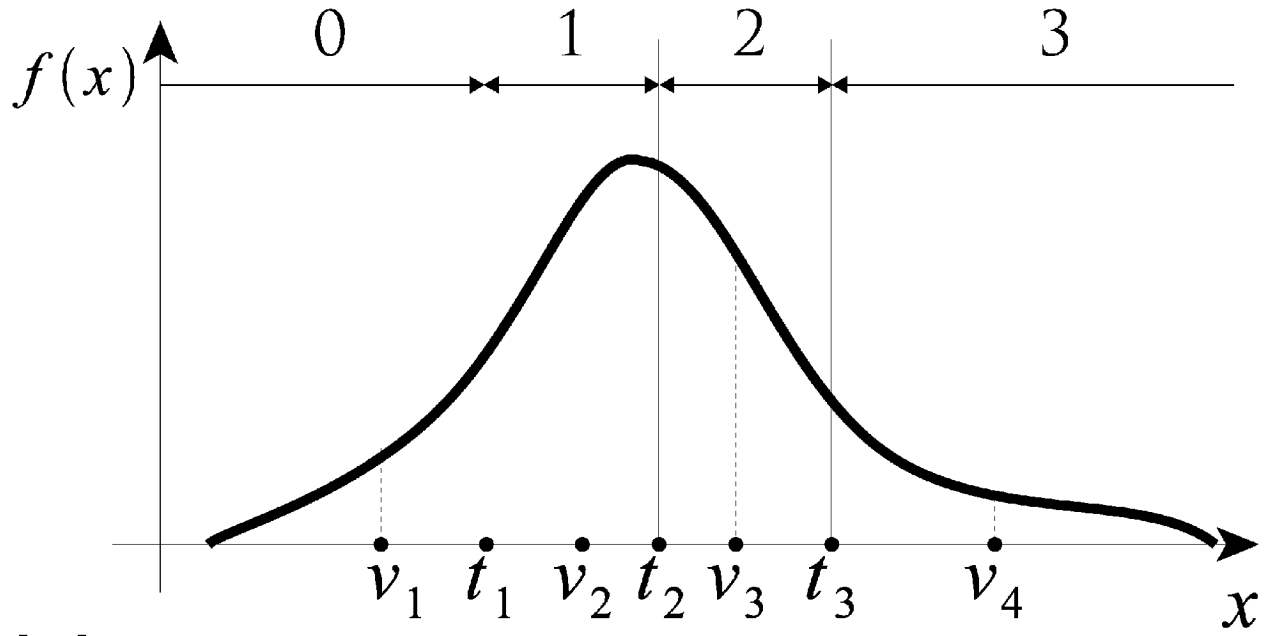
複数の画像が前記画像データベースに登録され、問題とする物体を示す一つの画像がクエリーとして与えられたとき、前記抽出部が登録されるべき画像と同様に前記クエリーから複数のクエリー局所記述子を抽出し、前記スカラー量子化部が登録されるべき画像と同様に各クエリー局所記述子を量子化した後、前記画像データベースに登録されたベクトルの中から各クエリー局所記述子の近傍ベクトルとしてのベク

トルを前記近似最近傍探索のアルゴリズムを用いて検索し、前記近傍ベクトルに付加された画像識別子を取得し、取得された画像識別子に基づいて問題とする物体を示す少なくとも一つの画像を決定する検索部、

の各部としてコンピュータを機能させ、

前記スカラー量子化工程は、各ベクトルの次元を所定ビット数のスカラー値に量子化する物体認識用画像データベースの処理プログラム。

[図1]



[図2]



(a)

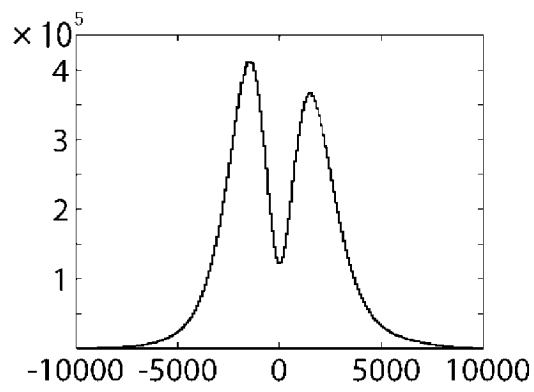


(b)

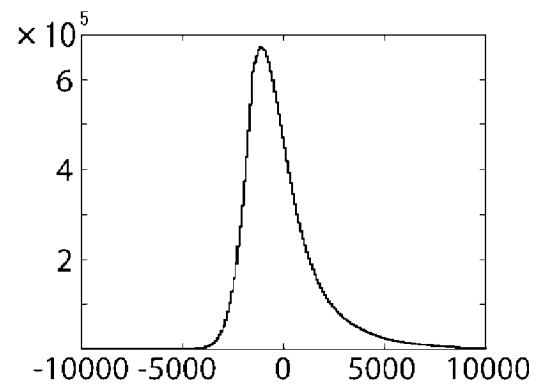


(c)

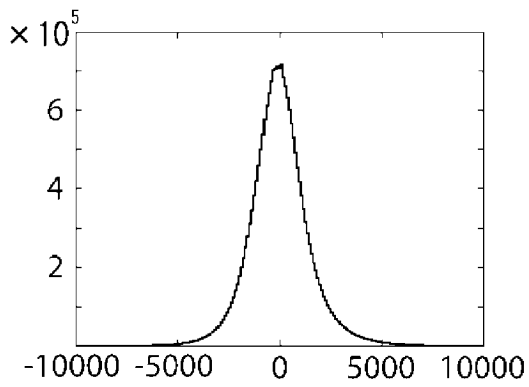
[圖3]



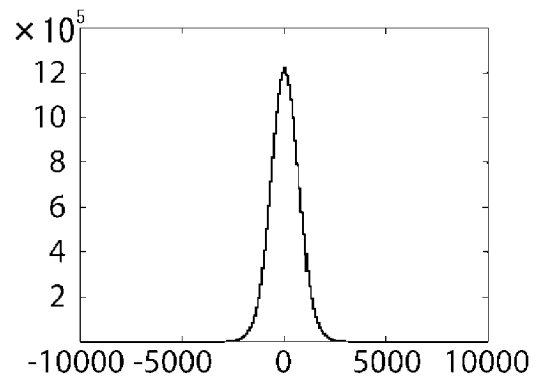
(a) 第1次元



(b) 第2次元

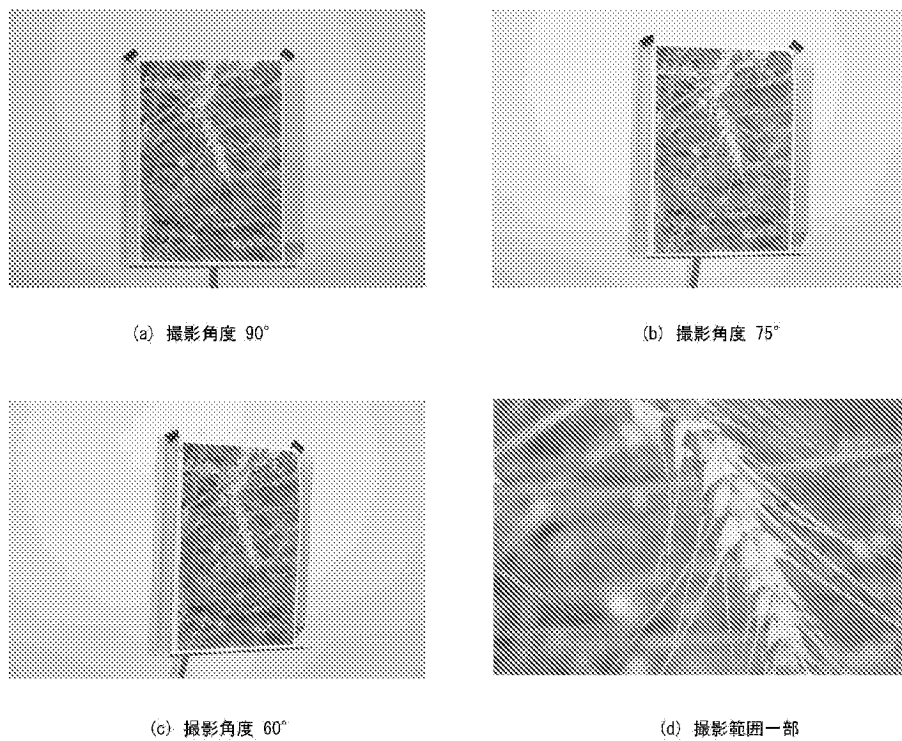


(c) 第3次元

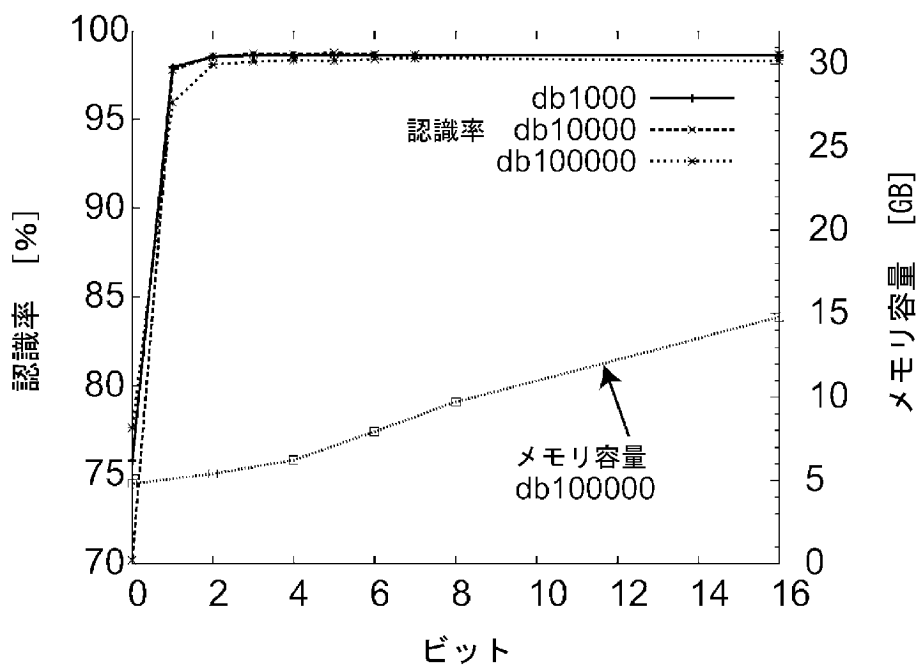


(d) 第36次元

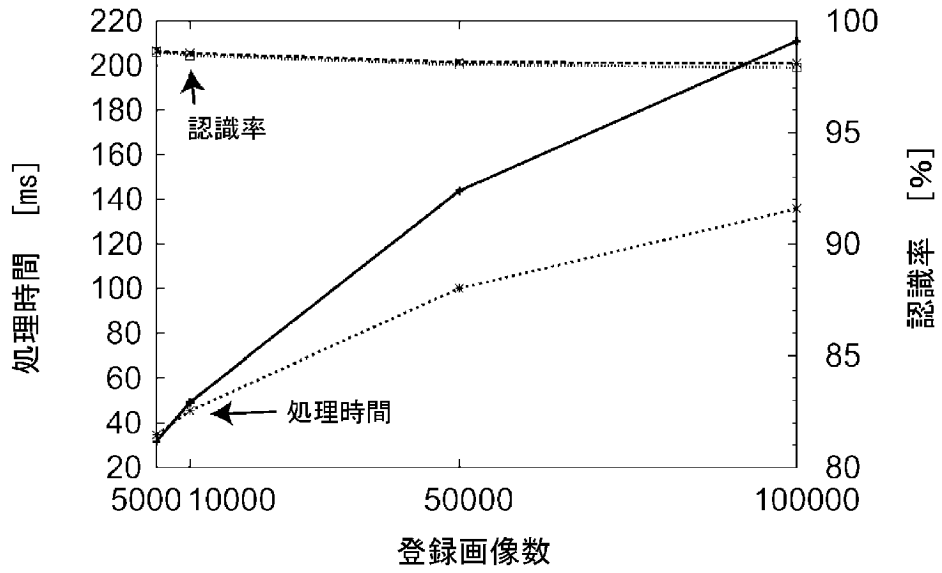
[図4]



[図5]

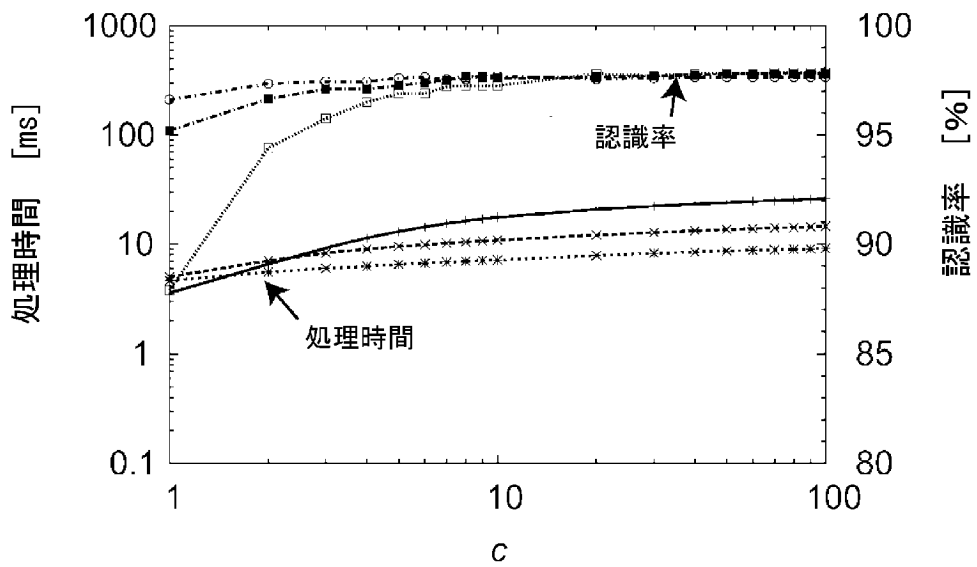


[図6]



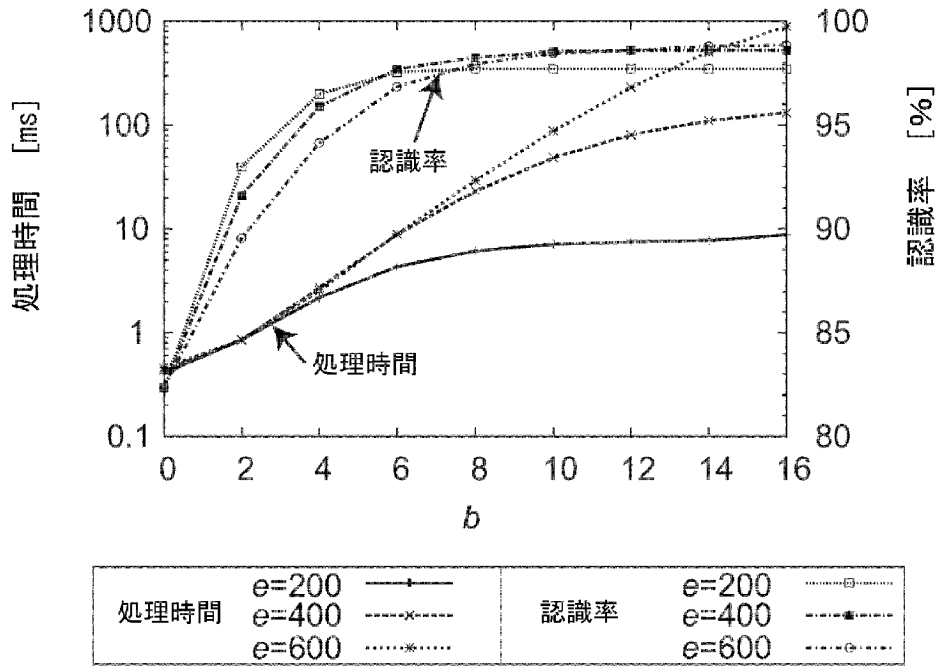
	2bit	16bit
処理時間	—●—	---*---
認識率	---x---	---○---

[図7]

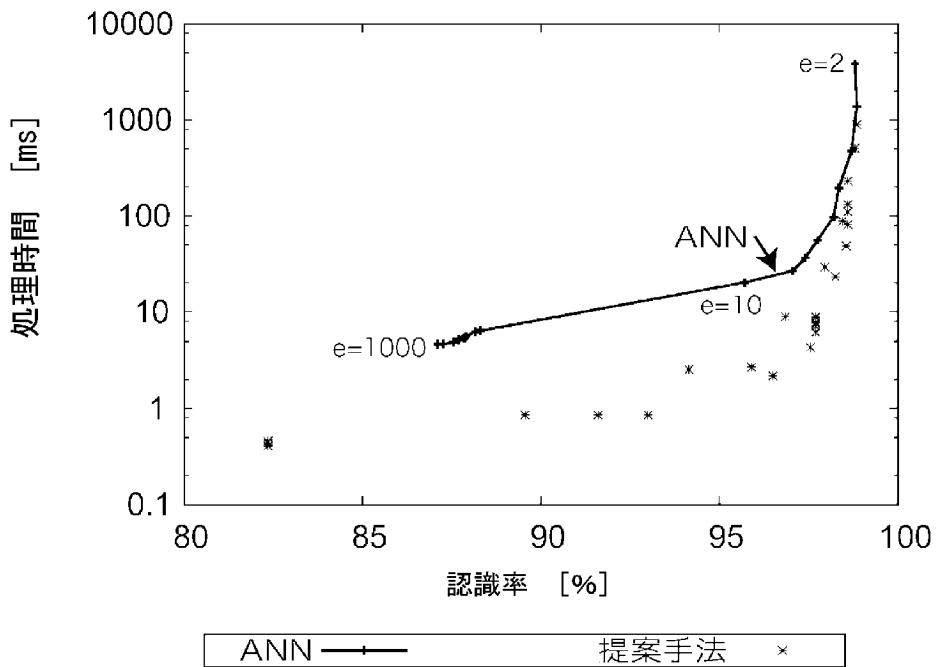


処理時間	d=24 —●—	認識率	d=24 ---○---
	d=26 ---x---		d=26 ---■---
	d=28 ---*---		d=28 ---△---

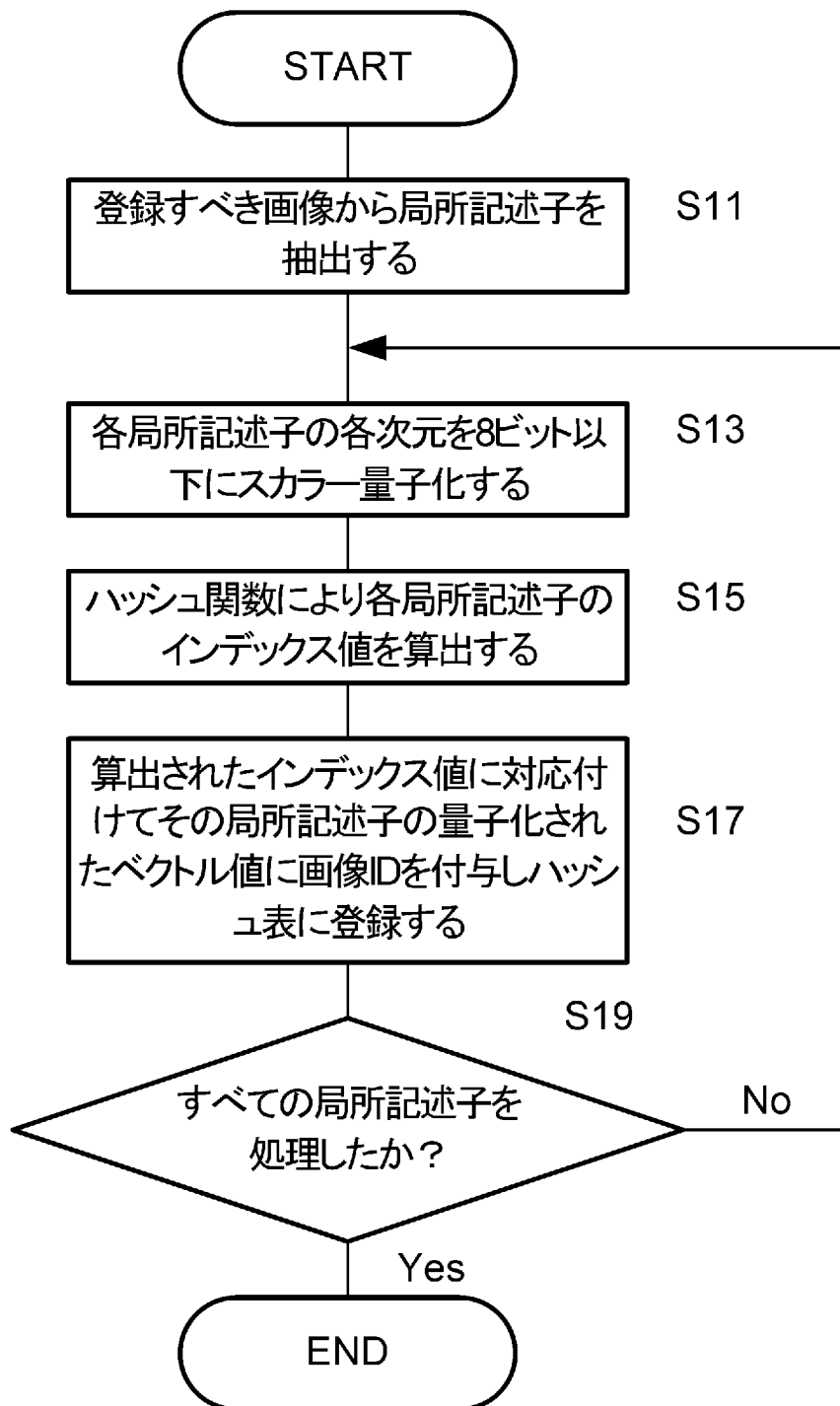
[図8]



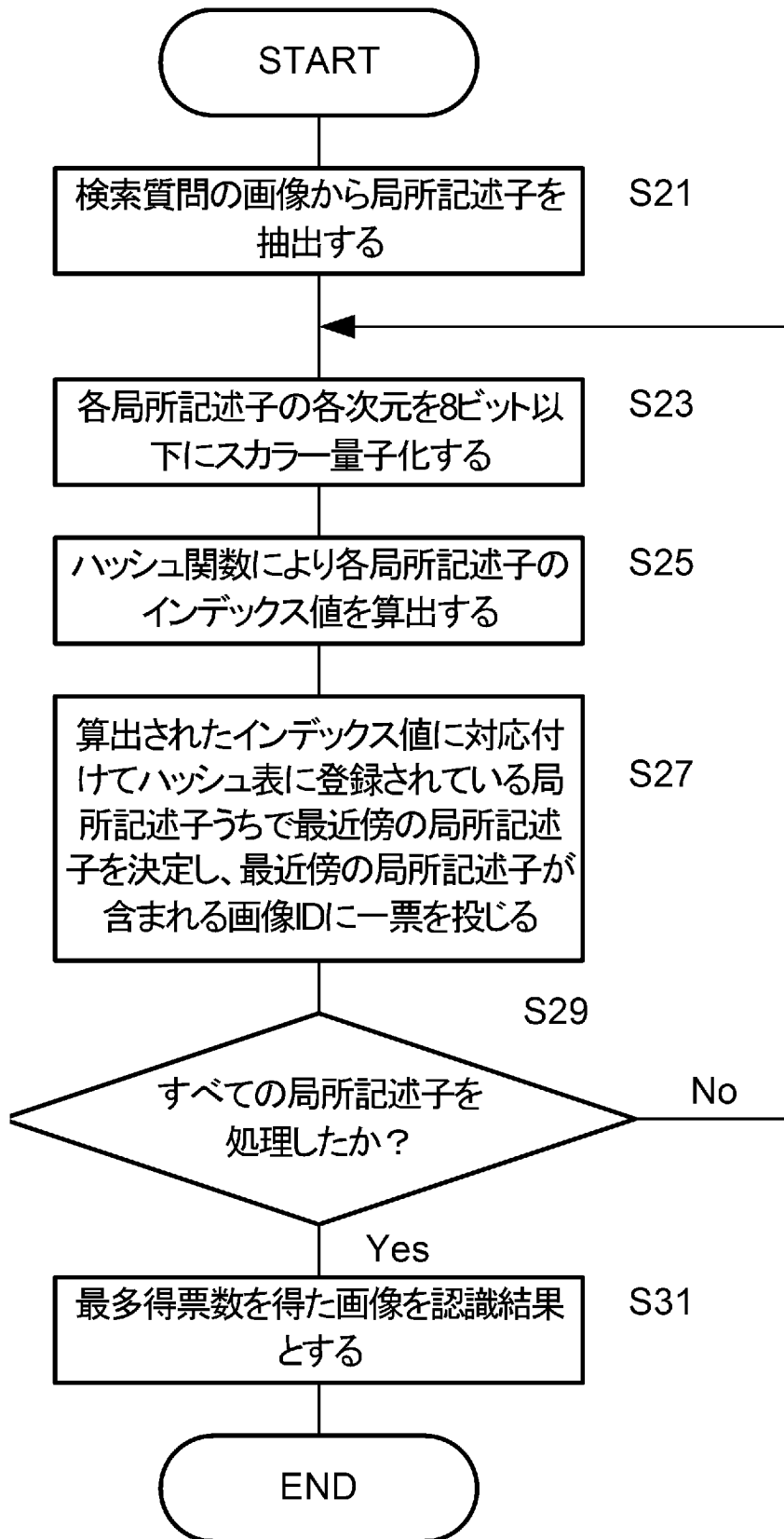
[図9]



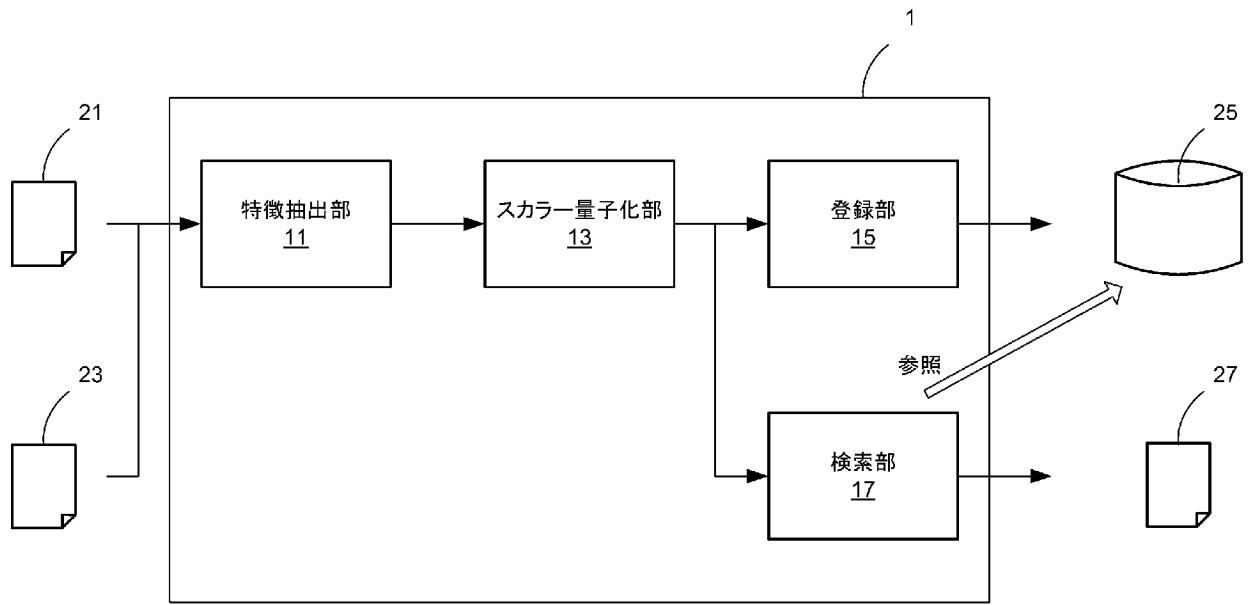
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2009/058285

A. CLASSIFICATION OF SUBJECT MATTER
G06T7/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06T7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2009
Kokai Jitsuyo Shinan Koho	1971-2009	Toroku Jitsuyo Shinan Koho	1994-2009

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/026414 A1 (Osaka Prefecture University), 06 March, 2008 (06.03.08), Full text; all drawings (Family: none)	1-7

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 20 May, 2009 (20.05.09)	Date of mailing of the international search report 02 June, 2009 (02.06.09)
--	--

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G06T7/00(2006.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. G06T7/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2009年
日本国実用新案登録公報	1996-2009年
日本国登録実用新案公報	1994-2009年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	WO 2008/026414 A1 (公立大学法人大阪府立大学) 2008.03.06, 全文, 全図 (ファミリーなし)	1-7

☐ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」口頭による開示、使用、展示等に言及する文献
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」同一パテントファミリー文献

国際調査を完了した日

20.05.2009

国際調査報告の発送日

02.06.2009

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

新井 則和

5H 8937

電話番号 03-3581-1101 内線 3531