

再帰的なフォトモザイクを用いた画像の関連性提示方法

佐藤 和 男^{†1} 藤澤 誠^{†1} 天野 敏之^{†1}
宮崎 純^{†1} 加藤 博一^{†1}

フォトモザイクは小さな写真を並べて一つの絵に見立てるモザイク画の一種である。本論文では、再帰的なフォトモザイクと写真に付加されたメタデータを用い、より多くの関連性のある写真をユーザに対して有効に提示させる方法を提案する。

A Method to Display Connellation Between Images by Using a Recursive Photomosaic System

KAZUO SATOH,^{†1} MAKOTO FUJISAWA,^{†1}
TOSHIYUKI AMANO,^{†1} JUN MIYAZAKI^{†1}
and HIROKAZU KATO^{†1}

Photomosaic consists of a mosaic in which a big picture is constructed using many small pictures. In this paper, we propose a method to effectively display correlated images by using a recursive photomosaic system and by adding metadata to the images.

1. はじめに

デジタルカメラやカメラ付きの携帯電話の普及によって、ユーザが写真をいつでもどこでも撮影することができ、なおかつSDカードやHDDなどに容易にかつ大量に保存できるようになった。また、Flickr¹³⁾ やフォト蔵¹⁴⁾ などの画像を共有するソーシャルネットワークサービス(SNS)サイトによって撮った写真を手軽にウェブに公開し、他のユーザが撮影

した写真を閲覧出来るようになっている。そして、それらのサイトはユーザの自由意思での画像に対しても主観的にタグを付加させることができる。これら大量の画像を単に見るだけでなく、タグや撮影日時などのメタ情報による関連性を考慮した上で提示することで、ユーザに対して新しい画像の発見を促せるのではないかと考えられる。しかし、ツリー表示やネットワーク表示のような画像同士の関連性をディスプレイに多く提示する場合、関連性は示せるもののそれらを表示する領域を確保するのに限界がある。一方、従来の画像を表示する方法として一般的なものは一次元のリスト表示や二次元のテーブル表示であるが、関連性を示すのは難しい。そこで、本研究では大量の画像をユーザに提示することができ、なおかつ図1のように画像間同士の関係性を単にネットワーク図として表すのではなく、それらの関係性を暗示させる新しい表示方法として芸術的に提示できるフォトモザイクを利用する。図2に示すように元画像を親ノード、小さなタイル(サムネイル)画像群は元画像の子ノードとして表す。

フォトモザイクは写真を要素としたモザイク画の一種であり、芸術作品として広く知られている。元の画像をタイルと呼ばれる小さな領域で分割し、そこにデータベースに格納された多くの写真(タイル画像)のあるルールに従ってあてはめることで、新しい画像を生成する。そして、そのあるルールとは主に色による近似によって、元の画像を近似することができる。また、藤澤らが開発したインタラクティブフォトモザイク¹⁵⁾ はモザイク中の小さなタイル画像の中からユーザが一つ選択することで、そのタイル画像が一枚の大きなモザイク画になる再帰的なモザイク画生成機能が備わっている。本研究では、インタラクティブフォトモザイクにタグの情報を付加することで拡張した。そのことにより、モザイク画とタイル画との写真間の関連性を暗示的に表示させることができ、さらに幾何学的な大規模ネットワークから一部分を断片的に取り出し、それを2次元的にコンパクトに表示させる技法として活用できる。

2. 関連研究

階層構造を2次元に表示させる手法としてJohnsonらが開発したTreemap⁶⁾が挙げられる。この方法はそれぞれのデータをさまざまな大きさの長方形で表現し、そしてそれらを入れ子として階層関係を示すものである。これにより、データの大きさや階層の深さと長方形の大きさに自然な対応付けがなされているため、全体の階層構造をユーザが認知しやすい表示になっている。そして、そのほかにもファイルの階層を3次元に表示し、なおかつ操作できるconetrees¹¹⁾や、Treemapを用いて画像のメタデータによる属性分類を表示させる、

^{†1} 奈良先端科学技術大学院大学
Nara Institute of Science and Technology
(630-0192 奈良県生駒市高山町 8916-5)

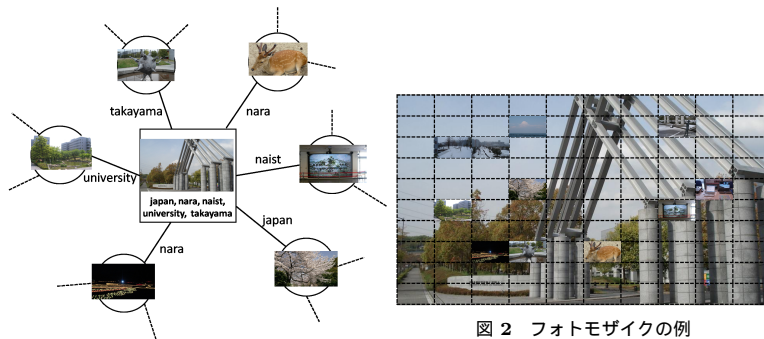


図 1 画像間の関連性によるネットワーク構造．
リンクの横にあるワードは各画像が保有
する共通のタグ情報である．

図 2 フォトモザイクの例

PhotoMesa⁴⁾などが挙げられる．しかし、大量の画像を効率的にかつコンパクトに提示するのは難しい．そこで、本研究では提示方法にフォトモザイクを用いる．

写真をタイルとして用いたフォトモザイクは Silvers ら¹⁰⁾によって開発された．そして、より高速なフォトモザイク生成のために、Di Blasi ら⁵⁾はタイル画像を9分割(3×3)し、各分割領域の平均色を計算することで得られるそれぞれの画像が表す27次元のベクトルを用いて画像のマッピングを行った．また、新しい表示手法として、区画の大きさを可変的にしてマッピングを行う手法⁵⁾や、長方形ではなく物体やキャラクターの形状を活用してジグソーパズルのようにマッピングさせる手法⁷⁾などが挙げられる．さらに、フォトモザイクを他のアプリケーションに応用する事例として、動画ナビゲーション⁸⁾やタイル画像が動画であるビデオモザイク⁹⁾などが挙げられる、

我々が提案する技法と一番類似しているものとして Vuillemot と Rumpler が提案する MozaiZ¹²⁾が挙げられる．この MozaiZ は親ノードの元画像とその画像と関連性のある子ノードのタイル画像群の関係性をモザイク画で表している．タイル画像を選択し、その画像を拡大表示する ZUI (Zoomable User Interface) 用いており、そしてその選択された画像のモザイク画が生成され、再帰的に繰り返し操作できる．本研究の手法と MozaiZ の手法との一番の相違点は画像のマッピング方法である．MozaiZ では、元画像となるモザイク画とタイル画群との関連性を重視しているため色の近似検索はされておらず、配置されたタイル画の上に元画像の色で加工された表示になっている．一方、本研究が提案する手法は、芸術

的な要素を加えることでユーザを惹き付ける効果を持たせるためにある程度の完成度が求められる点と、色の加工によってユーザが持つタイル画に対する印象に差異が生じないようにするため、各画像が持つ色の特徴ベクトルをもとに検索し、タイル画の配置を行っている．

3. インタラクティブフォトモザイクを用いた関連性提示方法

ここでは、どのように関連性を持たせたフォトモザイクを構成したかについて、各画像が保有しているタグ情報の取得方法とその情報を用いたフォトモザイクの生成方法を述べ、そして最後にどのようにインタラクティブな操作性を実現させたかについて述べる．

3.1 タグの取得方法

flickr¹³⁾などの画像投稿サイトにおいて、収集された各画像には不特定多数の人が入力されたタグがメタデータとして EXIF 形式 (Exchangeable Image File Format) に格納されている．そこで、OpenExif Toolkit¹⁾を用い、それらのタグ情報を抽出した．

3.2 タグ情報を用いたフォトモザイク生成

本研究で扱うフォトモザイクは、モザイク画となる元の画像を格子状に分割してタイル領域を作成する．そして、それらの各領域の色特徴ベクトルから類似したタイル画像を検索し、マッピング合成を行う．その際、色特徴ベクトルにはタイルを n^2 分割 ($n \times n$) したときの各分割領域の平均色 (r^{ij}, g^{ij}, b^{ij}) , $[i, j = 0, 1]$ からなる $3n^2$ 次元ベクトルを用いる．本研究では図4で示したように、一つの領域を4分割 (2×2) した．各分割領域における RGB の平均色を計算する必要があるため、12次元ベクトルとなる．

タグ情報を用いたフォトモザイク生成における全体の手順を図3で示す．最初に、前処理として元画像が保有しているタグ情報を用い、画像群の中からそれらのタグ情報を持つ画像だけでデータベースを構築する．そこで登録された画像の色特徴ベクトルを計算し、それを検索の高速化を行うために ANN (Approximate Nearest Neighbor) KD ツリー²⁾に格納する．全ての画像を用いてツリーを構築し、そこから関連性を含んだ画像を探索する手法もあるが、タグを用いて探索する際にヒットする画像の数に不安定さが生じ、探索範囲を広げる必要性が出てくるためそこにコストが生じる．本研究における方法でも、ユーザがタイル画像を選択するごとに必要となるタグ情報も変化するため、関連性を持つ画像のみでツリーを毎回構築するコストが生じるが、安定的に関連性のある画像が得られやすい．ここでは、画像間の関連性を重要視しているため、その手法を取り入れた．仮に、元画像が保有するタグの数が少ない場合はその次の孫ノードまで探索する必要はある．ここで言う孫ノードは元画像からのノード間距離が2の場合である．このとき、ノードをたどり過ぎると

ほぼ全部の画像を用いることとなってしまう、関連性の表示としては適切ではない。これを抑えるために探索範囲に制限を付ける。ユーザが指定する係数値 $R \in (0, 1]$ と最大最短経路 $Path_{max}$ をかけた値 $Search = Path_{max} \times R$ をノード間距離の最大とすることで探索範囲を制限する。

高速化のために、タイルの大きさがあらかじめ決定している場合は、各画像をタイルの大きさにリサイズしておく。そして図5中にタイル画像格納テクスチャとして示したようにリサイズされた画像を一次元配列に再配置し、データベース内の全てのリサイズ画像を一枚の画像として格納する。そのことによって、画像合成時においてタイル画像のアクセスが容易になった。

フォトモザイクを生成するときは、元の画像を読み込んだ後、その画像のタイル分割と各タイルにおける色特徴ベクトルの計算を行う。そして、前処理で構築されたツリーを用いて、最近傍探索により類似画像を取得し、最後にフォトモザイクの合成を行う。図5にGPUのピクセルシェーダを用いたフォトモザイク画像の合成方法を示す。シェーダは複数存在しており、各ピクセルごとに並列処理することができる。それを利用して、モザイクを行う元の画像と同じ大きさのビューポートを設定し、矩形ポリゴンを描画することで対応するテクスチャ座標から元画像上のピクセル位置を得る。そこから、タイル座標に変換し、マッチング結果を格納したテクスチャの座標値を参照することで対応する画像の番号を取得する。そうすることで最終的なピクセルの値を得ることができる。

3.3 インタラクティブな操作性

本研究では、新しい画像の発見を促すために再帰的なフォトモザイクアプリケーションを提案する。このアプリケーションでは、ディスプレイに表示されたフォトモザイクからユーザが任意の画像を選択する。そして、選択された画像を元に、新たにフォトモザイクをリアルタイムに生成し、画像を最大化する。このとき、ズームするようなエフェクトをかけたZUI (Zoomable User Interface)⁴⁾の機能を追加することで、ユーザはあたかも全ての画像がフォトモザイクで構成される感覚を持つ。さらに、フォトモザイクで生成された元となる画像をユーザが確かめられるように、キーボードの上下キーまたは左下にあるスクロールバーで α 値を好みに応じて操作できるようにした。

この再帰的なフォトモザイクを用いることで、ユーザが興味のあるタイルの選択を繰り返しながら最終的に必要な情報が得られる、エンターテインメント性を内包したナビゲーションシステムのような応用例が考えられる。

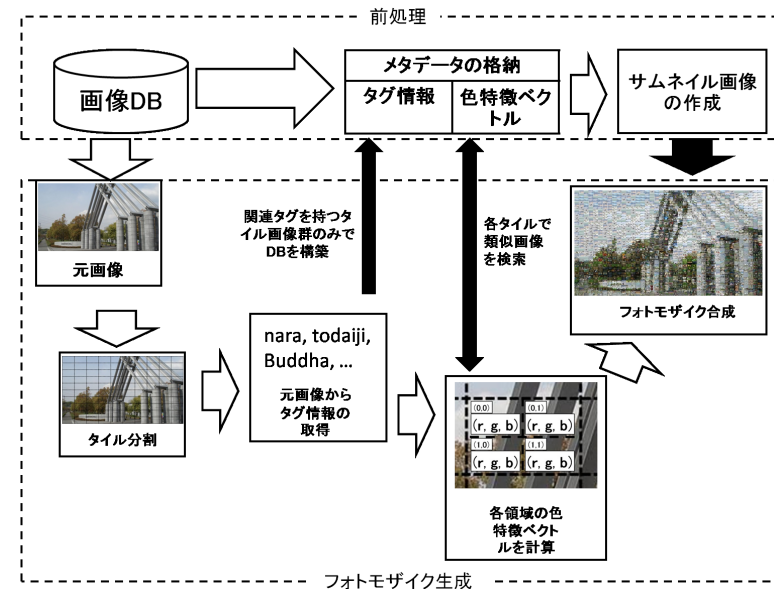


図3 フォトモザイクの生成手順

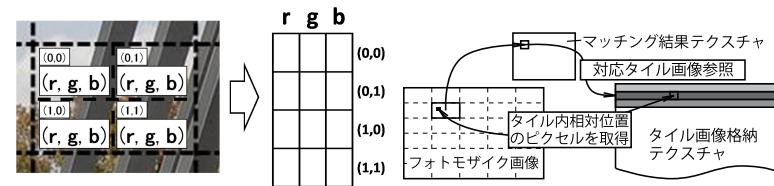


図4 色特徴ベクトルの計算 (4分割, $n = 2$ の場合)

図5 フォトモザイク合成

4. 結果

flickr等の画像投稿サイトから、'todaiji'や'yakushiji'などの奈良の観光名所に関連した言葉を主にクエリとし、画像の収集を行った。実験時に用いた全ての画像の枚数は2880枚であり、タグの種類は全2519種であった。このとき、大文字と小文字の区別はしておら

ず、また、例えば‘today-ji’のようにタグの文字列にハイフンやスペース等が入っている場合はその部分を除去し、‘todayji’として扱った。さらに、‘nara’というタグを所有している画像が2433枚、‘japan’が2202枚存在したため、それらのタグは検索の対象としなかった。ただし、‘奈良’や‘japon’など漢字や英語以外の他言語は意味が同義であっても除外しなかった。表1ではそのほかの上位のタグ情報の頻度を示す。

そして、Intel Core2 Duo 3.16GHz、ビデオカードとしてNVIDIA GeForce GTX 285（ビデオメモリ1GB）を搭載したPC上でアプリケーションを実行した。

図6では、中央の写真がフォトモザイク合成画像、右下が元画像、そして左下は各画像が保有しているタグ情報のリストを示している。そして、例えば左上にある鹿とアイスクリームの2つの写真が保有しているタグに注目すると、両方とも‘todayji’のタグを持っており、さらに下の紅葉の写真も‘todayji’タグを持っていることが分かる。本研究では一つでも同一のタグを持つ画像は関連性を持っていると定義し、この関連性を持つ画像のみでフォトモザイクを構築した。

写真間の関連性を分析するために、実験時に用いた画像からランダムにサンプル抽出し、大規模ネットワーク分析ツールであるPajek³⁾を用いて分析を行った。その結果、772枚の画像ノードには183261本のリンクが存在し、全ノードにおける平均最短経路は $Path_{ave} = 1.43654$ であり、最大最短経路は $Path_{max} = 5$ であった。このことから、ほとんど場合において元画像である親ノードから探索する場合、子ノードまでの1回のパスで十分な画像数を収集でき、孫ノードまでの2回のパスでほとんどの画像が収集できると考えられる。また、フォトモザイク生成に必要な画像数を安定して得られているのを確かめるために、ランダムにタイル画像を選択して自動的にフォトモザイクが生成されるように設定し、ANN-Kdツリーに格納した枚数をカウントした。このとき、ランダム選択を500回試行した。その結果、ツリーに格納された枚数は平均 $\mu = 1210.53$ であり、標準偏差は $\sigma = 532.86$ であった。これは、‘nara’と‘japan’を除去したものの、その次の‘todayji’

タグ情報	頻度	タグ情報	頻度
todayji	1040	shrine	418
temple	1021	buddha	331
奈良	862	ikoma	328
日本	530	東大寺	314
horyuji	478	法隆寺	309

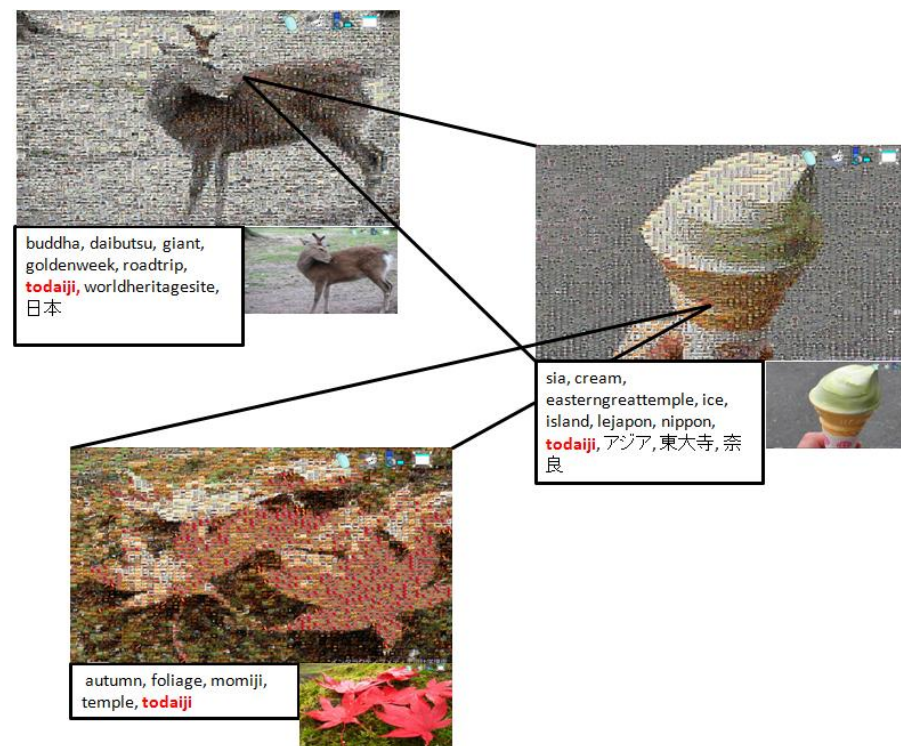


図6 実行結果

や‘temple’の単語がハブ(hub)となり、ショートカットの役割を果たしたために平均最短経路の値が低く抑えられ、それによって安定的に画像を獲得できたと考えられる。

5. おわりに

本研究では、従来のグリッドやネットワーク図などによる提示方法ではなく、タグのついた画像を用いたインタラクティブフォトモザイクを提案した。そこで、各画像内にあるメタデータからタグの情報を取り出し、タグと色特徴ベクトルを元にデータベースの構築を行うことによって、元画像とタイル画像群との関連性を表すフォトモザイクを実現した。今後の研究としては、画像選択回数や質問紙によるアンケート調査などの評価実験を行い、被験者

の反応を考察しながら実装修正に反映させることがあげられる。また、拡張として以下の2つが挙げられる。

- モザイク画の表示方法
5) のように、タイル画像の大きさをモザイク画の関連性に応じて可変にし、関連性が大きければそのタイル画像を大きめに表示するように、よりユーザが無意識に関連性を認知できるような表示方法が可能となると考えられる。
- ネットワーク構造
中心媒介性や属性コミュニティ間のつながりなどのネットワーク構造を考慮した画像の選別ルールを加えることで、ユーザが欲しい情報をより正確に提示できる可能性がある。

参 考 文 献

- 1) : OpenExif. <http://openexif.sourceforge.net/>.
- 2) Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R. and Wu, A.Y.: An optimal algorithm for approximate nearest neighbor searching fixed dimensions, *J. ACM*, Vol.45, No.6, pp.891–923 (1998).
- 3) Batagelj, V. and Mrvar, A.: Pajek. <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.
- 4) Bederson, B.B.: PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps, *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*, New York, NY, USA, ACM, pp.71–80 (2001).
- 5) di Blasi, G., Gallo, G. and Petralia, M.P.: Smart Ideas for Photomosaic Rendering, *Eurographics Italian Chapter Conference*, pp.267–272 (2006).
- 6) Johnson, B. and Shneiderman, B.: Tree-Maps: a space-filling approach to the visualization of hierarchical information structures, *VIS '91: Proceedings of the 2nd conference on Visualization '91*, Los Alamitos, CA, USA, IEEE Computer Society Press, pp.284–291 (1991).
- 7) Kim, J. and Pellacini, F.: Jigsaw image mosaics, *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, ACM, pp.657–664 (2002).
- 8) Kim, K., Essa, I. and Abowd, G.D.: Interactive mosaic generation for video navigation, *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, New York, NY, USA, ACM, pp.655–658 (2006).
- 9) Klein, A.W., Grant, T., Finkelstein, A. and Cohen, M.F.: Video mosaics, *NPAR '02: Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering*, New York, NY, USA, ACM, pp.21–28 (2002).
- 10) Silvers, R. and Hawley, M.: *Photomosaics*, H. Holt and Co (1997).

- 11) Robertson, G.G., Card, S.K. and Mackinlay, J.D.: Information visualization using 3D interactive animation, *Commun. ACM*, Vol.36, No.4, pp.57–71 (1993).
- 12) Vuillemot, R. and Rumpler, B.: MosaiZ: Interactive Image Mosaics, Technical Report RR-LIRIS-2009-020, LIRIS UMR 5205 CNRS/INSA (2009).
- 13) Yahoo: Flickr. <http://www.flickr.com/>.
- 14) ウノウ株式会社：フォト蔵. <http://photozou.jp/>.
- 15) 藤澤 誠，加藤博一：GPU を用いたインタラクティブフォトモザイク，映像情報メディア学会誌，Vol.64, No.9 (2010). (採録予定).