

光学ズームによる焦点変化を利用した映り込み除去手法

Reflection Removal Using Focal Length Shifts in Optical Zoom Systems

小田舞遥 宮本拓実 松岡諒
 Maharu Oda Takumi Miyamoto Ryo Matsuoka
 北九州市立大学
 The University of Kitakyushu

1 はじめに

ガラス越しの撮影では、前景の風景がガラスに映り込み、背景と重なってしまう場合がある。このような映り込みは、物体の認識や識別精度を大きく低下させる要因となる。特に、監視カメラや車載カメラ、ドローンといった安全性が求められるシステムにおいては、映り込みによる誤認識や見落としが重大なリスクにつながる。

近年、人工知能技術の進展に伴い、映り込みを効果的に分離・除去する技術の重要性が増している。とりわけ、自動運転や災害対応といった分野では、ガラス越しの映像からも正確な認識を行う必要があり、映り込みの除去は実用化の鍵を握る技術の一つである。

これまでに提案されてきた映り込み除去手法は、大別して深層学習アプローチと数理最適化アプローチに分けられる。深層学習に基づく手法は、高精度な除去結果を得られる一方で、大量の学習データや高性能な計算資源を必要とする点が課題である。一方、数理最適化アプローチでは、物理的・幾何的特徴に基づいて前景と背景を分離することにより、学習データに依存しない除去が可能である。

数理最適化に基づく代表的な研究としては、Li and Brown による、前景と背景の異なる勾配特徴を利用した単一画像からの映り込み除去手法 [1] や、Punnappurath and Brown らによる、デュアルピクセル (Dual Pixel: DP) センサを用いて視差情報を活用する手法 [2] がある。後者では、DP センサが持つ左右視差を利用し、前景と背景の奥行き差を検出することで映り込みの分離を実現している。

本研究では、ズームイン・ズームアウトによる焦点距離の変化に着目し、ガラス面の映り込みを除去する

新たな手法を提案する。具体的には、ズーム操作により生じる描写サイズや奥行きの変化を手がかりとして、観測画像から背景成分を推定する最適化問題を定式化する。また、パッチマッチによりズーム前後の画像間の動きや奥行きを解析し、背景のエッジを抽出する。得られたエッジには、ズーム倍率に応じたガウス分布を適用して重みマップを生成し、背景推定の精度を向上させる。評価実験では、一眼レフカメラで撮影した自然画像を用い、提案手法と既存の最適化ベース手法との比較を通じて、定性的および定量的にその有効性を検証する。

2 従来手法

2.1 DP センサを用いた映り込み分離

Punnappurath and Brown は、DP センサによる視差情報を活用して映り込みを分離する手法を提案している [2]。この手法では、背景はシャープで左右の DP ビュー間に視差がなく、一方、反射成分はボケており視差が生じるという前提に基づいている。

2.1.1 観測モデル

ガラス越しに撮影された画像 $\mathbf{y} \in \mathbb{R}^{3N}$ は、透過成分である背景 \mathbf{b} と、反射成分である前景 \mathbf{r} に分解でき、次式が成り立つ。

$$\mathbf{y} = \mathbf{b} + \mathbf{r} \quad (1)$$

ここで、 N は画像の総画素数を表し、 $\mathbf{y}, \mathbf{b}, \mathbf{r}$ は RGB 画像を 3 チャンネル連結した長さ $3N$ のベクトルとして表現している。

DP センサは、各撮像素子を 2 つのフォトダイオードに分割することで、左右視点に対応した微小視差画像を生成する。このとき、背景成分は焦点が合っているため視差が生じにくく、一方で、前景の反射成分は焦点が外れてボケが生じるため、視差が発生する。本研究ではこの性質を利用し、以下の観測モデルを構築する。

$$\mathbf{y}_L = \frac{\mathbf{b}}{2} + \mathbf{W}_L \mathbf{f} \quad (2)$$

$$\mathbf{y}_R = \frac{\mathbf{b}}{2} + \mathbf{W}_R \mathbf{f} \quad (3)$$

ここで、 $\mathbf{y}_L, \mathbf{y}_R \in \mathbb{R}^{3N}$ は左右それぞれの観測画像、 \mathbf{b} は背景成分、 \mathbf{f} は前景成分、 $\mathbf{W}_L, \mathbf{W}_R$ はボケフィルタである。これらの関係を用いることで、左右画像の合成結果として、元の観測画像 \mathbf{y} を背景成分と前景成分の和として表現する式 (1) が導かれる。

$$\begin{aligned} \mathbf{y} &= \mathbf{y}_L + \mathbf{y}_R = \mathbf{b} + \mathbf{r} \\ \mathbf{r} &= (\mathbf{W}_L + \mathbf{W}_R) \mathbf{f} \end{aligned} \quad (4)$$

2.1.2 視差マップの作成

視差推定には、小パッチ間の勾配差の二乗和 (SSD) を用い、さらに最小二乗フィッティングにより精度を高める。このフィッティングは以下の 2 次関数で表され、その最小値の位置 $s_i = -a_2/a_1$ がサブピクセル精度の視差となる。

$$\frac{1}{2} a_1 x^2 + a_2 x + a_3 \quad (5)$$

視差の信頼度 β_i は以下のように定義される ($\sigma_{a_1} = 5$, $\sigma_{a_3} = 256$)。

$$\beta_i = \exp \left(\frac{\log |a_{1_i}|}{\sigma_{a_1}} - \frac{a_{3_i}}{\sigma_{a_3}^2} \right) \quad (6)$$

この信頼度と視差に基づき、背景に属すると判断された画素には重み c_i を、それ以外の画素には微小値 ϵ_2 ($0 < \epsilon_2 < 1$) を割り当て、以下のように定義する。

$$c_i = \begin{cases} \rho \beta_i, & \text{if } |s_i| < \epsilon_1 \text{ and } \beta_i > 1 \\ \epsilon_2, & \text{otherwise} \end{cases} \quad (7)$$

ここで、 ρ はスケール係数であり、信頼度の高い画素に対して背景成分の推定を優先的に反映させるために $\rho = 5$ に設定する。

2.1.3 最適化問題

得られた重みマップ $\mathbf{C} = \text{diag}(c_i) \in \mathbb{R}^{5N \times 5N}$ および、画像の勾配成分を抽出する畳み込み演算行列 \mathbf{D} を用いて、背景成分 \mathbf{b} を推定するための最適化問題を以下のように定式化する。

$$\min_{\mathbf{b}} \|\mathbf{D}\mathbf{b}\|_p^2 + \lambda \|\mathbf{C}\mathbf{D}(\mathbf{y} - \mathbf{b})\|_2^2 \quad (8)$$

ここで、 $\|\cdot\|_p$ は p -ノルム (通常は $p \leq 1$) の p 乗を表し、 λ は正則化項と整合性項のバランスを調整する重み係数である。また、 \mathbf{D} は画像 \mathbf{b} に対して畳み込み演算を行うことで空間的な勾配成分を抽出する行列であり、水平方向および垂直方向の 1 階微分、ならびに 2 階微分を表現する畳み込み演算行列を縦に連結した縦長の行列である $\mathbf{D} = [\mathbf{D}_x^T, \mathbf{D}_y^T, \mathbf{D}_{xx}^T, \mathbf{D}_{xy}^T, \mathbf{D}_{yy}^T]^T$ 。

この最適化問題は、背景画像と観測画像 \mathbf{y} との整合性を考慮しつつ、信頼度に基づいてエッジの寄与を調整することで、背景成分 \mathbf{b} の空間的な滑らかさを促進しながら推定を行うことを目的としている。

3 提案手法

本章では、DP センサを用いた視差画像に基づく映り込み除去手法 [2] を拡張し、ズームイン・ズームアウトによる焦点距離の変化や描写サイズ・奥行きの違いを手掛かりとした新たな映り込み分離手法について述べる。図 1 に提案手法の全体の流れを示す。

提案手法は以下の手順で構成される。まず、ズーム操作に伴って生じる画像間の幾何的なズレを補正するために幾何補正を適用する。次に、入力画像とズームイン・ズームアウト画像間でパッチマッチを行い、SSD および ZNCC の 2 種類の指標により対応関係を評価する¹。得られた対応に基づき、サブピクセル精度での位置推定を行う。さらに、ズーム倍率に基づいたガウス分布によって重み付けを行い、それぞれの重みを統合して最終的な重みマップを生成する。最後に、この重みマップを用いて最適化問題を解くことで、背景成分と映り込み成分を分離する。

3.1 ズームイン・ズームアウトに基づく観測モデル

ズーム操作によって得られる画像では、焦点が合っている背景成分に対しては視差がほとんど生じない。一

¹SSD は照明変化に弱く、ZNCC の方が安定性が高いため、本研究では ZNCC を採用した。

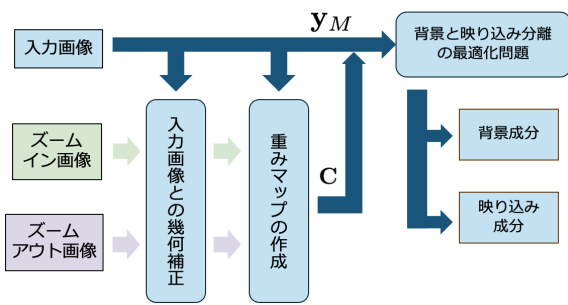


図1: 提案手法の処理手順

方で、前景の映り込み成分には焦点が合っておらず、ボケの影響によるズレが発生すると仮定する。この仮定に基づき、以下の観測モデルを定義する。

$$\mathbf{y}_M = \mathbf{b} + \mathbf{r} \quad (9)$$

$$\mathbf{y}_I = \mathbf{S}_I (\mathbf{b} + \mathbf{W}_I \mathbf{r}) \quad (10)$$

$$\mathbf{y}_O = \mathbf{S}_O (\mathbf{b} + \mathbf{W}_O \mathbf{r}) \quad (11)$$

ここで、 $\mathbf{y}_M, \mathbf{y}_I, \mathbf{y}_O \in \mathbb{R}^{3N}$ は、それぞれ適正画角で撮影された基準画像（入力画像）、および基準からズームイン・ズームアウトして撮影された画像を表す。また、 \mathbf{b} は背景成分、 $\mathbf{r} \in \mathbb{R}^{3N}$ は前景成分（映り込み成分）を示す。 $\mathbf{W}_I, \mathbf{W}_O \in \mathbb{R}^{3N \times 3N}$ はそれぞれ拡大・縮小に伴うボケ作用を表す作用素行列、 $\mathbf{S}_I, \mathbf{S}_O$ はズーム操作に伴って観測画像に生じる幾何変形を表す作用素行列である。

3.2 重みマップの生成

ズームイン・ズームアウト画像は、入力画像とピクセル単位の対応が異なるため、背景成分のエッジが整合するよう幾何補正を行う。その後、入力画像およびズーム画像の勾配を算出し、明るさ変化にロバストなZNCC (Zero-mean Normalized Cross-Correlation) 指標によりパッチ単位の類似度マッチングを行い、位置ずれの度合いを評価する。なお、参照した従来手法 [2] は、SSD 指標を用いて類似度計算を行っていたが、SSD は照明条件などの明るさの変化に弱いため、本研究では明るさの変化に影響を受けにくいZNCCをパッチの類似度計算に採用した。次に、類似度に基づき最小二乗フィッティングを行い、サブピクセル精度で位置ズレの距離を推定する。式 (6) により、位置ずれの信頼度スコアを算出し、さらに式 (7) により閾値処理を行い、重みマップを算出する。その後、ズームイン画像

に対してはガウス分布、ズームアウト画像に対しては反転ガウス分布を適用し、それぞれの画像における信頼度の高い領域に比重を置いたブレンドングにより、最終的な重みマップを生成する。

式 (8) で定義される映り込み分離のための最適化問題において、対角行列 \mathbf{C} の対角成分としてこの重みマップを用いることで、背景成分と映り込み成分の分離に寄与する。

4 実験結果

提案手法の有効性を示すために、一眼レフカメラ (Canon EOS kiss X7) を用いてポータブルガラスを設置して撮影した4シーンの実画像データに対して、従来手法および提案手法を適用し、その結果を定性的および定量的に比較した。

シーン1~3ではズームレンズの焦点距離を $\pm 2\text{mm}$ 、シーン4では $\pm 4\text{mm}$ ずらして撮影を行った。撮影時には、ズーム操作以外の要因による位置ずれや手ぶれを防ぐために三脚を用いてカメラを固定した。定量評価の理想画像としては、ポータブルガラスを物理的に取り外した状態で背景のみを撮影し、その際カメラのパラメータを固定して入力画像との位置ずれがないように調整した。

定量評価には、PSNR (Peak Signal-to-Noise Ratio)、SI (Sharpness Index) [3]、SSIM (Structural Similarity Index) [4] を使用した。PSNRは復元画像と理想画像との画素値の差異を反映する指標であり、値が高いほど復元精度が高いことを示す。SIはエッジ強度に基づいて画像のシャープさを評価する指標であり、細部構造の保持性能を測るのに有効である。SSIMは画像の局所構造・輝度・コントラストを統合的に評価する指標であり、人間の視覚特性に基づいた画像類似度を反映する。

すべての実験において、提案手法のパラメータは $\epsilon_1 = 3$ 、 $\epsilon_2 = 0.001$ 、 $\rho = 5$ に設定した。さらに、視覚的な映り込み除去精度を示すSSIMが最良となるように、シーンごとに正則化重み λ を調整した。また、従来手法のLi and Brown [1] と Yang et al. [5] についても、同様にSSIMが最良となるようにパラメータ調整を行った。表1に、従来手法 [1], [5]、および提案手法 (SSD, ZNCC) それぞれの定量評価結果を示す。いずれの評価指標においても、提案手法が従来手法を上回る性能を示していることが確認できる。

また、図2にシーン2における映り込み除去結果を示す。図2より、従来手法に比べて、提案手法では透

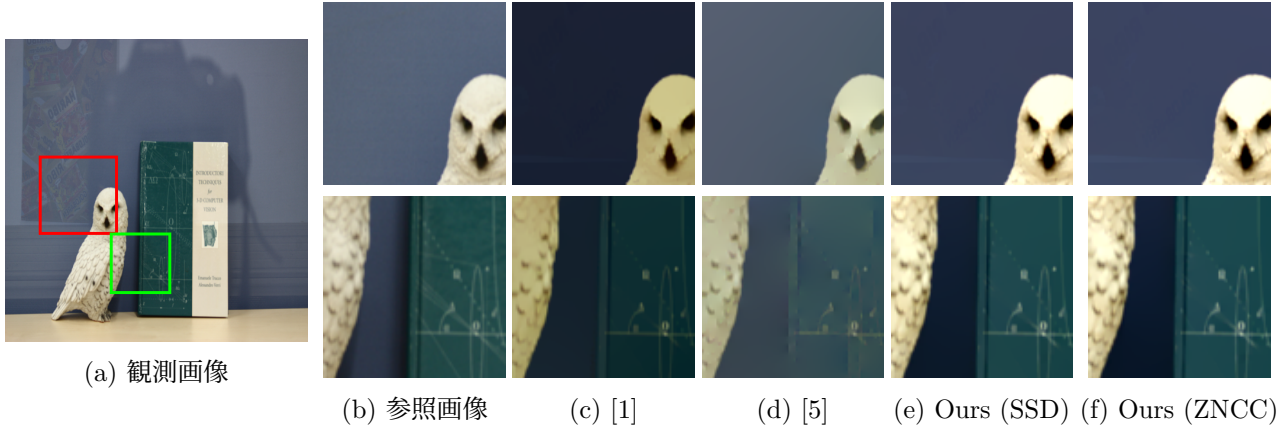


図 2: シーン 2 の映り込み除去結果 (透過成分) の比較

表 1: 各手法の定量評価結果 (PSNR / SSIM / SI)

	シーン 1	シーン 2	シーン 3	シーン 4
Li and Brown [1]	24.10 / 0.990 / 0.930	23.27 / 0.984 / 0.917	20.90 / 0.981 / 0.927	20.90 / 0.981 / 0.927
Yang et al. [5]	26.83 / 0.978 / 0.919	25.38 / 0.967 / 0.878	23.46 / 0.951 / 0.891	23.46 / 0.951 / 0.891
提案手法 (SSD)	26.88 / 0.990 / 0.950	26.18 / 0.986 / 0.943	25.31 / 0.984 / 0.953	25.41 / 0.983 / 0.947
提案手法 (ZNCC)	27.24 / 0.991 / 0.952	26.49 / 0.984 / 0.943	26.43 / 0.985 / 0.960	26.36 / 0.983 / 0.953

過成分のエッジがより明瞭に保持されたまま映り込みが効果的に除去されていることが分かる。

5 まとめ

本研究では、光学ズーム操作によって得られたズームイン・ズームアウト画像を参照し、映り込みが含まれる観測画像から背景成分と映り込み成分を分離する手法を提案した。ズーム操作に伴う焦点距離の変化により、背景成分と前景成分の描写特性に差異が生じるという仮定の下、観測画像の構造的整合性と背景のエッジ情報を活用する最適化問題を定式化した。

実画像を用いた実験により、提案手法は従来手法と比較して、より高精度に背景成分を復元し、映り込みの除去においても優れた性能を示すことが確認された。一方で、前景と背景の分離が完全でない場合や、奥行き・光源条件の変動によって推定精度が低下する場面も見られ、今後はより頑健な前景・背景判別手法の検討や、複数視点の活用による精度向上が課題として挙げられる。

謝辞

本研究の一部は、JSPS 科研費 24K15075 の助成を受けた。

参考文献

- [1] Yu Li and Michael S. Brown. Single image layer separation using relative smoothness. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2752–2759, 2014.
- [2] Abhijith Punnappurath and Michael S. Brown. Reflection removal using a dual-pixel sensor. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1556–1565, 2019.
- [3] Renjie Wan, Boxin Shi, Ling-Yu Duan, Ah-Hwee Tan, and Alex C. Kot. Benchmarking single-image reflection removal algorithms. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 3942–3950, 2017.
- [4] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [5] Yang Yang, Wenye Ma, Yin Zheng, Jian-Feng Cai, and Weiyu Xu. Fast single image reflection suppression via convex optimization, 2019.