

# 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

阿部 敬由<sup>†</sup> 豊浦 正広<sup>‡</sup> 茅 暁陽<sup>‡</sup>

<sup>†</sup> 山梨大学大学院医学工総合教育部

<sup>‡</sup> 山梨大学大学院医学工総合研究部

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

あらまし 我々は、指定した画家のスタイルを写真に転写する手法を提案する。提案手法によって、その画家が描いたような絵画調画像を写真から生成できる。写真が入力として与えられると、まず、指定した画家のデータベースから、入力した写真と似た構造を持った絵画が探索される。探索には、人間がシーンの大まかな構造を知覚するのに利用しているとされる GIST 特徴量を用いる。次に、探索された絵画の色遣いと筆使いを入力写真に転写することで、その画家のスタイルを反映した絵画調画像を生成する。この提案手法では、写真と画家を指定するだけで、画材や技法を意識することなく、自動的にその画家のスタイルの絵画調を生成することができる。

キーワード 絵画調画像生成, 描画スタイル転写, GIST

## Sample-Based Painterly Image Generation Using GIST

Noriyuki ABE<sup>†</sup> Masahiro TOYOURA<sup>‡</sup> and Xiaoyang MAO<sup>‡</sup>

<sup>†</sup> Department of Education, Interdisciplinary Graduate School of Medicine and

Engineering, University of Yamanashi

<sup>‡</sup> Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

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

**Abstract** We present a new technique for artistic style transferring by considering the perceptual structure matching between source image and target images. Given a photograph, our technique first searches for an existing art work on a similar scene through using the computational model of GIST perception and then transfers both the color and brush texture from it by considering the structure matching between two images. The proposed method is fully automatic and can be used for transferring any styles.

**Keyword** Painterly image generation, Artistic style transferring, GIST

### 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することが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。画家は描く対象に合わせて色や筆遣いを変えることで豊かな表現を実現している。このような描画特徴を捉えるためには、写真に写っているものと似ているシーンを描いた絵画から転写を行うことが望ましい。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知覚過程の初期瞬間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写っているもの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おう。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 節では関連研究について記述する. 3 節では構図が近い絵画を取得してやる方法について, 4 節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 節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.

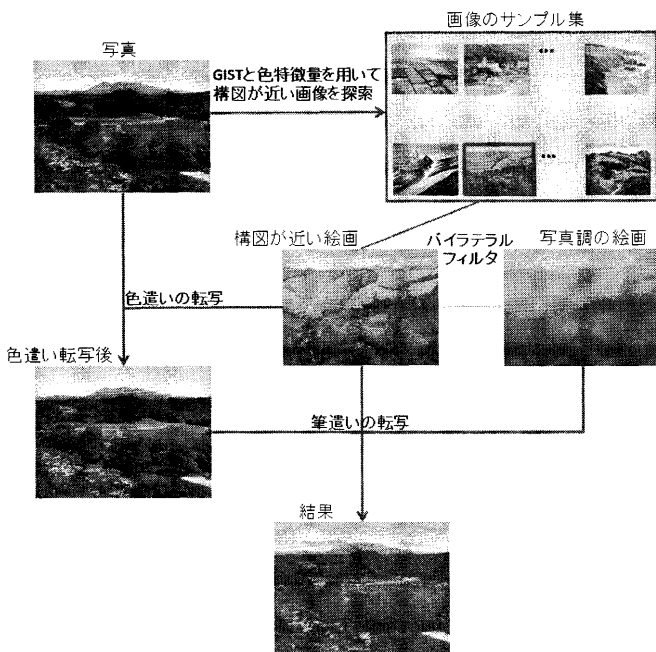


図 1 提案手法概要

## 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様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. 提案する手法は構図の似ている絵画から転写を行うことでこの問題の解決を試みる. 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, 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, グループ間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. この研究は 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できる点において本研究と類似している. しかし, Y. Chang らの手法では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

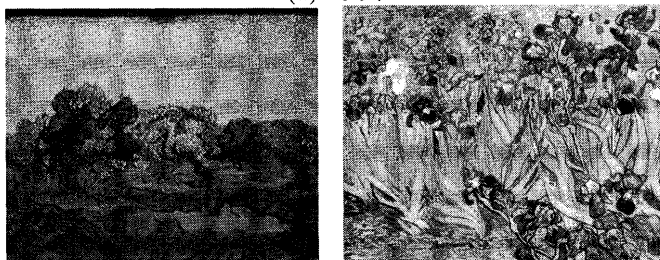
元で同じグループに分類される色がない場合は, やはり不自然な結果が得られてしまう可能性がある.

## 3. 構図が近い絵画の探索

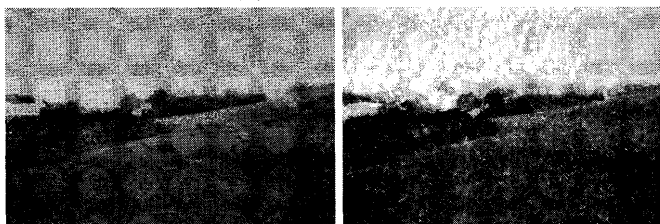
提案手法では絵画から写真へ色遣いと筆遣いを転写することによって絵画調画像を生成する. このとき写真と絵画は構図が近いことが好ましい. 写真に写っているものが絵画の中に存在しない場合は, 不自然な転写結果となる場合がある. 図 2 に構図が近い絵画から転写した結果と構図が異なる絵画から転写した結果を示す. 図 2 (a) は入力写真であり, 図 2 (b) の絵画から色遣いと筆遣いを図 2 (a) に転写した絵画調画像が図 2 (c) である. 左側は写真と絵画の構図が近い例である, 右側は構図が異なる場合である. 右側の生成結果では空にひび割れのようなものが生じ, またノイズの様な色が出てきている. 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 特徴ベクトル

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  $4 \times 4$  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。これにより各領域ごとに 30 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ないが、例えば紅葉している木と葉が緑の木を区別するように、色の特徴を捉えることはできない。この問題を解決するために色特徴量も併せて使用する。同じく入力画像のサイズに正規化した後  $4 \times 4$  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。これを各領域の色特徴ベクトルとし、最終的に探索に使用するベクトルは各領域ごとに  $30 + 12 = 42$  次元となる。

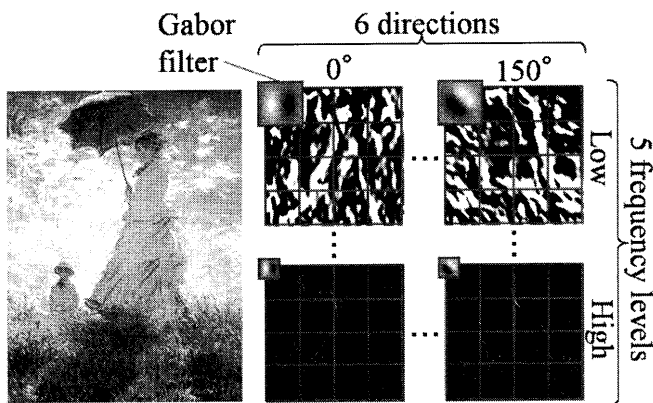


図 3 GIST 特徴量

### 3.2 シーンが似ている絵画の探索

入力である写真  $i$  と探索してきた絵画  $s$  とのシーンの類似度は下記のグローバルコスト  $G_g(i,s)$  を用いて評価する。

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

ここで、 $|\cdot|$  は  $L_2$  ノルムを表し、 $F_n^i$  と  $F_n^s$  はそれぞれ写真  $i$  と絵画  $s$  内の対応領域の 42 次元の特徴ベクトルである。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はほぼ同じである。式(1)のコストで検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、シーン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コストに局所的なマッピングを考慮する項を導入する。同じ位置にある領域同士ではなく、最も類似している領域同士で比較を行う。これをローカルコスト  $G_l$  と呼び、以下の式で算出する。

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

$F_m^s$  は入力画像の領域  $F_n^i$  と最も特徴ベクトルが近い領域である。式(2)は位置に関係なく、写真内の各領域と同じような領域が絵画に存在すれば、コストが小さくなる。しかし、画像や景色を構成する要素さえ似ていれば、位置がまったく対応していなくても低い値と

なるため、このローカルコストのみでは構図がまったく異なる絵画を検索してくる可能性もある。そこで、提案手法では式(1)により計算されるグローバルコストと式(2)により計算するローカルコストの重み付き和として最終コスト  $G$  を計算し、構図と構成要素の両方の類似度を測る。

$$G(i,s) = (1.0 - \rho)G_g(i,s) + \rho G_l(i,s) \quad (3)$$

係数  $\rho$  ( $0.0 \leq \rho \leq 1.0$ ) を変更することで構図と構成要素のどちらを重視するか調整することが可能である。 $\rho$  を大きくすることにより、異なるシーンでありながら、写真と同じような対象物が描かれている絵画が検出されやすくなる。今回実験用データベースを作成するにあたり、 $\rho = 0.6$  とした。

ローカルコストは最も類似している領域同士を対応付けて計算されるため、理論上の値域はグローバルコストと同じであっても、実際の値はグローバルコストより遥かに小さい。実利用における式(3)の有効性を保証するために、提案手法では統計的な手法を用いて両コストの実測値域を予測し、正規化を行った。写真 5 枚と絵画データベースにある 500 枚の絵画画像とローカルコスト、グローバルコスト合計 2500 枚の平均と分散を計算し、式(4)、式(5)を用いて 0.0 から 1.0 の範囲に正規化した。

$$\bar{G}_g(i,s) = \frac{1}{1 + \exp(-(G_g(i,s) - \tilde{G}_g)/\hat{G}_g)} \quad (4)$$

$$\bar{G}_l(i,s) = \frac{1}{1 + \exp(-(G_l(i,s) - \tilde{G}_l)/\hat{G}_l)} \quad (5)$$

ここで、 $\bar{G}_g(i,s)$ 、 $\bar{G}_l(i,s)$  はそれぞれグローバルコスト、ローカルコストの正規化後のコスト、 $\tilde{G}_g$ 、 $\tilde{G}_l$  はグローバルコスト、ローカルコストそれぞれに対して求めた平均値であり、 $\tilde{G}_g = 3.75 \times 10^{10}$ 、 $\tilde{G}_l = 7.11 \times 10^8$  となった。また、 $\hat{G}_g$ 、 $\hat{G}_l$  はグローバルコスト、ローカルコストそれぞれに対して求めた分散であり、 $\hat{G}_g = 2.72 \times 10^{10}$ 、 $\hat{G}_l = 8.04 \times 10^8$  となっている。最終的には、式(4)、(5)で計算する  $\bar{G}_g(i,s)$ 、 $\bar{G}_l(i,s)$  を式(3)中の  $G_g(i,s)$ 、 $G_l(i,s)$  に置き換えたコストを用いて、構図と構成要素の両方の類似度を測る。

## 4 スタイルの転写

提案手法では、第 3 節で述べた方法により探索した構図に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。

$$Cost_{p,q} = (\Delta E^*ab + |\sigma_p - \sigma_q|)(1 + |C_p - C_q|)^2 \quad (4)$$

ここで、 $\Delta E^*_{ab}$  は  $L^*a^*b^*$  色空間における色差であり、 $\sigma$  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。 $\gamma$  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。

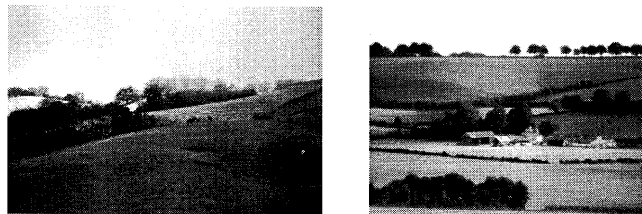
## 4.2 筆遣いの転写

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]法を利用する。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。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領域の輪郭を保存しながら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. 結果

図 4 に生成結果の例を示す。図 4(a)左に示す入力写真に対し“Gogh”の作品 300 枚が含まれている絵画データベースの中で最もシーンが似ている絵画として、図 4 (b)左に示す絵画が検索結果として得られた。写真と絵画の双方に、山、木、草原、空が含まれ、それらの位置関係も近いことが見て取れる。写真に絵画のスタイルを転写した結果が、図 4(c)左に示す出力画像である。写真の中の山、木、草原、空のそれぞれに対して、絵画の色遣いが転写されていることがわかる。また、写真ではほぼ均一の色をしていた空に対して、絵画で表現されているような細かな筆の跡が転写されている。しかし、写真の中の後ろの白い山と明るい茶色山については、検索された絵画の中には似た色の領域がないため、色は変化せず、筆遣いは空と同じものが転写されている。また、絵画の中比較的大きな面積を占める空と草原における筆遣いの違いも出力画像において表現できている。

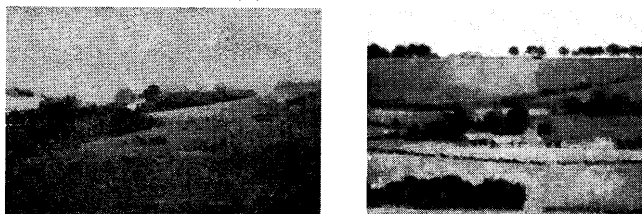
同様に、図 4(a)右に示す入力写真に対し“Monet”の作品 294 枚が含まれている絵画データベースの中で最もシーンが似ている絵画として、図 4 (b)右に示す絵画が検索結果として得られた。こちらも、写真と絵画の双方で、木、空、草原、家などが含まれていることがわかる。一方で、絵画にのみ、人物、赤い花、青空、雲が含まれている。写真に絵画のスタイルを転写した結果が、図 4(c)右に示す出力画像である。木や草原などで、色遣いと筆遣いの反映が成功していることがわかる。絵画にのみ含まれている人物、赤い花、青空の描き方は、出力画像にはほとんど反映されていない。これは、色遣いと筆遣いの転写の際に、位置と色の両方のコストを計算していることによるものである。



(a) 写真



(b) 構図が近い絵画



(c) 生成絵画

図 4 絵画調画像生成結果

## 6. おわりに

入力画像と絵画の構図の類似性を考慮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 ascan, 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.