

大量画像データベースへのアクセス手段の提案

岩崎 雅二郎[†] 両角 清隆^{††}

画像データベースからユーザが所望する画像を検索する方法として画像検索及び画像分類があるが、画像検索では問合わせの指定が容易ではなく、画像分類ではユーザの要求にあった精度の高い分類を生成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、その上で大量の画像を表示空間上に視覚化することで、総合的に画像アクセスの効率化を実現する画像アクセス方法を提案する。

Access Method for Huge Image Database

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

This paper describes a way to access images in a huge image database. Image retrieval and image classification are representative methods for accessing images in an image database. With image retrieval, it is difficult to specify a proper query image. With the image classification, it is difficult to achieve the high accuracy classification needed by users. We propose a method that makes it possible to access images efficiently using integration of image retrieval, image classification, and visualizing that can show many images in the image feature space.

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、画像データがオフィスにも家庭にも増加しつつある。このような環境の変化により、画像アクセスの研究開発は世界中で盛んに行われている。

一般的な画像アクセスの方法として、問合わせ画像から特徴量を抽出し、データベース中の画像の特徴量と比較し、類似する特徴量をもつ画像を検索結果とする方法がある。しかし、問合わせに用いる画像を如何にユーザに提示するかが問題となる。

一方、画像を予め自動的に分類しておく、その分類からユーザが所望するカテゴリ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いかったり、分類の精度が低いと希望の画像にたどりつけないという問題がある。

さらに総合的な画像アクセスのインターフェースを提供するためにはユーザの視認性を損わず大量の画像をユーザに如何に提示するかも重要な問題である。そこで、これらの問題を解決するために画像の検索及び分類を融合し、同時に画像空間を視覚化する方法を提

案する。

2. 従来の研究

画像特徴量による画像検索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出し検索を行う。QBIC では、問合わせ画像を与えて検索する方法の他に

- 画像の概略を描く
 - ユーザが画像のヒストグラムを入力する
- 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、ヒストグラムに関しても必ずしもユーザが適切に入力できるとは言えない上に負担も同様に大きい。

問合わせとして画像を指定して類似する画像を検索する方法では、問合わ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ユーザの所望する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、画像の分類方法としては、SOM により分類する方法^{2),3)} がある。また、カテゴリの制約条件を予め人為的に指定する分類方法として既存の画像データベースの検索条件指定を組合せてカテゴリの条件を指定する研究⁴⁾、同様に絵画に特化して自動的に分類

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

^{††} リコー 総合デザインセンター
Corporate Design Center, RICOH Co., Ltd.

する研究⁵⁾、個々のカテゴリ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⁶⁾がある。

このように分類をユーザーに提示しカテゴリを選択することでユーザーの所望する画像にアクセスする方法では、ユーザーはカテゴリを選択するだけでよく、また、データベースの画像空間が分類として表現され、ユーザーがデータベースの構成を把握し易いメリットがある。その反面ユーザーの要求する条件に適合する分類が提示されなければ、ユーザーはカテゴリを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カテゴリに属してしまっていると、ユーザーがその画像にアクセスすることが不可能になる。

カテゴリを代表する画像はデータベースを代表する画像なので、そのカテゴリを代表する画像を提示し問合わ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ーの要望と合わない場合や分類のみで所望の画像を取得できなかった場合に検索に移行できれば、ユーザーが適切な検索条件を指定することで所望する画像が検索でき、分類の問題点を回避することが可能である。そこで、画像特微量による検索及び分類機能の融合を行う。

分類及び検索の融合により画像への効率的なアクセス機能を提供することが可能だが、総合的な画像アクセスの効率化を目指すには分類中の画像集合や検索結果の画像集合から如何にユーザーが所望する画像を効率よく選択できるかが重要となる。提示する画像が少數であれば単にリストアップするだけよいが、100以上の画像を提示する場合には画像集合の視覚化を工夫しなければユーザーの負担が増大する。

情報の視覚化の研究としては、三次元空間での類似度に従いバネモデルによりインタラクティブに配置する方法⁷⁾がある。しかし、画像集合全体を表示空間に配置するので処理に時間を要する。

画像を何らかの尺度により空間に表示する方法は有效であるが、検索結果のように対話的に画像集合が生成される場合には画像空間の視覚化の処理時間が問題となる。そこで、画像分類及び検索を融合した上で類似尺度による簡便な画像空間の視覚化方法により総合的に効率的な画像アクセス手法を提案する。

3. 画像検索

本研究では画像から画像特微量を抽出し、画像特微量空間の中での特微量間の距離を類似度として画像の検索を行う。利用する画像特微量を以下に示す。

(1) 色：色のヒストグラムを特微量とする。

- (2) 配色：画像を格子状のブロックに分割し各ブロックの平均色を特微量とする。
- (3) エッジ分布：画像から抽出したエッジを格子状のブロックに分割し各ブロックごとに四方向（水平、垂直及び斜め二方向）のエッジのピクセル数を特微量とする。
- (4) テクスチャ：画像を格子状のブロックに分割し各ブロックの同時生起行列⁸⁾により求めた統計量を特微量とする。

$i = 1, \dots, 4$ で示される個々の特微量空間での特微量 f_{ai}, f_{bi} 間の距離を $D_i(f_{ai}, f_{bi})$ とすると、画像 I_a, I_b 間の距離 $D(I_a, I_b)$ は以下のように示される。

$$D(I_a, I_b) = \sum_{i=1}^4 w_i D_i(f_{ai}, f_{bi}) \quad (1)$$

w_i は各特微量の重み付けであり、特微量ごとに重み付けを行うことでユーザーの検索観点を反映することができる。

画像特微量はデータベースに画像を登録するときに画像から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わせ画像から画像特微量が抽出され、その特微量とデータベース中の特微量の距離を式 1 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リストを出力する。

4. 画像分類

画像検索と同じ四つの画像特微量を最も簡単なクラスタリング手法である NN(Nearest Neighbor) 法を用いてクラスタリングすることで分類する。より精度のよいクラスタリング手法も存在するが、分類を対話的に行うことの想定しているので、クラスタリングの処理に時間を要することは致命的な問題になる。そこで最も単純な NN 法を採用した。NN 法により得られたクラスタが分類を構成する各カテゴリとなり、各クラスタの中心となる特微量がカテゴリの代表画像となる。

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）が生成できる（図 1）。したがって、ユーザーは所望する画像と類似する代表画像を分類木のルートから順次辿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。

ただし、代表画像により構成される分類木からユーザーが所望の画像を得るには以下の条件を満足する必要がある。

- (1) 分類木の各ノードの代表画像はノード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。

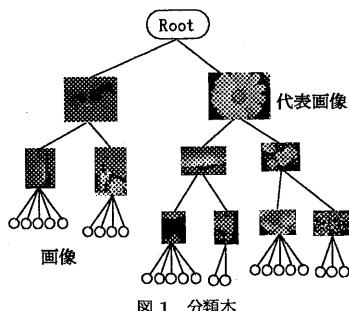


図 1 分類木

- (2) ノードを構成する画像集合のすべての画像はある一意の条件によって類似していなければならぬ。

ノード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ーはその代表画像を選択することはあり得ず、その所望する画像を探し出すことができない。したがつて、上記の第一の条件は画像分類においては必須条件となる。

しかし、分類木だけでなく検索を融合して画像にアクセスすることを前提とした場合には、分類木の選択により所望する画像に近い画像が一つでも見つかれば、それを検索の問合わせ画像として検索操作に移行することによって所望する画像が得られるので、第一の条件を緩和することができる。本研究では検索と分類の融合を前提としているので第一の条件は緩和される。

また、ノードの代表画像に対して、色の特徴が極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、テクスチャが極めてよく似ていて色はほとんど似ていない画像がそのノードに含まれていたとする。これは第一の条件を満足している。しかし、このように分類中に互いに類似点がない画像が存在するとユーザーはカテゴリの意味が理解できずに戸惑う。したがつて第二の条件は必須条件と言える。

以上のことから利用する四つの画像特徴による画像分類では以下の点を考慮する必要がある。

- 分類木のある階層での分類で複数の特徴量を利用することは第二の条件に反するので、各階層の分類では一つの特徴量しか利用できない。
- 四つの特徴量のうち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するので、大分類には向かない。つまり、第一の条件を満すようにエッジ分布またはテクスチャ分布を使って分類を行うとカテゴリ数が極めて多くなり、分類の意味が薄れてしまう。

- 逆に色や配色は比較的大まかな特徴であり、小分類には向かない。
- 色の特徴量に位置情報を追加したものが配色の特徴量であり、色と配色には包含関係がある。色の分類を細分化するために配色を利用できるが、その逆はあり得ない。したがって、分類木を構成する時には色の分類の下位に配色の分類がくることが必然的に決る。さらに、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捉えることができ、テクスチャ分布とエッジ分布も包含関係にあると言える。

これらの条件下では必然的に分類木は、ルートから色、配色、エッジ、テクスチャという順で分類を行うことが良いと判断できる。本研究では基本としてこの順で分類木を生成するが、ユーザーは自分の要求に合わせた分類木を対話的に作ることも可能である。

5. 画像特徴量空間の視覚化

分類の各ノード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 画像を越えることもある。100 を越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ーにとってかなり負担となる。ユーザーが画像を、よりブラウズし易くすることが重要となる。座標軸に意味を持たせつつ画像特徴量の多次元空間を二次元の表示空間に落すことは困難である。そこで、座標軸は意味を持たないが、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木構造で表現される部分空間に分割する。この木構造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二つのステップに分られる。

(1) 二分木生成

部分空間を二つに分割するには、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 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

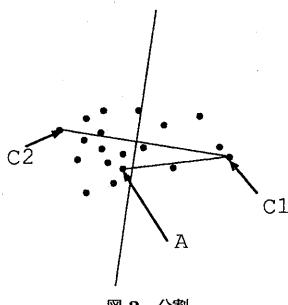


図 2 分割

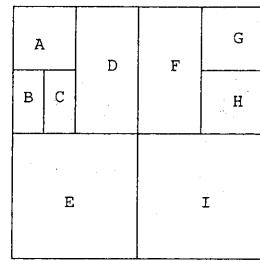


図 4 表示空間

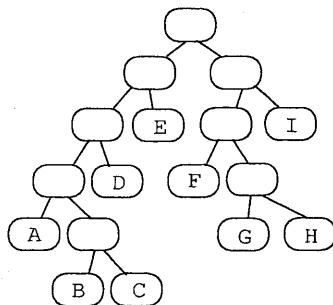


図 3 画像特徴量空間の二分木

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獲得する。獲得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に獲得した点を属させる。
- (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辿り、

- (a) 画面を前回の分割と異なる座標軸で等分に部分領域に分割する。初回の分割の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振り、子ノードがリーフノードの時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) 各子ノードを辿り (a)に戻る。辿る子ノードがなくなった時に処理を終了する。

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 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。

また、以下の表示方式も補足的に追加することで、ユーザーの要求に適したビューを選択することを可能とする。

● 単一類似ビュー

代表画像又は問合わ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。四つの特微量から任意の二つの特微量を水平軸及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、垂直軸を明度又は彩度とした二次元に各画像が配置する。

● 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。

6. 画像アクセスの実装

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。その画面を図 5 に示す。登録した画像データは 2,000 画像のフォト画像である。画面は四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ラーウィンドウ、右下がビューウィンドウである。

このツールでは、画像の検索、分類及びブラウズ機能を提供する。画像の検索や分類に利用する特微量は画像登録時に抽出される。また、分類木はルートから順にアクセスされた時点で下位の一階層分の分類を動的に生成する。分類時にユーザー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、ユーザーが自由に重み付けを指定することも可能である。生成された分類木はセーブ、ロードが可能であり、ユーザー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、ユーザーに提供することも可能である。

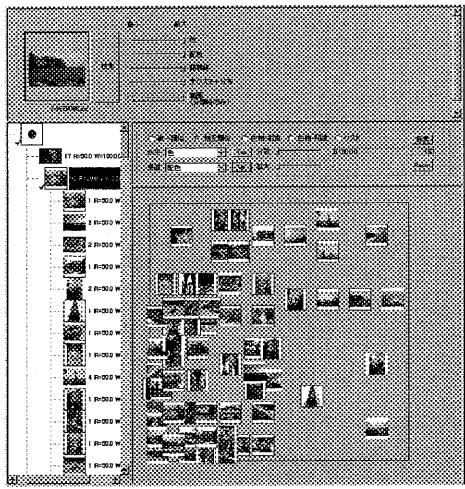


図 5 画像アクセスシステムの画面

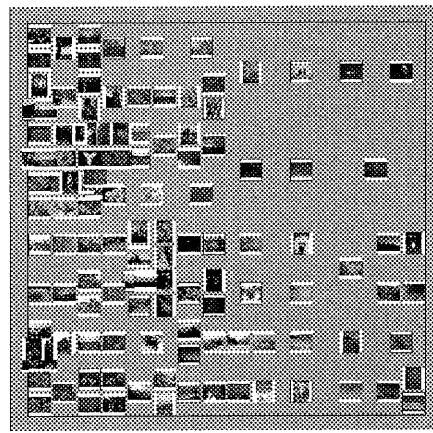


図 6 相互類似ビュー

示す。図の例では前述の 2,000 画像を分類し分類木の第一階層では 193 のノードが生成されており、その 193 の代表画像が表示されている。

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。また、リストビューを除く他のビューと比較して分散しており画像間の重なりも少ない。

図 7 に単一類似ビューを示す。水平軸及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、垂直軸が配色である。左下の気球の画像からの色特徴量の距離と配色特徴量の距離により各画像を配置している。従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わせ画像に類似する画像が特徴量ごとに明確となり、有効なビューと言える。

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象通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わせ画像に類似する順に並べられる。また、分類ノード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全くない点では視認性が良い。

6.1 検索ウィンドウ

問合わせ画像を元に画像の検索操作を行うウィンドウである。問合わ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。上位の四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、この四つのスライダによる重み付けは分類時にも利用する。新たに分類を行う時にはこのスライダー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がクラスタの半径として利用される。

6.2 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。画像は各ノードの代表画像である。チェックされた分類ノードは既に閲覧したことなどを示す。各ノードをクリックすると下位のノードが表示される。もし、下位のノードが未生成の場合には、クリックされると同時に分類処理を行い、下位のノードを一階層分生成する。

6.3 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。

6.4 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ノードの画像集合や検索結果の画像集合を二次元に表示する。分類木のノードが表示されている場合には、各画像は下位のノードの代表画像であり、その画像をクリックするとそのノード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

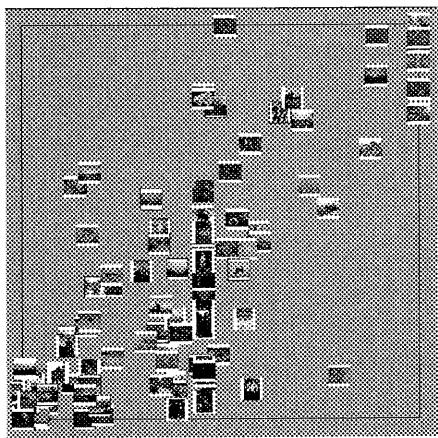


図 7 単一類似ビュー

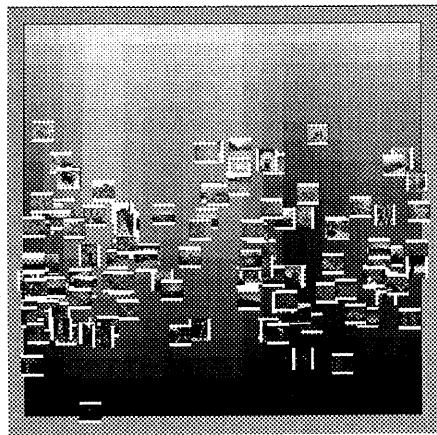


図 8 色空間ビュー（色相-明度）

7. 結論

画像検索では適切な問合わ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したことにより、分類の代表画像を画像検索での問合わ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、個々の分類及び検索結果である画像集合を簡便な手法により、二次元の画像空間として動的に視覚化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った。

本研究では画像検索に用いていた画像特徴量を基に分類を行っていたが、適切な分類を行うには、大分類

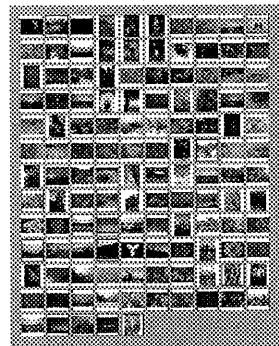


図 9 リスト

／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、分類木や画像表示空間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) M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele and P. Yanker, Query by Image and Video Content: The QBIC System, IEEE COMPUTER, Vol.28 No.9, pp.23-31, 1995
- 2) 波多野, 亀井, 田中, 映像自己組織化機構に基づく内容記述と類似シーン検索, 情報処理学会研究報告 データベースシステム 113-29, pp.173-178, 1997
- 3) 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 133-56, pp.335-340, 1997
- 4) 佐藤, 赤間, 山室, 大量画像概観のための分類木構築システム, DEWS'99, 1999
- 5) 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 IE97-89 pp.39-44, 1997
- 6) P. Lipson, E. Grimson, P. Sinha, Configuration Based Scene Classification and Image Indexing, Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.1007-1013, 1997
- 7) 館村, 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
- 8) 画像解析ハンドブック, 東京大学出版, pp.518-521, 1991