

## 画像変形による画像審美性の向上

内田 大智 金森 由博 遠藤 結城

筑波大学

## 1. はじめに

写真の美しさを決める要素の一つに構図がある。構図とは、画像内に写る物や人物の配置による構成のことである。プロの写真家は審美的な構図を考慮して写真の撮影や編集を行うが、初心者にとっては難しい。

そこで既存研究では、自動的に画像の構図を変更して、画像の審美性を高める方法が提案されてきた。例えば Hong ら [1] は、構図のルールの明示的なモデリングによる、画像のトリミング手法を提案した。また Zhong ら [2] は、画像の未撮影部分の外挿を行うことで、トリミングの範囲を画像の外側まで拡張した。しかしこれらの既存研究では、主に画像のトリミング操作を扱っており、審美性を高めるための自由度が低い。またこれらの既存研究では、ネットワークの訓練のために大量の正解画像やトリミング範囲や構図ラベルを必要とする。

本研究では、構図を変更する方法として、画像の変形を扱い、正解画像なしに入力画像の審美性を向上させる手法を提案する。具体的には、入力画像のエッジに着目した三角形メッシュを生成し、その三角形メッシュに入力画像を貼り付け、微分可能レンダリングにより画像を出力する。この出力画像を、学習済みの審美性スコア推定ネットワークに入力として与える。一連の処理は微分可能となっており、得られる審美性スコアが最大となるように、三角形メッシュの頂点座標を最適化することで、審美性を高めるような画像変形を実現する。

以上の方法により、既存手法のように大規模な正解データを用意する必要なしに、既存手法よりも自由度の高い探索空間で入力画像の審美性を高めることができる。本研究で使用したすべての画像に対して、審美性スコアが向上することを実験により確認した。

## 2. 提案手法

提案手法の概要を図 1 に示す。微分可能レンダリングにより一連の処理が微分可能となり、誤差逆伝播法により、三角形メッシュの頂点座標を最適化することで変形を実現できる。以下、手順について説明する。

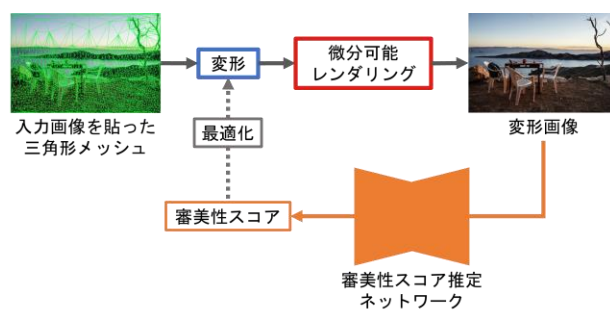


図 1: 提案手法の概要。

## 2.1 三角形メッシュの生成

まず、入力された画像のエッジに沿った三角形メッシュを生成する (図 2 の緑線)。これにより、画像中のエッジに沿った局所形状を保ちながら、画像を変形することが可能になる。

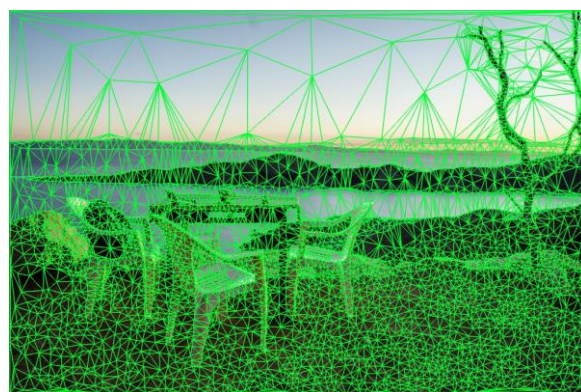


図 2: 画像のエッジに沿った三角形メッシュ。

## 2.2 三角形メッシュの最適化

次に、入力画像を貼り付けた三角形メッシュを微分可能レンダリングを用いて描画し、出力画像を学習済みの審美性スコア推定モデルの入力として与え、変形画像の審美性スコアを得る。

学習済みの審美性スコア推定モデルは、Zhang らによって提案された SAMPNet [3] を用いる。これは複数の構図パターンから画像内に写る物体の視覚的な配置を分析する構図評価ネットワークであり、スコアの範囲は [1, 5] である。最適化器は、この審美性スコア推定モデルから得られた審美性スコアが最大となるよう、三角形メッシュの頂点座標を最適化する。

### 3. 実装と実験

プログラミング言語として Python、ライブラリは PyTorch を用いて提案手法を実装した。微分可能レンダリングのライブラリには nvdiffrast を用いた。最適化の高速化と画像の歪みの抑制のために Nicolet らの手法 [4] を使用した。

最適化のパラメータについて、反復回数は 500、学習率は  $3e-2$ 、三角形メッシュの頂点数は 5,000 に設定した。学習率のスケジューラとして PyTorch の LambdaLR を使用し、1 エポックごとに  $0.97^2$  倍ずつ減衰させた。

入力画像として、Zhang ら [3] の CADB データセットから 2,151 枚の画像を使用した。GPU は NVIDIA RTX A5000 を使用し、一枚の最適化にかかった時間は平均で 19.0 秒であった。

### 4. 結果と考察

以下に実験結果を示す。まず、定量評価について説明する。実験によりすべての画像について審美性スコアが向上した。SAMPNet により得られた入力画像と出力画像の審美性スコアの差の平均値を計算したところ、0.789 であった。これは、審美性スコアが平均して 19.7%ポイント向上したことを意味する。

次に定性評価について説明する。以下に、実験により得られた変形画像を示す (図 3)。

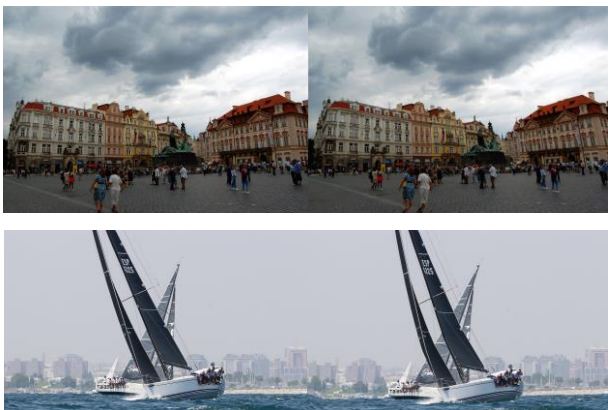


図 3: 実験結果 (左が入力画像、右が出力画像)。

図 3 上段の画像について、建物全体が右に移動するような変形が起こっている。図 3 下段の画像について、ヨットが拡大されるような変形が起こっている。入力画像と比較して、変形画像は変化が少ないことが分かる。これは、Nicolet らの手法 [4] の内部で計算されているメッシュラシアンラシアンの正則化により、画像の歪みが抑えられる代わりに変形が小さくなってしまっていることが原因として考えられる。

また、以下の画像では直線的な物体の歪みが目立ってしまっている (図 4)。この問題に対処するため、文献 [5] で提案されているような、画像内の直線的な物体を保持する工夫を取り入れる必要があると考えられる。

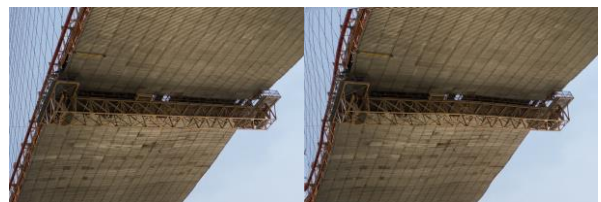


図 4: 直線的な物体が歪む例。  
(左が入力画像、右が出力画像)

### 5. まとめ

本研究では、大規模な正解データなしに、既手法よりも自由度の高い探索空間で入力画像の審美性を高めることができる、画像変形による画像審美性を向上させる手法を提案した。実験に用いたすべての画像で審美性スコアが向上することを確認できた。

### 参考文献

- [1] Hong *et al.*, “Composing Photos Like a Photographer,” In Proc. CVPR, 7057–7066, 2021.
- [2] Zhong *et al.*, “Aesthetic-guided Outward Image Cropping,” ACM Trans. Graph. 40, 6, Article No. 211, 2021.
- [3] Zhang *et al.*, “Image Composition Assessment with Saliency-augmented Multi-pattern Pooling,” arXiv preprint arXiv:1611.05916, 2021.
- [4] Nicolet *et al.*, “Large Steps in Inverse Rendering of Geometry,” ACM Trans. Graph. 40, 6, Article No. 248, 2021.
- [5] 宇都木ら, “直線性を保持する Seam Carving 拡張の提案,” 映情学誌, 66, 4, J108–J113, 2012.