

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 —

串間 和彦[†] 佐藤 路恵[†] 赤間 浩樹[†] 山室 雅司[†]

大量に蓄積されつつあるデジタル画像の閲覧方法として、色、形状等の画像特徴量に従ってそれら画像を階層的に分類し、本の体裁で画像目録として提示する方法について述べる。目録の作成では固定的な分類手順ではなく、画像内容、分類の目的に適した分類手順を実現するため、代表例に基づく分類法とクラスタリング法を併用した複合分類法を提案し、さらに有望そうな分類観点をシステムが提示する構築支援機能を作成した。構築した階層的分類木の表示には電子本を利用した。評価により、自動的なクラスタリングに比べて分類精度が高く、かつ内容検索に比べて希望する画像をより少ない手間で見つけられることを示した。

Hierarchical Classification Method for Overlooking Huge Amount of Image Contents - Making a Catalogue of Images -

KAZUHIKO KUSHIMA,[†] MICHIE SATOH,[†] HIROKI AKAMA[†] and MASASHI YAMAMURO[†]

In order to overview huge amount of image contents efficiently, the hierarchical classification method is proposed. The method is characterized by the combination of a template-based classification and an automatic clustering, each of which is based on the image features. A classification support environment is implemented, which treats several image features as classification viewpoints. According to the constructed classification tree, the images are displayed on the electronic book - called the CyberBook - by the means similar to the catalogue. It helps finding desired images. The experimental results reveals that the number of referred images for a finding target image is reduced by 25% compared with the case of content-based image retrieval.

1. はじめに

安いデジタル画像が大量に生成されつつある現在、画像にテキストの索引をつけ、キーワードで検索する手法に代わり、色や模様、形といった画像特徴に基づく内容検索手法が提案されている¹⁾。内容検索手法では、複数の検索対象画像、及びキーとして与えられる画像についてそれぞれ画像の特徴量を決定し、両者の類似度を計算することで所望の画像を発見する。しかし、内容検索技術には(1)画像集合全体が把握できないため、検索結果が正しいものであるか検証できない、(2)検索のキーとなる画像を指定しにくく、所望の画像に到達するのに手間がかかるといった問題がある。今後は、画像集合全体にどのような画像が存在

するかを俯瞰でき、特定のキーを与えなくとも欲しい画像を発見できる機能が必要である。本論文ではそのような機能を画像目録と呼ぶ。画像中から自動抽出可能な特徴に基づいて、代表例に基づく分類法²⁾とクラスタリング法を併用した複合分類法により、構築者の意図に沿った分類の構築を支援する手法を提案する。さらに分類結果の表示方法の一例として電子本を用いた実現方法を紹介する。

2. 従来の研究と課題

2.1 内容検索技術

大量の画像から希望するものを検索する方法として内容検索技術がある。検索対象として画像データベース(画像DB)中に格納された画像と、検索のキーとして与えられる画像のそれぞれについて色や模様、画像中のオブジェクトの形や位置といった画像特徴量を決定し、特徴量間の類似度を計算することで、希望する画像を検索する技術である。これまでにQBIC, Virage,

[†] NTTサイバースペース研究所
NTT Cyber Space Laboratories

VisualSEEk, ExSight 等のシステムが提案されている。詳細については文献⁹⁾を参照されたい。この内、ExSight⁹⁾はオブジェクトに基づく高速な検索を狙いとして、入力画像中からの自動的なオブジェクト抽出、多次元インデックスによる高速検索を特徴としている。しかし、ExSight を開発し、サービスに適用した経験から画像の内容検索には以下の問題があることが判明した。

(1) 画像の全体が把握できない

内容検索の結果として返されるのは、特定のキー画像に類似した画像 DB の一部分であり、それら以外にどのような画像がデータベース内に存在するかが分からない。このため、検索結果が正しいものであるか、特に検索もれがないかを判断できない。また、予め欲しい画像が分かっていることが前提であり、「何か良い物がないか」といった画像全体を見渡すような探し方はできない。

(2) 所望の画像を発見するためにはキー画像の指定が必要で手間がかかる

内容検索を行うためにはまず欲しい画像に類似した画像をキー画像として与える必要がある。このため、画像 DB から数件の画像をランダムに表示させながら希望する画像に近いキー画像を見つけるか、スケッチによって欲しい画像を描画する方法がとられる。しかし、前者では適当なキー画像を探すのに時間がかかる。また、後者では正確に検索キーをスケッチするのが困難といった問題がある。

2.2 情報の分類に関する従来手法

検索を補完する方法としては、情報を分類する方法がある。画像に関しては、画像特徴に基づいてクラスタリングにより画像集合を分類し、3D 化などのビジュアルライゼーションによって分類結果を表現する手法が多く提案されている。

波多野ら⁴⁾は動画像から切り出したカットについて DCT (離散コサイン変換) 係数に基づいて SOM⁶⁾ (Self-Organizing Map) により分類を行い、TV 番組のカットを分類している。その際 DCT 係数以外にカットに対する内容記述情報を併用し、精度を向上している。仲川ら⁸⁾はドキュメント内でのキーワードの共出現に関する解析方法を応用し、画像内のブロックを DCT 係数に従ってグルーピングした後、異種ブロック間の共出現ルールを SOM を利用して抽出することで、「移動する台風中心」等、画像内位置が変化する対象を分類している。Lipson ら⁷⁾は画像中の領域間の位置的、画像特徴的な関係に基づき Class Models という静的なテンプレートを作成し、自然写真の分類に適用して

いる。また、Ma ら⁸⁾は分類それ自体を目的とするものではないが、模様に着目した SOM により、テクスチャセンサーラスと呼ぶ類似した模様の画像集合を決定している。

これらの多くは DCT 係数やテクスチャといった単一の画像特徴に基づいている。しかし、一般的な写真や画像素材のように多様な画像集合を分類する場合には複数の観点からの分類が必要である。また、一階層の分類が基本であり、大量画像に適用する場合は各分類に属する画像数が大きくなり、全体を俯瞰しにくい。さらにクラスタリングを基本とする方式には人間の直感にあった分類ができにくい問題がある。例えば岩崎ら⁹⁾はクラスタリングによる階層的な分類と内容検索を組み合わせることでクラスタリング結果を補完する方式を提案している。

3. 画像特徴に基づく階層的分類手法

分類の手段は、対象とする画像の種類と分類の目的に大きく左右される。本論文では、我々の経験から分類に対する要望が強く、かつ一般的に用いられているクリップアート集、及び風景写真集を対象にし、それら画像集合の全体を俯瞰できる階層的な目録の構築を目的とした。さらに結果の定量的な評価を試みた。

3.1 提案手法の概要

画像から自動抽出可能な特徴に基づいて複数の分類観点を用意し、それらを再帰的に適用することで画像集合の階層的な分類を構築する。分類の構築に当たっては有効と思われる分類観定の候補をシステムが提示することで構築者を支援する環境を提供する²⁾。

構築した分類を目録として閲覧する方法には日常慣れ親しんだ情報の固まりである本を活用する。本には目次があり、それに沿ったばらめくりで本の概要を捉えることは誰もが日常良く行っていることである。そこで電子本である CyberBook¹⁰⁾を用い、階層的に分類した情報を本の章立てに対応させ、さらに目次にあたるページを設けて代表画像を表示する。

3.2 複合分類法

画像を分類する方法として、各画像特徴に基づいて自動的なクラスタリングを行う方法が考えられるが、その結果は人間から見て分かりやすい分類にならないケースが多い。そのため、予め代表的な分類の例(分類項目と呼ぶ)をいくつか決めておき、最も近い項目に個々の画像を割り当てる代表分類法と、k-means 法¹¹⁾によるクラスタリング法を併用した。この手法を複合分類法と呼ぶ。構築者はトップダウン的な分類において上位で代表分類法を用い、下位ではクラスタリン

グ法を用いることで無理のない詳細化を行うことができる。代表分類法，クラスタリング法の詳細は後述する。

複合分類法に従った分類木の作成は以下の手順で行う。

(1) 基本的アルゴリズム

入力 X: 画像集合。

D: 分類観点集合。

出力 T: 画像の分類木。先頭に分類観点を持つ順序付集合 [d, Y1, ..., YN] で表す。Yi は画像集合。

なお，下記で++,--は要素の追加，削除を表す。

makeCatalogueTree(X, D)

```

{
  if judgeFunc(X) return(X);
  // 再分類の必要性判定.
  d ← selectFunc(X,D); // D 内の観点 d を選択.
  [Y1,Y2,...YN] ← clustFunc(X, d);
  // X を観点 d で N 個の Yi に分類.
  T ← [d];
  forall Yi {
    T ← T ++ makeCatalogueTree(Yi, D -- d);
    // D から d を除いたものを観点とし，再帰.
  }
  return(T);
}

```

(2) 関数の説明

- judgeFunc(X): X の再分類を行うか，判定する関数。実装では X に含まれる画像数があるしきい値未満の（つまり再分類が不要な）場合に true を返す。
- selectFunc(X,D): X に関して，D 内の観点を 1 つ選択し d として返す関数。D としては表 1 に示す分類観点を用いる。かつ，D 内のどの観点を選択するかは構築者が判断する。
- clustFunc(X,d): X を観点 d で分類し，N 個の画像集合 [Y1,Y2,...YN] として返す関数。各観点毎に 3.3.1 以降に示す方法で画像集合を分割する。

(3) アルゴリズムの出力例

画像が X=[X1,...X10]，分類観点が D=[d1,...,d5] のとき，分類木は以下のように出力される。

T=[d1,[d2,[d3,X1,X2],[d4,X3,X4]],[d2,[d3,X5,X6],X7,[d5,X8,X9]],X10]

図 1 に上記に対応する分類木を示す。

3.3 画像特徴と分類観点

目録としての利用を考えると，どういう観点から分類したかが容易に把握できる画像特徴を選択する必要がある。そのため，理解が容易でかつ自動抽出可能な

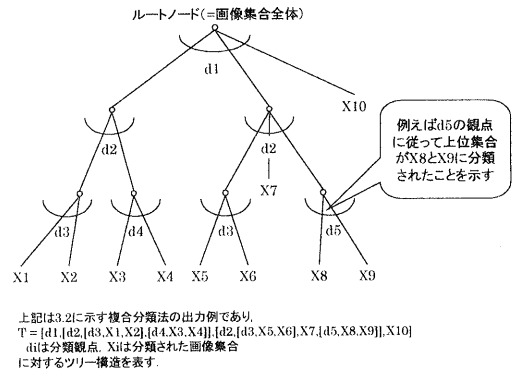


図 1 分類木の例
Fig.1 Example of Classification Tree

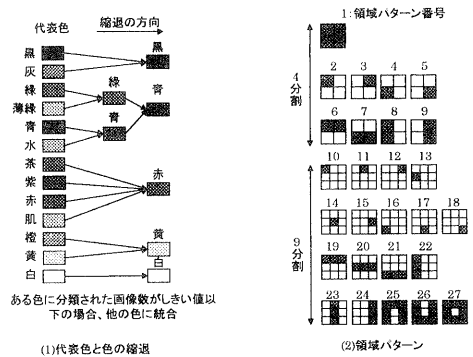


図 2 代表色と領域パターン
Fig.2 The Dominant Color and the Area Pattern

表 1 画像特徴と分類観点
Table 1 The Image Features and the Classification View-Point

画像特徴	分類観点	分類項目	分類方法
色	領域代表色	赤，緑等13色	代表分類
	第n主要色	赤，緑等13色	代表分類
	色クラスタリング	動的に決定	クラスタリング
形状	代表形状	丸，四角等8種	代表分類
	形状クラスタリング	動的に決定	クラスタリング
	アスペクト比	縦長，横長	代表分類
直線	主要直線	0本，1本以上	代表分類

色，形状，及び直線を採用した。分類の観点を表 1 に，詳細を以下に示す。

3.3.1 色に関する代表分類法

色は画像を分類する際の支配的な要素である。そこで，1 枚の画像について複数の領域毎に代表色を決定する方法と，画像全体について支配的な色を順序づけられた主要色として決定する方法を採用した。

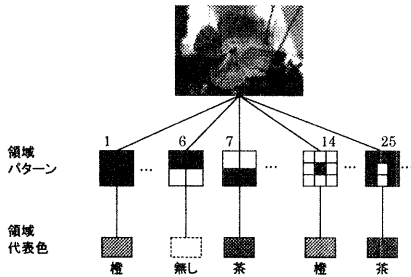


図3 領域代表色の例
Fig.3 Example of the Area-Color

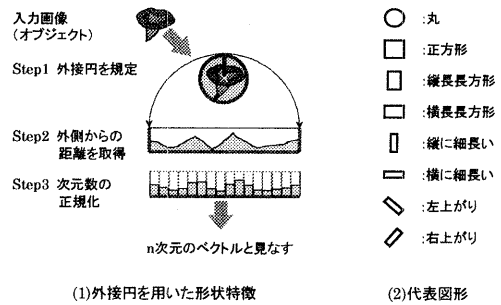


図4 形状の表現
Fig.4 Shape Representation

代表色としては人間の直感に沿った色として黒, 灰, 緑, 薄緑, 青, 水, 茶, 紫, 赤, 肌, 橙, 黄, 白の計 13 色を選択した (図 2(1)) .

領域色の決定では画像全体を 2×2, 及び 3×3 のブロックに分割する. 1 個以上のブロックをまとめることで画像全体, 上半分, 下半分, 中心等の計 27 種の領域パターンを定義する (図 2(2)). 画像をブロックに分割して色比較を行う方法は良く知られた方法であるが, 本手法では個々のブロックではなく, それらを組み合わせた 27 種の領域パターンに基づいて分類を行っている点が特徴である.

図 2(2)に示す領域パターンの内, 9 分割の場合は全てのブロックの組み合わせを満たしているものではない. これらの領域パターンは以下を目的として経験的に決定したが, 必要に応じて追加が必要である.

- ・ 1 ブロックを使用 (領域パターン番号 No.10 ~ No.18): 風景中の様々な箇所にある比較的小さな被写体に着目する.
- ・ 3 ブロックを使用 (No.19 ~ No.24): 空や地面といった上下の領域等, 比較的大きな被写体に着目する.
- ・ 7 もしくは 8 ブロックを使用 (No.26 ~ No.27): 中央の被写体をマスクして背景情報に着目する.

各領域パターン毎に 13 色を横軸とするカラーヒストグラムを作成し, ヒストグラムの高さがあるしきい値以上で最も高い色をその領域の代表色とする. しきい値以上の色が存在しない場合, 領域代表色は無しとする. なお, 人の目には暗くて地味な色より明るく派手な色の方が目につきやすいので, 明るい色の方が暗い色よりしきい値が低くなるよう設定した.

上記計算により, 各領域パターン毎に代表色を決定できる. 例えばある一枚の画像について, 領域パターン 1 の代表色は「橙」, 領域パターン 6 の代表色は「無

し」といった具合である (図 3). なお, 13 色への分類の結果, ある代表色に属する画像数が一定数 (実験では 10 枚) 未満の場合は図 2(1)にある縮退方向に従い, その集合を類似した色相の他の集合に統合する.

画像全体について見れば代表色は必ずしも一色ではなく, 複数の支配的な色が存在するケースが殆どである. そこで画像全体についての代表色ヒストグラムを作成し, 高いものから順に第 n 主要色までを決定する. 実装では第 5 主要色までを決定した.

3.3.2 形状に関する代表分類法

オブジェクトの形状は, 輪郭と外接円の距離を 15 度間隔で計算した 24 次元のヒストグラムとして表現することができる³⁾ (図 4(1)). 予め登録した代表図形と, 画像中のオブジェクトについて形状特徴を計算し, ヒストグラム間の距離計算によって, 各オブジェクトを最も類似する代表形状に分類する. 代表図形としては図 4(2)に示す基本的な 8 種類の形状を用意した. 対象とする画像の種類に応じた固有の図形を登録することも可能である. なお, 今回はクリップアートのようにオブジェクトが画像中に一つだけ存在する場合を対象とした. 画像中に複数のオブジェクトが存在する場合の扱いについては 5.5 で考察する.

3.3.3 支配的な直線に関する代表分類法

画像内には多くの直線が存在し, それらが構図を決定する場合が多い. 例えば風景写真では, 水平線や地平線のように上下を 2 分割するような直線, 或いは建物のように直立した直線等が構図を特徴付けている. 画像中の直線は Hough 変換¹²⁾を利用して抽出する (図 5). これは主要エッジ上の点を式 $\rho = x \cos \theta + y \sin \theta$ で表現し, 全エッジ点に対し (ρ, θ) 空間上で投票を行うことで, 直線を検出する方法である. 今回は構図を決定する支配的な直線に着目し, 抽出された直線群

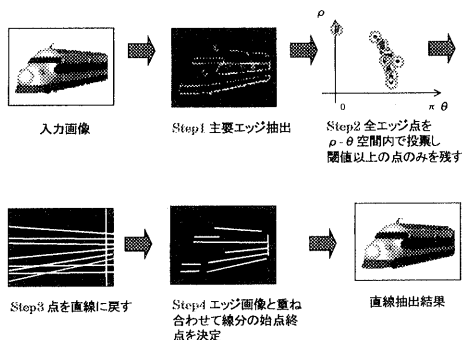


図 5 Hough 変換による主要直線の検出

Fig.5 Extraction of Dominant Lines By the Hough-transform

の内、直線の長さが画像長辺の 7 割以上のものについて本数が 1 本以上の場合と、0 本の場合に分類した。

3.3.4 クラスタリング法

色、形状に関しては k-means 法¹¹⁾によるクラスタリングを行う。k-means 法は広く知られたクラスタリング法であり、実装が容易であること、クラスタリングの個数を初期値として与えることが可能で構築者が集合を何分割するかを指定し易いことから採用した。

色に関しては画像全体を対象に HSI 尺度上での Hue に着目し、16 次元からなるカラーヒストグラムを生成する。次に、それを 16 次元の多次元空間に見立てて各画像に対する点をポイントし、k-means 法を適用する。また、形状に関しては 3.3.2 で述べた 24 次元のヒストグラムを多次元空間に見立てて k-means 法を適用する。なお、クラスタの個数は構築者が逐一指定できるが、実験では 3～5 個程度が多く用いられた。

3.4 分類観点提示による分類ナビゲーション

構築者は 3.2 で述べた分類観点のいずれかを選択し、画像全体をトップダウンに分割する操作を繰り返す。その際、複数の分類観点を用意したことで、選択の自由度は高まったが、逆に構築者はどの分類観点を選択していいかわかりにくくなる。そこで、次の分類のために有望であると考えられる観点を提示することで構築者をナビゲートする。テキスト情報の絞り込みの場合、次に指定されるであろう検索キーワードと、それが選択された場合にヒットする情報の個数を予め提示するナビゲーション手法が提案されている^{13),14)}。

同様の考えに従い、全ての分類観点毎に次にそれが選択された場合にどのような分類結果となるかを、次の分類に進む前に計算する。計算の結果は、バランス良く分類できた順に次の候補として構築者に提示する。この方法により、構築者は次の分類結果をある程

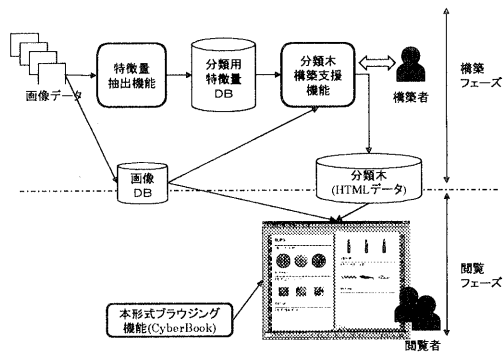


図 6 システム構成と処理の流れ

Fig.6 The System Configuration and the Process Flow

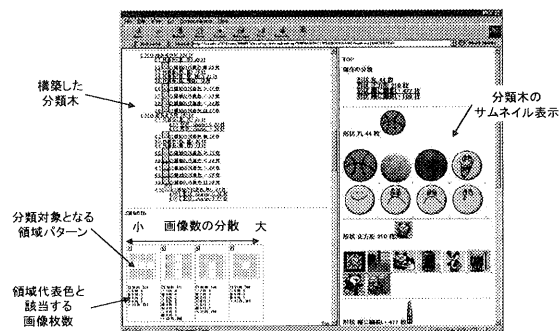


図 7 分類観点のナビゲーション例

Fig.7 Example of the Classification Navigation

度予想しながら最適な分類観点を選ぶことができる。

4. 画像目録システムの概要

前章の提案をもとに試作した画像目録システムの構成を図 6 に示す。

4.1 構築フェーズ

先ず代表分類法に属する、領域代表色、第 n 主要色、代表形状、代表直線に関して、画像毎に計算した結果を分類用の特徴量 DB に登録する。

次に分類木構築支援機能は領域色、形状といったそれぞれの分類観点ごとに候補を提示する (図 7 参照)。例えば「領域代表色」の場合は、領域パターン毎に各代表色に属する画像枚数を集計した後、集計された画像数の分散を領域パターン毎に計算することで、最もバランス良く分類可能と思われる領域パターンから順に構築者に提示する。

構築者が選択した分類観点に基づき、システムは分類した結果をサムネイルとともに構築者に提示する。

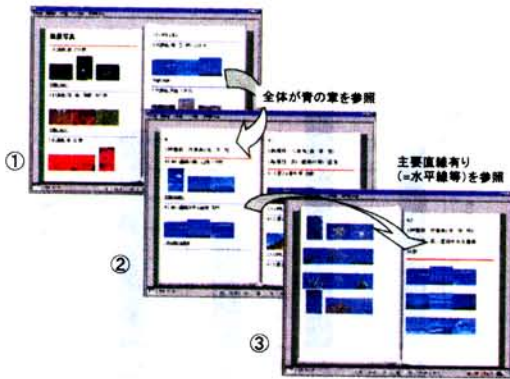


図8 サイバーブック上での閲覧例
Fig.8 The CyberBook Viewer

構築された分類木はブラウザ左部にディレクトリ風に表示される。個々のノードをクリックすることにより、分類された画像を閲覧したり、次の分類に進むことができる。これらの処理を繰り返して階層的な分類木を構築する。

4.2 閲覧フェーズ

作成した分類木は CyberBook¹⁰を用いて本の体裁で表示する。CyberBook は HTML 文書を電子本として表示する機能である。木構造にリンクされた HTML の場合にはルートの HTML を先頭ページに、それにリンクされたページを深さ優先で順次配置するため、現実の本に近い順序で情報を配置できる。この機能を利用し、分類の最上位階層を本の章、以下の階層をそれぞれ節とみなし、各章、節を代表画像数枚（実験では 3 枚）を用いて表現する（図 8 参照）。代表画像は代表分類法では基準に最も近い画像を、クラスタリング法ではクラスタリング中心に最も近いものを選んだ。最下位のレベルでは見開き 2 ページ当たり 30 枚程度を目安に集合に含まれる全画像をリスト表示する。各章や節へは URL に基づくリンクにより行き来できる。

5. 評価

5.1 評価の目的

提案手法により、2.に述べた二つの問題点が如何に改善されたかを評価する。

(1) 画像の全体が把握できない点に関して

内容検索の場合にはそもそも全体を概観することは目的ではないので、この問題が生じる。階層的な目録を作成し、それを上位から代表的な画像数枚で表現することで一般的には内容検索に比べて概観することは容易になる。ただし、その場合にも人間にとって理解し易い、つまり良い分類が作成できたかどうかで画像

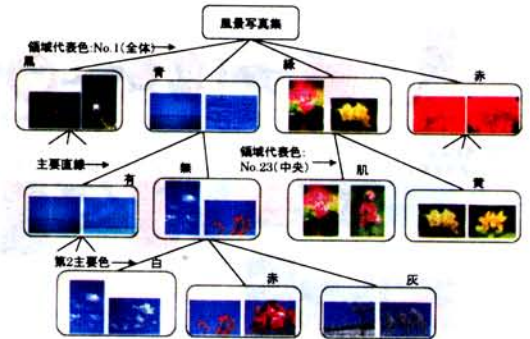


図9 構築した分類木の例
Fig.9 Example of the Actual Classification Tree

表 2 作成した分類木の詳細
Table 2 Details of the Constructed Tree

画像の種類	クリップアート	風景写真
画像数	900	1283
階層の最大値	3	4
中間ノード数	10	11
リーフノード数	38	56
全ノード数	48	67
最下位ノードの平均画像数	23.6	22.9
使用した分類観点	代表形状, 領域代表色, 形状クラスタリング	領域代表色, 第n主要色, 代表直線, アスペクト比

全体を巧く概観出来るかどうかが決まる。概観し易い分類が作成できたかは、各分類毎に人間の考える正解集合に近い画像が含まれているか、及び各階層の代表画像としてどのような画像が提示されているかが重要な要素と考えられる。本論文では前者について良い分類ができたかを再現率で評価することを試みた。後者の課題については今後検討が必要である。

(2) 所望の画像を発見するためにはキー画像の指定が必要で手間がかかる点に関して

キー画像が指定しにくいことにより、結果的に欲しい画像を見つけるのに手間がかかることが問題である。内容検索ではキー画像が必須であるのに対して、画像目録の場合は予め作成した分類に従って、欲しい画像が含まれているであろう部分を探索して行くことで欲しい画像を見つける。つまり、検索において必須であったキー画像の指定は目録の場合は分類に沿った探索で代替される。その際、閲覧者の希望する分類に必ずしも一致した分類とはなっておらず、結果的に多くの分類を見なければ目的の画像を見つけられないケースが考えられる。

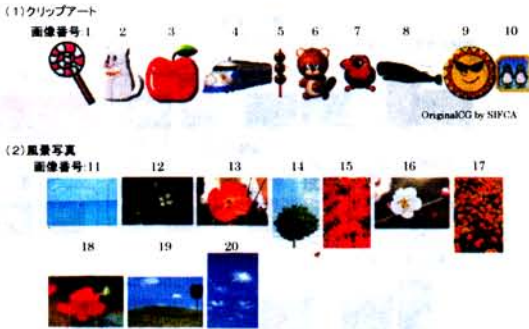


図 10 評価に使用した目的画像

Fig.10 Target Images used for the Evaluation

そこで最終的に欲しい画像を見つけるまでに、どの程度の手間がかかるかを参照画像数という尺度で計測し、欲しい画像を見つけるまでの手間を比較評価する。

なお、以下で被験者とは分類木の構築者とは別の人間である。

5.2 評価に使用した画像目録

素材としてはクリップアート集 900 枚¹⁵⁾、風景写真 1283 枚¹⁶⁾の画像を用いた。分類の構築は本機能に熟知した人間が行い、構築者が最も分かりやすいと判断した基準に従って行った。分類の粗さは目録としての表示を考慮し、最下位ノードの画像数が平均で 30 枚以下となるよう分類した。構築した目録の諸元を表 2 に、分類木の実現例を図 9 に示す。クリップアートの場合、形状が重要な分類要素であることから分類の階層は代表形状、領域代表色、形状クラスタリングと続く 3 階層で、全ノード数 48 個、最下層は平均画像数約 24 枚となった。風景写真では画像全体から受ける印象を重視し、領域代表色、第 n 主要色、代表直線等を用いることで 4 階層に分類した。分類に要した時間は双方とも約 2 時間であった。CyberBook 上には中間の階層については各 3 枚の代表画像を、また最下層については全画像を一覧で表示した。

5.3 再現率評価

人間が類似と判断した画像が分類結果に如何にもれなく含まれるかを再現率により評価する¹⁾。具体的には図 10 に示す計 20 種類の目的画像に対する正解集合を画像処理に全く知識のない被験者に選んでもらい、各目的画像が属する分類木の最下位ノード、及び最上位の階層²⁾を対象に再現率を評価した。実験は複合分類法とクラスタリング法の両者について行い、結果を比較した。クラスタリング法の分類観点は、複合分類法に出来る限り合わせることとし、クリップアートで

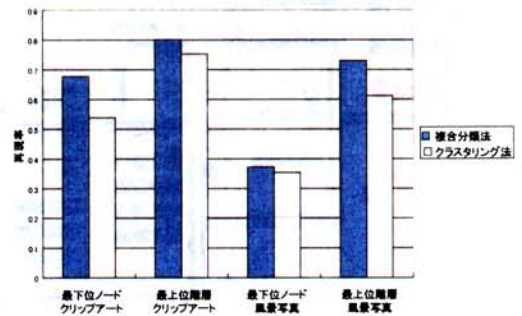


図 11 再現率の評価

Fig.11 Evaluation of Recall Rate

は形状、代表色、風景写真については代表色を用い、最下位ノードに含まれる画像数の平均が約 30 個となる分類木を作成した。評価結果を図 11 に示す。

最下位ノードに着目するとクリップアートの場合、複合分類法での平均 67%の再現率はクラスタリング法の 54%に比べて優位であり、30 枚程度の画像集合中に欲しい画像の多くを含んだ分類を構築できたことを示している。これに対し、風景写真の場合は複合分類法がやや勝っているものの、ほぼ同程度であった。これは複合分類法においても色が主要な分類観点となっており、同じ個数程度に分類を行うと最下位ノードは似通った結果になるためである。ただし、最上位の階層では複合分類法の再現率が高く、トップダウン的に全体を俯瞰する目録では優位であると言える。

なお、今回は再現率に関する評価を行ったが、5.1 で述べたように代表画像の選択方法も目録の概観にとって重要な要素であり、今後は両面からの検討が必要である。

5.4 参照画像数の評価

所望の画像を如何に効率良く見いだせるかは探したい画像がある程度イメージ出来ているような場合、特に重要である。目的画像を発見できるまでに目を通した画像数(参照画像数と呼ぶ)を、複合分類法で作成した目録と内容検索手法とで比較評価した。

5.4.1 内容検索機能

比較対照とする内容検索機能は、特徴量として画像全体の色相、クリップアートに関する全体形状、風景

*1) 適合率は本分類のように結果集合の個数が決まっている場合には正解個数の大小に依存する値であるため、評価からは除外した。

*2) 最上位の階層とはルート直下の第一階層。その配下にあるノードに含まれる全画像数を対象画像とした。

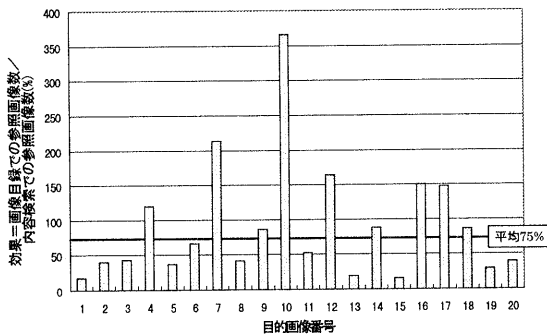


図 12 参照画像数の比較

Fig.12 Comparison of the Number of Referred Images

写真に関する直線情報（長さ、傾き、本数）をもつ。これらは分類で使用したのと同じである。キー画像とデータベース内画像の間でユークリッド距離に基づく距離計算を行い、類似順に上位 k 件（実験では 70 件）を返すことができる。どの特徴を重視するかは検索時に指定することができる。検索においてはデータベース内の画像から数枚（実験では 6 枚）をランダムに表示させ、それらの中から希望する画像に近いものをキー画像として選ぶか、検索結果として表示された画像の中からキー画像を選ぶことができる。

5.4.2 参照画像数

内容検索では、ランダムに画像を表示させ、その中から目的画像に近いものを選択し、それをキーとして検索を行い目的画像を発見する。画像目録では分類をトップダウンに見て行き、目的の画像が現れた時点で終了とする。いずれも所望の画像が見つかる迄に画面上で被験者が目を通した画像の数を参照画像数とする。

なお、検索の失敗や目録のたどり方の誤り等の場合にはやり直しが生じ、両者とも参照画像数が多くなる。

5.4.3 評価結果

図 10 の目的画像に対する参照画像数を評価した。検索に関してはランダム表示による揺らぎを平準化するために計 3 回の平均をとっている。計 20 種の目的画像に対する参照画像数を図 12 に示す。図 12 は内容検索での参照画像数を 1 としてあり、グラフの低い程、画像目録が優れていることを示す。結果は目的画像によりバラツキが見られる。目的画像 7, 10, 17 等は画像集合中に類似した画像が多く含まれるため、少ないランダム表示で近い画像が見つかり、内容検索に有利となっている。逆に類似した画像が少ない 1 や色により明確に分類できる 13, 15 では目録が有利である。今回の実験では全 20 例中、14 例で目録が優位であり、平

均すると約 75% の参照画像数で所望の画像を発見できた。なお、本評価は 1000 枚程度の比較的小規模の画像集合に関する実験である。数万枚規模の画像の場合にはランダムな表示で目的画像に近いものを見出すことはさらに困難になり、目録がより有利になると考えられる。

さらに参照画像数には陽に現れていないが、検索の場合には目的画像に近いキー画像をランダムな表示から選ぶ必要があり、全画像に丹念に目を通す必要があるが、目録の場合には分類を直感的に捉えることができ、同一参照数であっても見る手間は少なかった。

5.5 問題点と考察

評価実験を通して、本手法には以下のような問題があることが分かった。これらについては今後の検討が必要である。

(1) 風景写真での再現率が低い

5.3 で述べた再現率に関しては、クリップアートと風景写真の再現率を比較すると、全てのケースで風景写真の再現率が低い。特に最下位ノードでの再現率は 0.4 程度と最も低い。このことから風景写真の場合には分類の良さを上げるために画像内容をより良く反映した特徴量を追加することが必要である。本論文で述べた色、直線以外に模様 (texture)、構図 (structure)、さらには (4) で述べるような画像中のオブジェクトにもとづく詳細な分類観点を適用する必要がある。

(2) 予め決められた分類と違う観点からの分類・閲覧が困難

分類木を一つしか持たないことから、例えば色相-形状の順で分類されている際に、「すべての同一形状の」オブジェクトを探索する場合は、すべての色相階層ごとに同一形状に分類されるオブジェクトを閲覧していかなければならない。つまり静的に構築された分類から外れる画像を探すことは不得手である。この問題を解決するためには分割された画像集合に対し、複数のインデックス（例の場合には色相-形状の順、及び形状-色相の順）を持たせるといった機能追加が必要である。より一般的には分類のリーフノードに着目し、それらの組み合わせることで、自由なビューを定義する方法等が考えられる。

(3) 特定画像をキーとした画像発見が困難

代表分類法は予め決めた分類項目に沿って画像を分類する。そのため、内容検索では可能な「特定画像に色相が似ている他の画像を全集合から探す」といったような動的な画像発見は不得手である。画像分類が本の目次に相当し、画像集合の全体を俯瞰することに適していること、検索は索引に相当し、種々の観点から

特定画像に近いものを検索できることを考えると今後は両者の利点を組み合わせた検索手段が有効であると考えられる。

(4) 自然画像等、複数オブジェクトを含む画像に対して適用できない

実験に用いたクリップアートのケースでは、クリップアート = オブジェクトであったが、風景写真等の自然画像を対象とする場合には一枚の写真から(雑音を含む)複数のオブジェクトが抽出される³⁾。それら雑音を含むような複数オブジェクトを対象とした分類は本稿で提案した手法では不十分である。まずは抽出されたオブジェクトの内、どれを有意な分類候補とするかを決定しなくてはならない。かつ選択された複数オブジェクトの組み合わせを考慮した新たな分類観点を用意する必要があり、今後の検討が必要である。

(5) 電子本による閲覧では大量画像を一度に見ることができない

本稿では目次の概念に沿った一覧表示に適すると報告されている CyberBook をブラウザとして採用した。評価実験では目録のイメージが連想し易い、見方に対する詳細な説明無しでも被験者が簡単に探索できる等の実用的な利点はあったが、電子本では一度に見ることのできる情報(画像枚数)に限界がある、さらに階層間の関係を視覚的に理解しにくいといった問題がある。階層関係を強調しつつ、大量の情報を閲覧する方法として 3D を用いた Cat-a-Cone¹⁷⁾、納豆ビュー¹⁸⁾をはじめとする多くの手法が提案されており、今後これらの可視化手法の適用も考慮する必要がある。

6. おわりに

大量の画像情報を複数の分類観点に沿って分類することで、階層的な分類木を構築し、それを日常なじみの深い本の形式で目録として提示する手法を提案した。分類の観点としては画像内領域毎のカラーヒストグラム、オブジェクトの形状、及び画像内の主要直線を利用し、代表例を用いた分類とクラスタリングを併用する複合分類法を実装した。インタラクティブな分類支援機能のもとで階層的な分類をトップダウンに作成して行き、最終的な結果を CyberBook 上に本の章・節の形式で提示する。クリップアート、風景写真を対象とした評価により、(1) 複合分類法により、単純なクラスタリングに比べ質の高い分類を作成できること、(2) 20 種類の目的画像を対象とした今回の実験では、所望の画像を内容検索に比べて約 75%の手間で発見できたことを示した。

今後は、さまざまな画像集合に本システムを適用し

ていき、構築した分類木の情報を解析することで代表的な分類パターンが抽出できるのではないかと考えている。さらに分類と検索の有機的な結合にチャレンジしたい。

謝辞 本研究に当たり、CyberBook 上での目録の表示方法について貴重なアドバイスをいただいた NTT サイバソリューション研究所の小澤 英昭 主任研究員、並びに勝田 亮 社員に深謝します。また、多くの有益なコメントを戴いた査読者の方々に感謝します。

参考文献

- 1) 串間, 赤間, 紺谷, 山室: 色や形状等の表層の特徴に基づく画像内容検索技術, 情報処理学会論文誌: データベース, Vol.40, No. SIG3(TOD1), pp.171-184 (1999).
- 2) 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, 第 10 回データ工学ワークショップ (DEWS'99), 2B-1 (1999).
- 3) 串間, 赤間, 紺谷, 木本, 山室: オブジェクトに基づく高速画像検索システム: ExSight, 情報処理学会論文誌, Vol.40, No.2, pp.732-741 (1999).
- 4) 波多野, 亀井, 田中: 多段階自己組織化マップによるビデオ映像記述支援と類似シーン検索, 情報処理学会論文誌, Vol.39, No.4, pp.933-942 (1998).
- 5) T. Kohonen: Self-Organizing Map, Proc. IEEE, Vol.78, No.9, pp.1464-1480 (1990).
- 6) 仲川, 片山: 画像データベースのためのデータマイニング法の拡張, 98-DBS-115, 情報処理学会研究報告, Vol.98, No.34, pp.71-78 (1998).
- 7) P. Lipson, E. Grimson, P. Sinha: Configuration Based Scene Classification and Image Indexing, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.1007-1013 (1997).
- 8) W. Y. Ma, B. S. Manjunath: A Texture thesaurus for Browsing large aerial photographs, ECE Report #96-10, University of California, Santa Barbara (1996).
- 9) 岩崎, 両角: 大規模データベースへのアクセス手段の提案, 99-DBS-119, 情報処理学会研究報告, Vol.99, No.61, pp.42-48 (1999).
- 10) 小澤, 上野, 鈴木, 外村: CyberBook™: 本型ビューワで読む SGML/HTML ドキュメント, デジタル・ドキュメント・シンポジウム '98 論文集, pp.47-56 (1998).
- 11) 岩波情報科学事典, 岩波書店, pp.211 (1990).
- 12) 新版情報処理ハンドブック, 情報処理学会, pp.1292-1293 (1995).
- 13) 小西, 赤間: 関連情報連想システムにおける検索観点自動提示方式, 第6回データ工学ワークショップ (DEWS'95), pp.39-46 (1995).
- 14) 西村, 河野, 長谷川, 重み付き相関ルール導出アルゴリズムによる WWW データ資源の発見, 第7回データ工学ワークショップ (DEWS'96), pp.79-84 (1996).
- 15) インプレス社, Internet ホームページ用素材集.
- 16) メディアリパブリック社, ImagePhotoCatalogue, Vol.1, Vol.4.
- 17) M. A. Hearst, C. Karadi: Cat-a-cone: an interactive interface for specifying searches and

viewing retrieval results using a large category hierarchy, Proc. of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp.246-255(1997).

- 18) 塩澤, 西山, 松下:「納豆ビュー」の対話的な情報視覚化における位置づけ, 情報処理学会論文, Vol.38, No.11, pp.2331-2342(1997).

(平成 11 年 9 月 21 日受付)

(平成 11 年 12 月 27 日採録)

(担当編集委員 有澤 博)



串間 和彦 (正会員)

1980 年京都大学工学部電子工学科卒業。同年日本電信電話公社 (現 NTT) 入社。知識処理用プログラミング環境の研究, 大規模クライアントサーバシステムの実用化等を経て, 現在はマルチメディアデータベースの研究開発に従事。



佐藤 路恵 (正会員)

1995 年東北大学工学部電気工学科卒業, 1997 年同大学院電気・通信工学専攻修士課程修了, 同年日本電信電話株式会社入社。以来, マルチメディア情報整理の研究開発に従事。



赤間 浩樹 (正会員)

1988 年東海大学理学部情報数理学科卒業。1990 年同大学院理学研究科数学専攻修士課程修了。同年日本電信電話株式会社入社。以来, インテリジェント・ネットワーク向け DBMS の開発, ニュース・オン・デマンドの研究開発等を経て, 現在はマルチメディア情報検索の研究開発に従事。人工知能学会, 日本ソフトウェア科学会, ACM 各会員。



山室 雅司

1985 年早稲田大学理工学部数学科卒業。1987 年同大学院数学専攻修士課程修了。1990 年コロンビア大学大学院電気工学専攻修士課程修了。1999 年博士 (工学) 早稲田大学。1987 年日本電信電話株式会社入社。以来, ネットワークオペレーション情報モデル化・ビジュアル化, データベース設計法の研究に従事。現在, マルチメディア情報検索の研究に従事。1994 年電子情報通信学会学術奨励賞受賞。電子情報通信学会, 日本ソフトウェア科学会, IEEE-CS 各会員。