



実シーンの外観操作のための ニューラルスタイル変換による投影画像生成

中村天翼¹⁾, 入山太嗣¹⁾, 小室孝¹⁾

1) 埼玉大学 (〒 338-8570 埼玉県さいたま市桜区下大久保 255)

概要: 本研究では、プロジェクタによる映像投影で実シーンの外観を操作する手法を提案する。提案手法では、スタイル変換の枠組みを用いて、プロジェクタの加法的な出力制約と対象面の反射特性を踏まえて投影映像を最適化することで、映像の加算で実シーンの外観操作を実現する。実験では、加法的制約下において多様なスタイルが転送されるかを検証するため、画像の単純加算によるシミュレーションを行った。さらに、実環境における映像投影実験を行い、提案手法の有効性を示した。

キーワード: プロジェクションマッピング, 空間拡張現実, 画像スタイル変換

1. はじめに

近年、インタラクティブメディアやエンターテインメントの分野では、個人の嗜好に合わせて体験をカスタマイズする技術への注目が集まっている。絵画鑑賞や音楽鑑賞といった体験において、周囲の環境を自分の好みにアレンジすることができれば、より没入感のある体験にすることができる。物理環境をカスタマイズする方法の中でも、プロジェクタによる映像投影が注目されている。この技術は、物理的な加工をすることなく実環境の外観を操作することを可能にし、多様なビジュアルスタイルを試すことができる。

これまでに、映像投影で実物体の外観を操作する研究がいくつか行われてきた。例えば、壁や床、家具に映像コンテンツの延長部分を投影することで、空間的な没入感の向上を実現したものがある [1][2]。また、対象物体の周囲に特定の視覚パターンの映像を投影することで、静止物体に対する見かけの動きの付与や鑑賞者が知覚する色を変化させることを実現した研究もある [3][4]。しかし、これらの手法では、投影対象の形状や反射特性に合わせて手動で投影映像を作成する必要があるため、映像制作に大きな労力とコストを要する。

この問題に対処するため、映像投影対象の物体の物理的特性を推定し、それに応じて投影映像を自動的に補正する手法が提案されている [5][6]。これらの手法は、ニューラルネットワークを用いて投影対象の形状や反射特性を同時にモデル化し、その結果に基づいて投影画像を最適化することで、観測される画像の歪みや色ずれを最小限にした。一方で、これらは目的とする投影映像の再現を前提とした補償的アプローチであり、外観そのものを自在に変化させるような表現的操作には限界がある。

近年では、深層学習の発展により画像の印象を容易に変化させることができる技術が注目されている [7][8]。その中でもスタイル変換と呼ばれる技術は、画像の構造（コンテンツ）と画風（スタイル）の表現が畳み込みニューラルネッ

トワークの層において分離可能であることを利用し、スタイル画像から抽出したスタイルを、任意のコンテンツ画像に転写することで新しい画像を生成する。スタイル変換は元の構造的特徴との整合性を保ちながら印象のみを変化させられるため、投影映像の生成に用いることで、対象物体の外観を活かしながら柔軟で多様なビジュアル表現が可能になる。

本研究では、映像投影による実シーンの外観操作を実現するための、スタイル変換を用いた投影映像の生成手法を提案する。投影映像の生成にスタイル変換の枠組みを用いることで、実シーンの構造的特徴を保持したまま、外観の視覚的印象を変えることが可能となる。しかし、スタイル変換によって生成される画像には、対象面の輝度を下回る画素が含まれることがあり、光を加えることしかできないというプロジェクタの加法的な出力制約下ではこれを正確に再現できない可能性がある。提案手法では、スタイル変換のネットワークにおいて対象面の反射特性に基づいて投影後の視覚的外観をシミュレートし、加算的にしか情報を加えられないという制約の下で画像を最適化することで、投影の性質を踏まえた画像の生成を実現する。

2. 提案手法

2.1 概要

提案手法の概要を図 1 に示す。カメラで取得した投影対象面の画像 $I_c(x, y) \in \mathbb{R}^3$ と任意のスタイル画像 $I_s(x, y) \in \mathbb{R}^3$ を入力として、スタイル変換を行い生成された画像 $I_p(x, y) \in \mathbb{R}^3$ を対象面に投影することで外観操作を行う。この時、投影画像が正確に対象面に合うように、事前に幾何キャリブレーションを行い、カメラとプロジェクタ間の位置関係の補正を行う。また、対象面の反射特性や機器の入出力特性によって投影した画像の見た目が変化するため、撮像される画像 $I'_c(x, y) \in \mathbb{R}^3$ を、式 (2) で表される色の補正のための色変換行列 $M(x, y) \in \mathbb{R}^{3 \times 3}$ を用いて式 (1) で再現した画像 $\tilde{I}'_c(x, y) \in \mathbb{R}^3$ に近づけるように投影画像を最適化する。

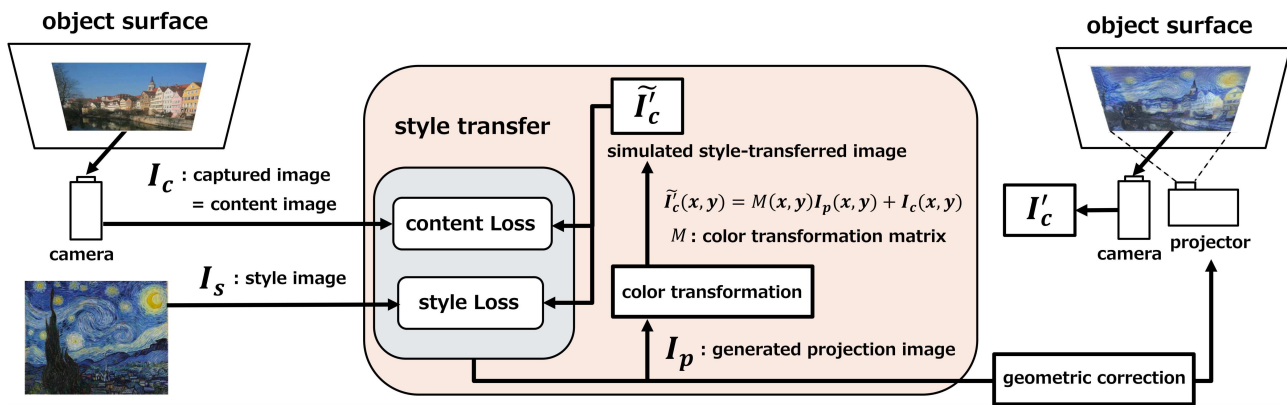


図 1: 提案手法の概要

ここで, $M(x, y)$ の各要素 $v_{ij}(x, y)$ は, プロジェクタのチャネル j の出力がカメラの観測画像におけるチャネル i に与える影響を表す. これにより, 実際の投影環境下でもスタイル転送後の外観が再現できるような画像の生成を実現する.

$$\tilde{I}_c(x, y) = M(x, y)I_p(x, y) + I_c(x, y) \quad (1)$$

$$M(x, y) = \begin{pmatrix} v_{RR}(x, y) & v_{RG}(x, y) & v_{RB}(x, y) \\ v_{GR}(x, y) & v_{GG}(x, y) & v_{GB}(x, y) \\ v_{BR}(x, y) & v_{BG}(x, y) & v_{BB}(x, y) \end{pmatrix} \quad (2)$$

2.2 スタイル変換のネットワーク

提案手法で用いるスタイル変換のネットワークを図 2 に示す. ネットワークは, Gatys ら [7] の手法を基盤とし, 事前学習済みの VGG-19 を用いて, コンテンツ損失とスタイル損失を最小化する. また, 生成される画像は, 式 (1) で計算される投影後の外観をシミュレートした画像にスタイル画像の視覚的特徴が反映されるように最適化される. さらに, プロジェクタには負の輝度値を持つ光を投影することができないという物理的制約があるため, 生成画像に対して ReLU 関数を適用することで非負制約を課す. これにより, 映像投影後の対象面の外観をより正確に再現し, 実空間の投影環境に適した画像生成が可能となる.

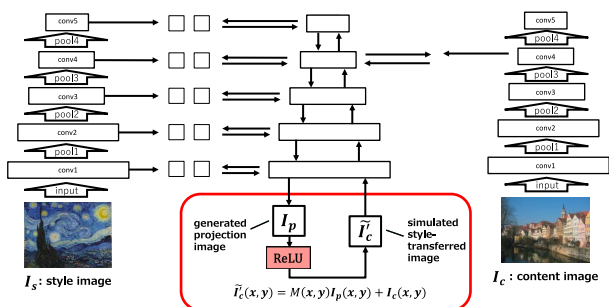


図 2: Gatys ら [7] の手法を拡張し, 最適化過程に制約を組み込んだネットワーク

3. 加法的制約下におけるスタイル変換の検証

映像投影における加法的制約下において, 対象の外観を多様なスタイルに変換できるかを検証するため, 投影画像とコンテンツ画像の単純加算によるシミュレーション実験を行った. 本実験では実際の投影は行っていないため, 事前の幾何キャリブレーションは行っておらず, 式 (1) の色変換行列 M は単位行列に設定した. スタイル変換のネットワークの最適化には, Adam を用いており, 学習率は 0.03, エポック数は 1000 とした. また, コンテンツ損失の重みを 1, スタイル損失の重みを 10^6 に設定した.

3.1 実験結果

図 3, 4 に提案手法におけるシミュレーションの結果と従来手法のスタイル変換の結果を示す. 加法的制約下においても, コンテンツ画像に適切なスタイルを転送することができた. また, 両者の結果には明るさや彩度に違いが見られたものの, スタイル画像から似たパターンが転送されていた. 一方で, 提案手法の結果では, 従来手法では見られなかった色使いが, 川や木々の領域において描写された. また, 提案手法では加法的制約の影響により, スタイル損失の値が従来手法より高くなったが, コンテンツ損失に大きな差が出ることは少なく, 構造的特徴を保ったままスタイル転送ができていたことが示唆された.

3.2 考察

提案手法の結果では, スタイル画像中の明るい領域が選択的に転送される傾向が見られた. これは, 提案手法におけるコンテンツ画像の画素値以上でしかスタイルの転送が行えないという加法的制約に起因していると推測できる. 一方で, 提案手法のネットワークでは, 画素値を増加させる方向にのみ最適化が可能のため, スタイル画像の暗い領域やコントラストの低い特徴の再現は困難であり, 意図しない色の変化やスタイル転送が生じる場合もあった.

4. 映像投影による実シーンの外観操作の実験

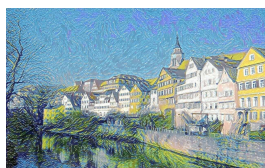
実環境において, 映像投影によって実シーンの外観を多様なスタイルに変換することができるとを検証するため, プロジェクタとカメラを用いた投影実験を行った.



(a) コンテンツ画像 (b) スタイル画像 (c) 投影画像



(d) 従来手法の結果 (コンテンツ損失: 2.6072, スタイル損失: 0.0064)



(e) 提案手法の結果 (a + c) (コンテンツ損失: 2.6796, スタイル損失: 0.0623)

図 3: 提案手法の結果で特有のハイライトが確認できた画像



(a) コンテンツ画像 (b) スタイル画像 (c) 投影画像



(d) 従来手法の結果 (コンテンツ損失: 4.5025, スタイル損失: 0.1084)



(e) 提案手法の結果 (a + c) (コンテンツ損失: 4.1748, スタイル損失: 1.9630)

図 4: 従来手法の結果と提案手法の結果で彩度に大きな差が確認できた画像

4.1 実験条件

実験は蛍光灯の照明下において行った。実験では、プロジェクタとカメラを使用し、図 5 のように投影対象面に対して正対させて配置した。対象面には、図 3 の (a) の画像を印刷した A4 用紙を使用した。カメラの解像度は 4000×3000 で、露光時間は $199955 \mu s$ 、ゲインは 20.0 dB に設定した。プロジェクタの解像度は 1280×800 で、明るさは $25/100$ 、コントラストは $25/100$ に設定した。また、スタイル変換におけるコンテンツ画像にはカメラで取得した投影対象面の画像を用いて、スタイル画像には、図 3, 4 と同様の画像を用いた。

4.2 色変換行列

プロジェクタから RGB の各チャンネルについて画素値を $0 \sim 255$ まで 15 ずつ段階的に変化させた単色画像を対象面に投影し、その反射光をカメラで取得した。次に、プロジェクタ空間の各座標に対応するカメラ画像の座標ごとの画素値を記録し、投影値と取得値の関係に対して線形回帰を適用した。この時、投影値が 255 に到達する前に取得



図 5: 実験環境

値が 255 に到達した場合は、センサの飽和の可能性があるため、それ以降に取得したデータは回帰分析から除外した。回帰分析によって得られた直線の傾きを利用して 3×3 の色変換行列 M を作成した。

4.3 幾何キャリブレーション

まず、プロジェクタを用いて投影対象面上の指定した 8 点の位置に円を投影し、その全体像をカメラで撮影した。次に、撮影画像から各円の中心座標を検出し、プロジェクタで指定した座標との対応点からホモグラフィ行列を求めた。この行列は、プロジェクタの各画素が投影された際にカメラ上で対応する位置を特定し、色補正処理において画素ごとに適切な変換行列を適用するために用いた。

4.4 実験結果

図 6, 7 に提案手法によって生成した画像を実シーンに投影した結果と、従来手法のスタイル変換の結果を実シーンにおいて映像投影で再現した結果を示す。提案手法によって生成した画像は、プロジェクタでの投影を考慮して最適化されているため、追加の色補正などを行わずそのまま投影しても、実シーンの外観のスタイルを適切に変えることができた。一方で従来手法では、生成画像に対してコンテンツ画像との差分を取り、単純に加算しただけでは、実シーンの外観を変えることはできなかった。

4.5 考察

提案手法の結果では、左下の暗部においても、水面の反射や木々の輪郭が強調されることで視認性が向上していた。これは、投影画像の最適化過程において対象面の反射率を考慮した局所的な輝度調整が有効に機能したためであると推測できる。一方で、白い外壁の建物では他の領域と比べてスタイル画像の視覚的特徴が十分に反映されなかった。これは、投影対象面における元々色が明るい領域では、プロジェクタの出力上限に達してしまったことで、色が飽和してしまったことが一因であると考えられる。

5. まとめ

本研究では、映像投影による実シーンの外観操作を実現するための、スタイル変換を用いた投影映像の生成手法を提案した。提案手法では、プロジェクタの加法的な出力制約と対象面の反射特性を踏まえて画像を最適化することで、画像の加算のみでスタイルの転送を実現する。

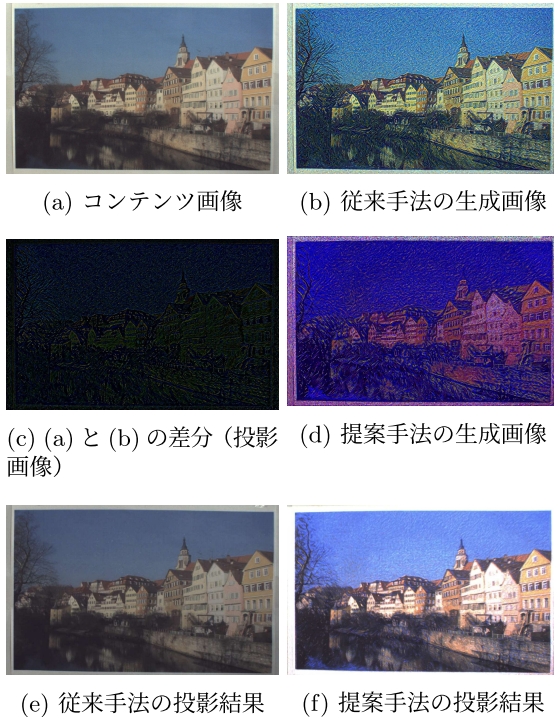


図 6: 提案手法の結果でスタイル画像の筆致が転送された画像

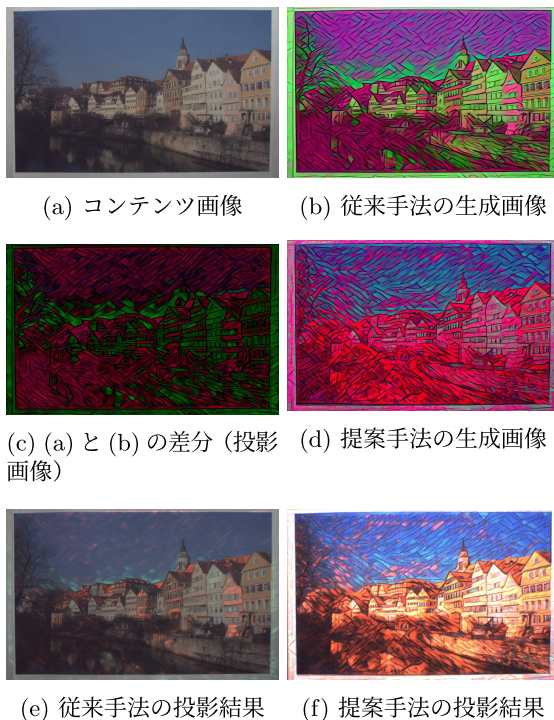


図 7: 提案手法の結果で高いコントラストと彩度が確認できた画像

実験の結果、提案手法は加法的制約下においてもスタイル変換を実行できることが確認された。結果からは、制限された画素値の範囲内でスタイルの転送を行い、スタイル画像に含まれる明るい領域の色が主にスタイル転送に寄与し

ていることが示された。また、映像投影実験から、生成画像をそのまま投影するだけで実シーンのスタイルを変えることができることを確認した。

参考文献

- [1] Brett R Jones, Hrvoje Benko, Eyal Ofek, and Andrew D Wilson. Illumiroom: peripheral projected illusions for interactive experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 869–878, 2013.
- [2] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. Roomalive: Magical experiences enabled by scalable, adaptive projector-camera units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, pp. 637–644, 2014.
- [3] Takahiro Kawabe, Taiki Fukiage, Masataka Sawayama, and Shin'ya Nishida. Deformation lamps: A projection technique to make static objects perceptually dynamic. *ACM Transactions on Applied Perception (TAP)*, Vol. 13, No. 2, pp. 1–17, 2016.
- [4] Ryo Akiyama, Goshiro Yamamoto, Toshiyuki Amano, Takafumi Taketomi, Alexander Plopski, Yuichiro Fujimoto, Masayuki Kanbara, Christian Sandor, and Hirokazu Kato. Illusory light: Perceptual appearance control using a projection-induced illusion. *Computers & Graphics*, Vol. 91, pp. 129–140, 2020.
- [5] Bingyao Huang, Tao Sun, and Haibin Ling. End-to-end full projector compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 6, pp. 2953–2967, 2021.
- [6] Yotam Erel, Daisuke Iwai, and Amit H Bermano. Neural projection mapping using reflectance fields. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 29, No. 11, pp. 4339–4349, 2023.
- [7] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2414–2423, 2016.
- [8] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *Proceedings of European Conference on Computer Vision*, pp. 694–711. Springer, 2016.