

(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2006-285570
(P2006-285570A)

(43) 公開日 平成18年10月19日(2006. 10. 19)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G O 6 F 17/30 3 5 O C	5 B O 5 O
G06T 1/00 (2006.01)	G O 6 F 17/30 1 7 O B	5 B O 7 5
G06T 7/00 (2006.01)	G O 6 F 17/30 3 2 O Z	5 L O 9 6
	G O 6 T 1/00 2 O O E	
	G O 6 T 7/00 3 O O F	
審査請求 未請求 請求項の数 10 O L (全 36 頁)		

(21) 出願番号	特願2005-103656 (P2005-103656)	(71) 出願人	899000068
(22) 出願日	平成17年3月31日 (2005. 3. 31)		学校法人早稲田大学
			東京都新宿区戸塚町 1 丁目 1 O 4 番地
		(74) 代理人	100080089
			弁理士 牛木 護
		(74) 代理人	100119312
			弁理士 清水 栄松
		(74) 代理人	100137800
			弁理士 吉田 正義
		(72) 発明者	松山 泰男
			東京都新宿区大久保 3 丁目 4 番 1 号 学校
			法人 早稲田大学理工学部内
		(72) 発明者	勝又 尚人
			東京都新宿区大久保 3 丁目 4 番 1 号 学校
			法人 早稲田大学理工学部内
			最終頁に続く

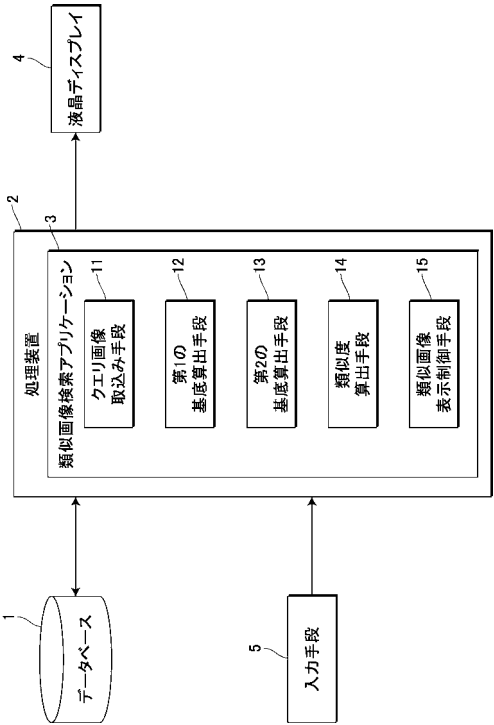
(54) 【発明の名称】 類似画像検索方法および類似画像検索装置

(57) 【要約】

【課題】 過学習や過汎化を引き起こすことなく、画像のもつ固有の情報を少ない冗長度で正確に表現できる類似画像検索方法および類似画像検索装置を提供する。

【解決手段】 1 枚のクエリ画像21から得られる基底と、1 枚の対比画像31から得られる基底とを直接比較することで、クエリ画像21に対する対比画像31の類似度を算出する。これにより、少ないサンプル数でありながら、クエリ画像21や対比画像31の特徴をよく反映した基底情報により、正しく類似画像を検索できる。しかも、1 枚のクエリ画像21に対して1 つの対比画像31の基底集合を学習させていることになるので、過学習や過汎化の問題も回避できる。

【選択図】 図 1



【特許請求の範囲】

【請求項 1】

基準となる画像の特徴を反映した第 1 の基底と、検索対象となる画像の特徴を反映した第 2 の基底とを比較することにより、前記検索対象となる画像の類似度を算出することを特徴とする類似画像検索方法。

【請求項 2】

選択された前記基準となる画像をクエリ画像として取り込み、

前記第 1 の基底を当該クエリ画像から算出すると共に、

前記第 2 の基底を前記検索対象となる画像から算出することを特徴とする請求項 1 記載の類似画像検索方法。

10

【請求項 3】

(a) 内積が最大となる前記第 1 の基底と前記第 2 の基底のペアを探し出し、

(b) 前記手順 (a) で探し出したペアを除く残りの前記第 1 の基底と前記第 2 の基底に対して、前記手順 (a) と同様で内積が最大となる前記第 1 の基底と前記第 2 の基底のペアを探し出し、求められた内積に加重をかけて加えて行き、

(c) 前記第 1 の基底と前記第 2 の基底の全てがペアとなるまで、前記手順 (a) と前記手順 (b) を繰り返して、前記検索対象となる画像の類似度を算出することを特徴とする請求項 1 または 2 記載の類似画像検索方法。

【請求項 4】

前記検索対象となる画像の最終的な類似度 S は、前記第 1 の基底と前記第 2 の基底とを比較することにより得られた第 1 の類似度 S_{basis} と、前記基準となる画像と前記検索対象となる画像との間の色に関する第 2 の類似度 S_{color} とを結合して得られることを特徴とする請求項 1 ~ 3 のいずれか一つに記載の類似画像検索方法。

20

【請求項 5】

前記検索対象となるの最終的な類似度 S は、加重パラメータ a を用いることにより、 $S = a S_{basis} + (1 - a) S_{color}$ (但し、 $0 \leq a \leq 1$) の式で算出され、前記加重パラメータ a が調整可能であることを特徴とする請求項 4 記載の類似画像検索方法。

【請求項 6】

基準となる画像の特徴を反映した第 1 の基底と、検索対象となる画像の特徴を反映した第 2 の基底とを比較することにより、前記検索対象となる画像の類似度を算出する類似度算出手段を備えたことを特徴とする類似画像検索装置。

30

【請求項 7】

選択された前記基準となる画像をクエリ画像として取り込むクエリ画像取込み手段と、

前記第 1 の基底を当該クエリ画像から算出する第 1 の基底算出手段と、

前記第 2 の基底を前記検索対象となる画像から算出する第 2 の基底算出手段と、

をさらに備えたことを特徴とする請求項 6 記載の類似画像検索装置。

【請求項 8】

(a) 内積が最大となる前記第 1 の基底と前記第 2 の基底のペアを探し出し、

(b) 前記手順 (a) で探し出したペアを除く残りの前記第 1 の基底と前記第 2 の基底に対して、前記手順 (a) と同様で内積が最大となる前記第 1 の基底と前記第 2 の基底のペアを探し出し、求められた内積に加重をかけて加えて行き、

(c) 前記第 1 の基底と前記第 2 の基底の全てがペアとなるまで、前記手順 (a) と前記手順 (b) を繰り返して、前記検索対象となる画像の類似度を算出するように、前記類似度算出手段を構成したことを特徴とする請求項 6 または 7 記載の類似画像検索装置。

40

【請求項 9】

前記第 1 の基底と前記第 2 の基底とを比較することにより得られた第 1 の類似度 S_{basis} と、前記基準となる画像と前記検索対象となる画像との間の色に関する第 2 の類似度 S_{color} とを結合して、前記検索対象となる画像の最終的な類似度 S を得るように、前記類似度算出手段を構成したことを特徴とする請求項 6 ~ 8 のいずれか一つに記載の類似画像検索装置。

50

【請求項 10】

前記検索対象となる画像の最終的な類似度 S は、加重パラメータ a を用いることにより、 $S = a S_{\text{basis}} + (1 - a) S_{\text{color}}$ (但し、 $0 \leq a \leq 1$) の式で算出され、前記加重パラメータ a を調整可能にする操作手段をさらに備えたことを特徴とする請求項 9 記載の類似画像検索装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、静止画像や動画像を対象となる情報源として、類似画像を検索する類似画像検索方法および類似画像検索装置に関する。 10

【背景技術】

【0002】

近年のマルチメディアの普及により、デジタルコンテンツの蓄積は益々増加の一途を辿っており、その一環をなす静止画像においては、特にその傾向が著しい。そして、このような大量のコンテンツを扱うには、検索やブラウジングといった技術が非常に重要となる。

【0003】

従来、静止画において最も一般的な手段として、テキストベースによる検索が知られていた。これは、画像を表現するテキストによって対象画像のラベル付けを行い、そのテキストを利用して画像の検索を行なうというものである。しかし、この方法では人手により分類して注釈をつけるラベル付けの作業を必要とし、時間と労力、ひいては費用が掛かりすぎる状況が生じてくる。さらに、ラベル付けを行なう際の感覚が人によって異なり、同じ画像であっても異なるラベル付けが行なわれてしまうので、意図したとおりの検索ができない。 20

【0004】

こうした問題を解決するために、内容による画像検索 (Content-Based Image Retrieval) 手法が提案されている。これは、キーワードを人手によって付加する代わりに、色、テクスチャ、物体形状といった画像の特徴を要素に持つ特徴ベクトル (feature vector) を作成し、その距離を画像の類似度として検索を行なう方法である。例えば非特許文献 1 には、空間分割と色情報に基づいて類似画像を検索する高度なシステム (QBICシステム: Query By Image Content System) が提案されており、現時点で既に商品化されているものもある。また別な手法として、似ている画像どうしでは画像を構成する色が似ているという仮定に基づき、画像中に色がどのような割合で含まれているのかを比較する色ヒストグラム法も知られている。 30

【0005】

上述した画像検索システムは、大きく分けて (1) 特徴量の選択, (2) 多次元のインデキシング, (3) システムデザインという 3 つの要素を持ち、どの要素に重点を置くのかによってシステムとしての個性があらわれる。殆どの画像検索システムでは、色、テクスチャ、形状などの画像の特徴を表わす特徴量をメタデータとして持っており、それらの幾つかを組み合わせるか、または全てを使用することで画像の検索を行なっている。つまり特徴量によって画像を表わすので、特徴量が増えればそれだけ画像表現の幅も広がる。しかし、これらの特徴量ベクトルの次元は数百に及ぶこともあるので、検索に用いる特徴量を増やし過ぎると、組み合わせや重み付けなどの多次元インデキシングの困難性が増加し、システムそのものとユーザーにとって負担となる。また、ユーザー毎に感覚や好み (形状や色の優先度など) が異なるので、一回の検索で所望の画像を抽出することは稀である。そのため、ユーザーからのフィードバックを次の検索へと反映することができる双方向性を持ったシステム設計が望まれる。 40

【0006】

一方、上記メタデータを利用した画像検索システムとは別に、主成分分析の基底 (Prin 50

Principal Component Analysis Bases: PCA 基底)、あるいは独立成分分析の基底 (Independent Component Analysis Bases: ICA 基底) を利用した画像 (顔) の検索方法が、例えば特許文献 1 などに開示されている。PCA や ICA は、画像の線形変換として知られるウェーブレット変換と同様に、画像を基底の重ねあわせとして表現するが、これらは基底関数そのものに源画像の情報が適応的に反映されるので、周波数、位相、位置、方向といったデザインパラメータを決定する必要がない。しかも ICA 基底を使う場合は、Gabor ウェーブレットのような哺乳類の視覚野と似た特性の基底が得られる。また、画像検索に利用される特徴量は、色、色レイアウト、テクスチャ、画像領域、物体形状に大きく分けることができるが、画像に ICA を適用することによって得られる ICA 基底は、画像のエッジやテクスチャであるといわれており、この ICA 基底を利用したものは、云わばテクスチャを特徴量とする画像検索システムに分類される。

10

【0007】

こうした特性を持つ ICA 基底からの応答を使う手法として、非特許文献 2 には Independent Spectral Representation 法 (独立スペクトル表現法) による類似画像の検索方法が提案されている。ここで図 19 を参照しながら、この方法の概念を説明すると、図 19 で 100 は情報源となる多数の画像で、先ずステップ S101 の手順では、それぞれの画像をサンプルデータとして ICA 基底集合を予め作成する。続くステップ S102 では、ICA 基底集合をフィルタバンクとして、各画像に対するフィルタリングを行なう。そしてステップ S103 において、各画像をフィルタリングして得られた応答のヒストグラムを特徴量として、検索対象となる画像に対する類似の判定を行なう。このように、独立スペクトル表現法では、ICA 基底をフィルタリングのために使用し、比較に際しては ICA 基底に対する係数 (重み) を使うことになる。

20

【特許文献 1】特開 2004 - 5456 号公報

【非特許文献 1】マイロン フリックナー、ハープリート ソーウニー、ウェーン ニブラック、ジョナサン アッシュレー、シャン ファン、バイロン ドム、モニカ ゴルガーニ、ジム ハフナー、デニス リー、ドラガティーン ベッチョーヴィック、デヴィッド スティール、ピーター ヤンカー (Myron Flickner, Harpreet Sawhney, Wayne Niblack, Jonathan Ashley, Qian Huang, Byron Dom, Monika Gorkani, Jim Hafner, Denis Lee, Dragutin Petkovic, David Steele, Peter Yanker: 「イメージおよびビデオコンテンツによる問合せ: QBIC システム (Query by Image and Video Content: The QBIC System)」(米国), IEEE コンピュータマガジン 28 巻 (IEEE Computer Magazine, Vol.28), 1995 年 9 月, p.23~32

30

【非特許文献 2】シューエン リウ、レイ チェン (Xiuwen Liu, Lei Cheng): 「認識画像の独立スペクトル表現法 (independent spectral representations of images for recognition)」(米国), ジャーナル オヴ ザ オプティカル ソサイアティ オヴ アメリカ 第 20 巻 7 号 (Journal of the optical society of America, Vol.20, No.7), 2003 年 7 月, p.1271~1282

【発明の開示】

【発明が解決しようとする課題】

【0008】

40

しかし、上記 Independent Spectral Representation 法は、ICA が学習アルゴリズムであるために、ICA 基底集合を作成する際に過学習 (over learning) と過汎化 (over generalization) を引き起こす問題が存在する。ここでいう過学習とは、学習モデルに対して学習サンプル数が少な過ぎるため、モデルがサンプルに特化してしまい、他のデータに対応できなくなってしまうことをいう。また過汎化とは、その逆に、データに対して学習モデルが小さ過ぎ、貧弱な識別能力しか持たないことをいう。つまり、ICA 基底集合を学習する時に、学習モデルの大きさと学習サンプルの数のバランスを取る必要がある。

【0009】

さらに画像データベースは、大量のデータを扱うことになるので、ICA による学習モデルも必然的に大きくなる。すると、学習に必要なサンプル数は指数関数的に増加し、基

50

底の学習に使うサンプルが相対的に少なくなるので、過学習が起こりやすくなる。逆に学習サンプル数にあわせて学習モデルを小さくすると、過汎化によって貧弱な識別能力しか持たないICA基底しか得られなくなる。

【0010】

加えて、独立スペクトル表現法では、ICA基底集合をフィルタとして使用する関係上、当該フィルタの選択が重要になる。実際、最適なフィルタを選択するために、フィルタ選択アルゴリズムを使う方法も知られてはいるが、大きな画像データに対してこうしたアルゴリズムを実行するには膨大な計算量が必要となり、実行は困難である。さらに、ICA基底集合を作成する際の上記学習と同様に、最適なフィルタの選択に使うトレーニングデータに関しても、問題が存在する。

10

【0011】

このように考えると、類似画像検索のための特徴量が優れたものであると見なせるためには、(1)画像のもつ固有の情報を、少ない冗長度で正確に表現できること、(2)ユーザーの感覚や好みを正しく且つ容易に反映できること、が必要になる。そして、このような要求を満たす類似画像検索方法や類似画像検索装置が望まれていた。

【0012】

本発明はこうした問題に鑑み、過学習や過汎化を引き起こすことなく、画像のもつ固有の情報を少ない冗長度で正確に表現することが可能になる類似画像検索方法および類似画像検索装置を提供することをその目的とする。

【課題を解決するための手段】

20

【0013】

本発明に係る請求項1の類似画像検索方法は、基準となる画像の特徴を反映した第1の基底と、検索対象となる画像の特徴を反映した第2の基底とを比較することにより、前記検索対象となる画像の類似度を算出することを特徴とする。

【0014】

本発明に係る請求項2の類似画像検索方法は、選択された前記基準となる画像をクエリ画像として取り込み、前記第1の基底を当該クエリ画像から算出すると共に、前記第2の基底を前記検索対象となる画像から算出することを特徴とする。

【0015】

本発明に係る請求項3の類似画像検索方法は、(a)内積が最大となる前記第1の基底と前記第2の基底のペアを探し出し、(b)前記手順(a)で探し出したペアを除く残りの前記第1の基底と前記第2の基底に対して、前記手順(a)と同様で内積が最大となる前記第1の基底と前記第2の基底のペアを探し出し、求められた内積に加重をかけて加えて行き、(c)前記第1の基底と前記第2の基底の全てがペアとなるまで、前記手順(a)と前記手順(b)を繰り返して、前記検索対象となる画像の類似度を算出することを特徴とする。

30

【0016】

本発明に係る請求項4の類似画像検索方法は、前記検索対象となる画像の最終的な類似度 S が、前記第1の基底と前記第2の基底とを比較することにより得られた第1の類似度 S_{basis} と、前記基準となる画像と前記検索対象となる画像との間の色に関する第2の類似度 S_{color} とを結合して得られることを特徴とする。

40

【0017】

本発明に係る請求項5の類似画像検索方法は、前記検索対象となる画像の最終的な類似度 S が、加重パラメータ a を用いることにより、 $S = a S_{basis} + (1 - a) S_{color}$ (但し、 $0 \leq a \leq 1$) の式で算出され、前記加重パラメータ a が調整可能であることを特徴とする。

【0018】

本発明に係る請求項6の類似画像検索装置は、基準となる画像の特徴を反映した第1の基底と、検索対象となる画像の特徴を反映した第2の基底とを比較することにより、前記検索対象となる画像の類似度を算出する類似度算出手段を備えている。

50

【 0 0 1 9 】

本発明に係る請求項 7 の類似画像検索装置は、選択された前記基準となる画像をクエリ画像として取り込むクエリ画像取込み手段と、前記第 1 の基底を当該クエリ画像から算出する第 1 の基底算出手段と、前記第 2 の基底を前記検索対象となる画像から算出する第 2 の基底算出手段と、をさらに備えている。

【 0 0 2 0 】

本発明に係る請求項 8 の類似画像検索装置は、(a) 内積が最大となる前記第 1 の基底と前記第 2 の基底のペアを探し出し、(b) 前記手順 (a) で探し出したペアを除く残りの前記第 1 の基底と前記第 2 の基底に対して、前記手順 (a) と同様で内積が最大となる前記第 1 の基底と前記第 2 の基底のペアを探し出し、求められた内積に加重をかけて加え
10
て行き、(c) 前記第 1 の基底と前記第 2 の基底の全てがペアとなるまで、前記手順 (a) と前記手順 (b) を繰り返して、前記検索対象となる画像の類似度を算出するように、前記類似度算出手段を構成したものである。

【 0 0 2 1 】

本発明に係る請求項 9 の類似画像検索装置は、前記第 1 の基底と前記第 2 の基底とを比較することにより得られた第 1 の類似度 S_{basis} と、前記基準となる画像と前記検索対象となる画像との間の色に関する第 2 の類似度 S_{color} とを結合して、前記検索対象となる画像の最終的な類似度 S を得るように、前記類似度算出手段を構成したものである。

【 0 0 2 2 】

本発明に係る請求項 10 の類似画像検索装置は、前記検索対象となる画像の最終的な類似度 S が、加重パラメータ a を用いることにより、 $S = a S_{basis} + (1 - a) S_{color}$ (但し、 $0 \leq a \leq 1$) の式で算出され、前記加重パラメータ a を調整可能にする操作手段をさらに備えたものである。
20

【 発明の効果 】

【 0 0 2 3 】

請求項 1 の類似画像検索方法または請求項 6 の類似画像検索装置によれば、1 枚の基準となる画像の基底と、1 枚の検索対象となる画像の基底とを直接比較することで、検索対象となる画像の類似度を算出している。そのため、少ないサンプル数でありながら、各画像の特徴をよく反映した基底情報により、正しく類似画像を検索することができる。しかも、1 枚の基準となる画像に対して 1 つの検索対象となる画像の基底集合を学習させてい
30
ることになるので、過学習や過汎化の問題も回避できる。つまり、過学習や過汎化を引き起こすことなく、画像のもつ固有の情報を少ない冗長度で正確に表現することが可能になる。

【 0 0 2 4 】

請求項 2 の類似画像検索方法または請求項 7 の類似画像検索装置によれば、複数の画像から基底を求めて、この基底を各画像にフィルタリングして得た応答を使って類似度を算出するのではなく、1 枚のクエリ画像から得られる第 1 の基底と、1 枚の検索対象となる画像から得られる第 2 の基底とを直接比較することで、クエリ画像に対する検索対象となる対比画像の類似度を算出している。そのため、少ないサンプル数でありながら、クエリ画像や検索対象となる画像の特徴をよく反映した基底情報により、正しく類似画像を検索
40
することができる。しかも、1 枚のクエリ画像に対して 1 つの検索対象となる画像の基底集合を学習させていることになるので、過学習や過汎化の問題も回避できる。つまり、過学習や過汎化を引き起こすことなく、画像のもつ固有の情報を少ない冗長度で正確に表現することが可能になる。

【 0 0 2 5 】

請求項 3 の類似画像検索方法または請求項 8 の類似画像検索装置によれば、類似度を算出するに際し、基準となる画像と検索対象となる画像の各基底集合について、その全ての基底のペアの内積を最大化する必要はない。すなわち、ここでは n 個の基底成分に対して n^2 オーダーの計算を行なえばよく、良好な近似手法を取り入れた現実的な類似度の算出を実現できる。また、類似度を算出するに際して基底成分の加重和を用いることにより、
50

基底集合に存在する非決定性を回避することができる。

【0026】

請求項4の類似画像検索方法または請求項9の類似画像検索装置によれば、検索対象となる画像の最終的な類似度が、基準となる画像および検索対象となる画像の基底に基づく類似度だけでなく、基準となる画像や検索対象となる画像がもつ色空間に関わる類似度を加味して判断される。そのため、物体の位置の移動により敏感な類似画像を検索することができる。

【0027】

請求項5の類似画像検索方法または請求項10の類似画像検索装置によれば、ユーザーが加重パラメータaの値を任意に調整できるので、類似画像の検索にユーザーの好みを簡単に反映させることができ、ユーザーにとって苦痛なパラメータ設定を少なくできる。

【発明を実施するための最良の形態】

【0028】

以下、添付図面に基づき、本発明における類似画像検索方法と、それを実現する装置の好ましい実施形態を詳しく説明する。図1は、システム構成を模式的に示したものであるが、この図において、1は静止画像若しくは静止画像の連続体としての動画像（以下、これらを単に画像という）の集合を記憶保存するデータベースで、このデータベース1は例えばコンピュータなどの処理装置2に少なくとも読み出し可能な状態に接続される。処理装置2は周知のように、本実施形態の特徴となる類似画像検索アプリケーション3を含んだ各種アプリケーションを備えており、必要に応じてデータベース1に蓄積された画像を表示手段である液晶ディスプレイ4で適宜表示できるようになっている。なお、データベース1は処理装置2に内蔵または外付けされる記憶媒体（ハードディスクなど）や、処理装置2に通信手段を介して接続されるサーバであってもよく、どのような形態であるかは特に限定されない。また処理装置2は、例えばマウスやキーボードなどの入力手段5を備えている。

【0029】

類似画像検索アプリケーション3は、入力手段5によってユーザーが選択したクエリ画像（query image）を取込むクエリ画像取込み手段11と、前記クエリ画像の小区画をサンプルデータとして、当該クエリ画像の基底を求める第1の基底算出手段12と、データベース1内から検索の対象となる画像を読み出し、この対象画像の小区画をサンプルデータとして、当該対象画像の基底を求める第2の基底算出手段13と、前記クエリ画像の基底と前記対象画像の基底を直接比較し、クエリ画像に対する対象画像の類似度を算出する類似度算出手段14と、前記類似度の高い順に前記データベース1中の画像を液晶ディスプレイ4に一乃至複数表示させる類似画像表示制御手段15と、をそれぞれ備えている。ここで利用できる基底としては、主成分分析基底（PCA基底）や独立成分分析基底（ICA基底）であることが好ましい。その理由は、これらの基底が認識あるいは分類しようとする情報源に依存した形で現れるからである。また、本実施形態においては、静止画像がもっとも適切な対象となるが、静止画像を連続化した動画像であっても構わない。さらに、クエリ画像はデータベース1内に保存される画像以外のものを利用してよい。

【0030】

図2は、上記構成に基づく類似画像の検索方法の処理手順を示したものである。同図において、21は入力手段5により特定され、クエリ画像取込み手段11に取込まれたクエリ画像で、このクエリ画像21は二次元状に配列された画素（ピクセル）の集合により構成される。第1の基底算出手段12はステップS1において、クエリ画像21を適宜分割して得た小区画22をサンプルデータとして、1枚のクエリ画像21からPCAまたはICAの基底を算出する。一方、31はデータベース1に蓄積された検索対象となる画像（対比する画像）で、これも二次元状に配列された画素の集合により構成される。第2の基底算出手段13はステップS2において、対比画像31を適宜分割して得た小区画32をサンプルデータとして、1枚の対比画像31からPCAまたはICAの基底を算出する。こうして、クエリ画像21と対比画像31の各基底が算出されると、類似度算出手段14は次のステップS3にて、双方の

10

20

30

40

50

基底どうしを比較し、類似している画像であるほど、クエリ画像21と対比画像31における各基底ベクトルの方向が似ていることに基づき、続くステップS4で類似度を算出する。類似画像検索アプリケーション3は、データベース1内の複数の対比画像31について、ステップS2～ステップS4の各手順を同様に行ない、類似画像表示制御手段15により類似度の高い対象画像31を液晶ディスプレイ4に表示させる。

【0031】

ここで注目すべきは、本実施形態では従来技術のように各画像をフィルタリングして得られた基底の応答を特徴量とするのではなく、クエリ画像21や対比画像31から得られた基底そのものを特徴量として、類似度の判断を行なっていることである。すなわち、本実施形態ではクエリ画像21に対する対比画像31の類似度算出に際して、PCAまたはICAの基底しか用いていない。

10

【0032】

本実施形態では、上記方法を採用するに当たり、tightフィッティングの概念を導入している。tightフィッティングとは、あえてある1つのクラスのみを学習することにより、学習モデルをそのクラスに特化させることである。こうすることにより、そのクラスの特徴をよく反映したモデルが得られ、こうして得られたモデルパラメータを比較することで、クラスの識別を行なうことができる。すなわちtightフィッティングでは、ただ1つのクラスを学習するだけでよいため、必要なサンプルの数が少なく済み、独立スペクトル表現法のような過学習や過汎化の問題を回避できる。これは、ICA基底の場合には、1枚の画像に対して1つの基底集合を学習させることに相当する。さらにモデルパラメータの比較は、得られた基底集合を比較することに相当する。こうして、本実施形態で採用する類似画像の検索方法や検索装置は、クエリ画像21と対比画像31との基底情報を比較することで、過学習や過汎化の問題を解決して、画像のもつ固有の情報を少ない冗長度で正確に表現することが可能になる。

20

【0033】

図3は、第1の基底算出手段12や第2の基底算出手段13が基底を算出するまでの処理手順を模式的に示したものである。同図において、41は前述のクエリ画像21や対比画像31に相当する1枚の画像で、ここではステップS11のように、画像41を64に等分割した縦横8ピクセルの小区画42が、画像41のサンプルデータとして用いられる。次に基底算出手段12, 13は、小区画42の各ピクセルを構成する $8 \times 8 = 64$ 次元（カラー画像の場合には、 $8 \times 8 \times 3 = 192$ 次元）のデータベクトル（各要素は、 $x_1, x_2, x_3, \dots, x_{64}$ からなる）を、縦に並べた列ベクトルとして各々行列 x のなかに組み入れる（ステップS12, S13）。基底算出手段12, 13は、64本の基底ベクトルからなる行列 W^{-1} と、同じく64本の重み付け係数（各要素は、 $y_1, y_2, y_3, \dots, y_{64}$ からなる）からなる行列との積が、前記データベクトルの行列 x に等しい（ $W^{-1}y = x$ ：すなわち $y = Wx$ で、 W は分離フィルタとなる）ことから、ステップS13, S14に示す式に基づいて、1枚の画像41から各基底ベクトルの集合を算出する。

30

【0034】

次に、上記処理手順において、主成分分析（PCA）と独立成分分析（ICA）の基本的な性質を説明する。まず、PCAとICAで取り扱うデータの前処理法について述べる。前処理は、画像の源信号 x が統計的性質の良いデータとなるように行なうもので、次の数1, 数2のように可逆である。

40

【0035】

【数1】

$$x \leftarrow x - E[x]$$

【0036】

【数2】

$$x \leftarrow x / \sqrt{|x|^2}$$

50

【 0 0 3 7 】

ここで、数 3 に示す記号は期待値をあらわす演算子である。K 個のサンプルデータが与えられたときに、数 4 に示すサンプルデータの平均によってこれを置き換える。

【 0 0 3 8 】

【 数 3 】

$$E[x]$$

【 0 0 3 9 】

【 数 4 】

$$1/K \sum_{k=1}^K x(k)$$

10

【 0 0 4 0 】

但し、上記数 4 において、 $x(k)$ は $m \times m$ の大きさの画像パッチをラスタースキャンして並べたものである。例えば、カラー画像を取り扱うデータでは、 $x(k)$ の次元が $3 \times m \times m$ となる。このサンプルデータの具体的な抽出方法については、後ほど説明する。

【 0 0 4 1 】

数 1 は、 x を平均 0 にする最も簡単な操作であるが、PCA と ICA においては、この操作により成分間に線形の依存が生じるため、有効な次元が 1 つ減るという副次的な効果を生じる。これにより得られる独立成分は 1 つ少なくなってしまうが、引いた平均値、すなわち直流成分は特殊な独立成分であると考えられることもできる。

20

【 0 0 4 2 】

第 1 の基底算出手段 12 や第 2 の基底算出手段 13 が画像の PCA 基底を算出する場合、基底算出手段 12, 13 は最初に数 5 に示す共分散行列の計算を行なう。

【 0 0 4 3 】

【 数 5 】

$$C = E[xx^T]$$

【 0 0 4 4 】

次に、対角行列 D と E の計算を行なう。D の対角成分は、前記数 5 の左辺 C の固有値を降順に並べた行列であり、行列 E の列ベクトルは、それらの固有値に対応する固有ベクトルである。このとき、PCA による変換行列 V は以下に示す数 6 のようになる。

30

【 0 0 4 5 】

【 数 6 】

$$V = D^{-1/2} E^T$$

【 0 0 4 6 】

上記数 6 を使用し、次の数 7 となるようにデータベクトル x を数 8 のように変換する。

【 0 0 4 7 】

【 数 7 】

$$E[zz^T] = I$$

40

【 0 0 4 8 】

【 数 8 】

$$z = Vx$$

【 0 0 4 9 】

これは白色化またはスフィアリングと呼ばれる処理である。これにより、変換されたデータベクトル x は、互いに無相関で分散が 1 という性質のよいデータとなる。これは、PCA に関する処理であるが、ICA の前処理とすることもできる。

【 0 0 5 0 】

画像を取り扱う場合、データベクトル x は高次となるため、基底算出手段 12, 13 は小さ

50

な固有値を切り捨ててデータの次元縮約も行なう。これはローパスフィルタの役割を果たすと同時に、計算時間とメモリを節約するのに有効である。上記数 6 , 数 8 より、次の数 9 に基づき縮約された z から、右辺の x (以後、英文字の上に付された $\hat{\cdot}$ や \wedge の記号は、便宜上横に並べて表記する) を回復する操作である。このように用いられる U^{\wedge}_{PCA} の各列ベクトルは、PCA 基底と呼ばれる。なお、これ以降は特に断りのない限り、 x を単に x と表記する。

【 0 0 5 1 】

【 数 9 】

$$(ED^{1/2})z \stackrel{\text{def}}{=} \hat{U}_{PCA} z = \bar{x}$$

10

【 0 0 5 2 】

上述した PCA 基底、すなわち U^{\wedge}_{PCA} の列ベクトルは直交であり、データの形状によっては制約となりすぎる場合がある。代わりに ICA を採用すれば、直交という条件はなく、その分だけよい情報源モデルを構成し得る。

【 0 0 5 3 】

この ICA 基底に関し、今観測された n 次元のデータベクトル x は、 $n \times n$ の未知の情報源ベクトル s の各成分を未知の混合行列 $A = [a_1, \dots, a_n]$ によって重ね合わせたものと仮定する。当該データベクトル x は、数 10 のように表わせる。

【 0 0 5 4 】

【 数 10 】

$$x = As = \sum_{i=1}^n a_i s_i$$

20

【 0 0 5 5 】

このときの独立成分の目的は、(a) $i \neq j$ のとき、情報源ベクトル s_i と s_j は、互いに独立である、(b) たかだか 1 つの s を除き、非ガウシアンである。以上の 2 つの仮定の下で、混合行列 A と信号源 s とを推定することである。このような独立成分の分析では、各成分 s_i と ($i = 1, \dots, n$) の独立性を様々なコスト関数の最適化に置き換えることで、混合行列 A と信号源 s の推定を行なうことができる。 z を前処理済のデータベクトルとすると、次のように表わすことができる。

30

【 0 0 5 6 】

【 数 11 】

$$\hat{s} = \hat{W}z = \hat{W}Vx$$

【 0 0 5 7 】

数 11 を前記数 10 と比較すると、混合行列 A は $(W^{\wedge}V)^{-1}$ と一致する。仮に、 x の代わりに z を考えるならば、混合行列 A に一致する行列は W^{\wedge} となる。このような手法には、(1) 最尤推定により尤度の最大化を行なう方法、(2) 平均相互情報量の最小化を勾配法により行なう方法、(3) 不動点法により非ガウス性の増大を行なう方法、(4) 高次統計量の最大化を行なう方法が知られており、それぞれの手法は全く異なるが、どの手法から得られた基底であっても、本実施形態における類似画像検索に利用することができる。

40

【 0 0 5 8 】

一つの例として、本実施形態では上記 (3) の方法に基づく Fast ICA を、画像から基底を得るアルゴリズムとして使用する。本願発明者は、上記 (2) の平均相互情報量を最小化する ICA アルゴリズムの一種として、モーメント法を用いた高速化を図り、この方法の短所である速度問題を補う - ICA アルゴリズムを提案しているが、超一次収束する - ICA アルゴリズムに対して、最大で 3 次収束する Fast ICA アルゴリズムの方がより高速な収束が可能である。

50

【 0 0 5 9 】

ここで F a s t I C A の概要について説明すると、不動点法に基づく F a s t I C A では独立性を、次のように " Negentropy " $J(y)$ で表わしている。

【 0 0 6 0 】

【数 1 2】

$$J(y) = H(y_{gauss}) - H(y) \geq 0$$

【 0 0 6 1 】

ここで、 y は s^{\wedge} に収束する確率変数であり、次の数 1 3 のように表わせる。また、 $H(y)$ はその微分エントロピーである。そして、 y_{gauss} は y と同じ分散のガウス分布をもつ確率変数であり、Negentropy $J(y)$ は、確率変数 y がガウス分布からどれだけ離れているのかをあらわしており、この量が大きい程、各成分は独立に近いと考えることができる。その理由は、中心極限定理により、独立な確率成分の和は急速にガウス分布に近づくからである。

10

【 0 0 6 2 】

【数 1 3】

$$y = Wz$$

【 0 0 6 3 】

$J(y)$ の各成分は、次のように近似できる。

20

【 0 0 6 4 】

【数 1 4】

$$J(y_i) \approx c \{E[G(y_i)] - E[G(v)]\}^2$$

【 0 0 6 5 】

ここで、 G は 2 次ではない非線形関数であり、 c は定数、 v は平均 0 , 分散 1 のガウス変数である。後述するオピニオンテストでは、多様な情報源に対して適用可能な次の関数 $G(u)$ を使用した。

【 0 0 6 6 】

【数 1 5】

$$G(u) = \log \cosh(u)$$

30

【 0 0 6 7 】

ここで、前記数 1 3 の w_i^T を W の行ベクトルとし、さらに $y_i = w_i^T z$ とする。このとき、1 つの独立成分を見つけるためには、次の数 1 6 に示す制約下で、数 1 7 に示す J_G を最大化すればよいことになる。

【 0 0 6 8 】

【数 1 6】

$$E \left[\left(w_i^T z \right)^2 \right] = 1$$

40

【 0 0 6 9 】

【数 1 7】

$$J_G(w_i) = \{E[G(w_i^T z)] - E[G(v)]\}^2$$

【 0 0 7 0 】

上記数 1 7 における J_G の最大化は、次の更新式を反復することで行なわれる。

【 0 0 7 1 】

【数 1 8】

$$w_i \leftarrow E \left[z g(w_i^T z) \right] - E \left[g'(w_i^T z) \right] w_i$$

50

【 0 0 7 2 】

ここで、 $g(u)$ は $G(u)$ を微分したもので、 $g'(u)$ は $g(u)$ を微分したものである。したがって、数 1 5 の例では、 $g(u) = \tanh(u)$ であり、 $g'(u) = \text{sech}^2(u)$ である。数 1 8 は i 番目の独立成分を求める式であるが、行列の形の更新式により複数成分を同時に求めることもできる。この場合、 w_i が制約条件を満たすように、毎回の更新の後で W を次の正規直交化を行なうだけでよい。

【 0 0 7 3 】

【 数 1 9 】

$$W \leftarrow (WW^T)^{-1/2} W$$

10

【 0 0 7 4 】

このようにして反復を繰り返すことで、 W は W^* に収束する。

【 0 0 7 5 】

最終的に、ICA 基底は次のようにして求めることができる。前記数 1 1 を次の数 2 0 のようにすると、これは行列 U^*_{ICA} を用いることにより、縮約された s^* から x を回復する操作となる。このように用いられる U^*_{ICA} の各列ベクトルは、ICA 基底と呼ばれる。なお、ICA の場合も特に断りのない限り、 x を単に x と表記する。

【 0 0 7 6 】

【 数 2 0 】

$$(\hat{W}V)^{-1} \hat{s} \stackrel{\text{def}}{=} \hat{U}_{ICA} \hat{s} = \bar{x}$$

20

【 0 0 7 7 】

このように PCA 基底や ICA 基底は画像の情報を反映するため、類似画像の検索に利用することができる。とりわけ ICA 基底では、画像のテクスチャ情報を反映するといわれているが、こうした PCA 基底や ICA 基底には、振幅と順序に関する非決定性という好ましくない問題がある。したがって、この非決定性を回復することが必要となる。PCA の場合、固有値を降順に並べることで基底の順序を回復することができる。また、変換行列 V を正規直行化すれば、振幅に関する非決定性の主要部分も回避できる。

【 0 0 7 8 】

ところが ICA では、順序非決定性が常に存在する。その理由は、混合行列 A と信号源 s が共に未知であるため、いずれか一方の順序または振幅が変化しても、他方によってその変化を打ち消すことができるからである。本実施形態では後述のように、こうした順序の非決定性の問題を回避することのできる基底集合の間の類似度を提案している。なお、振幅に関する非決定性は、数 2 の前処理における分散の正規化と、数 1 9 における反復中の正規直行化によって、その大部分は回避されている。

30

【 0 0 7 9 】

次に、PCA または ICA を用いた画像基底の導出について説明する。まず、画像から得られた小区画のカラーサンプルを、PCA と ICA の両手法で使われるデータベクトルへと変換する方法について述べる。前述のように、ある画像（クエリ画像 21 や対比画像 31）が与えられた時に、そこから次の数 2 1 に表わされる大きさのパッチと呼ばれる領域を取り出す。なお、当該領域は数 2 2 のように表わせる。前記数 1 のデータベクトル x は、 $I(x, y)$ をラスタースキャンした列ベクトルとなる

40

【 0 0 8 0 】

【 数 2 1 】

$$m \times m \stackrel{\text{def}}{=} n$$

【 0 0 8 1 】

【 数 2 2 】

$$\{I(x, y)\}_{x,y=1}^m$$

50

【 0 0 8 2 】

ここではカラー画像を扱っているので、各ピクセルは3つの値を持つ。扱える色空間としては、RGB, YIQ, HSVがあるが、それぞれの色空間において、あるピクセルを $[C_1(x, y), C_2(x, y), C_3(x, y)]$ という形で表わす。このときデータベクトル x は、次のようになる。

【 0 0 8 3 】

【 数 2 3 】

$$x = [C_1(x_1, y_1), C_1(x_2, y_1), \dots, C_1(x_m, y_m), \\ C_2(x_1, y_1), C_2(x_2, y_1), \dots, C_2(x_m, y_m) \\ C_3(x_1, y_1), C_3(x_2, y_1), \dots, C_3(x_m, y_m)]^T \\ \stackrel{\text{def}}{=} [x_{c_1}, x_{c_2}, x_{c_3}]^T$$

10

【 0 0 8 4 】

例えばRGB色空間の場合は、3nの次元を持つデータベクトル $x = [x_R, x_G, x_B]^T$ (Tは転置) という形で表わせる。

20

【 0 0 8 5 】

与えられた画像からPCA基底を得るには、先ず画像からN個の3n次元のデータベクトルを抽出した後に、前記数9中の U^{\wedge}_{PCA} を得るために、数5における共分散行列の計算を行なう。より具体的には、図4のフローチャートに示すように、ステップS21におけるサンプルデータの抽出では、画像の縦横の大きさが通常異なることから、仮想的に例えば $128 \times 128 = 16384$ の格子が画像上に存在していると考え、それぞれの格子上の点を基点として 8×8 の大きさのパッチ(小区画)を抽出する。この手順によって、 $N = 16384$ 個のデータベクトル x を抽出できる。次のステップS22では、データベクトル x の平均が0、分散が1となるように、上記数1および数2を利用した前処理を行なう。その次のステップS23では、数5に基づき共分散行列の固有値を計算し、値の大きいものから順に並べる。このうちの上位1/4の順位をもつ固有値と、それに対応する固有ベクトル E を用いて、数6の変換行列 V を算出する。これにより $8 \times 8 \times 3 = 192$ の次元を48にまで縮約できる。ここで次元を減少させたデータは、ICA基底を抽出する際にも利用できる。

30

【 0 0 8 6 】

一方、与えられた画像からICA基底を得るには、次の数24に示す混合行列 A^{\wedge} を計算するために、図5に示すフローチャートの各処理手順を実行する。

【 0 0 8 7 】

【 数 2 4 】

$$\hat{A} = (\hat{W}V)^{-1} = ED^{1/2}\hat{W}^T$$

40

【 0 0 8 8 】

同図において、ステップS21, S22の各手順はPCA基底抽出の場合と同じである。その次のステップS25では、行列 W の更新を行なう。より具体的には、最初のステップS25-1では、例えば同じ次元の単位行列 I を用いて、行列 W を初期化する。次のステップS25-2では、数18に従って行列 W を更新し、ステップS25-3において、その更新の毎に行列 W を正規直行化する。この行列 W の正規直行化に際しては、前述の数19を利用する。そして、行列 W がフーリエ変換した W^{\wedge} に収束するまで、ステップS25-2, S25-3の手順を反復実行する(ステップS25-4)。ここでの反復における収束条件は、次の数25を満たすことである。

【 0 0 8 9 】

50

【数 2 5】

$$\frac{1}{N_B} \sum_{i=1}^{N_B} \left| 1 - \left| w_i^{new} \cdot w_i^{old} \right| \right| < \varepsilon$$

【0 0 9 0】

上記式において、 \cdot は内積を表わす。また、 N_B は基底の数であり、 ε は収束の度合いを表わす定数である。一つの実験例として、 $\varepsilon = 5.0 \times 10^{-6}$ を使用し、数 2 5 の左辺がこの値未満になることを収束条件としている。

【0 0 9 1】

当該収束条件は、次のように解釈できる。行列 W の更新は、白色化された空間内で、 $w_i = 1$ の制約下で J_0 を最大にする w_i の方向を探すアルゴリズムである、したがって、更新前のベクトル w_i^{old} と更新後のベクトル w_i^{new} との方向が一致したときに、学習が収束したと判断することができる。これは、 $|w_i^{new} \cdot w_i^{old}| = 1$ の式と等価である。因みに、この式の左辺が絶対値となっているのは、 w_i も $-w_i$ も同じ方向を指していると考えられるためである。

【0 0 9 2】

その次のステップ S 2 6 では、数 2 4 を用いて混合行列 A を計算する。

【0 0 9 3】

ここで、ICA 基底に関する予備実験の結果を示す。図 6 における上段の画像 (a), (b) は、各々異なるテクスチャを持つ白黒の源画像で、画像 (a) は 756×504 ピクセル、画像 (b) は 768×512 ピクセルである。また、下段の画像 (a'), (b') は、図 5 に示すフローチャートの手順に基づき、自然画像 (a), (b) から得られた ICA 基底である。但し、源画像が白黒である関係上、次元数は $8 \times 8 = 64$ となる。この小区画は、次の数 26 に示すように、推定された混合行列 A^{\wedge} から得られた基底の列ベクトル a^{\wedge}_i の二次元表現である。

【0 0 9 4】

【数 2 6】

$$\hat{a}_i = \{a_i(x, y)\}_{x, y=1}^m, (i = 1, \dots, n)$$

【0 0 9 5】

列ベクトル a^{\wedge}_i のそれぞれの要素は、0 ~ 255 にスケールされている。また、有効な基底の数は $64 - 1 = 63$ となっているが、これは源信号 x の平均を 0 にする前処理を行っているからである。そして、源画像 (a), (b) と ICA 基底 (a'), (b') の各組を比較すると、それぞれの ICA 基底には、空間周波数が高いパターンを有する基底と、空間周波数の低いパターンを有する基底が存在することと、(a') と (b') の基底集合のパターンには明らかな違いが存在していることがわかる。

【0 0 9 6】

こうした基底集合の違いを定量的に調べるため、当該予備実験では Gabor ウェーブレットを利用した。これは Gabor-fit 法として知られており、ICA 基底と Gabor ウェーブレットとの誤差を最小にする Gabor ウェーブレットのパラメータを計算する手法である。

【0 0 9 7】

図 7 は、Gabor-fit 法を使った ICA 基底 (a'), (b') の比較結果を示している。同図において、中心角が基底の方向を $[0, \pi]$ の範囲であらわしており、中心からの距離が基底の周波数をあらわしている。また、それぞれの黒丸は、63 個ある (a') の基底の 1 つをあらわしており、同様に白い四角は、63 個ある (b') の基底の 1 つをあらわしている。黒丸の ICA 基底 (a') に注目すると、0 rad = π rad は同じ方向を指していると考えられ、ここでは明らかに 2 つのクラスが存在する。一方、白い四角の ICA 基底 (b') に注目すると、これはグラフ中の全体に拡がっている。これは、次のような源画像の特性による。画像 (a) では、葉と背景の部分で輝度の変化が少なく、画像中に明確なエッジが現れていない。逆に画像 (b) では、輝度の変化が激しく、明確なエッジを

10

20

30

40

50

捉えることができる。そのため、画像中の複雑なテクスチャパターンを最適に表現できるようにICA基底のパターンが得られる。以上より、次のことが云える。

(1) ICA基底は画像のエッジとテクスチャパターンの性質を反映する。

(2) 大きく異なる画像からは、異なるICA基底が得られる。このことは、ICA基底を類似画像検索に利用できるという考えを支持することになる。

(3) ICA基底は位置情報を持っておらず、当該位置情報は基底の重ね合わせの係数が持っている。最終的な類似度としてこの基底情報を重視するならば、画像中の物体の移動に対してロバストな検索となる。

【0098】

このように、ICA基底に関する予備実験では、単なる主観的な判断に頼らず、ICA基底をGabor-fit法により定量的に比較し、基底そのものが画像の分離性を持つことを実証している。 10

【0099】

次いで、類似度算出手段14における類似度の算出方法について、より詳細に説明する。本実施形態における類似度演算手段14は、クエリ画像21と対比画像31の各基底集合の間の類似度 S_{basis} のみならず、画像の部分領域の平均色に基づく類似度 S_{color} を結合して、最終的な類似度の判断を行なうのが好ましい。 S_{basis} は画像(クエリ画像21および対比画像31)のPCA基底またはICA基底に基づく類似度であり、それらの方向の一致度を類似度としている。図8はその概念を説明する図であるが、図8(a)のように分布が近い場合には、そこから得られる基底の方向も近くなり、基底どうしの一致度が高くなる。 20
一方、図8(b)のように分布が似ていなければ、得られる基底の方向も大きく異なる。つまり、分布が大きく異なっていくに従い、その一致度も低くなっていくと推測できる。

【0100】

しかし、ICA基底では前述した順序の非決定性が存在し、基底は規則正しく並んでいる訳ではないので、対応する基底の発見は難しくなる。PCA基底の場合、基底は直交であり、固有値の大きい順に順序付けされているが、単純に基底どうしを対応付けることはできない。このことは、画像のように高次の場合に当てはまるであろうことは容易に推測できる。ここでの S_{basis} は、順序の非決定性を考慮し、PCA基底とICA基底のいずれかの基底集合の比較も行なうことができる類似度となっている。

【0101】

S_{basis} は、数27、数28に示す2つの基底集合の全ての基底のペア(a_i^{11}, a_j^{12})の内積を最大化することで計算できる。ここで、k番目の画像から得られたi番目の基底を a_i^{1k} とする。この基底は、PCA基底とICA基底の何れであっても構わないが、この計算にはn!オーダーの組み合わせでの内積計算が必要であるため、計算量が膨大となる。本実施形態では、画像基底に存在するスケールや順序の問題に対し、PCA基底とICA基底のどちらにも有効な方法として、2つの基底集合のなかで、まず内積が最大となる一つのペアを選び出し、それを基準にして次々に内積が最大のペアを求めて重み付け加算を行なう類似度 S_{basis} を定義している。より具体的には、(a)内積が最大となる基底のペア a_i^{11}, a_j^{12} を探し出し、(b)探し出したペア a_i^{11}, a_j^{12} を除いた残りの $(n-1) \times (n-1)$ 個の基底に対して、手順(a)と同様に計算を行って、内積が 40
最大となる基底のペアを探し出し、求められた内積に加重をかけて加算し、(c)全ての基底がペアとなるまで、手順(a)と手順(b)を繰り返して、2つの画像の間の類似度 S_{basis} を決定するようになっている。

【0102】

【数27】

$$\left\{ a_i^{11} \right\}_{i=1}^n$$

【0103】

【数 2 8】

$$\left\{ \mathbf{a}_i^{I_2} \right\}_{i=1}^n$$

【0 1 0 4】

上述した手順 (a) ~ (c) は、詳細な加重を含めて、次のような C 言語的コードで表現できる。すなわち、このようなプログラムを実行する類似度算出手段 14 を組み込んでおけば、 $n!$ オーダーの計算を n^2 オーダーで済ませる良好な近似手法を取り入れた類似度 S_{basis} の算出を実現できる。また、類似度 S_{basis} を算出するに際して基底成分の加重和を用いることにより、基底集合に存在する非決定性を回避することができる。

【0 1 0 5】

【数 2 9】

```
for ( k = 1 ; k <= n ; k++ ) {
```

$$\mathbf{a}_k^{I_1} = \mathbf{a}_k^{I_1} / \|\mathbf{a}_k^{I_1}\|; \quad \mathbf{a}_k^{I_2} = \mathbf{a}_k^{I_2} / \|\mathbf{a}_k^{I_2}\|;$$

}

$$S_{basis} = 0;$$

```
for ( k = 1 ; k <= n ; k++ ) {
```

$$S_{basis} += (n - k + 1) \times \max_{i,j} (|\mathbf{a}_i^{I_1} \cdot \mathbf{a}_j^{I_2}|);$$

$$\mathbf{a}_i^{I_1} = \mathbf{a}_j^{I_2} = 0;$$

}

$$S_{basis} \times = 2 / (n \times (n + 1));$$

【0 1 0 6】

上記のコード中で、 n はそれぞれの基底集合中の基底の数をあらわしている。探し出した基底 $\mathbf{a}_i^{I_1}$ 、 $\mathbf{a}_j^{I_2}$ は、内積が計算される前にそれぞれ $\|\mathbf{a}_i^{I_1}\| = 1$ と $\|\mathbf{a}_j^{I_2}\| = 1$ に正規化される。基底は正規化された状態でデータベース 1 に格納されるので、この処理は普通必要とならない。内積の最大値の計算において絶対値が表われるのは、 $-\mathbf{a}_i^{I_1}$ であっても $\mathbf{a}_j^{I_2}$ であっても同じ方向を表していると考えられるためである。この手法を本実施形態の類似度算出手段 14 に取り入れることにより、数 2 7 や数 2 8 で示した 2 つの画像の基底集合との間において、その類似度 S_{basis} の計算に必要な計算量は数 3 0 に示す値まで減少する。

【0 1 0 7】

【数 3 0】

$$\sum_{i=1}^n i^2$$

【0 1 0 8】

本実施形態の類似度算出手段 14 は、前述のように画像の部分領域の平均色に基づく類似度 S_{color} を考慮に入れている。この画像の平均色に基づく類似度 S_{color} は、基底の類似度 S_{basis} とは異なる特性を有する。類似度 S_{color} を算出するに際しては、架空の格子を利用して 1 枚の画像を小区画に分割する。例えば前述の例では、 8×8 の大きさに 16384 個のサンプルが抽出される。次に、次の数 3 1、数 3 2 で平均色を計算すると共に、サンプルの前処理を行なう。

【0 1 0 9】

10

20

30

40

【数 3 1】

$$x_{DC} = \left(E[x_{c_1}], E[x_{c_2}], E[x_{c_3}] \right)$$

【0 1 1 0】

【数 3 2】

$$x \leftarrow x - E[x_{DC}]$$

【0 1 1 1】

ここでのサンプル数は16384個であるので、同数のデータベクトルのDC成分 x_{DC} が得られるが、計算量が多すぎるので、例えば $4 \times 4 = 16$ の重複のないブロックに色の情報を統合する。つまり、それぞれのブロックの平均色 I_{DCij} は次のように計算できる。

【0 1 1 2】

【数 3 3】

$$I_{DCij} = \frac{1}{N_{block(i,j)}} \sum_{b \in block(i,j)} x_{DCb}$$

【0 1 1 3】

ここで、 $N_{block(i,j)}$ は、 (i, j) 番目のブロック中のサンプルパッチの数である。このとき、ベクトル I_{DC} は全体的な平均色と位置の情報を持っている。以上より、大きさの異なる画像間の色に関する類似度 S_{color} を、次のように定義することができる。

【0 1 1 4】

【数 3 4】

$$S_{color} \stackrel{\text{def}}{=} \frac{1}{N_B} \sum_{i,j} \frac{I_{DCij}^{(1)} \cdot I_{DCij}^{(2)}}{|I_{DCij}^{(1)}| \cdot |I_{DCij}^{(2)}|}$$

【0 1 1 5】

上記式において、 N_B はブロックの数($4 \times 4 = 16$)であり、 $I_{DC}^{(k)}$ は画像 $I^{(k)}$ の平均色を表すベクトルである。基底の類似度 S_{basis} と同様に、色に関する類似度 S_{color} も、正規化された類似度であるため、ユーザーが調整できる加重パラメータ a を使って、類似度 S_{basis} 、 S_{color} を次のように結合することができる。

【0 1 1 6】

【数 3 5】

$$S = aS_{basis} + (1-a)S_{color}, \quad 0 \leq a \leq 1$$

【0 1 1 7】

加重パラメータ a は、操作手段に相当する例えば図1の入力手段5を利用してユーザーが任意に調整変更することができ、もし大きい値を選択して基底の類似度 S_{basis} を強調した場合には、画像のエッジやテクスチャパターンがより重視され、最終的な類似度 S は画像中の物体の位置の移動に対してロバストになる。逆に、加重パラメータ a の値を小さくするに従い、画像中の物体の位置情報を持つ類似度 S_{color} が次第に重視され、最終的な類似度 S は画像中の物体の移動に敏感なものとなる。こうした特性を理解すれば、ユーザーが入力手段5によりデザインパラメータ a を選択するだけで、類似画像の検索にユーザーの好みを簡単に反映させることができ、ユーザーにとって苦痛なパラメータ設定を少なくできる。

【0 1 1 8】

類似度 S_{color} を定義する上で使用される色空間としては、RGB, YIQ, HSVがある。RGBはそれぞれのピクセルを{Red(赤: R), Green(緑: G), Blue(青: B)} [0, 1]³で表わす最も標準的な色空間で、光の3原色に基づき、R, G, Bの

10

20

30

40

50

加法混色によって色を表わす。それぞれの要素の数値が増える毎に色は白に近づいてゆく。通常はそれぞれの要素が1バイトで表わされ、実際の色空間の大きさは $[0, 255]^3$ となる。

【0 1 1 9】

Y I QはNTSCテレビ放送により標準化されている色空間で、輝度情報を持つYと、色差情報を持つI（オレンジ - シアン）、Q（緑 - マゼンダ）の3要素からなる。RGB空間とY I Q空間は線形変換でき、この変換には幾つかのバージョンがあるが、後述するオピニオンテストでは、次の変換式のものを使用した。

【0 1 2 0】

【数 3 6】

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$I = 0.596 * R - 0.275 * G - 0.321 * B$$

$$Q = 0.207 * R - 0.497 * G + 0.290 * B$$

【0 1 2 1】

H S Vは、色の種類をあらわす色相（Hue angle, H）、色の純粋さをあらわす彩度（Saturation, S）、色の明るさをあらわす輝度（Value, V）の3要素からなる色空間である。RGB空間とH S V空間との変換は非線形変換である。あるピクセルの色の要素の最大値と最小値をそれぞれ $MAX = \max(R, G, B)$ 、 $MIN = \min(R, G, B)$ とする。このとき、H S Vの各要素は次のように表わせる。

【0 1 2 2】

【数 3 7】

$$H = H_R, \text{ or } H_G, \text{ or } H_B$$

$$S = (MAX - MIN) / MAX$$

$$V = MAX$$

【0 1 2 3】

但し、上式において、 $\{H_R, H_G, H_B\}$ は次の通りである。

【0 1 2 4】

【数 3 8】

$R = MAX$ のとき：

$$H_R = \left(0 + \frac{G - B}{MAX - MIN} \right) \times 60$$

$G = MAX$ のとき：

$$H_G = \left(2 + \frac{B - R}{MAX - MIN} \right) \times 60$$

$B = MAX$ のとき：

$$H_B = \left(4 + \frac{R - G}{MAX - MIN} \right) \times 60$$

【0 1 2 5】

このように画像の類似性を判断する場合に、画像の基底はエッジやテクスチャを反映するだけのものであるので、別な要素として部分領域に関する平均色を考えるのが望ましい。本実施形態では、仮想グリッドを用いた前処理で得られるパッチの平均値に基づく部分領域色の類似度 S_{color} を定義している。そして、画像基底の類似度 S_{basis} と色の類似度

10

20

30

40

50

S_{color} との重み付け加算を行ない、その結果を総合の類似度 S として判断している。

【0126】

その他に本実施形態では、クエリ画像21や対比画像31をフィルタリング処理するフィルタ手段（図示せず）を、処理装置2の内部に備えてもよい。ここでの「フィルタリング」とは、 3×3 の平均値フィルタのことを云う。フィルタ手段が行なうフィルタリング処理によって、平均化による画像ノイズの除去が可能になる。

【実施例】

【0127】

次に、上述した類似画像検索の方法および装置に関し、その性能評価を行なったオピニオンテストの結果を以下に記述する。ここでは類似度として、上記実施形態におけるPCA基底による方法と、ICA基底による方法の他に、従来法として色ヒストグラム法と、独立スペクトル表現法を比較対象のためにそれぞれ採り上げ、色に関する類似度 S_{color} を定義する色空間として、RGB, YIQ, HSVをそれぞれ採り上げる。さらに、フィルタリング処理を行なった場合と、行なわない場合についてもテストを行なう。こうして、類似度、色空間、フィルタリング処理についてそれぞれ組み合わせを行ない、合計 $4 \times 3 \times 2 = 24$ 種類の類似画像検索システムを構築した。

【0128】

なお、ここで使用するPCA基底法は、総合類似度 S が上記数35で定義され、基底の類似度 S_{basis} として数9の U^{PCA} を用いている。またICA基底法では、総合類似度 S が同様に上記数35で定義され、基底の類似度 S_{basis} として数20の U^{ICA} を用いている。また色空間に関し、RGB色空間では、それぞれのピクセルが $x = [R, G, B]^T$ によって表現される3次元のベクトルとなり、YIQ色空間では、RGB空間を線形変換した数36によって表現される3次元のベクトルとなり、HSV色空間では、RGB空間を非線形変換した数37、数38によって表現される3次元のベクトルとなる。さらにフィルタリング処理により、例えば 3×3 の平均値フィルタが画像に施される。

【0129】

一例として、検索対象となる画像集合は、ワシントン大学が公開しているgroundtruthのデータベースを使用する。このデータベースには1100枚を超える多様な自然画像が含まれているが、予め上述のような手法でPCA基底およびICA基底を抽出すると共に、数31に基づきデータベクトルのDC成分 x_{DC} を算出しておく。ユーザーは類似画像検索システムに対して1枚のクエリ画像 I_{query} を提示する。システムが当該クエリ画像 I_{query} を取込むと、類似度に従ってデータベース内の画像を順位付けし、類似度の高いもののうち幾つかをユーザーに提示する。このときシステムの性能は、ユーザーが「似ている」と判断できる画像は上位に含まれているかによって決められる。

【0130】

図9は、PCA基底法、HSV色空間を採用し、フィルタリング処理を行なわない場合（{PCA基底法, HSV, フィルタなし}：以下、{ }で記述する）のシステムを用いた実際の類似画像検索の結果である。ここでは、川を人が渡ろうとしているクエリ画像に対して、システムが類似度の高い上位5位までの画像を検索結果として提示している。また、ここでの2つの類似度 S_{basis} , S_{color} を結合する要素 a は、 $a = 0.3$ としている。システムが提示した検索結果の画像は、渡河の画像であるか、似たようなトポロジーを持つ地形の画像であることがわかる。図9に示す検索結果の上位3件は、 S_{basis} , S_{color} 共に高い類似度となっている。したがって、加重パラメータ a をどのように設定しても、当該画像が上位3件にあらわれる。この検索例では、類似画像として報告された各画像は、縦横比の違いを上手く吸収しており、類似画像検索システムとして望ましい性質を十分に満足している。

【0131】

次に上記24種類の類似画像検索システムにおけるオピニオンテストの内容について、より詳しく説明する。被験者の各人に課するオピニオンテストの内容は、以下の通りである。

- (a) 本テストは、画像の類似度に関して各人の判断を含んだものでなければならない。
- (b) あまりに過剰に個人の嗜好が反映されてはいけない。
- (c) 上記(a), (b)を満たした上で、被験者は1100枚以上の画像を見なければならぬため、その負担を軽くするために極力簡単なものでなければならない。

【0132】

オピニオンテストでは、次のステップに従って、上記(a)~(c)の性質を満たすように注意深くデザインされている。

ステップ1: データベース中の全1100枚の画像から、各被験者がクエリ画像と対比画像のペア ($\{I_{query}, I_{target}\}$) を100組選び出す。

ステップ2: ステップ1で選択された100組のペアからランダムに50組のペアを選び出し、トレーニングデータとする。残りの50組は、性能を評価するためのテストデータとする。

ステップ3: トレーニングデータを使用し、各被験者にとって検索成功率が最大となる類似度結合加重 a を0.01刻みで求める。ここで検索成功率とは、クエリ画像 I_{query} がデータベース中の画像の上位1%に順位付けられることとする。これは、PCA基底またはICA基底を使う検索システムにとって重要なステップである。

ステップ4: ステップ3で得られた加重 a とテストデータを用い、検索成功率を調べる。

ステップ5: 以上のステップ2~ステップ4の手順を1セットとし、これを100回繰り返して平均成功率を求める。

【0133】

このようなオピニオンテストを20代の男女10人に対し行なった。なお、上記各ステップにおける各数字は一例であり、適宜変更可能である。

【0134】

図10と図11は、上記オピニオンテストの結果を示すグラフである。図10の棒グラフは、PCA基底法, ICA基底法, 色ヒストグラム法のそれぞれについて、その検索成功率を示したものである。グラフは横軸に沿って、色空間とフィルタリングで区別した6種類の手法の集団をあらわしている。各集団において、左側点線の棒は色ヒストグラム法, 中央破線の棒はPCA基底法, 右側実線の棒はICA基底法によるもので、RGB, YIQ, HSVの各色空間について、フィルタなしとフィルタ有りの場合の検索成功率がそれぞれ示されている。また、縦軸は10人の被験者の検索成功率の平均をとったものである。

【0135】

同様に図11は、PCA基底法, ICA基底法, 独立スペクトル表現法(ISR)のそれぞれについて、その検索成功率を示したもので、左側点線の棒は独立スペクトル表現法, 中央破線の棒はPCA基底法, 右側実線の棒はICA基底法に対応している。これらの各図から、次のことが結論できる。

【0136】

性能については、{ICA基底法, HSV色空間, フィルタなし}の組み合わせが、最も優れた検索成功率(85.3%)となった。次いで、{ICA基底法, RGBまたはYIQまたはHSV色空間, フィルタ有り}と、{PCA基底法, HSV色空間, フィルタ有りまたはフィルタなし}の組み合わせが、検索成功率の高いものとして続いている。

【0137】

ICA基底法は、同じ組み合わせ条件(色空間やフィルタ有無)でのPCA基底法に比べて、検索成功率が常に高い。その理由は、PCA基底が直交基底であるのに対して、ICA基底は非直交基底であるため、より画像情報を良好に反映できるためと思われる。

【0138】

PCA基底法やICA基底法を用いる場合、RGB色空間とYIQ色空間とでは性能の差が殆ど見られない。これは前記数36にあるように、RGB色空間とYIQ色空間は線形変換しただけのものであるため、基底を線形に結合するというモデルから得られるPCA基底法やICA基底法では、差が出にくくなるからと思われる。これに対して、RGB

10

20

30

40

50

色空間とHSV色空間の変換は、人の色感に近づけるために非線形変換となっており、この効果がHSV色空間における優位差となって現れているものと考えられる。

【0139】

本実施形態で提案したPCA基底法やICA基底法に比べて、従来の色ヒストグラム法や独立スペクトル表現法は、検索成功率がはるかに低い。さらに、独立スペクトル表現法は色スペクトル法よりも成績が悪くなっている。

【0140】

以上より、{ICA基底法, HSV色空間, フィルタなし}の組み合わせがシステムとして最も優れているといえる。この場合、加重パラメータ a の最適値を調べてみると、0.25~0.35の範囲であった。本システムをユーザーが初めて使う場合には、加重パラメータ a の初期値として、0.3を設定するような処理装置2の構成であることが好ましい。

10

【0141】

上記オピニオンテストの結果を踏まえ、最適な色空間の選択について考察する。HSV色空間を採用した類似度 S_{color} が最も優れているという図10の結果は、以下の詳細なデータによって裏付けられる。

【0142】

表1は、加重パラメータ $a = 0$ とした場合の各被験者の平均検索成功率を示したものである。つまりここでの最終的な類似度 S は、基底の類似度 S_{basis} に関する情報を利用していない。この表から次のことがわかる。

【0143】

20

【表1】

	RGB		YIQ		HSV	
被験者	フィルタなし	フィルタあり	フィルタなし	フィルタあり	フィルタなし	フィルタあり
1	67.0	67.3	68.1	68.5	64.1	63.7
2	70.0	70.6	68.8	68.5	62.4	61.5
3	77.9	76.7	75.7	73.4	67.1	67.2
4	75.7	75.1	71.9	72.1	71.3	73.6
5	71.5	72.4	71.2	71.4	73.0	73.8
6	74.6	74.3	72.8	73.3	77.4	77.3
7	78.3	78.8	77.2	77.9	78.6	80.2
8	83.3	82.9	81.8	82.6	85.3	85.7
9	79.5	79.2	78.0	77.6	75.7	77.8
10	81.2	80.8	78.4	78.5	81.4	82.9
平均	75.9	75.8	74.4	74.4	73.6	74.4

30

【0144】

40

最もよい検索成功率となる色空間は、個人によって異なる。これは、次のような理由による。被験者の選び出す画像の組は、その被験者の好みや心理状態を反映するが、この画像の組の距離は色空間に依存して決まる。なぜならば、色空間ごとに色の分離性や凝集性などに違いがあり、画像間の色分布の僅かなずれが大きく影響する色空間もあれば、そうでない色空間もあるためである。また、違う色空間であっても、被験者全員の平均値は殆ど変わらない。これは成績の良い色空間が個人毎に異なり、突出した色空間が出ないためと考えられる。さらに、フィルタ手段の有無による差は殆ど見られない。これは、使用したフィルタが平均値フィルタであるため、平均色と共に色空間による類似度 S_{color} には影響が殆どないためである。以上のことから、表1における類似度 S_{color} による検索成功率からは、色空間の選択に関して何らかの結論を出す必要はない。

50

【 0 1 4 5 】

表 2 は、ICA 基底法を採用し、加重パラメータ $a = 1$ とした場合の各被験者の平均検索成功率を示したものである。つまりここでの最終的な類似度 S は、色空間による類似度 S_{color} に関する情報を利用していない。この表から次のような傾向をみることができる。

【 0 1 4 6 】

【表 2】

被験者	RGB		YIQ		HSV	
	フィルタなし	フィルタあり	フィルタなし	フィルタあり	フィルタなし	フィルタあり
1	52.3	50.8	52.8	51.2	67.7	54.9
2	53.3	61.7	56.9	58.7	67.2	64.1
3	63.5	69.7	61.6	65.5	71.8	59.0
4	56.2	69.8	62.9	67.5	78.3	70.3
5	64.0	67.3	56.5	73.6	72.5	61.9
6	65.2	72.3	63.8	71.2	76.9	68.8
7	68.9	74.3	68.2	74.1	74.4	67.9
8	61.0	75.4	63.9	71.4	77.8	66.2
9	62.6	75.2	66.7	72.1	77.7	67.7
10	60.4	68.2	64.0	69.6	68.8	62.6
平均	60.7	68.5	61.7	67.5	73.3	64.3

10

20

【 0 1 4 7 】

殆どの被験者にとって、{ HSV 色空間, フィルタなし } の組み合わせが最も高い検索成功率を示し、被験者全員の検索成功率の平均値も、他のものと比べてはるかに優れている。したがって、表 2 における基底の類似度 S_{basis} による検索成功率からは、{ HSV 色空間, フィルタなし } のシステムを構築するのが最も好ましいといえる。これは、図 1

30

【 0 1 4 8 】

次に、最適なフィルタ手段の選択について考察する。源画像からデータを抽出する場合は、偏りのある分布が望ましい。それは、類似度 S_{basis} の算出に際して、基底どうしの方向を比較しているが、分布に偏りがあるほど推定される基底の信頼性が増すからである。したがって、局所的でない偏りを適切に作り出せるフィルタリング処理を行なうのが望ましい。

【 0 1 4 9 】

ここでは、移動平均値フィルタと移動中央値フィルタについて、その検索成功率をテストした。移動平均値フィルタは、フィルタの中心をピクセル値の加重平均で置き換えるフィルタであり、移動中央値フィルタは、フィルタの中心をピクセル値の中央値で置き換えるフィルタである。中央値フィルタは、平均値フィルタに比べてエッジを保存し易いという性質をもつ。図 12 の棒グラフは、フィルタなし, 移動中央値フィルタ, 移動平均値フィルタのそれぞれについて、その検索成功率を示したものである。このような平滑化を目的としたフィルタリング処理は、平均値をとる類似度 S_{color} には殆ど影響を及ぼさないで、ここでは類似度 S_{basis} のみの検索成功率について示している。横軸は、色空間で区別した 3 種類の集団をあらわしており、それぞれの集団の中で、左側点線の棒はフィルタなし, 中央破線の棒は移動中央値フィルタ, 右側実線の棒は移動平均値フィルタの場合を示している。また、縦軸は検索成功率をあらわしている。

40

【 0 1 5 0 】

50

図12のグラフより、特にRGB色空間やYIQ色空間の場合に、平均値フィルタを施したときの結果がよくなっている。これは中央値フィルタに比べて平均値フィルタの方がぼかし効果が高く、分布が偏るのが理由であると思われる。また、フィルタの大きさを3×3より大きくすると、図13に示すように崩れた基底が得られることがあり、検索成功率が悪化する。これはフィルタリングの効果が強すぎて、分布が過度に偏ったためであると考えられる。なお、図13(a)はフィルタなしの場合、(b)は5×5の移動平均値フィルタを施した場合の基底を示している。

【0151】

次に、フィルタの有無と各色空間との関連性について、図14および図15に示すグラフに基づき説明する。これらの各図において、フィルタリングによる局所的でない分布の増加と、色空間の変換による分布の変化を可視化して提示する。本実施形態における新規な方法は、前述のようにクエリ画像21と対比画像31の各基底の方向を比較するものなので、信頼性の高い基底を推定できる偏りのある分布が望まれる。ここでは、フィルタリングと色空間変換の選択によって、そのような分布の変化が起きていることを確認する。可視化のために、同じフルカラー画像から各色空間に変換したサンプルを取り出し、その次元を192次元から2次元にまで縮約した。このときのデータは白色化されている。

【0152】

図14は、フィルタをかけない場合の各色空間のサンプルの分布である。同図(a)はRGB色空間、(b)はYIQ色空間、(c)はHSV色空間におけるデータの分布と、その分布から得られる独立成分をあらわしている。図14(a)に示すRGB色空間では、データの分布が空間上に均等に広がっており、正しい基底の推定は困難である。同様に、図14(b)に示すYIQ色空間も、RGB色空間の場合と殆ど分布が変わらない。その理由は、RGB色空間とYIQ色空間が線形変換の関係であるためである。一方、図14(c)に示すHSV色空間では、RGB色空間とHSV色空間が非線形変換の関係であるため、他の色空間とは形状はかなり異なり、偏りの強いサンプルの分布となっている。このため、独立成分の方向もはっきりとしている。

【0153】

図15は、移動平均値フィルタをかけた場合の各色空間のサンプルの分布である。ここでは違いを明確にするために、5×5の大きさで平均値フィルタをかけている。同図(a)はRGB色空間、(b)はYIQ色空間、(c)はHSV色空間におけるデータの分布と、その分布から得られる独立成分をあらわしている。RGB色空間どうしを比較すると、図15の分布の左右には、図14の分布よりもサンプルの集中している領域がフィルタリング処理によって存在するため、独立成分の推定はより信頼性の高いものとなる。しかし、HSV色空間で比較した場合は、RGB色空間とは逆の結果となっている。すなわち、フィルタリングによって全体的に偏りが強くなり、2つの独立成分のうち一つが弱まっている。このことは、画像の正確な特徴抽出が難しくなることを意味する。以上の考察によって、図10や図11に示すように、{ICA基底法, HSV色空間, フィルタなし}の組み合わせが、最も優れたシステムであるという結論が裏付けられた。

【0154】

本実施形態における方法では、数8に示すようなサンプルデータの白色化と同時に、次元の縮約を行なっている。次元を縮約すると計算時間を削減できるが、それにより持っている情報も少なくなるため、計算時間と情報量のトレードオフとなる。ここでは、性能に影響を与えずに、どこまで次元を縮約できるかをテストする。

【0155】

$n \times n$ の行列の固有値を $\lambda_1, \dots, \lambda_n$ としたときに第 $m < n$ 主成分までの累積寄与率は、次の数39にてあらわせる。累積寄与率は、元の情報をどれだけ保持しているのかという指標になる。

【0156】

10

20

30

40

【数 3 9】

$$\sum_i^m \lambda_i / \sum_i^n \lambda_i$$

【0 1 5 7】

図 1 6 は、縮約する次元と累積寄与率との関係を示している。この寄与率は、データベース中の画像の寄与率を求め、その平均を採ったものである。横軸はサンプルの次元数で、縦軸は平均の累積寄与率である。このグラフに示すように、サンプルの次元数が 30 よりも低くなると、累積寄与率が急激に減少する。また、次の表は幾つかの次元を抜き出したものである。

【0 1 5 8】

【表 3】

次元の縮約率	サンプルの次元	寄与率 %
1/12	16	70.0
1/6	32	83.0
1/4	48	89.2
1/3	64	92.9
1/2	96	96.9
2/3	128	98.8

10

20

【0 1 5 9】

図 1 7 は、フィルタありの場合と、フィルタなしの場合について、縮約する次元と検索成功率との関係を示したもので、この図では、元々 192 あったサンプルの次元を、1 / 12 (16 次元)、1 / 6 (32 次元)、1 / 4 (48 次元)、1 / 3 (64 次元)、1 / 2 (96 次元)、2 / 3 (128 次元) まで縮約した検索成功率をプロットしている。但し、類似度 S_{cor} は次元の縮約の影響を受けないので、ここでは類似度 S_{basis} のみを使っている。さらに、図 1 7 (a) はフィルタをかけない場合、図 1 7 (b) はフィルタをかけた場合をあらわし、それぞれの実線は H S V 色空間、破線は Y I Q 色空間、点線は R G B 色空間を示している。

30

【0 1 6 0】

同図から明らかなように、縮約する次元により多少のばらつきは見られるものの、1 / 4 (48 次元) の縮約率までは、検索成功率がほぼ変わらないことがわかる。一方、サンプルの次元の縮約率を 1 / 12 (16 次元) とすると、検索成功率が明らかに低下する。1 / 6 (32 次元) の縮約率では評価が難しいが、図 1 6 に示す累積寄与率は平均値であるため、画像によってはこれよりも情報量が少なくなり、検索成功率が悪化することが考えられる。そのためサンプルの次元の縮約率は、1 / 4 (48 次元) までとするのが好ましい。

【0 1 6 1】

縮約する次元が変わることで検索成功率にばらつきが見られることの理由としては、得られる基底の方向が変化することが原因の一つと考えられる。例えば、3 次元の場合を例にして考えると、第 3 主成分に平行な直線と交差する 3 本の独立成分があったとすると、この独立成分は第 3 主成分に直交する成分からは重なって 1 つの独立成分として見える。ここで第 3 主成分が無くなったとすると、1 つの独立成分はそのまま同じ方向を向くが、重なってしまったもう一つの独立成分は、異なる方向を探すことになる。これは不安定な基底となるため、検索成功率が悪化する。逆に、次元を減らすことで不安定な基底が減ると、検索成功率は向上する。このような原因により、検索成功率にばらつきが起こると考えられる。とりわけフィルタ手段を備えた場合には、前述のように分布の偏りが起こるため、検索成功率のばらつきが強くあらわれると考えられる。

40

【0 1 6 2】

次に、画像から抽出するサンプルの大きさを変更した時の影響について考察する。画像

50

から抽出するサンプルの大きさは、学習モデルの大きさに相当する。従来の独立スペクトル表現法では、学習モデルが大きくなると学習に必要なサンプル数が増大し、過学習する可能性が高くなる。一方、tightフィッティングの場合は、学習モデルを1つのクラスに対応することが目的であるため、tightフィッティングした学習モデルの領域の占める範囲は、そのモデルの大きさで学習可能な解空間の中で相対的に狭くなる。そのため、サンプルパッチ（小区画）のサイズを大きくすると学習モデル間の分離度が上がり、検索成功率が高くなると考えられる。

【0163】

図18は、サンプルパッチのサイズを変化させた場合の検索成功率について調べたグラフである。ここでは各サンプルパッチのサイズにおいて、当該サンプルの次元を1/3と1/4に縮約している。またこの場合のフィルタリング処理は、検索成功率に影響しないため、ICA基底法でフィルタ無しの場合の類似度 S_{basis} を比較した。図18の横軸は、RGB, YIQ, HSVの各色空間と、縮約した次元との組み合わせで区別し集団であり、縦軸は検索成功率をあらわしている。また、それぞれの集団の中で、左側点線の棒はサンプルパッチの大きさが4, 中央破線の棒はサンプルパッチの大きさが8, 右側実線の棒はサンプルパッチの大きさが12の場合を示している。この図に示すように、色空間と縮約した次元をどのように組み合わせた場合でも、サンプルパッチの大きさが大きいほど検索成功率が高くなっていることがわかる。すなわち、基底を画像にtightフィッティングさせるため、画像から抽出するパッチを大きくするほど画像クラス間の分離度が上昇し、検索成功率がよくなることが判った。さらに、tightフィッティングの場合は過学習が存在しないため、時間とメモリの許す限り大きなサンプルを使うほうがよいと結論できる。

【0164】

以上のように本実施形態では、選択された画像をクエリ画像21として取り込むクエリ画像取込み手段11と、クエリ画像21の特徴を反映した基底をこのクエリ画像21から算出する第1の基底算出手段12と、検索対象となる対比画像31の特徴を反映した基底をこの対比画像から算出する第2の基底算出手段13と、クエリ画像21の基底と対比画像31の基底とを比較することにより、クエリ画像21に対する対比画像31の類似度を算出する類似度算出手段14とを備えている。

【0165】

こうすると、複数の画像から基底を求めて、この基底を各画像にフィルタリングして得た応答を使って類似度を算出するのではなく、1枚のクエリ画像21から得られる基底と、1枚の対比画像31から得られる基底とを直接比較することで、クエリ画像21に対する対比画像31の類似度を算出している。そのため、少ないサンプル数でありながら、クエリ画像21や対比画像31の特徴をよく反映した基底情報により、正しく類似画像を検索することができる。しかも、1枚のクエリ画像21に対して1つの対比画像31の基底集合を学習させていることになるので、過学習や過汎化の問題も回避できる。つまり、過学習や過汎化を引き起こすことなく、画像のもつ固有の情報を少ない冗長度で正確に表現することが可能になる。

【0166】

そしてこれは、選択された画像をクエリ画像21として取り込み、クエリ画像21の特徴を反映した基底をこのクエリ画像21から算出し、検索対象となる対比画像31の特徴を反映した基底をこの対比画像31から算出し、クエリ画像21の基底と対比画像31の基底とを比較することにより、対比画像31の類似度を算出する方法でも達成される。

【0167】

また本実施形態では、(a)内積が最大となるクエリ画像21の基底と対比画像31の基底のペアを探し出し、(b)手順(a)で探し出したペアを除く残りのクエリ画像21の基底と対比画像31の基底に対して、手順(a)と同様で内積が最大となるクエリ画像21の基底と対比画像31の基底のペアを探し出し、求められた内積に加重をかけて加えて行き、(c)クエリ画像21の基底と対比画像31の基底の全てがペアとなるまで、手順(a)と手順(b)を繰り返して、対比画像31の類似度を算出するように、類似度算出手段14を構成して

いる。

【0168】

こうすると、類似度を算出するに際し、クエリ画像21と対比画像31の各基底集合について、その全ての基底のペアの内積を最大化する必要がなくなる。すなわち、ここでは n 個の基底成分に対して n^2 オーダーの計算を行なえばよく、良好な近似手法を取り入れた現実的な類似度の算出を実現できる。また、類似度を算出するに際して基底成分の加重和を用いることにより、基底集合に存在する非決定性を回避できる。

【0169】

また本実施形態では、クエリ画像21の基底と対比画像31の基底とを比較することにより得られた第1の類似度 S_{basis} と、クエリ画像21と対比画像31との間の色に関する第2の類似度 S_{color} とを結合して、対比画像31の最終的な類似度 S を得るように、類似度算出手段14を構成している。

10

【0170】

この場合、対比画像31の最終的な類似度 S が、クエリ画像21および対比画像31の基底に基づく類似度 S_{basis} だけでなく、クエリ画像21や対比画像31がもつ色空間に関わる類似度 S_{color} を加味して判断される。そのため、物体の位置の移動により敏感な類似画像を検索することができる。

【0171】

また本実施形態では、前記対比画像の最終的な類似度 S は、加重パラメータ a を用いることにより、 $S = a S_{basis} + (1 - a) S_{color}$ の式で算出され、加重パラメータ a を0

$a = 1$ の範囲で調整可能にする操作手段としての入力手段5をさらに備えている。すなわちこの場合は、ユーザーが加重パラメータ a の値を任意に調整できるので、類似画像の検索にユーザーの好みを簡単に反映させることができ、ユーザーにとって苦痛なパラメータ設定を少なくできる。

20

【0172】

さらに上記実施形態では、類似画像の検索を行なうたびに、クエリ画像21や対比画像31についてPCA基底またはICA基底を直接算出するようにしているが、クエリ画像21の代わりにPCA基底やICA基底をそのまま与える検索も可能である。また、データベース1中の対比画像31に予めPCA基底またはICA基底をラベル（付加情報）として付け加えてもよい。つまり、基準となる画像（クエリ画像21）の特徴を反映した第1の基底と、検索対象となる画像（対比画像31）の特徴を反映した第2の基底とを比較することにより、前記対比画像の類似度を算出するような方法、またはそれを実現する類似度算出手段14を備えていれば、1枚の基準となる画像の基底と、1枚の検索対象となる画像の基底との直接的な比較により、検索対象となる画像の類似度を算出できる。そのため、この場合も少ないサンプル数でありながら、各画像の特徴をよく反映した基底情報により、正しく類似画像を検索することができる。しかも、1枚の基準となる画像に対して1つの検索対象となる画像の基底集合を学習させていることになるので、過学習や過汎化の問題も回避できる。

30

【0173】

本発明は上記実施形態に限定されるものではなく、本発明の要旨の範囲において種々の変形実施が可能である。

40

【産業上の利用可能性】

【0174】

本発明は類似画像検索という実例に基づいているが、画像のみならず文書や音楽といった種々のファイル検索を、パソコンのデスクトップ上で統合するデスクトップ検索の動きが現在活発となっているため、こうしたデスクトップ検索技術に本発明の方法や装置を適用することは非常に重要となる。さらには今後の発展（例えば離散データの連続値化など）により、本発明の概念を広範囲に適用することが可能になる。

【図面の簡単な説明】

【0175】

50

【図 1】本発明の好ましい実施形態を示す類似画像検索装置のブロック構成図である。

【図 2】本実施形態における類似画像検索方法の処理手順を示すフローチャートである。

【図 3】本実施形態における基底算出の処理手順を示す説明図である。

【図 4】本実施形態において P C A 基底を抽出する処理手順を示すフローチャートである。

。

【図 5】本実施形態において I C A 基底を抽出する処理手順を示すフローチャートである。

。

【図 6】(a) , (b) は源画像を示し、(a ') , (b ') は (a) , (b) の源画像から得られた I C A 基底を示す図である。

【図 7】Gabor - fit法を使った図 6 における I C A 基底 (a ') , (b ') の比較結果を示したグラフである。 10

【図 8】I C A 基底における分布の形態を示す図である。

【図 9】類似画像の検索結果を示す図である。

【図 10】P C A 基底法 , I C A 基底法 , 色ヒストグラム法における検索成功率を示したグラフである。

【図 11】P C A 基底法 , I C A 基底法 , 独立スペクトル表現法における検索成功率を示したグラフである。

【図 12】フィルタなし , 移動中央値フィルタ , 移動平均値フィルタのそれぞれについて、その検索成功率を示したグラフである。

【図 13】フィルタなしの場合と、5 × 5 の移動平均値フィルタを施した場合の基底をそれぞれ示すグラフである。 20

【図 14】フィルタをかけない場合の各色空間のサンプルの分布を示すグラフである。

【図 15】移動平均値フィルタをかけた場合の各色空間のサンプルの分布を示すグラフである。

【図 16】縮約する次元と累積寄与率との関係を示すグラフである。

【図 17】フィルタありの場合と、フィルタなしの場合について、縮約する次元と検索成功率との関係を示すグラフである。

【図 18】抽出するサンプルの大きさと検索成功率との関係を示すグラフである。

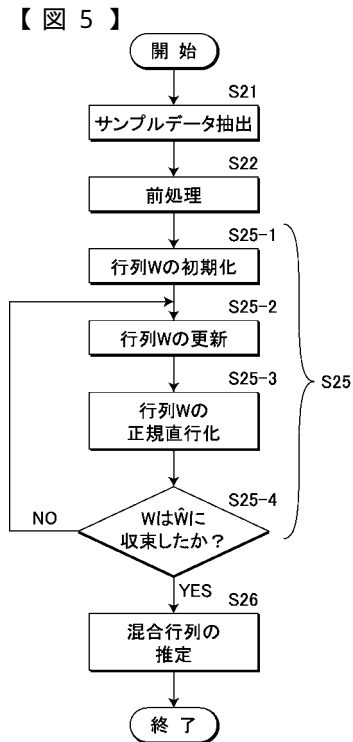
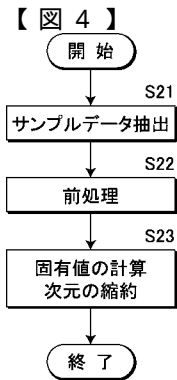
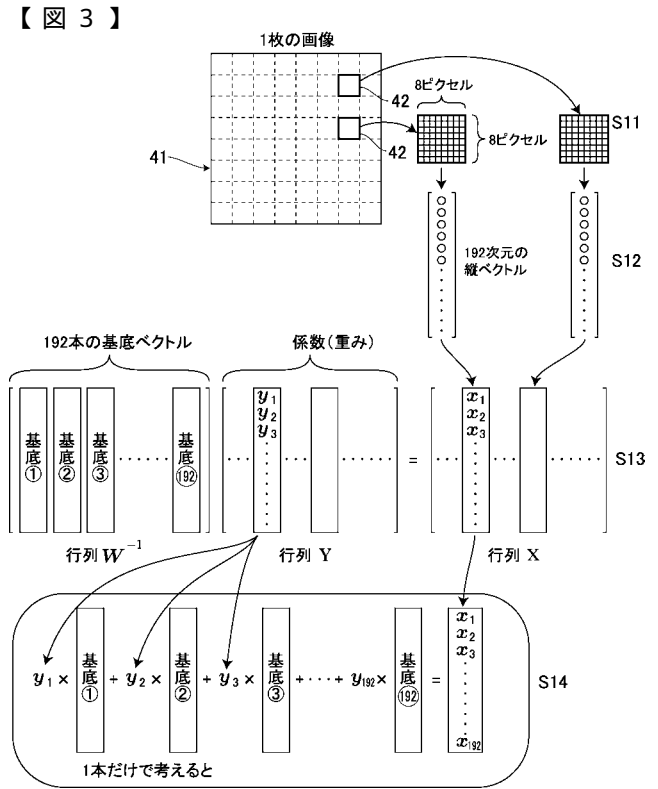
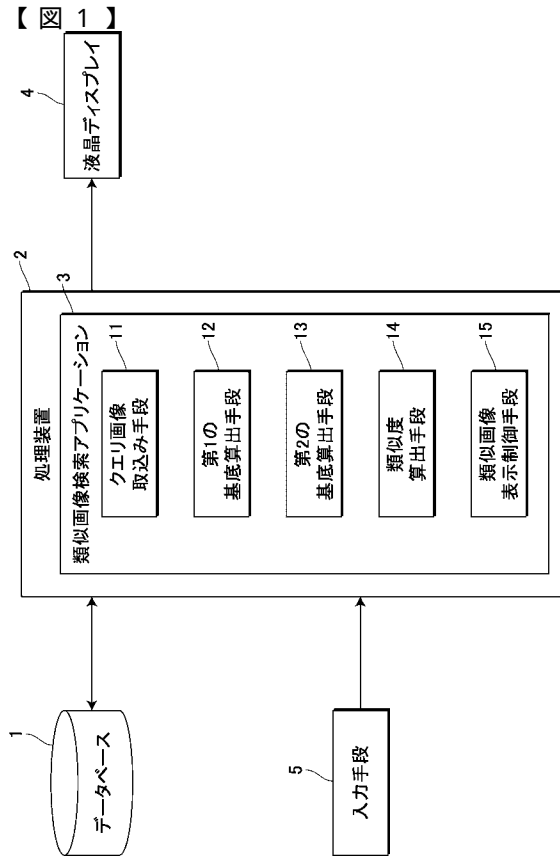
【図 19】従来の類似画像検索方法における処理手順を示すフローチャートである。

【符号の説明】

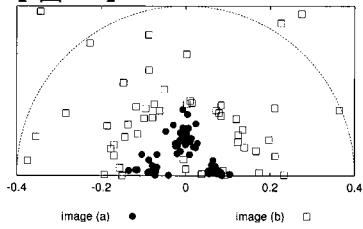
30

【 0 1 7 6 】

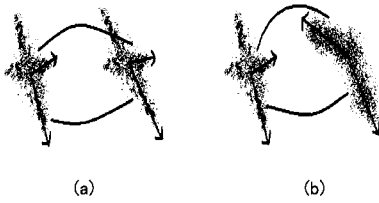
- 5 入力手段 (操作手段)
- 11 クエリ画像取込み手段
- 12 第 1 の基底算出手段
- 13 第 2 の基底算出手段
- 14 類似度算出手段
- 21 クエリ画像 (基準となる画像)
- 31 対比画像 (検索対象となる画像)



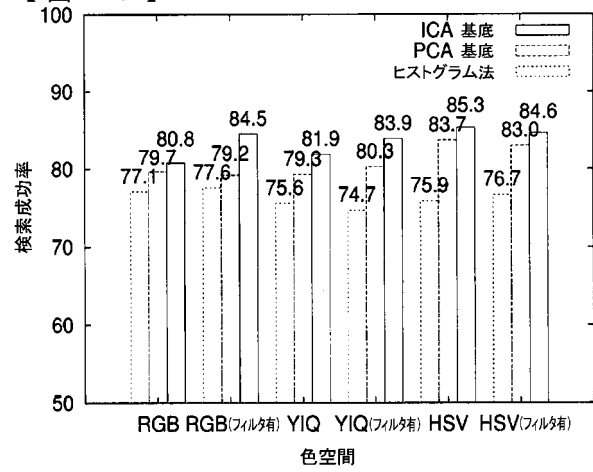
【図 7】



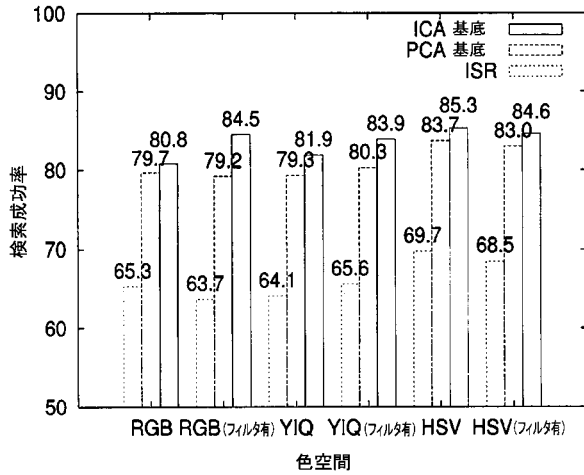
【図 8】



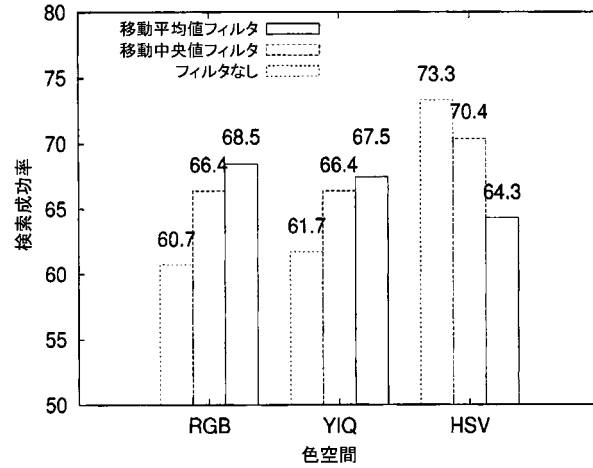
【図 10】



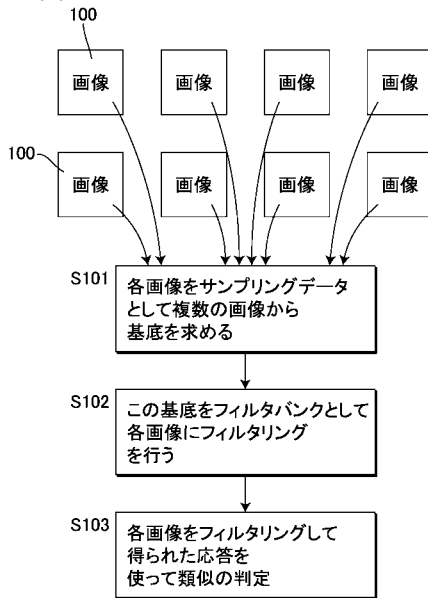
【図 11】



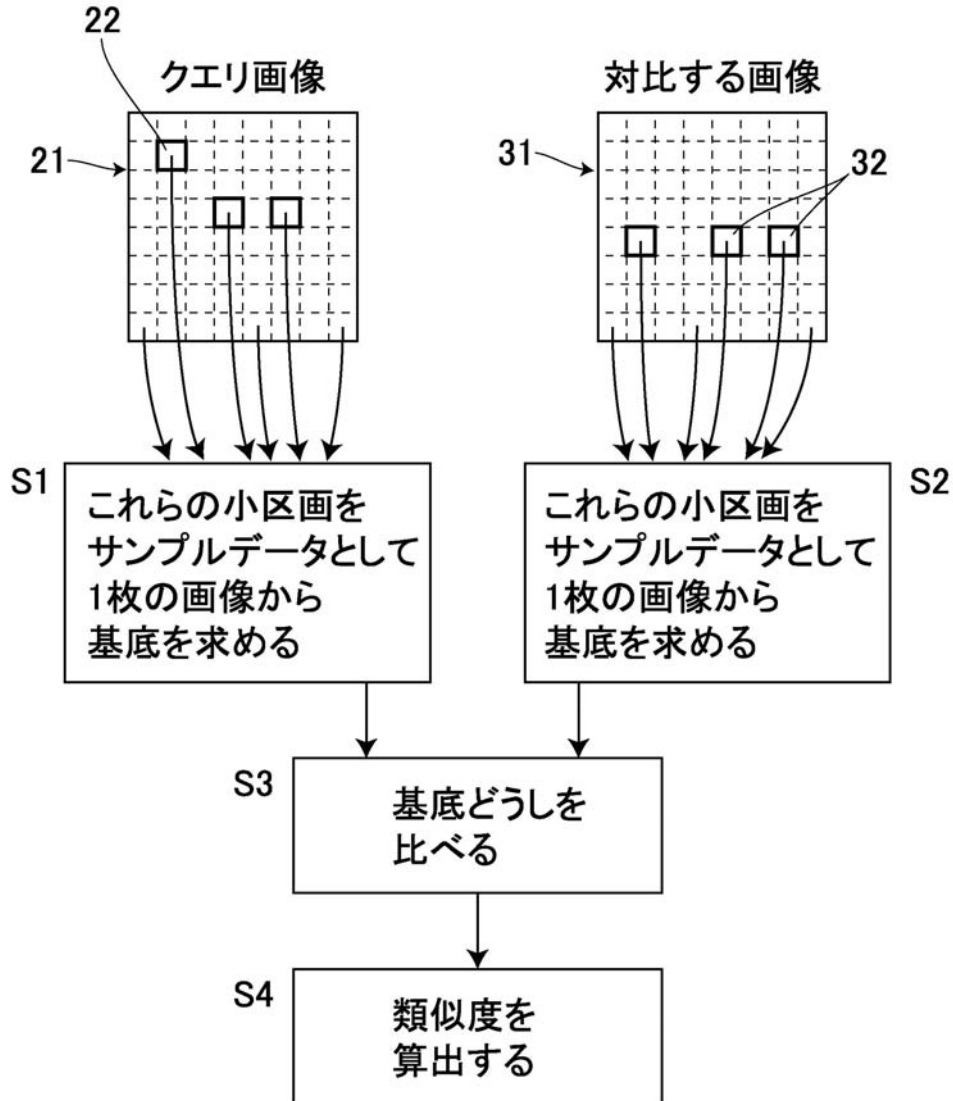
【図 12】



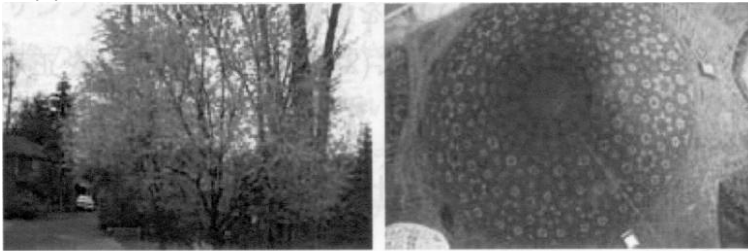
【図 19】



【図 2】

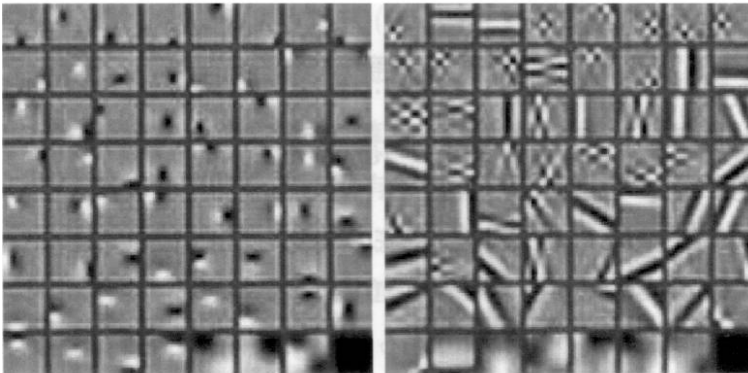


【 図 6 】



(a)

(b)



(a')

(b')

【 図 9 】



クエリ画像

第1位

第2位

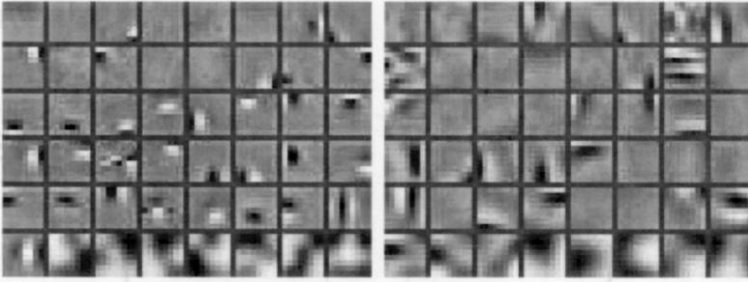


第3位

第4位

第5位

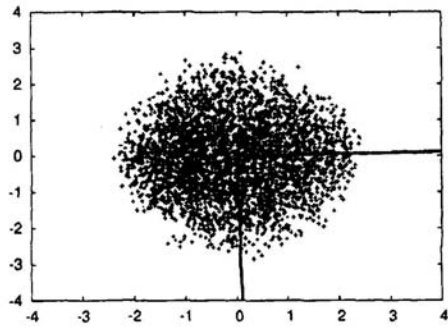
【 図 1 3 】



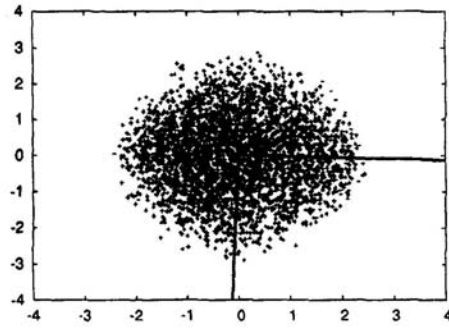
(a)

(b)

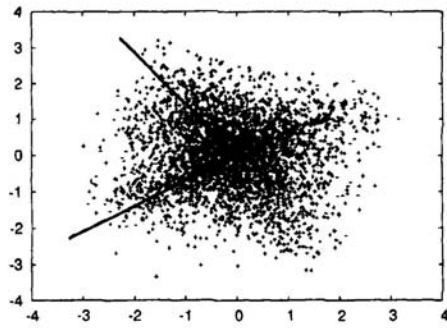
【 図 1 4 】



(a) RGB

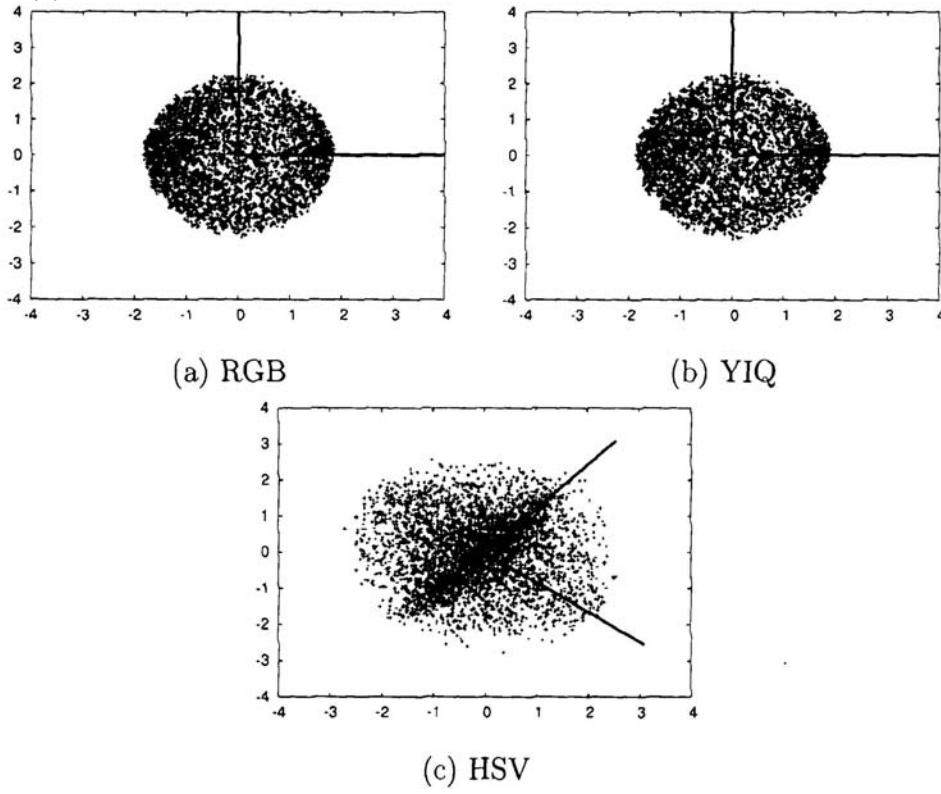


(b) YIQ

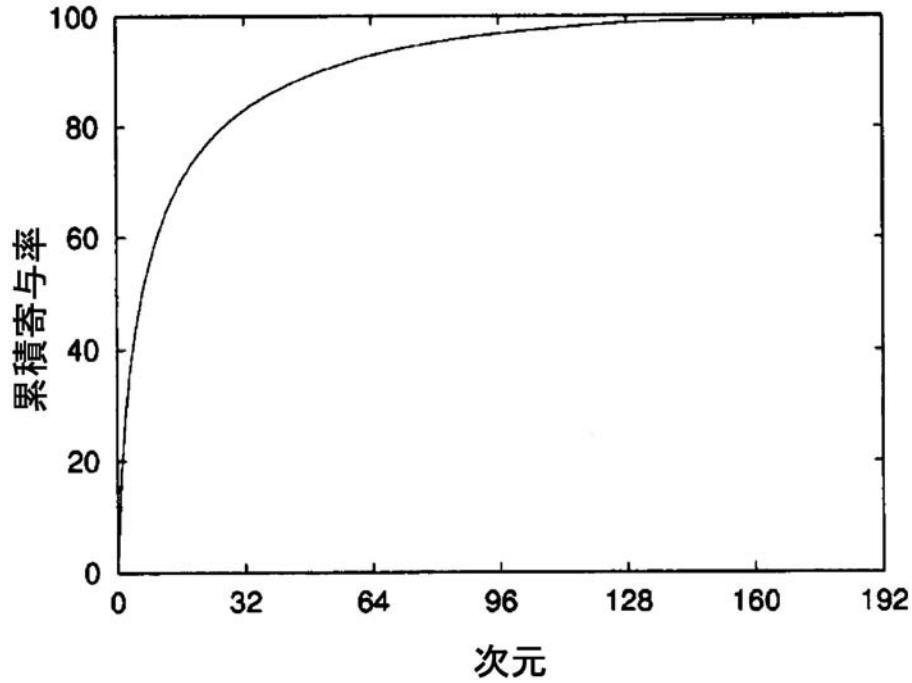


(c) HSV

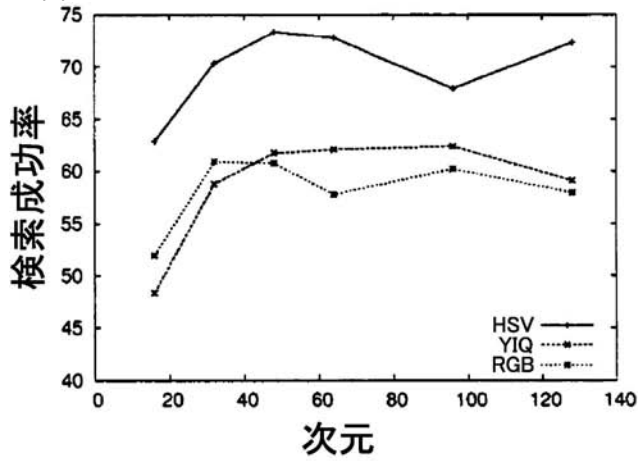
【 図 1 5 】



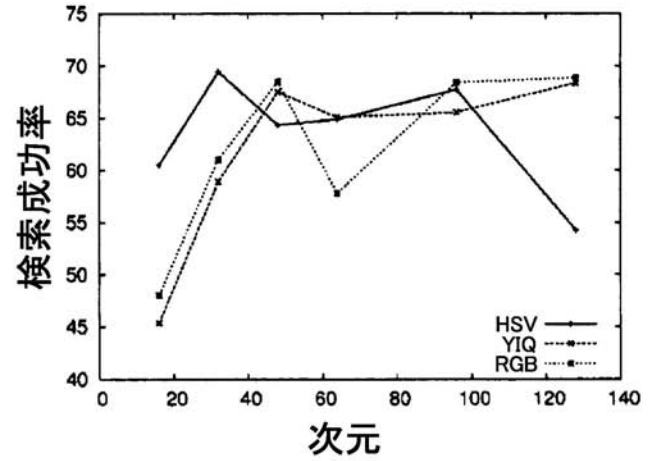
【 図 1 6 】



【図 17】

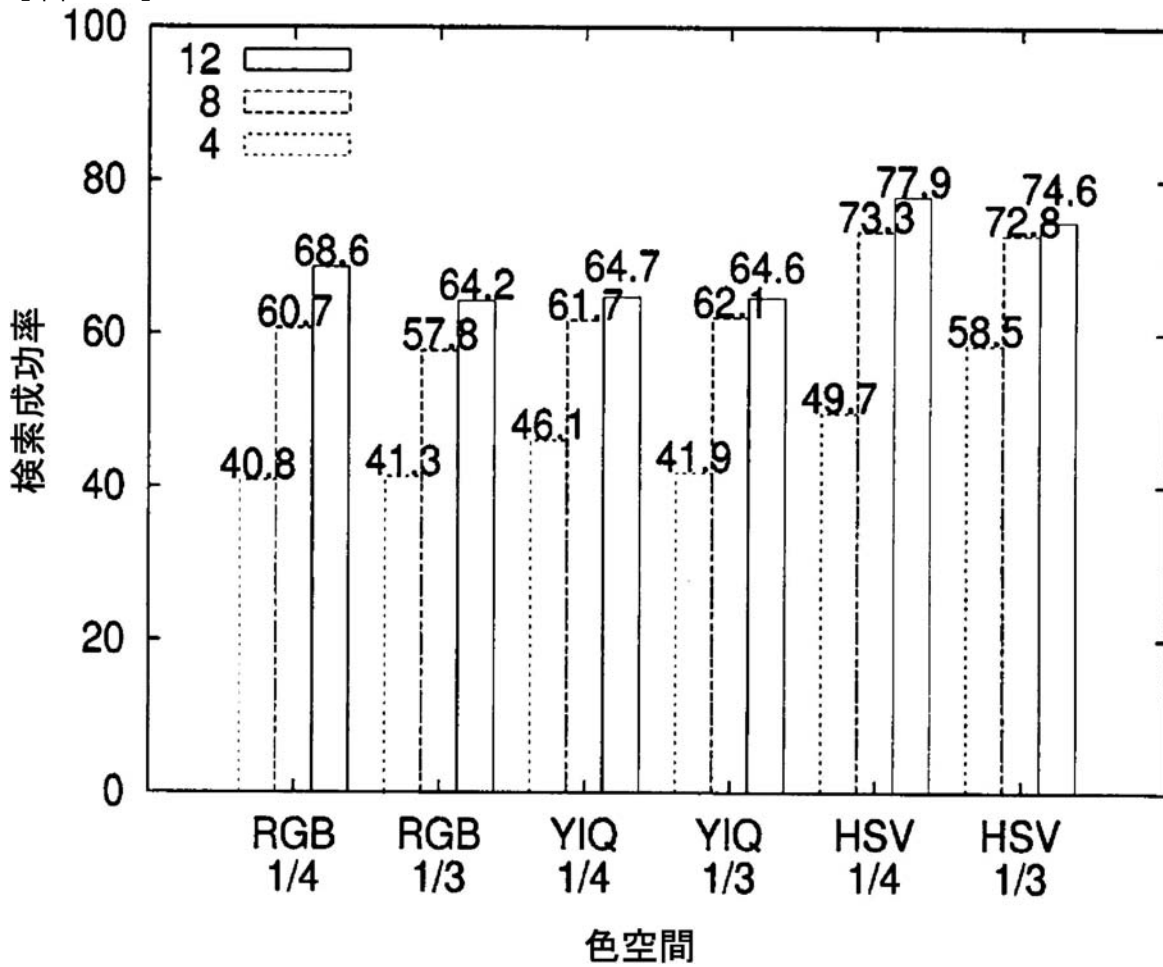


(a) フィルタなし



(b) フィルタあり

【図 18】



フロントページの続き

F ターム(参考) 5B050 AA06 AA09 BA10 BA15 CA07 DA02 DA04 EA09 EA15 EA16
EA18 FA02 FA08 GA08
5B075 ND06 PP10 QM05
5L096 HA08 JA03 JA11 KA04 MA07