

CAT:大量画像の一覧可視化と詳細度制御のための GUI

五味 愛[†] 宮崎 麗子[†] 伊藤 貴之[†](正会員)
ジ ア リ^{††}[†] お茶の水女子大学大学院
^{††} ペンシルバニア州立大学

あらまし 本論文では、大量画像のクラスタリング手法、およびその一覧表示と詳細度制御を持ち合わせた可視化手法 CAT(Clustered Album Thumbnails) を提案する。CAT は、前処理として大量画像を多段階にクラスタリングし、各クラスタから代表画像を選出する。そして構築された階層構造を、階層型データ可視化手法「平安京ビュー」を用いて表示する。CAT の特徴は、ズーム率に合わせた詳細度制御を設けている、という点である。ズームアウト時には各クラスタの代表画像を表示し、ズームイン操作によって局所的に各々の画像を表示する。この操作により、人間の視覚能力とディスプレイの解像度に応じて表示枚数を調節し、大量画像中の注目部分をスムーズな操作で可視化できる。キーワード：階層型データ可視化、画像閲覧システム、詳細度制御、長方形充填。

Summary This thesis proposes CAT (Clustered Album Thumbnail), a technique for clustering and browsing large number of images. It also provides a user interface for controlling the level of details.

As a preprocessing, CAT first hierarchically clusters images and selects representative images for each cluster. And then, CAT visualizes the tree structure applying a hierarchical data visualization technique HeiankyoView. As a characteristic of CAT, it provides a graphical user interface for the zooming operation to effectively browse images. It selectively displays representative images while zooming out, or individual images while zooming in, by the mouse operation. This feature realizes high frame rates, and display of adequate number of images.

Key words: Hierarchical data visualization, Image browser, Level-of-detail control, Rectangle packing.

1. はじめに

デジタルカメラの普及により近年、簡単にデジタル画像を作成することができるようになった。それゆえ、画像を扱うさまざまな分野において、大量の画像が蓄積されている。そのため、それら大量画像の分類、閲覧、検索のための画像閲覧システムの必要性が増している。

我々はしばしば、個人写真を扱った画像閲覧ソフトや、インターネット上による画像検索サイトなどで、大量の

画像を目にすることがある。これらのシステムでは、ファイル名順、撮影時刻順、閲覧回数順、といった1次元的な順列に基づいて画像を整列表示することが多い。それに対して我々は、大量の画像を分類保存している機会が多い。そのため、順列よりも分類に基づいて大量の画像を閲覧できるシステムが有用である、という場面も多いと考えられる。これは例えば、個人写真の多くは計算機上でフォルダごとに分類されていること、また医用画像やリモートセンシングなどの分野で画像の自動分類技術が多用されていること、などからも明らかである。

我々は、計算機に蓄積された大量画像をどのように探索するか、といった点について予備調査を行い、以下のような結果を得た。

“CAT: A Graphical User Interface for Visualization and Level-of-Detail Control for Large Scale Image Collections” by Ai GOMI, Reiko MIYAZAKI, Takayuki ITOH, (Member), (Graduate School of Humanities and Sciences, Ochanomizu Universityshi), and Jia LI, (The Pennsylvania State University).

- ファイルの探索操作に関する質問について、回答者の83%が、ファイルシステムのフォルダを上位から探索する、あるいは特定フォルダに保存されている画像ファイルを順に閲覧する、などの形でファイルシステムの探索 GUI を利用している、と回答した。
- ファイル探索の手がかりとして何が重要か、という質問に対して、回答者の50%がフォルダ名などのキーワードと回答し、50%が撮影日時や撮影場所などのメタ情報と回答した。
- 画像内容（色分布や構図など）はファイル探索の手がかりとして有用か、という質問に回答者の100%が、有用だがキーワードやメタ情報より重要度が低い、と回答した。

この予備調査結果を踏まえて本論文では、キーワードやメタ情報だけでなく画像内容にも基づく画像のクラスタリング手法、およびファイルシステムと同様に階層構造化された大量画像群の一覧表示と詳細度制御のためのインタフェースを持ち合わせた、可視化手法 CAT(Clustered Album Thumbnails) を提案する。

各画像に1個以上のキーワードが付与されていると仮定した上で、まず CAT は前処理として、キーワードと画素情報から2段階クラスタリングを適用し、画像群を階層的に分類するとともに、各クラスタに対して代表画像を選択する。また、キーワードにおけるクラスタリングを行う際に、2種類のクラスタ構築方法を搭載し、画像閲覧の用途に合わせていずれかを選択できるようになっている。1つ目の方法は、GUIでキーワードを選択させ、特定キーワードを含む画像群からクラスタを構築する方法、2つ目の方法は、画像群全体からクラスタを構築する方法である。続いて CAT は、「平安京ビュー³⁾⁴⁾という大規模階層型データ可視化手法を用いて、図1に示すように、各クラスタを示す長方形領域ごとに画像群を表示する。さらに CAT では、ズームアウト時には長方形領域に代表画像をマッピングして表示することにより、大量画像全体のサマリーとなるような表示結果を提供する(図1(上)参照)。そしてズームイン操作に伴って、CAT は低階層クラスタの代表画像を表示する(図1(中)参照)。さらにズームイン操作を進めることで、CAT は低階層クラスタの各々の画像サムネイルを表示する(図1(下)参照)。このように、ズーム率に基づく詳細度制御によって CAT は、クラスタリング結果全体の概略表示と、クラスタリング内部の局所表示を、シームレスに実現する。

大量画像の一覧表示ブラウザに関する先行研究¹⁾⁵⁾¹¹⁾と比較して、CAT は階層化された画像群に対するズーム

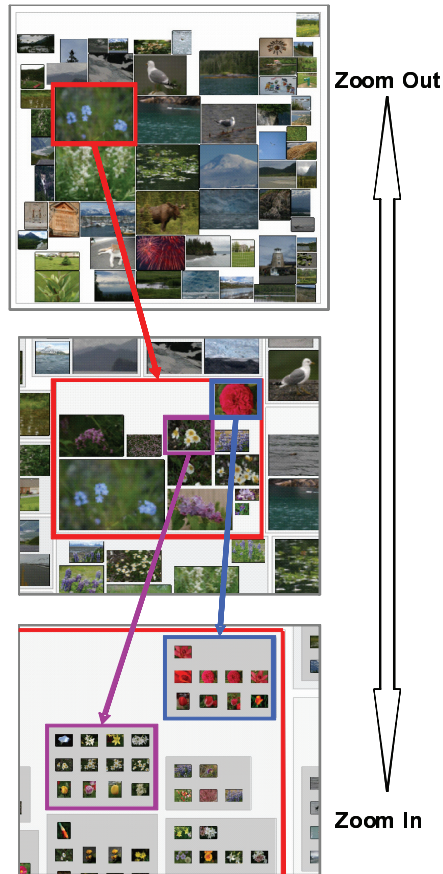


図1 CAT の概観。
Fig. 1 Overview of CAT.

操作によって、直観的に画像を絞り込みながら閲覧できるだけでなく、人間の視覚能力とディスプレイの解像度に応じた適切な枚数で画像を表示できる、という点が特徴としてあげられる。また、この詳細度制御において重要な点は、各クラスタを表現する領域が、代表画像の縦横比に同等な形状の長方形で表現される、という点である。これにより、画像本来の縦横比とほぼ同等な形状で代表画像を表現できる。先行研究の中でも PhotoMesa¹⁾は、CAT と同じくクラスタごとに長方形領域を生成して画像を表示する手法であるが、これに対して CAT は、クラスタを表現する長方形領域の形状を制御できること、不均一な深さをもつ階層構造にも対応できること、などの点で優位性があると考えられる。

2. 関連研究

2.1 画像検索

有名な画像検索エンジンの多くは、ウェブページ上にあるテキスト記述を参照して画像に注釈をつけ、その注

積に基づいて検索結果を提供する。しかし、これらのテキスト記述から、必ずしも画像の内容を適切に解説する注釈が得られるとは限らない。よってこの仕組みでは、十分な検索結果を得られるとは限らない。

そこでこの問題を解決するために、画像内容検索(CBIR)の研究が取り組まれている。CBIR 関係技術の多くは、画像の特徴量に基づいて画像内容を検索する。参考文献¹⁰⁾では、CBIR において色からの特徴量、続いて周波数特性から得た特徴量が頻繁に使われていると報告している。CAT においても、色と周波数特性から得た特徴量を画像クラスタリングに用いている。

一方でここ数年、画像内容に基づく自動画像注釈の研究が活発であり⁸⁾、画像検索エンジンの問題点を解決しうる技術として注目されている。CAT では画像にキーワードが付与されていることを前提としているが、自動画像注釈技術が充実すれば、画像へのキーワードも自動付与できることが期待される。

2.2 大量画像ブラウザ

ウェブサイトの画像検索エンジンなどの検索結果を表示するインタフェースの多くは、検索結果である画像群を、単純に類似性の高い順に並べて表示する。それゆえ、この種のインタフェースでは、ユーザの望む画像を常に効果的に素早く見つけられるとは限らない。この問題を解決する一例として Kang らは、さまざまな質問入力機能や詳細表示機能を搭載した、大量画像の検索用ブラウザを提案している⁶⁾。

大量画像ブラウザに関する研究の多くは、画像間の類似性を考慮した画像サムネイルのレイアウトに着目している。その中でも初期の多くの手法では、多次元尺度構成法(MDS)や主成分分析(PCA)を活用し、低次元ユークリッド空間に画像サムネイルを配置する。

木構造やグラフ構造などを対象とした情報可視化技術のいくつかは、大量画像ブラウザの研究にも適用されている。Jankun-Kelly らは、局所拡大表示型のグラフレイアウト技術を大量画像ブラウザに適用している⁵⁾。Bederson らは、2.3 節で紹介する Quantum Treemap²⁾という階層型データ可視化手法を採用した大量画像ブラウザ PhotoMesa¹⁾を提案している。

画像注釈に基づく大量画像の閲覧手法として、Yang らは SIB(Semantic Image Browser)¹¹⁾を提案している。SIB では画像の木構造化やグラフ化は実現しておらず、Yang ら自身も画像の階層化や詳細度制御表示を今後の課題と位置づけている。

これらの関連手法と比べて CAT は、任意の階層構造に基づいて画像群を表示できるので、より直観的な画像探索が実現できると考えられる。

2.3 階層型データの可視化

階層型データの情報可視化手法の中でも有名なものに、木を表示する手法と、画面空間充填による手法があげられる。CAT は画面空間充填による手法に属する。

CAT は画面空間充填による階層型データ可視化手法「平安京ビュー」^{3),4)}をベースにした手法である。「平安京ビュー」では図 2 のように、入れ子型の長方形の枠で階層を表し、葉ノードをアイコンで表示する。「平安京ビュー」は階層構造の最下位に属する葉ノード群を、クリックしやすく等しいサイズのアイコンとして、しかも互いに重ならずに表示できる、という特徴がある。このアイコンを画像サムネイルに置き換えることで、CAT は大量の画像を、等しいサイズで互いに重ならずに一覧表示できる。また「平安京ビュー」を用いることで、複数のクラスタに属する画像群を同時に表示できる。よって、従来のファイルシステムと同様な表示法では困難であった、周辺クラスタ(類似クラスタ)との画像の比較が可能となる。

2.2 節で紹介した PhotoMesa¹⁾が用いている Quantum Treemap²⁾も、画面空間充填による階層型データ可視化手法であり、「平安京ビュー」に類似した特徴を持つ。文献¹⁾では Quantum Treemap と「平安京ビュー」の比較評価結果を掲載している。この比較評価結果において「平安京ビュー」は、類似する階層型データのレイアウトの安定性と、長方形領域の縦横比において、よい結果を得ている。また Quantum Treemap と違って「平安京ビュー」は、不均一な 2 段階以上の階層を有する任意の階層構造を表現できる。これらの長所は、そのまま CAT の PhotoMesa に対する優位性としてあげられる。

なお、文献^{3),4)}で述べるように「平安京ビュー」は、画像配置の理想位置を記述したテンプレートによって、長方形の位置を操作する機能を有する。これを活用することで、視覚的類似性の高い画像を画面上の近い位置に配置する、といったことも実現可能である。

3. CAT の処理手順

3.1 概要

CAT はまず、以下の 2 段階処理によって画像をクラスタリングし、多段階の階層構造を構築する。

1. まずキーワードに基づいて画像をクラスタリングする。この処理結果として得られるクラスタを、高階層クラスタと称する。
2. 各々の高階層クラスタに属する画像を、画素情報でクラスタリングする。この処理結果として得られるクラスタを、低階層クラスタと称する。

CAT は上記の手順で階層構造化された画像群を、「平安京ビュー」と同様に、入れ子状の長方形領域の集合で

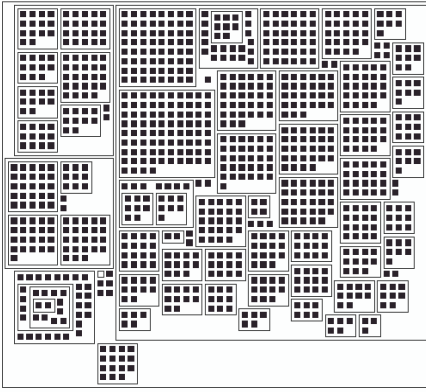


図2 「平安京ビュー」の概観。
Fig. 2 Overview of HeiankyoView.

表現する。そしてズームアウト時には各クラスタの代表画像を長方形領域にマッピングして表示し、ズームイン時には各クラスタ内の各画像を表示する。

なお CAT では、以下の 2 種類の高階層クラスタ構築方法を搭載し、画像閲覧の用途に合わせて両者のいずれかを選択できるようになっている。

高階層クラスタ構築方法 (1): 全く同一なセットのキーワード群を有する画像を同一クラスタに格納し、GUI でキーワードを選択させる方法。特定のキーワードで絞り込みながら画像を探索する目的に向いている。

高階層クラスタ構築方法 (2): キーワード間距離を算出し、距離の近いキーワードを有する画像群を同一クラスタに格納する方法。各クラスタの保有画像数が均一になるようにクラスタリングを行うことで、クラスタの視認性を高め、画像群の階層構造全体の見渡しをよくすることから、画像群全体から注目画像を探す目的に向いている。

以下、CAT を構成する各処理工程について論じる。

3.2 高階層クラスタ構築方法 (1)

この方法では、初期画面として、キーワードと一緒に代表画像を表示させたキーワード選択画面を設ける (図 3)。このキーワード選択画面において、ユーザはマウス操作によって 1 個以上のキーワードを選択するものとする。

この方法では、全く同一のキーワード群を有する画像を同一クラスタに格納する。このとき、クラスタに属する画像数に制限を設けない。ここで、クラスタ C_i のキーワードセット W_i を、以下のように示す。

$$W_i = \{w_{i,1}, \dots, w_{i,m_i}\}, w_{i,j} \in V \quad (1)$$

ここで V は画像群に用いられる全てのキーワード、 $w_{i,j}$ は W_i の j 番目のキーワード、 m_i はクラスタ C_i のキーワード数を示す。また、ユーザが選択したキーワードセッ

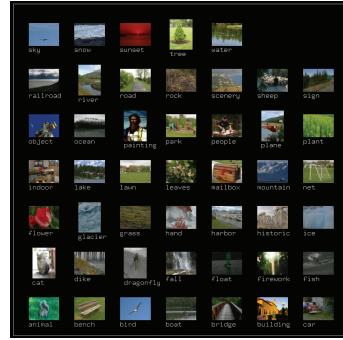


図3 キーワード選択画面。

ト S を、以下のように示す。ここで M は、ユーザが選択したキーワードの数を示す。

$$S = \{s_1, \dots, s_M\}, s_i \in V \quad (2)$$

ユーザがキーワードを選択すると CAT は、動的にユーザが選択したキーワードを含む画像の部分木を生成する。 W_i を構成するキーワードの最低 1 個が、 S を構成するキーワードの最低 1 個と一致した場合、クラスタ C_i を部分木に取り入れる。さもなければ、 C_i は部分木に取り入れられない。以上の方法により構築された部分木による階層構造を画面表示することにより、少数に絞られた画像群が表示の対象となる。よってこの方法を用いることで、画像探索所要時間の短縮が期待できる。

ただし、キーワードの総数が膨大になった場合、各画像に多数のキーワードが付加されている場合や、あるいはユーザが多数のキーワードを選択した場合においては、クラスタ数が極めて多くなってしまふことが懸念される。解決の一手法として、3.3 節に述べるクラスタの再帰的合併法を組み合わせるなどして、クラスタを併合することが考えられる。今後の課題として検討したい。

3.3 高階層クラスタ構築方法 (2)

この方法では最初に、文献⁹⁾と同様な手法で、キーワードに基づく画像間の距離を定義する。ここで画像 X_i のキーワードセットを W_i とし、以下に示す。

$$W_i = \{w_{i,1}, \dots, w_{i,m_i}\}, w_{i,j} \in V \quad (3)$$

なお m_i は、画像 X_i に割り当てられたキーワード数である。このとき、2 画像 X_i と X_j の距離を、以下の式⁹⁾によって算出する。

$$D_w(W_i, W_j) = \frac{1}{2m_i} \sum_{k=1}^{m_i} \overline{d}_w(w_{i,k}, W_j) + \frac{1}{2m_j} \sum_{l=1}^{m_j} \overline{d}_w(w_{j,l}, W_i) \quad (4)$$

ここで、 D_w は 2 つのキーワードセット W_i と W_j 間の距離、 $\overline{d}_w(w_{i,k}, W_j)$ はキーワード $w_{i,k}$ とキーワードの

セット W_j の間の距離を示している。我々の実装では、ソフトウェアパッケージ WordNet Similarity¹⁾ によってキーワード間の距離を算出している。

続いて、この方法では、画像間の距離に基づいて、画像をクラスタリングする。ここではボトムアップ形式のリンケージクラスタリングアルゴリズムを使用する。ただし CAT では、各々の高階層クラスタに属する画像数の均等化を試みるために、クラスタリング過程において以下の処理も同時に行う。

ここで、 $s_1 < s_2$ となるような閾値 s_1 と s_2 を定義する。あるクラスタに属する画像数が s_1 よりも少なかったら、CAT では積極的に他のクラスタと合併させる。一方、あるクラスタに属する画像数が s_2 よりも多かったら、そのクラスタはこれ以上の合併を避ける。そして、クラスタ間の最小距離値が閾値を超えるまで、距離が最小であるクラスタを再帰的に合併する。

3.4 低階層クラスタ構築方法

CAT では、キーワードに基づいて生成された各クラスタ内の画像を、画素情報に基づいて再度クラスタリングする。我々の実装では、色情報および周波数情報に基づく特徴量から、各画像の特徴ベクトルを算出し、そのベクトル間の余弦に基づいてクラスタを生成する。色情報に基づく特徴量として、現時点で我々は、画像空間を格子分割し、各々の領域における代表画素値を算出し、この結果を多次元ベクトル化して用いている。周波数情報に基づく特徴量として、現時点で我々は文献⁷⁾と同様な手法を用いている。この手法ではウェーブレット変換を施して高周波成分を算出し、その高周波成分が非零である画素を抽出する。続いて抽出された画素に対し、高次局所自己相関関数を用いて特徴ベクトルを得る。参照点 r でのエッジ画像の輝度値を I_r とすると、 N 次自己相関関数は、参照点周りの N 個の変位 (a_1, \dots, a_N) に対して、

$$x(a_1, \dots, a_N) = \int I(r)I(r + a_1) \dots I(r + a_N) dx \quad (5)$$

で定義される。次数 N を高々 2 までとし、変位の有無を画素ごとに 2 値化すると、参照点 r の周りの局所的な 3×3 画素の領域におけるパターンは、平行移動により等価な特徴を除くと、全部で 25 個のテンプレートのいずれかで表現される。特徴量は画像の各位置においてこの 25 個のテンプレートに対応する画素の輝度値の積を計算し、それらを画像全体で足し合わせることで求められる。

この 2 種類の特徴ベクトルを併合したベクトルを用いて CAT では、余弦値が大きい順にボトムアップ的に低階層クラスタを生成する。この処理は、2 クラスタ間の

最大余弦値が閾値より小さくなるまで反復される。この過程で前節と同じように、各々の低階層クラスタに属する画像数の均等化を試みる。

3.5 代表画像の選択

クラスタリング完了後に CAT では、高階層クラスタ、低階層クラスタ、の両方に対して代表画像を選択する。現時点で我々は、以下の 2 種類の選択方法を採用している。

1 つ目の手法は、画素情報に基づく手法である。特徴ベクトル空間のクラスタの中心に最も近い画像が、クラスタの平均的な画像であるといえるので、その画像はクラスタの代表として望ましいと判断できる。

2 つ目の手法は、キーワードに基づく方法である²⁾。ここで、キーワード $w_{i,j}$ を持つ画像の数を $k_{i,j}$ とする。このとき $k_{i,j}$ が他よりも大きければ、 $w_{i,j}$ は重要なキーワードと想定できる。よって $k_{i,j}$ は、クラスタ内のキーワードの重要度を表すと考えられる。ここで以下の式により、画像 X_i のキーワードの重要度合計を定義する：

$$K_i = \sum_{j=1}^{m_i} k_{i,j} \quad (6)$$

多くの場合において、 K_i 値が最大である画像は、クラスタの特徴を表していると言えるので、その画像はクラスタの代表画像として望ましいと判断できる。

3.6 階層型画像のクラスタの可視化

続いて CAT では、以下の 3 ステップにより、大量画像を画面空間に配置する(図 4)。2.3 節で述べた通り、この手法は「平安京ビュー」をベースにした手法である。

1. 低階層クラスタに属する画像サムネイルを格子状に配置し、長方形でそれらを囲む。全ての低階層クラスタに、同様の過程を繰り返す。
2. 低階層クラスタを表現する全ての長方形を隙間なく配置し、高階層クラスタを表現する長方形で囲む。全ての高階層クラスタに、同様の過程を繰り返す。
3. 高階層クラスタを示す長方形を隙間なく配置する。

CAT は「平安京ビュー」と同様、文献^{3),4)}に示す以下の値を最小化するように長方形群を配置する。

$$aA + rR + dD \quad (7)$$

ここで a, r, d は、正の定数でありユーザーが指定する。 A は長方形配置における前後間の長方形の面積比、 R は長方形配置における前後間の長方形の縦横比、 D は長方形の実際の位置とテンプレートに描かれた理想的な位置との間の距離である。CAT では変数 R を、理想的な縦横比 (= 代表画像の縦横比) に対する長方形領域の縦横比の悪化の度合い、として計算する。

1) <http://www.d.umn.edu/~tpederse/similarity.html> 参照。

2) この手法は高階層クラスタ構築手法 (2) を用いた場合にのみ有効である。

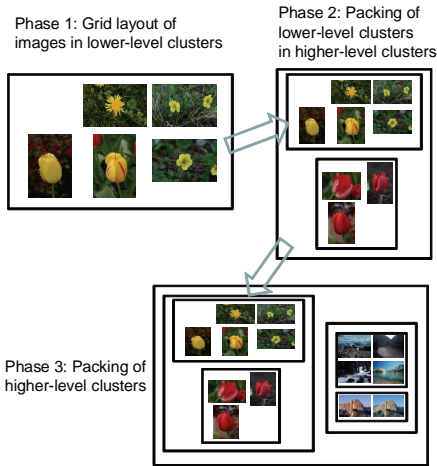


図4 画像配置のための3ステップ.

CATでは、類似度の高いクラスタを、画面上で隣接配置できる。まず2.2節で説明したMDSやPCAに基づく方法で、クラスタの2次元空間上での理想位置を算出する。続いて2.3節で説明した方法で、理想位置の算出結果をテンプレートとして各クラスタの画面上の位置を操作し、類似度の高いクラスタを画面上で隣接配置する。

3.7 ユーザインタフェースと詳細度制御

我々のCATの実装では、マウスの左クリックにズームイン、右クリックにズームアウト、の各操作を割り当てている。また、特定の長方形の背景をクリックすると、その長方形が画面中央に移動する。以上の操作によりCATは、特定のクラスタに対してスムーズなズームイン操作を実現し、また1クリックで簡単にもとのズームアウト状態に戻すこともできる。

また我々の実装では、ズームアウト状態でCATを起動し、まず代表画像だけをハードディスクから主記憶に読み込む。そしてズームイン状態に伴い、フォーカス部位の画像サムネイルだけを読み込む。さらに、表示対象から外れた画像を、随時主記憶から解放する。以上の操作によってCATは、主記憶使用量を節約する。

4. 評価実験

我々は、GNU gcc 3.4 および OpenCV を用いてクラスターリング部を実装し、Visual C++を用いてブラウザ部を実行した。テスト環境には Windows XP を搭載した PC を用い、横 384 画素、縦 256 画素の JPEG 画像 2360 枚を活用した³⁾。以下、10 人の学生被験者による、主観評価のためのユーザテスト結果を示す。

³⁾ <http://www.stat.psu.edu/~jiaili/index.download.html> に、本実験に用いた 2360 枚の画像が公開されている。

表1 主観的概観評価

順位	1	2	3	4	5
クラスタなし	0	0	1	0	9
低階層クラスタ・代表画像なし	0	0	0	10	0
低階層クラスタ・代表画像あり	0	6	3	0	1
2階層クラスタ・代表画像なし	2	2	6	0	0
2階層クラスタ・代表画像あり	8	2	0	0	0

表2 特定画像を探し出す所要時間(秒).

	方法(1)	方法(2)
最短	9	8
平均	47	83
最長	152	308

4.1 主観的な外観評価

まず本実験では、5種類の異なる表示条件を課したブラウザを用意した「クラスタなし」では、全ての画像を、単純に格子状に表示する「低階層クラスタ・代表画像なし」では、画素情報に基づく1階層のクラスタを生成し、代表画像は表示しない。「低階層クラスタ・代表画像あり」では、画素情報に基づく1階層のクラスタを生成し、代表画像を表示する。「2階層クラスタ・代表画像なし」では、キーワードと画素情報に基づく2階層のクラスタを生成し、代表画像は表示しない。「2階層クラスタ・代表画像あり」では、キーワードと画素情報に基づく2階層のクラスタを生成し、代表画像を表示する。

この5種類のブラウザを被験者に数分間操作してもらい、視覚的印象と使いやすさの観点から、1位が最高評価、5位が最低評価として、主観的に順位をつけてもらった。表1に結果を示す。この結果より、2階層クラスターリングおよび代表画像の有効性がわかる。特に代表画像の表示機能に高い評価が集中している。

この実験で被験者から得た意見として、代表画像の存在は、詳細度制御による画像表示サイズの適正化だけでなく、表示速度の観点からも重要である、という意見があった。代表画像を表示しない場合、初期操作で非常に多くの画像を表示するため、表示速度と主記憶使用量の観点で不利であることが、被験者にも実感されていることがわかる。

4.2 特定の画像を探し出す所要時間

続いて、3.2節で提案したキーワード選択画面を伴う方法(1)と、3.3節で提案した全クラスタを表示する方法(2)とで、紙に印刷した特定の画像3枚を被験者に提示し、2360枚の画像の中から探し出す時間を計る実験を行った。

この実験において、代表画像と特定の画像が類似している場合、方法(1)も方法(2)でも、探索時間は短くなる傾向にあった。一方で、方法(2)では、複数のキーワード

表 3 縦横比率の統計結果

縦横比の区間	高階層	低階層
1.0 ~ 1.1	6	10
1.1 ~ 1.2	3	6
1.2 ~ 1.3	18	38
1.3 ~ 1.4	14	41
1.4 ~ 1.5	7	19
1.5 ~ 1.6	2	16
1.6 ~ 1.7	0	12
1.7 ~ 1.8	1	2
1.8 ~ 1.9	1	3
1.9 ~ 2.0	4	4
2.0 ~	5	0
平均値	1.336	1.383

表 4 長方形領域の統計

	$s_1 = 1$	$s_1 = 10$
最小値	9.28	45.43
平均値	81.46	178.44
最大値	833.17	833.17
最大値と最小値の比	89.78	18.34

を連想しやすい画像に関しては、関連する多くのクラスタを閲覧することで、大きな時間を要する場合が多かった。

表 2 では、時間の測定結果を表示した。最短、平均、最長時間をそれぞれの手法において示している。表 2 の結果からもわかるように、画像やユーザによって探索時間は変化するものの、方法 (2) に比べ方法 (1) では、最長時間と平均時間で大幅な探索時間の短縮が見られる。方法 (2) では、すべての高階層クラスタの代表画像の一覧から特定のクラスタを探さなければならないが、方法 (1) では、特定のクラスタをキーワードにおいて絞り込むことで、効率よく探索できることが理由として挙げられる。

4.3 長方形の縦横比

続いて CAT による画面配置結果から、各クラスタを表現する長方形領域の縦横比を算出し、クラスタの代表画像の縦横比にどれくらい近いかを評価した。本実験に用いた全ての画像の縦横比は $4/3$ なので、本実験における最適な縦横比は、全ての長方形領域において $4/3$ である。

表 3 に縦横比の算出結果を示す。ここでは、高階層と低階層クラスタの縦横比を個々に表示した。

高階層、低階層ともに、多くのクラスタの縦横比は、理想値に非常に近いといえる。また高階層クラスタにおいて、式 7 の定数 a を調節することによって、縦横比を改良することが出来る。しかしながら、これによって空領域比が悪化する可能性もある。式 7 の最も望ましい値の設定については、実験的に、そして主観的に議論する必要があると考えられる。

4.4 長方形領域

最後に、全ての長方形領域について面積を計算し、CAT におけるクラスタリング手法が長方形領域の面積均一化に貢献していることを実証した。

CAT のクラスタリング手法は、 s_1 の値を適切に設定することで、小さいクラスタを合併させ、バランスのよいクラスタリング結果を得る。ただし $s_1 \leq 1$ である場

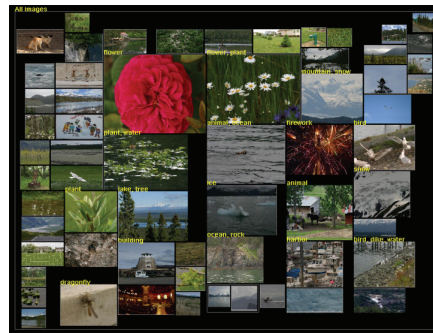


図 5 長方形領域の比較。(上) $s_1 = 1$ 。(下) $s_1 = 10$ 。

合には、クラスタの併合は一切行われぬ。このクラスタ併合によって CAT では、ズームアウト操作によって代表画像の表示サイズが小さくなったときにも、視認できる十分な大きさで代表画像を表示する。

表 4 では $s_1 = 1$ と $s_1 = 10$ の 2 値について、高階層クラスタを構成する長方形領域面積の最小値、平均値、最大値、最大値と最小値の比、の各々を算出した結果を示している。図 5 では、高階層クラスタの代表画像を表示した結果を示している。本研究の目的において、この算出結果は、長方形領域面積の最小値が大きく、最大値と最小値の比が小さいことが望ましい。表 4 からわかるように、我々の実験では $s_1 = 10$ のときのほうが、望ましい長方形領域面積が得られている。この結果から、提案手法によるクラスタの併合が効果的に働いていることがわかる。

5. 結 論

本論文では、キーワードやメタ情報だけでなく画像内容にも基づく画像のクラスタリング手法、およびファイルシステムと同様に階層構造化された大量画像群の一覧表示と詳細度制御のためのインタフェースを持ち合わせた、可視化手法 CAT(Clustered Album Thumbnails) を提案した。また評価実験結果として、CAT の 2 段階クラスタリングおよび代表画像表示が高い主観評価を得たこと、キーワード選択画面を含む手法が画像探索において良好な結果を得たこと、少数クラスタリングの併合処理や長方形領域の縦横比制御処理が良好な数値評価結果を得たこと、を示した。

今後の課題として、何十万もの画像を用いた膨大なデータでの実験を行いたい。また、式 7 の a, r, d の最適な値を求めるための主観評価を行い、よりよい階層構造の構築手法の考察をしていきたい。GUI に関しては、ズームインタフェース機能や focus+context インタフェース機能を備えた可視化技術との融合を加えていきたい。また、画像検索において、代表画像の大きさを均一にした方が検索しやすいという意見もあるため、代表画像の均一化に関する手法を検討中である。

参 考 文 献

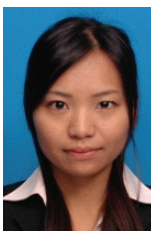
- 1) Bederson B., B., PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps, *User Interface Software and Technology*, pp. 71-80, 2001.
- 2) Bederson B., Schneiderman B., Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies, *ACM Trans. on Graphics*, Vol. 21, No. 4, pp. 833-854, 2002.
- 3) Itoh T., et al., Hierarchical Data Visualization Using a Fast Rectangle-Packing Algorithm, *IEEE Trans. on Visualization and Computer Graphics*, Vol. 10, No. 3, pp. 302-313, 2004.
- 4) 伊藤, 山口, 小山田, 長方形の入れ子構造による階層型データ視覚化手法の計算時間および画面占有面積の改善, 可視化情報学会論文集, Vol. 26, No. 6, pp. 51-61, 2006.
- 5) Jankun-Kelly T. J., Ma K.-L., MoireGraphs: Radial Focus+Context Visualization and Interaction for Graphs with Visual Nodes, *IEEE Information Visualization*, pp. 59-66, 2003.
- 6) Kang H., Shneiderman B., Visualization Methods for Personal Photo Collections: Browsing and Searching in the Photofinder, *IEEE Intl. Conf. on Multimedia and Expo 2000*, pp. 1539-1542, 2000.
- 7) Kubo M., Aghbari Z., Makinouchi A., Image Retrieval by Edge Features Using Higher Order Autocorrelation in a SOM Environment, *IEICE, Trans. on Information and Systems*, Vol. E86-D, No. 8, pp.1406-1415, 2003.
- 8) Li J., Wang J. Z., Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach, *IEEE Trans. on Pattern Analysis and Machine Intelligence* Vol. 25, No. 9, pp. 1075-1088, 2003.
- 9) Li J., A Mutual Semantic Endorsement Approach to Im-

age Retrieval and Context Provision, *ACM SIGMM, Multimedia Information Retrieval*, pp. 173-182, 2005.

- 10) Veltkamp R. C., Tanase M., Content-Based Image Retrieval Systems: A Survey, *UU-CS-2000-34, Institute of Information and Computing Science, University Utrecht*, 2000.
- 11) Yang J., et al., Semantic Image Browser: Bridging Information Visualization with Automated Intelligent Image Analysis, *IEEE Visual Analytics in Science and Technology*, pp. 191-198, 2006.

(2008 年 x 月 x 日受付)

五 味 愛



宮 崎 麗 子

2006 年お茶の水女子大学理学部情報科学科卒業。2008 年お茶の水女子大学大学院人間文化研究科博士前期課程修了。お茶の水女子大学大学院人間文化創成科学研究科博士後期課程在学中。情報処理学会会員。



伊 藤 貴 之 (正会員)

2002 年お茶の水女子大学理学部情報科学科卒業。2004 年お茶の水女子大学大学院人間文化研究科博士前期課程修了。同年ソニー(株)入社。お茶の水女子大学大学院人間文化研究科博士後期課程在学中。情報処理学会会員。



ジ ア リ

1992 年早稲田大学大学院理工学研究科修士課程修了。同年日本アイ・ビー・エム(株)入社。1997 年博士(工学)。2000 年カーネギーメロン大学客員研究員。2003 年京都大学大学院情報学研究科 COE 研究員(客員助教授相当)兼任。2005 年お茶の水女子大学理学部情報科学科准教授。ACM, IEEE CS, 他会員。



1995 年米国スタンフォード大学電気工学専攻修士課程修了。1999 年米国スタンフォード大学電気工学専攻博士後期課程修了。現在米国ペンシルバニア州立大学統計学専攻准教授。IEEE 会員。