

大量画像データベースへの効率的アクセスを可能とする 統合画像アクセスインタフェース

岩崎 雅二郎[†] 両角 清隆^{††}

本稿では画像データベースから所望の画像へ効率的にアクセスする手法を提案する。類似画像検索や画像分類は画像アクセスの代表的な方法であるが、類似画像検索では適切な問合せを与えることが困難であり、また、画像分類はユーザの要求に適合した分類を提供することが困難である。そこで、画像検索と画像分類を融合した GUI を提供することで両者の問題点を解決し、画像特徴量の特性に適した階層分類を提案する。さらに総合的な画像アクセス効率を向上させるために高速な特徴量空間のビュー生成の方法を提案する。そのうえで本手法を画像アクセスシステムとして実装し、その有効性を示す。

Integrated Image Access Interface Enabling Effective Access of Large Image Database

MASAJIRO IWASAKI[†] and KIYOTAKA MOROZUMI^{††}

This paper presents an approach to effectively access desired images in a large image database. Similarity-based image retrieval and image classification are representative methods for accessing images in an image database. With similarity-based image retrieval, it is difficult to specify a proper query image. The image classification cannot achieve the high accuracy classification that users need. We integrate image retrieval and image classification to overcome the above issues and propose hierarchical classification, which is suited to the image features we developed. We also propose a method of fast image feature space visualization. Beyond that, we show the usability of an image access system that integrates the image retrieval and classification method with the visualization method.

1. はじめに

今日、画像を処理する環境が整備されてきたことにより、画像データが一般オフィス、そして家庭にも増加しつつある。このような環境の変化により、画像内容に基づいた大量画像への画像アクセスの研究がさかんに行われている。

画像アクセスの1つとして、画像特徴量に基づいた画像検索の手法がある。問合せ画像から抽出した特徴量、または、直に指定された特徴量と、データベース中の画像の特徴量とを比較し、類似する特徴量を持つ画像を検索結果とする方法が一般的である。しかし、アプリケーションによってはユーザが問合せ画像を有しているとは限らず、特徴量を指定するにしてもユー

ザへの負担が大きい。

一方、画像をあらかじめ自動的に分類しておき、その分類体系からユーザが所望する分類を選択し画像を取得する方法がある。この方法の場合には分類がユーザの要求に適合していなかったり、分類の精度が低いと所望の画像にたどりつけなかったりするという問題がある。

このように従来の手法では実際のアプリケーションに適用するには問題がある。さらに、総合的な画像アクセスの効率化を目指す場合には画像空間の視覚化による画像ブラウズの効率化は欠かせない。そこで本稿では検索や分類だけでなく視覚化といった手法を組み合わせることで所望する画像への効率的なアクセスを実現する画像アクセスの方式を検討し提案する。

2. 従来の研究

画像特徴量を用いた画像検索には様々な研究^{1)~4)}があるが、その代表的な研究として QBIC¹⁾がある。QBICは画像から色、配色、形状といった特徴量を抽

[†] 株式会社リコーソフトウェア研究所
Software Research Center, RICOH Co., Ltd.

^{††} 東北工業大学工業意匠学科
Department of Industrial Design, Tohoku Institute of
Technology

出し検索を行う。QBIC では、問合せ画像を与えて検索する手段のほかに

- 画像の概略を描く
- ユーザが画像の構成色を入力する

といった手段を提供している。しかし、概略画を描くことはユーザにとって負担が大きい。また、構成色を指定するにしても、色に関するユーザの記憶は曖昧なのでユーザが適切に入力できるとはいえないうえに負担も同様に大きい。

問合せとしての画像(問合せ画像)を指定して類似する画像を検索する方式では、問合せ画像がシステム側から提示されていればユーザは画像を選択するだけでよく負担が少ない。そこで、データベース中の画像をランダムに提示する方法があるが、単にランダムに画像を提示するだけでは所望の画像に類似する画像が提示されるとは限らず効率が悪い。

また、実際には1回の問合せでは所望の画像を得られることは少なく、検索結果の画像を用いて何度か問合せを繰り返すことで画像特徴量空間を探索し、所望の画像にたどり着ける。しかし、この操作の繰返しによってローカルミニマムに落ち込み、所望の画像にたどり着けない場合がある。さらによくないことには、データベース空間を把握する手掛かりが検索結果だけであるので、データベース空間を概観することができない。つまり、データベースの構成が把握できないので、ユーザはローカルミニマムに陥ったことすら気づかない。

一方、こうした問題点を解決する手段として画像分類があり、その1つとしてSOMを用いて分類する方法^{5)~8)}がある。また、分類の制約条件をあらかじめ人為的に指定する分類とクラスタリングによる方法を組み合わせた研究⁹⁾、絵画に特化して自動的に分類する研究¹⁰⁾、個々の分類の条件としてテンプレートを利用して自然画像を分類する研究¹¹⁾がある。

これらの分類手法ではユーザに提示した分類を選択するだけでユーザは所望する画像にアクセスできるばかりでなく、データベースの画像空間が分類として表現され、ユーザがデータベースの構成を把握しやすい。その反面ユーザの要求する条件に適合する分類が提示されなければ、ユーザは個々の分類を適切に選択することができない。また、分類の精度が悪く所望する画像が誤って他の分類に属してしまっていると、ユーザがその画像にアクセスすることが不可能になる。

このように一般的な画像アクセス手段である検索や分類手法は、それぞれ問題があり、実際のアプリケーションに適用することが困難である。さらに、分類、

検索の機能が充実していても、やはり最終的には分類や検索結果として、ある程度の数の画像を実際にブラウズする必要がある。したがって、総合的な画像アクセスの効率化を目指すにはいかにユーザが所望する画像を効率良くブラウズできるかが重要となる。提示する画像が少数であれば単にリストアップするだけでよいが、100以上の画像を提示する場合には画像空間の視覚化によるビューを提供する必要がある。

情報の視覚化の研究として三次元空間での類似度をパネモデルにより表現し配置する方法^{12),13)}や、Multidimensional Scaling(MDS)を利用する方法¹⁴⁾がある。さらに、ボトムアップにクラスタリングすることでビューを階層化する方法¹⁵⁾がある。しかし、いずれも計算量が多いので大量の画像をビューへ高速に配置することが困難である。

以上のように画像アクセスの研究は大きく画像検索、画像分類、画像空間視覚化という3つの側面から研究される場合が多いが、それぞれ前述のような問題がある。我々は距離空間インデックスを用いて画像検索を高速化することで検索時の対話性を高める¹⁶⁾とともに総合的な画像アクセス機能の研究¹⁷⁾を行ってきた。本稿では、画像検索、画像分類、画像空間視覚化といった手法を統合した画像アクセス方式を提案する。

第1に検索と分類の機能を統合したGUIを提供する。つまり、分類を代表する画像(代表画像)はデータベースを代表する画像なので、その代表画像をシステムから提示し問合せ画像として利用することにより、画像検索の問題点を解決することが可能である。また、画像の分類がユーザの要望と合わない場合や分類のみで所望の画像にアクセスできない場合に即座に検索に移行できるGUIを提供し、ユーザの視点を指定して検索することで所望の画像にアクセスでき、分類の問題点を回避することが可能である。

第2に検索と分類の機能を明確にし、特徴量の特性に従った最適な分類構成を決定付けたうえでクラスタリングを再帰的に用いた分類木を提供する。これは、画像検索に特化した最適な分類木の構成を提供する機能だけでなく、ユーザの要求に従って対話的に分類木を生成する機能をも提供することで、ユーザの要求に適合した分類木を生成することを可能とする。

第3に検索結果の画像だけでなく分類された画像および代表画像、さらには検索結果の画像を画像特徴量空間上に視覚化するビューを提供することで、画像のブラウズ効率を向上させるだけでなく、画像データベース全体の把握を容易にする。

3. 画像アクセスの前提条件

本稿で対象とする画像データベースの条件として、対象とする画像、画像から抽出する画像特徴量、そして、その画像特徴量を用いた画像の類似度の算出方法について述べる。

3.1 対象画像

画像内容による画像アクセスでは画像全体をアクセス対象とする手法と画像中のオブジェクトをアクセス対象とする手法とがある。画像全体をアクセス対象とする場合には画像全体から特徴量を抽出すればよいが、画像中のオブジェクトを対象とする場合にはオブジェクトの抽出処理が必要となる。

人が画像を手がかりに検索する場合に画像中のオブジェクトを検索したい場合は少なからずある。しかし、人が検索したいと思うオブジェクトの単位は多種多様であり、ある人は植物の花だけに着目して検索したいし、ある人は葉だけで検索したく、また別の人は植物の全体で検索したいかもしれない。このように、人により、また、目的により検索オブジェクトの単位が変動する場合に、それに対応してすべてのオブジェクトを抽出することは現実的ではない。したがって、オブジェクトを対象とする場合ではオブジェクトの単位を限定するためにアプリケーションをかなり限定する必要がある。

一方、写真素材や絵画データベースのように画像全体がアクセス対象となり、画像全体の構図によるアクセスが比較的多いアプリケーションがある。本稿ではこのようなアプリケーションをターゲットとし、基本的に画像全体をアクセス対象とする。

3.2 画像特徴量

本稿では画像から画像特徴量を抽出し、画像特徴量空間中での特徴量間の距離を画像間の類似度とする。画像を n 個のサブ領域に分割し、その各サブ領域の局所特徴量 f_i は画像内のサブ領域 i における領域内の位置不変量であり、画像全体の特徴量は $F = \{f_1, f_2, f_3, \dots, f_n\}$ で表される。局所特徴量を抽出するサブ領域の設定方法は様々考えられるが、位置不変量をどのように設定するか、つまりは、位置ずれをどのように吸収したいかであり、対象画像集合が限定された時点で、その画像集合の特性に基づいてサブ領域の形状を決定すればよい。

本稿ではアクセス画像の種別を特に限定しないので、最も汎用性があると考えられる格子状に等分した $n = N \times N$ 個の領域をサブ領域とする。なお、画像サイズが大きい場合の処理速度の低下や画像のアスペ

クト比の相違による処理の煩雑を回避するために、特徴抽出処理では画像全体を 200×200 ピクセル以下（最適な特徴量抽出を行うために個々の特徴量単位に正規化のピクセルサイズが異なる）の正方形画像に正規化したうえで人の類似感覚に即している以下の特徴量の抽出を行っている。

- (1) 色：色のヒストグラムを特徴量とし、画像全体の色合いを示す（サブ領域： 1×1 ）。
- (2) 色分布：平均色を特徴量とし、色に基づく構図を示す（サブ領域： 10×10 ）。
- (3) エッジ分布：4方向（水平、垂直および斜め2方向）に対するエッジのピクセル数を特徴量とし、輪郭線に基づく（色に基づかない）構図を示す（サブ領域： 5×5 ）。
- (4) テクスチャ分布：領域の同時生起行列¹⁸⁾により求めた統計量を特徴量とし、質感に基づく構図を示す（サブ領域： 5×5 ）。

カラー画像を対象とした場合、一般に人は色を重視する傾向があるので、色に関してはサブ領域を 1×1 と 10×10 とした場合の2つを用意した。これは後述する分類に必要な特徴量空間での精度にも関連する。また、色では表されないテクスチャ（質感）を特徴量として採用した。さらに、単一属性で構成される各領域の境界に人は鋭敏に反応する視覚的特性があるので、エッジ分布も特徴量として採用した。このように、これらの特徴量は人が画像を一見したときの印象に基づいた特徴量であり、その特徴量空間は人間の類似感覚に即した空間となる。

3.3 画像の類似度

前述の画像特徴量は画像特徴量空間をなし、その画像特徴量空間での画像特徴量間の距離が画像の類似度（相違度）となる。サブ領域の特徴量により構成される上記4種の特徴量における同一種の特徴量 F_a, F_b 間の距離は

$$D(F_a, F_b) = \frac{1}{n} \sum_{i=1}^n d(f_{ai}, f_{bi}) \quad (1)$$

と表される。なお、 $d(f_{ai}, f_{bi})$ はサブ領域の特徴量間の距離であり、 i はサブ領域を示す識別子である。

さらに、 $j = 1, \dots, 4$ を4種の特徴量の識別子とし、画像 I_a, I_b から抽出した4種の特徴量を F_{aj}, F_{bj} 、4種の特徴量間の距離を $D_j()$ とすると、画像 I_a, I_b 間の距離は

$$D(I_a, I_b) = \sum_{j=1}^4 w_j D_j(F_{aj}, F_{bj}) \quad (2)$$

と表される． w_j は各特徴量の重み付けであり，特徴量ごとに重み付けを行うことでユーザの観点を反映することができる．本稿ではユーザが対話的に分類木を生成したり，検索を行うときにはこの重み付けはユーザによって指定される．また，システムにより自動的に分類木を生成する場合には後述の基準に従って重み付けが決定される．

4. 画像アクセスにおける検索と分類の統合

画像検索と分類は異なる画像アクセス手段として研究される場合が多いが，データベース中の画像への効率的なアクセスを考えた場合には，その一方のみでユーザの要求を完全に満足することは現状では困難である．しかし，その両者を統合することで，両者の欠点を補完することが可能である．本稿での画像アクセスシステムでは以下に述べる画像検索および画像分類を統合し，画像アクセスの効率化を目指した．

4.1 画像検索

本稿での画像検索では一般的な類似画像検索と同様にデータベースへ画像を登録する時点で画像から画像特徴量が抽出されデータベースに登録される．検索時には与えられた問合せ画像から画像特徴量が抽出され，その特徴量とデータベース中の特徴量の距離を式 (1)，(2) に基づき逐一計算し検索結果として距離（類似度）順に並べられた画像のランキングを出力する．

4.2 画像分類

分類にはカテゴリゼーションとクラスタリングの2種類がある．カテゴリゼーションは人手により事前に設定されたカテゴリに個々のデータを割り振ることであり，クラスタリングはデータ集合よりクラスタを導き出すものである．カテゴリゼーションは人がカテゴリを設定するのでカテゴリが人間の知覚に合致する反面，個々のカテゴリを設定しなければならず，また，対象とするデータ集合に対して適切にカテゴリを設定しなければカテゴリ単位のデータ数の偏りが激しく，画像アクセスの観点から不適切な分類となる．

一方，本画像分類で利用する画像特徴量は人の類似感覚に即して生成されており，つまり特徴量空間での距離が人の類似感覚に即しているため，特徴量空間でのクラスタリングにより人間の感覚に即した分類の自動生成が可能である．したがって，本稿ではクラスタリングの手法を採用する．また，本稿では生成される分類と人が要求する分類との相違を検索や視覚化により補完できることを想定している点から自動分類による精度の低下の問題点を解消できると考える．

前述の4種の画像特徴量からなる特徴量を式 (2) の

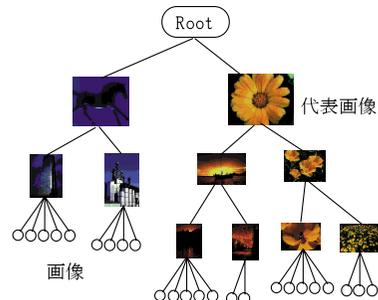


図 1 分類木

Fig. 1 Clustering tree.

距離式を用いてクラスタリングすることでクラスタ（分類）を生成し，さらに，個々のクラスタを再帰的にクラスタリングすることにより階層的なクラスタ構造（分類木）を生成する（図 1）．なお，重み付けは後述の基準により設定する．ユーザはこのようにして生成された分類木のルートから所望する画像と類似する代表画像を順次たどることによって所望する画像に到達することが可能である．

クラスタリングにより再帰的に分類することで分類木を生成すること自体は容易なことであるが，その分類木が画像アクセスを目的にした場合には，つまり，その分類木のルートから各クラスタの代表画像を選択することでユーザが所望する画像を探し出せるためには以下の条件を満足する必要がある．

- (1) 分類木の各クラスタの代表画像はクラスタを構成する画像集合のすべての画像と類似していなければならない．
- (2) クラスタを構成する画像集合のすべての画像はある一観点によって類似していなければならない．

クラスタ中の画像に所望する画像があったとしてもその所望する画像と代表画像が類似していない場合にはユーザはその代表画像を選択することはありえず，その所望の画像を探し出すことができない．したがって，上記の第 1 の条件は画像分類においては必須条件となる．

また，クラスタの代表画像に対して，色の特徴がきわめてよく似ている反面，テクスチャはほとんど似ていない画像と，テクスチャがきわめてよく似ているが色はほとんど似ていない画像がそのクラスタに含まれていたとする．これは第 1 の条件を満足しているが，その 2 つの画像に類似点がない．このようにクラスタ中で互いに類似点がない画像が存在するとユーザはクラスタの分類観点を理解できずにクラスタを適切に選択できない．したがって第 2 の条件も必須条件とい

える。

クラスタの代表画像の特徴量を F_r , クラスタを C , クラスタの半径を R とすると, 第 1 の条件は以下の式で与えられる。

$$D(F_r, F) < R, \forall F \in C \quad (3)$$

つまり, 画像特徴量空間でのクラスタ半径が固定であるクラスタリング手法である NN (Nearest Neighbor) 法が適していることが分かる。また, この手法は処理がシンプルなので比較的高速にクラスタリングを実行することが可能であり, この点からも分類を対話的に行うことを想定している本稿での手法に適している。NN 法により得られたクラスタが分類を構成する各分類となり, 各クラスタの中心となる特徴量が分類の代表画像となる。NN 法での半径 R は第 1 の条件により決定されるが, つまり, 利用する画像特徴量の特性に完全に依存することを意味し, その半径の妥当性は人間の主観評価によってのみ確認される。

本稿では 4 種の画像特徴量を用いているが第 2 の条件は各階層におけるクラスタリングに利用する特徴量は 1 種であることを意味する。すなわち, 式 (2) における重み付け w_j の要素は 1 つを除いてつねに 0 であることを意味する。

4.3 分類と検索の統合

以上のような画像検索と画像分類は, いずれも画像アクセスを目的としているが, それぞれ問題点がある。画像検索では, 問合せ画像がシステム側から提示されていればユーザは画像を選択するだけでよく負担が少ない。そこで, データベース中の画像をランダムに提示する方法があるが, 単にランダムに画像を提示するだけでは所望の画像に類似する画像が提示されるとは限らず効率が悪い。

また, 1 回の問合せでは所望の画像を得られることは少なく, 検索結果の画像を用いて何度か問合せを繰り返すことで画像特徴量空間を探索し, 所望の画像にたどり着ける。しかし, この操作の繰返しによってローカルミニマムに落ち込み, 所望の画像にたどり着けない場合がある。さらによくはないことには, データベース空間を把握する手がかりが検索結果だけであるので, データベース空間を概観することができない。つまり, データベースの構成が把握できないので, ユーザはローカルミニマムに陥ったことすら気づかない。

しかし, 上記の分類木の代表画像はデータベースの代表画像でもあり, 問合せ画像として適切である。また, 分類木はデータベースの構成を体現しており, ユーザは分類木をたどることで所望の画像をアクセスすることができ, ローカルミニマムに陥ることはもち

ろんない。このように分類は検索の問題点を補える。

一方, 静的であれ, 動的にであれ生成された分類が完全にユーザの要求に適切に適合していることはありえず, 所望の画像がデータベースに存在していても, その画像を分類木をたどることで見つけ出せない場合がある。その場合でも, 所望の画像に類似する画像を問合せ画像としてユーザの検索観点を重み付けで指定したうえで検索を行うことで, 所望の画像に到達することが可能である。

このように, 検索や分類を統合することでそれぞれの問題点を補完し, 画像アクセスの効率化を実現することが可能である。

5. 画像特徴量の性質に基づく画像分類木の生成

利用する画像特徴量は人が知覚する画像の類似性に即し, 4 つの観点での特徴量となっており, その特徴量をどのように利用して分類木を生成するかにより, 分類を用いたアクセス効率がまったく異なる。そこで, 本稿での分類と検索の機能や, 画像特徴量の性質を明確化したうえで画像分類木の生成方法を決定付ける。

5.1 画像分類と画像検索の機能定義

一般の検索では条件に合致するか否かであるが, 本画像検索では前述のように類似度によるランキング検索である。一方, 本画像分類では分類の代表画像に類似する画像を 1 つの分類とする方法である。特に NN 法によるクラスタリングであるので, 個々の分類はその代表画像を問合せ画像として検索した結果と言い換えることができる。さらに, 分類木における下位ノードは絞り込み検索に対応する。このように考えると本手法での画像分類と画像検索はきわめて近い機能を提供していることが分かる。そこで本手法における画像分類と画像検索との違いを明確にする。

分類では分類を決定付ける条件が静的に (事前に) 決定されており, 検索では検索条件が動的に検索者により決定される。

本稿での分類を決定付ける条件とは分類の代表画像とクラスタの半径であり, 検索での問合せ画像と検索範囲にあたる。分類を構成する要素が動的に生成されるか静的に生成されるかはシステムの実装に依存し, 検索と分類を区別する基準にはならない。分類を画像検索の範疇で言い換えると「分類はそのデータベース中のデータに対して最適であろう検索条件をあらかじめ検索者に提示している検索方式」となる。

静的に決定できる条件として, 全データ集合から適切な分類条件の決定が可能であり, 検索者が検索条件

を探索するよりも効率的な分類条件をあらかじめ設定できる場合である。したがって、データベースの画像集合に対して適切な分類条件を設定することのみが分類の本質となる。そこで、本稿では画像特徴量の特性を解析したうえで画像特徴量空間における適切な分類条件を決定付ける。

5.2 画像特徴量の性質

検索では特徴量空間中で問合せ画像の近傍を検索するだけなので、特徴量空間中で比較的離れた特徴量間の距離が正しくなくても問題がない。つまり、比較的類似していない画像間の距離が人間の類似感覚を正しく反映していなくても検索ではあまり問題にならない。しかし、大量画像を分類する場合にはクラスタ数を抑えるためにどうしても大分類が必要となり、つまりは、特徴量空間で比較的大きな領域を大分類とする必要がある。したがって、大分類では特徴量空間中で比較的離れた特徴量間の距離の精度が高い必要がある。さらに、分類木の場合には特徴量の持つ情報の包含関係も考慮しなければならず、包含関係に従った分類の階層関係が保たれなければならない。

このことから利用する特徴量の性質を適切に把握する必要があり、その性質を考慮したうえで分類を生成することが重要となる。我々の画像特徴量の特性から以下の制約条件が生ずる。

- エッジ分布とテクスチャ分布は画像の細部の特徴を抽出しており、特徴量の近傍における距離精度は高いが、比較的離れた特徴量間の距離精度は低い。したがって、大分類には向かない。
- 色や色分布は特徴量空間全般において比較的適正な距離精度ではあるが、エッジ分布やテクスチャ分布と比較すると近傍の距離精度が低く、大分類に向く。
- 色の特徴量に位置情報を追加したものが色分布の特徴量といえるので、色と色分布には包含関係がある。色の分類を細分化するために色分布を利用できるが、その逆はありえない。つまり、分類木で色が色分布の上位に位置しなければならない。
- テクスチャ分布はエッジ分布をさらに細分化したものととらえることができ、テクスチャ分布とエッジ分布も包含関係にあり、分類木でエッジ分布はテクスチャ分布の上位に位置しなければならない。

4.2 節で示したように 1 つの階層での分類は 1 種の画像特徴量を用いることと、これらの条件により必然的に分類木は、ルートから階層ごとに色、色分布、エッジ分布、テクスチャ分布という順で特徴量を用いて分類を行うことが最適であると判断できる。しかし、

ユーザの要求によっては別の順が好ましい場合も考えられるので、本稿ではユーザが対話的に分類木を生成しない場合には上記順により分類木を生成するが、ユーザが対話的に分類木を生成する場合には、個々の階層単位にユーザの分類する観点を式 (2) における重み付け w_j および式 (3) におけるクラスタリング R の半径を指定することで分類木を生成することも可能である。

6. 画像特徴量空間の視覚化

分類の各クラスタの画像集合や検索によって得られた画像集合は時として 100 画像を超えることもある。100 を超える画像集合をランダムにリストアップした場合には、その中から所望の画像を選択することはユーザにとってかなり負担となるので、ユーザが画像を、よりブラウズしやすくすることが重要な課題となる。本稿の視覚化では、類似する画像をビュー上で近傍に配置することで所望の画像を効率良く探索可能とすることを目的とし、画像集合の解析を行うことを目的としない。ただし、検索結果の画像集合のビューとしても利用するので動的にビューが生成される必要がある。したがって、対話性を確保するために処理の高速性が要求される反面、高い精度は要求されない。

多次元データにおいてそのデータ間の類似度や関連度に基づき二次元や三次元の空間に視覚化する手法として、物理現象のシミュレーションとしてバネモデルの手法や、また、主成分分析や MDS により次元数を削減する手法があるが、計算量がきわめて多く対話的な処理には向かない。そこで、視覚化の精度は落ちるが類似する画像を表示空間上で近い位置に配置し、かつ、処理の負担が比較的少ない方法を提案する。

まず特徴量の空間の構造を二分木で表現される部分空間に分割する。この二分木上では類似する特徴量は近くに配置され、画像空間上での画像の位置関係を表している。この木構造を画面の表示空間に配置することで仮想的に特徴量の空間を表示空間(相互類似ビュー)に表現することができる。相互類似ビューの生成過程は以下の二分木生成と二次元空間への配置の大きく 2 つのステップに分られる。

(1) 二分木生成

部分空間を 2 つに分割するには、まず以下のよう
にノードの中心点を獲得する。

- (a) 空間内で任意の点 A を選択する。
- (b) 選択した点 A から最遠の点をノードの中心点 C1 とする。
- (c) C1 から最遠の点を 2 番目のノードの中

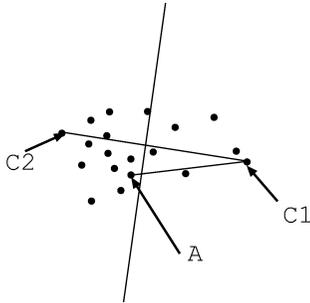


図 2 画像特徴量空間の分割

Fig. 2 Splitting of an image feature space.

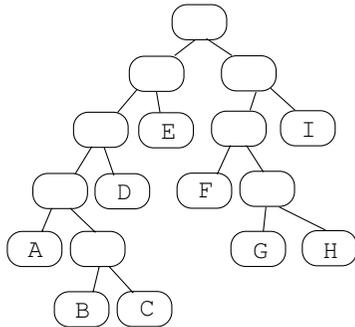


図 3 画像特徴量空間の二分木

Fig. 3 Binary tree of an image feature space.

心点 C2 とする。

このようにして 2 つのノードの中心点 C1 と C2 を得ることができる (図 2) . 次に各点の分配をする。

- (a) 画像集合から任意の点を選択する。選択する点がなくなったら処理を終了する。
- (b) 距離の近い方のクラスターに選択した点を属させる。
- (c) (a) に戻る。

このようにして部分空間を二分することができ、これを再帰的に行うことで二分木 (図 3) を生成することができる。

(2) 二次元空間への配置

得られた二分木を二次元平面に配置する。二分木をルートからたどり、

- (a) 画面を前回の分割と異なる座標軸で等分に領域を分割し部分領域を生成する。初回の分割時は任意の方向で分割する。
- (b) 各部分領域に子ノードを割り振り、子ノードがリーフノードのときにはリーフノードの部分領域の中心に画像を配置する。
- (c) 各子ノードをたどり (a) に戻る。たどる

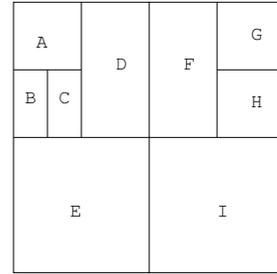


図 4 二分木の表示空間へのマップ

Fig. 4 Mapping a binary tree to a view space.

子ノードがなくなったときに処理を終了する。

このように再帰的に処理することで木構造のすべてのリーフノードが表示空間に配置される。

図 4 に図 3 の二分木を配置した様子を示す。図 4 のアルファベット部分に画像が配置される。

このようにして生成された相互類似ビューは二分木生成処理時の最初の点 A の選択により画像の配置は大きく変化することになるが、本稿での視覚化の目的が前述のように「類似する画像をビュー上で近傍に配置する」ことであり、点 A をどのように選択したとしてもこの目的には合致するので点 A は任意の点で構わない。

また、以下のビューも補足的に追加することで、ユーザの個々の要求に適したビューを選択可能とする。

- 単一類似ビュー

代表画像または問合せ画像と各画像の距離により各画像は二次元に配置される。4 つの特徴量から任意の 2 つの特徴量を水平軸および垂直軸として選択することができる。

- 平均色空間ビュー

画像全体の平均色により水平軸を色相、垂直軸を明度または彩度とした二次元に各画像が配置される。

- リストビュー

単純に画像をリストアップする。検索結果の場合には類似度順に並ぶ。

7. 画像アクセスシステムの実装

上記の機能を提供する総合的な画像アクセスシステムを試作した。図 5 にそのシステムの GUI を示す。登録した画像データは 2,000 画像のフォト画像である。画面は 4 つのウィンドウから構成される。上部が検索ウィンドウ、左下が分類木ウィンドウ、右中央がビューコントロールウィンドウ、右下がビューウィンドウで

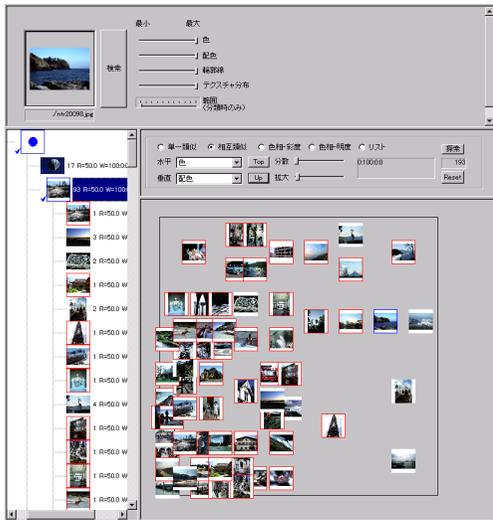


図 5 画像アクセスシステムの画面
Fig. 5 System interface window.

ある．このシステムでは，画像の検索，分類およびブラウズ機能を提供する．

7.1 検索ウィンドウ

検索ウィンドウは問合せ画像を元に画像の検索操作を行うウィンドウである．問合せ画像は左のウィンドウに表示され，検索ボタンによって検索を実行し結果はビューウィンドウに表示される．上位の4つのスライダにより検索時の各特徴の重み付けが可能である．なお，この4つのスライダによる重み付けは分類時にも利用する．新たに分類を行うときにはこのスライダで指定された重み付けにより分類され，最下位に位置するスライダの値はクラスタの半径として利用される．

7.2 分類木ウィンドウ

分類木ウィンドウでは分類木が Windows のエクスプローラと同様にツリー形式で表示される．画像は各クラスタの代表画像である．チェックマークのある分類クラスタはすでに閲覧したことを示す．

分類木はルートから順にアクセスされた時点で下位の1階層分の分類を動的に生成する．各クラスタをクリックされると同時に分類処理を行い，下位のクラスタを1階層分生成する．一度生成されたクラスタを再度クリックした場合にはすでに生成されている下位クラスタが表示される．この分類時にユーザが重み付けを特に指定しない場合は，前述の分類規則に従い分類木が生成されるが，ユーザが自由に重み付けや分類時のクラスタの半径を検索ウィンドウで指定することも可能である．なお，ユーザが指定しない場合のクラスタの半径には本画像特徴量に最適な半径を主観評価よ

り決定した値を用いている．

生成された分類木は二次記憶へのセーブ，ロードが可能であり，ユーザが自分の視点にあった分類木を複数生成しておき，選択的に利用することが可能である．もちろん，データベース提供者が分類木を複数作成しておき，ユーザに提供することも可能である．

分類木ウィンドウで表示される分類木はデータベース全体を表しているため，ユーザは分類木によりデータベースの空間を容易に概観することが可能となる．また，個々の分類階層においてビューウィンドウにより画像空間が視覚化されるので，その効果は大きい．

7.3 ビューコントロールウィンドウ

ビューコントロールウィンドウではビューウィンドウに表示するビューの選択や，ビューの拡大縮小といったビューウィンドウの操作が可能である．

7.4 ビューウィンドウ

ビューウィンドウは分類木ビューで選択された分類木のクラスタの画像集合や検索結果の画像集合を二次元に表示する．分類木のクラスタが表示されている場合には，各画像は下位のクラスタの代表画像であり，その画像をクリックするとそのクラスタの画像集合が表示される．各ビューにより画面上に表示した様子を図6~9に示す．図の例では前述の2,000画像を分類し，分類木の第1階層では125のクラスタが生成されており，その125の代表画像が表示されている．ビューウィンドウに表示された任意の画像を問合せ画像として画像検索に移行できるので，分類木の探索に行き詰まった場合に有効である．

図6に相互類似ビューを示す．画像間の類似性を考慮して配置しているので類似する画像が他のビューと比較して近くに配置されており，ユーザの画像選択が容易である．

図7に単一類似ビューを示す．水平軸および垂直軸はそれぞれビューコントロールウィンドウで指定された特徴量である．この例では，水平軸が色，垂直軸が色分布である．左下の気球の画像からの色特徴量の距離と色分布特徴量の距離により各画像を配置している．したがって気球に近い画像ほど左下に配置され，この例では青っぽい画像は左下に配置されている．しかし，各画像間の類似性は考慮されていないので類似しない画像が近くに配置される場合が少なからずある．また，画像間の重なりも多く視認性が悪い．しかし，検索結果の場合には問合せ画像に類似する画像が特徴量ごとに明確となり，有効なビューといえる．

図8に画像の平均色空間ビューを示す．水平軸が色相，垂直軸が明度である．いずれも画像全体の色の印

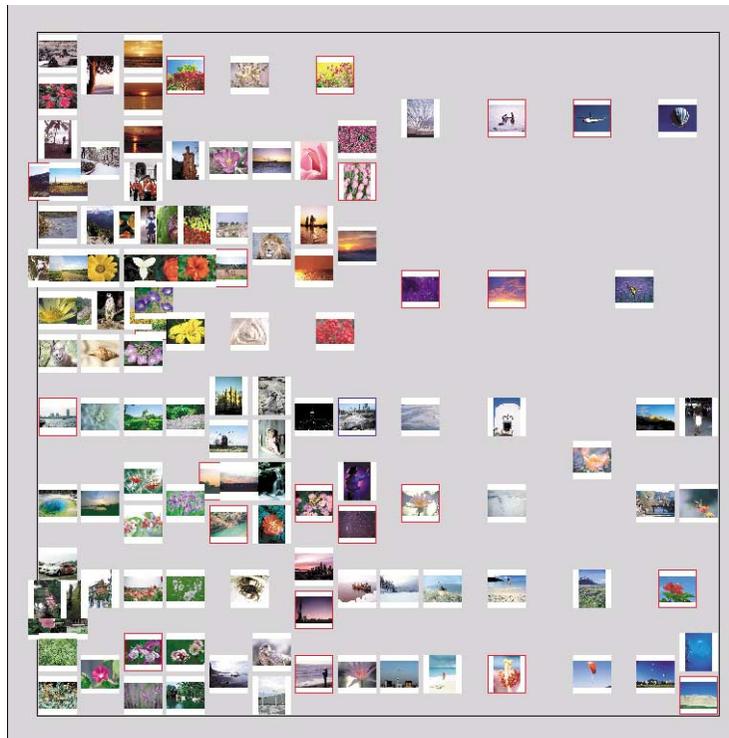


図 6 相互類似ビュー

Fig. 6 Mutual Similarity View.



図 7 単一類似ビュー

Fig. 7 Single Similarity View.

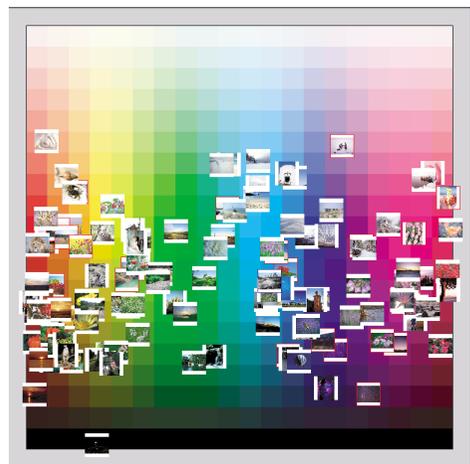


図 8 色空間ビュー (色相-明度)

Fig. 8 Color Space View (Hue vs. Brightness).

象どおりに二次元空間に配置されているので、色により画像を選択したい場合には有効なビューである。

図 9 にリストビューを示す。これは画像を単純にリストで表示したビューである。検索結果の画像集合の場合には問合せ画像に類似する順に並べられる。また、クラスタの画像集合を表示させた場合にはリストの順には意味がない。単純ではあるが画像の重なりがまっ

たくない点では視認性は良い。

7.5 画像分類と画像検索の統合

図 5 で示されるように画像検索と画像分類の 2 つの機能が 1 つの GUI で提供され、ユーザはいつでもその 2 つの機能を利用できる。ユーザは分類木やビューによりデータベースを概観し、データベースの把握が容易であり、その分類木をたどることで所望の画像を



図 9 リストビュー
Fig.9 List View.

探索することができる。分類中に所望の画像が見つけれなくとも、分類中にある類似の画像を問合せ画像として、ただちに画像検索を行うことが可能である。このように分類と検索を融合した GUI を提供することで、一方の機能を提供する場合よりも画像アクセスの効率化が可能である。

7.6 分類木

図 10 に下位の分類の相互類似ビューを示す。最上位の分類は画像全体の色の特徴量のみを利用して分類しているため図 10 を見ても分かるように全体の色合いが類似している画像が 1 つの分類になっており、画像特徴量の特性に基づき、適切な分類木が生成されていると判断する。

7.7 特徴量空間の視覚化

図 5 から類似する画像がビュー上で近接していることが分かり、ユーザは所望の画像に類似する画像群を容易に見つけ出すことができる。また、本システムを Pentium II 300 MHz で動作させた場合に、画像のロード時間も含め 4 秒以内に図 6 のビューを表示することが可能であり、この処理時間ならばユーザはストレスなく対話的に画像探索を行えると予想する。

本画像アクセスシステムではデータベース内の全画像を 1 つのビューで表示することはなく、このように分類の代表画像または検索結果画像のみなので、通常 100 前後の画像であり、対話的な処理に問題はないと予想する。ただし、大量に画像が登録され、かつ、その画像集合に片寄りが多い場合やきわめてばらつきのある場合には、本分類方法では分類に含まれる画像数や分類数が極端に多くなる可能性も考えられ、この場合には対話性が損なわれる可能性がある。こうした場合のブラウズ性能を損わないために、対話的に分類木を生成する機能を用いることで、画像数や分類数が極端に多くならないようにすることが可能である。

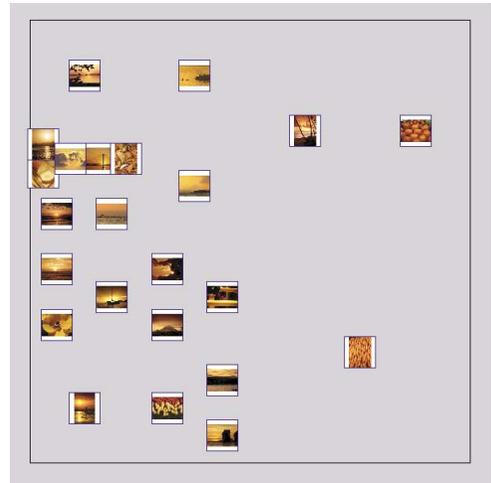


図 10 下位の相互類似ビュー
Fig.10 Lower Mutual Similarity View.

また、画像の重なりが生じているが、マウスのクリックにより下側の画像を見られる。それでも視認性が低下していることは事実であり、課題として画像が重ならないように実装する必要があるであろう。

8. 結 論

画像検索では適切な問合せ条件の指定が困難である。また、画像分類ではユーザの要求にあった分類を生成することが困難であったり、十分な分類精度を得ることができないという問題があった。そこで、画像の分類と検索を融合することにより、分類の代表画像を画像検索での問合せ画像として利用でき、また、分類木をたどって所望の画像に到達できなかった場合には適宜検索に移行することで両者の問題点を補完し画像アクセスの効率を高めることができた。さらに、個々の分類および検索結果である画像集合を二次元の画像空間へ動的に視覚化する手法を提案しユーザの画像の視認性を高めただけでなくデータベースの把握が容易となり、結果として総合的に画像アクセスの効率向上を実現できた。

本稿では本来画像検索に用いていた画像特徴量を基に分類を行ったが、適切な分類を行うには、大分類/小分類といったそれぞれの分類階層に適した特徴が必要である。また、分類木やビューの生成は処理速度を考慮して簡便な方法を選んだために若干精度が低い。今後こういった問題を解決していく予定である。

参 考 文 献

- 1) Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P.: Query by image and video content: The QBIC system, *IEEE COMPUTER*, Vol.28, No.9, pp.23-32 (1995).
- 2) Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R. and Shu, C.: The Virage image search engine: An open framework for image management, *SPIE Storage and Retrieval for Image and Video Databases IV*, Vol.2670, pp.76-87 (1996).
- 3) Pentland, A., Picard, R.W. and Sclaroff, S.: Photobook: Tools for content-based manipulation of image databases, *SPIE Storage and Retrieval for Image and Video Databases II*, Vol.2185, pp.34-47 (1994).
- 4) Smith, J.R. and Chang, S.F.: Tools and techniques for color image retrieval, *SPIE Storage and Retrieval for Image and Video Databases IV*, Vol.2670, pp.426-437 (1996).
- 5) 呉, 金子, 牧之内, 上野: 自己組織化特徴マップに基づいた類似画像検索システムの設計・実装と性能評価, 電子情報通信学会技術研究報告, Vol.100, No.31, pp.9-16 (2000).
- 6) Zhang, H. and Zhong, D.: A scheme for visual feature based image indexing, *SPIE Storage and Retrieval for Image and Video Databases III*, Vol.2420, pp.36-46 (1995).
- 7) Hatano, K., Qing, Q. and Tanaka, K.: A SOM-based information organizer for text and video data, *Proc. 5th International Conf. on Database Systems for Advanced Applications*, pp.205-214 (1997).
- 8) 保木, 片山, 仲川, 小西: 協調マルチメディア情報収集法, 情報処理学会研究報告データベースシステム, Vol.133-56, pp.335-340 (1997).
- 9) 串間, 佐藤, 赤間, 山室: 大量画像の閲覧を目的とする階層的分類支援機能—画像目録の実装と評価, 情報処理学会論文誌データベース, Vol.41, No.SIG1(TOD5) (2000).
- 10) 斎藤, 村上: 絵画の自動分類法に関する一検討, 電子情報通信学会技術研究報告, Vol.IE97-89, pp.39-44 (1997).
- 11) Lipson, P., Grimson, E. and Sinha, P.: Configuration based scene classification and image indexing, *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.1007-1013 (1997).
- 12) 舘村: インタラクティブ視覚化による文献集合からの情報獲得支援, 日本ソフトウェア科学会第13回大会 (1996).
- 13) Chalmers, M. and Chitson, P.: Bead: Explorations in information visualization, *Proc. ACM SIGR'92*, pp.330-337 (1992).
- 14) Rubner, Y., Guibas, L. and Tomasi, C.: The earth mover's distance, multi-dimensional scaling, and color-based image retrieval, *Proc. ARPA Image Understanding Workshop* (1997).
- 15) Chen, J.Y., Bouman, C.A. and Dalton, J.C.: *Similarity pyramids for browsing and organization of large image databases* (1998).
- 16) 岩崎: 類似画像検索を実現する距離空間インデックスの実装及び評価, 情報処理学会論文誌データベース, Vol.40, No.SIG3(TOD1), pp.24-33 (1999).
- 17) Iwasaki, M., Takahashi, N. and Morozumi, K.: Access Method for Image Database, *SPIE Storage and Retrieval for Media Databases 2000*, Vol.3972, pp.21-29 (2000).
- 18) 高木, 下田 (編): 画像解析ハンドブック, pp.518-521 (1991).

(平成 12 年 6 月 20 日受付)

(平成 12 年 9 月 27 日採録)

(担当編集委員 西尾 章治郎)



岩崎雅二郎 (正会員)

1987 年早稲田大学理工学部工業経営学科卒業。1989 年同大学院工学研究科機械工学専攻修士課程修了。同年日本電気(株)入社。1990 年退社。同年(株)リコー入社。現在同社ソフトウェア研究所に勤務。全文検索の研究開発, 構造化文書データベースの研究を経て画像検索の研究開発に従事。



両角 清隆

1978 年金沢美術工芸大学美術工芸学部卒業。1980 年筑波大学大学院芸術研究科デザイン専攻修了。同年いすゞ自動車(株)入社。1983 年退社。同年(株)リコー入社。総合デザインセンターに勤務。1992~1995 年リコー在職中に千葉大学大学院自然科学研究科環境科学専攻デザイン科学講座に入学・修了〔博士(学術)〕。2000 年退社。同年東北工業大学工業意匠学科に勤務。プロダクトデザイン, インタラクティブデザインの教育・研究に従事。