

# キーワードの共通性を配置結果に反映した大量画像の一覧可視化手法

安田 理紗 (指導教員: 伊藤 貴之)

## 1. はじめに

近年のデジタルデータの増加に伴い、画像ブラウザに関する研究が活発に発表されている。画像ブラウザにおけるキー技術は大量画像の画面配置手法であり、「構造型」と「非構造型」に大別されると考える。構造型の先行研究として、CAT(Clustered Album Thumbnails)[1]がある。CATは階層構造により画像を管理し、その階層構造に基づいて画像を画面上に配置する手法である。CATの特徴的な点は、階層構造に分類された画像群からクラスごとに選んだ代表画像を表示することで全体像を表し、ズームイン操作により局所的に各画像を表示する、という操作性にある。一方、非構造型の先行研究として、PhotoSurfing[2]がある。構造に依ることなく画像を空間内にちりばめる手法である。

本研究では、画像群の全貌を一目で把握でき、かつ類似性のある画像が画面上で近くに配置される、という両特徴を同時に満たすことに重点を置いた画像ブラウザの開発を考える。本研究では、類似した画像は共通のキーワードを有すると考え、予めキーワードが与えられた画像群を想定する。そしてCATと同様に、キーワードに基づいて構築した階層構造により画像を画面配置する。しかしCATでは、キーワードの共通性を画面配置の考慮に入れていないため、共通したキーワードを有する画像が離れて配置されるという問題が起こりうる。そのため、局所表示時に画像を見落としてしまい、興味の対象となる画像群を探すのに時間を要する恐れがある。

本報告では、CATの改良手法として、共通したキーワードを有する画像が近くに配置されるような画面配置手法を提案する。局所表示時の画像の見落としを防ぎ、興味の対象となる画像群を短時間で容易に見つけることができると期待される。

## 2. 関連研究

### 2.1. CAT

CATとは、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である。前処理として、大量画像に対してキーワード、画素情報に基づく2段階のクラスタリングを適用し、階層型データを構築する。そして、各クラスから代表画像を選出する。続いて、この階層型データに、階層型データ可視化手法「平安京ビュー」[3]の配置アルゴリズムを適用し、画像群を一覧可視化する。このアルゴリズムにより、以下の2条件

**【配置条件1】** 画像同士、クラス同士の遮蔽回避

**【配置条件2】** 画像群の配置占有面積の低減  
を満たす画面配置を実現する。

CATではズーム率に合わせた詳細度制御を設けており、ズームアウト時は各クラスの代表画像を、ズームイン時には局所的に各画像を表示する。

### 2.2. FRUITS Net

本手法で適用するネットワーク可視化手法FRUITS Net (Framework and User Interface for Tangled Segments Network)[4]は、1個以上のアイテムが各ノードに付加されたネットワーク

を対象とした可視化手法である。FRUITS Netではノードをアイテム毎に色分けし、ノード間の連結をエッジで表す。そして、力学モデルと空間充填モデル[3]を併用した配置アルゴリズムによりノード配置を決定する。これにより、複数のアイテム情報を有するネットワークの全体像を一画面上で表すことが可能となる。FRUITS Netでは上記の【配置条件1】【配置条件2】に加えて、

**【配置条件3】** 共通アイテムを有するノード群を近くに配置

**【配置条件4】** エッジ長の総計、およびエッジ交差数の低減も同時に満たす画面配置を実現する。

## 3. 提案内容

本研究では、予め各々の画像に1個以上のキーワードが与えられた画像群を想定する。

### 3.1. 階層型データの構築

本手法では、CATと同様に、大量画像に対し、キーワード、画素情報に基づく2段階のクラスタリングを適用し、階層型データを構築する。

キーワードによるクラスタリングでは、まず、キーワードに基づく画像間の距離を定義する。画像 $X_i$ の持つキーワードセットを $W_i$ とする。

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

$n_i$ は画像 $X_i$ に与えられたキーワード数、 $w_{i,k}$ は画像 $X_i$ の $k$ 番目のキーワード、 $V$ は画像群に用いられている全てのキーワードを表す。このとき、画像 $X_i$ と $X_j$ 間の距離 $D(X_i, X_j)$ を以下の式により算出する。

$$D(X_i, X_j) = \frac{1}{2m_i} \sum_{k=1}^{n_i} d_w(w_{i,k}, W_j) + \frac{1}{2m_j} \sum_{l=1}^{n_j} d_w(w_{j,l}, W_i) \quad (2)$$

$d_w(w_{i,k}, W_j)$ はキーワード $w_{i,k}$ とキーワードセット $W_j$ 間の距離を表す。キーワード間の距離は、ソフトウェアパッケージWordNet Similarityにより算出している。得られた距離に基づいて画像をクラスタリングする。この距離を用いることで、全く同一のキーワードセットを有する画像群だけでなく、非常に意味の近いキーワードセットを有する画像群を、同一クラスに所属させることができる。こうして生成したクラスを高階層クラスと呼ぶ。

続いて、高階層クラス内の画像に対し、画素情報によるクラスタリングを行う。色情報および周波数情報による特徴量から、各画像の特徴量ベクトルを算出する。色情報による特徴量として、画像空間を格子分割し、各領域における代表画素値を算出し、結果を多次元ベクトル化して用いる。周波数情報による特徴量として、ウェーブレット変換を施して高周波成分を算出し、非零画素およびその隣接画素のパターンを検出し、各パターンの頻度から得られるヒストグラムを多次元ベクトル化して用いる。これら2種類の特徴ベクトルを併合したベクトルの余弦に基づき、値が大きい順にボトムアップ式にクラスタを生成する。上記の処理を、2クラス間の最大余弦値が閾値より小さくなるまで反復する。こうして生成したクラスタを低階層クラスと呼ぶ。

### 3.2. 階層型データの画面配置

本手法では、高階層クラスタに対し、FRUITS Net の配置アルゴリズムを適用し、各クラスタを画面配置する。

処理過程において、階層型グラフ  $G=(C,E)$  を生成する。 $C=\{c_1, \dots, c_{n_C}\}$  は各々が 1 枚以上の画像から成るクラスタ群を表す。 $n_C$  はクラスタの総数である。各々のクラスタ  $c_i$  は  $n$  次元のブーリアン変数  $c_i=\{b_{i1}, \dots, b_{in}\}$  を有する。 $n$  はキーワードの総数である。 $E=\{e_1, \dots, e_{n_E}\}$  は次の条件により生成されたクラスタ間のエッジを表す。 $n_E$  はエッジの総数を表す。

**[条件]**  $c_i, c_j$  において、 $b_{ik}$  が真かつ  $b_{jk}$  も真、という関係が少なくとも 1 次元以上成立するならば、 $c_i, c_j$  間にエッジを生成

続いて、生成された各エッジ  $e$  に対し、以下の式により算出された重み  $w$  をつける。

$$\begin{cases} w = \alpha r & (4) \\ r = n_{AND} / n_{OR} & (n_{OR} > 0) \\ r = 0.0 & (n_{OR} = 0) \end{cases} \quad (5)$$

$\alpha$  は定数、 $n_{AND}$  は  $b_{ik}$  と  $b_{jk}$  が共に真であるブーリアン変数の次元数を表し、 $n_{OR}$  は  $b_{ik}$  と  $b_{jk}$  のうち少なくともいずれかが真であるブーリアン変数の次元数を表す。

こうして与えた重みを基に、エッジにバネの力学モデルを仮想し、エッジの長さを最適化する。ここでクラスタの質量を  $m$ 、クラスタの加速度を  $a$ 、エッジのバネ定数を  $k$ 、エッジの現在の長さを  $l$ 、エッジの安定状態の長さを  $l_0$  とし、以下の運動方程式を用いてクラスタを移動させることができる。 $l_0$  はエッジの重み  $w$  から算出されるものとする。

$$ma = -k(l - l_0) \quad (6)$$

この処理を各クラスタに対して反復することで、安定なクラスタ配置を得ることができる。この手法により [配置条件 3] を満たし、共通したキーワードを持つクラスタを近くに配置することが実現される。力学モデルに基づく画面配置では、初期位置からの反復処理により結果を算出するため、計算時間が大きくなることが懸念される。本手法では、画像単位ではなくクラスタ単位で計算を行うことで、計算時間の増加を抑える。

しかし、この配置結果が [配置条件 1] [配置条件 2] を満たしているとは限らない。そこで、その修正のために空間充填モデルを適用する。まず、力学モデルの結果を座標値としてテンプレートに記述する。テンプレートを参照し、空間充填モデルによりクラスタの配置を決定する。テンプレートの座標値を理想位置とし、できるだけ近い位置に、クラスタの遮蔽を回避し、かつ配置占有面積の増加を抑えながら画面配置結果を再算出する。以上により、2 章で論じた [配置条件 1] ~ [配置条件 3] を同時に満たす画面配置結果を実現する。

## 4. 実行結果

我々は、ペンシルバニア州立大学 Jia Li 准教授よりご提供いただいた 2360 枚の画像を対象に可視化を試みた。各画像は 1~3 語のキーワードを有し、キーワード総数は 49 である。本報告の実験データでは、低階層クラスタリングは適用していない。生じたクラスタ総数は 136 である。

可視化結果を図 1 に示す。画像 1 枚が 1 クラスタに該当し、画像群の全貌を一目で把握することができる (図 1(左))。ス

ケール操作、シフト操作が可能であり (図 1(中央))、局所的に各画像を表示し、任意の画像上で左ボタンをクリックすると選択した画像が拡大表示される (図 1(右))。



図 1: 可視化結果

本手法の可視化結果を図 2(左)に、CAT の可視化結果を図 2(右)に示す。赤枠で囲まれた画像はキーワード "grass" を有するクラスタを表す。両者を比較してみると、キーワード "grass" を有するクラスタは、CAT では四隅に点在しているが、本手法では画面中央付近に固まって配置されている。確かに、共通したキーワードを有する画像、つまり類似した画像が近くに配置されている。以上から、局所表示時に画像の見落としを防ぎ、興味の対象となる画像群を短時間で容易に見つけることができるという点に優位性のある配置結果が得られているのではないかと見える。



図 2: キーワード "grass" を有するクラスタ配置の比較

## 5. まとめ

本報告では、キーワードの共通性を配置結果に反映した大量画像の一覧可視化手法について提案した。

今後の課題としてユーザテストの実施、GUI の機能の拡充、クラスタの配置占有面積の改善、別の配置アルゴリズムに置き換えた場合との比較などに着手したい。

## 謝 辞

本研究を進めるにあたり、画像データをご提供いただいた、ペンシルバニア州立大学 Jia Li 准教授に感謝の意を表す。

## 参考文献

- [1] 五味, 宮崎, 伊藤, Li, "CAT: 大量画像の一覧可視化と詳細度制御のための GUI," 画像電子学会誌, Vol. 38, No. 4, pp. 1108-1115, 2008.
- [2] H. Horibe, T. Itoh, "PhotoSurfing: A 3D Image Browser Assisting Association-Based Photograph Browsing," NICOGRAPH International 2009.
- [3] 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可視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.
- [4] T. Itoh, C. Muelder, K.-L. Ma, J. Sese, "A Hybrid Space-Filling and Force-Directed Layout Method for Visualizing Multiple-Category Graphs," 2009 IEEE Pacific Visualization Symposium, pp. 121-128, 2009.