

大局的特徴量 GIST を用いた作品例に基づく 絵画調画像生成 - Sample-Based Painterly Image Generation Using GIST -

阿部 敬由[†] 豊浦 正広[‡] 茅 暁陽[‡]

Noriyuki Abe[†] Masahiro Toyoura[‡] and Xiaoyang Mao[‡]

[†]山梨大学大学院医学工総合教育部 [‡]山梨大学大学院医学工総合研究部

[†]Department of Education, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

[‡]Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

E-mail: {g11mk002, mtoyoura, mao }@yamanashi.ac.jp

1. はじめに

一般の人にとって絵を描くことは難しい。近年、パソコンやインターネットの普及に伴い、市販またはフリーの画像処理ソフトを利用して手軽に絵画調画像を作成できるようになった。しかし、このように作った絵画調画像のほとんどは入力画像全体を何らかのフィルタに一通りに通したものに過ぎず、本物の絵画とは大きく異なる。また、模倣できる絵画のスタイルも限られている。より本物に近い絵画調画像を生成すること目的として、色鉛筆、水彩、水墨など、各種画材や描画技法をコンピュータ上でシミュレーションする研究が報告されている。これらの技術を利用して絵画調画像を生成するには多くのパラメータを調整する必要があり、ソフトウェアを操作する知識だけでなく、絵画に関する知識も要求される場合が少なくない。

一方、画材や技法が同じであっても、作者によって絵画のスタイルは様々であり、芸術作品の価値はアーティストの個性にあるとも言える。本研究では画材や技法を限定せず、指定した画家のスタイルを反映した絵画調画像の自動生成を試みる。

本研究では、色遣いと筆遣いを画家のスタイルを形成する主な特徴の二つと考え、既存絵画の色遣いと筆遣いを写真に転写することによって指定した画家のスタイルを反映した絵画調画像を生成する。しかし、絵画から写真へ特徴を転写させる際にはいかなる絵画でも良い結果が得られる訳ではない。画家は描く対象に合わせて色や筆使いを変えることで豊かな表現を実現している。絵画に描いてあるシーンと写真に写り込んでいるシーンが大きく異なる場合は不自然な転写結果が得られてしまう。一方、シーンを構成する要素としては空間的な構造と色の二つが上げられる。我々は近年様々な分野でその応用が注目され始めている人間の視覚過程の初期瞬間視知覚をモデリングする GIST 特徴量[1]が画像の知覚上の構造を捉えるのに有効であることに着目し、GIST 特徴量と色特徴量を組み合わせることにより写真に移っているシーンと似ているシーンを描いている絵画を絵画サンプル集から自動取得することに成功した。

提案手法では、まず、ユーザに写真と画家名を入力してもらう。そして、入力された写真に構造と色（以降合わせて構図と呼ぶ）がもっとも近い絵画を、入力された画家の絵画サンプル集から探索する。最後に、その絵画の色遣いと筆遣いを入力写真に転写させ、出力画像とする(図 1)。

2 章では関連研究について記述する。3 章では構図が近い絵画を取得してくる方法について、4 章では絵画のスタイルを転写する方法について記述する。5 章では実験結果を紹介するとともに、関連研究との比較を行う。



図 1 提案手法概要

2. 関連研究

入力画像を指定した絵画のスタイルに変換する研究として Wang らのパッチベーステクスチャ合成に基づく方法がある [2]。彼らの手法はユーザにストロークテクスチャのサンプルとして使う領域を参照絵画内で指定してもらい、サンプルテクスチャを入力画像の各領域にタイリングすることで絵画調画像を生成する。それに対して、Chang らは mean-shift 法を用いて絵画を色で領域分割し、各領域の平均色に応じたストロークのサンプルテクスチャを自動生成する手法を提案した[3]。サンプル画像さえ用意すれば、後は自動で生成できるという利点があるが、Chang らの手法では領域の平均色に基づいてストロークサンプルの作成と探索を行うため、色が近い領域は必ず同じストロークテクスチャとなり、また領域内は変化がなく一律なストロークテクスチャに変換されてしまう。同じ青色でも空と海は異なる筆遣いで描かれることがあることから分かるように、筆遣いは対象物の種類や位置にも依存する。そこで、本研究で提案する手法はまず指定した画家のサンプル絵画集から構図の似ている絵画を探索し、その絵画の色遣いと筆遣いを入力画像に移すことにより指定した画家のスタイルに

近い絵画調画像を生成する。構図が似ている絵画の探索には GIST と呼ばれる画像の大局的特徴量を用いる[5]。色遣いと筆遣いは絵画を特徴づけるもっとも重要な要素と考えられるので、この二つを模倣できれば、指定した画家のスタイルに近い絵画調画像を生成できると考えられる。また各画素周辺の輝度と色の分布を考慮して転写を行うため、領域内の輝度と色の変化も保存される。

また、提案手法と同様に参照画像から色の特徴を入力画像へ転写させる研究として Y. Chang らの手法[4]がある。これは色合いを転写するとき、直接参照画像から対象画像へ移すのではなく、人間の知覚に基づき作成した色のグループを用い、グループ中で比較して近い色を転写することによって入力画像に不自然な色が転写されることを無くしている。また Image Analogies と呼ばれるテクスチャ転写法との組み合わせにより、筆遣いを絵画から写真へ転写する点においても本研究と類似している。しかし、Y. Chang らの手法では指定した対象画像と参照画像を入力として、その 2 枚の間で色合いの転写を行う。もし対象画像と参照画像の構図が大きく異なり、色の転写先、転写元で同じグループに分類される色がない場合は、やはり不自然な結果が得られてしまう可能性がある。

本研究では筆遣いの転写に画像間のテクスチャ転写を行う技術として広く利用されている image analogies[5]法を利用している。image analogies 法はある画像 A と A を変換して得られる画像 A' を教師画像ペアとして、別の画像 B が与えられた時、B に同様な変換を施した結果を生成することができる。本研究の場合、筆遣いの転写元となる絵画は A' に相当し、転写先の入力写真は B に相当する。提案手法では絵画にバイラテラルフィルタを施し、絵画が描いているシーンの写真 (A に相当する) を模倣する画像を生成する。

3. 構図が近い絵画の探索

提案手法では絵画から写真へ色遣いと筆遣いを転写することによって絵画調画像を生成する。このとき写真と絵画は構図が近いことが好ましい。写真に移っているものが絵画の中に損残しない場合は、不自然な転写結果となる場合がある。図 2 に構図が近い絵画から転写した結果と構図が異なる絵画から転写した結果を示す。図 2 (a) は入力写真であり、図 2 (b) の絵画から色遣いと筆遣いを図 2 (a) に転写した絵画調画像が図 2 (c) である。左側は写真と絵画の構図が近い例である、右側は構図が異なる場合である。右側の生成結果では空にひび割れのようなものが生じ、またノイズの様な色が出現している。左の画像の方が違和感なくより実際の絵画に近いことが見て取れる。



(a) 写真



(b) 参照する絵画



(c) 生成した絵画調画像

図 2 構図が近い絵画と遠い絵画から転写した結果

3.1 GIST と色特徴量

人間があるシーンに目を向けた瞬間 (100ms 以下) におきる知覚のことを GIST Perception といい、GIST Perception において人間はシーンの大まかの構造を知覚することができると言われてている。Oliva らは GIST の計算学的モデルとして、多方向多解像度ガボールフィルタを提案した[1]。ブロックに区切った画像に対するこれらのフィルタの出力からなる特徴ベクトルはシーンの大局的な構造を表す。森本らはこの GIST 特徴ベクトルを利用してインターネットからグレースケール画像と構図の似ているカラー画像の探索に成功した[6]。

指定した画家の絵画サンプル集から入力画像と構図の似ている絵画を探索するために、本研究はまず絵画画像を入力画像と同じサイズに正規化した後 4×4 の領域に分け、領域ごとに 6 方向 5 段階の周波数成分を検出するガボールフィルタをかける。これにより 480 次元の特徴ベクトルを得る(図 3)。しかし、この特徴ベクトルは画像の構造を検出することはできるが、例えば紅葉している木と葉が緑の木を区別することができない。この問題を解決するために色特徴量も併せて使用する。同じく入力画像と同じサイズに正規化した後 4×4 の領域に分割し、それぞれの領域に対し HSV 色空間における色相(Hue)のヒストグラム (12 色) を作成する。色特徴は $4 \times 4 \times 12 = 192$ 次元のベクトルとなり、最終的に探索に使用するベクトルは $480 + 192 = 672$ 次元である。

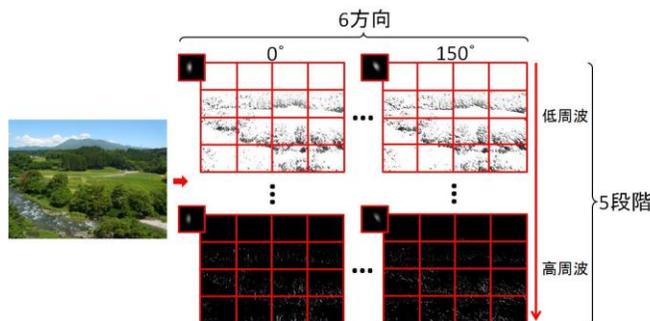


図 3 GIST 特徴量

図4に特徴量をGISTのみを用いた場合と色特徴量も併せて用いた場合の探索結果を示す. 図4(a)を入力写真として構図が近い絵画をGISTのみを用いた探索結果が図4(b), GISTと色特徴量を併せて用いた探索結果が図4(c)である. 左側の画像を見みると, 図4(b)と図4(c)には木や草など図4(a)の写真と同じ被写体が写り込んでいる. しかし, 図4(b)は地面が主に薄い茶色でまた山の部分にも茶色が多く混じっている等, 図4(a)とは色合いが少々違う. 一方図4(c)は木や草, 川等同じ被写体が含まれており, かつ構成する色合いも近い色となっている. 右側の画像では図4(b)と図4(a)はほぼ違う被写体で構成されており, 構図が近いとは言いがたい. しかし図4(c)には空, 木, 山, すずきなど図4(a)の写真と同じ被写体が写り込んでおり, 色合いも近いものとなっている. このように色特徴量も加えることで被写体の色や画像全体の色が近い絵画を取得することができる.



図4 GISTと色特徴を用いて探索した構図が近い絵画

3.2 ブロックマッチング法

構図が近い絵画を探索してくるために, 3.1章で求めた特徴量を用いる. 入力である写真 i とデータベースの絵画 d との構図の類似度は下記のグローバルコストを用いて評価する.

$$Cost_{global}(i,s) = \sum_{n=0}^{15} |F_n^i - F_n^s| \quad (1)$$

ここで, F_n^i と F_n^s はそれぞれ入力画像 i と絵画 s 内の対応領域の特徴ベクトルである. 画像を 4×4 の領域に分割し, 同じ位置にある領域同士でコストを計算している.

しかし, 車のあるシーンを例にとってみると, 背景が大きく変わらなければ, 車の位置が少し移動しても, 人間から見たシーンの構図はほぼ同じである. しかし, 式(1)のコストで絵画の探索を行うと車の位置が1ブロック以上ずれていれば, コストは大きくなり, 構造が類似していないと

判断されてしまう. このようなケースにも対応できるように提案手法ではブロックマッチングという方法を導入する. ブロックマッチング法では, 単純に同じ位置にある領域同士で比較してコストを決めるのではなく, 入力である写真とサンプル集の1枚の絵画の中から最も特徴に近い2つの領域のペアを探し出し, その2つの領域同士で比較してコストを決める. これにより, 局所的にでも似たような被写体が写っている絵画が選択される可能性が高くなる. ブロックマッチング法によるローカルコストは以下の式で計算する.

$$Cost_{local}(i,s) = \sum_{n=0}^{15} \min_m |F_n^i - F_m^s| \quad (2)$$

F_m^s は入力画像の領域 F_n^i と最も特徴ベクトルが近い絵画画像の領域である. 式(1)のように同じ位置の領域同士で比較するのではなく, 入力画像の各領域に対して最も特徴ベクトルが近い絵画の領域をペアにしてコストを計算する. これにより被写体が移動した様な絵画でもコストが小さくなるため, 構図が近い絵画に選ばれる可能性が高くなる. 最終的に検索に利用するコストは, グローバルコストとローカルコストの重み付き和である.

$$Cost_{total}(i,s) = Cost_{global}(i,s) + (Cost_{local}(i,s) \cdot \rho) \quad (3)$$

ρ はローカルコストに対する重みで, 値を小さくするほど構図の制限が厳しくなる. 図5に式(1)のコストで画像を探索した結果と式(3)のコストで画像を探索した結果を示す. 探索に使用した絵画のサンプル集には画像投稿サイト Flickr から "Oilpainting", "Landscape" という検索ワードで検索した最新の画像約4000枚を用いた.



図5 2種類の方法を用いた構図が近い絵画の探索結果

図5(b)は式(1)を用いた結果で, 確かに同じ位置に同じような色合いや輪郭が存在する絵画が選ばれている. 一方図5(c)の絵画は同じ位置に同じような色合いや輪郭が必ず存在するわけではないが, 同じような被写体が写り込んでいて, 図5(b)の絵画よりも図5(a)の写真に構図が近い印象がある.

4 絵画のスタイルの転写

提案手法では、構図が近い絵画から色遣いと筆遣いを転写することによって絵画調画像を生成している。

4.1 色遣いの転写

画像間の色転写を行う技術はいくつも開発されているが、本研究は高速に処理できるという観点から森本らのカラー化技術を利用した[6]。森本らの手法ではグレースケール画像のカラー化が目的であったため、画素の類似度の比較に、輝度と周辺の輝度の標準偏差に加えて位置情報を用いた。本研究ではカラー画像同士の比較を行うため、輝度の代わりに $L^*a^*b^*$ 色空間における距離を使用した。写真中の画素 p と絵画中の画素 q 間の距離コスト $Cost_{p,q}$ を以下のように定義した。

$$Cost_{p,q} = (\Delta E^*ab + |\sigma_p - \sigma_q|)(1 + |C_p - C_q|)^\gamma \quad (4)$$

ここで、 ΔE^*ab は $L^*a^*b^*$ 色空間における色差であり、 σ は近傍の輝度の標準偏差、 C は 0.0~1.0 に正規化された画素の位置である。 γ は色のコストと位置のコストの重みである。画素 p に対しこのコストが最も小さくなる画素 q を探索し、その画素の色を転写する。

図 6 は写真に対し、絵画から色遣いを転写した様子である。図 6 (a) の写真に対して式(3)を用いて探索してきた構図が近い絵画が図 6 (b) である。図 6 (c) は図 6 (b) の構図が近い絵画から図 6 (a) の写真に色遣いを転写した画像である。左側の画像では図 6 (b) の木、草、川、空の色が図 6 (a) の対応する被写体に転写されている。右側の画像では図 6 (b) のすすき、木、空の色が図 6 (a) の対応する被写体に転写されている。



(a) 写真



(b) 構図が近い絵画



(c) 色遣いを転写した結果
図 6 色遣いの転写の様子

4.2 筆遣いの転写

筆遣いの転写には Image Analogies[5]を用いる。Image Analogies は A と A' の教師画像間に見られるテクスチャ特徴の対応関係を B に転写し入力画像に対し同じような対応関係を持った B' を作成することができる。つまり絵画が描いているシーンの写真があれば、その写真と絵画を教師画像のペアとして Image Analogies を適用し、絵画のストロークテクスチャを入力写真に転写することができる。しかし、サンプルとなる絵画は過去に描かれたものであり、対象のシーンの写真は通常存在しない。本研究では絵画にバイラテラルフィルタをかけることで、写真の代替画像を作成する。絵画ではストロークを用いて色やトーンを表現するため、写真において色や輝度が滑らかに変化する領域もストロークテクスチャの高周波成分を含む。バイラテラルフィルタはこのような領域内の高周波成分を無くすことができる。

5. 実験結果

この章では提案手法による生成結果を示す。いずれの絵画のサンプル集も画像投稿サイト Frickr からそれぞれの画家名のタグで検索をして集めた。図 8 は図 8 (a) の写真と「モネ」を入力とし、絵画調画像生成を行った結果である。サンプル集にはモネの作品 295 枚が含まれている。図 8 (a) の入力画像に対して、式(3)のコストによって探索した構図が近い絵画が図 8 (b) である。図 8 (c) は図 8 (b) から図 8 (a) へ色遣いを転写した画像である。図 8 (b) の木、草、空から図 8 (a) の対応する被写体に色遣いが転写されている。図 8 (d) は図 8 (b) から図 8 (c) へ筆遣いを転写した結果である。色遣いと同じように対応している被写体へ筆遣いが転写されている。



(a) 写真



(b) 構図が近い絵画



(c) 色遣い転写後



(d) 筆遣い転写後

図 8 モネを画家に指定した場合

図 9 は図 9 (a) の写真と「ゴッホ」を入力とし、絵画調画像生成を行った結果である。サンプル集にはゴッホの作品 300 枚が含まれている。図 9 (a) の入力画像に対して、式(3)のコストによって探索した構図が近い絵画が図 9 (b) である。図 9 (c) は図 9 (b) から図 9 (a) へ色遣いを転写した画像である。図 9 (b) の木、草、道路から図 9 (a) の対応する被写体に色遣いが転写されている。図 9 (d) は図 9 (b) から図 9 (c) へ筆遣いを転写した結果である。色遣いと同じように対応している被写体へ筆遣いが転写されている。



(a) 写真



(b) 構図に近い絵画



(c) 色遣い転写後



(d) 筆遣い転写後

図9 ゴッホを画家に指定した場合

6. おわりに

絵画調画像生成に関する研究が数多く報告されている。その多くは特定の画材や技法のシミュレーションまたは模倣に焦点を当てていた。一方、画材や技法が同じであっても、作者によって絵画のスタイルは様々であり、芸術作品の価値はアーティストの個性にあるとも言える。本研究では画材や技法を限定せず、指定した画家のスタイルを反映した絵画調画像の生成法を提案・実装した。

現在では特徴の1つとして筆遣いを取り上げ、また筆遣いを転写する方法として Image Analogies を用いている。Image Analogies は画像の構造を一切考慮せず、テクスチャとしての局所的な連続性のみを考慮して画素の転写を行うため、人物など形状や構造の正確な表現が要求される画像への適用は困難である。画像の構造を考慮した新しい転写方法の開発が今後の課題の一つとなる。

参考文献

- [1] A. Oliva and A. Torralba, "Building the gist of ascean, The role of global image features in recognition", In Visual Perception, Progress in Brain Research, Vol. 155, pp. 23-36, 2006.
- [2] B. Wang, W. Wang, H. Yang, and J. Sun, "Efficient Example-Based Painting and Synthesis of 2D Directional Texture, " IEEE Transactions on Visualization and Computer Graphics, vol. 10, no. 3, pp. 266-277, May/June, 2004.
- [3] C. Chang, Y. Peng, Y. Chen, and S. Wang, "Artistic Painting Style Transformation Using Example-based Sampling Method, ", Journal of information science and engineering, vol. 26, no. 4, pp. 1443-1458, 2010.
- [4] Y. CHANG, S. SAITO, K. UCHIKAWA, M. NAKAJIMA, "Example-Based Color Stylization of Images," ACM Transactions on Applied Perception, Vol. 2, No. 3, July 2005.
- [5] A. Hertzmann, C. Jacobs, N. Oliver, B. Curless, and D. Salesin, "Image analogies," In Proc. SIGGRAPH, pp. 327-340, 2001.
- [6] 森本悠嗣, 苗村健, "無数の画像群に着目したモノクロ画像の自動 Colorization", 電子情報通信学会技術研究報告, vol. 110, No. 35, pp. 93-98, 2010.