

参照画像を用いた高精度画像復元に関する研究

2017年3月

馬場 達也

要旨

本論文では参照画像を用いた画像復元手法を提案する。画像処理において、雑音、ブレおよび色あせなどで劣化した画像から劣化のない原画像を復元することは重要な課題である。近年では画像復元の際、原画像の先験情報を定式化に取り入れ、最小化問題を解くことで画像復元を達成する手法が多く提案され、大きな成果を挙げている。しかし、劣化の激しい画像は有用な情報が少なく、単一画像からの画像復元手法では復元精度が著しく低下する問題があった。そこで、参照画像を用いることで未知の原画像のもつ情報をより多く扱うことで精度の高い画像復元を提案する。

参照画像を用いた画像復元は特に、暗所で撮影された劣化画像の復元において顕著な成果を取めている。これは ISO 感度を高めて撮影した高感度画像、もしくは露光時間を長くして撮影した長時間露光画像に対し、フラッシュを炊いて撮影したフラッシュ画像を参照画像とすると前者の高感度画像と長時間露光画像は色彩豊かであり、後者のフラッシュ画像は鮮明な勾配をもつという異なる画像特徴を有効に活用できるためである。しかし、画像全域においてフラッシュ画像の勾配情報が正しいとする従来の手法は、フラッシュが全域に到達しないような夜間の屋外と言った空間的に広がりを持つシーンにおいては前提が成り立たず、復元結果が不鮮明となる。また、従来の暗部における参照画像を用いた画像復元手法は画像ペアの位置ずれを許容しない。そのため、撮影時に三脚などの固定装置を必要とし、実用化の際、大きな問題となる。従来の位置合わせ手法を用いて位置ずれをおおよそ解消できるが、画素の完全一致を必要とする従来の画像復元手法では性能が悪化する。

1章では本論文の背景と提案について述べ、2章、3章では本論文の関連研究と理解するにあたり必要な基礎知識を紹介する。4章において、フラッシュ画像は一部暗部をもつが雑音をもたないという仮定を用いた雑音除去手法を提案する。また、フラッシュが到達し、鮮明な勾配をもつカメラから近い箇所は雑音で劣化した高感度画像の勾配をフラッシュ画像の勾配に近似する最適化問題を用いて新たに定式化する。5章では位置ずれのある画像ペアにおいて、位置合わせの性能が低い場合においてもフラッシュ画像の鮮明な勾配を保ちつつ、おおよそ位置合わせされた色彩豊かな長時間露光画像の色に近似することで位置ずれのある画像ペアを用いた画像復元をおこなう。暗部のみならず、背景や光源色により色が劣化した人物画像において、前述した手法の知見を基に参照画像を用いて色補正する手法を6章にて提案する。最後に7章にて本論文における提案手法を総括し、まと

めを簡潔にのべる。

Abstract

This paper presents image restoration methods using reference images. Image restoration is one of the important issues in image processing. Recently, many image restoration methods use optimization techniques with effective prior and brought great success. However, it is still challenging to restore an image with strong deterioration. Consequently, a single image restoration tends to have limited performance. In the thesis, we propose image restoration methods with higher accuracy by using a reference image to handle more information of the original image.

For dark scenes, many methods using the reference images have been proposed. They restore a degradation-free image using a no-flash image as an input image and a flash image as the reference image, where the no-flash image is taken by high-ISO sensitivity or long-exposure time and the flash image is taken with strobe. Although the no-flash image has low contrast due to high-ISO sensitivity or long-exposure time, it has a vivid color. Meanwhile, although the flash image has an unnatural color due to strobe, it has high contrast. The many methods integrate a vivid color of the no-flash image and a sharp detail of the flash image to generate the degradation-free image. The conventional methods assume that the flash images have the sharp detail in the whole image, so that, a restoration result is blurred in spatial scenes, which has the region that strobe cannot reach and the contrast is low. Additionally, they require well-aligned image pairs, which are taken with a tripod. The requirement is disadvantageous on a practical side.

In Chapter 1, we describe a background of this paper, while Chapter 2 and 3 introduce related works and basic knowledge necessary for understanding this paper. In Chapter 4, we propose two image restoration methods via convex optimization for high and low-contrast regions in the flash image. In the high-contrast region of the flash image, which is close to the camera and has a sharp detail, we formulate a new approach to transfer the detail from the flash image to the no-flash image. Meanwhile, in the low-contrast region of the flash image, we remove a noise of the no-flash image using a detail of the noise-free flash image. The two region results are appropriately integrated to be the final result. In Chapter 5, we introduce an integration method for a misaligned image pair by locally transferring the colors of the flash images using a small fraction of the corresponding pixels in the

long-exposure images. We formulate image integration as a convex optimization problem with the local linear model. The proposed method makes it possible to integrate the color of the long-exposure image with the detail of the flash image without causing any harmful effects to its contrast, where we do not need perfect alignment between the images by virtue of our new integration principle. We also propose a color correction method not only image restoration for dark scenes. Portrait photographs sometimes have undesirable color around face due to scene light or strong background colors. In Chapter 6, we address the color correction issue using the knowledge on chapter 5. Finally, this research is concluded in Chapter 7.

目次

第1章	序論	12
1.1	本論文の背景	12
1.1.1	画像復元	14
1.1.2	多くの画像復元手法が抱える問題	15
1.2	本論文の目的	15
1.2.1	凸最適化を用いたフラッシュ／高感度画像の合成手法	16
1.2.2	位置ずれのある画像ペアを用いた画像復元手法	16
1.2.3	参照画像を用いた人肌色の自動補正手法	17
1.3	本論文の構成	19
第2章	関連研究	20
2.1	従来の参照画像を用いた画像復元手法	20
2.1.1	フラッシュ／ノンフラッシュ画像合成	21
2.1.2	色変換手法	22
2.2	画像間の対応点探索と位置合わせ	24
第3章	基礎知識	26
3.1	凸最適化問題	26
3.1.1	画像処理において有用な凸関数	26
3.1.2	近接写像	26
3.1.3	凸最適化アルゴリズム	28
3.2	エッジ保存型平滑化フィルタ	32
3.2.1	バイラテラルフィルタ [1]	33
3.2.2	ノンローカルミーニングフィルタ	33
3.2.3	ガイドドフィルタ	35
3.2.4	ジョイントフィルタリング	36
第4章	凸最適化を用いたフラッシュ／高感度画像合成による画像復元	38
4.1	概要	38
4.2	関連手法	39
4.3	提案手法	39
4.3.1	高露光領域用処理	41

4.3.2	低露光領域用処理	43
4.3.3	高露光領域と低露光領域の合成	45
4.4	実験および考察	45
4.4.1	実画像を用いた定性評価	47
4.4.2	復元精度の定量評価	49
第5章	位置ずれのある画像ペアを用いた画像復元手法	51
5.1	概要	51
5.2	局所線形性を用いた画像統合	51
5.2.1	提案手法の概要	51
5.2.2	対応探索と位置合わせ	53
5.2.3	信頼できる対応画素の選択	53
5.2.4	凸最適化を用いた画像統合	54
5.2.5	凸最適化	56
5.3	実験	57
5.3.1	実験におけるパラメータ設定	58
5.3.2	従来法との比較	59
5.3.3	異なる画像ペアを用いた画像復元	61
5.3.4	備考	63
第6章	参照画像を用いた自動肌色補正	64
6.1	概要	64
6.2	提案手法	65
6.2.1	顔検出	66
6.2.2	肌色抽出	67
6.2.3	文献 [2] によるカラーグレーディング	68
6.2.4	最適化を用いたハイブリッド GF	69
6.3	実験および考察	71
6.3.1	自動年鑑写真生成	74
6.3.2	従来手法との比較	77
6.3.3	半自動色補正	79
6.3.4	提案法における失敗例	79
第7章	結論	81
7.1	凸最適化を用いたフラッシュ／高感度画像の合成	81
7.2	位置ずれのある画像ペアの画像合成	81
7.3	参照画像を用いた人の顔の自動肌色補正	82
7.4	提案手法の問題点と将来への展望	82
付録		83

参考文献	85
謝辞	93

表 目 次

4.1 実験2における各処理のパラメータ	50
4.2 画像ペア (d)-(g) の PSNR による評価	50

目次

1.1	屋外撮影時のフラッシュ画像の問題	17
1.2	屋外撮影されたフラッシュ／高感度画像ペア. それぞれ左が高感度画像, 右がフラッシュ画像.	17
1.3	複数枚画像を扱う際の問題点	17
1.4	位置ずれを含むフラッシュ／長時間露光画像ペア. それぞれ左が長時間露光画像, 右がフラッシュ画像.	17
1.5	背景色や光源色による画像劣化	18
1.6	背景色の反射や光源色による顔肌色劣化画像	18
2.1	フラッシュ／高感度画像の統合による雑音除去	20
2.2	フラッシュ／高感度画像の統合によるホワイトバランス補正	20
2.3	フラッシュ／長時間露光画像の合成	22
2.4	Reinhard らの色変換	23
2.5	Pitié らの色変換 [2] とその発展手法 [3]	23
2.6	カラー化 [4] を用いた色変換	24
2.7	HaCohen らの手法による色変換	24
3.1	移動平均フィルタの処理結果	32
3.2	ガウシアンフィルタのフィルタ形状とフィルタ結果	33
3.3	バイラテラルフィルタのフィルタ形状とフィルタ結果	34
3.4	バイラテラルフィルタを複数回処理した結果. 各パラメータは $\sigma_s = 3.2$, $\sigma_r = 0.05$, フィルタサイズは 21×21 である.	34
3.5	ノンローカルミンフィルタによる雑音除去. 各パラメータは $\sigma = 0.1$, フィルタサイズは $ \mathcal{N}_1 = 21 \times 21$, 類似度を測る際の局所領域サイズは $ \mathcal{N}_2 = 3 \times 3$ である.	35
3.6	異なるパラメータ設定によるガイドドフィルタの結果	37
3.7	各ジョイントフィルタリングによる雑音除去. BLF はバイラテラルフィルタ, NLMF はノンローカルミンフィルタ, GIF はガイドドフィルタであり, JBLF, JNLMF, および JGIF は各ジョイントフィルタを意味する. 各フィルタリングのパラメータは結果が良好になるように適宜設定している.	37
4.1	提案手法のフローチャート	40

4.2	高感度画像（上段）とフラッシュ画像（下段）の各成分	42
4.3	フィルタ設計のパラメータを変動した提案法の平滑化 (4.10) と joint-BLF の結果画像. 式 (4.10) のパラメータを $\lambda = 0.5$ とし, joint-BLF の反復回数は 5 とした)	46
4.4	フラッシュマップの例	47
4.5	実験に用いた高感度画像とフラッシュ画像	48
4.6	実験 1 と実験 2 における提案法と従来法 [5] [6] のフラッシュ画像と高感度画像の合成法の比較:(*-1) フラッシュ画像, (*-2) 高感度画像, (*-3) Petschnigg らの結果 [5], (*-4) Shirai らの結果 [6], (*-5) 提案法	48
4.7	単一画像の雑音除去手法 [7,8] との比較	48
5.1	提案手法のフローチャート	52
5.2	画像間の位置合わせと画素選択: (a) フラッシュ画像, (b) 長時間露光画像, (c) 位置合わせされた長時間露光画像, (d) (c) における画素集合, (e) 選択された画素集合. (d),(e) において灰色は位置合わせされた領域と選択された画素を意味する.	54
5.3	各画像の画素値例. 提案手法では, フラッシュ画像 (緑色の線), および長時間露光画像の選択された画素 (青色の点) から赤い線で示される画像を生成する.	55
5.4	画像位置合わせ結果. (a) フラッシュ画像, (b) 長時間露光画像, (c) <i>SIFT-flow</i> (5.1) による位置合わせ. 異なるパラメータ ($\tau, \mathcal{N}_1 $) による画素選択の結果. 青の画素は取り除かれた画素を意味する.	59
5.5	入力画像: 左はフラッシュ画像, 右は長時間露光画像	60
5.6	従来手法との比較. 各行の画像は図 5.5(I1) から (I6) を入力画像とした時の各手法の結果.	60
5.7	フラッシュ/高感度画像ペアにおける画像復元. 左から: (a) フラッシュ画像, (b) 参照画像としての高感度画像, (c) 提案手法における復元結果, (d) 各画像の拡大図.	62
5.8	NRDC [9] の画像ペアにおける色変換. 原画像の色を各手法において参照画像に近似している.	62
6.1	提案手法の流れ	66
6.2	顔検出. (左) 元画像, (中央) 検出された顔領域 (候補), (右) 元画像の顔領域. 青い四角は顔の候補を示す.	67
6.3	各領域の例. (a) 原画像 Y , (b) 肌色領域 Ω_{skin} , (c) 膨張した肌色領域 Ω'_{skin} , and (d) 画像内における各領域. 桃, 赤, 黄, および橙はそれぞれ Ω_{skin} , Ω'_{skin} , Ω_{bg} , and Ω_{bdr} を意味する	67
6.4	抽出された顔肌色領域およびそれらの RGB 色分布. (上段の左から): 目標画像, 原画像, 補正結果. (下段): RGB 色空間における各画像の色分布. 目標画像は所望の肌色を持つ	69

6.5	ハイブリッド GF の入出力. (左から): 原画像, 目標画像, 各領域の Ω_{bg} の Ω_{skin} 目標色, ハイブリッド GF 結果.	70
6.6	輝度補正. 入力画像とフィルタリングされた画像はそれぞれ (6.11) によって分解された強度成分と色成分であり, 入力画像の強度成分とフィルタリングされた画像の色成分が (6.12) によって結合される.	72
6.7	顔肌色補正結果. (左から): (a) 目標画像, (b) 原画像, (c) 顔肌色補正結果. ターゲット画像は Shin らのデータセット [10] であり, 入力画像は Gehler データセット [11] を用いた.	73
6.8	フラッシュ画像上の顔の肌の色補正. (左から): (a) 目標画像, (b) フラッシュ画像, (c) 顔肌色補正結果.	73
6.9	提案手法による自動年鑑写真生成. (a) 目標画像, (b) 原画像, (c) 我々のアルゴリズムによる年鑑写真生成結果. 背景色を有する顔肌色が参照画像の肌色に変換されている.	74
6.10	自動年鑑写真生成のフローチャート. 赤枠は前処理を示し, 青枠は後処理を示す.	75
6.11	欠損領域の修復における類似手法との比較. (左から): (a) 原画像, (b) 参照画像, (c) 各 RGB レイヤーへの Colorization [12], (d) 各 RGB レイヤーへの JBU [13], および (e) 提案手法 (6.5). (b) の黒い画素は欠損を表し, 各赤矢印はアーティファクトを示す.	76
6.12	既存の方法との比較. (左から): (a) 目標画像, (b) 原画像, (c) [3] with [2], (d) NRDC [9], (e) Jaesik <i>et al.</i> [14], および (f) 提案手法.	77
6.13	背景置換結果および [10] との比較. (a) 目標画像, (b) 原画像, (c) 背景置換結果, (d) [10] によるスタイル変換結果, (e) 提案手法アルゴリズムによる年鑑写真生成.	78
6.14	半自動色補正. (a) 目標画像, (b) 目標画像の各領域, (c) 原画像, (d) 原画像の各領域, および (e) 提案法結果. (b) と (d) の色は図 1 のように各領域を示している. 原画像と目標画像は https://pixabay.com より入手した.	80
6.15	皮膚色抽出段階に起因する我々の方法の悪い結果. これは画像中の顔と髪の色が類似しているが原因である.	80
7.1	領域拡張による matting	84

第1章 序論

本章では、本論文で扱う参照画像を用いた画像復元手法に関する研究背景について述べ、従来の参照画像を用いた画像復元手法の問題点を指摘するとともに本研究の主張点を述べる。

1.1 本論文の背景

カメラの普及、特にカメラ機能付き携帯電話の普及により、目で捉えたシーンをカメラで撮影し記録することが容易となった。またそれはSNSの普及も相まってあらゆるシーンにおいて人々はカメラを手に取り、撮影するようになった。現在普及しているデジタルカメラでは、シャッターを押した瞬間から、レンズを通してカメラ内部に取り込まれた光をセンサが逐次溜め込んでいき、フォトダイオードにより電気信号に変換された信号がRAWデータ（生データ）となり、デジタル画像の原料となる。結局、カメラ内部で信号処理された後、デジタル信号として保存され、任意のディスプレイを通されたものが我々が目にしている「画像」である。カメラ機器が高価であるほど、得られる画質は高いが、一部の専門家以外は専用の高価な機器を用いず、一般に普及している汎用デジタルカメラやスマートフォンといったデバイスで撮影するため、撮影環境によっては著しく品質の劣化した画像しか取得できない。

上述したように、センサに取り込まれた光が電気信号となり、適切に処理されることで画像データが生成される。しかし、夜間屋外などの暗所ではシーンの光量が少ないため、日中と同じ設定で撮影すると画像が暗く不鮮明となる。センサ感度を高め、取り込んだ光を電気信号へ変換する際に増幅させることで明るい画像を取得できる。センサ感度を高めての撮影は、光量とともに雑音まで増幅するため雑音で劣化した画像が得られる。センサ感度を上げるのではなく、センサが光を取り込む時間を長く設定することでも明るい画像を取得可能である。しかし、この撮影方法では被写体ブレや手ブレを招くため、ブレた画像が得られる。一方で暗部撮影の際、補助光、つまりカメラのフラッシュを用いてシーンの光量を増幅させる撮影方法も有効である。しかし、この撮影方法はシーンに人工的な光が照射されるため、シーン本来の色を損なう。本論文ではこれら三つの撮影方法で得られる画像を次のように呼ぶ。

- 高感度画像：センサ感度を高めて撮影した画像

- 長時間露光画像：露光時間を長く設定して撮影した画像
- フラッシュ画像：カメラのフラッシュを用いて撮影した画像

便宜上、高感度画像と長時間露光画像の総称をノンフラッシュ画像とする。暗部以外にも画像が劣化する環境はある。白色光源下で撮影された画像は、被写体各々がもつ固有の色を鮮やかに表現するが、有色光源下で撮影された画像では被写体の色に加え光源色が画像全体に分布し、コントラストの低い画像となる。また、強い色の背景があるシーンや、水中画像においても色劣化が生じる。

ここで画像処理の代表的手法を二つに大別する。

画像復元

- 雑音除去。雑音により不鮮明となった画像から雑音を取り除く手法。
- ブレ除去。手ブレや被写体ブレによりコントラストの低下した画像からブレのない高コントラスト画像へ復元する手法。
- 色補正。光源色の影響により歪んだ色分布を歪みのない色分布へ補正する手法。

画像生成

- ハイダイナミックレンジ画像生成。異なる露光で撮影された画像を生成することで実シーンに近いダイナミックレンジをもつ画像を生成する手法。
- パノラマ画像生成。左右、もしくは上下に広がるシーン撮影の際、異なる角度で撮影された画像をつなげる手法。
- 構図編集。傾きを持つ画像の補正や、不要被写体を除去することで画像構図を整える手法。

劣化した画像より元の劣化のない画像を復元する「画像復元」とユーザが望む任意の画像を生成する「画像生成」である¹。暗所のように劣悪な撮影環境により劣化した画像に求められるのは、「画像復元」である。本論文では夜間暗部のように劣悪な撮影環境において撮影され劣化した画像から、高品質な画像を取得する画像復元を提案する。

¹画像内の被写体は何か、重要な被写体はどの個体であるのかなどの画像認識は画像処理分野で重要な課題ではあるが、画像を直接処理する技術ではなく、本論文の趣旨とは異なるため割愛する。

1.1.1 画像復元

ここでは撮影環境により劣化した画像から未知の原画像を推定する画像復元とそのアプローチを簡潔に述べる。

雑音除去手法

古くから研究されており、画像処理の主要研究の一つとして挙げられるのが雑音除去である。実シーンにおける雑音画像は、カメラ内での処理過程において、センサから発せられたセンサ雑音で不鮮明になったものを指す。雑音除去の関連研究は現在も多く提案されている [7, 15–21]。ここでは古典的ではあるが、有効な雑音除去手法である Non-local mean (NLM) 法 [17] について述べる。一般的な雑音除去では画像の平滑化処理で達成される。画像は隣り合う画素同士の画素値が近い性質があり、NLM 法はこの性質を画素から局所領域に拡張し、隣接する局所領域は類似するパターンをもつことに着目している。各画素の平滑化は近傍の局所領域の中心画素を畳み込むことで達成される。NLM 法は局所領域の類似度による重み付けで画像の詳細を残した平滑化を実現している。NLM 法と同様に、局所領域の類似度を用いた雑音除去は多く提案されておりその有効性が確認されている [7, 18]。

ブレ除去

ブレ除去は手ブレや被写体ブレをもつ画像からブレのない画像を復元する手法である [22–27]。ブレの原因である手ブレや被写体ブレの軌跡をカーネルといい、カーネルと原画像の畳み込みによりブレ画像は生成される。そのためブレ除去は (i) カーネル推定、(ii) 推定カーネルを用いたブレ除去の二ステップで構成される。(i) と (ii) のステップを一度繰り返すだけではブレ除去は完全に処理されないため、多くの手法では (i) と (ii) を交互に繰り返し処理することで高精度なブレ除去を達成している。また、ブレ画像においてグリッド化された局所領域ごとにカーネルを推定し、ブレ除去する手法も提案されている [28–31]。

色補正手法

色補正は二つに大別される。一つは有色光源下で撮影された画像に対し、光源色を推定し、推定した光源色で各画素の色チャネルを補正するホワイトバランス補正である。もう一方はカメラの性能などで色あせた画像に対し、所望の色をもつ画像を用いて色彩豊かな画像に補正する色変換に基づく手法である [3, 14, 32–37]。多くの場合、色変換手法は色を形成する RGB やその他の色空間での分布近似で達成される。

画像復元へのアプローチ

画像復元のアプローチにはヒューリスティックな手法、原画像が満たすべき性質を表した式を最小化する最適化問題を用いた手法がある。後者の画像復元を最適化問題に帰着させる手法は近年盛んに議論されている [8, 21, 38–44]。これは画像処理において扱いの難しかった有用な関数や、複数の関数からなる最適化式の効率的な解法が提案されたことが背景にある。更に、ある条件下では、初期値に依存せずに大域的最適解が得られるため、式設計の善し悪しの判断が容易である点、原画像の先見情報を関数として表せるため、設計した関数を他の問題に流用できる点も最適化問題が画像復元で広く普及した要因といえる。

1.1.2 多くの画像復元手法が抱える問題

画像復元手法が抱える問題とその理由を簡潔に述べる。画像復元は、取得した一枚画像より劣化のない原画像へ復元する取り組みが大半を占める。特に最近、上述した画像の滑らかさや色分布などの原画像の先見情報を基に、最適化問題として定式化する手法が活発に議論されている。一枚画像からの画像復元は高い精度で達成されることもあるが、劣化度合いにより、原画像がもつべき詳細や高いコントラストなどを損なう問題は未だ解決されていない。

1.2 本論文の目的

本稿では参照画像を用いた高精度な画像復元を目的とする。復元対象画像に対し、同シーンや類似シーンで撮影された画像を補助として用いる、「参照画像を用いた画像復元手法」が多く提案されている。例えば雑音除去やブレ除去の際、復元対象画像は高周波成分が劣化している。そこで、高周波成分に関して劣化はないが、低周波成分に劣化が見られる画像を補助として、復元対象画像と参照画像の利点を組み合わせることで良好な復元画像が得られる。色補正であれば、色が劣化した画像と類似したシーンで撮影され、尚且つ色彩鮮やかな画像を参照画像とすることで色彩補正する。これら参照画像を用いた画像復元は劣化画像がもつ欠点を補う参照画像を導入しつつも、劣化画像の利点はそのまま利用する。また、これらにはヒューリスティックな手法が多く、更には雑音除去やブレ除去の際は復元対象と参照の画像ペア間で画素が完全に一致している必要がある。色補正であれば補正したい被写体のみに焦点を当てた色補正手法は未だ提案されていない。

本論文では、参照画像ベースの画像復元を最適化問題に帰着させる手法をいくつか提案すると共に実験においてその有効性を示す。

1.2.1 凸最適化を用いたフラッシュ／高感度画像の合成手法

雑音により劣化した高感度画像の雑音除去のため，フラッシュ画像を補助として用いる雑音除去手法を提案する．暗所撮影の際に生じる雑音を画像処理により除去する方法として，フラッシュ画像（参照画像）を用いて，ノンフラッシュ画像の画質劣化を改善する方法がある．従来法の処理は平滑化フィルタに基づく簡易な方法であり，二重輪郭などのアーティファクトが生じる問題があった．本論文では画像をストラクチャ成分とテクスチャ成分に分離し最適化問題を適用することで，コントラストをより鮮明に保持した暗所の復元を実現する．更に，従来法において考慮されていないフラッシュ画像（参照画像）に含まれる低露光領域の取り扱いについて考える．すなわち，屋外のような空間的な広がりをもつシーンでは，フラッシュ光の届かない領域があり，光の届いた領域との相対的なコントラストの差によりノンフラッシュ画像よりも画質が劣化する．図 1.1 は空間的に広がりがあるシーンにおけるフラッシュ画像撮影のイメージ図であり，図 1.2 は実際のフラッシュ画像と高感度画像の例である．本論文ではジョイントバイラテラルフィルタを最適化問題に組み込んだ新しい鮮鋭化手法を低露光領域に適用する．提案手法は，各領域を判別した後，それぞれの領域に前述の処理を適用する．各領域の結果画像は，最終的に入力画像ペアの輝度値から作成されたアルファマップで合成され，全領域において鮮鋭な画像となる．定性評価においてアーティファクトの低減，テクスチャの復元精度を確認する．更に，正解画像を用意した定量評価によって有効性を示す．

1.2.2 位置ずれのある画像ペアを用いた画像復元手法

長時間露光画像で撮影された画像を補助としたフラッシュ画像の色補正によるブレ除去を行う．暗所で高品質な画像得るため，フラッシュ画像と長時間露光画像のための新しい画像統合技術を提案する．従来法のフラッシュ／ノンフラッシュ画像を用いた画像復元は，位置ずれのない画像を必要とするが，実用上好ましくない．フラッシュ／ノンフラッシュ画像ペアを撮影する際には図 1.3 が示すような被写体の動きやカメラの位置ズレが生じ，図 1.4 のような画像ペアが撮影される．提案法では，入力画像ペアの対応する画素のごく一部を使用し，フラッシュ画像の色を局所領域毎に変換することでこの問題を解決する．局所領域毎の変換では，局所的線形性を用いた凸最適化問題としてが画像復元を定式化する．これは長時間露光画像の色をフラッシュ画像のデティールと合成することを可能にし，コントラストに悪影響を及ぼすことなく，更に，画像間の完全な位置合わせを必要としない．実験において，参照画像を用いた色変換の最先端手法よりも優れていることを示す．

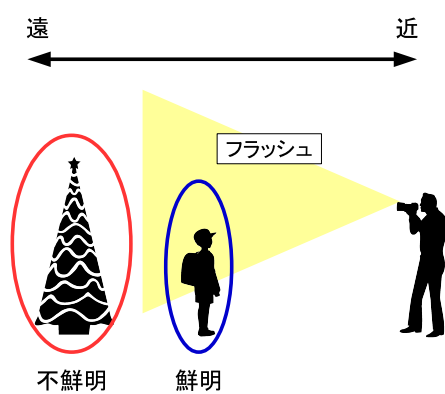


図 1.1: 屋外撮影時のフラッシュ画像の問題



図 1.2: 屋外撮影されたフラッシュ／高感度画像ペア．それぞれ左が高感度画像，右がフラッシュ画像．

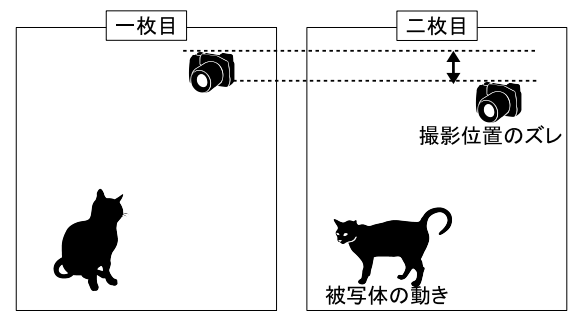


図 1.3: 複数枚画像を扱う際の問題点

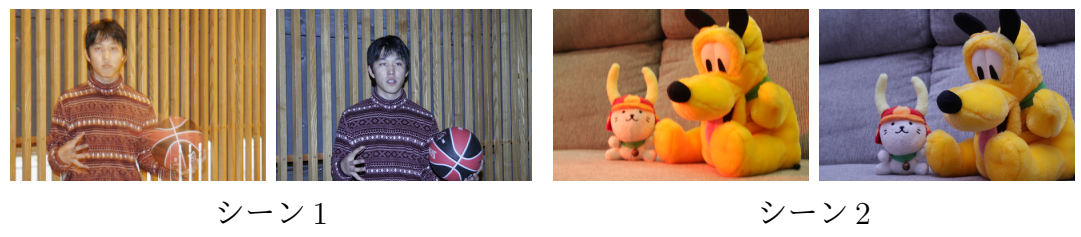


図 1.4: 位置ずれを含むフラッシュ／長時間露光画像ペア．それぞれ左が長時間露光画像，右がフラッシュ画像．

1.2.3 参照画像を用いた人肌色の自動補正手法

背景色や光源色による肌色変化を抑制し，一貫した顔の肌色を保つため顔写真の自動画像補正手法を提案する．顔写真では，環境光（例えば、色付きの壁から反

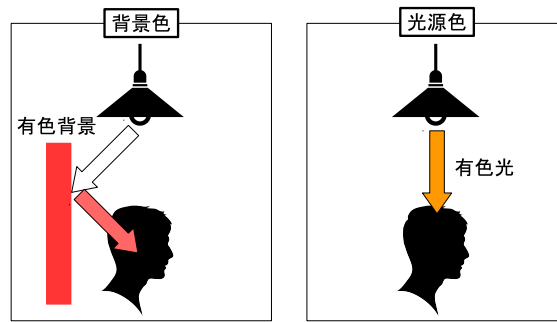


図 1.5: 背景色や光源色による画像劣化



(a) 背景色の反射による劣化



(b) 光源色による劣化

図 1.6: 背景色の反射や光源色による顔肌色劣化画像

射された照明およびカメラストロボ) によって肌色が歪み, 特に画像の背景を入れ替えた場合などに不自然な画像となる. 図 1.5 は問題となるシーンのイメージ図であり, 図 1.6 は実際に顔肌色が劣化した画像である. 図 1.6(a) は同一人物の画像であるが, 背景色の反射や背景色により異なる顔肌色が確認できる. 図 1.6(b) では光源色により色が歪み, 不自然な顔肌色となっている. 提案するアルゴリズムでは, 顔領域を大まかに抽出し, 色空間内の肌色分布を補正した後, それを用いて画像の顔周辺の色や明るさの補正を行い, 輝度の影響を受けない顔画像の色補正を実現する. 色補正の従来のアルゴリズムとは異なり, 最終結果は, 参照画像を用いた色補正処理によって達成される. 特に, 色補正のための我々の参照画像フィルタリングは, Heらによって提案されたガイドドフィルタで必要とされる完全に位置合わせされた参照画像を必要としない. 実験では, 本手法は従来の手法に比べ, 顔写真だけでなく, 自然なシーン (例えば, 背景が複雑な画像) の写真に対しても, より自然な結果が得られることが示す. また, 年鑑写真を自動生成するアプリケーションも提案する.

1.3 本論文の構成

2章では取り扱う各画像復元に対する関連研究を紹介する。3章では、提案する画像復元において、重要な技術である最適化問題の設計と最適化アルゴリズムおよび本研究を理解するにあたり必要な基礎知識を紹介する。4, 5および6章ではそれぞれ1.2.1節, 1.2.2節, 1.2.3節の詳細を述べる。最後に7章で本論文で提案した手法を総括し、本論文の成果を述べる。

第2章 関連研究

本章では4章以降にて述べる提案手法に関連する研究を紹介する。

2.1 従来の参照画像を用いた画像復元手法

本節では、従来の参照画像を用いた画像復元手法をいくつか紹介する。参照画像を用いた画像復元はフラッシュ／ノンフラッシュ画像合成に対して多く提案されている [5, 6, 45–49]。これらの多くはフラッシュ画像の鮮明な勾配とノンフラッシュ画像の鮮やかな色彩を合成することでその両方の利点をあわせもつ画像を生成する。また、色補正手法は画像ペアにおいて色合いの異なる色を一方の色に近似する手法である [2, 9, 12, 32, 33, 36, 37, 50, 51]。



(a) 高感度画像 (b) フラッシュ画像 (c) Petschnigg et al. [5] (d) He et al. [46]

図 2.1: フラッシュ／高感度画像の統合による雑音除去



(a) 高感度画像 (b) フラッシュ画像 (c) Elmar et al. [45]

図 2.2: フラッシュ／高感度画像の統合によるホワイトバランス補正

2.1.1 フラッシュ／ノンフラッシュ画像合成

フラッシュ画像と高感度画像を合成し、鮮明な高感度画像を生成する手法が提案されている [5,6,46]. ここで“合成”とは、各画像の低周波数成分と高周波数成分の分離・組換え・合成処理の他に、平滑化による雑音除去などを含み、最終的に一枚の鮮明な画像を作り出す処理とする。代表的な合成法としては、フラッシュ画像の模様（テクスチャ成分）と高感度画像の色と輪郭（ストラクチャ成分）を合成する方法 [5] と、フラッシュ画像の模様を保持したまま色を変化させる方法 [6,46] がある。図 2.1 にフラッシュ／高感度画像ペアと代表手法の処理結果を示す。前者の [5] では、フラッシュ画像を基準画像として各画像をジョイントフィルタリングを用いて平滑化した後、原画像と差し引くことで各成分に分離し、フラッシュ画像のテクスチャ成分と高感度画像のストラクチャ成分を組合せる。単純な方法であり計算量も短いですが、結果画像の精度が平滑化法のエッジ保存精度に強く依存する問題があり、エッジ保存力や平滑化力の弱いバイラテラルフィルタ (BLF) [1] では輪郭周りにハロや二重輪郭が生じたり、色ムラが残るなどの問題が生じる。後者の [6,46] はガイドドフィルタ (GF) と呼ばれ、画像の局所領域の色分布が直線的または平面的に広がる色線形性 [52] を利用する。フラッシュ画像と高感度画像の対応の取れた領域では色分布の形状が似ており、回転とオフセット補正により容易に色変換が可能となる。問題点としては、局所領域毎に最適化問題を解く必要があり計算量が多くなる点がある。また、基準画像であるフラッシュ画像のテクスチャ成分は処理されないため、フラッシュ画像の模様が不鮮明な場合は結果画像の模様も不鮮明となる。また、Elmar ら [45] は強い色かぶりをもつシーンにおいて撮影されたフラッシュ画像と高感度画像に対し、色成分と輝度成分に分離し、更に各成分をテクスチャ成分とストラクチャ成分に分離後、フラッシュ画像の色成分のテクスチャ成分と高感度画像の輝度成分のストラクチャ成分を合成することで光源色を効果的に除去する手法を提案している。図 2.2 は Elmar らの手法の結果であり、コントラスト低下の原因である高感度画像の強い光源色を除去できている。

一方で、フラッシュ画像と長時間露光画像を合成し、長時間露光の色味を持ちつつ、フラッシュ画像の鮮明な勾配をもつ画像を復元する手法も提案されている [47,48]. これらは [46] でフラッシュ画像と長時間露光画像を平滑化後、平滑化された長時間露光画像にフラッシュ画像と平滑化されたフラッシュ画像の差分を足し込む操作を反復処理することで、徐々に長時間露光画像のブレの原因である高周波成分を補正する手法である。図 2.3 の結果が示すようにブレがなく、長時間露光画像の鮮やかな色彩をもつ画像が生成可能である。単一画像でのブレ除去のようにカーネル推定後、推定したカーネルを用いてブレ除去する手法の際、フラッシュ画像を補助として用いる手法も提案されている [53].

また、文献 [49] では、露光不足で暗く撮影された画像を長時間露光画像を補助として用いた暗所撮影画像の復元手法が提案されている。



図 2.3: フラッシュ／長時間露光画像の合成

これらの従来法に共通する問題点が二つ挙げられる．一つ目は，フラッシュ画像を絶対的な基準画像とみなす点があげられる．すなわち，実際に撮影したペア画像では，ノンフラッシュ画像よりもフラッシュ画像が不鮮明となる場合があり，結果画像に影響が生じる．例えば，被写体と背景の建造物の間に空間的に広がりがある場合，フラッシュ光の届く近距離の被写体は光を強く反射するが，遠方の背景は暗いままとなる．なぜなら，露光量は前者を基準として調整され，相対的に後者はより暗く写り不鮮明となる．二つ目は，画像ペア間で位置ずれを許容しない点である．フラッシュ／長時間露光画像ペアはズレが生じているように見えるが，実際にはフラッシュ画像の画素位置を中心としてブレが生じており，大きくズレが生じた際には画像復元の性能が悪化する．

2.1.2 色変換手法

色補正手法は色変換手法で達成されるため，ここでは従来の色変換手法を紹介する．Reinhard ら [32] は，二つの画像間の色に関する統計的特徴を一致させることによって色の特性を転送する色補正（カラーグレーディングとも呼ばれる）を導入した．これは原画像の Lab 色空間における，明度 L と色度 a, b の一次元分布を平行移動とスケールで参照画像の分布へ近似するもので図 2.4 は [32] による色変換結果である．平均色の調整といった単純な色補正であれば Reinhard らの手法が有効である．Pitié ら [33] は色のヒストグラムマッチングを用いた色変換を実現している．これは原画像の RGB 色空間の三次元分布を，ラドン変換を用いて参照画像の分布へ近似する．この手法は色分布を完全に一致可能であるが，画像の空間情報を損なうため，図 2.5(c) からわかるように空や家などに疑似エッジのようなアーティファクトが生じる．これはゲインアーティファクトと呼ばれ，Rabin ら [3] はこの問題に対し，原画像と平滑化された原画像の差分から生成されるテクスチャを，エッジ保存型平滑化フィルタで Pitié らの結果を平滑化した画像へ足し込むことで解決した．図 2.5(d) は Rabin らの手法を図 2.5(c) に施した結果であり，ゲインアーティファクトの除去が確認できる．後に幾つかの発展手法も提案されてる [2, 37]．ヒストグラム近似に基づくこれらの手法は，画像ペアに位置ずれや画像サイズの制約がないため多くのシーンに柔軟に適用することができる．しか

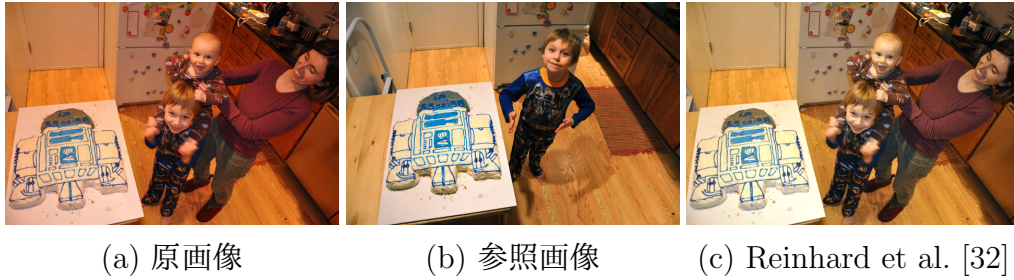


図 2.4: Reinhard らの色変換



図 2.5: Pitié らの色変換 [2] とその発展手法 [3]

し，画像の空間的な色の変化において十分な検討がなされておらず，結果として有色照明が空間的に変化する場合，生成される画像は不自然または偽色を有する傾向がある。

Levin ら [4] は，部分的に色情報が与えられたグレースケール画像に対するカラー化手法を提案した．カラー画像をグレースケールし，ユーザが任意の色情報を与えることで色変換として扱うことができ，文献 [4] の手法を用いた色変換結果を図 2.6 に示す．色付けが手動であった Levin らの手法に対し，文献 [50,51] では，参照画像ベースのカラー化が提案されている．参照画像ベースのカラー化は，画像ペア間で類似のテクスチャを見つけ，そのテクスチャの持つ色をグレースケール画像に着色する．参照画像ベースのカラー化手法は自動であるが，セグメント化された参照テクスチャが必要であることと十分な着色量が必要となる．また，これらのカラー化手法は代表的な色を正確に特定する必要があり，色の不一致が結果として得られる画像に違和感を与える可能性がある。

HaCohen ら [9] は NRDC と呼ばれるのカラーグレーディングのための手法を提案した．NRDC は，二つの入力画像間（図 2.7(a)(b)）の密な対応（図 2.7(c)）を探し，見つかった対応関係を使って RGB チャンネルのトーンカーブ（図 2.7(d)）を設計して色変換をおこなう．処理結果を図 2.7 に示す．色変換結果は参照画像の色味を有し，高い精度での色近似が確認できる．文献 [36] では，NRDC 同様に密な画素対応を必要とし，各色がどの色に変換されるかを確率的モデルで推定することで正確な色変換を実現する．これらの両手法は，画像ペアの画素が完全に一致していなくてもうまく機能する．しかし，色の差が大きい画像や照明の局所的な変化では，両手法とも失敗する．

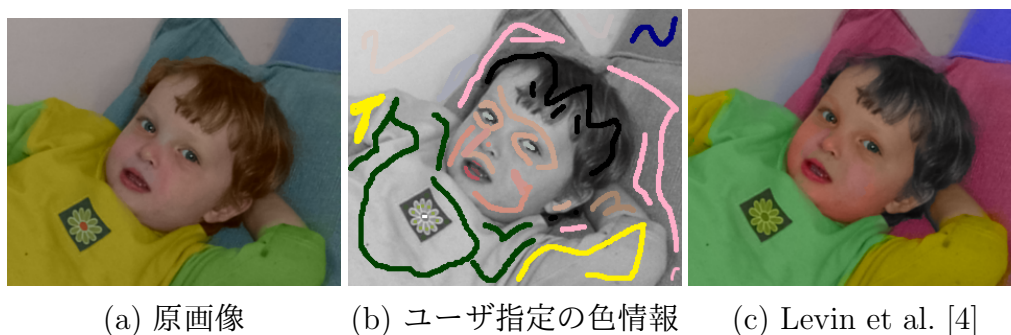


図 2.6: カラー化 [4] を用いた色変換

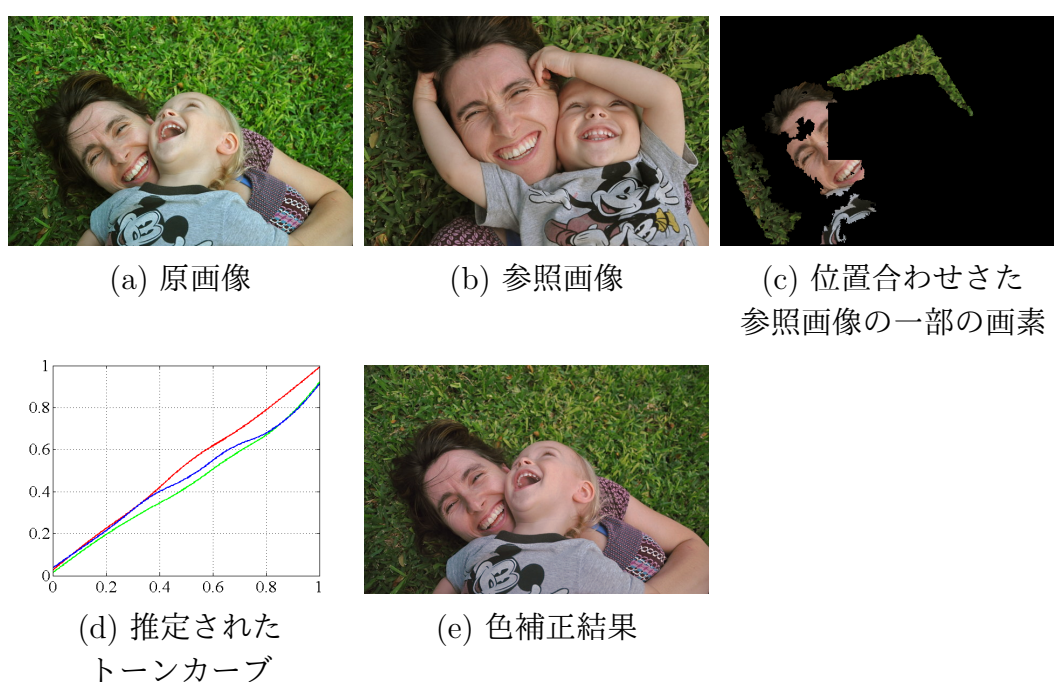


図 2.7: HaCohen らの手法による色変換

2.2 画像間の対応点探索と位置合わせ

5章の位置ずれのある画像ペアにおいて重要な技術である画像間の位置合わせ手法について紹介する. 対応点探索 [54-58] は, 位置ずれのある画像ペアにおいて重要な技術である. SIFT特徴量 [59] のような古典的アプローチは, 特徴ベクトルを用いて画像内の局所的特徴を表現する. SIFT 記述子はスケーリングと回転に対して不変であり, 文献 [59] で実験的に確認されているように, さまざまな照明条件に対して堅牢に動作する. 他の密な対応点探索 [9, 54, 55, 60] は, 特に大きな変位や非剛体の動きを持つ画像ペアの場合, 疎な対応付け方法よりも堅牢な対応探索を達成する. NRDC [9] は, 各画素を中心とする局所領域の対応探索のため [56, 61] をより一般化した色変化に適応した手法であり, 二つの画像間の密接な対応を見

つけ、非剛体の動きを持つ画像を扱うことができるが、照明が空間的に変化し、色のバランスが局所的に変化するときにはうまく機能しない。画像間に色変化があり、色変化が均一でない場合は SIFT 特徴量に基づく [54] が有効な位置合わせ手法である。

第3章 基礎知識

ここでは本手法を理解するに辺り重要な技術である凸最適化問題に関する基礎知識とその解法，本章において重要な画像処理技術について紹介する。

3.1 凸最適化問題

近年では多くの文献が画像復元問題を凸最適化問題に帰着させることで画像復元の性能を向上させた [8, 21, 38–44].

3.1.1 画像処理において有用な凸関数

ここでは画像処理において有用かつ本論文で用いる凸関数を幾つか紹介する。任意の $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{X} \in \mathbb{R}^{n \times m}$ に対し以下のように定義する関数を本論文で用いる。

- l_2 ノルム : $\|\mathbf{x}\|_2 = \sqrt{\sum_i^N x_i^2}$.
 l_2 ノルムを用いて $\|\mathbf{x}\|_2^2 = \sum_i^N x_i^2$ も画像処理で頻出する関数である。
- l_1 ノルム : $\|\mathbf{x}\|_1 = \sum_i^N |x_i|$
- フロベニウスノルム : $\|\mathbf{X}\|_F = \sqrt{\sum_i^n \sum_j^m x_{i,j}^2}$
- 指示関数 : $\iota_{\mathcal{C}}(\mathbf{x}) := \begin{cases} 0 & \text{if } \mathbf{x} \in \mathcal{C}, \\ +\infty & \text{otherwise,} \end{cases}$

ここで, \mathcal{C} は凸集合であり, 上述した関数で構成される最小化問題は次節以降紹介する近接写像 [62] と凸最適化アルゴリズムを用いることで最適解が得られることが保証される。

3.1.2 近接写像

前節で紹介した凸関数は近接写像 [62] が効率的に計算可能である。近接写像は以下のように表される。

$$\text{prox}_{\kappa G}(y) := \arg \min_x G(x) + \frac{1}{2\kappa} \|y - x\|_2^2. \quad (3.1)$$

ここで、 $G(\cdot)$ は前述した非可微分であるが、近接写像が効率的に計算可能な関数であり、 $\kappa > 0$ はパラメータである。

l_1 ノルムの近接写像は次のようにソフト閾値処理として表される。

$$[\text{prox}_{\kappa\|\cdot\|_1}(\mathbf{x})]_i = \text{sign}(x_i)\max(|x_i| - \kappa, 0) \quad (3.2)$$

指示関数の近接写像は集合 \mathcal{C} に依存するが、本論文で扱うのは $\mathbf{x} \in \mathbb{R}^N$ と任意の変数 $\mathbf{o} \in \mathbb{R}^N$ との誤差を η まで許容する際に用いる、次式で表される集合である。

$$\mathcal{C}_{\{\mathbf{o}, \eta\}} = \{\mathbf{x} \in \mathbb{R}^N \mid \|\mathbf{x} - \mathbf{o}\|_2 \leq \eta\} \quad (3.3)$$

このとき、 $\iota_{\mathcal{C}_{\{\mathbf{o}, \eta\}}}$ に対する近接写像は次式のような距離射影となる。

$$P_{\mathcal{C}_{\{\mathbf{o}, \eta\}}}(\mathbf{x}) = \text{prox}_{\iota_{\mathcal{C}_{\{\mathbf{o}, \eta\}}}}(\mathbf{x}) = \begin{cases} \mathbf{x} & \text{if } \mathbf{x} \in \mathcal{C}_{\{\mathbf{o}, \eta\}}, \\ \mathbf{o} + \eta \frac{\mathbf{x} - \mathbf{o}}{\|\mathbf{x} - \mathbf{o}\|_2} & \text{otherwise,} \end{cases} \quad (3.4)$$

\mathbf{x} が条件を満たさない時、 \mathbf{o} を中心に η の半径をもつ球状の側面に射影することから、 l_2 ボール射影と呼ばれる。また、変数 \mathbf{x} の全要素ではなく、任意の要素での誤差を測る場合も同じように処理することができ、対角に $\{0, 1\}$ を持つマスク行列 $\mathbf{M} \in \mathbb{R}^{N \times N}$ を用いて次式のようなになる。

$$P_{\mathcal{C}_{\{\mathbf{o}, \eta, \mathbf{M}\}}}(\mathbf{x}) = \text{prox}_{\iota_{\mathcal{C}_{\{\mathbf{o}, \eta, \mathbf{M}\}}}}(\mathbf{x}) = \begin{cases} \mathbf{x} & \text{if } \mathbf{x} \in \mathcal{C}_{\{\mathbf{o}, \eta, \mathbf{M}\}}, \\ (\mathbf{U} - \mathbf{M})\mathbf{x} + \mathbf{M} \left(\mathbf{x} + \eta \frac{\mathbf{M}(\mathbf{x} - \mathbf{o})}{\|\mathbf{M}(\mathbf{x} - \mathbf{o})\|_2} \right) & \text{otherwise,} \end{cases} \quad (3.5)$$

ここで、 $\mathcal{C}_{\{\mathbf{o}, \eta, \mathbf{M}\}} = \{\mathbf{x} \in \mathbb{R}^N \mid \|\mathbf{M}(\mathbf{x} - \mathbf{o})\|_2 \leq \eta\}$ 、 $\mathbf{U} \in \mathbb{R}^{N \times N}$ は単位行列である。つまり、マスク行列で抽出される画素の誤差が許容誤差を上回った場合においても、マスク行列で抽出されない画素は入力をそのまま帰す操作となる。また、推定する画像の上限値と下限値を設ける場合にも指示関数による近接写像用いられる。この時、集合を次のように定義する

$$\mathcal{C}_{\underline{w}, \bar{w}} = \{\mathbf{x} \in \mathbb{R}^N \mid \underline{w} \leq x_i \leq \bar{w} \text{ for } i = 1, 2, \dots, N\} \quad (3.6)$$

指示関数 $\iota_{\mathcal{C}_{\underline{w}, \bar{w}}}$ に対する近接写像は次式で与えられる。

$$[P_{\mathcal{C}_{\underline{w}, \bar{w}}}(\mathbf{x})]_i = [\text{prox}_{\iota_{\mathcal{C}_{\underline{w}, \bar{w}}}}(\mathbf{x})]_i = \begin{cases} \underline{w} & \text{if } x_i < \underline{w} \\ \bar{w} & \text{if } x_i > \bar{w} \\ x_i & \text{otherwise} \end{cases} \quad (3.7)$$

上述した凸関数以外にも、近接写像が計算可能であり有用な凸関数は多数あるが本論文では使用しないため割愛する。その他の凸関数に関しては [63] を参照されたい。

3.1.3 凸最適化アルゴリズム

凸最適化アルゴリズムは解く問題に合わせて適したアルゴリズムを選択することで効率よく解が得られる．ここでは本論文で用いる三つの凸最適化アルゴリズム [64–66] を紹介する．凸最適化アルゴリズムは初期値に依存することなく大域的な最小解に収束することが保証されている．また，各アルゴリズムの $F(\cdot)$ は微分可能な凸関数， $G(\cdot)$ ， $H(\cdot)$ は微分不可であるが近接写像が効率的に計算可能な凸関数を意味する．

MFISTA [64]

ここで紹介する MFISTA は ISTA の発展系であり，解ける問題は少ないが，高速に収束することで知られている．MFISTA は以下の問題に対して適用可能である．

$$\min_{\mathbf{x} \in \mathbb{R}^N} F(\mathbf{x}) + G(\mathbf{x}) \quad (3.8)$$

与えられた $\mathbf{x}^0 =: \mathbf{v}^1 \in \mathbb{R}^N$ と $t^1 := 1$ に関して，MFISTA の反復は，以下の五つのステップで構成される：

$$\left\{ \begin{array}{l} \hat{\mathbf{z}}^k := \mathbf{v}^k - \frac{1}{L} \nabla F(\mathbf{v}^k) \\ \mathbf{z}^k := \text{prox}_{G/L}(\hat{\mathbf{z}}^k) \\ \mathbf{x}^k := \arg \min \{ F(\mathbf{z}) \mid \mathbf{z} \in \{\mathbf{z}^k, \mathbf{x}^{k-1}\} \} \\ t^{k+1} := \frac{1 + \sqrt{1 + 4(t^k)^2}}{2} \\ \mathbf{v}^{k+1} := \mathbf{x}^k + \frac{t^k}{t^{k+1}}(\mathbf{z}^k - \mathbf{x}^k) + \frac{t^k - 1}{t^{k+1}}(\mathbf{x}^k - \mathbf{x}^{k-1}) \end{array} \right. \quad (3.9)$$

ここで， $1/L$ はステップサイズである．

次式を MFISTA で求解する手順を説明し，各ステップを確認する．

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{K}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad (3.10)$$

ある信号 $\mathbf{y} \in \mathbb{R}^{N_1}$ を辞書 $\mathbf{K}^{N_1 \times N_2}$ とスパースな係数 $\mathbf{x} \in \mathbb{R}^{N_2}$ で表現するという定式化となっている．ここで，各関数を $F(\mathbf{x}) = \frac{1}{2} \|\mathbf{K}\mathbf{x} - \mathbf{y}\|_2^2$ ， $G(\mathbf{x}) = \lambda \|\mathbf{x}\|_1$ とおく．一ステップ目と三ステップ目に関して次式を計算する

$$\begin{aligned} F(\mathbf{x}) &= \frac{1}{2} \|\mathbf{K}\mathbf{x} - \mathbf{y}\|_2^2 \\ &= \frac{1}{2} (\mathbf{K}\mathbf{x} - \mathbf{y})^\top (\mathbf{K}\mathbf{x} - \mathbf{y}) \\ &= \frac{1}{2} (\mathbf{x}^\top \mathbf{K}^\top \mathbf{K} \mathbf{x} - 2\mathbf{x}^\top \mathbf{K}^\top \mathbf{y} + \mathbf{y}^\top \mathbf{y}) \\ \nabla F(\mathbf{x}) &= \mathbf{K}^\top \mathbf{K} \mathbf{x} - \mathbf{K}^\top \mathbf{y} \end{aligned} \quad (3.11)$$

二ステップ目は l_1 ノルムの prox であり, $\kappa = \frac{\lambda}{L}$ として $\text{prox}_{\frac{\lambda}{L}\|\cdot\|_1}(\mathbf{x})$ を計算すればよい. 結局各ステップは以下のように書き換えられる.

$$\left\{ \begin{array}{l} \hat{\mathbf{z}}^k := \mathbf{v}^k - \frac{1}{L}(\mathbf{K}^\top \mathbf{K} \mathbf{v}^k - \mathbf{K}^\top \mathbf{y}) \\ \mathbf{z}^k := \text{prox}_{\frac{\lambda}{L}\|\cdot\|_1}(\hat{\mathbf{z}}^k) \\ \mathbf{x}^k := \arg \min \left\{ \frac{1}{2}(\mathbf{z}^\top \mathbf{K}^\top \mathbf{K} \mathbf{z} - 2\mathbf{z}^\top \mathbf{K}^\top \mathbf{y} + \mathbf{y}^\top \mathbf{y}) \mid \mathbf{z} \in \{\mathbf{z}^k, \mathbf{x}^{k-1}\} \right\} \\ t^{k+1} := \frac{1 + \sqrt{1 + 4(t^k)^2}}{2}, \\ \mathbf{v}^{k+1} := \mathbf{x}^k + \frac{t^k}{t^{k+1}}(\mathbf{z}^k - \mathbf{x}^k) + \frac{t^k - 1}{t^{k+1}}(\mathbf{x}^k - \mathbf{x}^{k-1}) \end{array} \right. \quad (3.12)$$

ADMM [65]

ADMM は画像復元において TV 正則化など幅広い問題を解くことができる. ADMM が適用可能な問題は次式で表される.

$$\min_{\mathbf{x} \in \mathbb{R}^{N_1}, \mathbf{z} \in \mathbb{R}^{N_2}} F(\mathbf{x}) + G(\mathbf{z}) \text{ s.t. } \mathbf{z} = \mathbf{A}\mathbf{x} \quad (3.13)$$

ここで, 各変数 $\mathbf{x} \in \mathbb{R}^{N_1}$, $\mathbf{z} \in \mathbb{R}^{N_2}$ はそれぞれ所望信号と補助変数であり, $\mathbf{A} \in \mathbb{R}^{N_1 \times N_2}$ は線形作用素である. 任意の $\mathbf{x}^0, \mathbf{z}^0$ に対し, ADMM は以下の三ステップで構成される.

$$\left\{ \begin{array}{l} \mathbf{x}^{k+1} := \arg \min_{\mathbf{x}} F(\mathbf{x}) + \frac{1}{2\gamma} \|\mathbf{z}^k - \mathbf{A}\mathbf{x} - \mathbf{d}^k\|_2^2 \\ \mathbf{z}^{k+1} := \arg \min_{\mathbf{z}} G(\mathbf{z}) + \frac{1}{2\gamma} \|\mathbf{z} - \mathbf{A}\mathbf{x}^{k+1} - \mathbf{d}^k\|_2^2 \\ \mathbf{d}^{k+1} := \mathbf{d}^k + \mathbf{A}\mathbf{x}^{k+1} - \mathbf{z}^{k+1} \end{array} \right. \quad (3.14)$$

ここで, $\mathbf{d} \in \mathbb{R}^{N_2}$ は双対変数である.

TV 正則化を例にとり, ADMM の各ステップの例を示す. TV 正則化には様々な種類があるが, ここで用いるのは最も古典的であり, 水平垂直微分の l_1 ノルムを最小化する TV 正則化である. 劣化信号 $\mathbf{y} \in \mathbb{R}^N$ より原信号 $\mathbf{x} \in \mathbb{R}^N$ を TV 正則化で求める問題は次式で与えられる.

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{D}\mathbf{x}\|_1 \quad (3.15)$$

ここで, $\mathbf{D} = [\mathbf{D}_h; \mathbf{D}_v] \in \mathbb{R}^{2N \times N}$ であり, $\mathbf{D}_h \in \mathbb{R}^{N \times N}$ と $\mathbf{D}_v \in \mathbb{R}^{N \times N}$ はそれぞれ水平, 垂直方向の微分を算出するフィルタ行列である. まず, 補助変数 $\mathbf{z} \in \mathbb{R}^{2N}$ を導入する.

$$\min_{\mathbf{x}, \mathbf{z}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{z}\|_1 \text{ s.t. } \mathbf{z} = \mathbf{D}\mathbf{x} \quad (3.16)$$

各関数と変数を $F(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2$, $G(\mathbf{z}) = \|\mathbf{z}\|_1$, $\mathbf{A} = \mathbf{D}$ として ADMM に適用可能である。ここで、さらに拡張ラグランジュにより制約なしの問題へとする。

$$\min_{\mathbf{x}, \mathbf{z}} \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{z}\|_1 + \frac{\mu}{2}\|\mathbf{z} - \mathbf{D}\mathbf{x} - \mathbf{d}\|_2^2 \quad (3.17)$$

\mathbf{x} , \mathbf{z} に対しての部分問題がそれぞれ ADMM の一ステップ目、二ステップ目に相当することに注意されたい。 \mathbf{x} に対する部分問題は

$$\begin{aligned} \mathbf{x}^* &:= \arg \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 + \frac{\mu}{2}\|\mathbf{z} - \mathbf{D}\mathbf{x} - \mathbf{d}\|_2^2 \\ &:= \arg \min_{\mathbf{x}} \frac{1}{2}(\mathbf{x} - \mathbf{y})^\top(\mathbf{x} - \mathbf{y}) + \frac{\mu}{2}(\mathbf{z} - \mathbf{D}\mathbf{x} - \mathbf{d})^\top(\mathbf{z} - \mathbf{D}\mathbf{x} - \mathbf{d}) \\ &:= \arg \min_{\mathbf{x}} \frac{1}{2}(\mathbf{x}^\top\mathbf{x} - \mathbf{x}^\top\mathbf{y} + \mathbf{y}^\top\mathbf{y}) \\ &\quad + \frac{\mu}{2}(\mathbf{z}^\top\mathbf{z} - 2\mathbf{z}^\top\mathbf{d} + \mathbf{d}^\top\mathbf{d} - 2\mathbf{x}^\top\mathbf{D}^\top(\mathbf{z} - \mathbf{d}) + \mathbf{x}^\top\mathbf{D}^\top\mathbf{D}\mathbf{x}) \\ &:= (\mathbf{U} - \mu\mathbf{D}^\top\mathbf{D})^{-1}(\mathbf{y} + \mu\mathbf{D}^\top(\mathbf{z} - \mathbf{d})) \end{aligned} \quad (3.18)$$

ここで、 $\mathbf{U} \in \mathbb{R}^{N_1 \times N_1}$ は単位行列である。 \mathbf{x} の部分問題は逆行列演算が必要であるが、 \mathbf{D} が巡回行列であり、高速フーリエ変換により対角化可能であるため、次式を用いて高速に演算可能である。

$$\mathbf{x}^* = \mathcal{F}^{-1} \left\{ \left(\mathcal{F}(\mathbf{y}) + \mu\overline{\mathcal{F}(\mathbf{D})} \otimes \mathcal{F}(\mathbf{z}) \right) \circ \left(\mathcal{F}(\mathbf{U}) + \mu\overline{\mathcal{F}(\mathbf{D})} \otimes \mathcal{F}(\mathbf{D}) \right) \right\} \quad (3.19)$$

ここで、 $\mathcal{F}(\cdot)$, $\mathcal{F}^{-1}(\cdot)$ はフーリエ変換と逆フーリエ変換、 $\overline{\mathcal{F}(\cdot)}$ は複素共役を表し、演算子 \otimes, \circ は要素毎の積と商を表す。 \mathbf{A} が巡回行列ではない場合、一ステップ目において、共役勾配法 [67] のような反復的な解法が必要で内部ループの出現により、求解に多くの時間を要する。二ステップ目は $\kappa = \frac{\lambda}{\mu}$ として $\text{prox}_{\frac{\lambda}{\mu}\|\cdot\|_1}(\mathbf{A}\mathbf{x} + \mathbf{d})$ を計算すればよく、各ステップは以下のように書き換えられる。

$$\begin{cases} \mathbf{x}^{k+1} &:= \mathcal{F}^{-1} \left\{ \left(\mathcal{F}(\mathbf{y}) + \mu\overline{\mathcal{F}(\mathbf{D})} \otimes \mathcal{F}(\mathbf{z}^k) \right) \circ \left(\mathcal{F}(\mathbf{U}) + \mu\overline{\mathcal{F}(\mathbf{D})} \otimes \mathcal{F}(\mathbf{D}) \right) \right\} \\ \mathbf{z}^{k+1} &:= \text{prox}_{\frac{\lambda}{\mu}\|\cdot\|_1}(\mathbf{D}\mathbf{x}^{k+1} + \mathbf{d}^k) \\ \mathbf{d}^{k+1} &:= \mathbf{d}^k + \mathbf{D}\mathbf{x}^{k+1} - \mathbf{z}^{k+1} \end{cases} \quad (3.20)$$

PDS [66]

PDS は紹介する凸最適化アルゴリズムの中で最も広い範囲の問題を扱える¹。PDS が適用可能な問題は次式で表される。

$$\min_{\mathbf{x} \in \mathbb{R}^{N_1}} F(\mathbf{x}) + G(\mathbf{x}) + H(\mathbf{A}\mathbf{x}) \quad (3.21)$$

¹実際には PDS [66] で解ける問題は ADMM [65] で求解可能である。しかし、ADMM の収束速度は行列 \mathbf{A} に依存する。そのため \mathbf{A} に依存することなく、効率的な計算が行える PDS を最も広い範囲の問題を扱える凸最適化アルゴリズムとした。

ここで $\mathbf{A} \in \mathbb{R}^{N_1 \times N_2}$ は線形作用素である。PDS は以下の二ステップの反復演算により解を求める。

$$\begin{cases} \mathbf{x}^{k+1} := \text{prox}_{\gamma_1 G}(\mathbf{x}^k - \gamma_1 \nabla F(\mathbf{x}^k) - \gamma_1 \mathbf{A}^T \mathbf{z}^k) \\ \mathbf{z}^{k+1} := \text{prox}_{\gamma_2 H^*}(\mathbf{z}^k + \gamma_2 \mathbf{A}(2\mathbf{x}^{k+1} - \mathbf{x}^k)) \end{cases} \quad (3.22)$$

ここで, $\gamma_1, \gamma_2 > 0$ は PDS の収束に関与する任意のパラメータであり $\frac{1}{\gamma_1} - \gamma_2 \|A\|^2 > 0.5\beta$ を満たすように設定し, $\mathbf{z} \in \mathbb{R}^{N_2}$ は補助変数, $\text{prox}_{\gamma H^*}$ は $H(\cdot)$ の共役関数に対する近接写像であり以下のように定義される。

$$\text{prox}_{\gamma H^*}(\mathbf{x}) = \mathbf{x} - \gamma \text{prox}_{H/\gamma}(\mathbf{x}/\gamma) \quad (3.23)$$

媒介変数 $\hat{\mathbf{z}}$ を導入することで以下のように書き直す

$$\begin{cases} \mathbf{x}^{k+1} := \text{prox}_{\gamma_1 G}(\mathbf{x}^k - \gamma_1 \nabla F(\mathbf{x}^k) - \gamma_1 \mathbf{A}^T \mathbf{z}^k) \\ \hat{\mathbf{z}}^{k+1} := \mathbf{z}^k + \gamma_2 \mathbf{A}(2\mathbf{x}^{k+1} - \mathbf{x}^k) \\ \mathbf{z}^{k+1} := \hat{\mathbf{z}}^{k+1} - \gamma_2 \text{prox}_{H/\gamma_2}(\hat{\mathbf{z}}^{k+1}/\gamma_2) \end{cases} \quad (3.24)$$

TV 正則化に加え, $[0, 1]$ の範囲制約付きの問題を PDS で解く際の例を示す。

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{D}\mathbf{x}\|_1 \quad \text{s.t. } \mathbf{x} \in [0, 1]^N \quad (3.25)$$

制約があり, \mathbf{x} が $[0, 1]$ の範囲であることに注意されたい。制約なしの問題にするため, 指示関数を導入し次式を得る。

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{D}\mathbf{x}\|_1 + \iota_{[0,1]}(\mathbf{x}) \quad (3.26)$$

$F(\mathbf{x}) = \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$, $\mathbf{G}(\mathbf{x}) = \iota_{C_{0,1}}(\mathbf{x})$, $H(\cdot) = \lambda \|\cdot\|_1$, $\mathbf{A} = \mathbf{D}$ として PDS で解ける形である。一ステップ目に関して。

$$\nabla F(\mathbf{x}) = \mathbf{x} - \mathbf{y} \quad (3.27)$$

$$\text{prox}_{\gamma_1 G}(\cdot) = P_{C_{0,1}}(\cdot) \quad (3.28)$$

三ステップ目は l_1 ノルムの近接写像であり, $[0, 1]$ の範囲制約付き TV 正則化の問題は PDS を用いて次式で求解できる

$$\begin{cases} \mathbf{x}^{k+1} := P_{C_{0,1}}(\mathbf{x}^k - \gamma_1(\mathbf{x}^k - \mathbf{y} - \gamma_1 \mathbf{D}^T \mathbf{z}^k)) \\ \hat{\mathbf{z}}^{k+1} := \mathbf{z}^k + \gamma_2 \mathbf{D}(2\mathbf{x}^{k+1} - \mathbf{x}^k) \\ \mathbf{z}^{k+1} := \hat{\mathbf{z}}^{k+1} - \gamma_2 \text{prox}_{\frac{\lambda}{\gamma_2} \|\cdot\|_1}(\frac{\hat{\mathbf{z}}^{k+1}}{\gamma_2}) \end{cases} \quad (3.29)$$

PDS の特筆すべきは各ステップが行列演算と近接写像で計算でき, 一回の反復が高速である点である。つまり, ADMM のように逆行列演算を必要とせず, 行列 \mathbf{A} はどのような形であってもよい。ただし, ADMM や MFISTA に比べ収束に多くの反復回数を要することに注意されたい。



(a) 原画像 (b) フィルタ結果

図 3.1: 移動平均フィルタの処理結果

3.2 エッジ保存型平滑化フィルタ

ここではフィルタリング、なかでもエッジ保存型平滑化フィルタとして画像処理でよく用いられ、本稿でも登場するバイラテラルフィルタ [1], ノンローカルミーニングフィルタ [17], ガイデッドフィルタ [46, 68] を紹介する. エッジ保存型平滑化フィルタ紹介の前に画像処理におけるフィルタを簡潔に述べる. フィルタを用いた平滑化などを一般的にフィルタリングというが, この処理は各画素を中心とした画像の局所領域において, 任意のフィルタがもつ係数を対応する画素にかけ合わせ, 足し込む操作を意味する. 入力画像 $\mathbf{x} \in \mathbb{R}^N$ に対するフィルタ結果 $\mathbf{y} \in \mathbb{R}^N$ とすると, 各画素の出力は以下のように算出される.

$$y_i = \sum_{j \in \mathcal{N}(i)} w_{i,j} x_j \quad (3.30)$$

ここで, j は着目画素 i の近傍画素のインデックス, $w_{i,j}$ は画素 i に対する画素 j のフィルタ係数である. 一番単純なフィルタとして移動平均フィルタがあげられる. これは局所領域内でフィルタの係数がすべて同値であり局所領域の平均値を算出するフィルタであり, フィルタ係数は以下のように各周辺画素において一定である.

$$\begin{aligned} w_{i,j} &= \frac{1}{\sum_{j \in \mathcal{N}(i)} 1} \\ &= \frac{1}{|\mathcal{N}|} \end{aligned} \quad (3.31)$$

$|\mathcal{N}| = \sum_{j \in \mathcal{N}(i)} 1$ は各画素のフィルタリングの際に使用する局所領域の総和である.

移動平均フィルタは過度な平滑化となり, 画像のディテールを損なう (図 3.1 参照). 画像はなめらかな輝度変動を持ち, 近傍の画素の値が似た値を持つことを用いたガウシアンフィルタがある. フィルタ係数は標準偏差 σ のガウス関数 $g(x, \sigma) = \exp(-\|x\|^2/\sigma^2)$ を用いて

$$w_{i,j} = \frac{g(i-j, \sigma)}{\sum_{j \in \mathcal{N}(i)} g(i-j, \sigma)} \quad (3.32)$$

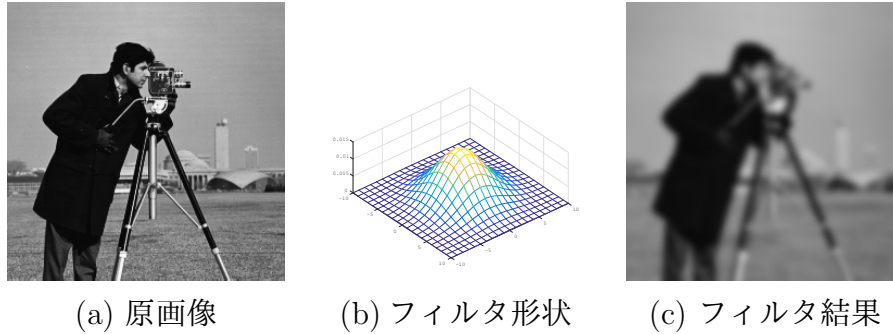


図 3.2: ガウシアンフィルタのフィルタ形状とフィルタ結果

図 3.2 は 21×21 のガウシアンフィルタの例であり，中心から外側にむけて係数が低くなっていることが確認できる．これはフィルタの各係数がフィルタの中心画素からの空間距離に即した正規分布となっており，移動平均フィルタに比べて画像のディテールを崩すことなく平滑化が行える．以降，バイラテラルフィルタ，ノンローカルミーンフィルタ，ガイドドフィルタについて述べる．

3.2.1 バイラテラルフィルタ [1]

バイラテラルフィルタはエッジ保存型平滑化フィルタとして幅広く画像処理に用いられている．バイラテラルフィルタはガウシアンフィルタで用いられた画素距離に加え，対象画素の輝度値の差（以降，輝度距離）をフィルタの係数算出の際に用いる．バイラテラルフィルタは以下のように定義される．

$$w_{i,j} = \frac{g(i-j, \sigma_s)g(p_i - p_j, \sigma_r)}{\sum_{j \in \mathcal{N}(i)} g(i-j, \sigma_s)g_r(p_i - p_j, \sigma_r)} \quad (3.33)$$

ここで， σ_s ， σ_r は画素距離と輝度距離のガウス関数それぞれに対する標準偏差である．図 3.3 は (a) の中心の赤画素を i とした際のバイラテラルフィルタの係数である．上段の場合，図 3.3(b) からわかるように局所領域内で輝度変動がないため，ほぼガウシアンフィルタの形状であるが，下段では着目画素の右下側は着目画素と大きくはなれた画素値を持つことがわかる．このときバイラテラルフィルタはその画素の係数が低くなり，下段右のようにガウシアンフィルタを切断した形状となる．バイラテラルフィルタを一度施すだけでは十分な平滑化が得られないため，図 3.4 のように複数回処理することで十分な平滑化が可能である．

3.2.2 ノンローカルミーンフィルタ

ノンローカルミーンフィルタは輝度距離ではなく，局所領域の類似度を基にフィルタ係数を算出するフィルタである．画素 i および j を中心としてその周辺画素

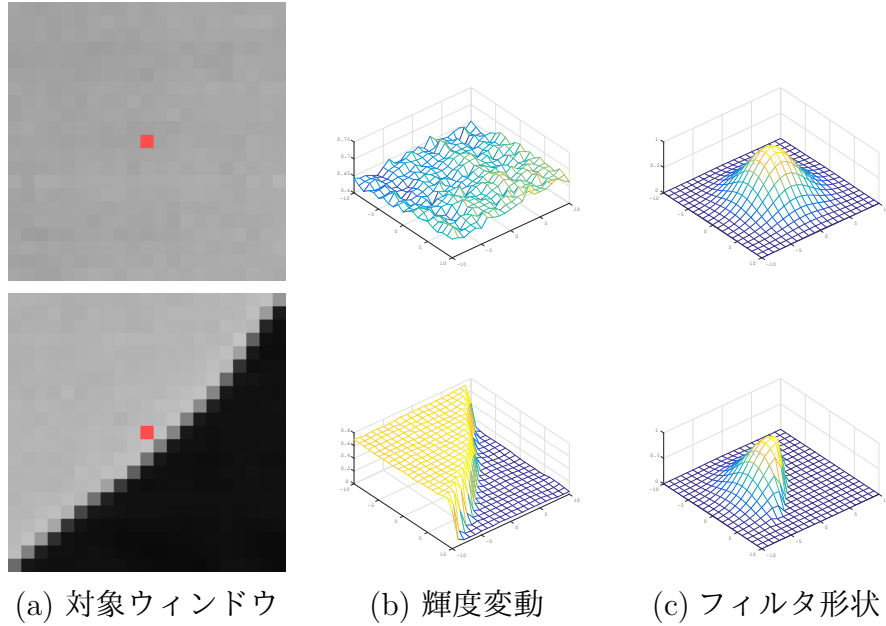


図 3.3: バイラテラルフィルタのフィルタ形状とフィルタ結果

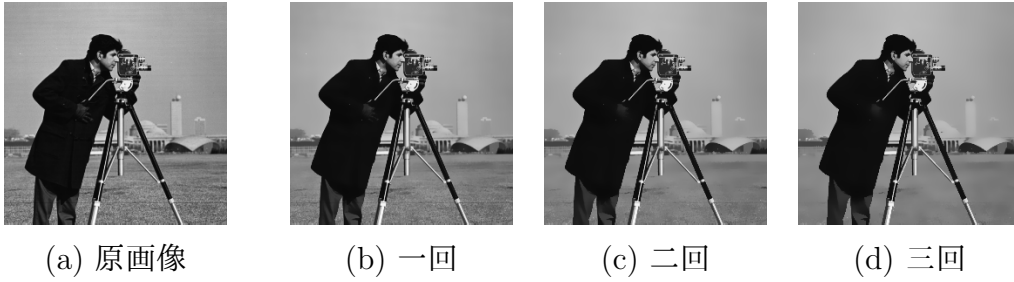


図 3.4: バイラテラルフィルタを複数回処理した結果. 各パラメータは $\sigma_s = 3.2$, $\sigma_r = 0.05$, フィルタサイズは 21×21 である.

$\mathcal{N}_2(i)$ もしくは $\mathcal{N}_2(j)$ からなる局所領域を抜き出す関数を P_i , P_j とするとその類似度は以下の $SSD_{i,j}$ で表せる.

$$SSD_{i,j}(\mathbf{x}) = \|P_i(\mathbf{x}) - P_j(\mathbf{x})\|_2^2 \quad (3.34)$$

$SSD_{i,j}$ は画像認識において類似度の尺度として用いられる各局所領域の二乗誤差である. このとき, $SSD_{i,j}$ が大きいほど $P_i(\mathbf{x})$ と $P_j(\mathbf{x})$ の類似度は低く, 小さければ類似度が高い. 局所領域の類似度 $SSD_{i,j}$ を用いてノンローカルミーンフィルタの各係数は以下で表される.

$$w_{i,j} = \frac{g(SSD_{i,j}(\mathbf{x}), \sigma)}{\sum_{j \in \mathcal{N}_1(i)} g(SSD_{i,j}(\mathbf{x}), \sigma)} \quad (3.35)$$

バイラテラルフィルタは画素 i, j の輝度距離を測っているため, フィルタ係数が雑音により大きく変動するために雑音除去としての性能は低い. 一方で, 局所領域



図 3.5: ノンローカルミーンフィルタによる雑音除去. 各パラメータは $\sigma = 0.1$, フィルタサイズは $|\mathcal{N}_1| = 21 \times 21$, 類似度を測る際の局所領域サイズは $|\mathcal{N}_2| = 3 \times 3$ である.

の類似度でフィルタ係数を測るノンローカルミーンフィルタは各画素のフィルタ係数決定の際に雑音の影響を受けにくく, 雑音除去性能が高い. 図 3.5 に示すように細部の情報を残しつつ雑音を除去できる.

3.2.3 ガイデッドフィルタ

バイラテラルフィルタやノンローカルミーンフィルタに対し, 局所領域毎の変換によりフィルタリングを達成するのがガイデッドフィルタである. ガイデッドフィルタは局所領域ごとに以下の変換式で構成される

$$y_j = a_i x_j + b_i, \forall i \in \mathcal{N}(i) \quad (3.36)$$

ここで, a_i, b_i はスケーリングとオフセットであり, 線形式により y_j を表現する. a_i, b_i は以下の誤差を最小化する値である.

$$E(a_i, b_i) = \sum_{j \in \mathcal{N}(i)} ((a_i x_j + b_i - y_j)^2 + \varepsilon a_i^2) \quad (3.37)$$

ここで, ε はチコノフ正則化のバランスパラメータである. a_i, b_i の部分問題はそれぞれ二次式となるため, 微分し 0 を取る点を求める. まず, b_i は以下のように求める.

$$\begin{aligned} \frac{\partial}{\partial b_i} \sum_{j \in \mathcal{N}(i)} ((a_i x_j + b_i - y_j)^2 + \varepsilon a_i^2) &= 0 \\ \sum_{j \in \mathcal{N}(i)} (a_i x_j + b_i - y_j) &= 0 \\ \sum_j (1) b_i &= \sum_j y_j a_i \sum_j x_i \\ b_i &= \frac{\sum_j y_j}{\sum_j 1} - a_i \frac{\sum_j x_j}{\sum_j 1} \\ &= \bar{y}_i - a_i \bar{x}_i \end{aligned} \quad (3.38)$$

ここで、 \bar{x}_i および \bar{y}_i は局所領域 $\mathcal{N}(i)$ における各平均値である。求めた b_i を式 (3.37) に代入後、 a_i に関しても同様に微分し 0 を取る点を求める。

$$\begin{aligned} \frac{\partial}{\partial a_i} \sum_{i \in \mathcal{N}(k)} ((a_i(x_j - \bar{x}_i) - (y_j - \bar{y}_i))^2 + \varepsilon a_k^2) &= 0 \\ \sum_j (x_j - \bar{x}_i)^2 a_i - \sum_j (x_j - \bar{x}_i)(y_j - \bar{y}_i) + \varepsilon a_i &= 0 \\ \left(\sum_j (x_j - \bar{x}_i)^2 + \varepsilon \right) a_i &= \sum_j (x_j - \bar{x}_i)(y_j - \bar{y}_i) \\ a_i &= \frac{\sum_j (x_j - \bar{x}_i)(y_j - \bar{y}_i)}{\left(\sum_j (x_j - \bar{x}_i)^2 + \varepsilon \right)} \end{aligned} \quad (3.39)$$

ここで分母第一項は $\mathcal{N}(i)$ の分散を意味するので $\sum_j (x_j - \bar{x}_i)^2 = \sigma_i$ 。式を整理して次式が得られる。

$$\begin{aligned} a_i &= \frac{\frac{1}{|\mathcal{N}|} \sum_{j \in \mathcal{N}(i)} x_j y_j - \bar{x}_i \bar{y}_i}{\sigma_i^2 + \varepsilon} \\ b_i &= \bar{y}_i - a_i \bar{x}_i \end{aligned} \quad (3.40)$$

ただし、これは $\mathcal{N}(i)$ に関してのみの係数であり、周辺の各局所領域は互いに重なり合うため、画素 i に対するフィルタ結果は次式のようにオーバーラップする局所領域の変換係数 a_i , b_i の平均を用いる

$$\begin{aligned} y_i &= \frac{1}{|\mathcal{N}|} \sum_{i|j \in \mathcal{N}(i)} (a_j x_i + b_j) \\ &= \frac{x_i}{|\mathcal{N}|} \sum_{i|j \in \mathcal{N}(i)} a_j + \frac{1}{|\mathcal{N}|} \sum_{i|j \in \mathcal{N}(i)} b_j \\ &= \bar{a}_i x_i + \bar{b}_i \end{aligned} \quad (3.41)$$

実装の際には \bar{a}_i , \bar{b}_i の算出でフィルタリングが達成されるが、各フィルタ係数は以下で表される。

$$w_{i,j} = \frac{1}{|\mathcal{N}|} \sum_{k:(i,j) \in \mathcal{N}(k)} \left(1 + \frac{(x_i - \mu_k)(x_j - \mu_k)}{\sigma_k^2 + \varepsilon} \right) \quad (3.42)$$

ガイドドフィルタはフィルタ係数を画素毎の積を足しこむ操作で算出できるため、積分フィルタを用いることでフィルタサイズに依存しない高速な処理が可能である。また、チコノフ正則化のパラメータを小さく設定した場合には式 (3.37) の第一項を忠実に満たすため、入力画像 \mathbf{x} のディテールを保ち、大きく設定した場合には強い平滑化フィルタとなる。図 3.6 では ε を変化させたガイドドフィルタの処理結果であり、パラメータ設定による平滑化度合いが確認できる。また、変換係数 a, b を用いた局所領域ごとの線形式によるフィルタリングは次節のジョイントフィルタリングの際に非常に有効な定式化である。

3.2.4 ジョイントフィルタリング

上記のエッジ保存型平滑化フィルタは全て画像の画素値によりフィルタ係数が決定される。そのため、位置ずれのない参照画像がある場合にはこれらのフィル

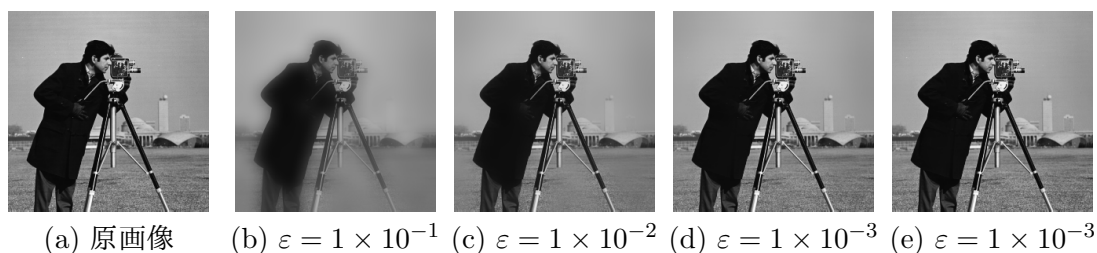


図 3.6: 異なるパラメータ設定によるガイドドフィルタの結果

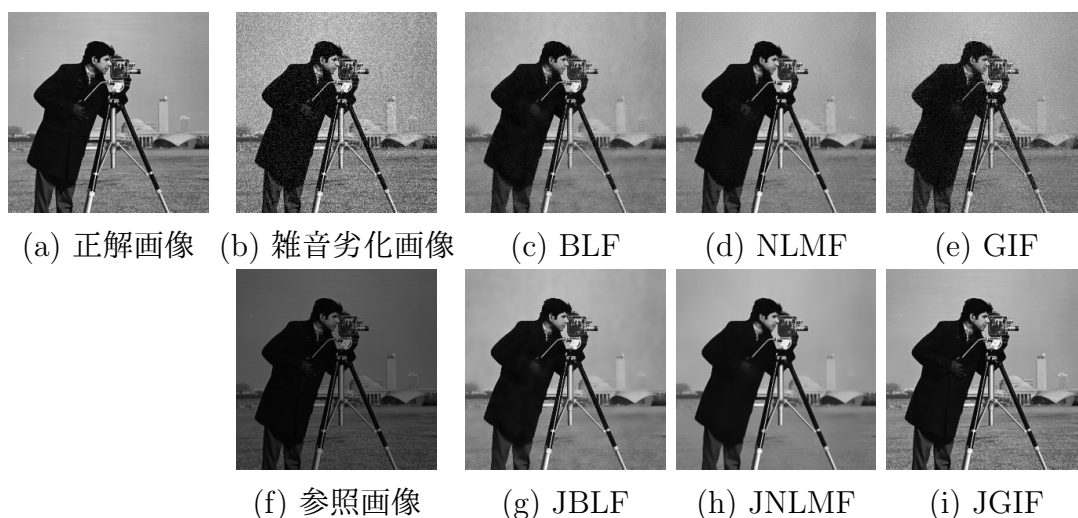


図 3.7: 各ジョイントフィルタリングによる雑音除去. BLF はバイラテラルフィルタ, NLMF はノンローカルミーンフィルタ, GIF はガイドドフィルタであり, JBLF, JNLMF, および JGIF は各ジョイントフィルタを意味する. 各フィルタリングのパラメータは結果が良好になるように適宜設定している.

タ係数を参照画像より算出し, 原画像にフィルタリングするジョイントフィルタリング (クロスフィルタリングと呼ばれることもある) が有効な手段である. ここでは紹介したエッジ保存型平滑化フィルタのジョイントフィルタリングによる雑音除去の各結果を示す. 図 3.7 上段は各フィルタ係数を雑音画像 (b) から算出し, フィルタ処理した結果であり, 下段はコントラストは低い雑音がない画像 (f) を参照画像として, フィルタ係数の算出に用いてフィルタ処理した結果である. 各フィルタ処理結果から, 下段のジョイントフィルタリングは上段の雑音画像のみをつかうフィルタ処理に比べ, 雑音除去性能が高いことが確認できる. ここでは原画像のコントラストを下げたものを参照画像として用いたが, 実際のシーンで撮影された画像に対しては雑音のないフラッシュ画像を参照画像とすることで高い雑音除去効果が確認されている [5].

第4章 凸最適化を用いたフラッシュ ／高感度画像合成による画像 復元

4.1 概要

フラッシュ画像と高感度画像の合成手法の問題点の一つに、フラッシュ画像を絶対的な基準画像とみなす点があげられる。すなわち、実際に撮影したペア画像では、高感度画像よりもフラッシュ画像が不鮮明となる場合があり、結果画像に影響が生じる。例えば、被写体と背景の建造物の間に空間的に広がりがある場合、フラッシュ光の届く近距離の被写体は光を強く反射するが、遠方の背景は暗いままとなる。なぜなら、露光量は前者を基準として調整され、相対的に後者はより暗く写り不鮮明となる。

本提案法では、従来法 [5, 45] と同様にテクスチャ成分とストラクチャ成分の合成を用いて、雑音のない鮮明な高感度画像の生成を目的とする。フラッシュ画像においてフラッシュ光の届いた領域と届かない領域（以降、高露光領域と低露光領域と呼ぶ）を判別し、各領域の性質にあわせて適切な合成法を施す。各合成法は従来法で採用されていた単純な BLF を用いたジョイントフィルタリング (JBLF) とは異なる。従来法では JBLF により分離されたテクスチャ成分とストラクチャ成分は、対応する画素の和として合成される。提案法では各成分の合成を（微分できない）凸最適化問題として定式化している。このようなアプローチをとる利点として、i) 事前にわかっている合成画像が満たすべき諸性質（滑らかさ、入力画像との整合性等）を柔軟に活用できる、ii) 各合成ステップにおいて最適性が保証できる、といったことが挙げられる。前者に関しては、実際に、強力な平滑化能力を有する TV 正則化や BLF の行列表現を組み込んだ新しい正則化項を利用することで、望ましい画像合成（二重輪郭の低減等）を実現している。また、後者によって提案法全体における各合成ステップの役割が明確となり、提案法の性能検証や他の応用に際した変更等を見通しよく行うことができると考えられる。なお、提案法の各ステップで解かなければならない凸最適化問題は、その非可微分性のため反復解法を要する比較的難しい問題となっているが、4.3 節で説明するように、式変形の後に適切な凸最適化アルゴリズムを適用することで、効率的に最適解を求めることができる。

以降、4.2節で関連手法、4.3節で提案法について述べる。最後に、4.4節における定性評価と定量評価で提案法の有効性を示す。

4.2 関連手法

ここでは、直接の目的としてはフラッシュ画像の合成を扱わないが、同様の処理が可能な手法について述べる。手法は大きく分けて、雑音除去法、画像合成法に分類される。

雑音除去法としては、BM3D [7] などの高い雑音除去能力をもつ non-local mean (NLM) 系の手法を単に高感度画像に対して施す、または、フラッシュ画像を補助画像としてジョイントフィルタリングする方法があげられる [3]。これら NLM 系の手法は良好な高感度画像を生成するが、基本的には平滑化法であるので、もともと存在したテクスチャが薄れる問題があり、原画像以上に鮮明なテクスチャが復元されることはない。また、BM3D のように周波数空間を介する処理を行う場合は、幾何学的な模様が生じる問題がある。

画像合成法としては、多重解像度での合成法である exposure fusion [69] が高露光画像と低露光画像の合成を扱っており、フラッシュ画像と高感度画像の合成に近い処理となる。しかし、もともとなる image fusion は原画像に雑音が含まれないことを前提としており、合成画像を鮮明とするために合成用の画像の高周波数成分を優先的に用いるため、雑音が強調される問題がある。他の手法として、補助画像を用いて反射成分分離を行う [70] においてフラッシュ画像と高感度画像が用いられる。勾配空間処理と呼ばれ、画像を一度微分した後に、二つの画像の勾配量を組み替え、再度画像を復元することで、広い範囲にわたって生じる写り込みや陰影などのアーティファクトを除去できる。ただし、この方法も画像中に雑音が含まれないことを前提としており、雑音を含む際は復元過程において雑音が強調される問題がある。

以上より、フラッシュ画像と高感度画像の合成を目的とする場合は、前章にて述べた二種類の手法が有効であると考えられる。

4.3 提案手法

提案法では、高露光領域と低露光領域に対して、異なる合成法を施す。図 4.1 に提案法のフローチャートを示す。まず、高露光領域と低露光領域を分ける処理が必要となる。なお後続の処理の簡略化のため、最初に各領域を区別せず、全画像領域に対して各領域用の処理（図中左と中央のブロック）を施した後、各領域を判別し（図中右のブロック）、最終的にマスク処理により各領域を合成する。

まず、フラッシュ画像を p 、高感度画像を q と表す。フラッシュ画像は受光感度を低めに設定してフラッシュ光を用いて撮影する。一方、高感度画像は受光感度

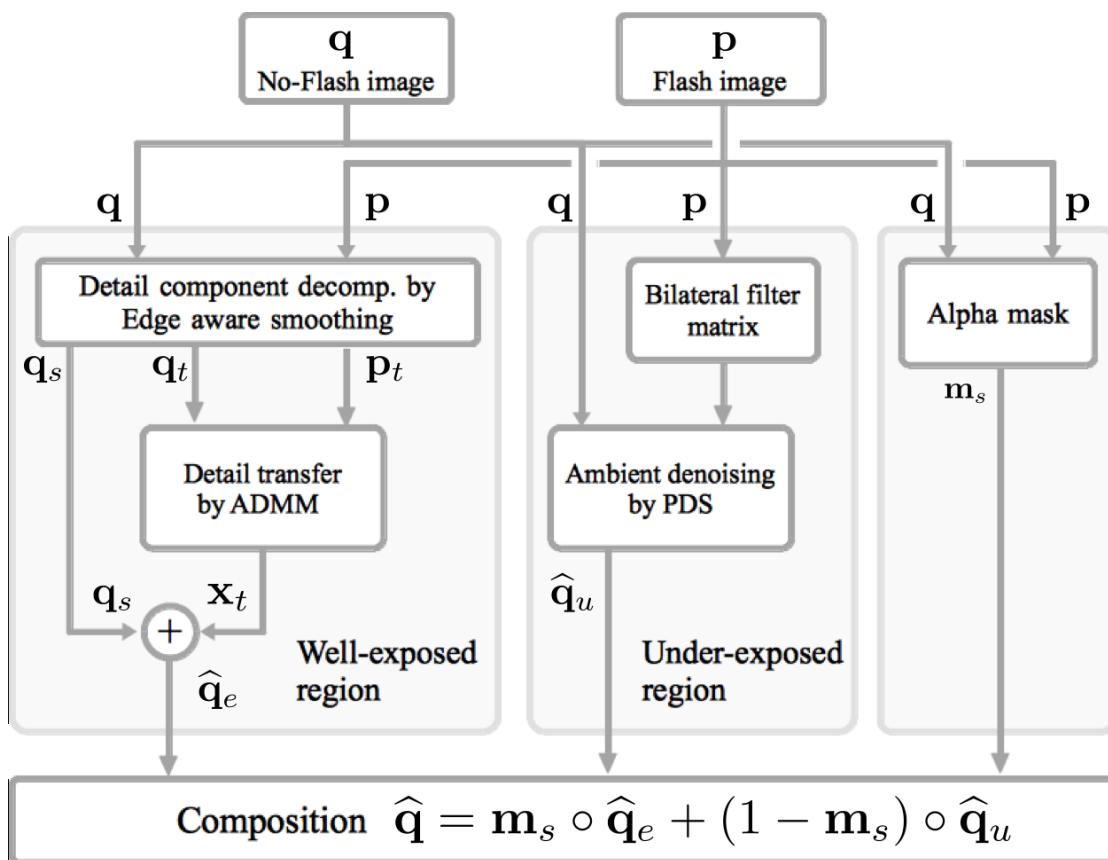


図 4.1: 提案手法のフローチャート

を高く設定して撮影する。フラッシュ光の届かない領域を撮影すると、フラッシュ画像は、高感度画像の撮影時よりも受光感が低いため、暗くなり不鮮明となる。

各処理の概要は次のようになる。

- 高露光領域に対する処理（図 4.1 左ブロック）では、各画像をテクスチャ成分 \mathbf{p}_t と \mathbf{q}_t とストラクチャ成分 \mathbf{p}_s と \mathbf{q}_s に分けた後、テクスチャ成分から新たなテクスチャ \mathbf{x}_t を生成し、新たな画像 $\hat{\mathbf{q}}_e = \mathbf{q}_s + \mathbf{x}_t$ を生成する（4.3.1 節参照）。
- 低露光領域に対する処理（図 4.1 中央ブロック）では、フラッシュ画像 \mathbf{p} から JBLF 用のフィルタ行列を生成した後、高感度画像 \mathbf{q} に対して平滑化を施し、新たな画像 $\hat{\mathbf{q}}_u$ を生成する（4.3.2 節参照）。
- 高露光領域と低露光領域を判別する処理（図 4.1 右ブロック）では、高露光領域を表すマスク \mathbf{m}_s を生成する（4.3.3 節参照）。
- 合成処理では（図 4.1 下段ブロック）では、高露光領域のマスク \mathbf{m}_s を用いて画像 $\hat{\mathbf{q}}_e$ と $\hat{\mathbf{q}}_u$ を合成し、雑音のない高感度画像 $\hat{\mathbf{q}}$ を生成する（4.3.3 節参照）。

各処理の詳細については以降の節にて述べる。

4.3.1 高露光領域用処理

高露光領域においては、フラッシュ画像の模様（テクスチャ成分）は鮮明に撮影されるが、強いフラッシュ光により色彩（ストラクチャ成分）が変化した可能性がある。一方、高感度画像は雑音を含むものの理想的な色彩をもつ。そのため、高感度画像を平滑化し、復元画像の基準色をもつストラクチャを生成した後、差分情報であるテクスチャ成分をフラッシュ画像および高感度画像から生成する。

テクスチャ成分とストラクチャ成分の分離

フラッシュ画像を $\mathbf{p} \in \mathbb{R}^N$ 、高感度画像を $\mathbf{q} \in \mathbb{R}^N$ として画素数 N の次元をもつベクトルとして表す。次に、各画像をテクスチャ成分とストラクチャ成分に分離するため、まずエッジ保存平滑化を行い、ストラクチャ成分 \mathbf{p}_s と \mathbf{q}_s を得る。

$$\mathbf{q}_s = g(\mathbf{q})^k, \quad \mathbf{p}_s = g(\mathbf{p})^k \quad (4.1)$$

ここで $g(\cdot)^k$ はエッジ保存平滑化を k 回施すことを表し、本手法では BLF と同等の性能をもち、高速に動作する domain transform filter (DTF) [71] を用いる。DTF

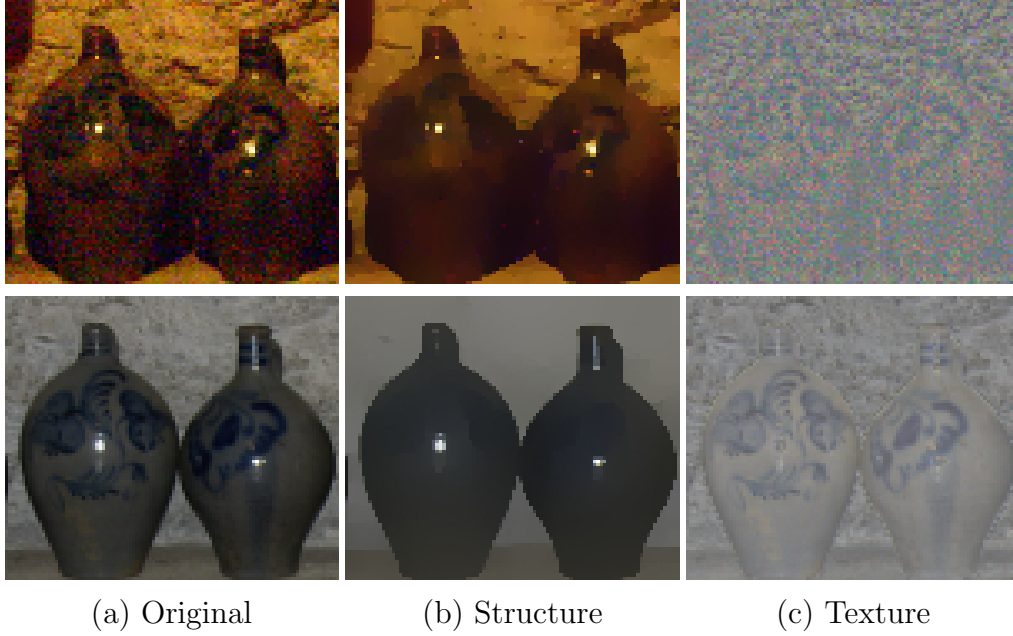


図 4.2: 高感度画像（上段）とフラッシュ画像（下段）の各成分

の各パラメータ σ_s と σ_r , および k の決り方方法は 4.4 節で詳細を述べる. テクスチャ成分を原画像とストラクチャの差分として定義する.

$$\mathbf{p}_t = \mathbf{p} - \mathbf{p}_s, \quad \mathbf{q}_t = \mathbf{q} - \mathbf{q}_s \quad (4.2)$$

図 4.2 に得られた各成分の例を示す.

高露光領域用の合成モデル

得られた各成分画像を用いて合成を行う. ここでは高露光領域用の合成モデルを次式で表す.

$$\hat{\mathbf{q}}_e = \mathbf{q}_s + \arg \min_{\mathbf{x}_t} \left\{ \frac{1}{2} \|\mathbf{x}_t - \mathbf{q}_t\|_2^2 + \gamma \|\mathbf{D}\mathbf{x}_t - \mathbf{D}\mathbf{p}_t\|_1 \right\} \quad (4.3)$$

\mathbf{q}_s は復元画像の基準色となる高感度画像のストラクチャ成分を表し, 最適化で得られたテクスチャ成分 \mathbf{x}_t が足し込まれる形となる. $\{\cdot\}$ 内部の評価関数 3.1.1 節で紹介した凸関数からなり, ここで用いられる $\mathbf{D} = [\mathbf{D}_h; \mathbf{D}_v] \in \mathbb{R}^{2N \times N}$ はそれぞれ水平方向と垂直方向に一次微分を行うフィルタ行列を表す. 第一項目は高感度画像の色彩を保つ役割を果たす. 第二項目は Total Variation (TV) による正則化を意図しており, フラッシュ画像の勾配に近づける役割を果たす.

最適な $\hat{\mathbf{x}}_t$ を得ることを考える. 微分可能な l_2 ノルムと近接写像可能な TV ノルムで構成されるため, ここでは凸最適化手法である ADMM [65] を用いて方程式を解く. 上式の $\{\cdot\}$ 内の l_1 ノルムを媒介変数 $\mathbf{z} := [\mathbf{z}_h; \mathbf{z}_v] \in \mathbb{R}^{2N}$ と, 双対変数 $\mathbf{d} := [\mathbf{d}_h; \mathbf{d}_v] \in \mathbb{R}^{2N}$ を用いて以下の拡張ラグランジュの形で書き換え,

$$\frac{1}{2} \|\mathbf{x}_t - \mathbf{q}_t\|_2^2 + \gamma \|\mathbf{z}\|_1 + \frac{\rho}{2} \|\mathbf{D}(\mathbf{x}_t - \mathbf{p}_t) - \mathbf{z} + \mathbf{d}\|_2^2 \quad (4.4)$$

以下に示す繰り返しアルゴリズムにしたがって解の更新を行う。

$$\begin{cases} \mathbf{x}_t^{k+1} := \arg \min_{\mathbf{x}_t} \frac{1}{2} \|\mathbf{x}_t - \mathbf{q}_t\|_2^2 + \frac{\rho}{2} \|\mathbf{D}(\mathbf{x}_t - \mathbf{p}) - \mathbf{z}^k + \mathbf{d}^k\|_2^2 \\ \mathbf{z}^{k+1} := \arg \min_{\mathbf{z}} \gamma \|\mathbf{z}\|_1 + \frac{\rho}{2} \|\mathbf{D}(\mathbf{x}_t^{k+1} - \mathbf{p}_t) - \mathbf{z} + \mathbf{d}^k\|_2^2 \\ \mathbf{d}^{k+1} := \mathbf{d}^k + (\mathbf{D}(\mathbf{x}_t^{k+1} - \mathbf{p}_t) - \mathbf{z}^{k+1}) \end{cases} \quad (4.5)$$

一ステップ目は式 (4.4) の \mathbf{x}_t に関する項を最小二乗法で解いており、第一項目と第三項目を \mathbf{x}_t で微分し、ゼロを与える \mathbf{x}_t を求める。式を整理すると次式を得る。

$$\mathbf{x}_t = (\mathbf{I} + \rho \mathbf{D}^T \mathbf{D})^{-1} (\mathbf{q}_t + \rho \mathbf{D}^T \mathbf{D} \mathbf{p}_t + \rho \mathbf{D}^T (\mathbf{z}^k - \mathbf{u}^k)) \quad (4.6)$$

ここで $\mathbf{I} \in \mathbb{R}^{N \times N}$ は単位行列を表す。 \mathbf{D} を巡回フィルタ行列 (BCCB 行列) とした場合、 \mathbf{D}^T と $\mathbf{D}^T \mathbf{D}$ も BCCB 行列となるため、高速フーリエ変換により対角化でき、効率よく解ける。

$$\mathbf{x}_t = \mathcal{F}^{-1} \left\{ \left(\mathcal{F}(\mathbf{q}_t) + \rho (\overline{\mathcal{F}(\mathbf{D}_h)} \otimes \mathcal{F}(\mathbf{z}_h) + \overline{\mathcal{F}(\mathbf{D}_v)} \otimes \mathcal{F}(\mathbf{z}_v)) \right) \oslash (\mathcal{F}(\mathbf{U}) + \rho \mathcal{F}(\mathbf{L})) \right\} \quad (4.7)$$

ここで $\mathbf{z}_h = \mathbf{z}_h^k + \mathbf{D}_h \mathbf{p}_t - \mathbf{d}_h^k$ と $\mathbf{z}_v = \mathbf{z}_v^k + \mathbf{D}_v \mathbf{p}_t - \mathbf{d}_v^k$ とする。 $\mathcal{F}(\cdot)$ はフーリエ変換を表し、 $\mathcal{F}^{-1}(\cdot)$ は逆フーリエ変換を表す。 $\overline{\mathcal{F}(\cdot)}$ は複素共役を表し、演算子 \otimes と \oslash はそれぞれ要素毎の積と商を表す。分母の $\mathcal{F}(\mathbf{U})$ は全画素において 1 となり、 $\mathcal{F}(\mathbf{L}) = \mathcal{F}(\mathbf{D}_h^T \mathbf{D}_h) + \mathcal{F}(\mathbf{D}_v^T \mathbf{D}_v)$ はラプラシアンフィルタを表す。 $\mathcal{F}(\mathbf{L})$ は 0 以上の正の実数となるため、分母は常に ρ 以上となり、解が保証される。

式 (4.5) の二ステップ目は式 (4.4) の \mathbf{z} に関する l_1 の近接写像のため、式 (3.2) を用いて \mathbf{z} の解は次のように求まる。

$$\mathbf{z}^{k+1} = \text{prox}_{\frac{\gamma}{\rho} \|\cdot\|_1} (\mathbf{D} \mathbf{x}_t^{k+1} - \mathbf{p}_t + \mathbf{d}^k) \quad (4.8)$$

最後に、求まった主変数 \mathbf{x}_t^{k+1} と \mathbf{z}^{k+1} を用いて、双対変数 \mathbf{d}^{k+1} を更新する。式 (4.3) 中の凸最適化問題および式 (4.4) の式変形は ADMM の収束に必要なある条件を満たすため、式 (4.5) で更新される \mathbf{x}_t^k は当該最適化問題の解を任意の精度で近似可能である。詳細は [65, Theorem 8] を参照されたい。

4.3.2 低露光領域用処理

低露光領域においては、十分な露光量が確保できずフラッシュ画像のコントラストは著しく低下する。一方、高感度画像は鮮明なエッジをもつが、雑音を含む。ただし、フラッシュ画像は不鮮明ではあるが、雑音の影響を受けずに輝度変化を保持するため、フラッシュ画像を用いた JBLF により高感度画像の平滑化を行う。

バイラテラルフィルタ行列

合成処理に先立って、まず BLF の重みを用いた JBLF を考える。フラッシュ画像 \mathbf{p} が与えられた際、画素 i とその近隣画素 $j \in \mathcal{N}(i)$ の画素値 p_j の重み $w_{i,j}$ は式 (3.33) を用いて次式で表される。

$$w_{i,j} = \frac{g(i-j, \sigma_s)g(p_i - p_j, \sigma_r)}{\sum_{j \in \mathcal{N}(i)} g(i-j, \sigma_s)g(p_i - p_j, \sigma_r)} \quad (4.9)$$

得られた重み $w_{i,j}$ を i 行 j 列目の要素値としてもつ行列 $\mathbf{B} \in \mathbb{R}^{N \times N}$ をバイラテラルフィルタ行列 (BLF 行列) と呼ぶこととする。 $\mathbf{B}\mathbf{p}$ の演算が通常の BLF を表し、 $\mathbf{B}\mathbf{q}$ の演算が JBLF を表す。また、テクスチャの分離を行う $(\mathbf{p} - \mathbf{B}\mathbf{p}) = (\mathbf{U} - \mathbf{B})\mathbf{p}$ (ここで $\mathbf{U} \in \mathbb{R}^{N \times N}$ は単位行列) の演算の $(\mathbf{U} - \mathbf{B})$ の部分を高域通過型の BLF 行列と呼ぶ。

低露光領域用の合成モデル

新たに定義した高域通過型の BLF 行列を用いて、高感度画像に対する JBLF を以下のようにモデル化する。

$$\hat{\mathbf{q}}_a = \arg \min_{\mathbf{x}_a} \frac{1}{2} \|\mathbf{x}_a - \mathbf{q}\|_2^2 + \lambda \|(\mathbf{U} - \mathbf{B})\mathbf{x}_a\|_1 \quad (4.10)$$

第一項目は高感度画像の色彩を保つ役割を果たす。第二項目は高域通過型の BLF を用いた正則化項となり、高感度画像のコントラストを維持したまま JBLF を行う意図がある。更に、グラデーション部分の損失を抑制する効果がある。

上述の式 (4.10) は式 (4.3) と同様に、微分可能な項と近接写像可能な項で構成される凸最適化問題であり、ADMM により最適解を得ることができる。ただし、 $\mathbf{B}^T\mathbf{B}$ は巡回行列ではなく、対角化が困難であるため、高露光領域用の処理で用いたフーリエ変換を用いた高速化を適用できない。そこで、ここでは逆行列演算を必要としない PDS [66] を用いて計算を行う。式 (4.10) に対して PDS を適用する場合は、式 (3.21) に対して各関数と変数を次のように設定すればよい。

$$\begin{aligned} F(\mathbf{x}) &= \frac{1}{2} \|\mathbf{x} - \mathbf{q}\|_2^2, & G(\mathbf{x}) &= 0, \\ H(\mathbf{A}\mathbf{x}) &= \lambda \|(\mathbf{U} - \mathbf{B})\mathbf{x}\|_1 \end{aligned} \quad (4.11)$$

PDS の各ステップに各関数を設定すると、以下の更新式が得られる。最終的に以下の繰り返しアルゴリズムに帰着する。

$$\begin{cases} \mathbf{x}_a^{k+1} := \mathbf{x}_a^k - \tau_1(\mathbf{x}_a^k - \mathbf{q}) - \tau_1(\mathbf{U} - \mathbf{B})^T \mathbf{z}^k \\ \tilde{\mathbf{z}}^{k+1} := \mathbf{z}^k + \tau_2(\mathbf{I} - \mathbf{B})(2\mathbf{x}_a^{k+1} - \mathbf{x}_a^k) \\ \mathbf{z}^{k+1} := \tilde{\mathbf{z}}^{k+1} - \tau_2 \text{prox}_{\frac{\lambda}{\tau_2} \|\cdot\|_1} \left(\frac{\tilde{\mathbf{z}}^{k+1}}{\tau_2} \right) \end{cases} \quad (4.12)$$

単純な JBLF [5] と提案した式 (4.10) の平滑化能力の差異を示すため、フィルタ設計に用いるパラメータ σ_s と σ_r を変化させて生成した結果を図 4.3 に示す。JBLF については反復的に 5 回フィルタ処理を行った。ここではフラッシュ画像と高感度画像に対して同一の画像を用いた。単純な JBLF の結果に比べ、提案法の平滑化はコントラストを保持していることがわかる。

4.3.3 高露光領域と低露光領域の合成

ここまでに、式 (4.3) から高露光領域での合成結果 $\hat{\mathbf{q}}_e$ を、式 (4.10) から低露光領域における合成結果復 $\hat{\mathbf{q}}_u$ を得た。これらを合成して最終的な合成結果とする。しかし、高露光領域と低露光領域を厳密に判別することは困難である。本論文では、単純な方法ではあるが、ペア画像を撮影した際の ISO 感度の違いから生じた輝度値の増幅を補正した後、高感度画像とフラッシュ画像の輝度値を比較し、明るい領域を高露光領域、暗い領域を低露光領域と定める。

$$m_i = \begin{cases} 1 & \text{if } p_i \geq q_i \\ 0 & \text{otherwise} \end{cases} \quad (4.13)$$

ここで画像 $\mathbf{m} \in \mathbb{R}^N$ をアルファマップとして合成に用いると、領域の境界で急激に値が $\{0, 1\}$ と変化するため、二重輪郭などが生じる。そこでこれらをガウシアンフィルタを用いて平滑化した後、合成に用いる。

$$\mathbf{m}_s = \mathbf{Gm} \quad (4.14)$$

ここで $\mathbf{G} \in \mathbb{R}^{N \times N}$ はガウシアンフィルタ行列とし、標準偏差を $\sigma = 2$ と定めた。図 4.4 にフラッシュ画像と高感度画像の比較により生成されたアルファマップの一例を示す。

最終的に画像 $\hat{\mathbf{q}}_e$ と $\hat{\mathbf{q}}_u$ の合成をアルファマップ \mathbf{m}_s を用いて以下のように行う。

$$\hat{\mathbf{q}} = \mathbf{m}_s \otimes \hat{\mathbf{q}}_e + (1 - \mathbf{m}_s) \otimes \hat{\mathbf{q}}_u \quad (4.15)$$

ここで \otimes は要素毎の積を表す。

4.4 実験および考察

本章では従来法 [5] [6] と比較し提案法の有効性を示す。提案法では従来法で考慮されていない低露光領域に対しても式 (4.10) で有効に処理できる。しかし式 (4.10) 内のフィルタ係数行列 \mathbf{B} の輝度距離で決定される重みはフラッシュ画像の輝度変化で構成されるため、フラッシュ画像が輝度変化を持たない場合、輝度に対する重みが一定となり BLF として機能せず、フィルタサイズの大きなガウシアンフィ

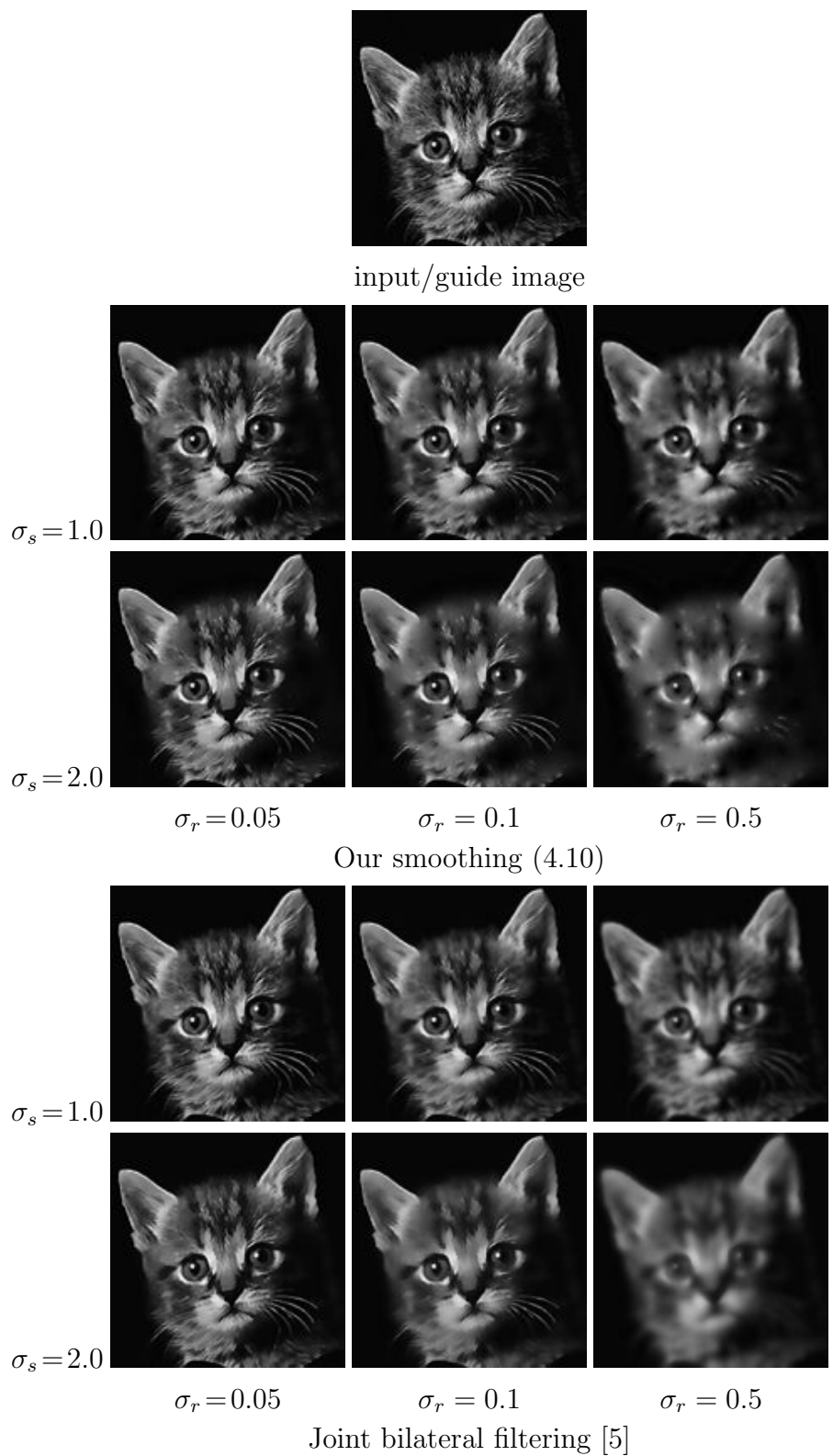


図 4.3: フィルタ設計のパラメータを変動した提案法の平滑化 (4.10) と joint-BLF の結果画像. 式 (4.10) のパラメータを $\lambda = 0.5$ とし, joint-BLF の反復回数は 5 とした)

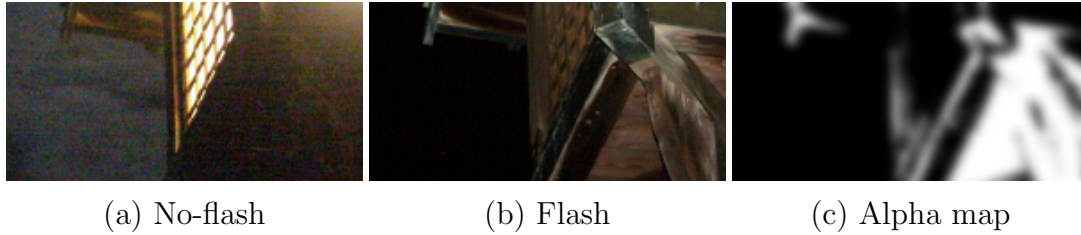


図 4.4: フラッシュマップの例

ルタとなる．結局，フラッシュ画像の低露光領域が輝度変化をもたない場合，式 (4.10) ではコントラストを保持できない．以上のことからフラッシュ画像の低露光領域には輝度変化が生じているという撮影状況に限定して実験を行う．

4.4.1 実画像を用いた定性評価

まず，実験 1 として実画像を用いた画像合成を行い，各手法の復元画像を定性評価する．実験に用いる画像は Canon EOS 20D を用いて撮影し，F 値を 4.5 とし，ISO 感度は高感度画像では 1600，フラッシュ画像では 100 と設定した．輝度値の範囲を $[0,1]$ で正規化して用いる．また，本手法ではカメラ内でガンマ補正された画像を実験に用いる¹．図 4.5 に実験に用いた高感度画像とフラッシュ画像の全景を示す．図 4.6 は従来法 [5] [6] と提案手法の復元結果となる．高露光領域の結果を示したものが図 4.6(a) と (b) である．更に低露光領域の結果を示したものが図 4.6(c) である．高露光領域処理用のパラメータは，成分分離に用いる DTF [71] の標準偏差を $\sigma_s = 10$ ， $0.05 \leq \sigma_r \leq 0.075$ で $k = 3$ 回施し，式 (4.4) 内の γ と ρ を 0.1 と設定した．低露光領域処理用のパラメータは，BLF 行列の標準偏差を $\sigma_s = 2$ ， $\sigma_r = 0.05$ とし，式 (4.10) の正則化パラメータを $0.02 \leq \lambda \leq 0.065$ と設定した．従来法のパラメータは雑音除去の効果が等しくなるように調整を行った．手法 [5] は BLF を用いたためいくつかの色の階調領域に対して，偽色やぼけ等が発生している．更に領域の境界においてテクスチャの有無により二重輪郭による不自然さが生じている．図 4.6(a-3) では他の手法に比べ，輪郭が直線的に復元されていないことがわかる．手法 [6] ではテクスチャの鮮明さがフラッシュ画像に依存するため，フラッシュ画像が劣化した低露光領域ではコントラストの低下が生じている．一方，提案手法は偽色やぼけの発生を抑制しつつ，コントラストを保持した合成を実現している．手法 [5] [6] の結果画像のテクスチャと提案法の結果画像のテクスチャを比較すると，提案法のテクスチャ復元性能が最も高いことが確認できる．特に図 4.6(b) のように，画像ペアの陰影の差が少なく，テクスチャの分布が近い場合に提

¹ガンマ補正により，高輝度領域のダイナミックレンジは圧縮される．しかし，人間の視覚特性上，高輝度領域での繊細な描写は必要でないことが多い．更に，本研究で問題とされる高感度ノイズは主に低輝度で顕著となる．以上の観点よりガンマ補正された画像を実験に用いる．

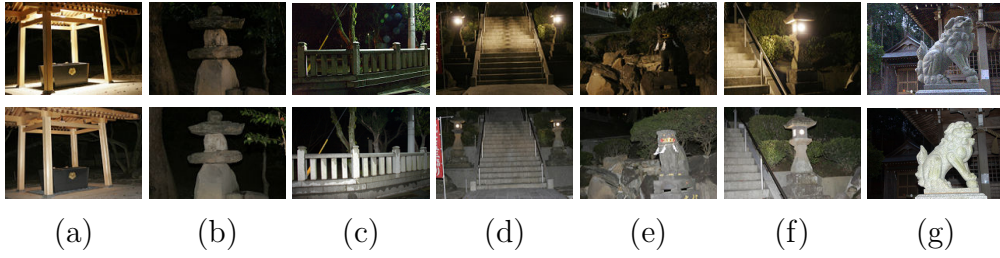


図 4.5: 実験に用いた高感度画像とフラッシュ画像

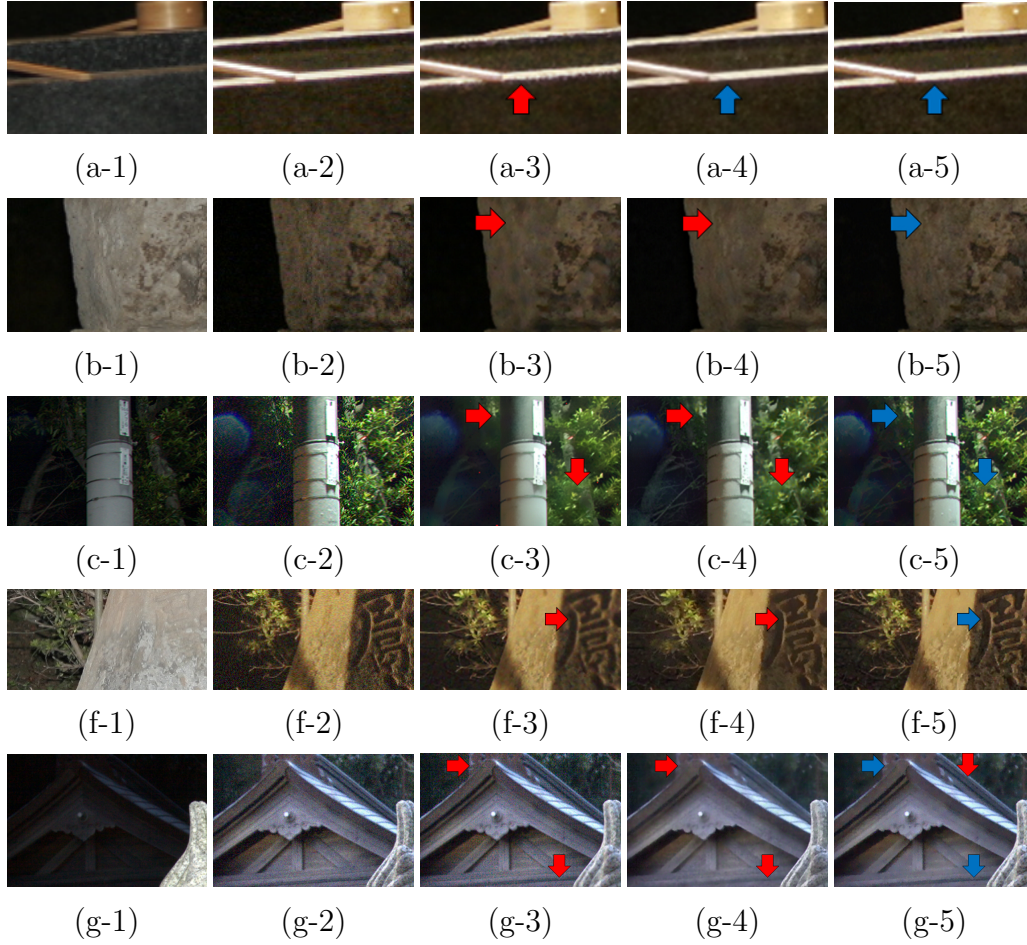


図 4.6: 実験 1 と実験 2 における提案法と従来法 [5] [6] のフラッシュ画像と高感度画像の合成法の比較:(*-1) フラッシュ画像, (*-2) 高感度画像, (*-3) Petschnigg らの結果 [5], (*-4) Shirai らの結果 [6], (*-5) 提案法



図 4.7: 単一画像の雑音除去手法 [7,8] との比較

案法は非常に有効である。これは式(4.3)の正則化項で、フラッシュ画像のテクスチャを強く反映させたためである。

低露光領域で提案法は、フラッシュ画像の微量な輝度変化で設計したBLFを正則化項に組み込むことで、コントラストを保持しつつ、雑音除去する。図4.6(c-5)より提案法が意図した結果となったことを確認できる。

補助画像を用いて雑音除去する提案法の有効性を示すため、単一画像の雑音除去手法との比較を行う。比較対象に、提案法と同様に凸最適化手法であるColor TV [8]、高性能雑音除去手法であるBM3D [7]を選択し、復元した画像の拡大図を図4.7に示す。また、復元対象は図4.5(b)の高感度画像であり、提案手法は補助画像にフラッシュ画像を使用した。単一の雑音除去手法は原画像以上に鮮明なテクスチャを復元することは困難であり、復元画像にボケを引き起こす(図4.7(c)(d)赤矢印部分)。更に、BM3Dは4.2節で指摘したように幾何学的な模様が発生している(図4.7(c)黄矢印部分)。一方、提案法はフラッシュ画像の勾配に近づける処理が含まれるため、単一画像での処理では困難な、ボケを抑制した詳細復元を実現している(図4.7(e)青矢印)。補助画像を用意する必要があるが、雑音除去において提案法は有効であるといえる。

4.4.2 復元精度の定量評価

実験2では、合成画像の復元精度をはかるために、長時間露光により撮影した画像と比較を行う。本実験における正解画像は三脚を用いて手ブレや被写体ブレのない長時間露光画像、更に正解画像に人為的にガウス雑音を加え、劣化したものを高感度画像(以降、劣化画像と呼ぶ)とする。劣化画像を各手法で処理した画像と、正解画像のピーク信号対雑音比(PSNR)を用いた定量評価を行う。フラッシュ画像は実験1と同様に撮影したものをを用いて復元画像を取得し、正解画像との誤差を測定した。ガウス雑音のパラメータは平均を $\mu = 0$ とし、標準偏差を $\sigma = 0.025$ と 0.05 の二値とした。各処理のパラメータについては表4.1に示す。4.3.1節におけるDTFの σ_s と σ_r 、処理回数 k については \mathbf{a}_s から完全に雑音が除去されるよう設定し、正則化パラメータである γ は雑音の標準偏差に近いものを選択した。 ρ は γ と同値としている。4.3.2節では正則化項に用いるBLFの設計が重要となる。実験により、 σ_s を大きく設定し、 σ_r は雑音の標準偏差に近い値を選択することで高い復元精度となることを確認した。表4.2に実験に用いた図4.5(d-g)の画像から得られたPSNRを示す。表4.2左側がガウス雑音の標準偏差 $\sigma = 0.025$ 、表4.2右側が $\sigma = 0.05$ の結果であり、最もPSNRが高いものを太字表記した。低露光領域における雑音量の変化による復元精度の定性評価を行う。フラッシュ画像に雑音が含まれる場合、手法[5]では復元画像に雑音を残す結果となった(図4.6(g-3))。[6]では雑音の増加により、不鮮明さが目立つ結果となった。また、提案法ではガウス雑音の標準偏差を大きく設定した場合エッジ付近に雑音が目立つ結果となった。JBLFにおいて、フラッシュ画像のエッジ近辺では g_r の重みが低くなり、結果と

表 4.1: 実験 2 における各処理のパラメータ

σ of noise	params of Sec.4.3.1				params of Sec.4.3.2			
	σ_s	σ_r	k	γ	ρ	σ_s	σ_r	λ
0.025	10	0.1	3	0.01	0.01	2	0.025	0.02
0.05	10	0.2	3	0.03	0.03	2	0.05	0.05

表 4.2: 画像ペア (d)-(g) の PSNR による評価
 雑音の標準偏差 $\sigma = 0.025$ 雑音の標準偏差 $\sigma = 0.05$

Image	[5]	[6]	Our	Image	[5]	[6]	Our
(d)	33.30	31.70	35.69	(d)	32.48	30.82	32.88
(e)	32.37	31.04	33.98	(e)	31.88	29.29	31.85
(f)	31.85	30.66	34.52	(f)	31.53	28.81	31.82
(g)	27.65	29.84	33.42	(g)	28.49	28.22	29.94

して式 (4.10) による平滑化が行われないことが原因である。高露光領域では雑音の増加により、三手法ともに色ムラが目立つ結果となった。なお、色ムラに対しては色ムラ除去法 [72] を併用することで改善できると考える。ガウス雑音の標準偏差 $\sigma = 0.025$ における画像の拡大図を図 4.6(f)(g) に示す。図 4.6(f) は従来法に比べ提案法の結果がより鮮明であるのが確認できる。図 4.6(g) のフラッシュ画像は少量のテクスチャしか保持しておらず、若干の雑音を含むシーンである。この場合でも提案法は従来法に比べてテクスチャを保持した雑音除去を実現した。

定量評価では、ガウス雑音の標準偏差が 0.025 の時、手法 [5,6] に比べ提案法の PSNR の数値が大幅に上回る結果となった (表 4.2 左側)。特に [6] はフラッシュ画像を局所領域における色変換で高感度画像に近似するため、フラッシュ画像の依存度が高く、雑音量の少ない画像の復元には適さない。一方、提案法は高露光領域で式 (4.3) によりストラクチャ成分とテクスチャ成分を分離した処理となる。フラッシュ画像の依存はテクスチャ成分のみであり、ストラクチャ成分を保護した結果、精度の高い復元を実現できたと考えられる。低露光領域では式 (4.10) の第二項により、劣化画像を基準にした処理のため微量の雑音では非常に有効となる。また、ガウス雑音が $\sigma = 0.05$ においても従来法に比べ、提案法が有効であることを示す結果となった (表 4.2 右側)。

第5章 位置ずれのある画像ペアを用いた画像復元手法

5.1 概要

雑音除去やブレ除去に対し、多くの手法が提案されているが、雑音やボケが大きな画像から鮮明な画像を復元することは、以前として困難な問題である。シンプルかつ効率的な取り組みとして複数枚画像の使用が挙げられる。文献 [5,6,47,48,73–75] では、画像の特徴を組み合わせ、ノンフラッシュ画像の鮮やかな色とフラッシュ画像の鮮明なコントラストを統合する。これらの手法は、特に暗い照明条件下で十分なぼけ除去またはノイズ除去能力を有するが、完全に位置合わせされた画像を必要とする。これは三脚などの固定装置でカメラを固定する必要があり、シーンが静止していなければならない、実際にこの条件を満たすことは困難である。文献 [9,36] では、位置がずれた画像ペアのカラーグレーディング方法を提案している。これらは非剛体の被写体でもうまく機能するが、画像に大きな照明や色の違いがあると失敗することがある。

本章は、位置合わせされていない画像ペアのための新しい画像統合手法を提案する。提案手法では、長時間露光画像とフラッシュ画像の二種類の画像を用いて、暗い照明条件下で鮮やかな色および高いコントラストを有する画像を復元する。提案手法は、低照度環境で鮮明な画像を得るために、両方の画像に好ましい特徴を組み合わせることを目的とする。また、新しい統合アルゴリズムの恩恵で、画像間の完全な位置合わせを必要とせず、画像内の照明が大きく異なる場合でも、自然な色と鮮明なコントラストを有する画像を生成することができる。

5.2 局所線形性を用いた画像統合

5.2.1 提案手法の概要

提案手法のフローチャートを図 5.1 に示す。提案するアルゴリズムは入力としてフラッシュ画像 P と長時間露光画像 Q の画像ペアを使用する。三脚を用いずに撮影されるため、これらの二つの画像は図 5.2(a)(b) に示すように位置ずれがある。

本手法の目的は、前述の画像集合 P と Q から高品質画像 X の再構成である。画素の位置合わせから始めて、二つの画像間の対応を探索する。対応にははしばし

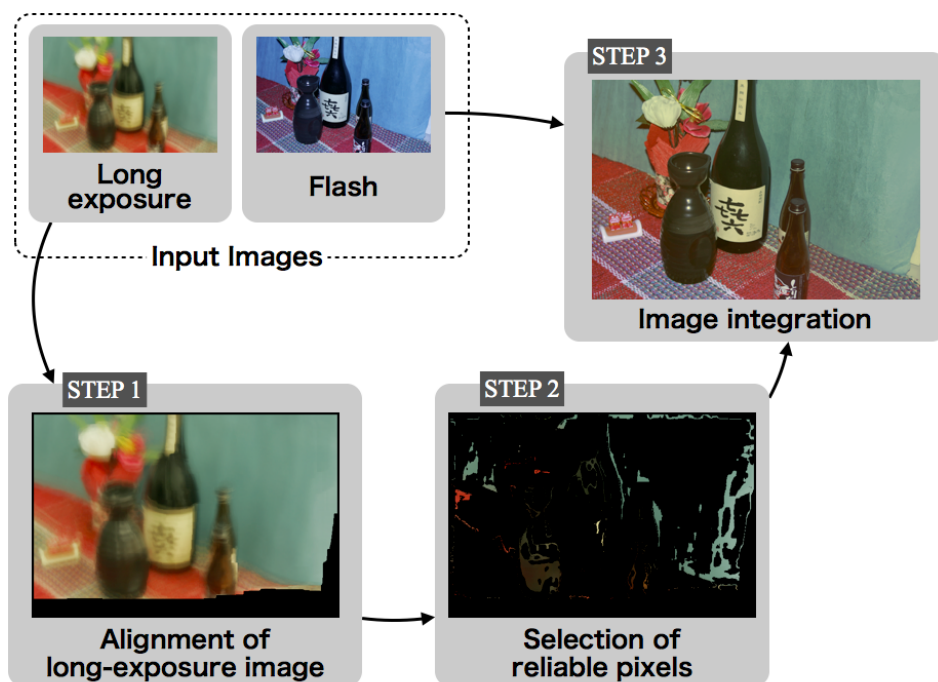


図 5.1: 提案手法のフローチャート

ば失敗が含まれるため，信頼のおける対応のみを用いる．次に，信頼のおける対応に基づいてフラッシュ画像全体の色を変換する．この処理は，局所照明変化に対するロバスト性を保証し，信頼できる画素対が一部のみであったとしてもうまく機能する．

本手法のアルゴリズムは以下のように要約される．

1. 対応探索：画像間の密接な対応関係を探索する．対応関係に基づいて，画像セットを位置合わせするために長時間露光画像を変換する（図 5.2(c)）．この処理には *SIFT flow* [54] を採用する．
2. 画素選択：位置合わせされた画像ペアから信頼できる画素のを選択し，後段の処理の性能を高める．
3. 画像統合：最後に，二つの画像を統合して高品質イメージを取得する．画像統合問題は，局所線形モデル (LLM) に基づいており，凸最適化問題として述べられる．この最適化手順は，長時間露光画像の色情報とフラッシュ画像の局所的特徴（詳細）を維持しながら統合する．

これら三つのステップはそれぞれ 5.2.2 節，5.2.3 節，5.2.4 節で述べられる．

5.2.2 対応探索と位置合わせ

フラッシュ/長時間露光の画像ペアが与えられた場合，まず二つの画像間の対応関係を探索する．パッチベースの [55,56] や特徴量ベースのメソッド [54] など，多くの対応探索アルゴリズムが提案されている．扱う画像ペアは，二つの画像が異なる光源を含んでおり，照明に堅牢な特徴量ベースの *SIFT flow* [54] で位置合わせする．*SIFT flow* は高密度対応アルゴリズムであり，オプティカルフローなどの画素に基づく対応探索よりも一般に信頼性が高い．

長時間露光画像のフラッシュ画像への位置合わせを次のように表す．

$$\mathbf{Q}' := T_P(\mathbf{Q}) \quad (5.1)$$

$T_P(\cdot)$ は *SIFT flow* 処理の演算子である．図 5.2 は，*SIFT flow* による位置合わせ例を示している．*SIFT flow* により変換された長時間露光画像 \mathbf{Q}' (図～5.2 (c)) がフラッシュ画像 \mathbf{P} (図～5.2 (a)) に大まかに位置合わせされたことが確認できる．図 5.2 (d) の灰色の画素は，図 5.2 (c) の位置あわせした領域の画素集合を示す．変換された \mathbf{Q}' において，位置合わせされた領域の画素からなる集合 \mathcal{S}' (図 5.2 (d) の灰色の画素) を定義する．

SIFT flow アルゴリズムは，対応関係を堅牢に見つけ，おおよその画像位置合わせが出来るが，いくつかの位置ずれに失敗した画素は依然として残っている．画像の間の変位が大きい場合，得られる位置合わせ結果画像は不一致をもち，これは後段の画像復元処理の品質を下げる．そこで，次の節で説明するように，位置合わせされた長時間露光画像から信頼できる画素を選択する．

5.2.3 信頼できる対応画素の選択

5.2.2 節で得られたおおよかに位置合わせされた画像は，通常，多くの不一致をもつ．従って，画像統合の性能を向上させるには，不要な領域を削除する必要がある．集合 \mathcal{S}' 内の信頼できる画素対照を選択するために，フラッシュ画像 \mathbf{P} と位置合わせした長時間露光画像 \mathbf{Q}' の局所分散の差を計算する．画素 i を中心とする正方形の局所領域 $\mathcal{N}_1(i)$ における，分散の差 $v_i^{p,q'}$ は次式で計算される．

$$v_i^{p,q'} := |\sigma_{p^Y,i}^2 - \sigma_{q'^Y,i}^2| \quad (5.2)$$

ここで， $\sigma_{p^Y,i}^2 := \sum_{j \in \mathcal{N}_1(i)} (p_j^Y - \bar{p}_i^Y)^2 / |\mathcal{N}_1|$ は局所領域における分散であり， $\bar{p}_i^Y := \sum_{j \in \mathcal{N}_1(i)} p_j^Y / |\mathcal{N}_1|$ は i を中心とする局所領域の平均値である． p_i と q'_i はそれぞれ \mathbf{P} と \mathbf{Q}' の i 番目の画素を意味する．上付き文字 Y はグレイスケール画像を意味する¹．局所領域 $\mathcal{N}_1(i)$ における画素数 $|\mathcal{N}_1|$ は全ての画素 i に対して固定である．位置合わせされた長時間露光画像の各画素は，(5.2) に大きな差がある場合には削

¹画像のグレイスケール化は次式を用いる $Y = 0.2989R + 0.5870G + 0.1140B$ ．

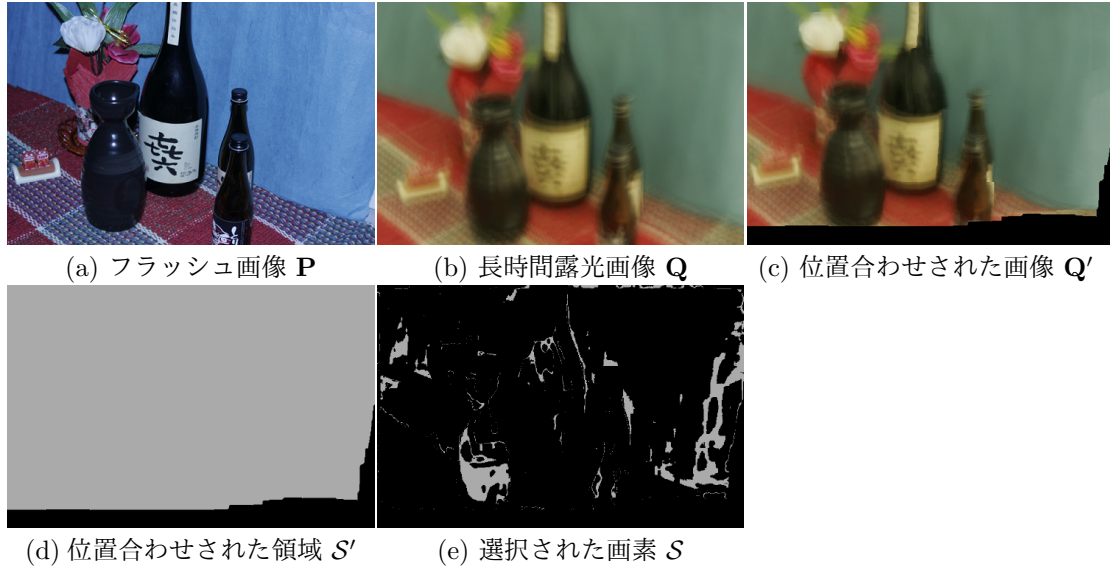


図 5.2: 画像間の位置合わせと画素選択: (a) フラッシュ画像, (b) 長時間露光画像, (c) 位置合わせされた長時間露光画像, (d) (c) における画素集合, (e) 選択された画素集合. (d),(e) において灰色は位置合わせされた領域と選択された画素を意味する.

除される. 残った信頼できる画素が, 画像統合方法の代表画素として 5.2.4 節で使用される. 残りの信頼できる画素からなる S' の部分集合を

$$S := \{s \in S' \mid v_s^{p,q'} \leq \tau\} \quad (5.3)$$

図 5.2 (e) は信頼における画素集合 S を意味する. 上記で使用するパラメータの詳細は 5.3.1 節にて述べる.

5.2.4 凸最適化を用いた画像統合

色補正の最も簡単なアプローチの一つは, フラッシュ画像 P 全体に均一な色変換を処理し Q' に適合させることである. 画像全体における変換は目的を達成することもあるが, 局所的な色の変化には適応できない. これは特に, 画素の対応がわずかしか見つからない場合に重大な欠点である. 本節では, 局所線形変換に基づく手法を導入することで問題を解決する.

提案するの画像統合モデルは, 各色チャンネルを独立して処理する. \mathbf{x} , \mathbf{p} , そして $\mathbf{q}' (\in \mathbb{R}^N)$ はそれぞれ, ベクトル化された所望画像 \mathbf{X} , フラッシュ画像 \mathbf{P} , 長時間露光画像 \mathbf{Q} , N は画素数を意味する. フラッシュ画像の詳細と長時間露光画像の色を維持するために, 局所線形モデル (LLM) に基づく画像統合モデルを提

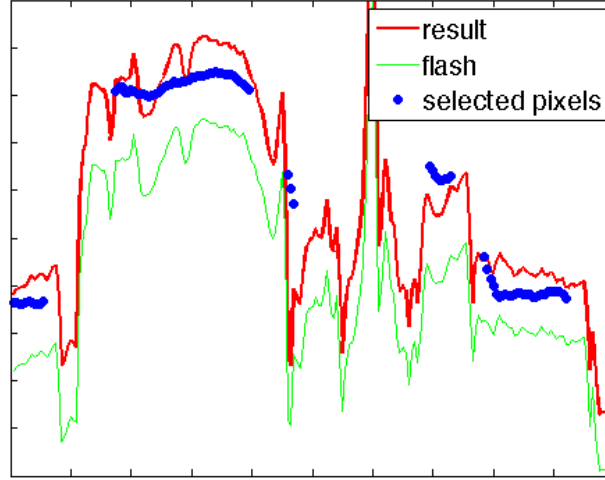


図 5.3: 各画像の画素値例。提案手法では、フラッシュ画像（緑色の線），および長時間露光画像の選択された画素（青色の点）から赤い線で示される画像を生成する。

案する。提案する画像統合モデルは次式の最小化問題で表される。

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{a}, \mathbf{b}} \quad & f(\mathbf{x}, \mathbf{a}, \mathbf{b}) \\ \text{s.t.} \quad & \sum_{s \in \mathcal{S}} (\bar{x}_s - q'_s)^2 < \eta \end{aligned} \quad (5.4)$$

目的関数 $f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ は LLM に基づいている。LLM は [76] のマッピングのために提案され，次式で表される。

$$f(\mathbf{x}, \mathbf{a}, \mathbf{b}) := \sum_i \left(\sum_{j \in \mathcal{N}_2(i)} (x_j - a_i p_j - b_i)^2 + \varepsilon a_i^2 \right) \quad (5.5)$$

ここで， (a_i, b_i) は局所領域 $\mathcal{N}_2(i)$ で一定と仮定される線形係数であり，第二項はチコノフ正規化項である。式 (5.4) で $s \in \mathcal{S}$ は 5.2.3 節で与えられた対応画素のインデックスあり， \bar{x} は \mathbf{x} の局所領域 $\mathcal{N}_2(i)$ における平均値である。

LLM (6.5) は，画像 \mathbf{P} をスケーリングとオフセットのパラメータ a_i と b_i で線形変換することを目的とする。二つのパラメータは各局所領域で異なるため，局所変換が実現される。式 (5.4) の制約は，領域 \mathcal{S} 内の結果画像 \mathbf{X} の画素が長時間露光画像の色を持つことを保証する。ここで，長時間露光の画像は手ぶれや動きのためにブレているため， \bar{x} がブレ画像と一致するように， x の代わりに \bar{x} を用いている。最後に，提案するモデル (5.4) は，結果画像のブレたものが領域 \mathcal{S} の \mathbf{Q}' と一致するように線形変換を局所的に適用する。

図 5.3 は，提案する画像統合モデル (5.4) の仕組みを示している。緑色の線はフラッシュ画像 \mathbf{p} を示し，青い点は長時間露光画像 \mathbf{q}' の選択された画素 q'_s である。

図中の赤い線は、図 5.3 は結果を示し、フラッシュ画像を長時間露光画像へ局所領域毎に近似したことで \mathbf{p} の大まかな形を維持しながら、結果として得られた画像が選択された画素に合わせて伸縮により近似されていることがわかる。

5.2.5 凸最適化

式 (5.4) を次のような制約のない問題に再定式化する。

$$\min_{\mathbf{x}, \mathbf{a}, \mathbf{b}} f(\mathbf{x}, \mathbf{a}, \mathbf{b}) + \iota_{C_{\{\mathbf{q}', \eta, \mathbf{M}\}}}(\mathbf{G}\mathbf{x}) \quad (5.6)$$

ここで、 $\iota_{(\cdot)}$ は 3.1.1 節で紹介した指示関数、 $C_{\{\mathbf{q}', \eta, \mathbf{M}\}} = \{\mathbf{x} \in \mathbb{R}^N \mid \|\mathbf{M}(\mathbf{x} - \mathbf{q}')\|_2 \leq \eta\}$ 、許容誤差は $\eta = \sqrt{\sum_{s \in \mathcal{S}} \hat{\eta}^2}$ であり、ここで $\hat{\eta}$ は、ユーザー定義のパラメータである。行列 $\mathbf{M} \in \{0, 1\}^{N \times N}$ は、画素を抽出するための対角行列であり、抽出画素は集合 \mathcal{S} に含まれるインデックス、局所領域 $\mathcal{N}_2(i)$ の平均値を計算する平滑化フィルタ行列を $\mathbf{G} \in \mathbb{R}^{N \times N}$ とした。式 (3.21) における PDS の各関数を $F := f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ 、 $G := 0$ 、 $H := \iota_{C_{\{\mathbf{q}', \eta, \mathbf{M}\}}}$ 、 $\mathbf{A} := \mathbf{G}$ とすると、(5.6) に対する PDS の各ステップでは次のように表される（上付き文字 k は反復番号を示す）。

$$\begin{cases} (\mathbf{x}^{k+1}, \mathbf{a}^{k+1}, \mathbf{b}^{k+1}) := \mathbf{x}^k - \gamma_1 \nabla f(\mathbf{x}^k, \mathbf{a}^k, \mathbf{b}^k) - \gamma_1 \mathbf{G}^\top \mathbf{z}^k \\ \hat{\mathbf{z}}^{k+1} := \mathbf{z}^k + \gamma_2 \mathbf{G}(2\mathbf{x}^{k+1} - \mathbf{x}^k) \\ \mathbf{z}^{k+1} := \hat{\mathbf{z}}^{k+1} - \gamma_2 P_{C_{\{\mathbf{q}', \eta, \mathbf{M}\}}} \left(\frac{\hat{\mathbf{z}}^{k+1}}{\gamma_2} \right) \end{cases} \quad (5.7)$$

ここで、 $P_{C_{\{\mathbf{q}', \eta, \mathbf{M}\}}}(\cdot)$ は式 (3.5) で定義した近接写像を意味する。以下では、 $\nabla f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ と近接写像の計算方法を述べる。

$\nabla f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ の算出方法

まず、解 b_i は、 $f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ 導関数がゼロをとる点で与えられる

$$b_i^* = \bar{x}_i - a_i \bar{p}_i, \quad (5.8)$$

ここで、 \bar{x}_i, \bar{p}_i は局所領域の平均値である。(5.8) と同じ手順を用い、 a_i の最適解を次のように求める。

$$a_i^* = \frac{\sum_{j \in \mathcal{N}_2(i)} (x_j - \bar{x}_i)(p_j - \bar{p}_i)}{\sum_{j \in \mathcal{N}_2(i)} (p_j - \bar{p}_i)^2 + \varepsilon} \quad (5.9)$$

ここで、(6.5) を行列式で表すために二つのベクトル $\boldsymbol{\delta}_{(i)} \in \mathbb{R}^N$ と $\boldsymbol{\mu}_{(i)} \in \mathbb{R}^N$ は導入し、 k 番目の $\boldsymbol{\delta}_{(i),k}$ と $\boldsymbol{\mu}_{(i),k}$ は以下のように定義される。ベクトル $\boldsymbol{\delta}_{(i)}$ はベクトル化された画像 \mathbf{x} の i 番目の要素を抽出する。

$$x_i = \boldsymbol{\delta}_{(i)}^\top \mathbf{x}, \quad \text{where } \boldsymbol{\delta}_{(i),k} = \begin{cases} 1, & \text{if } k = i \\ 0, & \text{otherwise} \end{cases}$$

ベクトル $\boldsymbol{\mu}_{(i)}$ は、局所領域 $\mathcal{N}_2(i)$ の画素の平均を取る：

$$\bar{x}_i = \boldsymbol{\mu}_{(i)}^\top \mathbf{x}, \text{ where } \boldsymbol{\mu}_{(i),k} = \begin{cases} \frac{1}{|\mathcal{N}_2(i)|}, & \text{if } k \in \mathcal{N}_2(i) \\ 0, & \text{otherwise} \end{cases}$$

このベクトルを用いて二つの行列 $\widehat{\mathbf{M}}_{(i)} \in \mathbb{R}^{N \times N}$ と $\widehat{\mathbf{G}}_{(i)} \in \mathbb{R}^{N \times N}$ は次のようになる。

$$\widehat{\mathbf{M}}_{(i),k} = \begin{cases} \boldsymbol{\delta}_{(k)}^\top, & \text{if } k \in \mathcal{N}_2(i) \\ \mathbf{0}^\top, & \text{otherwise} \end{cases}$$

$$\widehat{\mathbf{G}}_{(i),k} = \begin{cases} \boldsymbol{\mu}_{(i)}^\top, & \text{if } k \in \mathcal{N}_2(i) \\ \mathbf{0}^\top, & \text{otherwise} \end{cases}$$

ここで、 $\widehat{\mathbf{M}}_{(i),k} \in \mathbb{R}^N$ と $\widehat{\mathbf{G}}_{(i),k} \in \mathbb{R}^N$ はそれぞれ $\widehat{\mathbf{M}}_{(i)}$ と $\widehat{\mathbf{G}}_{(i)}$ の k 番目列ベクトルであり、 $\mathbf{0} \in \mathbb{R}^N$ は要素が全てゼロのベクトルである。前述の \mathbf{M} と \mathbf{G} と区別するために、記号 $\widehat{\cdot}$ が導入されている。 k 番目の画素が i を中心とする局所領域に含まれている場合、行列 $\widehat{\mathbf{M}}_{(i)}$ は 1 を持つ対角行列となり、マスキング操作を実行する。

二つの行列の差によって与えられる行列 $\mathbf{V}_{(i)} \in \mathbb{R}^{N \times N}$ をさらに紹介する。 $\mathbf{V}_{(i)}$ は局所領域 $\mathcal{N}_2(i)$ のピクセル値を抽出するための局所中心化行列で、次に平均値を減算する。

$$\mathbf{V}_{(i)} = \widehat{\mathbf{M}}_{(i)} - \widehat{\mathbf{G}}_{(i)}$$

式 (6.5) は $\mathbf{V}_{(i)}$ を用いて行列として $\nabla f(\mathbf{x}, \mathbf{a}, \mathbf{b})$:

$$f(\mathbf{x}, \mathbf{a}, \mathbf{b}) := \sum_i \|\mathbf{V}_{(i)} \mathbf{x} - a_i^* \mathbf{V}_{(i)} \mathbf{p}\|_2^2 + \varepsilon a_i^{*2} \quad (5.10)$$

結局、関数 $f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ の勾配は、次のように与えられる。

$$\begin{aligned} \nabla f(\mathbf{x}, \mathbf{a}, \mathbf{b}) &= \sum_i \mathbf{V}_{(i)}^\top (\mathbf{V}_{(i)} \mathbf{x} - a_i^* \mathbf{V}_{(i)} \mathbf{p}) \\ &= \sum_i \mathbf{V}_{(i)} \mathbf{x} - \sum_i a_i^* \mathbf{V}_{(i)} \mathbf{p} \end{aligned} \quad (5.11)$$

ここで、行列の性質を用いて $\mathbf{V}_{(i)} = \mathbf{V}_{(i)}^\top \mathbf{V}_{(i)}$ としている（詳細は付録 A を参照）。 $\nabla f(\mathbf{x}, \mathbf{a}, \mathbf{b})$ の計算は演算量が多く見えるが、積分フィルタ [77] を用いて $\mathcal{O}(N)$ で計算可能である。

最後に、(5.7) の処理を使用して $\mathbf{x}, \mathbf{a}, \mathbf{b}, \mathbf{z}$ を更新し、すべての色チャンネルから得られた解 \mathbf{x} を統合結果とする。アルゴリズム 1 は本手法の手順を示す。

5.3 実験

まず、パラメータの設定について説明し、従来法との比較を行い、本手法の有効性を検証する。

Algorithm 1 Algorithm of our method.

Require: \mathbf{P} : Flash image, \mathbf{Q} : long-exposure image

procedure IMAGE REGISTRATION(\mathbf{P}, \mathbf{Q})

$\mathbf{Q}' := T_p(\mathbf{Q})$

 Define a set that consists of the pixels in the aligned region: \mathcal{S}'

end procedure

procedure PIXEL SELECTION($\mathbf{P}, \mathbf{Q}', \mathcal{S}'$)

for $i \leftarrow 1, N$ **do**

$v_i^{p,q'} := |\sigma_{p_i}^2 - \sigma_{q'_i}^2|$

end for

$\mathcal{S} := \{s \in \mathcal{S}' \mid v_s^{p,q'} \leq \tau\}$

end procedure

procedure IMAGE INTEGRATION($\mathbf{P}, \mathbf{Q}', \mathcal{S}$)

 Set γ_1 and γ_2 .

 Choose $\mathbf{x}^{(0)}, \mathbf{a}^{(0)}, \mathbf{b}^{(0)}, \mathbf{z}^{(0)}$.

while A stopping criterion is not satisfied **do**

$(\mathbf{x}^{k+1}, \mathbf{a}^{k+1}, \mathbf{b}^{k+1})$

$:= \mathbf{x}^k - \gamma_1 \left(\sum_i \mathbf{V}_{(i)} \mathbf{x}^k - \sum_i a_i^* \mathbf{V}_{(i)} \mathbf{p} + \mathbf{G}^\top \mathbf{z}^k \right)$

$\hat{\mathbf{z}}^{k+1} := \mathbf{z}^k + \gamma_2 \mathbf{G} (2\mathbf{x}^{k+1} - \mathbf{x}^k)$

$\mathbf{z}^{k+1} := \hat{\mathbf{z}}^{k+1} - \gamma_2 P_{\gamma_2 C_{\{\mathbf{q}', \eta, \mathbf{M}\}}} \left(\frac{\hat{\mathbf{z}}^{k+1}}{\gamma_2} \right)$

end while

end procedure

5.3.1 実験におけるパラメータ設定

5.2.3節の画素選択では、 τ と局所領域 $|\mathcal{N}_1|$ のサイズを指定する必要がある。実験的に、図5.4のパラメータの影響を調べた。除去された領域は閾値が小さく設定された時と局所領域サイズが小さい時に拡大する（それぞれ図5.4の中段、下段参照）。本実験では $\tau = 10^{-2}$ と $|\mathcal{N}_1| = 21 \times 21$ を採用した。

式(6.5)において局所領域サイズ $\mathcal{N}_2(i)$ とチコノフ正則化のために ε を指定する必要がある。提案式(5.6)では、理想的な画像 \mathbf{x} はフラッシュ画像 \mathbf{p} の詳細を保持しつつ長時間露光画像 \mathbf{q} の輝度値に近似される。LLMは局所領域 $|\mathcal{N}_2|$ が大きい値に設定されていると、合成画像はより正確にフラッシュ画像の詳細を維持する。私たちは、様々画像に対する実験で $|\mathcal{N}_2| = 31 \times 31$ がうまく動作すること確認した。チコノフ正則化のためのパラメータは $\varepsilon = 10^{-5}/|\mathcal{N}_2|$ と設定した。この値が大きいほどガイドドフィルタ[46]の正則化と同様に、画像にボケが生じる。更

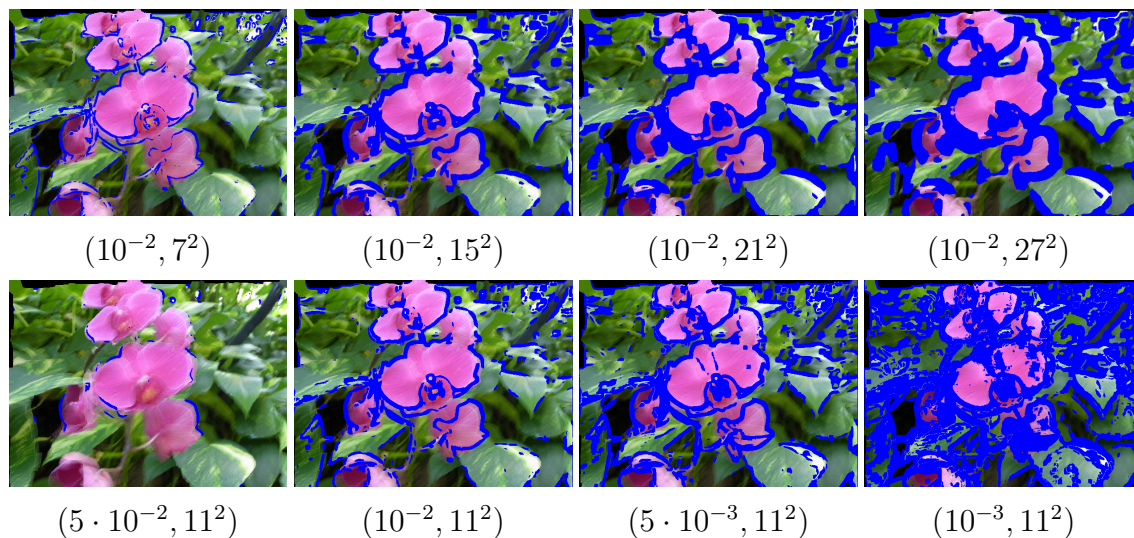


図 5.4: 画像位置合わせ結果. (a) フラッシュ画像, (b) 長時間露光画像, (c) *SIFT-flow* (5.1) による位置合わせ. 異なるパラメータ (τ , $|\mathcal{N}_1|$) による画素選択の結果. 青の画素は取り除かれた画素を意味する.

に, ε の値が小さすぎると, 推定が不安定になることを確認した.

PDS で使用されるパラメータは, $\gamma_1 = 10^{-4}$ および $\gamma_2 = 1/(12\gamma_1)$, 許容誤差 η のパラメータ $\hat{\eta}$ は $5 \cdot 10^{-2}$ に設定した.

5.3.2 従来法との比較

二つの画像を連続して撮影する場合, 二つの画像は類似しているが, 手ブレによる画像と物体の動きには違いがある. 提案法はフラッシュ画像の局所的な詳細と, 長時間露光画像の自然な色を持つ画像の取得が目的であり, この点について実験結果を確認する.

提案手法をいくつかの従来手法と比較する. 有効性を検証するために従来法に Pitié の方法 [2], HaCohen らの方法 [9], Hwang らのメソッド [36] を用いた. 図 5.6 は, 図 5.5 のフラッシュ長時間露光画像のペアから得られた結果を示す. 最初の二つの画像 (I1) と (I2) は同じシーンであるが, (I2) は (I1) よりもブレが大き

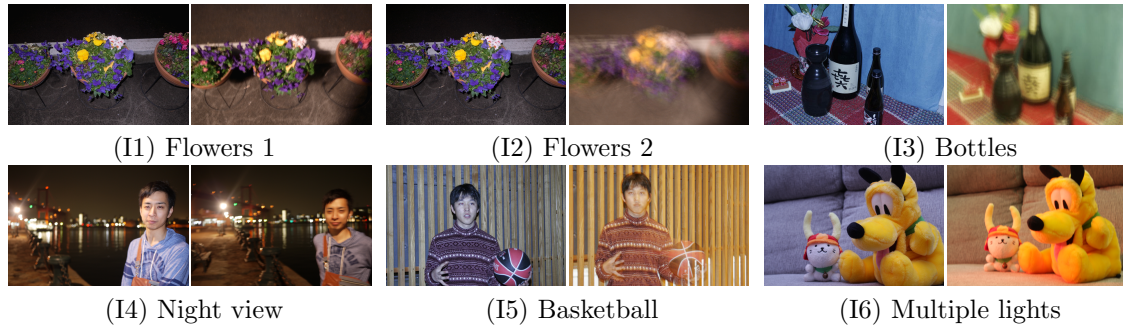


図 5.5: 入力画像：左はフラッシュ画像，右は長時間露光画像

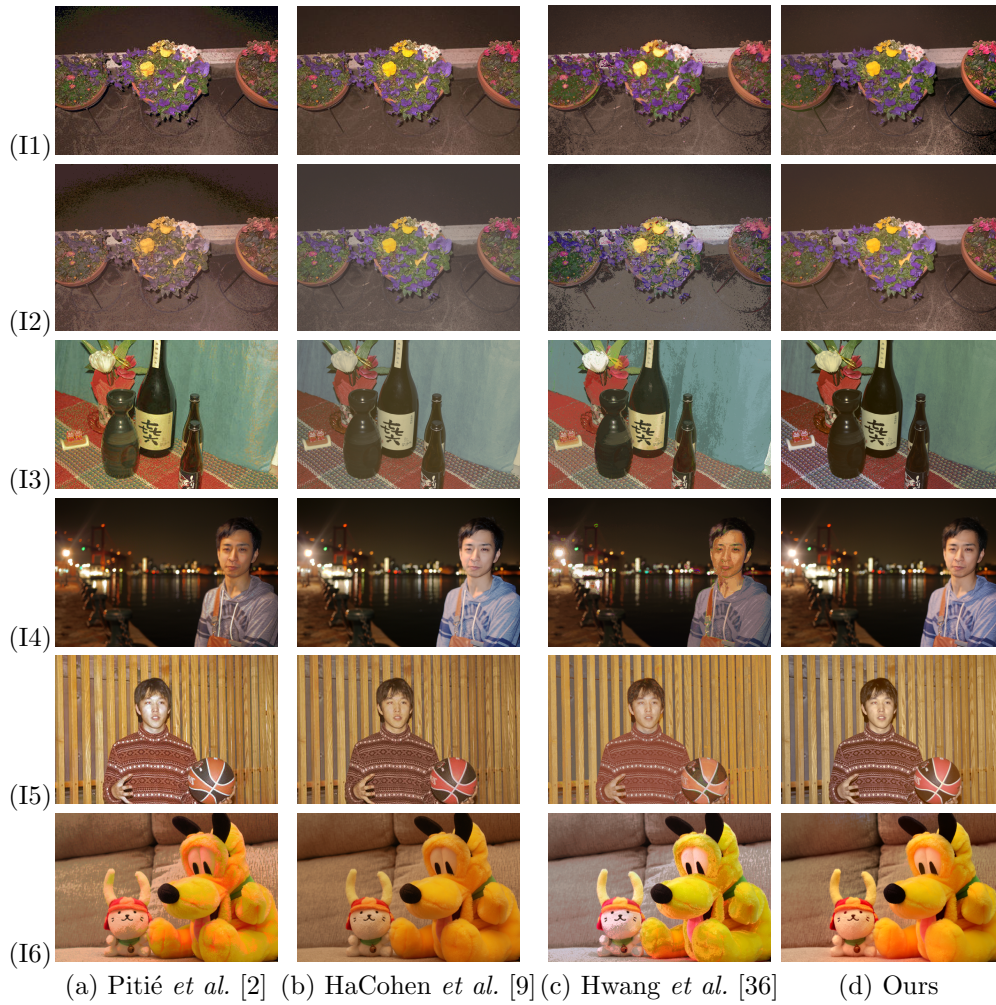


図 5.6: 従来手法との比較. 各行の画像は図 5.5(I1) から (I6) を入力画像とした時の各手法の結果.

い。第三の画像セット (I3) は色の差が大きく, (I4) および (I5) は非剛体運動を有する。六番目のシーン (I6) には複数の照明がある (赤い光がオブジェクトの左から照射されている)。著者らが公表したソースコードを用いて, [9] と [2] の結果を得た (図 12~図 13 の (I2-b) および (I3-b) [9] は色の差が大きくに画像間の対応関係えられなかったため, SIFT flow の位置合わせよりトーンカーブを推定した。文献 [36] は独自に実装したが, 文献内で用いる画像探索は本実験でうまく動作しないため, NRDC の対応結果を使用した。

Pitié [2] の結果は, 部分的にフラッシュ画像と長時間露光画像の自然な色合いの高いコントラストを持つ。しかし, 図 7~8 に示すように, いくつかのゲインアーティファクトを生じる。HaCohen らの手法 [9] は, (I1-b) と (I5-b) において良好な結果であるが, 画像に大きなぼけがあり, フラッシュ/長時間露光画像ペアのカラーコントラストを減少させる。画像対の間の対応する領域に基づく全体的な色変換を使用するので, (I6-b) に示す混合照明による画像対に対してはうまく動作しない。Hwang らの手法 [36] は, (I1-c) と (I5-c) に対しても良好な結果だが, 動きのある領域または部分的な照明変化が (I4) と (I6) のような長時間露光画像ではぼやけている。提案手法は, フラッシュ画像の細部を維持しながら, 長時間露光画像からフラッシュ画像への色の統合に成功し, 画像は大きなぼやけおよび照度差を有するにもかかわらず, 他の方法より高いコントラストおよび自然な色を有する。

5.3.3 異なる画像ペアを用いた画像復元

ここでは追加のアプリケーションとして, フラッシュ/高感度画像ペア, 類似シーン画像ペアを用いたときも提案法が有効であることを実験で示す。

雑音画像を参照画像とした実験

本章では, ブレ画像 (すなわち, 長時間露光画像) を参照とした復元を示してきた。この実験では, 高い ISO 感度で撮影された雑音の多い高感度画像を参照として提案手法による画像統合を試みる。フラッシュ画像と高感度画像との統合を実現するために, 式 (5.6) を以下のように組み替える

$$\min_{\mathbf{x}, \mathbf{a}, \mathbf{b}} f(\mathbf{x}, \mathbf{a}, \mathbf{b}) + \iota_{C_{q', \eta, M}}(\mathbf{x}) \quad (5.12)$$

(5.6) の第二項では, 長時間露光画像がぼけているため, 結果画像のブレたものが長時間露光画像に近づくように $\mathbf{G}\mathbf{x}$ を用いた。しかし, 雑音画像が参照の場合, この定式化は品質を低下させるため, $\mathbf{G}\mathbf{x}$ ではなく, (5.12) の第二項で \mathbf{x} を直接使用する。

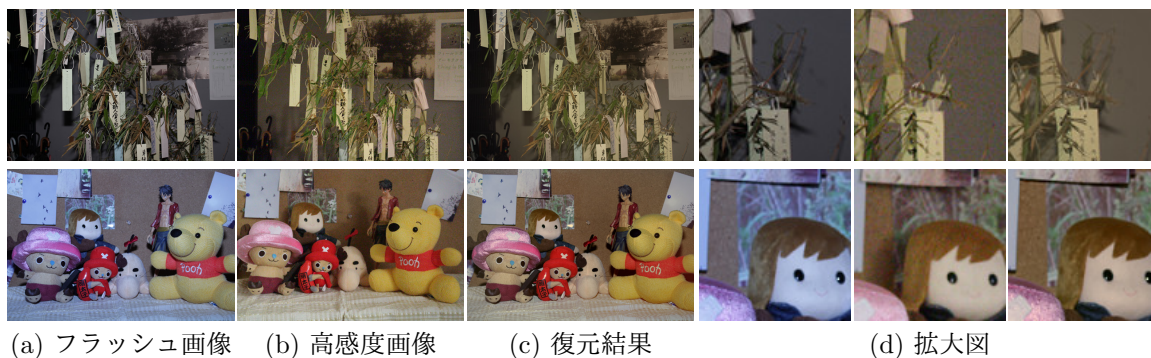


図 5.7: フラッシュ／高感度画像ペアにおける画像復元. 左から：(a) フラッシュ画像, (b) 参照画像としての高感度画像, (c) 提案手法における復元結果, (d) 各画像の拡大図.

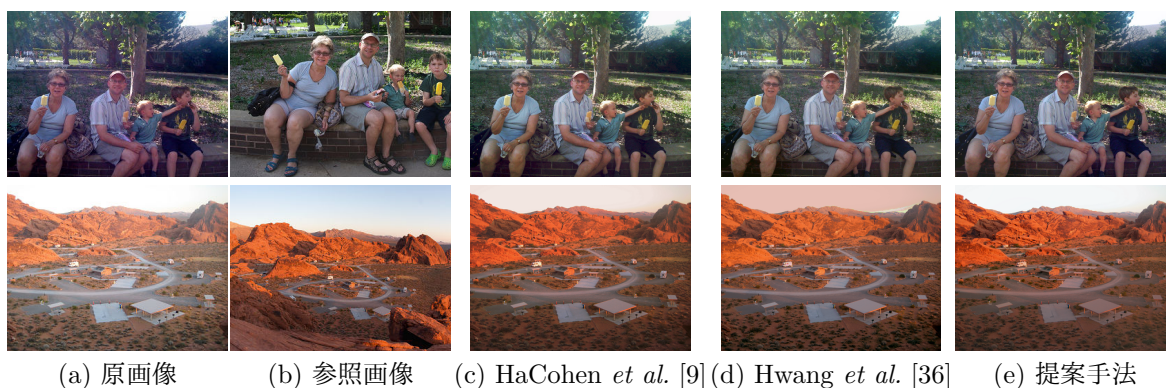


図 5.8: NRDC [9] の画像ペアにおける色変換. 原画像の色を各手法において参照画像に近似している.

図 5.7 は、位置ずれのあるフラッシュ/高感度画像ペアを使用した場合の結果画像を示す. 図 5.7 (d) では、高感度画像からの鮮やかな色とフラッシュ画像の鮮明なディテールをもつ画像が得られている. ここで、[5] のようなフラッシュ/高感度画像を扱う従来の方法では、完全に画素が一致している画像に対して提案されているため、本実験画像では得られる画像が大きく劣化することに注意されたい.

類似シーンを参照画像とした色変換

式 (5.12) は、色変換としても機能する. 図 5.8 は、いくつかのシーンの色変換結果を示す. 私たちは、このような入力画像に対して信頼できる対応画素を得るために、*SIFT flow* の代わりに HaCohen らの対応アルゴリズム [9] を使用した. この実験は、原画像が参照画像の色を持つようにすることを目的としている. 従来の方法でも良好な結果が得られるが、提案手法においても良好な色変換を達成し

ており，その性能は従来法に匹敵するか，またはそれより優れていることが確認できる．

5.3.4 備考

提案法は対応探索と最適化に基づくアプローチを使用しているので，計算の複雑さは他の方法 [2,9] よりも高い．例として，3.2GHz の i5 CPU を搭載した計算機では，画素の位置合わせ (*SIFT flow*) に 39 秒かかり，最適化問題の解決には 25 秒．一方，HaCohen らの手法 [9] は約 20 秒，Pitie らの手法 [2] は 8 秒を要した．計算の複雑さを減らすことが将来的に求められる．

第6章 参照画像を用いた自動肌色補正

6.1 概要

人物写真は背景色の反射によって望ましくない色になることがあるため、写真ごとに肌の色や明るさを統一する必要がある。このような状況では、通常、特別な装置を用いて照明条件を調整するが、写真家ではない人にとっては、同じ品質を達成するためには照明条件の調整は困難である。このため、複数の画像間の色や明るさを統一するための特別な装置を必要としない方法が求められている。

本章では、様々な色や明るさをもつ顔画像の補正の際、補助画像を用いた手法を提案する。提案手法では、コストをかけずに高品質な顔画像を含む年鑑などの写真リストを作成するなど、さまざまなアプリケーションを提供できる。異なるカメラで異なる照明条件で撮影した写真を収集しようとする、既製の画像処理ソフトウェアを使用しても画像を修正することは膨大な時間を要する。顔写真補正の際、考えられる状況では、顔の肌の色、明るさなどいくつかの特徴を統一する必要がある。自動修正は次の二ステップで処理できる。

- (1). 顔検出と顔の肌の色抽出.
- (2). 顔の肌の色補正.

これらの技術の単純な組み合わせは、以下の理由により画像を補正することができないことが多い。

- 抜き出した肌色のみを補正すると境界で差異が生じる。
- 各カメラには独自のカメラ応答感度があり、色空間における肌色の分布は、使用されるカメラにも依存する。
- 背景色は顔に反映されるため、色が歪む。
- 複数の色の服を着用すると、画像領域全体の色が肌色を歪ませ、肌色補正が衣服を変色させる。

したがって、顔領域のみに影響を及ぼす色補正処理が必要である。

画像間の色（各色のトーンカーブや色分布）を一致させるカラーグレーディング法 [2,9,35] は、複数の画像間の色補正に有効である。しかし、これらの既存の

技術は、画像全体を補正する手法であり、背景および衣服までも補正されてしまう。別のアプローチとして、補助画像に理想的な肌色を持つ画像用いた補助画像フィルタリングに基づく色変換があげられる [3,5,45,68]。しかし、補助画像フィルタリングは補助画像と原画像間で、位置ずれのない完全に位置合わせされた画像を用意する必要があり、これは多くの状況に当てはまらない。

顔画像補正法として、皺除去法 [78] が提案されているが、目的は顔面色の色補正とは異なる。文献 [10] で提案された方法は、原画像のスタイルを参照画像に変換することができる。しかし、顔写真に対してのみの手法であるとともに衣服をも参照画像に近似してしまう。

本章で紹介するアルゴリズムでは、ハイブリッド GF を使用して顔の肌の色と明るさの補正を行う¹。領域抽出と色転送は別の手順であるため、ある領域の色補正は、原画像の色と変換部分との間に色の急激な変化による違和感が生まれる。この問題を解決するために、色変換と変換領域の判定を同時に行うハイブリッド GF (図の 6.1 の赤枠) を提案する。手法では顔領域のみが抽出され (図 6.1(1))、色補正による調整が行われた後、ハイブリッド GF は、補正された画像を参照画像として、補正領域以外の色を維持しつつ入力画像の顔領域を補正する (図 6.1(2)) つまり、ハイブリッド GF は、画像の一部を境界に違和感を生じさせずに補正可能である。He らの GF [68] が完全に位置合わせされたガイド画像を必要であることに對し、提案する GF は、おおまかに位置合わせされた画素しか必要としない。

6.2 提案手法

提案するアルゴリズムは、一般的な色変換方法とは異なり、入力画像の一部の色を目標の色に変換し、変換された周辺箇所を補正する。図 6.1 は本手法の方法のフローチャートを示す。提案手法は以下に示す二つの処理で構成される。

- (1). 顔の検出と顔の肌の色抽出 (図.6.1 黄色枠) : 顔の部分を検出し、顔の肌の色を抽出する。
- (2). 肌の色補正 (緑枠) : ステップ (1) で肌領域を抽出した後、画像 (a) を用いて顔の肌色の分布を修正する。また、画像 (b) をガイド画像として顔領域の色を変更する²。

また、本章では、肌色と顔肌色の二つの用語を使用する。肌色は全身の肌色を意味し、顔肌色は顔のみの肌色である。提案手法の新規性は (2) にあり、色補正は画像の一部のみに影響する。顔検出のような他のステップのためにいくつかの修

¹前章における知見を用いており、式が類似するが、式設計の柔軟性や各色チャンネルではなく、RGB 色空間における定式化となっている。

²本章では分けられた範囲を領域と区域で分けて説明する。区域は検出された顔区域を 6.2.1 節で示すように矩形で形成され、領域は抽出された顔の肌色領域を 6.2.2 節で示し分けられた範囲の形は問わない。

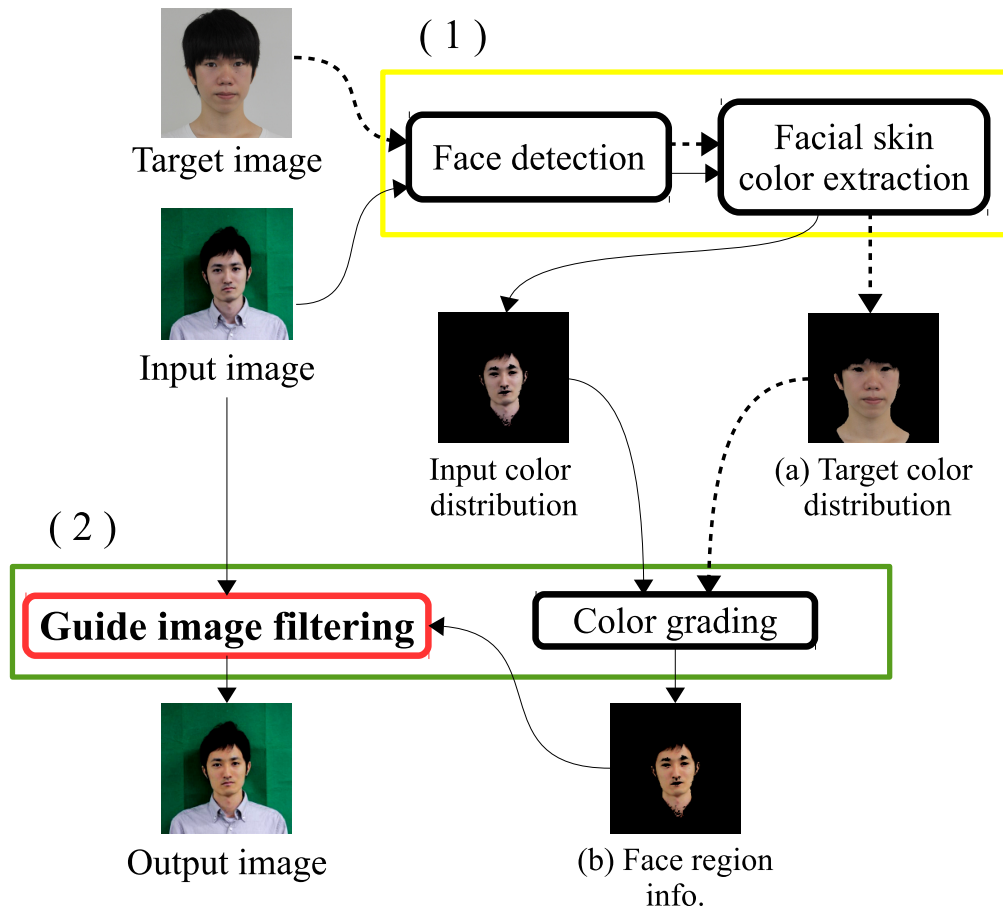


図 6.1: 提案手法の流れ

正を加えた従来の画像処理手法を使用する。各手順は、以降の節でより詳細に説明する。

6.2.1 顔検出

まず、顔検出のための Haar-Cascade 検出 (Viola-Jones アルゴリズムとも呼ばれる) を使用し、頭から肩までの区域を検出する。顔区域は、重心座標のセット (x, y) と検出された候補矩形のサイズ (w, h) とを有する複数の矩形ウィンドウによって記述される図.6.2 (中央)。我々は、 x, y, w, h の中央値を以下のように求める。

$$\forall_{c \in \{x, y, w, h\}} \hat{c} := \text{median}(\{c_i\}) \quad (6.1)$$

ここで、 i は候補のインデックスである (図 6.2 の各青枠が検出されたいくつかの候補を示す)。矩形は、重心座標 (\hat{x}, \hat{y}) の中央値を最終顔領域として採用する。顔領域のサイズ (\hat{w}, \hat{h}) は画像に依存するため、顔領域のサイズは次のように調整さ

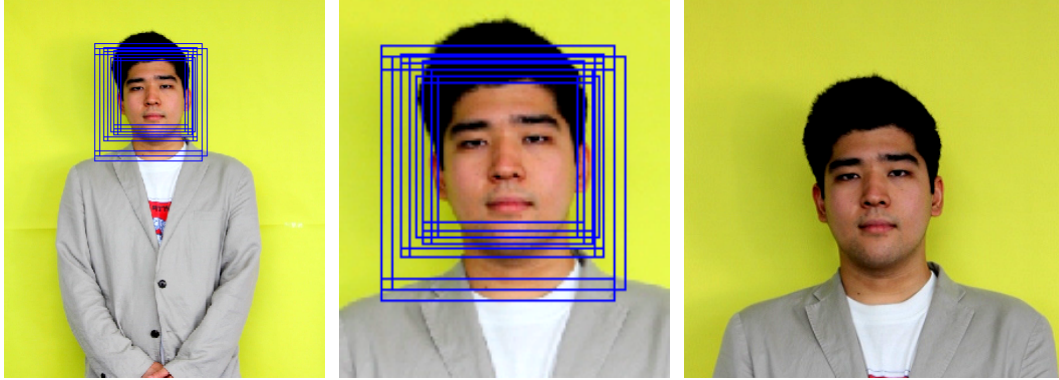


図 6.2: 顔検出. (左) 元画像, (中央) 検出された顔領域 (候補), (右) 元画像の顔領域. 青い四角は顔の候補を示す.

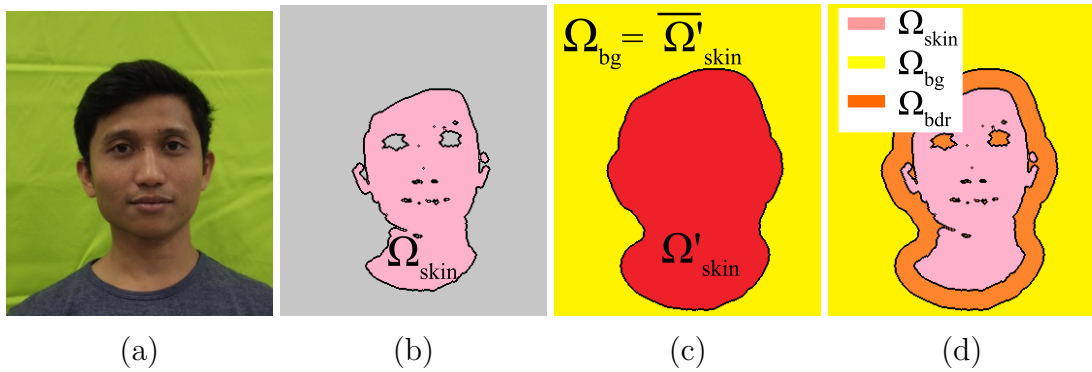


図 6.3: 各領域の例. (a) 原画像 Y , (b) 肌色領域 Ω_{skin} , (c) 膨張した肌色領域 Ω'_{skin} , and (d) 画像内における各領域. 桃, 赤, 黄, および橙はそれぞれ Ω_{skin} , Ω'_{skin} , Ω_{bg} , and Ω_{bdr} を意味する

れる.

$$2\hat{r} + 1 \times 2\hat{r} + 1 \quad \text{where} \quad \hat{r} := s\hat{w} \quad (6.2)$$

ここで, 上式は任意のスケール変数 s を用いて計算されることに注意されたい.

6.2.2 肌色抽出

顔区域における肌色領域の抽出は, モルフォロジー演算と色空間でのクラスタリングによって達成される.

- (i) HSV 色空間内の画像全体の色分布をクラスタリングすることによって各画素の色を分類し, 各画素について, 各画素が属するクラスターのラベルを割り当てる.
- (ii) 一部の領域は, 画像空間内の同じラベルで隣接する画素を連結することによって生成され, 次に, 検出された顔領域内の主要クラスターの属する領域 Ω_{rect}

(Eq. (6.2)) を抽出し、顔肌色領域に設定する.

検出された顔領域内 (6.2.1 節) の肌色は既知であるので、色分布をクラスタリングするために k-means 法 [79] のようなより簡単な手法を用いることによって満足のいく結果を得ることができる.

本論文では、簡単化のため、HSV 色空間における *hue* 値の一次元分布に k 平均クラスタリングを適用し、上記 (i) と (ii) の手順を実行して顔領域 Ω_{skin}^{hue} を検出する. 次に、すべての画素の HSV 色空間における *saturation* および *value* を区域 Ω_{skin}^{hue} で閾値処理を実行する. 肌色領域 Ω_{skin} と肌色の条件を画素 p で実験的に以下のように定義する.

$$\Omega_{skin} \ni p : \begin{cases} p \in \Omega_{skin}^{hue} \cap s_p \in [\hat{s} - 0.2, \hat{s} + 0.2] \\ v_p \in [0.15, 0.95] \end{cases} \quad (6.3)$$

ここで、 \hat{s} は顔領域の *saturation* の中央値であり、*saturation* と *value* のしきい値は $[0, 1]$ に正規化されている. クラスタの数は陰影などの影響を受け、シーンによって異なるが本手法では、 $k=4$ と統一する.

ハイブリッド GF ため、モルフォロジー演算による領域拡張によって Ω_{bg} および Ω_{bdr} を定義する. Ω_{bg} および Ω_{bdr} は次のように与えられる.

$$\begin{aligned} \Omega'_{skin} &= dilate(\Omega_{skin}) \\ \Omega_{bg} &= \overline{\Omega'}_{skin} \\ \Omega_{bdr} &= \Omega'_{skin} \setminus \Omega_{skin} \end{aligned} \quad (6.4)$$

ここで、 $dilate(\cdot)$ は、20 画素の半径の円からなる構造要素による、モルフォロジー演算の拡張関数である. 各領域の一例を図 6.3 に示す.

6.2.3 文献 [2] によるカラーグレーディング

図 6.4(a) において、抽出された領域 Ω_{skin} の肌色 (より具体的には、色分布の形状³) が目標の肌色に近づくように、色変換が行われる (図 6.4). 図 6.4 の下段は、上段の画像の RGB 色分布を示す. カラーグレーディングは、色分布の形状 (RGB の三次元座標) を変換するため補正結果の分布が目標画像の色分布となる. 各 RGB 色成分を修正するため、トーンカーブ推定などのより単純な技術を使用してもよいが、その場合は画像間の画素対応を必要とし、異なる人の顔の間で色変換させる本手法の場合には適していない. Pitie らの方法 [2] は画素対応を必要としない. この手法は、特定の軸に沿った色分布の幅を考慮し、指定された軸に一致するように反復的に近似していき、色分布の非線形変換を実現する.

³前のセクションで使用した HSV 色空間の代わりに RGB 色空間を使用することに注意されたい. これは、カラーグレーディング法 [2] が RGB 色空間に対して提案され、次節で使用される色線形性画像特徴が RGB 色空間に関連する特徴のためである.

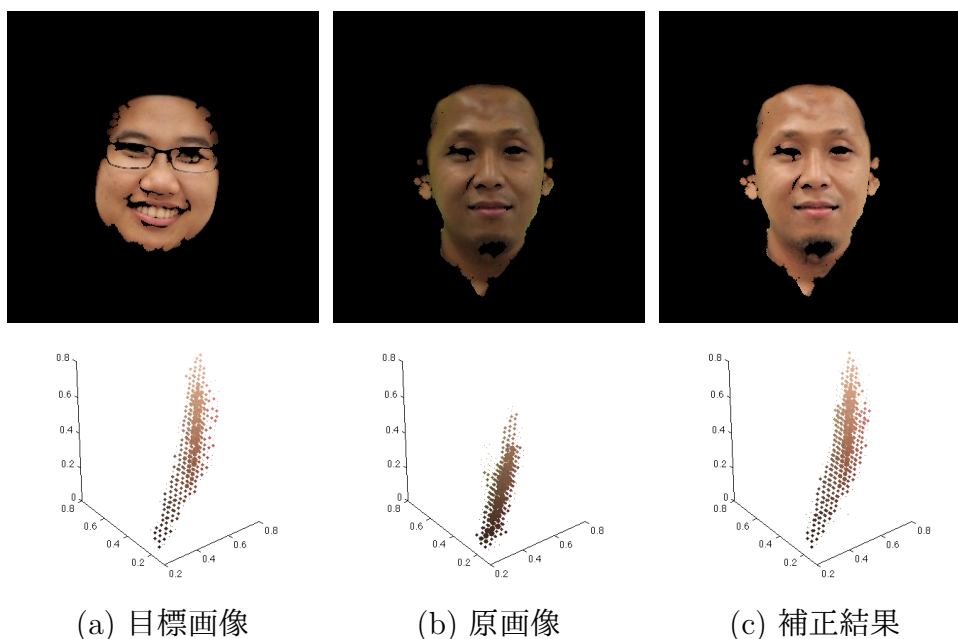


図 6.4: 抽出された顔肌色領域およびそれらの RGB 色分布. (上段の左から): 目標画像, 原画像, 補正結果. (下段): RGB 色空間における各画像の色分布. 目標画像は所望の肌色を持つ

6.2.4 最適化を用いたハイブリッド GF

本節では, 参照画像を用いた GF を最適化ベースの方法として提案し, 前節での色補正結果を入力画像に適応させる. 提案する目的関数は, 局所線形性をデータ忠実項に取り入れるるため, matting [76, 80] に類似するが, 設計目標は異なる. 色線形性は高ダイナミックレンジ撮像のレンジ圧縮手法であるトーンマッピングにおいても有効な画像特徴であることが証明されている [81]. 提案する GF は, 入力画像 Y とカラーグレーディング結果を参照画像として使用し, 顔肌色領域が矯正された色を有する画像 X を出力する. RGB 色空間 (\mathbb{R}^3) において, 推定画像の画素 j における画素値を \mathbf{x}_j , 入力画像を \mathbf{y}_j , 6.2.3 節のカラーグレーディングによって与えられる色補正された参照画像を \mathbf{g}_j とする. 図 6.5 に各画像を示す. それらを用いて, 最小化問題を次のように定式化する.

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{A}, \mathbf{b}} \sum_i \left(\sum_{j \in w_i} \|\mathbf{x}_j - \mathbf{A}_i \mathbf{y}_j - \mathbf{b}_i\|_2^2 + \varepsilon \|\mathbf{A}_i\|_F^2 \right) \\ + \iota_{\{\mathbf{g}, \eta_{skin}, \mathbf{M}_s\}}(\mathbf{x}) + \iota_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}}(\mathbf{x}) \end{aligned} \quad (6.5)$$

ここで, $\mathbf{A}_i \in \mathbb{R}^{3 \times 3}$ および $\mathbf{b}_i \in \mathbb{R}^3$ は, 各画素 i について \mathbf{y}_j から \mathbf{x}_j に近似するスケーリング行列およびオフセットベクトルであり, w_i は画素 i を囲む正方形の局所領域であり, $\iota_{(\cdot)}$ は 3.1.1 節で紹介した指示関数, $\mathbf{M}_s \in \mathbb{R}^{3N \times 3N}$ および $\mathbf{M}_b \in \mathbb{R}^{3N \times 3N}$ は顔肌色領域 Ω_{skin} と Ω_{bg} に属する画素に 1 をもちそれ以外は 0 を

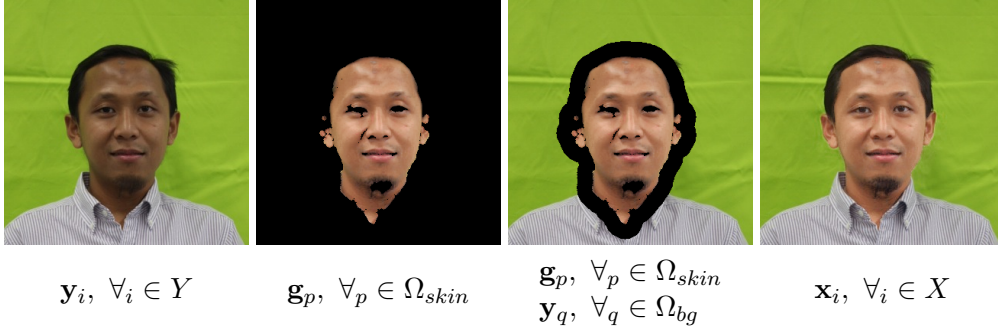


図 6.5: ハイブリッド GF の入出力. (左から): 原画像, 目標画像, 各領域の Ω_{bg} の Ω_{skin} 目標色, ハイブリッド GF 結果.

もつマスク行列である. 式 (6.5) において, 一番上の行の項は [80] に対応し, \mathbf{x} へ \mathbf{y} の局所領域のテクスチャとコントラストを反映するデータ忠実度として使用する. 第二項は顔肌色を参照画像の色に近づけ, 三番目項は入力画像の背景色を維持する働きをもつ. 各画素における誤差が η を超えないという制約を加えることで, 補助画像と元画像の差異から生じるアーティファクトを低減する. ここで, $\eta_{skin} = 5|\Omega_{skin}| \times 10^{-4}$, $\eta_{bg} = 5|\Omega_{bg}| \times 10^{-10}$ を実験において使用し, $|\Omega|$ は各領域に含まれる画素の数を表す.

目的関数 (6.5) の凸最適化問題では, MFISTA [64] を最適化アルゴリズムに採用する. 画像復元で最適化問題を扱う際に, 広く用いられている ADMM や PDS といった最適化アルゴリズムにおいても解を得られるが, 計算上効率的かつ高速な収束のため MFISTA を採用した. 式 (3.8) における MFISTA の各関数は (6.5) の第一項を $F(\cdot)$, 第二項を第三項を $G(\cdot)$ として扱う.

MFISTA の手順に必要な差分 $\nabla F(\mathbf{x})$ は [82] が提案したフィルタリングによる各画素の計算を用いるが, 文献 [82] では $[\nabla F(\mathbf{x})]_i \in \mathbb{R}^1$ に対し, 本手法では $[\nabla F(\mathbf{x})]_i \in \mathbb{R}^3$ であることに注意されたい. また, これは [80] の高速化版である. 次のように各画素の差分 $[\nabla F(\mathbf{x})]_i \in \mathbb{R}^3$ を計算する.

$$\mathbf{A}_i^* = \Delta_i^{-1} \left(\frac{1}{|w_i|} \left(\sum_{j \in w_i} \mathbf{y}_j \mathbf{x}_j \right) - \bar{\mathbf{y}}_i \bar{\mathbf{x}}_i \right) \quad (6.6)$$

$$\mathbf{b}_i^* = \bar{\mathbf{x}}_i - \mathbf{A}_i^* \bar{\mathbf{y}}_i \quad (6.7)$$

$$[\nabla F(\mathbf{x})]_i = |w_i| \mathbf{y}_i - \left(\left(\sum_{j \in w_i} \mathbf{A}_j^* \right) \mathbf{y}_i + \left(\sum_{j \in w_i} \mathbf{b}_j^* \right) \right) \quad (6.8)$$

ここで, $\Delta = \mathbf{m}\Sigma_i + \frac{\varepsilon}{|w_i|} \mathbf{U} \in \mathbb{R}^{3 \times 3}$, $\bar{\mathbf{x}}_i = [\bar{x}_i^R \ \bar{x}_i^G \ \bar{x}_i^B]^\top$ および $\bar{\mathbf{y}}_i = [\bar{y}_i^R \ \bar{y}_i^G \ \bar{y}_i^B]^\top$ は局所領域 w_i に含まれる画素の各色の平均値ベクトルであり, $|w_i|$ はウィンドウ内 w_i の画素数, $\mathbf{m}\Sigma_i$ は共分散行列, \mathbf{U} は単位行列である. MFISTA の一ステップ目より $\hat{\mathbf{z}}^k = \mathbf{v}^k - 1/L \nabla F(\mathbf{v}^k)$ を得た後, 第二項, 第三項に対する近接写像は $\hat{\mathbf{z}}^k$ を用

い、 $\Omega_{skin} \cap \Omega_{bg} = \{0\}$ であるため次式のように与えられる.

$$\mathbf{z}^k := \text{prox} \left({}^{\iota}C_{\{\mathbf{g}, \eta_{skin}, \mathbf{M}_s\}}, {}^{\iota}C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}} \right) (\hat{\mathbf{z}}^k) = P_{(C_{\{\mathbf{g}, \eta_{skin}, \mathbf{M}_s\}}, C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}})} (\hat{\mathbf{z}}^k) =$$

$$\left\{ \begin{array}{ll} \hat{\mathbf{z}}^k & \text{if } \hat{\mathbf{z}}^k \in C_{\{\mathbf{g}, \eta_{skin}, \mathbf{M}_s\}} \wedge \hat{\mathbf{z}}^k \in C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}} \\ (\mathbf{U} - \mathbf{M}_s) \hat{\mathbf{z}}^k + \mathbf{M}_s \left(\eta_{skin} \frac{\hat{\mathbf{z}}^k - \mathbf{g}}{\|\mathbf{M}_s(\hat{\mathbf{z}}^k - \mathbf{g})\|} \right) & \text{if } \hat{\mathbf{z}}^k \notin C_{\{\mathbf{g}, \eta_{skin}, \mathbf{M}_s\}} \wedge \hat{\mathbf{z}}^k \in C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}} \\ (\mathbf{U} - \mathbf{M}_b) \hat{\mathbf{z}}^k + \mathbf{M}_b \left(\eta_{bg} \frac{\hat{\mathbf{z}}^k - \mathbf{y}}{\|\mathbf{M}_b(\hat{\mathbf{z}}^k - \mathbf{y})\|} \right) & \text{if } \hat{\mathbf{z}}^k \in C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}} \wedge \hat{\mathbf{z}}^k \notin C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}} \\ (\mathbf{U} - \mathbf{M}_{sb}) \hat{\mathbf{z}}^k + \mathbf{M}_s \left(\eta_{skin} \frac{\hat{\mathbf{z}}^k - \mathbf{g}}{\|\mathbf{M}_s(\hat{\mathbf{z}}^k - \mathbf{g})\|} \right) \\ \quad + \mathbf{M}_b \left(\eta_{bg} \frac{\hat{\mathbf{z}}^k - \mathbf{y}}{\|\mathbf{M}_b(\hat{\mathbf{z}}^k - \mathbf{y})\|} \right) & \text{otherwise} \end{array} \right. \quad (6.9)$$

ここで、 $\mathbf{M}_{sb} = \mathbf{M}_s + \mathbf{M}_b$ である. $\mathbf{x}^0 =: \mathbf{v}^1 \in \mathbb{R}^N$, $t^1 := 1$ として MFISTA の各更新は以下となる.

$$\left\{ \begin{array}{l} \hat{\mathbf{z}}_i^k := \mathbf{v}_i^k - \frac{1}{L} \nabla F(\mathbf{v}^k), \\ \mathbf{z}^k := P_{(C_{\{\mathbf{g}, \eta_{skin}, \mathbf{M}_s\}}, C_{\{\mathbf{y}, \eta_{bg}, \mathbf{M}_b\}})} (\hat{\mathbf{z}}^k) \\ \mathbf{x}^k := \arg \min \{ F(\mathbf{z}) \mid \mathbf{z} \in \{\mathbf{z}^k, \mathbf{x}^{k-1}\} \}, \\ t^{k+1} := \frac{1 + \sqrt{1 + 4(t^k)^2}}{2}, \\ \mathbf{v}^{k+1} := \mathbf{x}^k + \frac{t^k}{t^{k+1}} (\mathbf{z}^k - \mathbf{x}^k) + \frac{t^k - 1}{t^{k+1}} (\mathbf{x}^k - \mathbf{x}^{k-1}), \end{array} \right. \quad (6.10)$$

6.3 実験および考察

本設では、提案手法により得られた結果を示す. 実験において、RGB 値の範囲は $[0, 1]$ で正規化し、6.2.4 節での局所領域サイズはそれぞれ 19×19 とする.

提案する補助画像フィルタリングは、影によって引き起こされる入力画像のグラデーションを平らにするため、図 6.6 の上部に示すように、しばしば不自然な画像が得られる. 図 6.6 において、入力画像はアルゴリズムの入力であり、フィルタリングされた画像は GF の結果である. 輝度補正のために、画像の各画素値 $\mathbf{x}_i^c = [x_i^R, x_i^G, x_i^B]^\top \in \mathbb{R}^3$ は、以下のように色成分 $\mathbf{x}_i^{c'} = [x_i^{R'}, x_i^{G'}, x_i^{B'}]^\top \in \mathbb{R}^3$ と強度成分 $x_i^I \in \mathbb{R}^1$ とに分解される

$$x_i^I = \frac{(x_i^R)^2}{x_i^R + x_i^G + x_i^B} + \frac{(x_i^G)^2}{x_i^R + x_i^G + x_i^B} + \frac{(x_i^B)^2}{x_i^R + x_i^G + x_i^B} \quad (6.11)$$

$$\mathbf{x}_i^{c'} = \frac{\mathbf{x}_i^c}{x_i^I}$$

この分解手順は [83] と同じである.

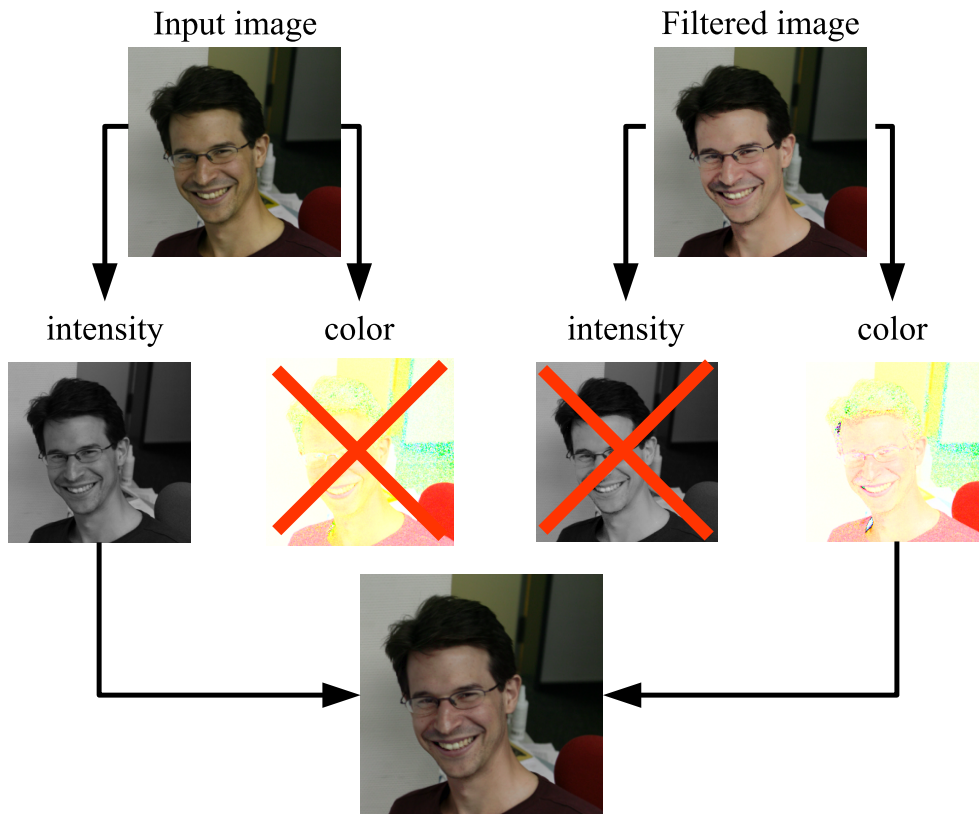


図 6.6: 輝度補正. 入力画像とフィルタリングされた画像はそれぞれ (6.11) によって分解された強度成分と色成分であり, 入力画像の強度成分とフィルタリングされた画像の色成分が (6.12) によって結合される.

入力画像 Y と GF 結果画像 X は (6.11) により二つの成分に分解され, 入力画像の強度成分 y_i^I とフィルタリング画像の色成分 $x_i^{c'}$ とが以下のように合成される

$$\mathbf{x}_i'^c = y_i^I \mathbf{x}_i^{c'} \quad (6.12)$$

図 6.6 はこの手順とその有効性を示している.

図 6.7 は, 肌色補正の結果を示す. 環境光や背景色のため, 入力画像の顔周囲の色が歪んでいるが, 結果は, 目標の顔の肌の色としてホワイトバランスのとれた顔の肌の色をしている.

暗いシーンで写真の撮影には光量を補うためにフラッシュが使用されるが, カメラのフラッシュによって不自然な顔色になることがある (図.6.8(b)). そこで, 提案手法を暗いシーンで撮影されたフラッシュ画像に適用し, 人工光の望ましくない影響を低減する. 図 6.8(c) に本手法によるフラッシュ画像補正結果を示す. 結果より, フラッシュ画像の不自然な色が自然な色に補正されていることが確認できる.

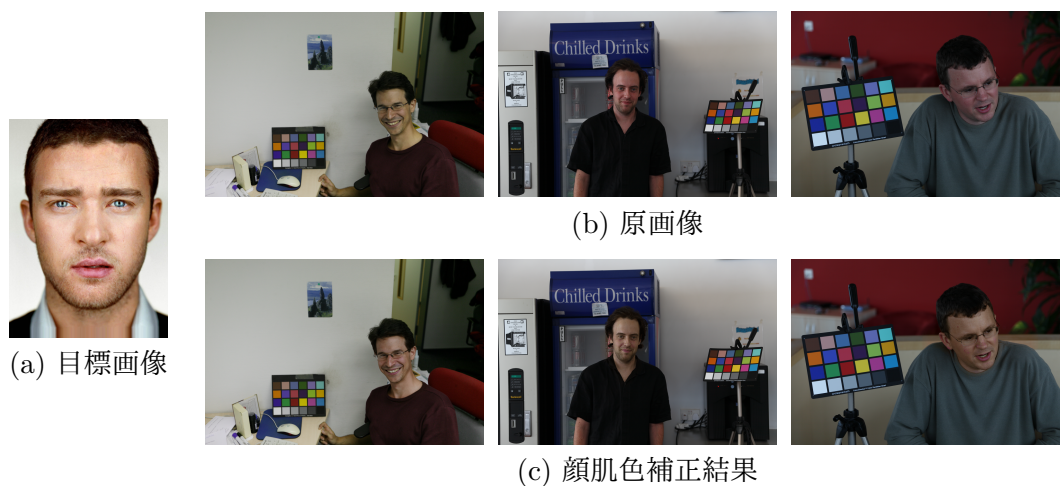


図 6.7: 顔肌色補正結果. (左から): (a) 目標画像, (b) 原画像, (c) 顔肌色補正結果. ターゲット画像はShinらのデータセット [10] であり, 入力画像はGehlerデータセット [11] を用いた.



図 6.8: フラッシュ画像上の顔の肌の色補正. (左から): (a) 目標画像, (b) フラッシュ画像, (c) 顔肌色補正結果.

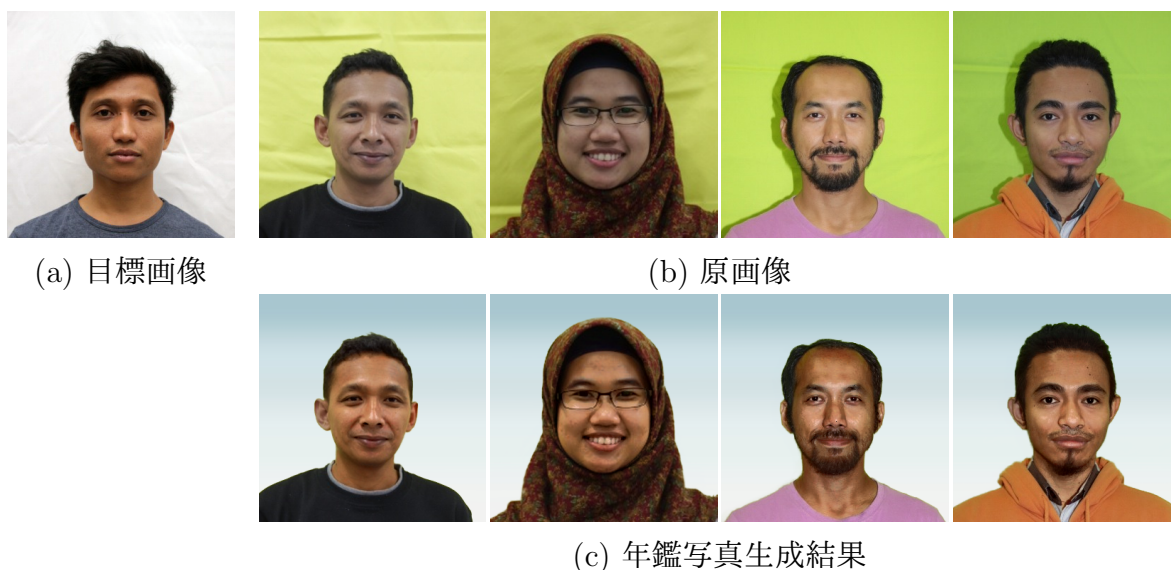


図 6.9: 提案手法による自動年鑑写真生成. (a) 目標画像, (b) 原画像, (c) 我々のアルゴリズムによる年鑑写真生成結果. 背景色を有する顔肌色が参照画像の肌色に変換されている.

6.3.1 自動年鑑写真生成

この節では、顔肌色補正法に前後処理を加えて、年鑑写真の自動写真生成法を紹介する. 年鑑写真を作る際、人手による写真補正には長い時間がかかるが、自動写真生成では実時間で処理可能である.

まず前処理として、顔検出 (図 6.10 の赤枠) を使い写真をトリミングし、顔の肌の色を修正する. 後処理として背景をクリアするため、用意した背景画像に置き換える (図 6.10 の青枠).

顔区域のトリミング

6.2.1 節において、顔区域を検出している. 検出された顔区域をトリミングし、画像のサイズを拡大縮小して統一する. トリミングサイズは、画像サイズに基づき調整する. 実験において、トリミングしリサイズされた画像サイズは 320×320 ($h \times w$) とする.

アルファブレンディングによる背景置換

画像情報の前景領域と背景領域を分離し、それらに値 $\alpha_i := [0, 1]$ を割り当てる. これらの値をラベルとして使用すると、前景 f_i 、背景 b_i (ここでの背景、前景の意図は式 (6.5) と異なる), および各画素における原画像 y_i の関係は、以下のよう

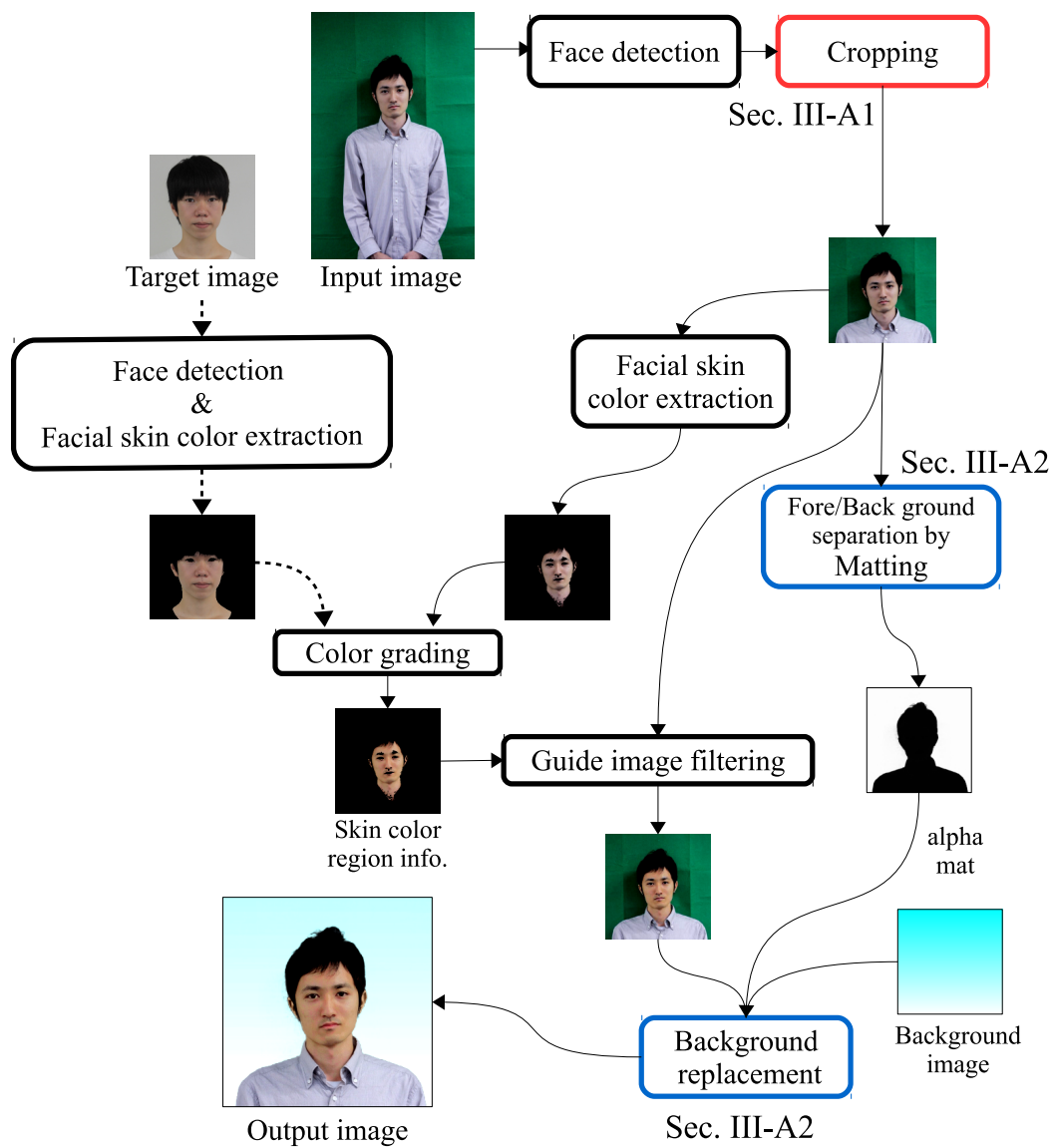


図 6.10: 自動年鑑写真生成のフローチャート. 赤枠は前処理を示し, 青枠は後処理を示す.

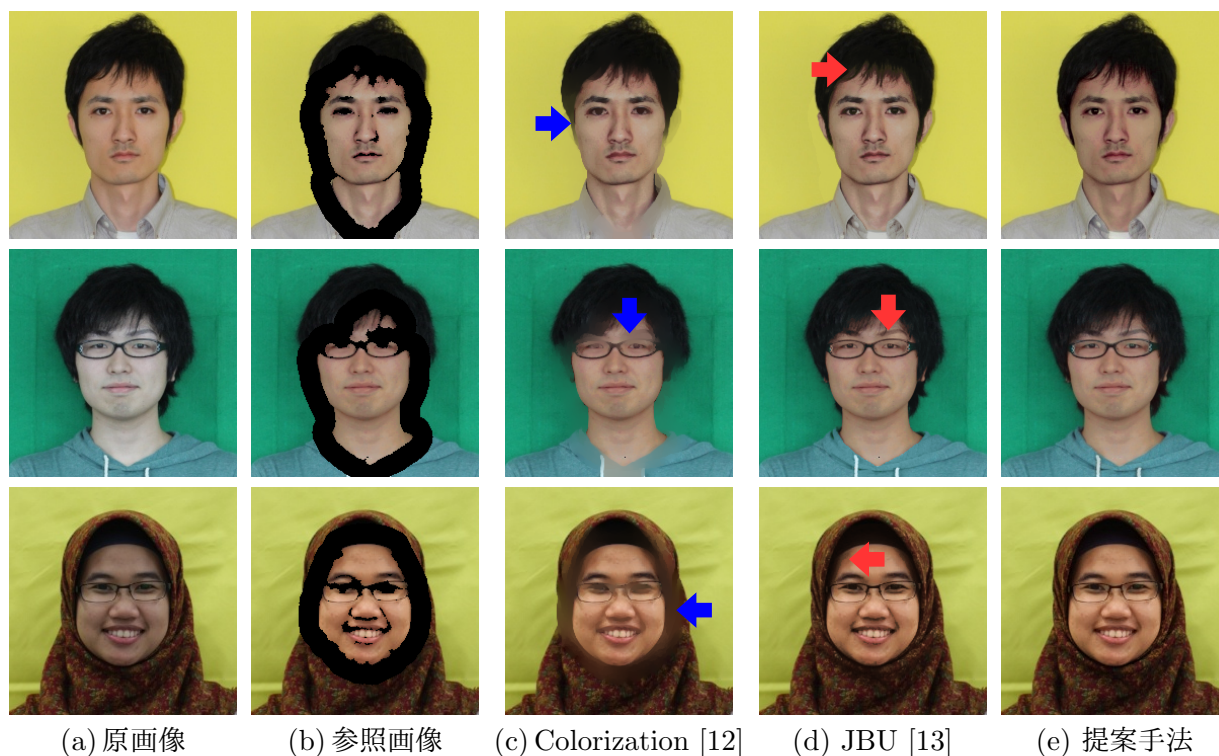


図 6.11: 欠損領域の修復における類似手法との比較. (左から): (a) 原画像, (b) 参照画像, (c) 各 RGB レイヤーへの Colorization [12], (d) 各 RGB レイヤーへの JBU [13], および (e) 提案手法 (6.5). (b) の黒い画素は欠損を表し, 各赤矢印はアーティファクトを示す.

に与えられる.

$$\mathbf{y}_i := \mathbf{f}_i + \mathbf{b}_i, \quad \mathbf{f}_i := (1 - \alpha_i)\mathbf{y}_i, \quad \mathbf{b}_i := \alpha_i\mathbf{y}_i. \quad (6.13)$$

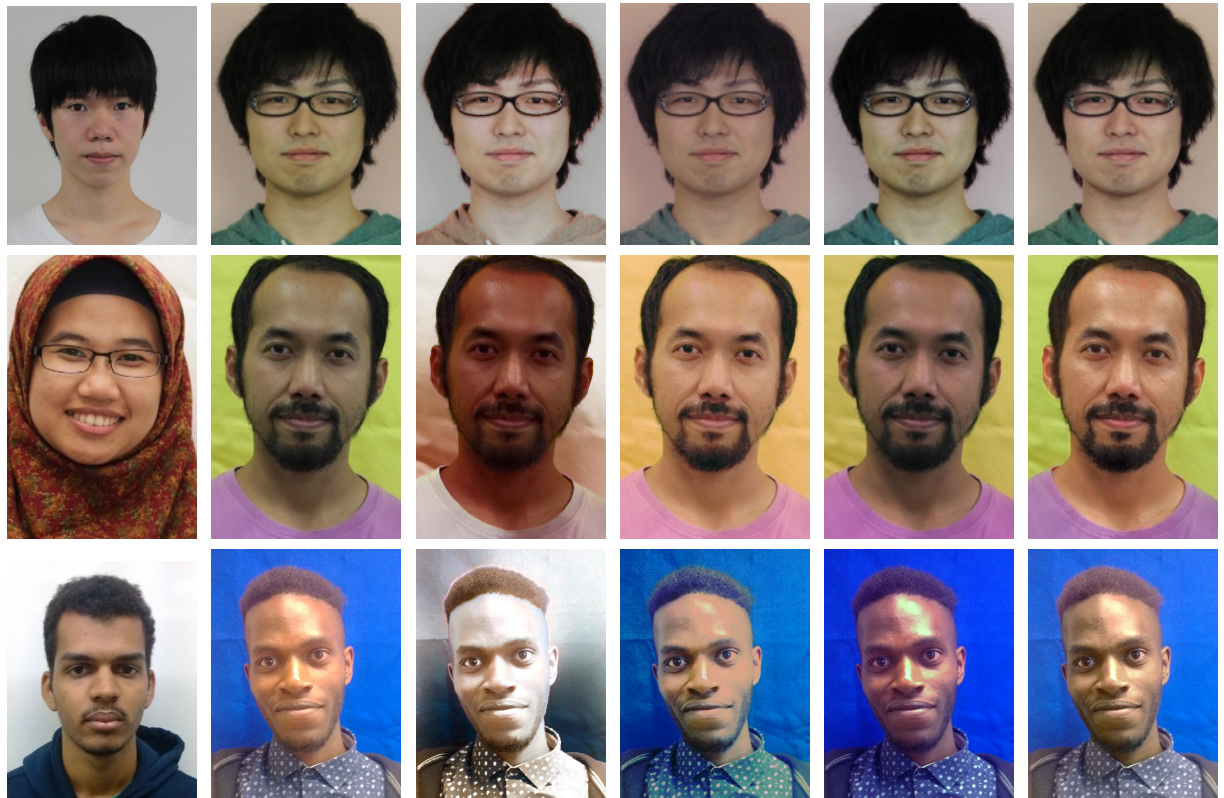
各画素のラベル α はブレンド率で, アルファマットと呼ばれる. 次式で背景を別の背景 \mathbf{z} に置き換える.

$$\mathbf{y}'_i := \mathbf{f}_i + \alpha_i\mathbf{z}_i. \quad (6.14)$$

アルファマットの推定は付録 B に記載されている.

図 6.9 は, 年鑑写真を自動生成した結果である.

MATLAB と OpenCV (C++) でアルゴリズム全体を実装し, 3.20GHz コア i5 CPU における合計実行時間は 11 秒以内である. 例として, 顔検出 (6.2.1 節) は 5 秒, 顔色抽出 (6.2.2 節) は 1 秒, カラーグレーディング (6.2.3 節) は 2 秒, GF (6.2.4 節) は 2 秒, マッチング (6.3.1 節) は 1 秒である.



(a) 目標画像 (b) 原画像 (c) [3] with [2] (d) [9] (e) [14] (f) 提案手法

図 6.12: 既存の方法との比較. (左から): (a) 目標画像, (b) 原画像, (c) [3] with [2], (d) NRDC [9], (e) Jaesik *et al.* [14], および (f) 提案手法.

6.3.2 従来手法との比較

提案するハイブリッド GF は図 6.3(d) における境界の欠損領域 Ω_{bdr} を違和感なく修復する手法とも言える. そこで, 提案手法と既存のガイド画像を用いた欠損領域の修復手法 [12, 13] を図 6.11 で比較する. 文献 [12, 13] については, YUV 色空間の UV 色成分に対して提案されているが, 本実験では RGB 色空間における各カラーチャンネルへ処理する. Colorization [12] は境界まで色を塗りつぶすだけなので, 元の画像の詳細を保持することはできない (図 6.11(c)). Joint Bilateral Upsampling (JBU) [13] は, 補助画像からバイラテラル重み係数を計算し, その係数を用いたフィルタリングによって欠損領域の画素値を算出する. 図 6.11 の上段と中段のコントラストの低下を見ることができるが, これはバイラテラルフィルタによる平滑化の結果である. 青い矢印と赤い矢印は, 各手法のアーティファクトを示している. 一方, 提案手法では, ガイド画像の色を維持しながら, 元の画像の鮮明な詳細を持つ画像が得られた.

図 6.12 は, 従来の変換手法との比較を示す. 結果 (c) は, Pitiéらの手法 [2] で入力画像の色分布を目標画像に変換し, Rabinらの手法 [3] により色変換された

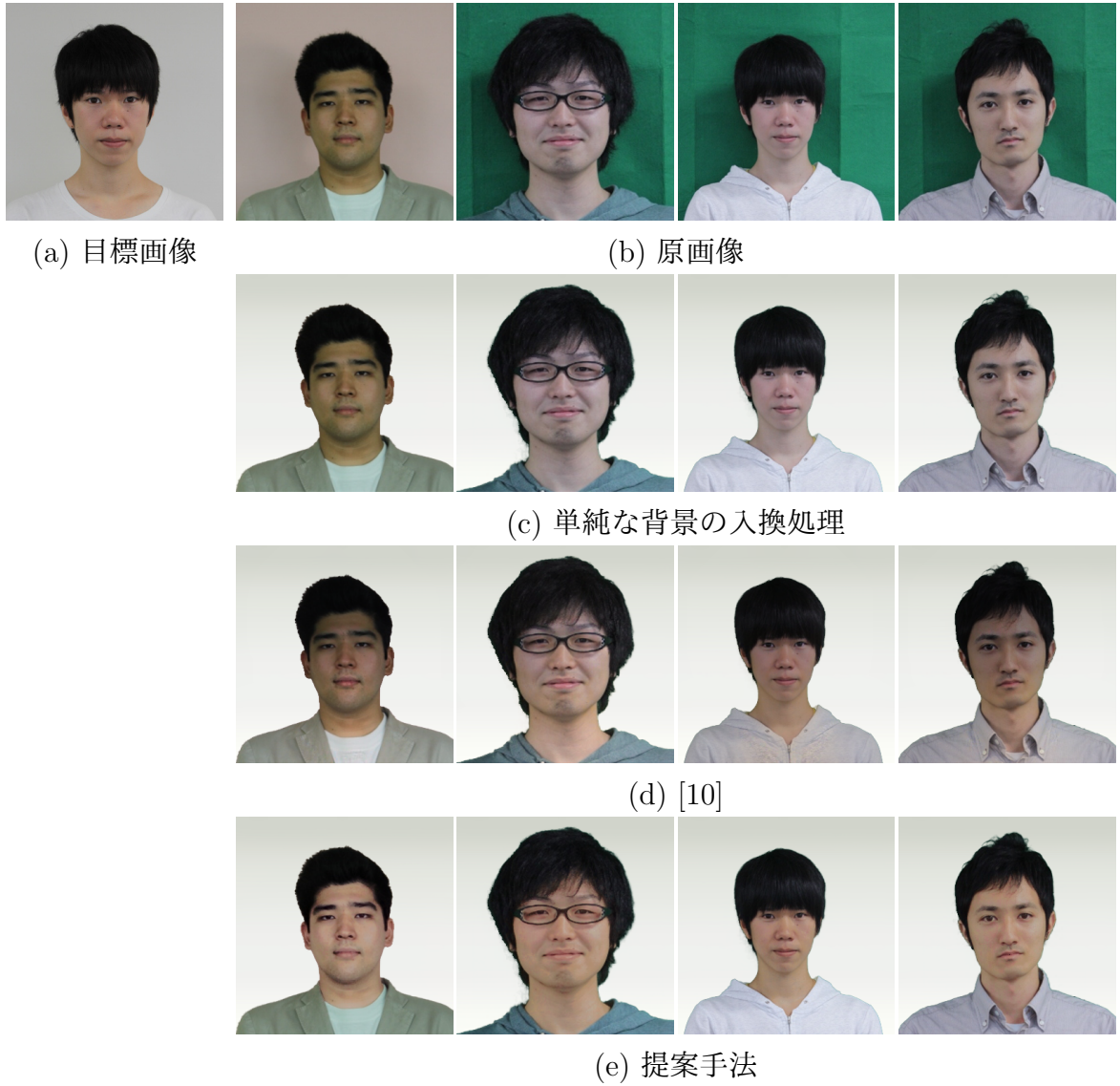


図 6.13: 背景置換結果および [10] との比較. (a) 目標画像, (b) 原画像, (c) 背景置換結果, (d) [10] によるスタイル変換結果, (e) 提案手法アルゴリズムによる年鑑写真生成.

入力画像の利得雑音アーティファクトを除去する (2.1.2 節参照). 文献 [2,3] を用いたカラーグレーディングの結果では, 均一な色変換であるために, 得られる画像の顔と背景は同じ色を有する. また, 反復的なバイラテラルフィルタリングが必要なため, 多くの処理時間を有する. NRDC [9] および Jaesik らの手法 [14] では, 顔色の改善が小さく, 服の周りの領域は変色する. 一方, 我々の方法 (e) では, 色補正が顔肌色のみで達成され, より自然に見える. 以上の結果から, 提案手法は良好な結果を得ることが確認できた.

図 6.13 は自動年鑑写真生成における, 単純な背景置換結果および [10] との比較を示す. 単純な背景置換の結果は, 顔肌色補正をしておらず, 背景色の影響で顔肌色が歪んでいる. Shin らの手法は, 入力画像のスタイルを目標画像のスタイルに変換する. 結果の顔肌色は目標画像とほぼ同じであるが, 眼鏡をかけた人の画像を対象画像の肌色に補正することはできない. 本手法は, 人が眼鏡をかけたとしても, 目標の顔肌色を持つ画像を生成出来ている.

6.3.3 半自動色補正

領域指定を手動で行うことで, 提案手法は動物の写真などの他の状況における画像の修正も可能である. 私たちの色補正方法では, 領域 Ω_{skin} を前景, 領域 Ω_{bg} を背景領域としている. 顔領域を検出するための多くの方法が提案されているが, 自然なシーンでの正確な物体検出は依然として課題である.

図 6.14(c) のような自然なシーンでの色補正では, 各領域を図 6.14(b)(d) のように手動で作成した. ここで各色は図 6.3 に対応する領域を示す. 最後に, 原画像上のオブジェクトの色を自動的に目標色に調整する. 図 6.14(e) は自然なシーンの色補正結果を示し, 赤枠は本手法が境界で違和感のない色補正が可能であることを示している. ここでは半自動手法を紹介したが, 任意の自動オブジェクト検出手法の導入により, このアプリケーションは自動アプリケーションとなる.

6.3.4 提案法における失敗例

図 6.15 は, 提案手法がうまくいかない例の画像を示す. 顔と髪の色が似ているため, VB-GMM などの高価な手法を採用しても, 顔肌色だけでなく髪の毛の領域も抽出される (図 6.15(c) 参照). その結果, 我々の方法は, 顔と髪の毛の間に類似した色の画像を出力する. 実際, 顔肌色抽出処理は, 被験者が白肌の人でなくても, 余分な領域を取得することがあり, 顔画像と抽出された余分な領域との間で同じ色が出力されてしまう.

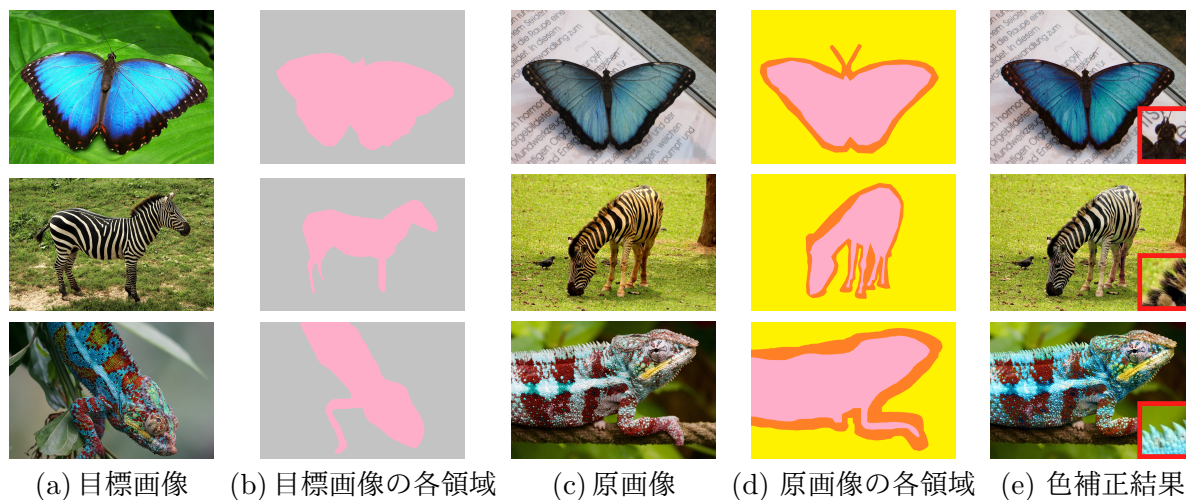


図 6.14: 半自動色補正. (a) 目標画像, (b) 目標画像の各領域, (c) 原画像, (d) 原画像の各領域, および (e) 提案法結果. (b) と (d) の色は図 1 のように各領域を示している. 原画像と目標画像は <https://pixabay.com> より入手した.

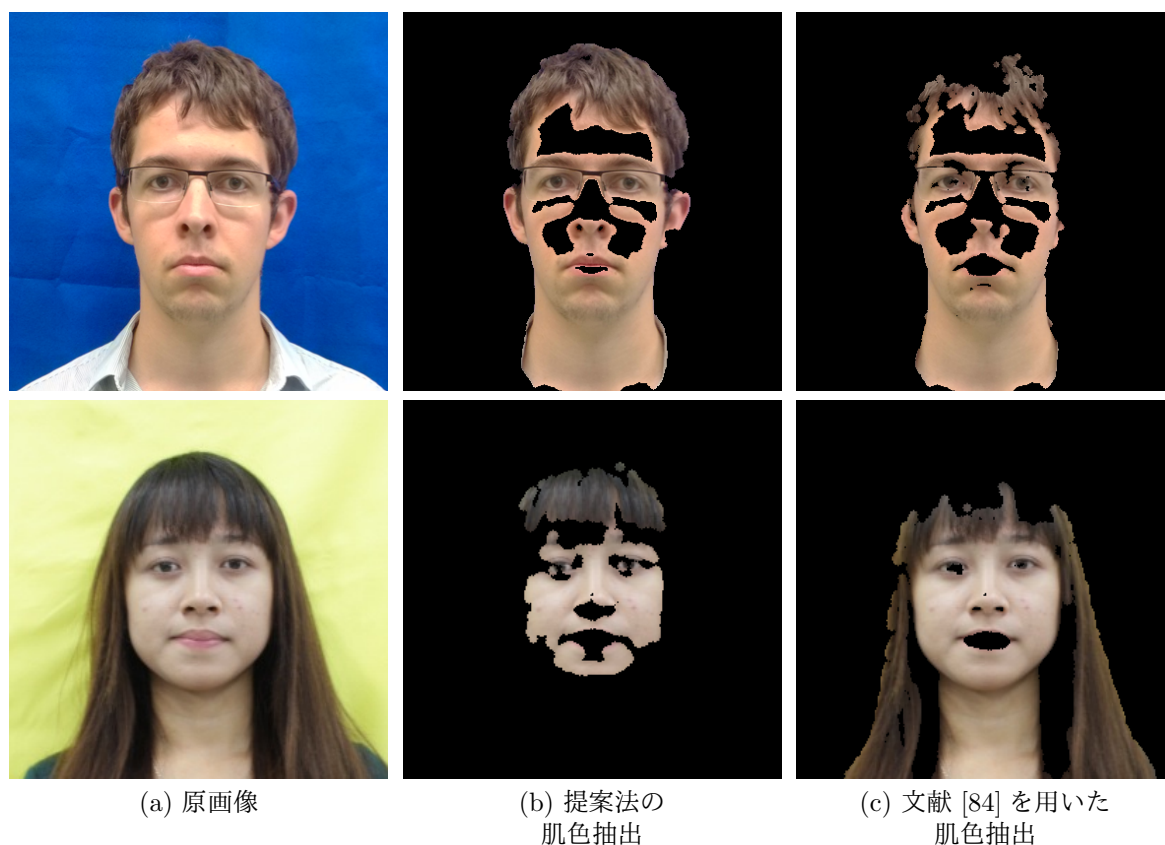


図 6.15: 皮膚色抽出段階に起因する我々の方法の悪い結果. これは画像中の顔と髪の色が類似しているが原因である.

第7章 結論

本論文では、参照画像を用いた高詳細な画像復元手法を提案した。本章ではそれらを総括し、最後に将来の展望について述べる。

7.1 凸最適化を用いたフラッシュ／高感度画像の合成

カメラでの暗所の撮影の際に、雑音で不鮮明となる高感度画像の復元のため、参照画像にフラッシュ画像を用いた合成法を試みた。従来法では、テクスチャ成分とストラクチャ成分を補正処理せずに合成するため、アーティファクトが生じる問題があげられた。提案法では凸最適化問題により整合性を保った合成を実現した。また、撮影状況によりフラッシュ画像が劣化し、従来法での合成が不完全である問題を指摘した。本手法では、フラッシュ画像の劣化領域を別の最適化問題としてモデル化した。具体的には、フラッシュ画像が劣化した領域の鮮鋭化のため、JBLFを最適化問題に組み込んだ新しい手法を提案した。更に、提案法では、二つの領域の合成画像を輝度値より算出したアルファマップで統合し、全領域で良好な画像合成を実現した。

7.2 位置ずれのある画像ペアの画像合成

位置ずれのあるフラッシュ／長時間露光画像ペアにおいてフラッシュ画像の詳細を保存する画像統合手法を提案した。従来の色変換方法は、不自然な画像を生成し、コントラスト低下が見られた。本手法では画像の構造を維持しつつ色変換を行えるため、自然かつコントラストの高い結果となった。ただし、実験において失敗する例がいくつかあった。これらは画像ペア間における影などの照度差が原因であり、フラッシュ／ノンフラッシュ画像合成において重要な課題と言える。今回は色の誤差を測る際に二乗誤差を用いたが照度差による画像間の差異をより外れ値にロバストな絶対値和による誤差、つまり l_1 を導入することで性能改善が見込められ、今後実験によりその性能差を確認したい。

7.3 参照画像を用いた人の顔の自動肌色補正

顔肌色の自動補正手法を提案した。従来のカラーグレーディング方法は、色空間における色補正方法と考えることができる。一方、提案したハイブリッドGFは、画像空間における参照画像を用いた補正方法と考えることができる。画像復元においては画像空間の情報を考慮したほうが視覚的に良い結果が得られることが確認できた。人物の顔は認識が比較的容易であるために肌色抽出に成功し、良好な色補正を行えた。ただし、肌色抽出処理は閾値処理の恩恵が大きく、よりよい判別手法を取り入れることで実験で失敗するような画像においても良好な結果が得られると考える。また、実験において人物画像ではない画像に対する補正を示した。補正結果は良好であったが画像中のオブジェクト認識や前述したように補正対象色の正確な処理は未だ困難な問題であり、解決すべき問題といえる。

7.4 提案手法の問題点と将来への展望

本論文では劣化した画像を復元する際、参照画像を用いた画像復元を試みた。従来の単一画像における画像復元では情報の欠落により復元精度が低下する問題があったが、色彩やディテールなど、劣化画像において欠落した情報を参照画像で補うことで高い精度での画像復元を実現した。更に、参照画像を扱う手法の多くはヒューリスティックな手法が多い中、提案法では問題を凸最適化問題に帰着させ十分な性能がでることを示した。参照画像を用いた画像復元は補助として類似シーンや同シーンの画像を必要とすることや、画素の一致など問題はあるものの、画像復元において有効なアプローチであることが本研究を通して伝われば幸いである。

問題設定の際、空間的な広がりを持つシーンや画像ペアの位置ずれを許容するなど、実用的な画像を対象とした画像復元を念頭に本研究は進められた。実シーンにおけるいくつかの課題を解決した本研究が産業発展に貢献できることを願い、本論文を結ぶ。

付録

付録 A: 5章における $\mathbf{V}_{(i)}^T \mathbf{V}_{(i)}$ について

$\widehat{\mathbf{M}}_{(i)}$ 局所領域 $\mathcal{N}_2(i)$ 内の画素を抽出する行列である。 $\widehat{\mathbf{G}}_{(i)}$ は局所領域内の平均値を算出する行列であるため、

$$\widehat{\mathbf{G}}_{(i)} = \frac{1}{|\mathcal{N}_2|} \mathbf{J} \widehat{\mathbf{M}}_{(i)},$$

ここで、 \mathbf{J} は $\widehat{\mathbf{M}}_{(i)}$ と同じサイズの全ての要素が 1 の行列である。 また、

$$\mathbf{V}_{(i)} = \widehat{\mathbf{M}}_{(i)} - \widehat{\mathbf{G}}_{(i)} = \left(\mathbf{U} - \frac{1}{|\mathcal{N}_2|} \mathbf{J} \right) \widehat{\mathbf{M}}_{(i)}$$

\mathbf{U} は単位行列である。 故に

$$\begin{aligned} \mathbf{V}_{(i)}^T \mathbf{V}_{(i)} &= \widehat{\mathbf{M}}_{(i)}^T \left(\mathbf{I} - \frac{1}{|\mathcal{N}_2|} \mathbf{J} \right)^T \left(\mathbf{I} - \frac{1}{|\mathcal{N}_2|} \mathbf{J} \right) \widehat{\mathbf{M}}_{(i)} \\ &= \widehat{\mathbf{M}}_{(i)}^T \left(\mathbf{I} - \frac{2}{|\mathcal{N}_2|} \mathbf{J} + \frac{1}{|\mathcal{N}_2|^2} \mathbf{J}^T \mathbf{J} \right) \widehat{\mathbf{M}}_{(i)} \\ &= \widehat{\mathbf{M}}_{(i)}^T \left(\mathbf{I} - \frac{1}{|\mathcal{N}_2|} \mathbf{J} \right) \widehat{\mathbf{M}}_{(i)} \\ &= \widehat{\mathbf{M}}_{(i)}^T \widehat{\mathbf{M}}_{(i)} - \widehat{\mathbf{M}}_{(i)}^T \widehat{\mathbf{G}}_{(i)} \\ &= \widehat{\mathbf{M}}_{(i)}^T \widehat{\mathbf{M}}_{(i)} - \widehat{\mathbf{M}}_{(i)} \widehat{\mathbf{G}}_{(i)}. \end{aligned}$$

行列 $\widehat{\mathbf{M}}_{(i)}$ が対応する画素を抽出する時、明らかに $\widehat{\mathbf{M}}_{(i)}^T \widehat{\mathbf{M}}_{(i)} - \widehat{\mathbf{M}}_{(i)} \widehat{\mathbf{G}}_{(i)} = \widehat{\mathbf{M}}_{(i)}^T \widehat{\mathbf{M}}_{(i)} - \widehat{\mathbf{M}}_{(i)} \widehat{\mathbf{G}}_{(i)}$ が成り立つため、 $\mathbf{V}_{(i)}^T \mathbf{V}_{(i)} = \mathbf{V}_{(i)}$ である。

付録 B: 6章の matting による前景背景分離

自然な合成がおこなえうアルファマットを得る際、6章で説明されており流用できるために *closed form matting* 手法 [82] を用いる。 この手法はラベルが $[0, 1]$ の実数として得られる連続最適化に基づいている。 ただし、文献 [82] の手法は前景と背景の境界付近にユーザ指定の情報を必要とする。 ユーザ支援を避けるため、matting における従来法である [85] で使用される領域拡張法を用いて繰り返し処理でアルファマットを取得する。 領域拡張の様子は図 7.1 で確認でき、以下のような手順で達成される。



図 7.1: 領域拡張による matting

- (a) 前景 Ω_F の初期情報として, 肌色領域 Ω_{skin} に加え, 顔の上の髪の毛の領域 (黒い大きな領域を大まかに選択) および顔の下にある服の領域が指定される. 背景 Ω_B として, 顔の左右の2つの矩形領域が指定される.
- (b) 前処理付き共役勾配法をもちいて matting 処理する (文献 [82] では単純な共役勾配法が用いられている). 結果として各画素の $\alpha_i \in [0, 1]$ のラベルが与えられる.
- (c) 前景および背景として0付近および1付近である画素は次の反復処理のため, (a) で生成される初期領域に次のように追加される:
 $\Omega_B^+ := \{i \mid \alpha_i \leq 0.2\}$, $\Omega_F^+ := \{i \mid 0.8 \leq \alpha_i\}$ この時,
 $\{\alpha_i := 0 \mid i \in \Omega_B \cup \Omega_B^+\}$, $\{\alpha_i := 1 \mid i \in \Omega_F \cup \Omega_F^+\}$.
- (d) ステップ (b) および (c) を数回繰り返す (本実験では4回). 局所領域のサイズは反復毎に半分に縮小される. この処理は境界付近の大きな物体 (耳など) と細かな物体 (髪の毛など) の両方に良好な結果をもたらす.
- (e) 中間に属するソフトラベルを0もしくは1へ近づけるため, シグモイド関数がアルファマットに適用される:
 $\alpha_i := (1 + \exp(-10(\alpha_i - 0.5)))^{-1}$.

参考文献

- [1] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” Proc. IEEE ICCV, pp.839–846, 1998.
- [2] F. Piti’e, A. Kokaram, and R. Dahyot, “Automated colour grading using colour distribution transfer,” Elsevier J. Comput. Vision and Image Understanding, vol.107, no.1–2, pp.123–137, 2007.
- [3] J. Rabin, J. Delon, and Y. Gousseau, “Regularization of transportation maps for color and contrast transfer,” Proc. IEEE ICIP, pp.1933–1936, 2010.
- [4] A.L. Dani, D. Lischinski, and Y. Weiss, “Colorization using optimization,” ACM T. Graph., vol.23, pp.689–694, 2004.
- [5] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, “Digital photography with flash and no-flash image pairs,” ACM T. Graph. (Proc. SIGGRAPH), vol.23, no.3, pp.664–672, 2004.
- [6] K. Shirai, M. Ikehara, and M. Okamoto, “Noiseless no-flash photo creation by color transform of flash image,” Proc. IEEE ICIP, pp.3437–3440, 2011.
- [7] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3d transform-domain collaborative filtering,” IEEE T. Image Process., vol.16, no.8, pp.2080–2095, 2007.
- [8] P. Blomgren and T.F. Chan, “Color TV: Total variation methods for restoration of vector valued images,” IEEE Trans. Image Processing, vol.7, pp.304–309, 1996.
- [9] Y. HaCohen, E. Shechtman, D.B. Goldman, and D. Lischinski, “Non-rigid dense correspondence with applications for image enhancement,” ACM T. Graph. (Proc. SIGGRAPH), vol.30, no.4, pp.70:1–70:10, 2011.
- [10] Y. Shih, S. Paris, C. Barnes, W.T. Freeman, and F. Durand, “Style transfer for headshot portraits,” ACM Trans. Graph., vol.33, no.4, pp.148:1–148:14, July 2014.

- [11] A.B.T.M.T.S. Peter Vincent Gehler, Carsten Rother, “Bayesian color constancy revisited,” Proc. IEEE Conf. CVPR, pp.1–8, 2008.
- [12] A. Levin, D. Lischinski, and Y. Weiss, “Colorization using optimization,” ACM TOG, vol.23, no.3, pp.689–694, 2004.
- [13] J. Kopf, M.F. Cohen, D. Lischinski, and M. Uyttendaele, “Joint bilateral upsampling,” ACM TOG, vol.26, no.3, 2007.
- [14] S.S.I.S.K. Jaesik Park, Yu-Wing Tai, “Efficient and robust color consistency for community photo collections,” Proc. IEEE Conf. CVPR, p.1, June 2016.
- [15] A. Buades, B. Coll, and J.M. Morel, “A review of image denoising algorithms, with a new one,” SIAM Multiscale Modeling & Simulation, vol.4, no.2, pp.490–530, 2005.
- [16] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” IEEE T. Image Process., vol.15, no.12, pp.3736–3745, 2006.
- [17] A. Buades and B. Coll, “A non-local algorithm for image denoising,” In CVPR, pp.60–65, 2005.
- [18] H. Talebi and P. Milanfar, “Global image denoising,” IEEE Transactions on Image Processing, vol.23, no.2, pp.755–768, Feb 2014.
- [19] C. Knaus and M. Zwicker, “Dual-domain image denoising,” 2013 IEEE International Conference on Image Processing, pp.440–444, Sept 2013.
- [20] C. Knaus and M. Zwicker, “Progressive image denoising,” IEEE Transactions on Image Processing, vol.23, no.7, pp.3114–3125, July 2014.
- [21] S. Ono and I. Yamada, “Decorrelated vectorial total variation,” 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp.4090–4097, IEEE, 2014.
- [22] P.C. Hansen, J.G. Nagy, and D.P. O’leary, Deblurring images: matrices, spectra, and filtering, SIAM, 2006.
- [23] D. Krishnan and R. Fergus, “Fast image deconvolution using hyper-laplacian priors,” in Advances in Neural Information Processing Systems 22, ed. Y. Bengio, D. Schuurmans, J.D. Lafferty, C.K.I. Williams, and A. Culotta, pp.1033–1041, Curran Associates, Inc., 2009.

- [24] Q. Shan, J. Jia, and A. Agarwala, “High-quality motion deblurring from a single image,” *ACM Trans. Graph.*, vol.27, no.3, pp.73:1–73:10, Aug. 2008.
- [25] L. Xu and J. Jia, “Two-phase kernel estimation for robust motion deblurring,” *Proceedings of the 11th European Conference on Computer Vision: Part I, ECCV’10, Berlin, Heidelberg*, pp.157–170, Springer-Verlag, 2010.
- [26] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, “Understanding blind deconvolution algorithms,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.33, no.12, pp.2354–2367, Dec. 2011.
- [27] T. Yue, S. Cho, J. Wang, and Q. Dai, “Hybrid image deblurring by fusing edge and power spectrum information.,” *ECCV (7)*, ed. D.J. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, *Lecture Notes in Computer Science*, vol.8695, pp.79–93, Springer, 2014.
- [28] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless, “Single image deblurring using motion density functions,” *ECCV ’10: Proceedings of the 10th European Conference on Computer Vision*, 2010.
- [29] Y.W. Tai, P. Tan, and M.S. Brown, “Richardson-lucy deblurring for scenes under a projective motion path.,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.33, no.8, pp.1603–1618, 2011.
- [30] M. Hirsch, C.J. Schuler, S. Harmeling, and B. Scholkopf, “Fast removal of non-uniform camera shake,” *Proceedings of the 2011 International Conference on Computer Vision, ICCV ’11, Washington, DC, USA*, pp.463–470, IEEE Computer Society, 2011.
- [31] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, “Non-uniform deblurring for shaken images,” *International Journal of Computer Vision*, vol.98, no.2, pp.168–186, 2012.
- [32] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, “Color transfer between images,” *IEEE Comput. Graph. Appl.*, vol.21, no.5, pp.34–41, 2001.
- [33] F. Pitié, A.C. Kokaram, and R. Dahyot, “N-dimensional probability density function transfer and its application to colour transfer,” *Proc. IEEE ICCV*, pp.1434–1439, 2005.
- [34] Y. HaCohen, E. Shechtman, D.B. Goldman, and D. Lischinski, “Non-rigid dense correspondence with applications for image enhancement,” *ACM TOG*, vol.30, no.4, pp.70:1–70:10, 2011.

- [35] Y. HaCohen, E. Shechtman, D.B. Goldman, and D. Lischinski, “Optimizing color consistency in photo collections,” *ACM TOG*, vol.32, no.4, pp.38:1–38:10, 2013.
- [36] Y. Hwang, J.Y. Lee, I.S. Kweon, and S.J. Kim, “Color transfer using probabilistic moving least squares,” *Proc. IEEE Conf. CVPR*, pp.3342–3349, 2014.
- [37] R. Nguyen, S. Kim, and M. Brown, “Illuminant aware gamut-based color transfer,” *Comput. Graphics Forum*, vol.33, no.7, pp.319–328, 2014.
- [38] V. Caselles, A. Chambolle, D. Cremers, M. Novaga, and T. Pock, “An introduction to total variation for image analysis,” *Theoretical Foundations and Numerical Methods for Sparse Recovery*, De Gruyter, Radon Series Comp. Appl. Math., vol.9, pp.263–340, 2010.
- [39] X. Zhang, M. Burger, X. Bresson, and S. Osher, “Bregmanized nonlocal regularization for deconvolution and sparse reconstruction,” *SIAM J. Imaging Sciences*, vol.3, no.3, pp.253–276, 2010.
- [40] S. Ono, T. Miyata, I. Yamada, and K. Yamaoka, “Image recovery by decomposition with component-wise regularization,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol.95, no.12, pp.2470–2478, 2012.
- [41] S. Ono, T. Miyata, and I. Yamada, “Cartoon-texture image decomposition using blockwise low-rank texture characterization,” *Image Processing, IEEE Transactions on*, vol.23, no.3, pp.1128–1142, 2014.
- [42] S. Ono and I. Yamada, “Color-line regularization for color artifact removal,” *IEEE Transactions on Computational Imaging*, vol.2, no.3, pp.204–217, 2016.
- [43] S. Lefkimmiatis, A. Roussos, P. Maragos, and M. Unser, “Structure tensor total variation,” *SIAM Journal on Imaging Sciences*, vol.8, no.2, pp.1090–1122, 2015.
- [44] S. Ono, K. Shirai, and M. Okuda, “Vectorial total variation based on arranged structure tensor for multichannel image restoration,” *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.4528–4532, IEEE, 2016.
- [45] E. Eisemann and F. Durand, “Flash photography enhancement via intrinsic relighting,” *ACM T. Graph. (Proc. SIGGRAPH)*, vol.23, pp.673–678, 2004.

- [46] K. He, J. Sun, and X. Tang, “Guided image filtering,” *Proc. ECCV*, pp.1–14, 2010.
- [47] S. Zhuo, D.G. 0001, and T. Sim, “Robust flash deblurring,” *Proc. IEEE Conf. CVPR*, pp.2440–2447, 2010.
- [48] H.J. Seo and P. Milanfar, “Robust flash denoising/deblurring by iterative guided filtering,” *EURASIP J. Adv. Sig. Process.*, vol.2012, no.3, pp.1–19, 2012.
- [49] J. Jia, J. Sun, C.K. Tang, and H.Y. Shum, “Bayesian correction of image intensity with spatial consideration,” *ECCV*, pp.III: 342–354, 2004.
- [50] R. Irony, D. Cohen-Or, and D. Lischinski, “Colorization by example,” *Proc. Eurographics Conf. Rendering Techniques*, pp.201–210, 2005.
- [51] R.K. Gupta, A.Y.S. Chia, D. Rajan, E.S. Ng, and H. Zhiyong, “Image colorization using similar images,” *Proc. ACM Int. Conf. Multimedia*, pp.369–378, 2012.
- [52] I. Omer and M. Werman, “Color lines: Image specific color representation,” *Proc. IEEE CVPR*, pp.946–953, 2004.
- [53] L. Yuan, J. Sun, L. Quan, and H.Y. Shum, “Image deblurring with blurred/noisy image pairs,” *ACM SIGGRAPH 2007 Papers, SIGGRAPH ’07*, ACM, 2007.
- [54] C. Liu, J. Yuen, A. Torralba, J. Sivic, and W.T. Freeman, “Sift flow: Dense correspondence across different scenes,” *Proc. ECCV*, pp.28–42, 2008.
- [55] S. Gould and Y. Zhang, “Patchmatchgraph: Building a graph of dense patch correspondences for label transfer,” *Springer LNCS (Proc. ECCV)*, pp.439–452, 2012.
- [56] C. Barnes, E. Shechtman, D.B. Goldman, and A. Finkelstein, “The generalized PatchMatch correspondence algorithm,” *Springer LNCS (Proc. ECCV)*, pp.29–43, 2010.
- [57] X. Shen, L. Xu, Q. Zhang, and J. Jia, “Multi-modal and multi-spectral registration for natural images,” in *Springer LNCV (Proc. ECCV)*, pp.309–324, 2014.
- [58] H.R. Su and S.H. Lai, “Non-rigid registration of images with geometric and photometric deformation by using local affine fourier-moment matching,” *Proc. IEEE Conf. CVPR*, pp.2874–2882, 2015.

- [59] D.G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vision*, vol.60, no.2, pp.91–110, 2004.
- [60] W.Y. Lin, L. Liu, Y. Matsushita, K.L. Low, and S. Liu, “Aligning images in the wild,” *Proc. IEEE Conf. CVPR*, pp.1–8, 2012.
- [61] C. Barnes, E. Shechtman, A. Finkelstein, and D.B. Goldman, “PatchMatch: A randomized correspondence algorithm for structural image editing,” *ACM T. Graph. (Proc. SIGGRAPH)*, vol.28, no.3, pp.24:1–24:11.
- [62] J.J. Moreau, “Fonctions convexes duales et points proximaux dans un espace hilbertien.,” *Comptes Rendus de l’Académie des Sciences (Paris), Série A*, vol.255, pp.2897–2899, 1962.
- [63] P.L. Combettes and J.C. Pesquet, “Proximal Splitting Methods in Signal Processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, ed. R.C.P.E.V.L.D.W.H.E. Bauschke, H.H.; Burachik, pp.185–212, Springer, 2011.
- [64] A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems,” *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2009.
- [65] J. Eckstein and D.P. Bertsekas, “On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol.55, pp.293–318, 1992.
- [66] L. Condat, “A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms,” *J. Optim. Theory Appl.*, vol.158, no.2, pp.460–479, 2013.
- [67] M.R. Hestenes and E. Stiefel, “Methods of conjugate gradients for solving linear systems,” *J. of research of the National Bureau of Standards*, vol.49, pp.409–436, 1952.
- [68] K. He, J. Sun, and X. Tang, “Guided image filtering,” *IEEE TPAMI*, vol.35, no.6, pp.1397–1409, 2013.
- [69] T. Mertens, J. Kautz, and F.V. Reeth, “Exposure fusion,” *Proc. Pacific Conf. Comput. Graph. and Applications*, pp.382–390, 2007.
- [70] A. Agrawal, R. Raskar, S.K. Nayar, and Y. Li, “Removing photography artifacts using gradient projection and flash-exposure sampling,” *ACM T. Graph. (Proc. SIGGRAPH)*, vol.24, pp.828–835, 2005.

- [71] E.S.L. Gastal and M.M. Oliveira, “Domain transform for edge-aware image and video processing,” *ACM T. Graph. (Proc. SIGGRAPH)*, vol.30, no.4, pp.69:1–69:12, 2011.
- [72] S. Ono and I. Yamada, “A convex regularizer for reducing color artifact in color image recovery,” *Proc. IEEE CVPR*, pp.1775–1781, 2013.
- [73] T. Baba, R. Matsuoka, S. Ono, K. Shirai, and M. Okuda, “Flash/no-flash image integration using convex optimization,” *Proc. IEEE ICASSP*, pp.1185–1189, 2014.
- [74] R. Matsuoka, T. Baba, M. Okuda, and K. Shirai, “High dynamic range image acquisition using flash image,” *Proc. IEEE ICASSP*, pp.1612–1616, 2013.
- [75] R. Matsuoka, T. Yamauchi, T. Baba, and M. Okuda, “Weight optimization for multiple image integration,” *Proc. IEEE ICIP*, pp.795–799, 2013.
- [76] A. Levin, D. Lischinski, and Y. Weiss, “A closed-form solution to natural image matting,” *IEEE T. PAMI*, vol.30, no.2, pp.228–242, 2008.
- [77] F.C. Crow, “Summed-area tables for texture mapping,” *Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ’84*, New York, NY, USA, pp.207–212, ACM, 1984.
- [78] N. Batool and R. Chellappa, “Detection and inpainting of facial wrinkles using texture orientation fields and Markov random field modeling,” *IEEE TIP*, vol.23, no.9, pp.3773–3788, 2014.
- [79] S.P. Lloyd, “Least squares quantization in PCM,” *IEEE T. Info. Theory*, vol.28, no.2, pp.129–137, 1982.
- [80] A. Levin, D. Lischinski, and Y. Weiss, “A closed-form solution to natural image matting,” *IEEE TPAMI*, vol.30, no.2, pp.228–242, 2008.
- [81] Q. Shan, J. Jia, and M.S. Brown, “Globally optimized linear windowed tone mapping,” *IEEE T. Visualization and CG*, vol.16, no.4, pp.663–675, 2010.
- [82] K. He, J. Sun, and X. Tang, “Fast matting using large kernel matting Laplacian matrices,” *Proc. IEEE Conf. CVPR*, pp.2165–2172, 2010.
- [83] E. Eisemann and F. Durand, “Flash photography enhancement via intrinsic relighting,” *ACM Trans. Graph.*, vol.23, no.3, pp.673–678, Aug. 2004.
- [84] C.M. Bishop, *Pattern recognition and machine learning*, Springer-Verlag New York, Inc., 2006.

- [85] J. Sun, J. Jia, C.K. Tang, and H.Y. Shum, “Poisson matting,” ACM TOG, vol.23, no.3, pp.315–321, 2004.

謝辞

本研究は、著者が北九州市立大学大学院 国際環境工学研究科 博士後期課程に在籍中に行ったものであり、本研究を進めるにあたり、ご指導、ご助言を厚く賜った指導教授及び本論文の主査である北九州市立大学 国際環境工学部 奥田 正浩 教授に深謝の意を表します。

ご多忙の中、本論文の副査を快く引き受けてくださり、ご丁寧なご助言を賜りました北九州市立大学 国際環境工学部 佐藤 雅之教授、孫 連明教授、松永 良一教授に深謝申し上げます。

更に、公私にわたり大変有益な御指導、ご助言に御尽力していただいた北九州市立大学 国際環境工学部 京地 清介 准教授、豊橋技術科学大学大学院 情報・知能工学系 神納 貴生 助教、香川大学 工学部 松岡 諒 助教、信州大学 工学部 白井 啓一郎 助教、東京工業大学 像情報工学研究所 小野 峻佑 助教、そして、北九州市立大学大学院 国際環境工学研究科 博士後期課程 Mia Rizkinia 氏を始めとする奥田研究室の全ての諸兄に厚く御礼申し上げます。

2017年3月