

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(O)

安田 理紗[†] 五味 愛[†] 伊藤 貴之[†]

[†]お茶の水女子大学理学部情報科学科 〒112-8610 東京文京区大塚 2-1-1

E-mail: † {risa,gomiai,itot}@itolab.is.ocha.ac.jp

あらまし 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。画像ブラウザの要件には、画像群の全貌を一目で把握できる、類似した画像が近くに配置される、といった点があるが、これらを両立するのは難しい。本報告では、類似した画像は共通のキーワードを有すると考え、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する。本手法では、1 個以上のキーワードを有する大量画像に対し、キーワード、画素情報と 2 段階のクラスタリングを行い、階層型データを構築する。構築された階層型データに、力学モデルと空間充填モデルを併用した可視化手法 FRUITS Net の配置アルゴリズムを適用し、クラスタの配置を決定し、一覧可視化する。共通したキーワードを持つ画像が近くに配置されることで、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。

キーワード 可視化、階層型データ、画像ブラウザ、力学モデル、空間充填モデル

Visualization of large image collections arranged by commonality of keywords(O)

Risa YASUDA[†] Ai GOMI[†] and Takayuki ITOH[†]

[†] Department of Information Sciences, Faculty of Science, Ochanomizu University

2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610 Japan

E-mail: † {risa,gomiai,itot}@itolab.is.ocha.ac.jp

Abstract Recently, image browsing techniques become an active research topic. There are several requirements for image browsers, such as all-in-one image browsing and similarity-based layout, however, it is generally difficult to satisfy all of such requirements. This paper proposes a visualization technique for browsing of large image collections arranged by commonality of keywords, based on the assumption that similar images have common keywords. As a preprocessing, the technique applies hierarchical clustering to the images based on their keywords and pixel values. Then, it places these clusters by a hybrid space-filling and force-directed layout algorithm. Placing images which have common keyword close, we expect that we can prevent oversight of images, and pick out images of interest easily in a short time.

Keyword Visualization, Hierarchical data, Image browser, Force-directed layout, Space-filling layout

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」に大別できると考える。

構造型ブラウザの先行研究として、PhotoMesa[1]やCAT(Clustered Album Thumbnails)[2]、MoireGraphs[3]が挙げられる。PhotoMesa, CAT は階層構造により、MoireGraphsはネットワーク構造により画像を管理し、そのデータ構造に基づいて画像を画面に配置する手法である。

続いて、非構造型ブラウザの先行研究として

PhotoSurfing[4]がある。これらは木構造やネットワーク構造といったものを導入せずに、画像を空間上にちりばめる手法である。

本研究では、画像群の全貌が一目で把握できる、類似した画像が近くに配置される、この両側面を同時に満たすことに重点を置いたブラウザを開発したいと考えた。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した。そこで、同じくキーワードが与えられた画像群を想定したCATのように、階層構造に基づいて画像を画面配置することを考える。CATの特徴として、階層構造に分類された画像群からクラスタごとに選んだ代表画像

を表示することで全体を表し、ズームイン操作により局所的に各画像を表示する、という操作性が挙げられる。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。

そこで本報告では、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、各クラスタ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」[5]の配置アルゴリズムを適用することで、画像群を一覧可視化する。このアルゴリズムによって、以下の2条件

[配置条件1] 画像どうし、クラスタどうしの遮蔽回避

[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タの代表画像を、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)[7]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[7]と空間充填モデル[5]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とにより、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では上述の[配置条件1][配置条件2]に加えて、

[配置条件3] 共通する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. 提案手法

本研究では、予め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、画素情報と2段階のクラスタリングを行い、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、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 Word Net 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、同一クラスタに所属させることができる。こうして生成したクラスタを高階層クラスタと呼ぶ。

続いて、高階層クラスタ内の画像に対し、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いて、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クラスタ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タと呼ぶ。

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。処理手順を図1に示す。

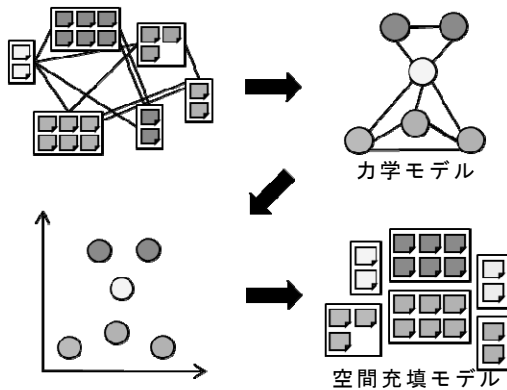


図1: 処理手順

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 $c_i = \{b_{i,1}, \dots, b_{i,n}\}$ 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、 $b_{i,k}$ が真かつ $b_{j,k}$ 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(図1左上)。

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

ここで、 α は定数、 n_{AND} は $b_{i,k}$ と $b_{j,k}$ が共に真であるブーリアン変数の次元数を表し、 n_{OR} は $b_{i,k}$ と $b_{j,k}$ 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もとに、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(図1右上)。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、この配置結果は[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(図1左下)。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(図1右下)。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、クラスタどうしの遮蔽を回避し、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。

3.3. 階層型データの可視化

本手法では、CATと同様に、詳細度制御を実現している。初期画面では、決定した画面配置に基づいて高階層クラスタの代表画像を表示する。ズームイン操作により、局所的に各画像を読み込み、表示する、

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。階層型データの構築により生じた高階層クラスタ数は136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。

可視化結果を図2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2(左))。スケール操作、シフト操作が可能であり(図2(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2(右))。



図2: 可視化結果

図3は、一部をズームインした結果である。枠で囲まれたクラスタは左から順に、

”grass, bird, water”, ”grass, water”, ”grass, bird”,

”grass, river”, ”grass, river, tree”
 というキーワードを有する画像群からなるクラスタである。共通したキーワード”grass”を有するクラスタが近くに配置されていることがわかる。同時に、例えば
 ”grass, bird, water”と”grass, water”,
 ”grass, river”と”grass, river, tree”
 のように、2 個以上のキーワードが共通したクラスタどうしが隣接していることも見てとれる。以上の観察結果から、共通したキーワードを有するクラスタが近くに配置されていることが確認できる。

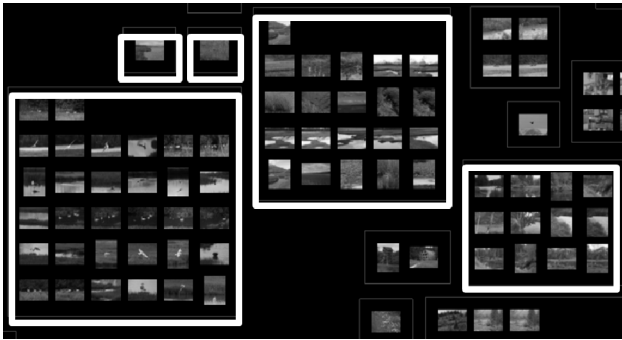


図 3：キーワード“grass”を有するクラスタ

本手法の可視化結果を図 4(左)に、本手法と同様の階層型データを用いた CAT の可視化結果を図 4(右)に示す。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している。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。画面上のクラスタ間距離を算出したところ、表 1 に示すように、距離の平均値、最大値ともに本手法のほうが小さくなった。

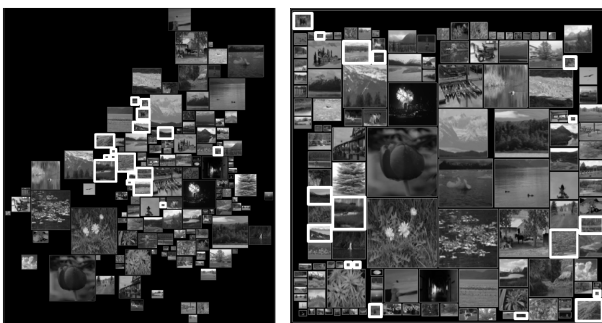


図 4：クラスタの画面配置の比較

表 1：キーワード”grass”を有するクラスタ間の画面上の距離

	平均値	最大値
CAT	1.328	2.636
本手法	0.347	0.843

前述の通り、CAT と本手法は、ズームイン操作により代表画像から局所的に各画像に表示を切り替える、

という点で共通している。CAT では、各クラスタを示す長方形の空間充填に重点を置き、キーワードの共通性を配置の考慮に入れていない。しかし、本手法では共通したキーワードを有する画像が近くに配置されていることから、局所表示時に画像の見落としを防止でき、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。

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、ユーザテストを実施するとともに、以下の機能の実装に取り組みたい。

- ・ キーワード選択時のクラスタ枠のハイライト
 - ・ 低階層クラスタ、リンク情報を用いた画面配置
- また、キーワード選択時のクラスタの再配置およびクラスタの画面占有面積の改善、FRUITS Net における力学モデルを別の配置アルゴリズムに置き換えた場合との比較にも着手したい。

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る。

参 考 文 献

- [1] B. Bederson, “PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps,” UIST 2001, ACM Symposium on User Interface Software and Technology, pp. 71-80, 2001.
- [2] 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.
- [3] T. J. Jankun-Kelly, K.-L. Ma, “MoireGraphs: radial focus+context visualization and interaction for graphs with visual nodes,” IEEE Information Visualization 2003, pp. 59-66, 2003.
- [4] H.Horibe, T.Itoh, “PhotoSurfing: A 3D Image Browser Assisting Association-Based Photograph Browsing,” NICOGRAPH International 2009,
- [5] 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専有面積の改善,” 可視化情報学会論文集, Vol. 26, No. 6, pp. 51-61, 2006.
- [6] 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.
- [7] 土井、伊藤, “力学モデルを用いた階層型グラフデータ画面配置手法の改良手法とウェブサイト視覚化への応用,” 芸術科学会論文誌, Vol. 3, No. 4, pp. 250-263, 2004.