

画像再構成問題に関する研究

– 超解像とぶれ除去 –

2018 年度

藤沢 貴典

学位論文 博士(工学)

画像再構成問題に関する研究
– 超解像とぶれ除去 –

2018 年度

慶應義塾大学大学院理工学研究科

藤沢 貴典

目次

第 1 章	序論	1
1.1	研究背景	1
1.2	既存研究	3
1.2.1	画像の劣化過程の表現	3
1.2.2	画像の超解像問題	3
1.2.3	画像のぶれ除去問題	6
1.3	研究目的	8
1.4	本論文の構成	10
第 2 章	基礎理論	11
2.1	本章の構成	11
2.2	画像の劣化過程の数学表現	11
2.2.1	画像の補間・超解像問題	12
2.2.2	ノイズの付加	12
2.2.3	ぶれに基づく画像の劣化過程	14
2.3	画像再構成問題	17
2.3.1	非線形問題による画像の再構成問題	17
2.3.2	非線形最適化問題によく用いられる制約条件	18
2.3.3	スパース表現に基づいた画像の再構成問題	20
2.3.4	確率理論に基づいた画像の再構成問題	22
2.4	最適化問題	23
2.4.1	最急降下法	23
2.4.2	近接勾配法	23
2.4.3	共役勾配法による高速な解法	24
2.5	従来の辞書ベース超解像手法	25

2.5.1	辞書ベースの超解像	26
2.5.2	特徴抽出空間による近似	27
2.5.3	L1 最適化による高解像度画像の再構成	28
2.5.4	辞書ベース超解像の流れ	28
2.5.5	グローバル再構成問題	30
2.5.6	事前辞書の学習	31
2.5.7	カラー画像に対する処理	32
2.6	従来のぶれ除去手法	33
2.6.1	Non-blind Deconvolution と Blind Deconvolution	33
2.6.2	カーネル推定問題のための画像の前処理	35
2.6.3	カーネル推定問題	36
2.6.4	逆畳み込み問題	37
2.7	Blind Deconvolution のための他のアプローチ	39
2.7.1	Fergus の手法	39
2.7.2	Levin の手法	39
2.7.3	Xu および Jia の手法	40
第 3 章	画像の超解像問題	41
3.1	本章の構成	41
3.2	従来手法の問題点と提案手法の改良点	41
3.3	L2 再構成問題による高速な再構成	43
3.4	自画像からのサンプリングによる辞書作成	43
3.5	実験と評価	45
3.5.1	2 倍拡大の PSNR 比較	47
3.5.2	実行時間の比較	47
3.5.3	画像サイズ・辞書サイズによる実行時間の変化	47
3.5.4	画像の視覚評価	51
3.6	まとめと今後の展望	51
第 4 章	画像のぶれ除去問題	57
4.1	本章の構成	57
4.2	従来手法の問題点と提案手法の改良点	57
4.3	特徴抽出を用いたカーネル推定	58

4.3.1	カーネル推定のための前処理の改善	58
4.3.2	特徴抽出フィルタを用いたカーネル推定処理	59
4.3.3	特徴抽出フィルタの設計	60
4.4	差分逆畳み込み	62
4.4.1	差分逆畳み込みのプロセス	62
4.4.2	周波数軸における共役勾配法を用いた解法	63
4.4.3	冗長な DFT 計算の削減による高速化	66
4.4.4	ガイド画像の生成	68
4.5	提案する Blind Deconvolution のフレームワーク	69
4.6	実験と評価	69
4.6.1	実験の設定	69
4.6.2	Non-blind deconvolution 手法との比較	74
4.6.3	Blind deconvolution 手法との比較	81
4.6.4	画像の視覚評価	83
4.6.5	実行時間の評価	88
4.7	まとめと今後の展望	88
第 5 章 結論		90
参考文献		93

囮 目次

1.1	画像情報の取得とその過程で入る劣化	2
1.2	画像の高解像度化のためのアプローチの分類	4
2.1	低解像度画像と高解像度画像の関係の数学表現	13
2.2	高解像度から低解像度へのサンプリング $s = 2$	13
2.3	ガウスノイズとインパルスノイズの例	14
2.4	カーネル行列によるモーションぶれ・焦点ボケの表現	15
2.5	画像とカーネルの畳み込み行列表現	16
2.6	空間軸と周波数軸における畳み込みの関係	17
2.7	スペースな行列とグループスペースな行列	19
2.8	辞書行列中の高解像度パッチ・低解像度パッチの対応	26
2.9	辞書ベース超解像の流れ	29
2.10	逆投影法の処理プロセス	30
2.11	学習用画像を用いた低解像度・高解像度辞書の学習	32
2.12	カラー画像に対する辞書ベース超解像処理	33
2.13	典型的な Blind Deconvolution のフレームワーク	35
2.14	従来法のカーネル推定のための画像の前処理	36
2.15	従来法の画像のしきい値処理	37
3.1	画像の自己相似性	42
3.2	L_1 と L_2 最小化によって得られる係数 α の違い	44
3.3	自画像からのパッチサンプルによる辞書作成	45
3.4	実験に用いる標準画像 24 枚	46
3.5	画像サイズと辞書サイズによる実行時間の変化	48
3.6	辞書サイズ 1024 時の実行時間の詳細	48

3.7	画像 Airplane の処理結果	53
3.8	画像 Houses の処理結果	54
3.9	画像 Tank の処理結果	55
3.10	画像 Room の処理結果	56
4.1	提案法の画像の前処理	59
4.2	カットオフ周波数 f_ω のハイパスフィルタ	61
4.3	提案法で用いる特徴抽出フィルタの設計	62
4.4	差分逆畳み込みによるリングングの抑制	63
4.5	空間軸処理と周波数軸処理のコスト比較	67
4.6	ΔY 画像端の不連続性を解消する方法	68
4.7	提案手法のカーネル推定・逆畳み込みのフレームワーク	71
4.8	実験で用いる 40 画像	72
4.9	実験で用いるサイズ 11×11 のカーネル 8 種	73
4.10	実験で用いるサイズ 35×35 のカーネル 5 種	73
4.11	画像 Bananas, サイズ 11 のカーネルの Non-blind 補正結果	77
4.12	画像 Snails, サイズ 11 のカーネルの Non-blind 補正結果	78
4.13	画像 Cards, サイズ 11 のカーネルの Non-blind 補正結果	79
4.14	画像 Multimeter, サイズ 35 のカーネルの Non-blind 補正結果	80
4.15	画像 Bananas の推定カーネルと補正結果	84
4.16	画像 Garden table の推定カーネルと補正結果	85
4.17	画像 Almonds の推定カーネルと補正結果	86
4.18	画像 Tools の推定カーネルと補正結果	87

表目次

1.1	劣化行列と加法性ノイズによる劣化過程と再構成手法	3
1.2	超解像手法の画像枚数, 速度, 事前情報の必要性における比較	9
3.1	2倍拡大画像の PSNR 比較	49
3.2	2倍拡大の実行時間比較	50
4.1	サイズ 11×11 のカーネルの逆畳み込み精度比較	75
4.2	サイズ 35×35 のカーネルの逆畳み込み精度比較	76
4.3	サイズ 11×11 カーネルの Blind Deconvolution 精度比較	82
4.4	カーネルごとの Blind Deconvolution 精度比較	83
4.5	Blind Deconvolution 実行時間の比較	88

第 1 章

序 論

1.1 研究背景

画像処理技術とは、人間が目で感じる視覚情報を画像情報として捉え、コンピュータ等の計算機器で処理する技術である。画像処理技術が必要な分野は、個人の PC、スマートフォンやタブレット PC 上で扱う写真・動画情報、ビジネス上でも産業や、テレビ等の通信やネットワークなどの多岐に渡っており、日常生活には欠かせないものとなっている。

近年、ハードウェアでの計算処理技術が大幅に進歩し、画像処理技術は目覚ましい発展をとげている。これは、ハードウェアの処理技術が高まることによって、大容量の画像を処理可能になったこと、処理の高速化に伴うリアルタイム処理化の幅が広まったこと、スマートフォンに代表される携帯機器にも搭載可能なより小型で高速なハードウェアが出現していることなどが理由である。画像処理技術の発展に伴い、4K テレビに代表されるような高解像度の画像処理技術や、モバイルデバイス上の画像処理の可能性が広まった。

画像情報に対するユーザーの価値観にも変化が現れている。それは、スマートフォン等の携帯撮像機器の普及によって、誰でも気軽に写真を取れるようになったこと、www や SNS 上にこれらの写真が数多く投稿され、第三者の取得した画像を簡単に閲覧できるようになったことが理由にあげられる。

画像情報は撮像機器によって、被写体からの視覚情報を撮影することで取得される。人間にとての視覚情報とは、光源からの光を物体が反射し、その光が人間の視神経に捉えられることで認識されるものである。その光の量が多いほど明るく、少ないほど暗く知覚される。デジタル画像はこの仕組みを模したものであり、格子状に配置された

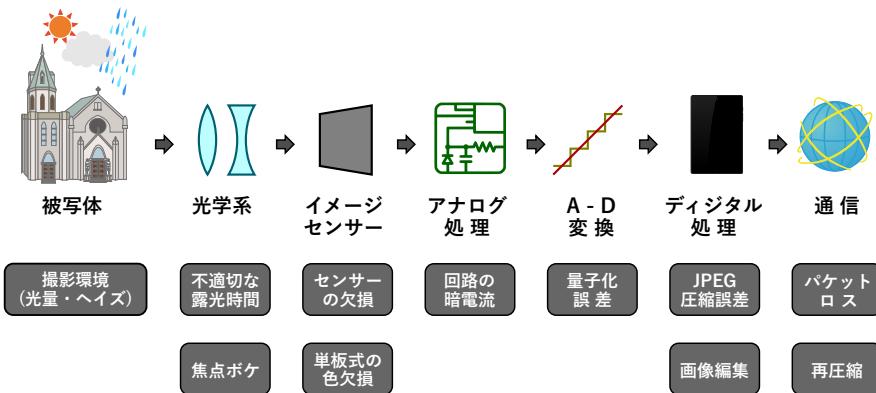


図 1.1 画像情報の取得とその過程で入る劣化

受光素子が物体から届く光の量を電圧に変換して計測する。受光素子は格子状に配置されているので得られる画像は 2 次元上に分布した値となる。ただし、実際にはコンピュータ上では連続値を取り扱えないで、受け取った光の量を 256 階調の離散値に量子化したものを画像情報として取り扱っている。

画像情報は図 1.1 上部の様に、被写体から撮像機器での各プロセス、画像処理を経る。この際に画像情報は同図の下部の様に様々な要因から劣化を受けたものとなる。被写体には霧や雨すじ等の特性が含まれる。被写体をレンズなどの光学系に通す際には、撮像機器の動きによるモーションぶれや、レンズの焦点位置に基づく焦点ぼけが発生する。光学系を通った光をイメージセンサーで取得する際に、センサーの欠損部分は画素の欠損となる。イメージセンサーから取得された電気信号を処理するアナログ回路においては暗電流に基づくノイズが混入する。アナログ信号からデジタル信号に変換する A-D 変換時には量子化誤差が、信号をユーザーが扱えるデータに変換するデジタル処理の際には JPEG 圧縮誤差や画像処理による劣化が混入する。

画像再構成技術は、これらの劣化過程をノイズとして画像から取り除き、本来の被写体の情報を得る技術であり、重要性が高い。画像の劣化過程は、撮影環境や焦点ぼけ、センサーの欠損などのノイズの混入要因に応じて異なったモデルに基づいており、これらのモデルの想定を間違うとノイズを除去できないのみでなく、本来の原画像の情報を損なう。そのため、処理したい画像の種類に応じて適切な劣化のモデルを選択する必要がある。

表 1.1 劣化行列と加法性ノイズによる劣化過程と再構成手法

劣化の種類	劣化行列 D	ノイズ n	再構成手法
焦点ボケ, 手ぶれ	畳み込み行列	加法性ノイズ	ぶれ除去
低解像度化	ダウンサンプル行列		画像補間, 超解像
ベイヤーパターン	マスク行列	加法性ノイズ	デモザイキング
カメラの暗電流		ガウスノイズ	ガウスノイズ除去
センサーの欠損		インパルスノイズ	インパルスノイズ除去
JPEG 圧縮誤差		ブロックノイズ	ブロックノイズ除去
画像の欠損	マスク行列		インペインティング

1.2 既存研究

1.2.1 画像の劣化過程の表現

既存の多くの研究では、劣化画像は画像に対して劣化行列をかけて、加法性ノイズを加算した形で劣化画像を表現する。すなわち、劣化のない画像のベクトル表現を x 、劣化画像を y 、行列の形で表現される劣化過程を D 、加法性ノイズを n としたときに、劣化画像は以下の表現で表すことができる。

$$y \simeq Dx + n \quad (1.1)$$

この表現において、劣化行列 D や加法性ノイズ n で何を表すかによって画像の劣化過程のモデルおよびその復元問題の種類が決定する。例えば、カメラの暗電流によって生じるガウスノイズや、センサーの欠損によって生じるインパルスノイズ、JPEG 圧縮誤差によって発生するブロックノイズなどは加法性ノイズ n によって表現することができる。画像の低解像度化によって発生するボケ、撮像機器の焦点ボケ・撮像機器の動きに基づくモーションぶれは劣化過程 D を用いて表現することができる。表 1.1 に劣化行列 D 、ノイズ n の表現と、それを除去する再構成手法の例を示した。

1.2.2 画像の超解像問題

画像信号を保存・伝送する際に低解像度化を行うと、画像にぼけが生じる。画像の高解像度化技術は、低解像度で撮像された画像を高解像度の表示機器に表示する際、画像補間によって生じるボケやジャギーの影響をなくし、より高繊細な画像を得る技術である。この技術は近年高解像度化の著しいディスプレイ等の画像表示機器上で、既存の映像資源をより高品質に表示するための技術として注目されている。図 1.2 に示すよう

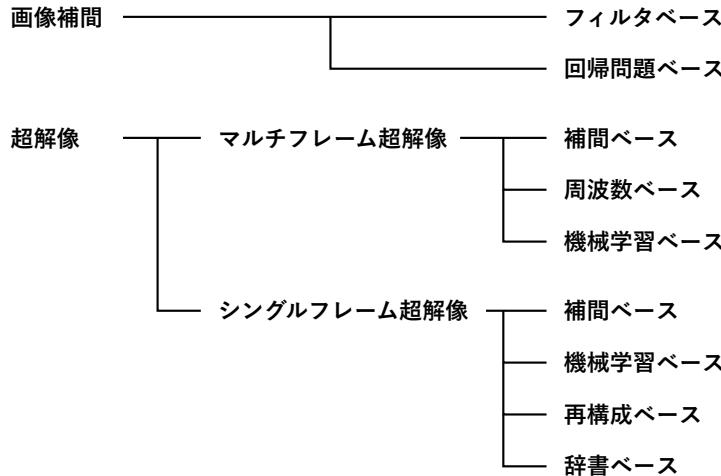


図 1.2 画像の高解像度化のためのアプローチの分類

に、画像の高解像度化のアプローチは、画像補間問題と超解像問題に分けられ、超解像問題はシングルフレーム超解像とマルチフレーム超解像に分類される。

画像補間問題は、得たい画素の近傍から線形フィルタ [1–3] や回帰問題 [4, 5] などを利用して画素値を得る手法であり、後述の超解像に比べて高速な手法が多い一方で、画像が本来持っていて、低解像度化の際に失われた情報の復元精度には限界があり、エッジ境界等でボケやジャギーを生じるものが多い。超解像問題は画像の高周波成分等の、入力低解像度画像に含まれなかった情報の復元を行う点で画像補間問題と異なっている。

超解像問題のアプローチには、入力画像の枚数によって、シングルフレーム超解像とマルチフレーム超解像に分かれる。マルチフレーム超解像は、複数の画像から高解像度画像を再構成する方式である。このような画像はビデオの連続するフレームや、露光時間の違う写真の組 [6]、異なるズームの画像の組 [7]、動画処理における異なるフレーム [8, 9] などを使用することで得られる。マルチフレーム超解像は、1枚の画像から高解像度画像を復元するシングルフレーム超解像に比べて、利用できる画像情報が多い分、高精度な画像を得やすく盛んに研究されており、補間を用いたものや [4]、画像の周波数特性を利用したもの [10]、MAP 法 (Maximum a Posteriori、最大事後確率) を用いたもの [11–14] が提案されている。一方で、マルチフレーム超解像は特に画像の拡大率が高い時に画像の詳細部が失われやすい問題がある [15]。

シングルフレーム超解像問題は一枚の画像から高解像度画像を得る手法であり、複数の画像を用意する必要がない分、マルチフレーム手法より広い範囲に適用できる。しかし、シングルフレーム超解像問題は得たい未知点の数が既知点の数よりも多い不良設定

問題となり、問題の不良設定性によって発生するぼけやジャギーなどをいかにして取り除くかが課題となる。手法 [16, 17] では画像補間技術を発展させ、補間の際のジャギーの発生を抑えている。手法 [18–20] では、画像の自己相似性に着目して、低解像度画像中の画素値に成り立つ関係式を高解像度画像に適用して超解像を行う。

画像再構成に基づくシングルフレーム超解像では、多数の学習データセットをもとに低解像度画像から高解像度画像を再構成する関係式を導出して超解像を実現する。手法 [21] では Markov Random Field を用いて低解像度・高解像度の関係を学習させる。手法 [22] はこのアプローチを拡張し、Primal Sketch Prior を用いてぼけたエッジ等を強調している。手法 [23] は Locally Linear Embedding を用いてより少ない学習データセットで学習を実現する方法を提案している。手法 [24, 25] ではハッシュ分けした画像領域ごとに個別に線形フィルタの形で低解像度・高解像度の関係を学習させている。手法 [26–29] では深層学習を用いて多数の学習データセットから低解像度画像から高解像度画像を再構成する関係式を導出する。

辞書ベースの超解像手法は画像再構成に基づくシングルフレーム超解像手法に分類され、事前に低解像度画像パッチと高解像度画像パッチの対応関係を辞書として保持しておき、低解像度画像を低解像度辞書で、高解像度画像を高解像度辞書で表現する。これは、辞書中の低解像度パッチと高解像度パッチの間にはぼけ・ダウンサンプルを含む低解像度・高解像度の関係が成り立っており、低解像度辞書から合成された画像と、高解像度辞書から構成された画像にも同様の関係が成り立っているという想定に基づいている。高解像度・低解像度の辞書のペアは、学習用の画像に対して、高解像度版と低解像度版を得て、それをもとに学習される。この方式は手法 [30] で初めて提案されている。画像を辞書の要素の線形結合で再構成する方式は、他の画像処理手法とも容易に統合することができ、詳細強調 [31] やノイズ除去 [32] 等に用いられている。辞書ベース超解像は高解像度画像の細かいテクスチャ部分の復元に関して他の手法と比べて有利である [33]。手法 [34] では事前に用意した高解像度・低解像度から切り出したパッチをもとに圧縮センシング [35, 36] に基づくスペース再構成問題を用いて超解像を行う。手法 [37] では、この手法の計算時間に着目し、組み合わせ辞書学習によって辞書の要素数を圧縮し、計算コスト削減を図っている。

画像劣化の観点では、低解像度画像は高解像度の画像を劣化させたものとみなし、その劣化特性を取り除くことで高解像度化を行う。低解像度画像は高解像度画像にダウンサンプルが入ったものとみなすことができる。また、信号に直接ダウンサンプルを行った時に高周波成分が回りこむ事によって発生するエイリアシングの影響をなくすために、原画像にぼけカーネル（点拡がり関数、PSF）の特性を想定する事が多い。そのた

め、超解像問題は低解像度画像をアップサンプルすると共に、ぼけカーネルの特性を取り除く問題と定義することができる。低解像度画像を \mathbf{y} 、高解像度画像を \mathbf{x} とし、ぼけカーネルを \mathbf{H} 、画素値のサンプリングを行う行列を \mathbf{S} としたときに、低解像度画像 \mathbf{y} は、高解像度画像 \mathbf{x} に、ぼけ \mathbf{H} 、サンプリング \mathbf{S} をこの順でかけた以下の形で表される。

$$\mathbf{y} = \mathbf{SHx} \quad (1.2)$$

これは、式 (1.1) において $\mathbf{D} = \mathbf{SH}$ とした表現とみなすことができる。しかし、この式を用いて \mathbf{y} から \mathbf{x} を求める問題は不良設定問題となるため、一意の \mathbf{x} を定めることはできない。

マルチフレーム超解像は、複数の画像を入力として用いることで問題の不良設定性を解消する。マルチフレーム超解像のアプローチとしては、すべての画像に対して式 (1.2) の関係が満たされると想定して、モーション行列 \mathbf{M}_k で位置合わせした高解像度画像と低解像度画像の間で以下の最小二乗問題を解いて高解像度画像を求める方法が取られる。

$$\mathbf{x} = \operatorname{argmin}_{\mathbf{x}} \sum_k \|\mathbf{y}_k - \mathbf{SHM}_k \mathbf{x}\|_2^2 + \rho(\mathbf{x}) \quad (1.3)$$

$\rho(\mathbf{x})$ は \mathbf{x} の事前分布によって決定される制約項であり、Total Variation ノルム [11,14], L_2 ノルム [12], L_1 ノルム [13] が用いられる。

1.2.3 画像のぶれ除去問題

撮像機器の手ぶれや被写体の動き、光学系の焦点距離と被写体とのずれ、光学系自身が持つ伝達関数によって画像にぶれが混入する。ぶれ除去問題はぶれを含む画像からぶれの影響を取り除いて本来のぶれ・ぼけのない画像を得る技術である。劣化画像であるぶれ画像は劣化のない原画像にカーネル (PSF) と呼ばれるぶれ特性が畳み込まれた画像とみなすことができる。手ブレ画像の補正は、この逆演算に相当し、逆畳み込み (Deconvolution) とよばれる。画像信号に対する既存の逆畳み込みのフレームワークは大きく 2 つに分けることができる。既知のカーネルを用いて画像を補正する Non-Blind Deconvolution と、ブレ画像のみを入力とし、カーネルと補正画像の両方を得る Blind Deconvolution である。Non-Blind Deconvolution の手法は古くから多くのアルゴリズムが提案されてきている。一方で Blind Deconvolution 問題は、求めたい画像とカーネルを得るために情報が入力のブレ画像から大きく欠落しているため、非常に困難な課題であり、Blind Deconvolution の具体的な手法が出てくるまでに、1990 年後半を待

つことになった。

画像の劣化過程の数学表現を考えた時、ぶれ画像はぶれのない画像とカーネルとの2次元畳み込みの関係で表現される。すなわち、ぶれ画像の2次元行列表現を \mathbf{Y} ぶれのない画像の2次元行列表現を \mathbf{X} 、カーネルの2次元行列表現を \mathbf{K} 、2次元畳み込みの演算を \otimes としたときに、 \mathbf{Y} は \mathbf{X}, \mathbf{K} を用いて以下のように表現される。

$$\mathbf{Y} = \mathbf{X} \otimes \mathbf{K} \quad (1.4)$$

また、ぶれ画像 \mathbf{Y} を1次元の列ベクトルにした表現を \mathbf{y} 、ぶれのない画像 \mathbf{X} を1次元の列ベクトルにした表現を \mathbf{x} としたときに、ぶれ画像 \mathbf{y} はぶれのない画像 \mathbf{x} に行列 \mathbf{C} を前からかけ合わせた表現として以下のように表される。

$$\mathbf{y} = \mathbf{C}\mathbf{x} \quad (1.5)$$

行列 \mathbf{C} は、カーネル行列 \mathbf{K} の要素を画像 \mathbf{x} に対応する位置に配置したもので、 \mathbf{K} の畳み込み行列とよばれる。畳み込み行列を用いることで、画像のぶれ・ぼけの特性は式(1.1)中の画像の劣化行列 \mathbf{D} を用いて表現できることがわかる。

ぶれ・ぼけのカーネルを未知とする Blind Deconvolution のアルゴリズムでは、カーネルの表現 \mathbf{K} も未知数として推定する問題となる。しかしながら、1枚のぶれ画像 \mathbf{Y} からカーネル \mathbf{K} および画像 \mathbf{X} の両方を推定する問題は不良設定問題となり、式(1.4)の関係を満たす画像 \mathbf{X} およびカーネル \mathbf{K} の組み合わせは一意に定まらない。問題の不良設定性を解消するアプローチの一つとして、カーネル \mathbf{K} に対して強い制約条件を課すアプローチが取られた。手法[38–42]では直線状のカーネル、手法[43–45]では折れ線状のカーネル、手法[46]では回転によって生ずるカーネルに限定して補正を行う。このような Blind Deconvolution のアプローチをパラメトリックなアプローチと呼ぶ。パラメトリックなぶれ軌跡推定法は、ブレ軌跡問題を大幅に簡略化できるものの、複雑なブレ軌跡には対応できない問題がある。

ぶれカーネルの形状に制約を加えないノンパラメトリックな推定手法も提案されている。ノンパラメトリックな推定手法はカーネルの形状に影響されない推定が可能になる他に、焦点ぼけを含むカーネル推定にも適用できる利点がある。この様なノンパラメトリックの推定手法は Fergus の手法[47]によって提案された。この手法は変分ベイズ推定によって \mathbf{K} と \mathbf{X} の推定値を得て、これによって得られたカーネルの推定値 \mathbf{K} と、ぶれ画像 \mathbf{Y} から、Richardson-Lucy 法[48, 49]を用いて最終的な補正画像を得るアルゴリズムである。Levin の手法[50]は Fergus の手法と同様にベイズ推定を用いてカーネルの推定値を得るが、Fergus の手法とは、画像 \mathbf{X} およびカーネル \mathbf{K} の事前分布の

設定方法に違いがある。Xu および Jia の手法 [51, 52] は最小二乗問題を用いたカーネル推定および逆畳込み手法を提案している。この手法ではショックフィルタを適用したぶれ画像 \mathbf{Y} からカーネル \mathbf{K} の推定値を得て、この推定値から補正画像 \mathbf{X} を得る。Cho の手法 [53] では、カーネル推定問題を周波数上の処理を用いて高速に処理するアルゴリズムを提案している。いずれのノンパラメトリック推定手法も、逆畳み込み問題を、カーネル \mathbf{K} の部分問題および画像 \mathbf{X} の部分問題に分割し、2つの問題を交互に実行する点に特徴がある。カーネル \mathbf{K} の部分問題では、 \mathbf{X} の部分問題で得られた補正画像 \mathbf{X} を固定して、ぶれ画像 \mathbf{Y} からカーネルの推定値を得る。 \mathbf{X} の部分問題では、カーネル推定問題によって得られたカーネル \mathbf{K} を固定して、ぶれ画像 \mathbf{Y} を Non-blind Deconvolution で逆畳み込みすることによって補正画像 \mathbf{X} を得る。式(1.4)の逆畳み込みを2つの部分問題に分割し、交互にこれらの問題を実行することで、各問題における不良設定性を解消することができる。

1.3 研究目的

本研究では画像の劣化除去のアプローチのうち、超解像問題とぶれ除去問題に焦点をあてる。超解像問題は、画像劣化のうち、低解像度化によって発生するぼけを除去する技術にあたる。近年、撮像機器やディスプレイ等の画像表示機器の解像度が大幅に向かっており、大型の表示機器への表示の他、スマートフォンやタブレット PC などの小型かつサイズの異なる表示機器への表示など、多岐にわたる画像の表示環境に応じて柔軟に画像のサイズ変更を行う需要は高まっている。前節で示した超解像のアルゴリズムを、画像枚数、速度、事前情報の有無で分類すると表 1.2 の様になる。辞書ベースの超解像手法は一枚の画像のみを必要とするシングルフレーム手法であり、エッジやテクスチャ部分の復元性能に優れているが、処理速度や、事前情報に依存する点に改良の見込みがある。そのため、本研究ではこの辞書ベース超解像手法をもとに改良を行う。従来の辞書ベース超解像手法は低解像度・高解像度の辞書のペアを予め用意しておく必要があり、事前辞書の容量や、パッチサイズ・辞書サイズ・拡大率等のパラメータに柔軟に対応できない問題があった。本研究では、入力低解像度画像から高解像度・低解像度の辞書の組み合わせを動的に生成し、従来法の問題であった、事前辞書を用意しておくことの欠点を解消した。また本研究では従来手法がスペース再構成問題を使用して超解像問題を行っていたのに対し、スペース再構成問題で用いられている L_1 正則化を Tikhonov 正則化に置き換えることで高速化を図った。 L_1 正則化を Tikhonov 正則化に置き換えることにより、再構成画像の精度には若干の低下が見られるが、実行時間の

表 1.2 超解像手法の画像枚数, 速度, 事前情報の必要性における比較

手法のベース	手法	画像枚数	速度	事前情報
マルチフレーム	MAP 法 [11–14]	複数	低速	不要
マルチフレーム 補間	Multi-Surface [4] Bilinear, Bicubic, DCC [16]	複数 一枚	高速 高速	不要 不要
補間	SAI [18], RSAI [19], BSAI [20]	一枚	低速	不要
機械学習	ANR [5], A+ [24]	一枚	低速	必要
機械学習	RAISR [25]	一枚	高速	必要
深層学習	SRCNN [26], VDSR [27]	一枚	低速	必要
辞書ベース	SRSR [34], ScSR [37]	一枚	低速	必要
辞書ベース	提案法	一枚	高速	不要

改善量はサイズ 512×512 の画像で 30 倍前後と非常に大きく, 実行時間短縮の利点が精度低下の欠点を上回ると考えられる.

ぶれ除去問題はぶれを含む画像からぶれの影響を取り除いて本来のぶれ・ボケのない画像を得る技術である. ぶれ特性は, 撮像機器の手ぶれや被写体の動き, 光学系の特性によって画像に混入するもので, 撮像環境の多様化によって多くの種類が存在するため, 直線や, 回転などの想定を加えないぶれ特性の除去手法が求められている. ノンパラメトリックな Blind Deconvolution 手法は撮像機器の手ぶれによるモーションブレ, レンズの収差や露光時間から生じる焦点ボケの両方を表現することができるため, 多様なぶれ特性に対応することができる. ノンパラメトリックなぶれ補正技術はカーネルを推定するカーネル推定問題と, 推定したカーネルから画像を補正する逆畳み込み問題を交互に行うことで実現され, カーネル推定問題・逆畳み込み問題ともに, 高速で高精度なアルゴリズムの選択が要求される. 本手法では, 高速なカーネル推定問題に特徴のある Cho の手法 [53] をもとにカーネル推定の精度向上および逆畳み込みの速度向上を行うことで, 高精度なぶれ補正をより実用的な処理時間で実現できるアルゴリズムの実現を図る. カーネル推定においては画像に特徴抽出フィルタを導入することで, 従来手法では用いられなかった領域でのカーネル推定を実現し, その精度向上を図った. 逆畳み込み問題では, 高精度な差分逆畳み込みを周波数軸上で実現する方法を考案し実行速度向上を図っている.

画像の劣化除去の観点から説明すると, 超解像問題は, ボケ行列とサンプリングを劣化過程に含む画像の再構成問題として取り扱うことができる. またぶれ除去問題はボケ行列に加えてモーションぶれを劣化過程に含む画像を再構成することができる. 両問題

ともぼけ行列のみを含む画像の再構成が可能である点が共通している。

1.4 本論文の構成

本論文は次のように構成される。まず2章で本研究の基礎となる画像の劣化過程について説明し、それを用いた簡単な画像再構成問題の説明と、画像再構成問題に用いられる最適化のアルゴリズムについて説明したのち、関連研究についての説明を行う。3章では画像超解像問題の提案手法として、事前辞書への依存の解消法およびL2再構成による高速化法の説明を行う。4章でぶれ除去問題の提案手法として、カーネル推定問題の高精度化法および逆畳み込み問題の高速化について説明を行う。5章で全体を総括し結論を述べる。

第 2 章

基礎理論

2.1 本章の構成

本章では提案する手法の理解に必要な、画像の劣化過程の数学表現とその再構成問題、本手法がベースとした画像の辞書ベース超解像問題およびぶれ除去問題の従来手法について説明を行う。まず 2.2 節において、画像の劣化過程としてよく用いられる数学表現について説明し、具体的な劣化過程の表現として画像の補間・超解像問題、ノイズの付加および画像のぶれ除去問題を取り上げる。次に 2.3 節において 2.2 節の劣化過程に基づく画像再構成手法として、非線形問題に基づく画像の再構成問題、スパース表現に基づく画像の再構成問題、確率理論に基づいた画像の再構成問題についての説明を行う。2.4 節ではこれらの再構成問題を実現するための最適化アルゴリズムについて説明を行う。2.5 節では画像の超解像問題の既存手法について説明を行い、2.6 節では画像のぶれ除去の既存手法についての説明を行う。

2.2 画像の劣化過程の数学表現

劣化のない画像のベクトル表現を $\mathbf{x} \in \mathbb{R}^{n_1 n_2}$ 、劣化画像を $\mathbf{y} \in \mathbb{R}^{n_1 n_2}$ としたときに、劣化画像は画像 $\mathbf{x} \in \mathbb{R}^{n_1 n_2}$ に対して、行列積の形で表現される劣化過程 $\mathbf{D} \in \mathbb{R}^{n_1 n_2 \times n_1 n_2}$ と、足し算の形で加わる加法性ノイズ $\mathbf{n} \in \mathbb{R}^{n_1 n_2}$ を用いて以下の表現で表すことができる。

$$\mathbf{y} \simeq \mathbf{D}\mathbf{x} + \mathbf{n} \quad (2.1)$$

劣化過程 \mathbf{D} および加法性ノイズ \mathbf{n} を変化させることで、超解像、ノイズ除去、手ブレ除去や、それらの組み合わせ等、多くの画像再構成問題に対応することができる。

2.2.1 画像の補間・超解像問題

画像の補間・超解像問題において、得たい高解像度画像を $\mathbf{x} \in \mathbb{R}^{n_h}$ 、低解像度画像を $\mathbf{y} \in \mathbb{R}^{n_l}$ とする。 n_h, n_l はそれぞれの画像の画素数で、拡大率を $s > 1$ とすると、 $n_h = n_1 n_2, n_h = s^2 n_l$ を満たす値である。このとき低解像度画像 \mathbf{y} は高解像度画像を \mathbf{x} にぼけカーネル \mathbf{H} とサンプリング行列 \mathbf{S} をこの順でかけた以下の形で表現できる。

$$\mathbf{y} = \mathbf{SHx} \quad (2.2)$$

ぼけカーネル $\mathbf{H} \in \mathbb{R}^{n_h \times n_h}$ は高解像度画像のナイキスト周波数より低いカットオフ周波数を持ち、サンプリングの際に発生するエイリアシングを除去する。行列 \mathbf{H} は畳み込み行列であり、図 2.1 右の様に、ぼけカーネルの係数を \mathbf{x} の対応する座標に配置したものである。サンプリング \mathbf{S} は、ぼかした高解像度画像 \mathbf{Hx} に対して、図 2.2 の様に、拡大率 s に応じた間引き操作を行う。行列 $\mathbf{S} \in \mathbb{R}^{n_l \times n_h}$ は、行列積の形でこの演算を行うもので、図 2.1 左の様に、高解像度画像のうち、残す画素を 1、それ以外の画素を 0 にした行列である。

マルチフレーム超解像問題を扱う際には、異なるフレーム間の空間的なずれを考慮するためのモーションを考慮する必要がある。すなわち、 k 番目の低解像度画像フレームを $\mathbf{y}_k \in \mathbb{R}^{n_l}$ としたとき。 \mathbf{y}_k は高解像度画像 \mathbf{x} に対して k 番目のモーション \mathbf{M}_k およびぼけカーネル \mathbf{H} 、サンプリング行列 \mathbf{S} が加わった表現として近似され、以下の表現で表される。

$$\mathbf{y}_k = \mathbf{SHM}_k \mathbf{x} \quad (2.3)$$

また、 \mathbf{y}_k をひとつの列ベクトルにまとめた表現を用いて大きな線形システムとして表現することもできる。

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} = \begin{bmatrix} \mathbf{SHM}_1 \\ \mathbf{SHM}_2 \\ \vdots \\ \mathbf{SHM}_K \end{bmatrix} \mathbf{x} \quad (2.4)$$

2.2.2 ノイズの付加

画像にノイズ成分が混入していると、画像処理アルゴリズムに対して好ましくない影響を与える。画像のノイズ成分は撮像機器の光学特性や電気ノイズ、イメージセンサーの欠損、JPEG 等の非可逆圧縮や通信の際のパケットロス等によって混入する。

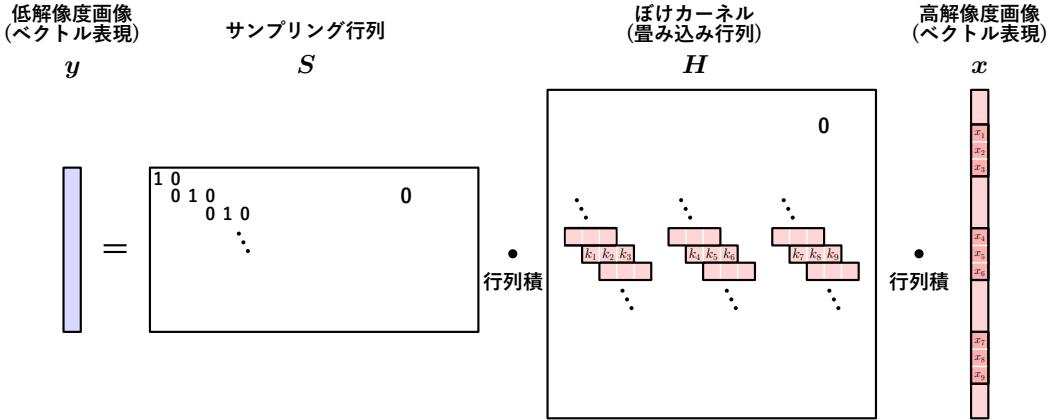


図 2.1 低解像度画像と高解像度画像の関係の数学表現

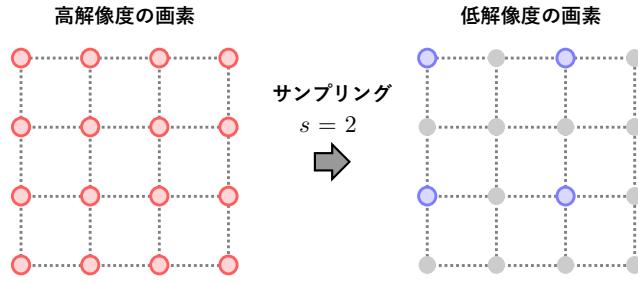


図 2.2 高解像度から低解像度へのサンプリング $s = 2$

ノイズには加法性ノイズと置き換えノイズがある。加法性ノイズはノイズのない画像 \mathbf{y} に対して、ノイズ成分 \mathbf{n} が加算された表現として表される。

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (2.5)$$

ノイズ成分 \mathbf{n} の種類として、 \mathbf{n} の要素 n_i がガウス分布を示すガウスノイズ $\mathcal{N}(0, \sigma^2)$ 、ガウスノイズが平滑化された平滑化ガウスノイズ（カラーノイズ）などが存在する。ガウスノイズは、加法性ノイズの中でも撮像機器から発生するノイズをよく近似することが知られており、[54–56] 等の手法でノイズ量を推定する手法が提案されている。

置き換えノイズはインパルスノイズとも呼ばれて、 \mathbf{x} 中の画素値 x_i が一定の確率 p で、ノイズの画素値 n_i に置き換わることで発生する。

$$y_i = \begin{cases} x_i & \text{with probability } p \\ n_i & \text{with probability } 1 - p \end{cases} \quad (2.6)$$

ここで、混合するノイズ n_i はガウシアン分布や一様分布 $U(0, 1)$ に従う。画像にガウスノイズおよびインパルスノイズを付加した例を図 2.3 に示した。



図 2.3 ガウスノイズ（標準偏差 25）とインパルスノイズ（混入率 15%）の例

ガウスノイズ除去手法としてウィナーフィルタ [57] や Non-local means フィルタ [58,59], BM3D [60–62], スパース近似 [63] などが使われる。インパルスノイズの除去にはしばしばガウスノイズとは異なる手法が適用される。これは、インパルスノイズにガウスノイズ除去の手法を適用すると、インパルスノイズが拡散して出力画像が大幅に劣化するからである。インパルスノイズ除去のための手法には中央値フィルタ [64], WMF [65], SMF [66], DWM [67] 等がある。SMF や DWM はインパルスノイズに対して、ノイズ検出とノイズ除去の 2 段階に分けた処理を行う特徴がある。

またガウス分布に従う加法性ノイズとインパルスノイズの両方が混入した混合ガウスノイズを考慮するアルゴリズムも考案されている [68–71]。

2.2.3 ぶれに基づく画像の劣化過程

ぶれ画像の 2 次元信号を $\mathbf{Y} \in \mathbb{R}^{N_1 \times N_2}$, これをぶれ除去した信号の 2 次元表現を $\mathbf{X} \in \mathbb{R}^{N_1 \times N_2}$, ぶれ軌跡（点拡がり関数, PSF, 以後カーネルと表記）の 2 次元表現を $\mathbf{K} \in \mathbb{R}^{M_1 \times M_2}$ と表現する。カーネル行列は図 2.4 の様に点の軌跡でモーションぶれ, 点の広がりで焦点ボケの両方を表現することができる行列である。この時, \mathbf{Y} は \mathbf{X} , \mathbf{K} の 2 次元畠み込みとして表現され, 以下の式が成り立つ。

$$\mathbf{Y} = \mathbf{X} \otimes \mathbf{K} \quad (2.7)$$

ここで, \otimes を 2 行列間の 2 次元畠み込みを計算する演算とした。座標 (n_1, n_2) における \mathbf{Y} , \mathbf{X} の画素値をそれぞれ $Y(n_1, n_2)$, $X(n_1, n_2)$ と表現する。この時に $Y(n_1, n_2)$

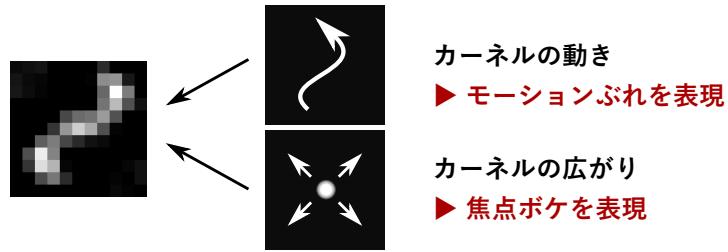


図 2.4 カーネル行列によるモーションぶれ・焦点ボケの表現

の値は (n_1, n_2) の周囲の値を \mathbf{K} の値で重み付き合計値として以下のように表される。

$$Y(n_1, n_2) = \sum_{k_1=1}^{N_1} \sum_{k_2=1}^{N_2} X(k_1, k_2) \cdot K(n_1 - k_1, n_2 - k_2) \quad (2.8)$$

以降の説明のために、画像 \mathbf{X} およびカーネル \mathbf{K} に対してベクトル表現および畳み込み行列を定義して、畳み込み表現 $\mathbf{X} \otimes \mathbf{K}$ を行列積の表現に変換する。

\mathbf{Y}, \mathbf{X} の要素を列ベクトル上に並べた表現をそれぞれ $\mathbf{y} \in \mathbb{R}^{N_1 N_2 \times 1}$, $\mathbf{x} \in \mathbb{R}^{N_1 N_2 \times 1}$ とする。また、カーネル \mathbf{K} の要素を列ベクトル上に並べた表現を $\mathbf{k} \in \mathbb{R}^{M_1 M_2 \times 1}$ と表記する。さらに、 \mathbf{K} の要素を 2 次元畳み込みの際に \mathbf{x} に対応する座標に配置した畳み込み行列を $\mathbf{C} \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$, \mathbf{X} の畳み込み行列表現を $\mathbf{A} \in \mathbb{R}^{N_1 N_2 \times M_1 M_2}$ とする。ここで、 \mathbf{A} の列数は出力画像のサイズに対応する $N_1 N_2$ としている。

この時に式 (2.8) の 2 次元畳み込み表現は、カーネルの畳み込み行列 \mathbf{C} と画像のベクトル表現 \mathbf{x} の積、あるいは、画像の畳み込み表現 \mathbf{A} とカーネルのベクトル表現 \mathbf{k} の積として以下の表現に置き換えることができる。

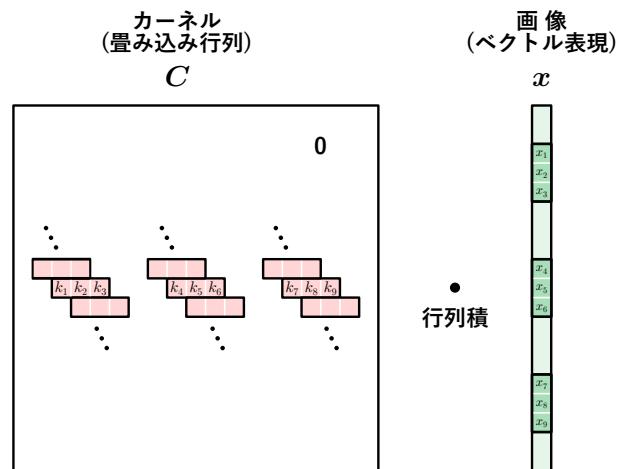
$$\mathbf{y} = \mathbf{Cx} \quad (2.9)$$

$$\mathbf{y} = \mathbf{Ak} \quad (2.10)$$

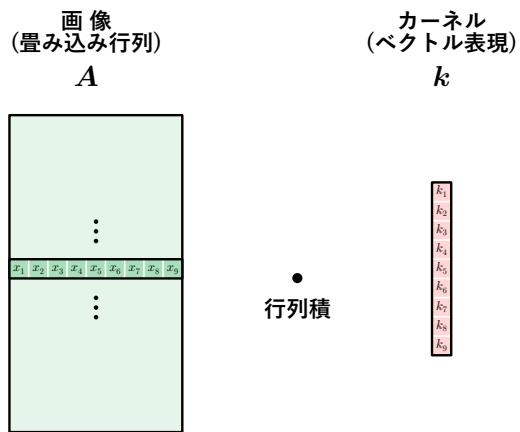
式 (2.9) の畳み込み行列 \mathbf{C} が、式 (2.1) の劣化過程 \mathbf{D} に対応する。画像・カーネルのベクトル表現と畳み込み行列表現の対応を図 2.5 にまとめた。



(a) 画像とカーネル行列の 2 次元畳み込み表現



(b) カーネルを畳み込み行列にした表現



(c) 画像を畳み込み行列にした表現

図 2.5 画像とカーネルの畳み込み行列表現

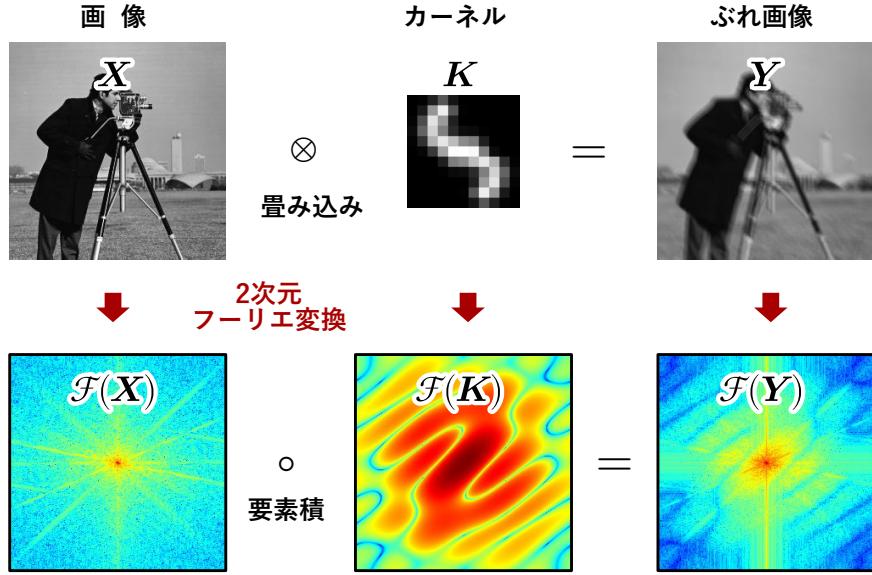


図 2.6 空間軸と周波数軸における畳み込みの関係

畳み込み定理に基づき, \mathbf{Y} , \mathbf{X} , \mathbf{K} の関係は周波数空間上でも表現できる. \mathbf{Y} , \mathbf{X} , \mathbf{K} を 2 次元フーリエ変換したものをそれぞれ $\mathcal{F}(\mathbf{Y})$, $\mathcal{F}(\mathbf{X})$, $\mathcal{F}(\mathbf{K})$ とするとき, $\mathcal{F}(\mathbf{Y})$ は $\mathcal{F}(\mathbf{X})$, $\mathcal{F}(\mathbf{K})$ を要素ごとに掛け算した以下の関係を満たす. ○は行列を要素ごとに掛け算する演算である.

$$\mathcal{F}(\mathbf{Y}) = \mathcal{F}(\mathbf{X}) \circ \mathcal{F}(\mathbf{K}) \quad (2.11)$$

図 2.6 に空間軸と周波数軸での畳み込みの関係を示した.

2.3 画像再構成問題

2.3.1 非線形問題による画像の再構成問題

このようにいくつかの画像の劣化過程は式 (2.1) の形で表現できることを示した. そのため, \mathbf{y} と劣化させた \mathbf{x} との間の最小 2 乗問題を解くという形で画像再構成問題を表現する事ができる.

$$\mathbf{x} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 \quad (2.12)$$

しかしながら, 式 (2.12) の問題は一般に劣決定問題であり, 上記の式を満たす \mathbf{x} は一意に定まらず, 得られる解も画像とはかけ離れたものとなる場合も存在する. 問題の劣決定性の解消や, \mathbf{x} をより自然な画像に近づけるために, \mathbf{x} に関する制約条件 $\rho(\mathbf{x})$ を

付加した以下の式を解く場合が多い.

$$\mathbf{x} = \underset{\mathbf{x}}{\operatorname{argmin}} \| \mathbf{y} - \mathbf{Dx} \|_2^2 + \rho(\mathbf{x}) \quad (2.13)$$

2.3.2 非線形最適化問題によく用いられる制約条件

ノルム制約

ベクトル \mathbf{x} に対して L_p ノルム ($p \geq 0$) を以下の式で定めることができる.

$$\| \mathbf{x} \|_p = \sqrt[p]{\sum_{n=1}^N |x_n|^p} \quad (2.14)$$

式 (2.14) のノルムで $p = 2$ としたノルムを L_2 ノルムという. L_2 ノルムに基づく正則化は Tikhonov 正則化とも呼ばれる. これは, 解 \mathbf{x} の集合のうち最もエネルギーの小さい解を選択する制約項であり, 問題の劣決定性を解消し, 最適化問題の収束を安定化させる目的で用いられる.

$$\| \mathbf{x} \|_2 = \sqrt{\sum_{n=1}^N x_n^2} \quad (2.15)$$

式 (2.14) のノルムで $p = 1$ としたノルムを L_1 ノルムという. L_1 ノルム最小化に基づく正則化はスパース正則化とも呼ばれ, x の要素の多くが 0 となり, 限られた要素のみが値を持つ表現となる. L_1 正則化を伴う線形回帰問題は LASSO と呼ばれ, これを対象とするアルゴリズムとして LARS [72] や GPSR [73] などが提案されている.

$$\| \mathbf{x} \|_1 = \sum_{n=1}^N |x_n| \quad (2.16)$$

画像再構成問題では, 画像の DCT 係数や DWT 係数に対するノルム制約を課すことが行われる. 画像 \mathbf{x} から DWT 係数を抽出する行列を \mathbf{W} とした時に, 画像の DWT 係数に対するノルム制約は以下の様に表現される.

$$\rho(\mathbf{x}) = \| \mathbf{Wx} \|_p \quad (2.17)$$

Total Variation ノルム

Total Variation (TV) ノルムは主に画像に対して多く用いられる制約項である [74–77]. これは画像 \mathbf{X} の微分成分に対して, L_1 や L_2 ノルムを適用したものである.

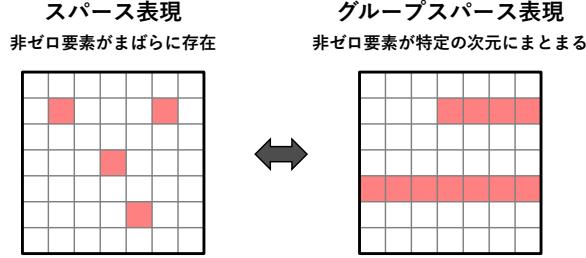


図 2.7 スパースな行列とグループスパースな行列

すなわち，画像に横方向微分を加えるカーネル $\partial_x = [1, -1]$ ，画像に縦方向微分を加えるカーネル $\partial_y = [1, -1]^T$ を定義した時に

$$\|\mathbf{X}\|_{\text{TV}} = \left\| \begin{bmatrix} \partial_x \otimes \mathbf{X} \\ \partial_y \otimes \mathbf{X} \end{bmatrix} \right\|_1 \quad (2.18)$$

と定義される。ここで，画像の微分値を加算するノルムは L_1 ノルムとした。

L1-L2 混合ノルム

行列 \mathbf{X} に対して，片方の次元に L_2 ノルムを，もう片方の次元に L_1 ノルムを適用した混合ノルムが提案されている。 L_1, L_2 混合ノルムは，要素 X_{ft} を持つ行列 \mathbf{X} に対して， t の次元で L_2 ノルム， f の次元で L_1 ノルムを取ったものに相当し，以下の式で表現される。

$$\|\mathbf{X}\|_{1,2} = \sum_{f=1}^F \sqrt{\sum_{t=1}^T X_{ft}^2} \quad (2.19)$$

L_1, L_2 混合ノルムを制約として解いた行列は図 2.7 右の様に行列中の非ゼロの要素が t の次元にまとまった分布を持つ。このような表現を行列のグループスパース表現という。

L_1, L_2 混合モデルを用いた画像再構成問題は [78] で述べられている。他にも JPEG ブロックノイズ除去 [79] 等で使用されている。

核ノルム

核ノルムは行列 \mathbf{X} に対して定義され，その行列のランク数 $\text{rank}(\mathbf{X})$ を最小化させる制約項として用いられる。行列 \mathbf{X} を特異値分解した時の特異ベクトル \mathbf{U}, \mathbf{V}^T ，特異値 $\sigma_1, \sigma_2, \dots, \sigma_N$ は以下の関係を満たす。

$$\mathbf{X} = \mathbf{U} \Sigma \mathbf{V}^T = \mathbf{U} \cdot \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_N) \cdot \mathbf{V}^T \quad (2.20)$$

この時に、行列 \mathbf{X} のランク数は以下の式で表される。

$$\text{rank}(\mathbf{X}) = \sum_{n=1}^N |\sigma_n|_0 = \|\boldsymbol{\sigma}\|_0 \quad (2.21)$$

ここで $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \dots, \sigma_N]$ とした。行列のランク数は特異値のベクトル $\boldsymbol{\sigma}$ の L_0 ノルムに相当し、これの最小化は NP 困難な問題となる。そのため、これの緩和のため、 L_0 ノルムを L_1 ノルムに緩和したノルムを以下のように定義する。

$$\|\mathbf{X}\|_* = \sum_{n=1}^N |\sigma_n| = \|\boldsymbol{\sigma}\|_1 \quad (2.22)$$

L_1 ノルムの他に L_p ノルムに緩和した Schatten- p ノルムも用いられる。

核ノルムを用いた画像再構成問題は、画像のノイズ除去 [80, 81] 等に用いられている。画像分野以外では、音声研究分野における楽曲からのボーカル分離 [82–84]、や多重音解析 [85] 等に使用されている。

2.3.3 スペース表現に基づいた画像の再構成問題

スペース表現はオリジナルの信号を限られた数の信号の線形結合で表現するための技術であり、画像再構成の他にパターン認識やコンピュータビジョンなどのアプリケーションで用いられている。

信号 $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K \in \mathbb{R}^n$ を K 個の既知サンプルとし、行列 $\mathbf{D} \in \mathbb{R}^{n \times K}$ をこれらのサンプルからなる行列とする。これは基底または辞書と呼ばれるもので、一般に過決定 ($n < K$) である。観測点を列ベクトル $\mathbf{x} \in \mathbb{R}^n$ とすると、すべての既知サンプルを用いて観測点は以下のように表現される。

$$\mathbf{x} = \mathbf{d}_1\alpha_1 + \mathbf{d}_2\alpha_2 + \cdots + \mathbf{d}_K\alpha_K \quad (2.23)$$

α_i は \mathbf{d}_i にかかる係数である。式 (2.23) は行列 \mathbf{D} とベクトル $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_K]^T$ を用いて以下の表現に書き換えられる。

$$\mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \quad (2.24)$$

しかしながら、式 (2.24) は劣決定な線形システムであり、 $\boldsymbol{\alpha}$ は事前情報や制約がない限り、单一の解が得られない。言い換えると式 (2.24) のみでは、観測点 \mathbf{x} を基底行列 \mathbf{D} で表現する方法は一つに定まらない。この問題に対処するために解 $\boldsymbol{\alpha}$ に対して適切な制約項を課すことが不可欠である。スペース表現では、係数 $\boldsymbol{\alpha}$ がスペースな解を持

つように解く。スパースとは、列ベクトルの要素の多くが 0 または 0 に近い値であり、僅かな数の係数が大きな値を持つ状態を表す。

最もスパースな解は (2.24) の線形システムを L_0 ノルム最小化制約と共に解くことで得られる。すなわち、(2.24) の問題は以下の最適化問題に書き換えられる。

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{\alpha}\|_0 \quad \text{s.t.} \quad \mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \quad (2.25)$$

ここで $\|\cdot\|_0$ はベクトル中の非ゼロ要素の数であり、ベクトルのスパース性を指し示す指標である。さらに観測点が辞書行列中の k 個の要素によって表現されているものとすると、問題 (2.25) は以下の問題と等価になる。

$$\mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \quad \text{s.t.} \quad \|\boldsymbol{\alpha}\|_0 \leq k \quad (2.26)$$

式 (2.25) の解を得るには、MP 法 [86] や OMP 法 [87] が用いられる。

L_0 最小化問題は $\boldsymbol{\alpha}$ のスパース解を得ることができるもの、この問題は NP 困難な問題であり、解を得ることが難しい。 L_0 最小化の NP 困難性を解消するために L_1 ノルムを用いたスパース再構成が考案されている。 L_1 ノルム最小化は Lasso 問題に基づいており、機械学習やパターン認識、統計分野で広く用いられている。 L_1 最小化制約によって得られた解はスパース性の要件を満たしており、 L_0 ノルム最小化によって得られる解と同等のものが得られることが示されている。さらに、 L_1 ノルム最適化は解析的な解が得られ、多項式時間で解くことができる。これらの理由から、 L_1 ノルム最小化はスパース再構成問題において L_0 ノルムの代わりに広く用いられている。 L_1 ノルム最小化に基づくスパース再構成問題は次の式で表される。

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{\alpha}\|_1 \quad \text{s.t.} \quad \mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \quad (2.27)$$

画像処理に辞書に基づいたスパース再構成を適用するには、画像をオーバーラップしたパッチに分割し、パッチに対して処理を行う方法が取られる。画像のパッチを \mathbf{x} とし、パッチを辞書 \mathbf{D} を用いて以下の表現で近似する。

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{\alpha}\|_p \quad \text{s.t.} \quad \|\mathbf{x} - \mathbf{H}\mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \varepsilon \quad (2.28)$$

これによって得られた係数 $\boldsymbol{\alpha}$ を用いて $\hat{\mathbf{x}} = \mathbf{D}\hat{\boldsymbol{\alpha}}$ より再構成パッチを得る。このパッチを画像の該当する場所に戻し、オーバーラップした箇所を平均する。これは再構成した画像の不連続性を解消し、より自然な出力を得るためである。

2.3.4 確率理論に基づいた画像の再構成問題

2.3.1 節の式 (2.12) において、劣化画像 \mathbf{y} とそのモデル $\mathbf{D}\mathbf{x} + \mathbf{n}$ の間に 2 乗誤差をとり、その最小化を行ったが、 \mathbf{y} や \mathbf{x} を確率変数と捉えることで、 \mathbf{x} に事前分布を導入したり、ガウスノイズ以外の分布をもつノイズを想定することができる。代表的な方法は \mathbf{x} に事前分布を導入し、その事後確率を最大化する最大事後確率 (MAP: maximum a posteriori) 推定で、画像のデノイズ [88, 89] やインペインティング [90] に使われている。

画像再構成問題において、画像 \mathbf{x} に関する事後確率を $p(\mathbf{x}|\mathbf{y})$ と定義する。これをベイズの定理によって、 \mathbf{y} に関する確率に書き換える。

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}) \cdot p(\mathbf{x}) \quad (2.29)$$

この時に、 $p(\mathbf{y}|\mathbf{x})$ が尤度関数、 $p(\mathbf{x})$ が画像信号の事前分布となる。 \mathbf{y} と \mathbf{x} の間に標準偏差 σ のガウスノイズを仮定すると $p(\mathbf{y}|\mathbf{x})$ は以下の形となる。

$$\begin{aligned} p(\mathbf{y}|\mathbf{x}) &= \prod_n \mathcal{N}(y_n|x_n, \sigma^2) \\ &= \prod_n \exp\left(-\frac{1}{2\sigma^2}(x_n - y_n)^2\right) \end{aligned} \quad (2.30)$$

よく使われる画像分布は、画像の微分成分に対するラプラス分布であり、以下の式で定義される。

$$p(\mathbf{x}) \propto \prod_n \exp(-|f \otimes x_n|^\alpha) \quad (2.31)$$

これらを用いると \mathbf{x} の再構成は以下の式を最大化する問題となる。

$$\max_{\mathbf{x}} \prod_n \exp\left(-\frac{1}{2\sigma^2}(x_n - y_n)^2\right) \cdot \prod_n \exp(-|f \otimes x_n|^\alpha) \quad (2.32)$$

確率変数の積の指数値は有限精度の計算機上では扱いづらいので、(2.32) 全体の対数を取る。これによって確率変数の積が和の形に変換される。

$$\max_{\mathbf{x}} \left(-\frac{1}{2\sigma^2} \sum_n (x_n - y_n)^2 - \sum_n |f \otimes x_n|^\alpha \right) \quad (2.33)$$

さらに、式全体の負値をとって最大化問題を最小化問題にする。

$$\min_{\mathbf{x}} \left(\frac{1}{2\sigma^2} \sum_n (x_n - y_n)^2 + \sum_n |f \otimes x_n|^\alpha \right) \quad (2.34)$$

この式は式 (2.13) の制約付き最小 2 乗問題と同じとなる。これは、 \mathbf{y} と \mathbf{x} の間の誤差にガウス分布を想定したためである。

2.4 最適化問題

2.4.1 最急降下法

式 (2.13) の様な最適化問題を実現するための最も単純なアルゴリズムは最急降下法である。最急降下法は、勾配降下法とも呼ばれ、目的関数の一階微分が計算可能な場合に適用できる。ベクトル \mathbf{x} に対して、目的関数を $f(\mathbf{x})$ と定義する。この時、 \mathbf{x} を以下の式で更新することで解に近づけていく。

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \gamma \nabla f(\mathbf{x}^{(k)}) \quad (2.35)$$

$\nabla f(\mathbf{x})$ は $f(\mathbf{x})$ の \mathbf{x} に関する勾配である。

$$\nabla f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x}^{(k)})}{\partial x_1^{(k)}}, \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_2^{(k)}}, \dots, \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_n^{(k)}} \right]^T \quad (2.36)$$

γ はステップ幅を表すパラメータとなる。ステップ幅が小さいほど収束が安定するが、解を得るのに多くの反復を必要とする。ステップ幅が大きいと最初は早く収束することが期待できるが、解が発散する場合もある。このため、問題に応じて適切なステップ幅を選ぶ必要があり、探索の初期には大きなステップ幅、ある程度収束したら小さなステップ幅が取られる。

最急降下法をベースにした手法は多くの分野で用いられている。例えば、最急降下法をオンライン学習に改良した確率的勾配降下法は深層学習の学習等に用いられている。

2.4.2 近接勾配法

最急降下法は目的関数の偏微分を計算できる必要があるが、凸ではあるが微分不可能な関数でも近接写像が計算可能な問題は近接勾配法を用いて解くことができる [91]。

目的関数を以下の関数 $f(\mathbf{x})$ と $g(\mathbf{x})$ の合計とする。

$$\operatorname{argmin}_{\mathbf{x}} f(\mathbf{x}) + g(\mathbf{x}) \quad (2.37)$$

ここで、 f は \mathbf{x} に関する勾配が導出可能な関数、 g は \mathbf{x} に関する近接写像が導出可能な関数である。近接写像とは

$$\operatorname{prox}_g(\mathbf{x}) = \operatorname{argmin}_{\mathbf{v}} (g(\mathbf{v}) + (1/2)\|\mathbf{x} - \mathbf{v}\|_2^2) \quad (2.38)$$

で定義される演算である。2.3.1 節の再構成問題で用いられる制約項には近接写像が計算可能なものが多い。これを用いて、近接勾配法は以下の式で \mathbf{x} の更新を行うことで解を得る。

$$\mathbf{x}^{(k+1)} := \text{prox}_{\lambda^{(k)} g}(\mathbf{x}^{(k)} - \lambda^{(k)} \nabla f(\mathbf{x}^{(k)})) \quad (2.39)$$

近接勾配法の改良として、収束速度を向上させた FISTA、拡張ラグランジュ法や ADMM 等が考案されている。

2.4.3 共役勾配法による高速な解法

共役勾配法 [92] とは、逆行列問題を解くためのアルゴリズムの一つであり、行列 \mathbf{A} 対称正定値行列であるときに、高速に解くことができるアルゴリズムである。正定値行列とは、エルミート行列 \mathbf{A} と、零ベクトルでない任意の列ベクトル \mathbf{z} に対して、

$$\mathbf{z}^H \mathbf{A} \mathbf{z} > 0 \quad (2.40)$$

が常に成り立つ行列 \mathbf{A} のことである。 \mathbf{z}^H はベクトル・行列 \mathbf{z} のエルミート転置である。

共役勾配法を用いて、正定値行列 \mathbf{A} に対し

$$\mathbf{y} = \mathbf{A} \mathbf{x} \quad (2.41)$$

を満たす \mathbf{x} を求めるプロセスは Algorithm 1 で表される。ここでベクトル $\mathbf{r}, \mathbf{p}, \mathbf{z}$ を中間表現、 e_i を残差、 x_{step} をステップ幅としている。共役勾配法は、行列 \mathbf{A} が正定値性を満たす必要があるが、 \mathbf{A} のサイズが $n \times n$ の時に、高々 n ステップで収束する。また零行列でない行列 \mathbf{A} に対し、行列 $\mathbf{A}^H \mathbf{A}$ は正定値性を満たすので、

$$\begin{aligned} \mathbf{A}^H \mathbf{y} &= \mathbf{A}^H \mathbf{A} \mathbf{x} \\ \mathbf{x} &= (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y} \end{aligned} \quad (2.42)$$

とすることで、正定値行列でない行列の逆問題を解くことも可能である [93]。

Algorithm 1 共役勾配法の反復プロセス

1. $\mathbf{x} = \mathbf{y}$
 2. $\mathbf{r} = \mathbf{y} - A\mathbf{x}$
 3. $\mathbf{p} = \mathbf{r}$
 4. $e_0 = \mathbf{r}^H \mathbf{r}$
 5. **for** $i = 1$ **to** converge **do**
 6. $\mathbf{z} = A\mathbf{p}$
 7. $x_{step} = \frac{e_{i-1}}{\mathbf{p}^H \mathbf{z}}$
 8. $\mathbf{x} = \mathbf{x} + x_{step} \mathbf{p}$
 9. $\mathbf{r} = \mathbf{r} - x_{step} \mathbf{z}$
 10. $e_i = \mathbf{r}^H \mathbf{r}$
 11. $\mathbf{p} = \frac{e_i}{e_{i-1}} \mathbf{p}$
 12. **end for**
-

2.5 従来の辞書ベース超解像手法

本節では画像の超解像アルゴリズムのうち、提案手法のベースとなっている Yang の辞書ベース超解像アルゴリズム [37] について述べる。辞書ベースの超解像手法は、高解像度のパッチを、高解像度・低解像度の辞書の組み合わせから再構成することで得る。この方式は [30] で初めて提案されている。Yang の辞書ベース超解像アルゴリズム [37] は ScSR (Sparse constrained super resolution) と呼ばれ、高解像度・低解像度の辞書のペア用いてスパース再構成問題によって高解像度パッチを再構成する。この際に、組み合わせ辞書学習を用いることで辞書のサイズを削減して、再構成の際の計算コストを削減した点に特徴がある。

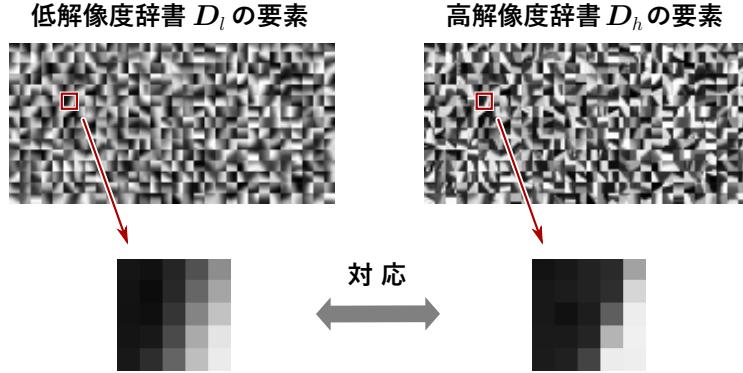


図 2.8 辞書行列中の高解像度パッチ・低解像度パッチの対応

2.5.1 辞書ベースの超解像

辞書ベース超解像のアルゴリズムは高解像度画像のパッチを辞書の要素とその要素にかかる係数の形に近似する。高解像度画像からサイズ $\sqrt{n} \times \sqrt{n}$ の領域を切り出した高解像度パッチ（ベクトル表現）を $x \in \mathbb{R}^n$, 低解像度画像の同様の領域を切り出した低解像度パッチを $y \in \mathbb{R}^n$ とする。この時に、辞書ベース手法では、高解像度パッチ x は、高解像度用の辞書行列 $D_h \in \mathbb{R}^{n \times K}$ とその係数 $\alpha \in \mathbb{R}^K$ を用いて、以下の形で近似する。

$$x = D_h \alpha \quad (2.43)$$

辞書ベースの超解像手法 [34, 37] では、式 (2.43) の関係式が低解像度パッチ y および低解像度様の辞書 D_l に対して同じ係数 α を用いて表現されるものと想定する。すなわち、低解像度パッチ y は、低解像度用の辞書 D_l と、高解像度と共に係数 α の積で表現される。

$$y = D_l \alpha \quad (2.44)$$

ここで、辞書 D_h と D_l の各列 d_h, d_l は、対応する高解像度画像と低解像度画像の組から、式 (2.43) と式 (2.44) が同じ係数 α を持つように学習されたもので、図 2.8 の様に画像のパッチの高解像度版と低解像度版の対応を含むものである。

問題の悪条件性を改善するために式 (2.43) に α の L_0 ノルム最小化の制約項を付加した以下の表現に置き換える。

$$x = D_h \alpha \quad \text{with } \|\alpha\|_0 \ll K \quad (2.45)$$

α_0 の L_0 ノルムは α 中の非ゼロの要素の数を表すため、式 (2.45) の表現は、高解像

度パッチ \mathbf{x} が高解像度辞書 \mathbf{D}_h のうちの限られた要素の組み合わせで近似されていることを表す。 α 中の非ゼロの要素数を少なくしたものをスパース表現という。 L_0 ノルム最小化の制約を低解像度パッチ \mathbf{y} にも適用する。

$$\mathbf{y} = \mathbf{D}_l \alpha \quad \text{with } \|\alpha\|_0 \ll K \quad (2.46)$$

2.5.2 特徴抽出空間による近似

前節の近似が、より画像の特徴的な部分にマッチする様に、入力低解像度パッチ \mathbf{y} に対して特徴抽出を適用する手法が提案されている [21, 23, 94]。特徴抽出フィルタは入力画像の構造を強調するもので、典型的にはハイパスフィルタが用いられる。これは高解像度の近似を行うためには、パッチの低周波成分よりも高周波成分のほうが重要であると考えられるためである。辞書ベース超解像では入力低解像度画像から切り出したパッチ \mathbf{y} に対して 4 つの特徴抽出フィルタ $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4$ を適用する。 $\mathbf{f}_1, \mathbf{f}_2$ は画像に横、縦の 1 階微分を行う演算子、 $\mathbf{f}_3, \mathbf{f}_4$ は画像に横、縦の 2 階微分を行う演算子である。これらの特徴抽出フィルタはそれぞれ以下の様に定義される。

$$\begin{aligned} \mathbf{f}_1 &= [-1, 0, 1] & \mathbf{f}_2 &= \mathbf{f}_1^T \\ \mathbf{f}_3 &= [1, 0, -2, 0, 1] & \mathbf{f}_4 &= \mathbf{f}_1^T \end{aligned} \quad (2.47)$$

$\mathbf{f}_1, \mathbf{f}_2$ は画素のずれを回避するために中間差分を用いている。また $\mathbf{f}_3, \mathbf{f}_4$ は $\mathbf{f}_1, \mathbf{f}_2$ をそれぞれ 2 回適用したものにあたるので 1 と -2 の間に 0 が入る形になる。

特徴抽出フィルタを画像再構成問題に組み込むために、特徴抽出行列 \mathbf{F} を考える。これは、特徴抽出フィルタ \mathbf{f}_1 から \mathbf{f}_4 の、行列 \mathbf{y} に対する畳込み行列を縦に並べたものである。

$$\mathbf{F}\mathbf{y} = \begin{bmatrix} \mathbf{f}_1 \otimes \mathbf{y} \\ \mathbf{f}_2 \otimes \mathbf{y} \\ \mathbf{f}_3 \otimes \mathbf{y} \\ \mathbf{f}_4 \otimes \mathbf{y} \end{bmatrix} \quad (2.48)$$

この特徴抽出行列を用いて、式 (2.46) の低解像度パッチの近似を以下の様に書き換える。

$$\mathbf{F}\mathbf{y} = \mathbf{F}\mathbf{D}_l \alpha \quad \text{with } \|\alpha\|_0 \ll K \quad (2.49)$$

2.5.3 L1 最適化による高解像度画像の再構成

式 (2.49) から係数 α を求める問題は $\mathbf{FD}_l\alpha$ と \mathbf{Fy} との 2 乗誤差を最小化する問題として、以下の様に表現することができる。

$$\alpha = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_0 \quad \text{s.t.} \quad \|\mathbf{FD}_l\alpha - \mathbf{Fy}\|_2^2 \leq \varepsilon \quad (2.50)$$

ε は $\mathbf{FD}_l\alpha$ と \mathbf{Fy} との間の誤差をどれくらい許容するかの尺度である。しかし、 L_0 最小化問題は NP 困難な問題であり、最適な解を探索するのに大きなコストがかかる。問題の NP 困難性を解消し、かつ α のスパースな解を得るために、 α の L_0 ノルムの代わりに L_1 ノルムの最小値を得る方法が取られる [95]。これを問題の緩和といい、上記の問題は L_1 ノルムを用いた以下の形に置き換えられる。

$$\alpha = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_1 \quad \text{s.t.} \quad \|\mathbf{FD}_l\alpha - \mathbf{Fy}\|_2^2 \leq \varepsilon \quad (2.51)$$

式 (2.51) を解析的に解くためにラグランジュ係数 λ を用いて以下の形を用いる。

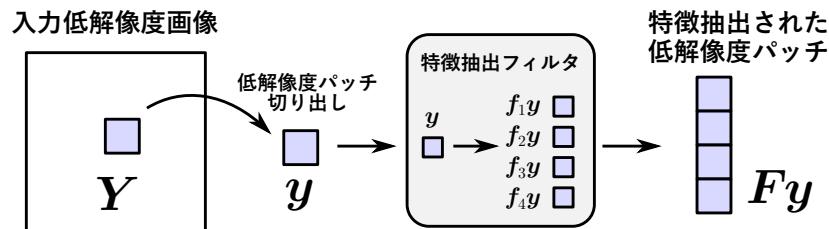
$$\alpha = \underset{\alpha}{\operatorname{argmin}} \lambda \|\alpha\|_1 + \frac{1}{2} \|\mathbf{FD}_l\alpha - \mathbf{Fy}\|_2^2 \quad (2.52)$$

λ は α のスパース性と、 $\mathbf{FD}_l\alpha$ 、 \mathbf{Fy} 間の 2 乗誤差を調整する係数とみなすことができる。(2.52) の前半部分をスパース制約項、後半部分を誤差項とよぶ。

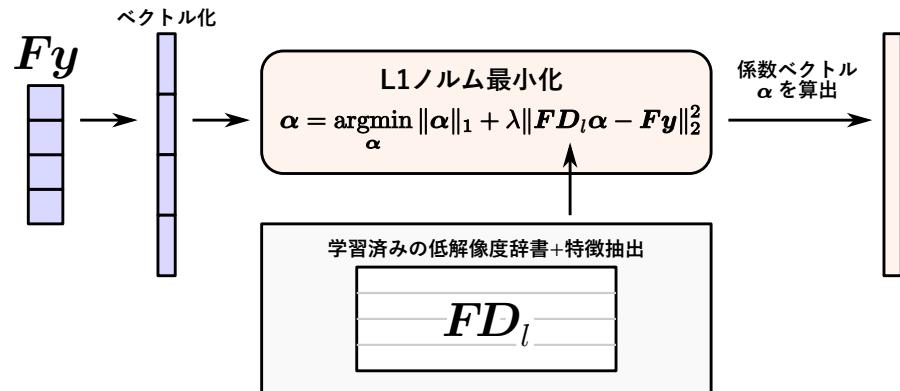
2.5.4 辞書ベース超解像の流れ

以上を踏まえて、低解像度画像から、辞書行列を用いて高解像度画像を再構成する手順は以下のようになる。

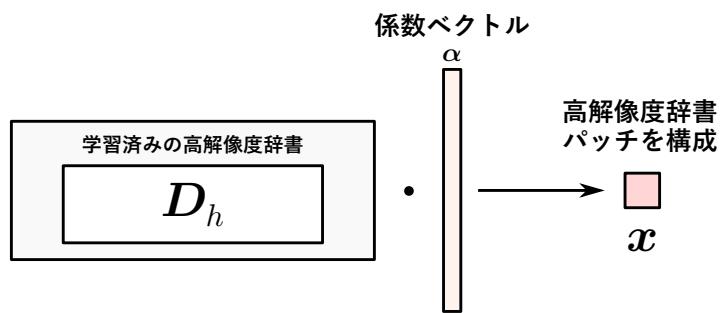
- 入力低解像度画像 \mathbf{Y} の特定の座標から低解像度パッチ \mathbf{y} を切り出す(図 2.9 (a))
- 低解像度パッチに対して特徴抽出フィルタをかけ \mathbf{Fy} を得る
- \mathbf{Fy} と辞書行列 \mathbf{FD}_l から式 (2.52) を用いて係数ベクトル α を導出(図 2.9 (b))
- 求めた係数ベクトルを用いて式 (2.43) から高解像度パッチ \mathbf{x} を再構成(図 2.9 (c))
- 画像の各座標に関して \mathbf{x} を算出して、高解像度画像 \mathbf{X} を合成



(a) 低解像度パッチの切り出しと特徴抽出



(b) 係数ベクトルの算出



(c) 高解像度パッチの合成

図 2.9 辞書ベース超解像の流れ

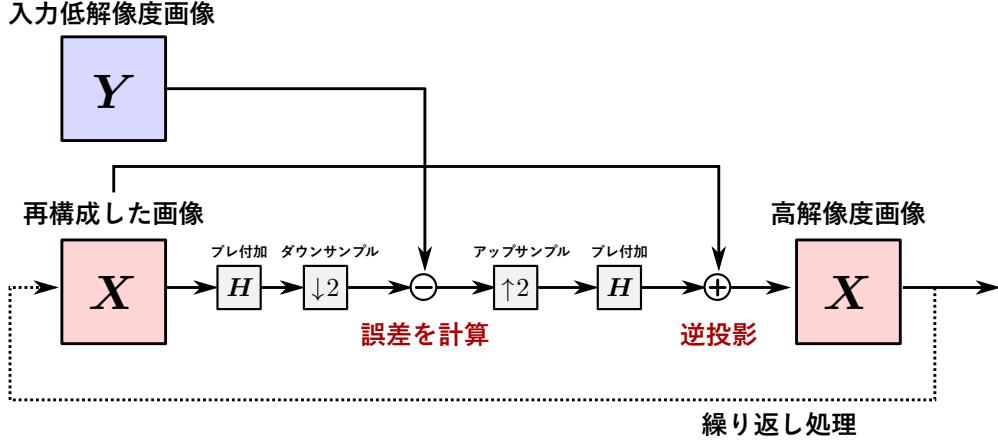


図 2.10 逆投影法の処理プロセス

2.5.5 グローバル再構成問題

画像の L_1 再構成問題は画像をパッチごとに処理するため、画像の境界等に不連続性を生じやすい。この問題を解消するために、[34], [37] の手法では、超解像問題の後処理として、以下に示すグローバル再構成問題を行っている。

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \| \mathbf{S} \mathbf{H} \mathbf{X} - \mathbf{Y} \|_2^2 + c \| \mathbf{X} - \mathbf{X}_0 \| \quad (2.53)$$

\mathbf{X}_0 は先の節によって再構成された高解像度画像である。

式 (2.53) の最適化は逆投影法と呼ばれる。これは、得られた高解像度画像と低解像度画像との誤差を高解像度画像に投影するものである。逆投影法は以下の式を用いて、高解像度画像 \mathbf{X}_t を \mathbf{X}_{t+1} に更新することで最終的な高解像度画像を得る。

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \mathbf{H}^T \mathbf{S}^T (\mathbf{Y} - \mathbf{S} \mathbf{H} \mathbf{X}_t) \quad (2.54)$$

\mathbf{S}^T は画像に s 倍のアップサンプル $\uparrow s$ を適用する処理。 \mathbf{H}^T は画像に \mathbf{H} と同様のぶれをかける処理に相当する。2 倍拡大 $s = 2$ の場合の逆投影法のプロセスは図 2.10 に示す。再構成された画像 \mathbf{X} に、ぶれ \mathbf{H} とダウンサンプル $\downarrow 2$ を加えた後、入力低解像度画像 \mathbf{Y} との誤差を計算する。この誤差 $\mathbf{Y} - \mathbf{S} \mathbf{H} \mathbf{X}$ に対して、2 倍のアップサンプルとぶれを加えたものを高解像度画像 \mathbf{X} に加える（逆投影する）ことで更新された高解像度画像 \mathbf{X}_{t+1} を得る。これを繰り返すことで出力高解像度画像の精度を高めていくことができる。パラメータ c は、式 (2.54) の更新の反復数に影響する。本手法ではこの反復を 20 回実行する。

2.5.6 事前辞書の学習

辞書ベース超解像問題で用いる辞書のペア $\{\mathbf{D}_l, \mathbf{D}_h\}$ は、学習用の高解像度画像と低解像度画像のペアから学習されたものである。[34] の手法では、 $\{\mathbf{D}_l, \mathbf{D}_h\}$ の要素に画像からサンプルしたパッチのペアを用いている。これは、辞書のサイズが大きくなり、計算コストがかかる問題点がある。そこで [37] の手法では、学習用の高解像度画像と低解像度画像のペアから比較的少数の要素をもつ辞書を学習させる方法を提案している。辞書学習の際には高解像度の辞書に対するスパース係数と、低解像度の辞書に対するスパース係数が同じものとなる必要がある。高解像度辞書の学習問題と、低解像度辞書の学習問題はそれぞれ以下の様に表される。

$$\mathbf{D}_h = \underset{\mathbf{D}_h, \mathbf{Z}}{\operatorname{argmin}} \|\mathbf{X}_t - \mathbf{D}_h \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z}\|_1 \quad (2.55)$$

$$\mathbf{D}_l = \underset{\mathbf{D}_l, \mathbf{Z}}{\operatorname{argmin}} \|\mathbf{Y}_t - \mathbf{D}_l \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z}\|_1 \quad (2.56)$$

\mathbf{X}_t は学習用の高解像度画像からサンプルされた高解像度パッチ群、 \mathbf{Y}_t は低解像度画像からサンプルされたパッチ群である。また、 \mathbf{Z} はこれらのパッチ群を現在の辞書 \mathbf{D}_h 、 \mathbf{D}_l で近似したときのスパース係数で、高解像度パッチの近似と、低解像度パッチの近似で共有されるものである。辞書のペアを学習する際にはこれら 2 つの問題を結合させる。

$$\{\mathbf{D}_l, \mathbf{D}_h\} = \underset{\mathbf{D}_l, \mathbf{D}_h, \mathbf{Z}}{\operatorname{argmin}} \|\mathbf{X}_t - \mathbf{D}_h \mathbf{Z}\|_2^2 + \|\mathbf{Y}_t - \mathbf{D}_l \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z}\|_1 \quad (2.57)$$

さらに \mathbf{D}_h 、 \mathbf{D}_l を縦に結合させた行列 \mathbf{D}_c を用いて以下の様に書き換える

$$\mathbf{D}_c = \underset{\mathbf{D}_c, \mathbf{Z}}{\operatorname{argmin}} \|\mathbf{X}_c - \mathbf{D}_c \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z}\|_1 \quad (2.58)$$

$$\mathbf{X}_c = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \end{bmatrix}, \quad \mathbf{D}_c = \begin{bmatrix} \mathbf{D}_h \\ \mathbf{D}_l \end{bmatrix} \quad (2.59)$$

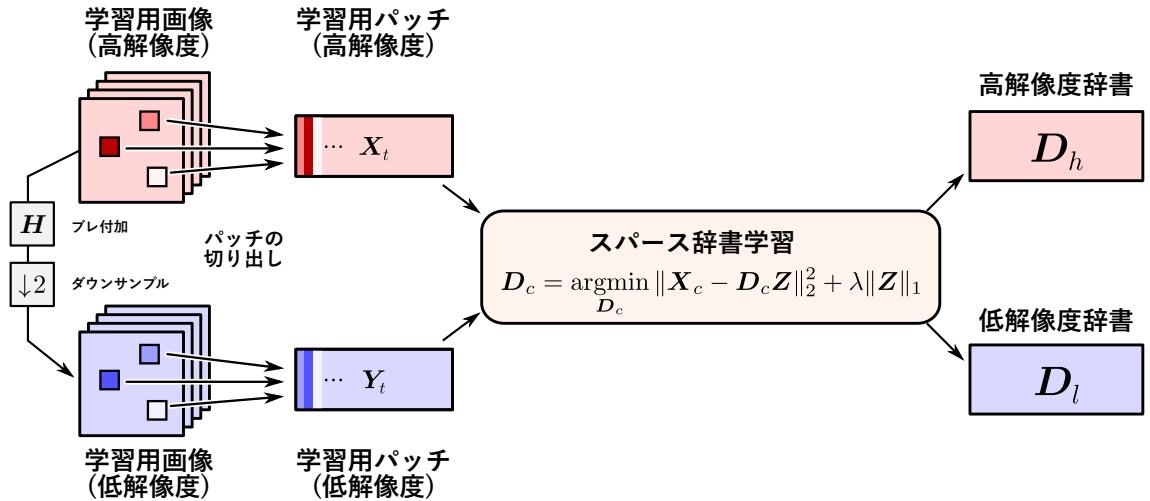


図 2.11 学習用画像を用いた低解像度・高解像度辞書の学習

式 (2.58) は、以下のプロセスで解くことができる。

1. D_c, Z を正規乱数で初期化する。
2. D_c を固定して、 Z を以下の式で求める

$$Z = \operatorname{argmin}_Z \|X - DZ\|_2^2 + \|Z\|_1 \quad (2.60)$$

3. Z を固定して D_c を以下の式で求める

$$D_c = \operatorname{argmin}_{D_c} \|X - D_c Z\|_2^2 \quad (2.61)$$

4. 過程 2, 3 を収束するまで繰り返す

2.5.7 カラー画像に対する処理

辞書ベース超解像は、グレースケールに対して成り立っているが、RGB 成分からなるカラー画像に対しても適用できる。これは、図 2.12 の様に RGB 成分を輝度 Y と色差 Cb, Cr に変換した後、輝度成分 Y に辞書ベース超解像を適用、色差成分 Cb, Cr に Bicubic 補間を適用し、輝度・色差成分を RGB 成分に逆変換することで実現できる。これは人間の知覚が輝度成分の変化には敏感で、色差成分の変化には鈍感である性質を利用したものである。

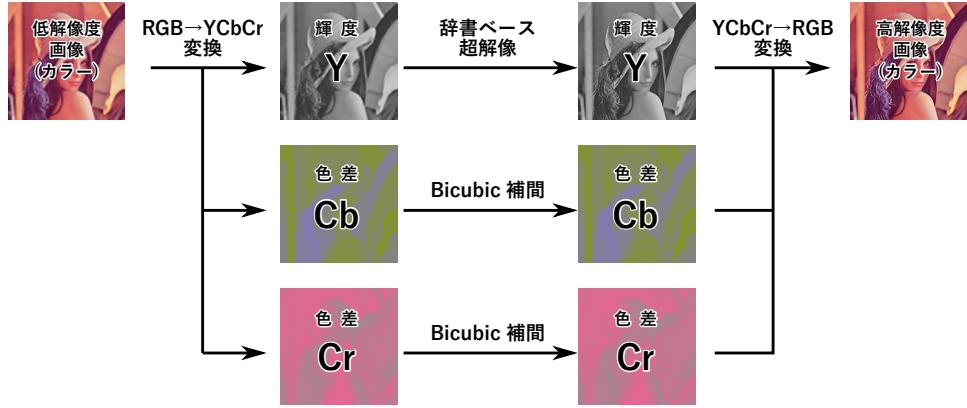


図 2.12 カラー画像に対する辞書ベース超解像処理

2.6 従来のぶれ除去手法

本節では画像の劣化除去として、モーションぶれや焦点ぼけを取り除くぶれ除去手法の既存手法について述べる。ぶれ除去技術は、ぶれカーネルを既知とするか未知とするかによって Non-blind deconvolution と Blind deconvolution に分かれ、Blind deconvolution はぶれカーネルの軌跡に想定を加えるパラメトリックな手法と、ノンパラメトリックな手法に分かれる。本節では、Non-blind deconvolution と Blind deconvolution の問題設定の違いを述べた後、ノンパラメトリックな Blind deconvolution 手法の一般的なフレームワークについて述べる。その後に、ノンパラメトリックな推定手法の中でも高速なカーネル推定を実現した Cho の Blind deconvolution アルゴリズム [53] についての説明を行う。

2.6.1 Non-blind Deconvolution と Blind Deconvolution

2.2.3 節で述べたとおり、ぶれ画像 \mathbf{Y} と、ぶれのない画像 \mathbf{X} 、カーネル \mathbf{K} の間には以下の 2 次元畳み込みの関係が成り立つ。

$$\mathbf{Y} = \mathbf{X} \otimes \mathbf{K} \quad (2.62)$$

ぶれ除去問題は \mathbf{X}, \mathbf{K} のうち、 \mathbf{K} を未知とするか、既知とするかによって、Non-blind Deconvoltion と Blind Deconvolution の 2 つに分類することができる。

Non-blind Deconvolution は \mathbf{K} が既知であるため、 \mathbf{X} を未知として \mathbf{Y} を $\mathbf{K} \otimes \mathbf{X}$

間の 2 乗誤差を最小化する問題として以下のように定義される.

$$\mathbf{X} = \operatorname{argmin}_{\mathbf{X}} \|\mathbf{Y} - \mathbf{K} \otimes \mathbf{X}\|_2^2 \quad (2.63)$$

\mathbf{K} を既知とみなせるケースは、カメラの焦点特性が分かっている時や、超解像問題をぶれカーネルを指定して解く場合が挙げられる。式 (2.63) は後述の Blind Deconvolution に比べて解きやすい問題ではあるが、ぶれ軌跡 \mathbf{K} を予め知っておく必要があるため、使える状況が限られる。

Blind Deconvolution は (2.7) の畳み込みを満たす \mathbf{X}, \mathbf{K} を、共に未知として推定する問題であり、以下のように表現することができる。

$$\{\mathbf{X}, \mathbf{K}\} = \operatorname{argmin}_{\{\mathbf{X}, \mathbf{K}\}} \|\mathbf{Y} - \mathbf{K} \otimes \mathbf{X}\|_2^2 \quad (2.64)$$

しかしこの式は、 \mathbf{X}, \mathbf{K} 共に条件が大きく欠落した劣決定問題のため、一意の解を得ることができない。ブレ除去の観点から言いかえれば、(2.64) の関係を満たす \mathbf{X}, \mathbf{K} の組み合わせは無数に存在するということになる。このことから、Blind Deconvolution は Nonblind に比べて解きづらいが、未知のモーションブレ以外にも焦点ボケ等を含むぶれ画像にも対応することができる。

問題の劣決定性を解消するために、Blind deconvolution 手法の多くは \mathbf{X}, \mathbf{K} のどちらか片方を既知とし、 \mathbf{K} を求めるカーネル推定問題 (2.65)、 \mathbf{X} を求める逆畳み込み問題 (2.66) を交互に解くことで、もっともらしいカーネル・補正画像の組を得る。

[53] の手法では Blind Deconvolution 問題を以下の 2 つの部分問題に分けて解く。

$$\mathbf{K} = \operatorname{argmin}_{\mathbf{K}} \|\mathbf{Y} - \mathbf{K} \otimes \mathbf{X}\|_2^2 + \beta \rho_k(\mathbf{K}) \quad (2.65)$$

$$\mathbf{X} = \operatorname{argmin}_{\mathbf{X}} \|\mathbf{Y} - \mathbf{K} \otimes \mathbf{X}\|_2^2 + \alpha \rho_x(\mathbf{X}) \quad (2.66)$$

すなわち、

1. ブレ画像 \mathbf{Y} と画像 \mathbf{X} を用いて、式 (2.65) からカーネルの推定値 \mathbf{K} を得る。
2. ブレ画像 \mathbf{Y} とカーネル \mathbf{K} を用いて、式 (2.66) から逆畳み込みした画像 \mathbf{X} を得る。
3. これによって得た画像を 1. のカーネル推定に再度利用して、 \mathbf{K} の精度を上げる。
4. 1~3 を何回か繰り返して最終的な推定カーネル \mathbf{K} とする。

というプロセスを経ることで、最終的なカーネル \mathbf{K} と、それを用いて逆畳み込みした出力画像 \mathbf{X} を得る。

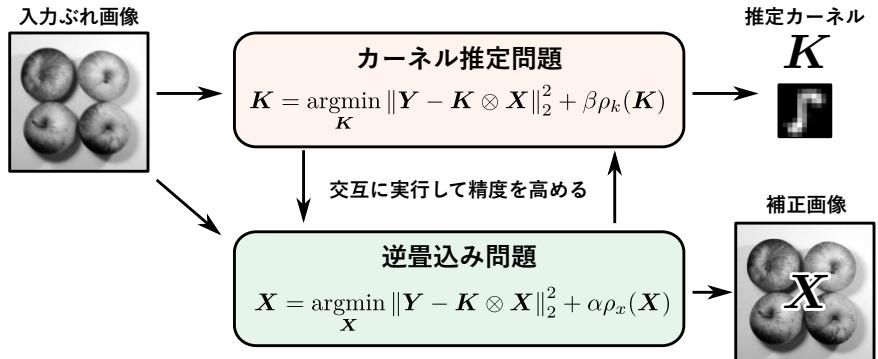


図 2.13 典型的な Blind Deconvolution のフレームワーク

2.6.2 カーネル推定問題のための画像の前処理

Blind Deconvolution のカーネル推定問題 (2.65)において、画像 \mathbf{X} は未知であるため、入力のブレ画像 \mathbf{Y} より何かしらの方法で擬似的な \mathbf{X} を推定しなければいけない。従来法では、ブレ画像よりしきい値処理にてエッジ成分を抽出した画像を作成し、カーネル推定問題の入力 \mathbf{X} として用いている。この擬似的な \mathbf{X} を生成するためのプロセスについて説明する。

従来法 [53] では、まず入力ブレ画像 \mathbf{Y} に対して Bilateral Filter [96, 97] を用いてブレ画像中の細かなノイズ、ブレ成分・テクスチャ成分を取り除く。ただし、このフィルタ操作によって本来抽出すべき画像のエッジ成分が弱まる問題があるため、これを補うために Shock Filter [98] を用いて失われたエッジ成分の復元を行う。これらのフィルタ処理を経た画像は図 2.14 の様に、画像の細かなブレ成分が消え、主要なエッジ成分が強調された画像となる。

従来法では、さらにここから、画像の勾配 $\partial_x \mathbf{X}, \partial_y \mathbf{X}$ に対しきい値処理を行うことで、画像のエッジ成分のみを残す画像 \mathbf{X} を生成する。具体的には図 2.15 の様に、画像の x 方向、 y 方向の微分画像を生成し、方向成分ごとに定めたしきい値を下回る勾配は 0 とする。さらに、これによって得られた $\partial_x \mathbf{X}, \partial_y \mathbf{X}$ を元に、元の画像 \mathbf{X} を生成し、カーネル推定プロセスの入力とする。2 回目以降のカーネル推定プロセスでは、ブレ画像 \mathbf{Y} の代わりに前段の逆畳み込みプロセスで得た補正画像 \mathbf{X} に対して同様の前処理フィルタリングとしきい値処理を行う。

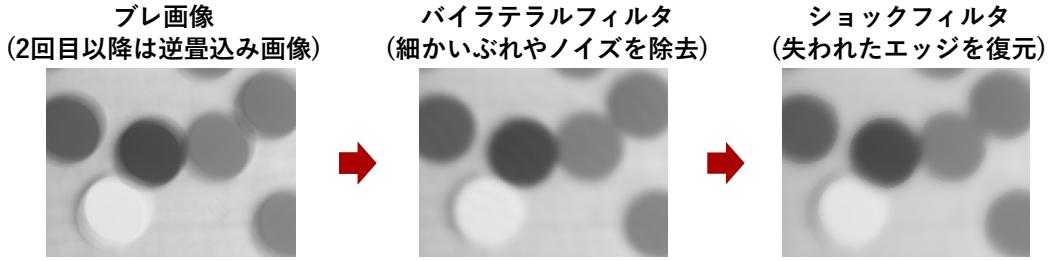


図 2.14 従来法のカーネル推定のための画像の前処理

2.6.3 カーネル推定問題

前節で前処理を行った補正画像 \mathbf{X} に対して、カーネル推定の非線形最適化問題を解き、カーネルの推定値を得る。従来法 [53] では高速化のために、式 (2.65) において、カーネルの制約項に L_2 ノルムを用いた以下の問題を用いる。

$$\mathbf{k} = \underset{\mathbf{k}}{\operatorname{argmin}} \sum_{*} w_* \|\mathbf{y}_* - \mathbf{A}_* \mathbf{k}\|_2^2 + \beta \|\mathbf{k}\|_2^2 \quad (2.67)$$

ここで \mathbf{k}, \mathbf{y} は求めたいカーネル、ぶれ画像のベクトル表現、記号 $*$ は画像に対して $\{\partial_x, \partial_y, \partial_{xx}, \partial_{yy}, \partial_{xy}\}$ の 5 種類の微分をかける微分方向、 \mathbf{y}_* はぶれ画像 \mathbf{y} に対して $*$ 方向の微分をかけた画像、 \mathbf{A}_* は補正画像 \mathbf{x} に対して $*$ 方向の微分をかけた画像の畳み込み行列表現である。各方向成分 $*$ は w_* によって重み付けされる。

上式は目的関数の \mathbf{k} に関する勾配を $\mathbf{0}$ とおくことで解くことができる。

$$\sum_{*} w_* \mathbf{A}_*^T (\mathbf{A}_* \mathbf{k} - \mathbf{y}_*) + \beta \mathbf{k} = \mathbf{0} \quad (2.68)$$

$$\therefore \mathbf{k} = \left(\sum_{*} w_* \mathbf{A}_*^T \mathbf{A}_* + \beta \mathbf{I} \right)^{-1} \left(\sum_{*} w_* \mathbf{A}_* \mathbf{y}_* \right) \quad (2.69)$$

行列 $\sum_{*} w_* \mathbf{A}_*^T \mathbf{A}_* + \beta \mathbf{I}$ は正定値行列であるため、式 (2.69) の逆行列計算は共役勾配法を用いて解くことになる。すなわち、 \mathbf{y} および \mathbf{A} からカーネル \mathbf{k} を得るプロセスは Algorithm 2 で表される。このアルゴリズムの 6 行目にある $\sum_{*} w_* \mathbf{A}_*^T \mathbf{A}_* + \beta \mathbf{I}$ の部分は **for** ループの外側で予め計算しておくことが可能である。

カーネルの制約項には重み β がかかる。式 (2.65) の二乗誤差の項は過決定な問題となるため、 β は解のエネルギーを最小化し、収束を安定させる値に調整する。

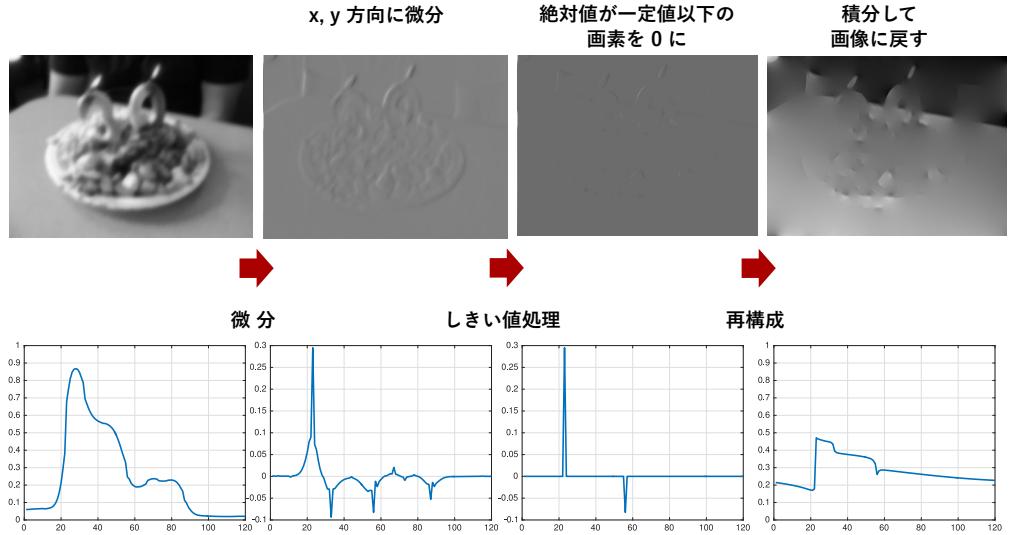


図 2.15 従来法の画像のしきい値処理

2.6.4 逆畳み込み問題

前節で推定したカーネルを元に、逆畳み込みを行い、補正画像を得る。この補正画像は再び先のカーネル推定問題の入力となり、より精度の高いカーネル推定が行われる。カーネル推定の時点ではそれほど高精度な逆畳み込み手法は要求されず、繰り返し実行されるために、高速なアルゴリズムが要求される。手法 [50, 53, 99] では以下の式を用いて逆畳み込みを行っている。

$$\mathbf{X} = \underset{\mathbf{X}}{\operatorname{argmin}} \sum_{\partial_*} \omega_* \|\partial_* \mathbf{Y} - \mathbf{K} \otimes \partial_* \mathbf{X}\|_2^2 + \alpha \|\Delta \mathbf{X}\|_2^2 \quad (2.70)$$

$\partial_* \in \{\partial_x, \partial_y, \partial_{xx}, \partial_{yy}, \partial_{xy}\}$ は、画像 \mathbf{X} の x 方向、 y 方向に 1 階微分、2 階微分をかける演算子である。また、 Δ は画像 \mathbf{X} に x 方向、 y 方向に 1 階微分をかける演算子である。

Algorithm 2 従来法のカーネル推定プロセス

1. 初期化 $\mathbf{k} = \mathbf{k}_0$
 2. $\mathbf{r} = (\sum_* w_* \mathbf{A}_* \mathbf{y}_*) - (\sum_* w_* \mathbf{A}_*^T \mathbf{A} + \beta \mathbf{I}) \mathbf{k}$
 3. $\mathbf{p} = \mathbf{r}$
 4. $e_0 = \mathbf{r}^T \mathbf{r}$
 5. **for** $i = 1$ **to** converge **do**
 6. $\mathbf{z} = (\sum_* w_* \mathbf{A}_*^T \mathbf{A} + \beta \mathbf{I}) \mathbf{p}$
 7. $k_{step} = \frac{e_{i-1}}{\mathbf{p}^T \mathbf{z}}$
 8. $\mathbf{k} = \mathbf{k} + k_{step} \mathbf{p}$
 9. $\mathbf{r} = \mathbf{r} - k_{step} \mathbf{z}$
 10. $e_i = \mathbf{r}^T \mathbf{r}$
 11. $\mathbf{p} = \frac{e_i}{e_{i-1}} \mathbf{p}$
 12. **end for**
-

2.7 Blind Deconvolution のための他のアプローチ

本節では、Cho の手法の他に提案法と比較を行った Blind deconvolution の手法についての説明を行う。

2.7.1 Fergus の手法

Fergus 他は変分ベイズ推定を用いたぶれ軌跡推定手法を提案している [47]。これにはカーネルと画像（差分信号） $\mathbf{K}, \Delta\mathbf{X}$ に関する事後確率最大化を行う。

$$p(\mathbf{K}, \Delta\mathbf{X} | \Delta\mathbf{Y}) \propto p(\Delta\mathbf{Y} | \mathbf{K}, \Delta\mathbf{X}) \cdot p(\Delta\mathbf{X}) \cdot p(\mathbf{K}) \quad (2.71)$$

尤度関数 $p(\Delta\mathbf{Y} | \mathbf{K}, \Delta\mathbf{X})$ にはガウス分布、 $\Delta\mathbf{X}$ の事前確率 $p(\Delta\mathbf{X})$ には混合ガウス分布、 \mathbf{K} の事前確率 $p(\mathbf{K})$ には指數混合分布を用いることで以下の表現を得る。

$$\begin{aligned} p(\mathbf{K}, \Delta\mathbf{X} | \Delta\mathbf{Y}) &\propto \prod_n \mathcal{N}(\Delta\mathbf{Y}(n) | \Delta\mathbf{X}(n) \otimes \mathbf{K}, \sigma^2) \\ &\cdot \prod_n \sum_{c=1}^C \pi_c \mathcal{N}(\Delta\mathbf{X}(n) | 0, \nu_c) \cdot \prod_m \sum_{d=1}^D \pi_d \mathcal{E}(\mathbf{K}(m) | \lambda_d) \end{aligned} \quad (2.72)$$

さらに事後確率 $p(\mathbf{K}, \Delta\mathbf{X} | \Delta\mathbf{Y})$ を 2 つの分布の積 $q(\mathbf{K}, \Delta\mathbf{X}) = q(\mathbf{K}) \cdot q(\Delta\mathbf{X})$ で近似する。 $q(\Delta\mathbf{X})$ にはガウス分布、 $q(\mathbf{K})$ にはガウス分布の正の値のみを用いる整流ガウス分布を用いる。最終的に、変分ベイズ推定を用いて、以下のカルバック・ライブラー距離を最小化することでカーネル推定を行う。

$$KL(q(\mathbf{K}, \Delta\mathbf{X}) \| p(\mathbf{K}, \Delta\mathbf{X} | \Delta\mathbf{Y})) \quad (2.73)$$

推定したカーネル \mathbf{K} を元に画像を逆畳み込みするアルゴリズムには Richardson-Lucy 法 [48, 49] を用いる。

2.7.2 Levin の手法

Levin 他の手法 [50] はカーネルと画像 \mathbf{K}, \mathbf{X} に関する事後確率最大化を行う

$$p(\mathbf{Y}, \mathbf{X}, \mathbf{K}) = p(\mathbf{Y} | \mathbf{X}, \mathbf{K}) \cdot p(\mathbf{X}) \cdot p(\mathbf{K}) \quad (2.74)$$

尤度関数に関しては、 y と $x \otimes k$ との誤差値がガウス分布に従うものとして、以下の様に定義する。

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{K}) = \frac{1}{(\sqrt{2\pi}\eta)^N} \exp\left(-\frac{1}{2\eta^2} \|\mathbf{K} \otimes \mathbf{X} - \mathbf{Y}\|^2\right) \quad (2.75)$$

画像の事前分布には、画像の縦、横の微分を混合ガウスモデル (MOG: Mixture of Gaussians) で近似したスペース近似を用いる。 γ は縦・横の微分方向を表し、 $f_{i,\gamma}(x)$ は \mathbf{X} を γ 方向に微分した画像の i 番目の画素を表す。

$$p(\mathbf{X}) = \prod_i \prod_\gamma \rho(f_{i,\gamma}(x)) \quad (2.76)$$

$$\rho(f_{i,\gamma}(x)) = \sum_j \frac{\pi_j}{\sqrt{2\pi}\sigma_j} e^{-\frac{1}{2\sigma_j^2} \|f_{i,\gamma}(x)\|^2} \quad (2.77)$$

また、カーネル \mathbf{K} の事前分布には一様分布を想定して制約を加えない。これらを統合した以下の式を用いることで画像 \mathbf{X} とカーネル \mathbf{K} の推定値を得る。

$$-\log p(\mathbf{Y}|\mathbf{X}, \mathbf{K}) = \frac{1}{2\eta^2} \|\mathbf{K} \otimes \mathbf{X} - \mathbf{Y}\|^2 - \sum_{i,\gamma} \log \rho(f_{i,\gamma}(x)) + c \quad (2.78)$$

2.7.3 Xu および Jia の手法

Xu および Jia の手法 [51, 52] ではそれぞれカーネル推定の手法と逆畳み込みの手法を提案している。カーネル推定には、ショックフィルタで処理した画像 \mathbf{X} に対して以下の式を最小化することで初期のカーネル \mathbf{K} を得る。

$$E(\mathbf{K}) = \|\Delta \mathbf{X} \otimes \mathbf{K} - \Delta \mathbf{Y}\|_2^2 + \gamma \|\mathbf{K}\|^2 \quad (2.79)$$

また、これによって得たカーネルに対し、Iterative support detection (ISD) 法 [100] によって補正を行う。

逆畳み込みでは画像の微分成分に対して L_0 制約をおいた以下の式を解く。

$$\min_{\mathbf{X}} \left\{ \frac{1}{\lambda} \|\mathbf{K} \otimes \mathbf{X} - \mathbf{Y}\|_2^2 + \sum_{*\in\{h,v\}} \sum_n \left\{ |l_{*,n}|^0 + \frac{1}{\varepsilon^2} (\partial_* x_n - l_{*,n})^2 \right\} \right\} \quad (2.80)$$

第 3 章

画像の超解像問題

3.1 本章の構成

本章では画像の劣化除去として、画像の低解像度化によって発生するぼけやサンプル行列の特性を取り除く超解像技術に焦点をあてる。辞書ベース超解像技術は高解像度画像の細かいテクスチャ部分等の復元性能が他の手法と比べて有利である一方、事前辞書が必要な欠点やスペース再構成の計算コストがかかる欠点が存在する。本研究はそれらの欠点を改善し、高速で事前情報に依存しない超解像手法を提案する。

本章の構成としては 3.2 節にて従来手法の問題点を述べそれに対する提案手法の解決策を述べる。3.3 節および 3.4 節にて、計算コストの改善および事前辞書の解消のための提案手法の詳細について述べる。3.5 節にて、従来法との超解像画像の比較を行い、提案手法の優位性について示す。

3.2 従来手法の問題点と提案手法の改良点

辞書ベースの従来法では、式 (2.58) の辞書学習を済ませた固定倍率の辞書をあらかじめシステムに用意しておく必要がある。これは、システム側に辞書を保存しておく容量を確保する必要がある。このため、パッチサイズの変化、辞書サイズの変化、拡大率の変化等のパラメータ変化に対応する辞書をすべて保持しておくことは難しい。また、多くのバリエーションを持つ画像に対して、固定の辞書を用いた場合、学習辞書の質に結果画像が依存する問題がある。

再構成問題においては、画像のパッチごとに式 (2.52) の最適化問題を解く必要がある。 L_1 制約の最適化は、行列の逆行列問題を含む 2 次計画法となるため、これを画像のパッチごとに個別に計算するのは非常に計算コストがかかる。

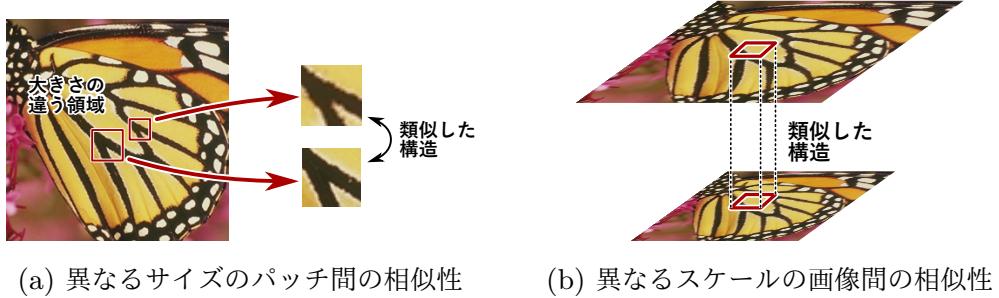


図 3.1 画像の自己相似性

事前辞書の問題と, L_1 再構成の計算コストを改善するために提案法では, 以下の 2 点において改良を行った.

- L_1 制約を L_2 制約に変更することでパッチ再構成の計算時間を短縮した
- 入力低解像度画像（自画像）から辞書ペアを動的に生成することで事前辞書を不要にした

高解像度パッチの再構成問題において, L_1 制約を L_2 制約に変更することで, 従来手法では二次計画法で解を得ていたものが, 閉形式の解で得られる. この閉形式の解は予め計算しておくことができる逆行列と低解像度パッチとの行列積で計算することができるため, 従来手法と比較して計算コストを削減することができる. また, 入力低解像度画像から辞書ペアを動的に生成することで, 従来法の, 事前辞書が必要である点を解消し, パッチサイズ, 辞書サイズ等に対して柔軟に辞書を構築できるようにする. 従来手法では, 学習用の高解像度・低解像度画像の組み合わせから, 高解像度・低解像度の対応関係を学習しているが, 提案手法では, 画像の自己相似性に基づいて, 入力低解像度画像一枚から同様の高解像度・低解像度の対応関係を取得する. 自己相似性とは, 画像を局所的に見たときに, 異なるスケール間で類似した構造を持つ特徴であり, 図 3.1 (a) の様に同じ画像から異なるサイズのパッチを取得したときや, 同図 (b) の様に異なるスケールに縮小した画像からパッチを取得した際にこれらのパッチ間に類似した構造が見られる. 提案手法の辞書ベース超解像では, 入力低解像度画像と, そこから更に縮小を行った画像からパッチをサンプルすることで, 画像の自己相似性の関係を抽出し, 超解像問題に適用する.

以後の節では, L_2 再構成を用いた高解像度パッチの再構成法と, 入力低解像度画像からの辞書の作成法について説明する.

3.3 L2 再構成問題による高速な再構成

再構成問題では、入力低解像度画像のパッチ \mathbf{y} は、低解像度辞書 \mathbf{D}_l の要素の線形結合で近似される。提案手法では式 (2.52) の最適化問題の L_1 ノルムを L_2 ノルムに変更する。

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{\lambda}{2} \|\boldsymbol{\alpha}\|_2^2 + \frac{1}{2} \|\mathbf{F}\mathbf{D}_l\boldsymbol{\alpha} - \mathbf{F}\mathbf{y}\|_2^2 \quad (3.1)$$

式 (3.1) の問題は以下の逆行列計算として容易に計算することができる。

$$\boldsymbol{\alpha} = ((\mathbf{F}\mathbf{D}_l)^T(\mathbf{F}\mathbf{D}_l) + \lambda\mathbf{I})^{-1} (\mathbf{F}\mathbf{D}_l)^T \mathbf{F}\mathbf{y} \quad (3.2)$$

上式によって得た係数ベクトル $\boldsymbol{\alpha}$ と高解像度辞書 \mathbf{D}_h を用いて高解像度パッチは以下の式で再構成できる。

$$\mathbf{x} = \mathbf{D}_h \boldsymbol{\alpha} \quad (3.3)$$

式 (3.2) のうち

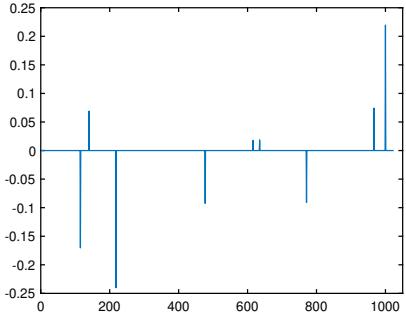
$$((\mathbf{F}\mathbf{D}_l)^T(\mathbf{F}\mathbf{D}_l) + \lambda\mathbf{I})^{-1} (\mathbf{F}\mathbf{D}_l)^T \quad (3.4)$$

の部分は入力低解像度パッチ \mathbf{y} に依存しないので、入力パッチの処理前に一度だけ計算しておけば良い。そのため、提案法の係数 $\boldsymbol{\alpha}$ 導出のための主な計算コストは式 (3.4) のサイズ $K \times K$ の行列の逆行列計算 1 回と、(3.2) の計算において、サイズ $K \times 4n$ の行列の掛け算がパッチ数の分必要になる。行列の掛け算は、(2.52) の L_1 最適化問題をパッチの数だけこなすのに比べて低い計算コストですむため、 L_2 ノルムへの変更によってアルゴリズム全体の処理時間を削減することができる。

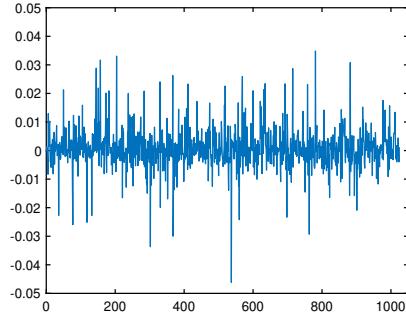
図 3.2 に L_1 再構成を行った時と L_2 再構成を行った時に得られる係数 $\boldsymbol{\alpha}$ の違いを示す。 L_1 制約を L_2 制約に変更することで、 $\boldsymbol{\alpha}$ のスパース性がなくなり、非ゼロの係数がベクトル中に万遍なく配置された形となる。提案法では、辞書ベース超解像において、 $\boldsymbol{\alpha}$ のスパース係数がなくても問題なく高解像度パッチを復元できることを示した。

3.4 自画像からのサンプリングによる辞書作成

超解像問題の再構成のために、事前辞書 $\mathbf{D}_h, \mathbf{D}_l$ の必要性を解消し、また事前辞書の質によって再構成画像の精度が左右される問題点を解決するために、本手法では、入力低解像度 \mathbf{Y} からのパッチサンプルによって辞書 $\mathbf{D}_h, \mathbf{D}_l$ を作成する方法を提案する。以降パッチサンプルに用いる入力低解像度画像自身のことを自画像と定義する。



(a) L_1 最小化によって得られる係数



(b) L_2 最小化によって得られる係数

図 3.2 L_1 と L_2 最小化によって得られる係数 α の違い

自画像からのパッチサンプルによる辞書作成の流れを図 3.3 に示す。まず入力低解像度画像 \mathbf{Y} から、与えられた拡大率 s に基づいて縮小を行い、更に低解像度の画像 \mathbf{Z} を作成する。さらにサンプルされる辞書のパターンを増やすために、 \mathbf{Y}, \mathbf{Z} に関して、オリジナルの画像の他に反転 2 種、回転 4 種、輝度反転 2 種を含む計 16 種類の変換画像を作る。ここから、ランダムな座標を選び、 \mathbf{Y} から切り出したパッチを高解像度辞書 \mathbf{D}_h の要素、 \mathbf{Z} の対応する座標から切り出したパッチを \mathbf{D}_l の要素として、辞書の対を構成する。これによって作成した低解像度辞書 \mathbf{D}_l には 2.5.2 節の特徴抽出行列 \mathbf{F} をかけて \mathbf{FD}_l とし、式 (3.4) の再構成問題に使用する。この操作によって自画像から高解像度画像 \mathbf{X} 再構成に必要な辞書 $\mathbf{D}_h, \mathbf{D}_l$ を大量の画像から予め学習させておく必要がなくなる。

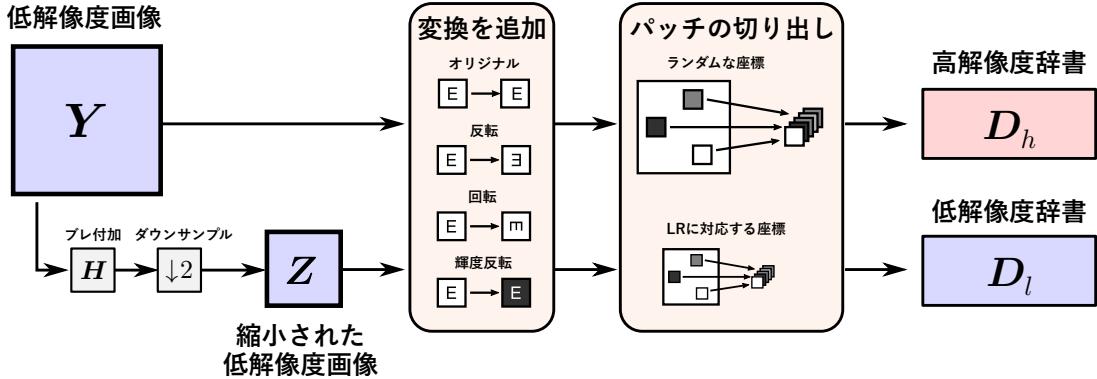


図 3.3 自画像からのパッチサンプルによる辞書作成

3.5 実験と評価

提案手法による辞書ベース超解像の優位性を示すために、サイズ 512×512 の標準画像 24 枚（図 3.4）を用いて評価を行った。

512×512 の画像を $1/2$ 倍の 256×256 に縮小、超解像アルゴリズムを用いて 2 倍の 512×512 に拡大したものを原画像と比較する。評価は PSNR による定量的評価と、結果画像と原画像を比較した視覚評価を行った。また、 L_2 再構成の計算時間評価のためにサイズ 128 から 1024 までの画像サイズに関して超解像を行ったときの計算時間を測定した。

比較対象の手法は標準の Bicubic 法、 L_1 ベースの超解像手法 [37] の他に、自画像辞書と L_2 再構成それぞれの効果を評価するため、自画像からサンプルした辞書を元に L_1 再構成を行う手法と、学習済みの辞書を元に L_2 再構成を行う手法の評価を行った。

パッチの要素数 n および辞書のサイズ K は従来法と提案法とともに $n = 5 \times 5 = 25$ 、 $K = 1024$ に設定した。従来法において、式 (2.52) の λ の値は 0.2 に、提案法において、式 (3.1) の λ の値は 0.001 に設定した。 λ の値は本テスト画像において良好な結果を示す値に調整した。

本手法の実装は MATLAB を用いて行い、計算時間評価は CPU Intel Core i7 4GHz、メモリ 24 GB 1867 MHz の環境で行った。



図 3.4 実験に用いる標準画像 24 枚

3.5.1 2倍拡大の PSNR 比較

表 3.1 に超解像 4 手法の PSNR 比較結果を示した。いずれの手法も Bicubic 法と比べると高い超解像性能を保っている。画像再構成精度は L_1 再構成の精度が最も優れているものの、 L_2 再構成では、 L_1 再構成に比べて大きく処理時間を短縮する一方、大きく劣化しない精度で再構成が可能である。

また、事前辞書の代わりに自画像からサンプルした辞書を用いる場合でも、 L_1 , L_2 再構成とともに PSNR を 0.1 dB から 0.2 dB 程度低下させるのみで、画質に大きな影響を与えていないことが分かる。

3.5.2 実行時間の比較

表 3.2 に、超解像 4 手法の実行時間の比較結果を示した。事前辞書を用いた L_2 再構成の手法は、従来法（事前辞書を用いた L_1 再構成）と比べて大幅に計算時間を削減している。これは、式 (3.1) によって、 L_1 再構成を L_2 再構成問題に緩和したことによって、パッチごとに線形計画問題を解く必要がなくなったことによる短縮である。提案手法（自画像辞書を用いた L_2 再構成）は、事前辞書を用いた L_2 再構成問題より 2.6 秒ほど多く計算時間がかかる。これは自画像からのサンプル時間によるものであるが、この時間を含めても L_1 再構成問題よりも高速に処理が可能である。

3.5.3 画像サイズ・辞書サイズによる実行時間の変化

図 3.5 に画像サイズ及び辞書サイズを変化させた時の提案法の処理時間計測結果を示した。また、図 3.6 に辞書サイズが 1024 の時の、提案法の各プロセスにかかる処理時間を示した。辞書のサイズ K は自画像辞書の作成時間および、式 (3.4) の逆行列算出時間に影響するが、逆行列算出は係数 α 導出前に 1 度だけ計算すれば良いので全体の計算時間に比べると非常に僅かな計算コストですむ。主な計算時間の差は画像サイズによるものであり、自画像からの辞書のサンプリングにかかる時間、係数 α の導出、高解像度パッチの再構成にかかる時間に影響する。いずれの処理時間も画像の画素数、すなわち画像の 1 辺のサイズの 2 乗に比例して計算時間が増加することが分かる。

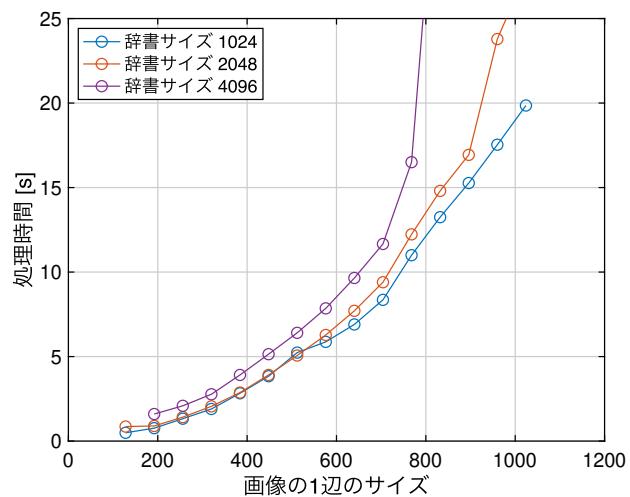


図 3.5 画像サイズと辞書サイズによる実行時間の変化

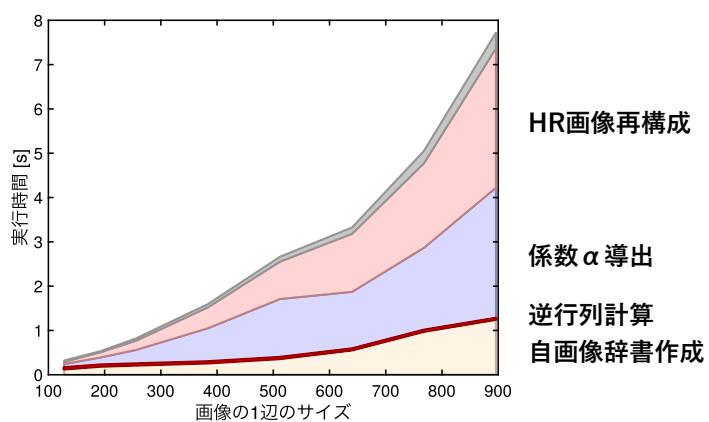


図 3.6 辞書サイズ 1024 時の実行時間の詳細

表 3.1 2 倍拡大画像の PSNR 比較 (単位 : dB)

画像名	Bicubic	従来法	自画像辞書	事前辞書	提案法
			L_1 再構成	L_2 再構成	
Aerial	24.742	26.009	25.782	25.658	25.658
Airplane	33.357	37.028	36.734	36.353	36.075
Barbara	25.210	27.132	26.984	27.040	27.045
Boat	29.904	31.498	31.038	30.928	30.893
Bridge	26.600	27.704	27.643	27.483	27.408
Building	25.950	26.973	26.925	26.792	26.776
Clown	32.766	35.091	34.880	34.628	34.646
Elaine	33.074	33.631	33.580	33.535	33.482
Finger	28.706	30.878	30.984	31.045	31.140
Finger2	30.639	33.047	32.902	33.089	33.029
Goldhill	31.389	32.510	32.364	32.208	32.218
Grass	27.763	29.216	29.075	29.107	29.104
Houses	23.682	25.143	24.934	24.681	24.703
Lena	32.953	37.480	37.186	36.951	37.022
Man	27.344	28.647	28.461	28.128	28.139
Mandril	23.263	25.968	25.923	25.865	25.884
Milkdrop	37.007	42.231	40.650	40.184	39.581
Nw-town	27.611	29.010	28.780	28.456	28.450
Pepper	32.222	36.936	36.397	35.948	35.823
Room	28.364	30.592	31.184	30.996	30.953
Station	29.581	32.075	32.153	31.826	32.110
Tank	32.850	33.745	33.608	33.619	33.519
Tank2	31.166	31.959	31.926	31.947	31.939
Watch	31.382	34.011	33.201	33.043	32.853
平均	29.480	31.605	31.387	31.230	31.185

表 3.2 2 倍拡大の実行時間比較 (単位 : 秒)

画像名	従来法	自画像辞書	事前辞書	提案法
		L_1 再構成	L_2 再構成	
Aerial	122.239	127.573	2.246	4.804
Airplane	131.274	132.876	2.397	4.912
Barbara	126.211	123.887	2.374	4.939
Boat	124.841	143.347	2.436	4.915
Bridge	128.061	129.478	2.429	4.895
Building	138.446	140.774	2.416	4.916
Clown	132.690	118.496	2.405	4.919
Elaine	124.587	126.865	2.378	4.935
Finger	128.238	142.195	2.381	5.204
Finger2	136.111	134.351	2.377	5.068
Goldhill	144.568	137.406	2.382	4.998
Grass	129.745	121.577	2.397	4.932
Houses	128.000	128.435	2.370	5.021
Lena	141.120	135.240	2.407	4.954
Man	143.691	134.479	2.395	4.955
Mandril	147.946	158.934	2.377	4.948
Milkdrop	142.401	139.635	2.392	5.011
Nw-town	128.264	106.811	2.397	4.932
Pepper	141.893	122.287	2.373	4.961
Room	126.123	95.950	2.394	4.929
Station	136.224	120.715	2.383	4.913
Tank	131.194	109.238	2.369	4.923
Tank2	155.508	143.778	2.394	4.930
Watch	125.533	121.217	2.377	4.935
平均	133.954	128.981	2.385	4.952

3.5.4 画像の視覚評価

図 3.7 に画像 Airplane の再構成結果を示す。 L_1 再構成、 L_2 再構成とともに文字の部分を高解像度化できている一方で低解像度画像に起因する文字の変形が両手法ともに見られる。また自画像から作成した辞書を用いた再構成では、文字の下の部分にアーティファクトが発生している。図 3.8 に示した画像 Houses の再構成結果でも自画像辞書では文字の下あたりにアーティファクトが発生し、これが PSNR 低下の原因となっていると考えられる。

図 3.9 に画像 Tank の再構成結果を示す。この画像は全体的に目立ったエッジが少ない画像であるが、自画像辞書を用いた再構成では星の部分等のエッジ成分を再構成できていることが分かる。

図 3.10 に示した画像 Room の再構成結果では、椅子のカバー等のエッジ部分をいずれの手法も再現できている。一方で画像左上のケーブル部分では、自画像辞書を用いた手法はブロック上のアーティファクトを発生させた。

3.6 まとめと今後の展望

本章では、従来の超解像手法における L_1 再構成の計算コストと、事前辞書を予め用意しておくことの欠点に着目して、 L_2 再構成と自画像辞書を用いた手法を提案した。

提案手法は、従来手法と同様に、低解像度パッチを低解像度辞書の線形結合で近似し、その近似係数を利用して高解像度パッチの再構成を行う。この係数導出の際に提案法は L_1 ノルム最小化問題を L_2 ノルム最小化問題に変更した。これによって係数算出問題が閉形式の解となり、辞書サイズに依存する逆行列計算の回数を 1 回に減らすことができる。また、提案手法の計算時間に影響する処理は、パッチ再構成問題に関する係数 α 導出に行列積 1 回高解像度パッチ x 再構成に行列積 1 回となり、二次計画法が必要な L_1 再構成に比べて低コストで計算が可能である。

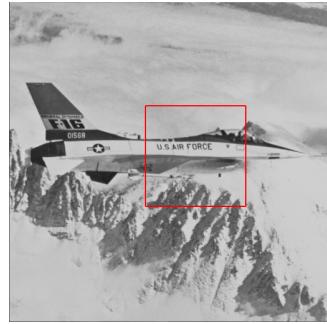
また、提案手法では自画像からのサンプルによって高解像度・低解像度の辞書のペア D_h, D_l を動的に作成する手法を提案した。これによって辞書サイズ・パッチサイズ・拡大率等のパラメータが異なる辞書を予め用意しておく必要がなくなり、柔軟に辞書を構築できる。

提案する超解像手法は、従来の L_1 再構成と事前学習した辞書を用いた手法と比べて若干の PSNR 低下が確認されたが、実行時間の面では L_2 再構成は L_1 再構成と比べて

大幅に計算コストの削減が可能である。実験結果では PSNR の低下は 0.3 dB 前後と比較的少ない低下量であるのに対し、計算コストの削減量は 512×512 の画像で 30 倍前後と計算コスト削減の利点は非常に大きいと考えられる。

一方で自画像からのサンプルされた辞書は、画像の種類によっては、縦横成分や、テクスチャ成分など、特定の種類のパッチのみが辞書にサンプルされ、画像中に少ないエッジ成分等にアーティファクトを発生させる結果が確認された。これを解決するためには、自画像のサンプル時に辞書の要素の偏りをなくし、方向成分・エッジ・テクスチャ部をバランスよく保持する自画像辞書の構築法の確立が望まれる。

また、自画像からのパッチのサンプル自体にかかる計算コストが画像の画素数に比例し、特に大きなサイズの画像を処理する時に辞書作成の計算コストが無視できなくなる問題が確認された。このため、辞書作成にあたり、画像の画素数が大きい時でも、画像中から分散の高いパッチ群をサンプルする手法の確立が課題となる。



(a) 拡大箇所



(b) 原画像



(c) 低解像度画像



(d) Bicubic 法



(e) 従来法 (事前辞書, L_1)



(f) 自画像辞書, L_1



(g) 事前辞書, L_2



(h) 提案法 (自画像辞書, L_2)

図 3.7 画像 Airplane の処理結果



(a) 拡大箇所



(b) 原画像



(c) 低解像度画像



(d) Bicubic 法



(e) 従来法 (事前辞書, L_1)



(f) 自画像辞書, L_1

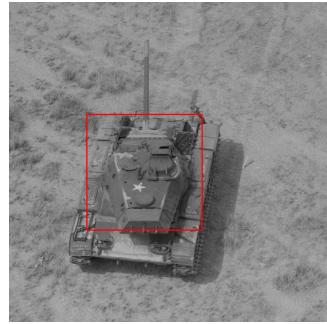


(g) 事前辞書, L_2



(h) 提案法 (自画像辞書, L_2)

図 3.8 画像 Houses の処理結果



(a) 拡大箇所



(b) 原画像



(c) 低解像度画像



(d) Bicubic 法



(e) 従来法 (事前辞書, L_1)



(f) 自画像辞書, L_1

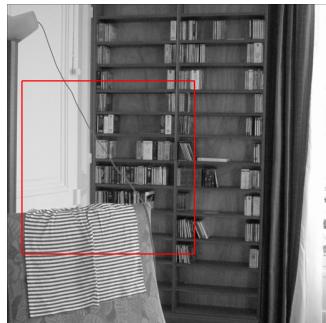


(g) 事前辞書, L_2



(h) 提案法 (自画像辞書, L_2)

図 3.9 画像 Tank の処理結果



(a) 拡大箇所



(b) 原画像



(c) 低解像度画像



(d) Bicubic 法



(e) 従来法 (事前辞書, L_1)



(f) 自画像辞書, L_1



(g) 事前辞書, L_2



(h) 提案法 (自画像辞書, L_2)

図 3.10 画像 Room の処理結果

第 4 章

画像のぶれ除去問題

4.1 本章の構成

本章では画像の劣化除去として、撮像機器の動きによるモーションぶれや撮像機器の光学特性、被写体との焦点距離のずれに基づく焦点ぼけを取り除く、ぶれ除去に焦点を当てる。2章で述べたとおり、ぶれ除去技術は、ぶれ軌跡の想定を加えるかどうかで、Non-blind deconvolution と Blind deconvolution に分けられ、Blind deconvolution はぶれカーネルの形状に想定を加えるかどうかでパラメトリックな手法とノンパラメトリックな手法に分類される。本研究はより実用的なアプリケーションへの応用を想定して、ノンパラメトリックな Blind deconvolution 手法に改良を加える。

2.6 節にて述べた Cho の手法 [53] は、最小二乗問題に基づくカーネル推定を高速化したアルゴリズムであり、実行速度面と精度の両立面で他の手法より優れているため、本研究ではこの手法をベースに実行速度のさらなる改善と補正画像の精度向上を図った。

本章では 4.2 節にて従来法の問題点とそれに対する提案手法の改良点の概要を述べる。次に 4.3 節にて提案手法のカーネル推定の改良点を、4.4 節にて提案手法の逆畳み込み手法の改良点について詳細に述べ、4.5 節にて提案手法全体のフレームワークについて説明する。最後に 4.6 節にて従来手法との補正画像の比較を行い、提案手法の優位性を示す。

4.2 従来手法の問題点と提案手法の改良点

Cho の従来手法 [53] はカーネル推定を主にエッジ領域で行うために、カーネル推定に用いるデータ点が減少し、推定精度が悪化する問題点がある。また、逆畳み込み手法は、反復法の繰り返しごとに畳み込みの計算を含む手法が用いられており、計算時間が

かかる問題点がある。そこで提案手法では以下の改良を行った。

- 画像の前処理に特徴抽出フィルタを用いることによるカーネル推定の改善
- 高速化した差分逆畳み込みによる逆畳み込みの実行時間改善

カーネル推定に用いる特徴抽出フィルタは、ぶれ画像 \mathbf{Y} と前処理画像 \mathbf{X} に対して適用し、この特徴空間上でカーネル推定を行う。これによって、従来法のしきい値処理では取り切れないエッジやテクスチャ成分もカーネル推定に用いることができ、推定精度を改善することが期待される。また、カーネル推定の最適化の誤差関数を周波数上で表現することで冗長な FFT の計算を省き、処理時間を改善した。

逆畳み込み問題では差分逆畳み込みをぶれ除去問題に適用する方法を提案した。差分逆畳み込みは、本来の画像の代わりに、画像から短時間露光画像などのガイド画像を引き、その差分を逆畳み込みすることで、補正画像に生じるリング等を削減する手法である。提案法では、逆畳み込みの過程で、ガイド画像を動的に生成し、ガイド画像がない状況でも可能な差分逆畳み込み手法を考案した。また、逆畳み込みの最適化問題の誤差関数を周波数上で表現して冗長な FFT 計算時間を削減し、処理時間を大きく短縮している。

4.3 特徴抽出を用いたカーネル推定

4.3.1 カーネル推定のための前処理の改善

従来法はボケ画像・補正画像の前処理としてバイラテラルフィルタとショックフィルタを適用した後、画像に縦横の微分を適用し、しきい値処理を行う。このうちバイラテラルフィルタ処理としきい値処理のしきい値決定に大きな計算コストがかかる。そこで提案法では、高速なエッジ保存平滑化フィルタである Guided フィルタ [101] を用いた画像の前処理手法を提案する。Guided フィルタはバイラテラルフィルタと同様のエッジ保存平滑化フィルタの一つであり、入力ノイズ画像とともにガイド画像を必要とする。

提案法では、ガウスフィルタで大まかにノイズを取り除いた画像をガイド画像として用いることで、リング等を増幅させることなしに、画像の細かいブレやノイズ成分を取り除く処理を実現した。ガウスフィルタの標準偏差はサイズ 11×11 のカーネルに対して 3.0 に設定した。Guided フィルタは入力画像に対して 2 から 3 回かける。また、Guided フィルタ処理後の画像に対してショックフィルタをかけることで失われ

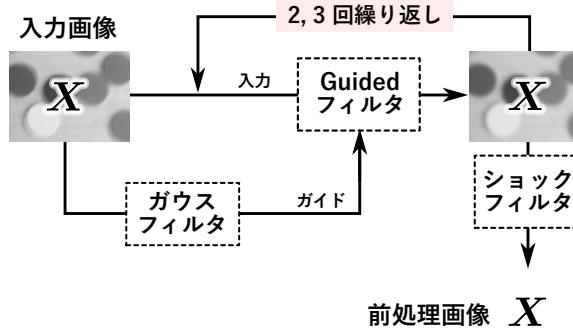


図 4.1 提案法の画像の前処理

たエッジ成分を復元させて、後段のカーネル推定問題に利用する。提案法の画像の前処理の流れを図 4.1 に示した。

4.3.2 特徴抽出フィルタを用いたカーネル推定処理

特徴抽出フィルタを \mathbf{P} とし、ブレ画像 \mathbf{Y} および一時的な補正画像 \mathbf{X} に対して、このフィルタを畳み込んだ $\mathbf{P} \otimes \mathbf{Y}$ および $\mathbf{P} \otimes \mathbf{X}$ に対して以下のカーネル推定問題を定義する。

$$\mathbf{K} = \underset{\mathbf{K}}{\operatorname{argmin}} \| \mathbf{P} \otimes \mathbf{Y} - \mathbf{K} \otimes \mathbf{P} \otimes \mathbf{X} \|_2^2 + \beta \| \mathbf{K} \|_2^2 \quad (4.1)$$

また、 \mathbf{P} の畳み込み行列表現を \mathbf{H} として、カーネルのベクトル表現 \mathbf{k} に対して以下の様に書き換える。

$$\mathbf{k} = \underset{\mathbf{k}}{\operatorname{argmin}} \| \mathbf{H} \mathbf{y} - \mathbf{H} \mathbf{A} \mathbf{k} \|_2^2 + \beta \| \mathbf{k} \|_2^2 \quad (4.2)$$

式 (4.1) を効率よく解くために提案手法では当式の二乗誤差項と制約項をそれぞれ周波数表現に置き換える。

$$\mathbf{f}_k = \underset{\mathbf{f}_k}{\operatorname{argmin}} \| \mathbf{F}_h \mathbf{f}_y - \mathbf{F}_h \mathbf{F}_x \mathbf{f}_k \|_2^2 + \frac{\beta}{N_1 N_2} \| \mathbf{f}_k \|_2^2 \quad (4.3)$$

$\mathbf{f}_k, \mathbf{f}_y$ はカーネルおよびぶれ画像の周波数特性 $\mathcal{F}(\mathbf{K}), \mathcal{F}(\mathbf{Y})$ の要素をそれぞれベクトル上に配置したもの、 $\mathbf{F}_h, \mathbf{F}_x$ は特徴抽出フィルタ、前処理画像の周波数特性 $\mathcal{F}(\mathbf{P}), \mathcal{F}(\mathbf{X})$ の要素を対角行列に配置したものである。フーリエ変換の線形性・直交性より $\| \mathbf{P} \otimes \mathbf{Y} - \mathbf{K} \otimes \mathbf{P} \otimes \mathbf{X} \|_2^2$ と $\| \mathbf{F}_h \mathbf{f}_y - \mathbf{F}_h \mathbf{F}_x \mathbf{f}_k \|_2^2$, $\| \mathbf{K} \|_2^2$ と $\| \mathbf{f}_k \|_2^2$ がそれぞれ比例す

Algorithm 3 提案法のカーネル推定アルゴリズム

1. **Initialize** \mathbf{f}_k to the vector of $\mathcal{F}(\mathbf{K}_0)$
 2. $\mathbf{r} = (\mathbf{F}_h \mathbf{F}_x)^H \mathbf{F}_h \mathbf{f}_y - ((\mathbf{F}_h \mathbf{F}_x)^H (\mathbf{F}_h \mathbf{F}_x) + \beta \mathbf{I}) \mathbf{f}_k$
 3. $\mathbf{p} = \mathbf{r}$
 4. $e_0 = \mathbf{r}^H \mathbf{r}$
 5. **for** $i = 1$ **to** converge e_i **do**
 6. $\mathbf{z} = ((\mathbf{F}_h \mathbf{F}_x)^H (\mathbf{F}_h \mathbf{F}_x) + \beta \mathbf{I}) \mathbf{p}$
 7. $k_{step} = e_{i-1} / (\mathbf{p}^H \mathbf{z})$
 8. $\mathbf{f}_k = \mathbf{f}_k + k_{step} \mathbf{p}$
 9. $\mathbf{r} = \mathbf{r} - k_{step} \mathbf{z}$
 10. $e_i = \mathbf{r}^H \mathbf{r}$
 11. $\mathbf{p} = (e_i / e_{i-1}) \mathbf{p}$
 12. **end for**
 13. **return** \mathbf{K} corresponding to \mathbf{f}_k
-

る。式(4.3)は \mathbf{f}_k に関して微分して $=\mathbf{0}$ とすることで解を得ることができる。

$$2(\mathbf{F}_h \mathbf{F}_x)^H (\mathbf{F}_h \mathbf{f}_y - \mathbf{F}_h \mathbf{F}_x \mathbf{f}_k) + \frac{2\beta}{N_1 N_2} \mathbf{f}_k = \mathbf{0}$$

$$\therefore \mathbf{f}_k = \left((\mathbf{F}_h \mathbf{F}_x)^H \mathbf{F}_h \mathbf{F}_x + \frac{\beta}{N_1 N_2} \mathbf{I} \right)^{-1} (\mathbf{F}_h \mathbf{f}_y) \quad (4.4)$$

式(4.4)は共役勾配法を用いて、Algorithm 3 の様に実行できる。

制約項 $\beta \|\mathbf{k}\|_2^2$ は従来法の式(2.69)と同じエネルギー最小化制約であり、従来法と同様に、収束を安定させる値に調整する。

4.3.3 特徴抽出フィルタの設計

提案手法で用いる特徴抽出フィルタ \mathbf{P} (行列表現 \mathbf{H}) の設計方法について説明する。提案法の特徴抽出フィルタは、ぶれ画像からエッジ成分を適切に抽出し、さらにリングングが発生しやすい周波数を抑制する。そのため、本手法のフィルタは、全方向に等しいカットオフ周波数を持つハイパスフィルタに微分フィルタを合成する形で設計を行う。

まずカットオフ周波数 f_ω のハニング窓を元に、窓関数法から次数 n_ω の1次元のFIR フィルタを作成する。カットオフ周波数 f_ω と次数 n_ω の値はそれぞれ $f_\omega = 0.2$,

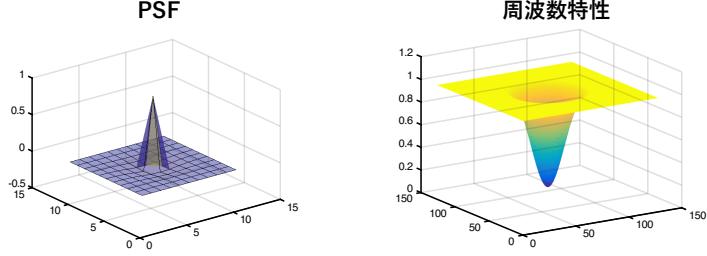


図 4.2 カットオフ周波数 f_ω のハイパスフィルタ

$n_\omega = 11$ とした. 次にこの 1 次元カーネルを以下のマクレラン変換行列を利用して円対称な 2 次元フィルタに変換する [102].

$$\mathbf{T}_M = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 2 & -4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (4.5)$$

この時点でのフィルタの PSF および周波数特性は図 4.2 のようになる. 提案法で用いる特徴抽出フィルタはこれに x および y 方向の微分カーネルを合成する. 微分カーネルは 1 次, 2 次の中間差分として以下の 5 種類を用いる.

$$\begin{aligned} \partial_x &= [1 \ 0 \ -1]^T, & \partial_y &= [1 \ 0 \ -1]^T, \\ \partial_{xx} &= [1 \ -2 \ 1]^T, & \partial_{yy} &= [1 \ -2 \ 1]^T, \\ \partial_{xy} &= \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \end{aligned} \quad (4.6)$$

図 4.2 のカーネルを \mathbf{P}_w とした時に, 提案法で用いる特徴抽出フィルタ \mathbf{P} は以下の式で合成される. w_1, w_2, w_3 は, 各微分成分の重みを表すパラメータで, 低次の微分成分と高次の微分成分のどちらを重視してカーネル推定を行うかを決定することができる. 提案法では経験的に良好なカーネル推定結果を得られる値として $w_1 = 0.5, w_2 = 1.4, w_3 = 1.4$ としている.

$$\mathbf{P} = (w_1 \partial_x + w_1 \partial_y + w_2 \partial_{xx} + w_2 \partial_{yy} + w_3 \partial_{xy}) \otimes \mathbf{P}_w \quad (4.7)$$

式 (4.7) で合成された提案法の特徴抽出フィルタは図 4.3 右側に示すような周波数特性となる.

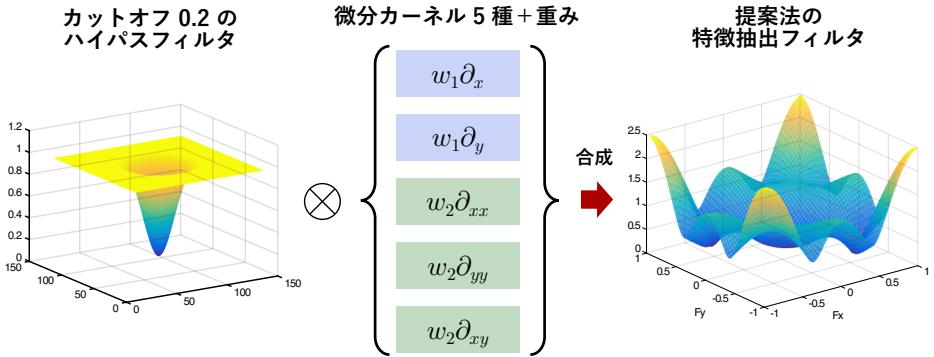


図 4.3 提案法で用いる特徴抽出フィルタの設計

4.4 差分逆畳み込み

4.4.1 差分逆畳み込みのプロセス

最終逆畳み込みのプロセスには、精度を高めるために Yuan の差分逆畳み込み [103] を改良して用いる。

差分逆畳み込みとはブレ画像 \mathbf{Y} を逆畳み込む際に、 \mathbf{Y} を直接処理するのではなく、 \mathbf{Y} とガイド画像 \mathbf{X}_G との残差 $\Delta\mathbf{Y} = \mathbf{Y} - \mathbf{X}_G \otimes \mathbf{K}$ に対して逆畳み込みを行い、補正画像 \mathbf{X} と \mathbf{X}_G の残差 $\Delta\mathbf{X} = \mathbf{X} - \mathbf{X}_G$ を得るというものである。ブレ画像と補正画像の残差 $\Delta\mathbf{Y}$, $\Delta\mathbf{X}$ の間には以下の関係が成立する。

$$\Delta\mathbf{Y} = \Delta\mathbf{X} \otimes \mathbf{K} \quad (4.8)$$

これは以下の過程から証明することができる。

$$\begin{aligned} \mathbf{X} \otimes \mathbf{K} &= (\mathbf{X}_G + \Delta\mathbf{X}) \otimes \mathbf{K} \\ &= \mathbf{X}_G \otimes \mathbf{K} + \Delta\mathbf{X} \otimes \mathbf{K} \\ &= \mathbf{X}_G \otimes \mathbf{K} + \Delta\mathbf{Y} \\ \therefore \Delta\mathbf{Y} &= \Delta\mathbf{X} \otimes \mathbf{K} \end{aligned} \quad (4.9)$$

差分逆畳み込みの目的は画像のエッジ部分のリングングの抑制にある。図 4.4 に 1 次元の例を示したが、ブレ画像とガイド画像の残差を取ることで、ブレ画像におけるエッジ近辺の画素値の変化が抑えられる。これを逆畳み込むと、元の画素値変動が大きい画像を直接処理した場合に比べ、リングングの幅が少なくなる。結果的に、ガイド画像を足して元の画像にした時のリングングが抑制される。

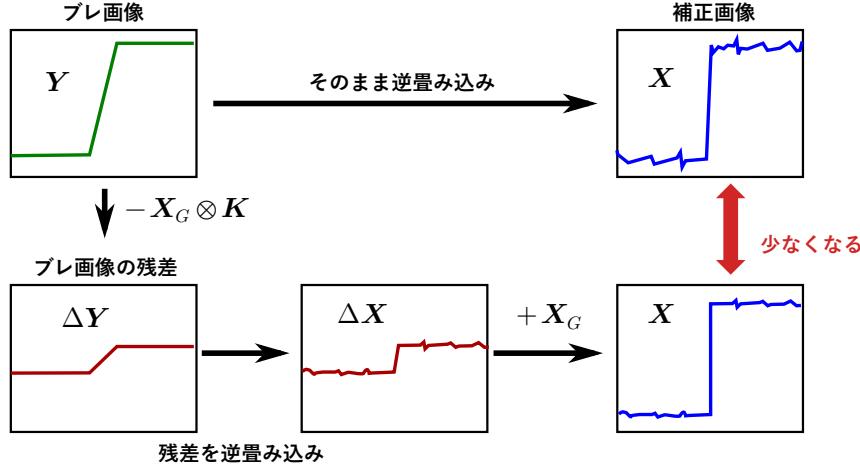


図 4.4 差分逆畳み込みによるリングングの抑制

式 (4.9) の関係を $\Delta \mathbf{X}$ に関する最小二乗問題とすることで、以下の形式で逆畳み込み問題を実現できる。

$$\Delta \mathbf{X} = \underset{\Delta \mathbf{X}}{\operatorname{argmin}} \| \Delta \mathbf{Y} - \Delta \mathbf{X} \otimes \mathbf{K} \|_2^2 \quad (4.10)$$

微分しやすいように $\Delta \mathbf{Y}$, $\Delta \mathbf{X}$ をベクトル表現の $\Delta \mathbf{y}$, $\Delta \mathbf{x}$, \mathbf{K} を畳み込み行列の \mathbf{C} に置き換えた以下の表現に書き換える。

$$\Delta \mathbf{x} = \underset{\Delta \mathbf{x}}{\operatorname{argmin}} \| \Delta \mathbf{y} - \mathbf{C} \Delta \mathbf{x} \|_2^2 \quad (4.11)$$

また、逆畳み込み中のリングング発生を抑制するために、 $\Delta \mathbf{x}$ に対して縦横の微分を行う行列 \mathbf{D} を用いて $\Delta \mathbf{x}$ の Total Variation $\| \mathbf{D} \Delta \mathbf{x} \|_2$ を制約として付加する。

$$\Delta \mathbf{x} = \underset{\Delta \mathbf{x}}{\operatorname{argmin}} \| \Delta \mathbf{y} - \mathbf{C} \Delta \mathbf{x} \|_2^2 + \alpha \| \mathbf{D} \Delta \mathbf{x} \|_2 \quad (4.12)$$

4.4.2 周波数軸における共役勾配法を用いた解法

提案手法は差分逆畳み込みの表現をフーリエ変換の線形性・直交性を利用して周波数軸上に移行する。提案法の逆畳み込みアルゴリズムは式 (4.12) の $\Delta \mathbf{x}$, $\Delta \mathbf{y}$, \mathbf{C} , \mathbf{D} を周波数変換した $\mathcal{F}(\Delta \mathbf{X})$, $\mathcal{F}(\Delta \mathbf{Y})$, $\mathcal{F}(\mathbf{K})$, $\mathcal{F}(\mathbf{D})$ に置き換えて以下の形式にする。

$$\mathcal{F}(\Delta \mathbf{X}) = \underset{\mathcal{F}(\Delta \mathbf{X})}{\operatorname{argmin}} \| \mathcal{F}(\Delta \mathbf{Y}) - \mathcal{F}(\mathbf{K}) \circ \mathcal{F}(\Delta \mathbf{X}) \|_2^2 + \alpha \| \mathcal{F}(\mathbf{D}) \circ \mathcal{F}(\Delta \mathbf{X}) \|_2 \quad (4.13)$$

フーリエ変換の直交性から (4.12) と (4.13) の二乗誤差項、制約項は共に比例するため、この 2 つは同じ逆畳み込み問題を実現することができる。 $\mathcal{F}(\Delta \mathbf{X})$, $\mathcal{F}(\Delta \mathbf{Y})$ の要素を

列ベクトルに配置した $\mathbf{f}_x, \mathbf{f}_y, \mathcal{F}(\mathbf{K}), \mathcal{F}(\mathbf{D})$ の要素を対角行列に配置した行列 $\mathbf{F}_k, \mathbf{F}_d$ を用いて式 (4.13) を以下の様に書き換える.

$$\mathbf{f}_x = \underset{\mathbf{f}_x}{\operatorname{argmin}} \| \mathbf{f}_y - \mathbf{F}_k \mathbf{f}_x \|_2^2 + \alpha \| \mathbf{F}_d \mathbf{f}_x \|_2 \quad (4.14)$$

これを \mathbf{f}_x で微分することで以下の式を得る.

$$\begin{aligned} & -2 \mathbf{F}_k^H (\mathbf{f}_y - \mathbf{F}_k \mathbf{f}_x) + \frac{\alpha}{\| \mathbf{F}_d \mathbf{f}_x \|_2} \mathbf{F}_d^H \mathbf{F}_d \mathbf{f}_x = \mathbf{0} \\ \therefore \mathbf{f}_x &= \left(\mathbf{F}_k^H \mathbf{F}_k + \frac{\alpha}{\| \mathbf{F}_d \mathbf{f}_x \|_2} \mathbf{F}_d^H \mathbf{F}_d \right)^{-1} \mathbf{F}_k^H \mathbf{f}_y \end{aligned} \quad (4.15)$$

式 (4.15) の右辺に \mathbf{f}_x が残っているため、この式で解の \mathbf{f}_x を得ることはできない。しかし提案手法は逆畳み込みアルゴリズムを繰り返し行うため、右辺の \mathbf{f}_x に前回の逆畳み込みの出力 \mathbf{f}'_x を用いることで、解を得ることができる。この方式に関しては後の 4.5 節で述べる。

$$\mathbf{f}_x = \left(\mathbf{F}_k^H \mathbf{F}_k + \frac{\alpha}{\| \mathbf{F}_d \mathbf{f}'_x \|_2} \mathbf{F}_d^H \mathbf{F}_d \right)^{-1} \mathbf{F}_k^H \mathbf{f}_y \quad (4.16)$$

式 (4.16) は共役勾配法を用いて Algorithm 4 の様に解くことができる。

Algorithm 4 提案手法の逆畳み込みアルゴリズム

1. Set \mathbf{f}_y to the vector of $\mathcal{F}(\Delta\mathbf{Y}) = \mathcal{F}(\mathbf{Y} - \mathbf{X}_G \otimes \mathbf{K})$.
 2. **Initialize** $\mathbf{f}_x = \mathbf{f}_y$
 3. $\mathbf{r} = \mathbf{F}_k^H \mathbf{f}_y - \left(\mathbf{F}_k^H \mathbf{F}_k + \frac{\alpha}{\|\mathbf{F}_d \mathbf{f}'_x\|_2} \mathbf{F}_d^H \mathbf{F}_d \right) \mathbf{f}_x$
 4. $\mathbf{p} = \mathbf{r}$
 5. $e_0 = \mathbf{r}^H \mathbf{r}$
 6. **for** $i = 1$ **to** converge e_i **do**
 7. $\mathbf{z} = \left(\mathbf{F}_k^H \mathbf{F}_k + \frac{\alpha}{\|\mathbf{F}_d \mathbf{f}'_x\|_2} \mathbf{F}_d^H \mathbf{F}_d \right) \mathbf{p}$
 8. $x_{step} = e_{i-1} / (\mathbf{p}^H \mathbf{z})$
 9. $\mathbf{f}_x = \mathbf{f}_x - x_{step} \mathbf{p}$
 10. $\mathbf{r} = \mathbf{r} - x_{step} \mathbf{z}$
 11. $e_i = \mathbf{r}^H \mathbf{r}$
 12. $\mathbf{p} = (e_i / e_{i-1}) \mathbf{p}$
 13. **end for**
 14. Set $\Delta\mathbf{X}$ to the \mathcal{F}^{-1} of \mathbf{f}_x 's matrix form
 15. **return** $\mathbf{X} = \mathbf{X}_G + \Delta\mathbf{X}$
-

4.4.3 冗長な DFT 計算の削減による高速化

前節で 2 乗誤差項と制約項を周波数軸に移行した理由は計算時間を削減するためである。図 4.5 に、Algorithm 4 の共役勾配法を空間軸の 2 乗誤差 $\|\mathbf{y} - \mathbf{Cx}\|_2^2$ に対して解いた場合と、周波数軸の 2 乗誤差 $\|\mathbf{f}_y - \mathbf{F}_k \mathbf{f}_x\|_2^2$ に対して解いた場合の比較を示した。

左の空間軸処理は、画像 \mathbf{x} を求める問題、右の周波数軸処理は画像の周波数特性 \mathbf{f}_x を解く問題となる。 $\mathbf{p}, \mathbf{r}, \mathbf{z}$ が空間軸処理では空間軸の値、周波数軸処理では周波数軸の値になる。また、周波数軸処理では $\mathbf{p}, \mathbf{r}, \mathbf{z}$ が複素数の値になるため、転置 T が複素共役転置 H に置き換わる。

右の提案手法は、 \mathbf{p}, \mathbf{z} を周波数軸の値としたことによって、7 行目 の $\mathbf{F}_k^H \mathbf{F}_k + \frac{\alpha}{\|\mathbf{F}_d \mathbf{f}'_x\|_2} \mathbf{F}_d^H \mathbf{F}_d$ の計算の際にフーリエ変換および逆フーリエ変換が不要になり、この分の計算時間を削減できる。8 行目から 12 行目の値が複素数となるため、計算量が 2 倍となるが、7 行目の計算時間削減と比較して影響は少ない。

空間軸の2乗誤差

$$\min_x \|y - Cx\|_2^2 + \alpha \|Dx\|_2$$

周波数軸の2乗誤差

$$\min_{f_x} \|f_y - F_k f_x\|_2^2 + \alpha \rho(F_d f_x)$$

Line 7 周波数軸に変換

$$p \xrightarrow{\text{DFT}} \mathcal{F}(p)$$

畳み込み・相關計算

$$\mathcal{F}(p) - F_k^H F_k + \frac{\alpha}{\|F_d f_x\|_2} F_d^H F_d \xrightarrow{\quad} \mathcal{F}(z)$$

空間軸に逆変換

$$\mathcal{F}(z) \xrightarrow{\text{IDFT}} z$$

Line 8 空間軸

$$x_{\text{step}} = e_{i-1} / (\mathbf{p}^T \mathbf{z})$$

Line 9 空間軸

$$x = x - x_{\text{step}} \mathbf{p}$$

空間軸の解を得る

Line 10 空間軸

$$\mathbf{r} = \mathbf{r} - x_{\text{step}} \mathbf{z}$$

Line 11 空間軸

$$e_i = \mathbf{r}^T \mathbf{r}$$

Line 12 空間軸

$$\mathbf{p} = (e_i / e_{i-1}) \mathbf{p}$$

↓
繰り返し

Line 7 周波数軸上で計算できる

$$p \xrightarrow{- F_k^H F_k + \frac{\alpha}{\|F_d f_x\|_2} F_d^H F_d} z$$

DFT 1回
IDFT 1回 削減

Line 8 周波数軸

$$x_{\text{step}} = e_{i-1} / (\mathbf{p}^H \mathbf{z})$$

Line 9 周波数軸

$$f_x = f_x - x_{\text{step}} \mathbf{p}$$

周波数軸の解を得る

Line 10 周波数軸

$$\mathbf{r} = \mathbf{r} - x_{\text{step}} \mathbf{z}$$

Line 11 周波数軸

$$e_i = \mathbf{r}^H \mathbf{r}$$

Line 12 周波数軸

$$\mathbf{p} = (e_i / e_{i-1}) \mathbf{p}$$

↓
繰り返し

図 4.5 空間軸処理と周波数軸処理のコスト比較

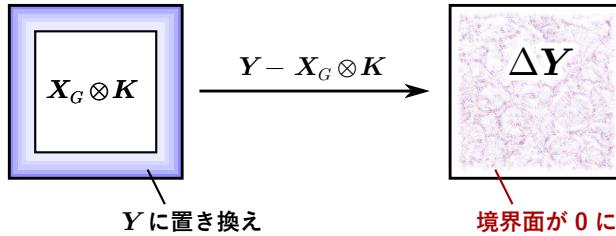


図 4.6 $\Delta\mathbf{Y}$ 画像端の不連続性を解消する方法

4.4.4 ガイド画像の生成

差分逆畳み込みの従来法では、ガイド画像 \mathbf{X}_G に短時間露光画像等の別画像を用いる必要があったが、提案法ではぶれ画像自身や、前回の逆畳み込み結果を利用して、ガイド画像がなくても逆畳み込みできる方式を考案する。提案手法は、ガイド画像がある場合とない場合で処理を分ける。

ガイド画像はある程度エッジ成分を持っているほうが画像のエッジ部分からのリングングの発生を抑制できると考えられる。一方で、ガイド画像がリングング成分を持っていると出力画像が $\Delta\mathbf{X} + \mathbf{X}_G$ であるため、 \mathbf{X}_G にリングングの影響があると最終出力画像にもリングングが残ってしまう。

そのため、前回の逆畳み込み出力をガイド画像として用いる場合、逆畳み込み出力からリングングを取り除き、かつ、エッジ成分を強調した画像をガイド画像として用いる。このために、提案手法では、入力ガイド画像に対して Guided フィルタでノイズやリングングの影響を取り除く。その後、ショックフィルタをかけて画像の主となるエッジ成分を復元させる。これによって得た \mathbf{X}_G を提案手法の $\Delta\mathbf{Y}, \Delta\mathbf{X}$ 算出に用いる。

ガイド画像がない場合入力ぶれ画像からガイド画像を動的に生成する。入力ぶれ画像にはモーションブレの影響がある一方で、リングング成分はないものと考えられる。そのため、入力ぶれ画像に対して、カーネル \mathbf{K} を上下左右反転させたミラーカーネルをかけてモーションぶれの影響を軽減させる。その後ショックフィルタをかけて画像の主となるエッジ成分を復元させたものをガイド画像 \mathbf{X}_G とする。

ぶれ画像 \mathbf{Y} とガイド画像の差分 $\mathbf{Y} - \Delta\mathbf{X}_G \otimes \mathbf{K}$ を算出する際には、図 4.6 の様に $\Delta\mathbf{X}_G \otimes \mathbf{K}$ の外側部分を \mathbf{Y} に置換する。これによって $\Delta\mathbf{Y}$ の外側部分が 0 となり、境界面での不連続性が解消されるため、画像端からのリングングの発生を抑えることができる。

4.5 提案する Blind Deconvolution のフレームワーク

4.3 節のカーネル推定および 4.4 節の逆畳み込みアルゴリズムを利用して提案法の Blind Deconvolution のフレームワーク図 4.7 の様に構築する。

カーネル推定の際に、カラー画像はグレースケールに変換して推定を行う。1 回目のカーネル推定はカーネルの発散を防ぐために式 (4.2) の空間軸におけるカーネル推定問題を使用する。また、前回の逆畳み込みの結果が存在しないため、式 (4.15) の逆畳み込みは $\alpha = 0$ すなわち TV 制約なしで解く。2 回目以降のカーネル推定では、周波数軸で二乗誤差を取った式 (4.3) を用いてカーネル推定を行い、前回の逆畳み込み結果を TV 制約に用いた式 (4.16) を用いて逆畳み込みを行う。逆畳みこみを行って得た補正画像を次のカーネル推定に用いて精度の向上をはかる。この繰り返しは初回のカーネル推定を含めて 5 回ほど行う。

最終的に得られたカーネル推定結果を元に元のカラー画像の逆畳み込みを行う。はじめに TV 制約・ガイド画像算出のために、式 (4.15) の逆畳み込みを $\alpha = 0$ 、すなわち TV 制約なしで解く。次に、これによって得た画像をガイド画像 \mathbf{X}_G と、TV 制約算出の \mathbf{f}'_x に用いて式 (4.16) で逆畳み込みを行う。これによって得た画像を最終出力画像とする。

4.6 実験と評価

4.6.1 実験の設定

提案手法の性能を評価するために、Tecnick TESTIMAGES [104] で提供されているサイズ 600×600 の画像 40 種類（図 4.8）に逆畳み込みアルゴリズムを適用し、比較を行った。

実験は、カーネルを既知として逆畳み込みアルゴリズムのみを行った Non-blind deconvolution の精度比較と、カーネルを未知としてカーネル推定・逆畳み込みの両方を行う Blind deconvolution の精度比較を行った。

Non-blind deconvolution の評価では、原画像にカーネルを畳み込みし、逆畳み込みアルゴリズムを適用した後、原画像との PSNR を算出し、比較を行う。カーネルには図 4.9 に示すサイズ 11×11 のカーネル 8 種および図 4.10 の 35×35 カーネル 5 種に関して評価を行った。結果の PSNR はカーネル 8 種を補正した際の結果を平均したもの

のを用いる。

比較対象の手法として、ウィナーフィルタ, Richardson Lucy 法 [49], Levin の手法の逆畳み込みプロセス [50], Levin の手法と同様にスペース制約を用いる Krishnan の手法 [105] を用いた。

Blind deconvolution の評価では Non-blind と同様に図 4.9 のサイズ 11×11 のカーネルを画像に適用し, カーネル推定と逆畳み込みを行う。Blind deconvolution アルゴリズムの比較対象手法には Fergus の手法 [47], Cho and Lee の手法 [53], Levin の手法 [50], Xu および Jia の手法 [51, 52] を用いた。

図 4.9 の比較に用いるサイズ 11×11 のカーネルは主にモーションぶれのみからなるカーネル (**1** から **4**) とモーションぶれと焦点ぼけの両方を含むカーネル (**5** から **8**) を含む。またそれぞれのカーネルにおいて, 直線ぶれからなるもの (**1, 2, 5**) と, より複雑なぶれ軌跡を持つものを含んでおり, それぞれのカーネルの補正結果を平均することで, カーネル推定手法および逆畳込み手法とともに, 多くの種類のカーネルに対応できる手法同士の比較が実現できると考えられる。Non-blind deconvolution 比較に用いた図 4.10 のサイズ 35×35 のカーネルは, より複雑なぶれ軌跡の補正を評価できる様, あらゆる方向成分を含むカーネルを用いた。

従来手法である Cho の手法において, 式 (2.69) カーネル推定問題の制約は $\beta = 5.0$ とした。提案手法においては式 (4.4) のカーネル推定問題の制約は $\beta = 5.0$, 式 (4.16) 逆畳み込み問題の制約は $\alpha = 0.001$ とした。

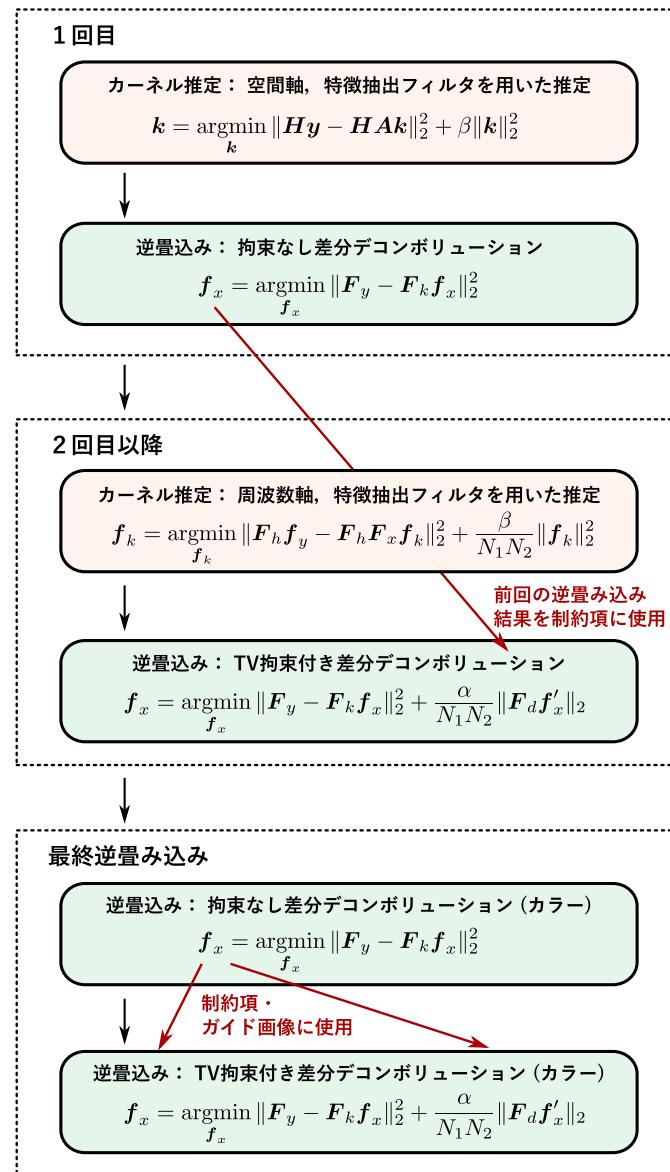


図 4.7 提案手法のカーネル推定・逆畳込みのフレームワーク



図 4.8 実験で用いる 40 画像

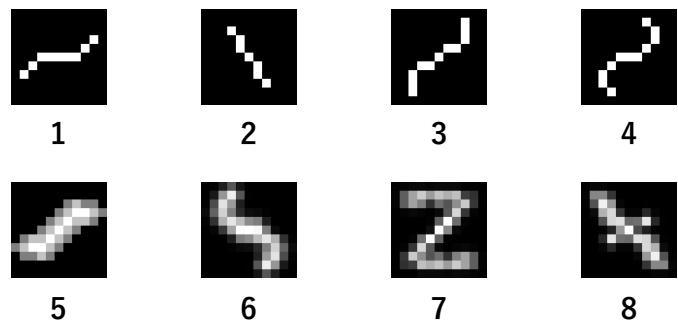


図 4.9 実験で用いるサイズ 11×11 のカーネル 8 種

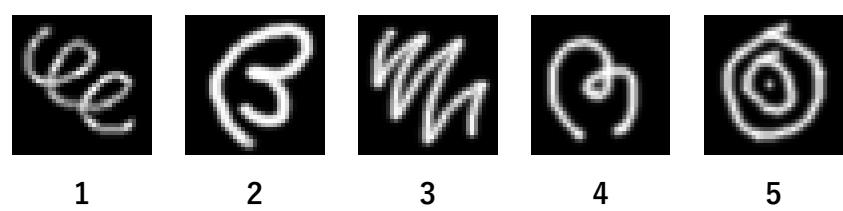


図 4.10 実験で用いるサイズ 35×35 のカーネル 5 種

4.6.2 Non-blind deconvolution 手法との比較

表 4.1 に 11×11 のカーネルを逆畳み込みした画像の PSNR 比較を示す。また表 4.2 に 35×35 のカーネルを逆畳み込みした画像の PSNR 比較を示す。サイズ 11×11 のカーネル補正結果では Levin の手法や Krishnan の手法と比べて PSNR で 3 dB 前後向上しており、差分逆畳み込み導入による画像補正精度の改善ができていることがわかる。サイズ 35×35 のカーネル補正結果では Krishnan の手法と比べて PSNR で 3 dB ほど改善しており本手法の逆畳み込みがより複雑なカーネルの補正にも対応できることが分かる。

図 4.11 に画像 Bananas にサイズ 11×11 のカーネルをかけて逆畳み込みした結果画像を示す。図の左上に原画像全体と用いたカーネル、画像の拡大位置を示し、(a) に原画像の拡大画像、(b) にぶれ画像の拡大画像、(c) から (g) に各手法の拡大画像を示している。各手法の補正精度は (a) の原画像と比較することで評価することができる。提案法は Levin や Krishnan の手法と比較して、ヘタの部分や背景のテクスチャ再現性に優れており、これが PSNR 向上に寄与したものと考えられる。Wiener フィルタは提案法同様とテクスチャ部の再現ができているが画像端からのリングングが復元精度を大きく悪化させている。

図 4.13 の画像 Snail では、目の部分のテクスチャを再現しつつ、Wiener フィルタや Richardson Lucy 法に比べてリングングの影響を低減させている。しかし、背景の平坦な部分では、スペース制約を用いる Levin, Krishnan 2 手法と比べてリングングが残っている問題がある。図 4.13 の画像 Cards も同様に、カードの絵柄のテクスチャ再現性を再現しつつも、平坦な部分にリングングを残す結果となっている。

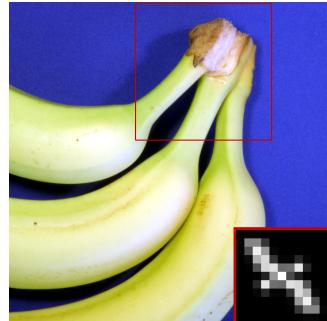
図 4.14 に画像 Multimeter に 35×35 のカーネルをかけて逆畳み込みした結果画像を示す。 35×35 の大きなカーネルではスペース制約を用いる Levin の手法と Krishnan の手法では Krishnan の手法のほうが復元性能が高い。提案法は文字の部分の復元性能を確保しつつ、Wiener フィルタで発生している激しいリングングの影響を低減させた復元ができている。

表 4.1 サイズ 11×11 のカーネルの逆畳み込み精度比較 (単位 : dB)

画像	Wiener	Lucy	Levin	Krishnan	提案法
almonds	29.576	31.687	30.815	30.889	33.737
apples	36.464	39.315	38.618	38.745	43.300
balloons	29.265	33.030	36.441	36.270	38.695
bananas	28.883	34.231	36.808	37.269	41.758
billiard_balls_a	29.322	34.699	38.938	38.710	40.443
billiard_balls_b	32.904	35.852	35.715	35.845	37.703
building	31.743	33.502	31.425	31.267	34.795
cards_a	25.749	27.162	26.718	27.135	29.392
cards_b	27.801	28.942	29.902	30.497	32.024
carrots	29.705	32.089	32.037	32.225	35.806
chairs	28.892	34.948	40.904	40.294	42.828
clips	25.059	27.801	26.658	26.796	28.301
coins	31.848	32.352	29.538	29.530	33.907
cushions	30.279	34.958	38.744	38.786	42.395
ducks	28.793	34.194	39.671	39.377	42.419
fence	27.109	30.644	32.700	33.138	34.269
flowers	27.672	30.359	29.793	29.792	32.513
garden_table	29.151	31.527	28.858	28.928	33.274
guitar_bridge	33.760	35.633	32.507	32.991	35.994
guitar_fret	28.538	31.996	33.231	32.957	35.625
guitar_head	31.399	32.450	29.098	29.036	32.835
keyboard_a	32.195	33.259	28.322	28.493	31.905
keyboard_b	32.526	30.058	31.229	31.507	33.465
lion	27.658	30.896	30.978	31.096	34.100
multimeter	33.965	34.440	32.431	32.595	35.924
pencils_a	31.227	32.884	33.215	32.778	36.168
pencils_b	27.191	31.423	33.585	33.511	36.729
pillar	31.339	33.252	31.471	31.245	35.221
plastic	33.859	36.850	35.128	35.506	40.905
roof	26.237	28.582	29.045	29.223	32.658
scarf	26.621	29.354	27.071	26.979	29.758
screws	28.510	27.411	24.589	25.095	27.350
snails	28.036	33.247	35.830	35.506	39.827
socks	29.505	31.948	28.867	29.007	33.693
sweets	28.949	32.716	34.438	34.340	36.987
tomatoes_a	29.699	34.374	38.844	39.006	41.433
tomatoes_b	32.124	37.103	39.067	38.980	40.780
tools_a	30.080	31.107	30.606	30.472	33.576
tools_b	29.836	34.346	35.668	35.781	39.132
wood_game	44.762	39.466	38.320	38.066	41.549
平均	30.206	32.752	32.946	32.992	36.079

表 4.2 サイズ 35×35 のカーネルの逆畳み込み精度比較 (単位 : dB)

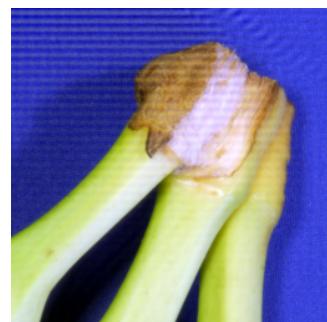
画像	Wiener	Lucy	Levin	Krishnan	提案法
almonds	24.784	24.876	18.361	24.002	26.594
apples	33.400	32.130	26.799	33.080	36.571
balloons	25.362	25.728	21.280	27.505	30.188
bananas	25.895	26.925	25.032	32.145	33.973
billiard_balls_a	25.819	27.180	21.865	29.268	31.202
billiard_balls_b	29.228	28.510	20.717	28.334	30.609
building	27.175	26.280	21.248	24.501	27.512
cards_a	22.137	20.743	15.579	20.898	23.089
cards_b	23.528	22.348	17.316	23.639	25.351
carrots	25.433	24.533	21.531	26.388	29.099
chairs	25.627	28.065	25.880	31.229	34.223
clips	20.339	19.467	11.866	19.072	21.049
coins	27.616	26.057	20.643	24.591	27.947
cushions	27.062	28.139	26.381	31.712	35.603
ducks	25.723	26.964	24.854	31.637	34.400
fence	23.304	23.513	16.592	24.935	27.215
flowers	23.427	23.243	17.254	22.985	25.777
garden_table	25.785	25.224	20.767	24.184	26.984
guitar_bridge	30.123	29.319	21.442	28.265	30.357
guitar_fret	25.024	25.039	20.326	26.385	30.383
guitar_head	29.256	27.286	19.726	24.351	28.164
keyboard_a	27.285	26.208	19.059	23.007	26.120
keyboard_b	28.051	25.439	18.183	25.195	28.254
lion	24.798	24.605	18.496	25.257	28.539
multimeter	29.921	28.198	20.969	27.010	30.248
pencils_a	24.907	25.165	20.193	25.021	30.367
pencils_b	22.996	23.553	19.405	25.043	29.381
pillar	27.771	26.973	21.173	25.557	29.707
plastic	30.629	30.184	27.282	31.903	34.834
roof	23.944	22.924	18.282	21.684	26.897
scarf	23.253	22.632	16.111	20.247	24.061
screws	24.763	21.933	14.856	18.965	22.889
snails	24.774	26.625	23.235	28.612	32.411
socks	25.046	25.041	18.792	24.701	27.627
sweets	25.017	25.331	20.416	26.339	29.101
tomatoes_a	26.041	26.290	22.745	29.948	31.721
tomatoes_b	26.790	28.757	22.092	28.882	31.482
tools_a	24.835	24.079	17.816	23.779	27.129
tools_b	25.717	26.963	21.765	28.315	30.612
wood_game	39.545	32.345	24.471	30.823	34.335
平均	26.303	25.870	20.520	26.235	29.300



(a) 原画像



(b) ぶれ画像



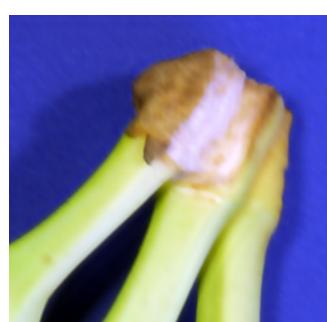
(c) Wiener



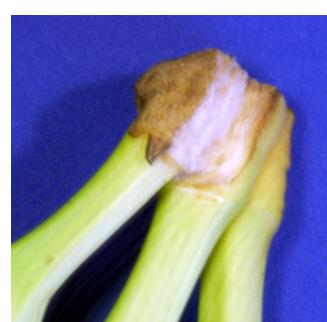
(d) Richardson Lucy



(e) Levin



(f) Krishnan

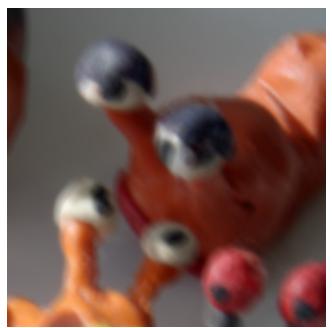


(g) 提案法

図 4.11 画像 Bananas, サイズ 11 のカーネルの Non-blind 補正結果



(a) 原画像



(b) ぶれ画像



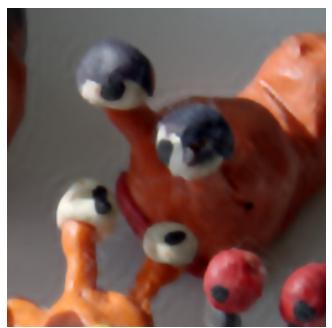
(c) Wiener



(d) Richardson Lucy



(e) Levin



(f) Krishnan

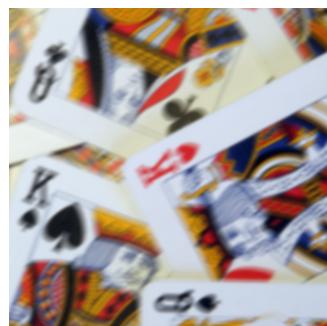


(g) 提案法

図 4.12 画像 Snails, サイズ 11 のカーネルの Non-blind 補正結果



(a) 原画像



(b) ぶれ画像



(c) Wiener



(d) Richardson Lucy



(e) Levin



(f) Krishnan

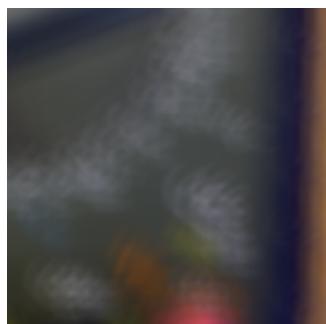


(g) 提案法

図 4.13 画像 Cards, サイズ 11 のカーネルの Non-blind 補正結果



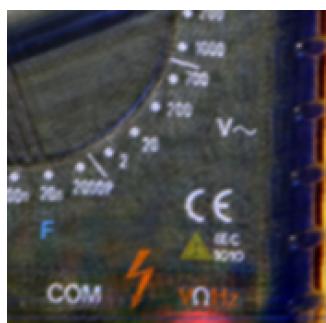
(a) 原画像



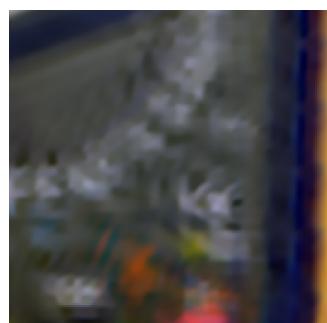
(b) ぶれ画像



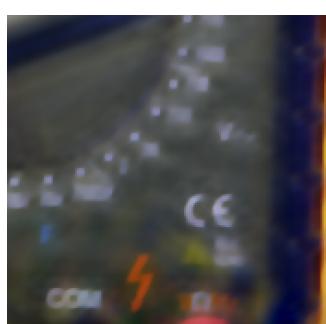
(c) Wiener



(d) Richardson Lucy



(e) Levin



(f) Krishnan



(g) 提案法

図 4.14 画像 Multimeter, サイズ 35 のカーネルの Non-blind 補正結果

4.6.3 Blind deconvolution 手法との比較

表 4.3 にサイズ 11×11 のカーネルをかけ、カーネル推定・逆畳込みを行った際の復元 PSNR の比較結果を示す。提案手法はベースとした Cho の手法と比べて PSNR を 2.1 dB 前後改善させている。またいくつかの画像では Levin の手法、Xu および Jia の手法より高い PSNR を示した画像がある。平均 PSNR では Xu および Jia の手法より 0.3 dB 劣っているが、この手法に対しては後述の実行時間の点で優位性がある。

表 4.4 にカーネルごとに補正した PSNR を算出した値を全画像で平均した値を示す。提案手法はベースとした Cho の手法と比べて、カーネルによらず補正 PSNR を改善できていることが分かる。また、提案手法は Levin の手法と比べて直線的なカーネルの補正に優れ、Jia の手法と比べてぼけを含むカーネルの補正に優れていることが分かる。Levin の手法や Jia の手法とは一部補正 PSNR が劣っている部分があるが、これらの手法に対しては実行時間の点で優位性がある。

Blind deconvolution アルゴリズムの精度は、カーネル推定プロセスにおいてどれくらい正確なカーネルが推定できたか、逆畳み込みプロセスにおいてどれほどリングングやアーティファクトを発生させずに画像補正を行うかに依存する。提案法のカーネル推定は周波数軸上の二乗誤差最小化によって推定を行い、カーネルの周波数特性を解として得るが、この方式でも問題なくカーネル推定を実現できている事がわかる。また、提案法の逆畳み込みアルゴリズムは Non-blind での検証の際にも述べたとおり、テクスチャ部の復元精度に優れる一方、平坦部にリングングが残りやすいアルゴリズムであるが、画像の二乗誤差基準で見た場合に全体的に復元誤差を改善できているものと考えられる。

表 4.3 サイズ 11×11 カーネルの Blind Deconvolution 精度比較 (単位 : dB)

画像	Cho	Fergus	Levin	XuJia	Proposed
almonds	23.594	21.871	28.984	27.382	27.865
apples	36.013	30.401	33.270	36.003	38.885
balloons	28.034	26.203	31.253	31.649	31.536
bananas	31.133	28.575	32.516	33.148	34.892
billiard_balls_a	31.293	27.105	31.771	32.331	34.812
billiard_balls_b	26.435	25.331	30.803	31.274	29.084
building	24.220	23.481	24.513	26.495	24.950
cards_a	19.510	17.753	23.992	23.420	22.338
cards_b	23.736	20.601	26.068	26.612	27.000
carrots	23.677	24.324	26.001	26.176	25.260
chairs	30.359	30.111	32.683	31.181	32.241
clips	16.107	14.066	18.174	21.423	21.993
coins	22.570	23.152	27.641	26.928	23.691
cushions	33.197	30.409	33.664	34.915	36.739
ducks	32.387	28.916	33.067	34.075	38.499
fence	23.473	21.967	25.279	27.542	28.145
flowers	21.577	20.969	25.840	26.077	26.191
garden_table	24.384	22.737	26.349	26.302	24.877
guitar_bridge	25.652	24.874	26.806	27.453	25.168
guitar_fret	23.940	24.768	26.064	25.120	23.047
guitar_head	22.707	22.458	24.262	25.311	22.944
keyboard_a	20.601	20.912	23.644	24.179	20.071
keyboard_b	22.372	22.108	22.739	25.654	22.744
lion	24.733	22.959	26.178	28.128	26.808
multimeter	27.009	24.706	28.115	29.084	27.022
pencils_a	23.211	25.901	25.148	24.147	23.163
pencils_b	25.343	22.715	26.968	27.716	25.013
pillar	23.647	24.028	27.918	27.756	26.163
plastic	27.789	29.380	31.431	27.867	27.397
roof	19.998	20.073	20.040	20.785	20.598
scarf	17.870	13.616	22.232	20.027	18.512
screws	15.359	14.258	18.423	17.671	15.261
snails	31.421	27.422	31.780	32.588	33.377
socks	22.357	20.314	24.302	23.793	24.876
sweets	28.767	24.781	30.416	30.695	32.421
tomatoes_a	30.328	27.499	30.861	32.682	34.444
tomatoes_b	31.366	27.199	31.153	32.993	35.142
tools_a	21.211	22.408	25.985	25.684	22.937
tools_b	29.661	26.147	30.101	30.252	32.016
wood_game	33.567	28.290	33.913	36.000	36.534
Average	25.515	23.870	27.509	27.963	27.616

表 4.4 カーネルごとの Blind Deconvolution 精度比較 (単位 : dB)

カーネル	Cho	Fergus	Levin	XuJia	Proposed
1	26.249	24.764	25.207	28.285	27.048
2	26.136	24.405	24.734	28.314	27.639
3	23.189	22.411	27.810	27.626	26.150
4	23.863	23.098	28.268	27.927	26.938
5	27.734	24.735	27.074	28.462	30.730
6	26.082	24.281	28.787	27.998	28.482
7	23.450	22.524	29.892	27.124	25.345
8	27.420	24.739	28.298	27.965	28.599
Average	25.515	23.870	27.509	27.963	27.616

4.6.4 画像の視覚評価

図 4.15 に画像 Bananas をカーネル **1** で畳み込みした画像の, 図 4.16 に画像 Garden Table をカーネル **3** で畳み込みした画像の, カーネル推定・画像補正結果を示す. 図の (a) は原画像, (b) には入力ぶれ画像を表しており, (b) の赤枠内には, 入力ぶれ画像を作成するのに用いたカーネルを示した. 図の (c) から (g) に, 従来手法 4 種の補正画像および推定カーネルを示した. それぞれの手法のカーネル推定結果は (b) の赤枠のカーネルと比較することで, 補正画像は (a) の原画像と比較することで評価することができる. カーネル **1, 3** はともにモーションのみからなり焦点ボケを含まないカーネルである. これらのカーネルに対しては, Cho の手法と提案手法がカーネルを比較的忠実に推定できていることが分かる. Cho の手法は空間軸上の二乗誤差, L_2 ノルム制約でカーネルを解く手法, 提案法は周波数軸上の二乗誤差, L_2 ノルム制約でカーネルを解く手法であり, いずれを用いてもカーネル推定問題を実現できることが分かる. 補正画像の比較では, 提案法は Bananas のヘタ部分や Garden Table の天板部分等, 主にテクスチャ部の復元に優れている. Cho の手法でも同等の傾向が見られるが, 提案手法ではよりリングングの影響が抑えられている.

図 4.17 に画像 Almonds をカーネル **5** で畳み込みした画像の, 図 4.18 に画像 Tools をカーネル **6** で畳み込みした画像のカーネル推定・画像補正結果を示す. 画像 Almonds では Cho, Levin, Xu, 提案法が, 画像 Tools では提案法が比較的高精度にカーネル推定を実現している. 補正画像では, Levin や Xu の手法ではテクスチャ部を平滑化させている一方で, 提案法は Almonds の表面や Tools のネジの部分などの復元性能が高い一方で, 背景部分などの平坦な部分にリングングの影響が出た出力となっている.

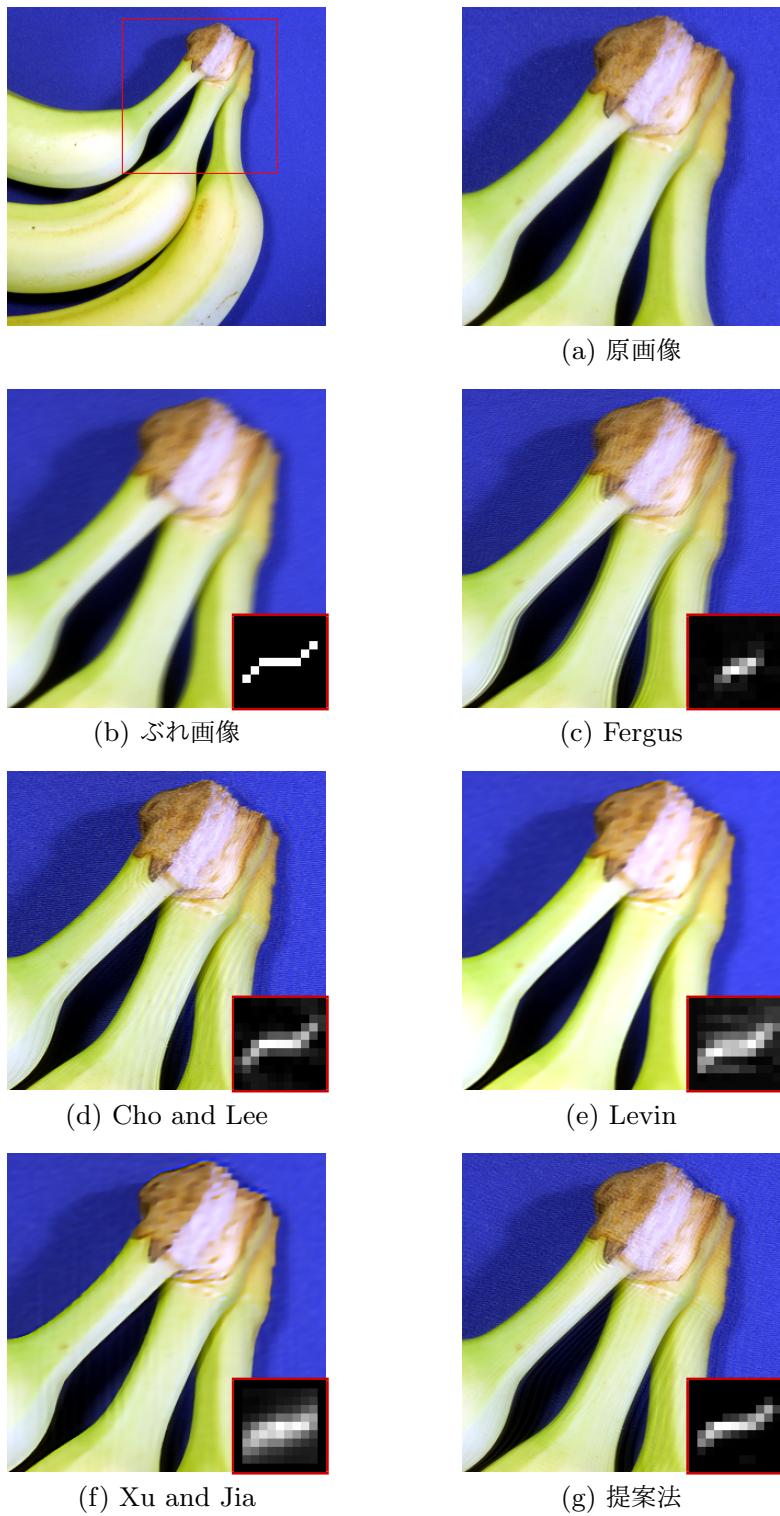


図 4.15 画像 Bananas の推定カーネルと補正結果



(a) 原画像



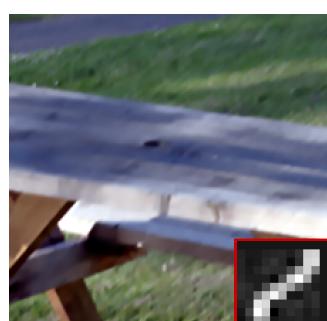
(b) ぶれ画像



(c) Fergus



(d) Cho and Lee



(e) Levin



(f) Xu and Jia



(g) 提案法

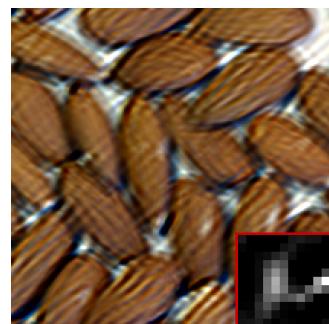
図 4.16 画像 Garden table の推定カーネルと補正結果



(a) 原画像



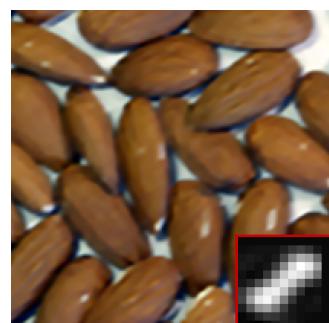
(b) ぶれ画像



(c) Fergus



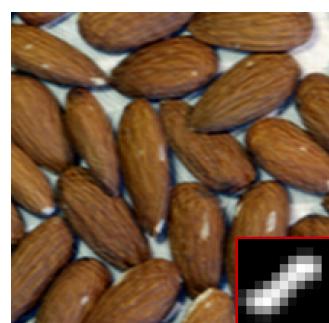
(d) Cho and Lee



(e) Levin

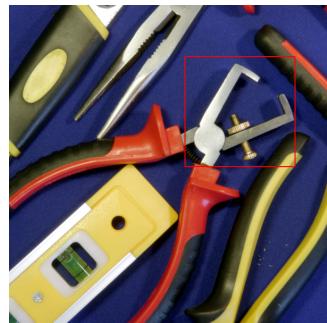


(f) Xu and Jia

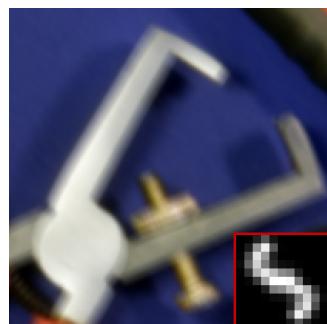


(g) 提案法

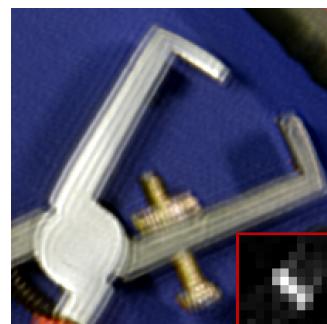
図 4.17 画像 Almonds の推定カーネルと補正結果



(a) 原画像



(b) ぶれ画像



(c) Fergus



(d) Cho and Lee



(e) Levin



(f) Xu and Jia



(g) 提案法

図 4.18 画像 Tools の推定カーネルと補正結果

表 4.5 Blind Deconvolution 実行時間の比較 (単位 : 秒)

画像サイズ	Fergus	Cho	Levin	XuJia	Proposed
300	42.39	3.15	33.32	177.27	2.51
480	96.70	6.86	92.49	351.26	4.60
600	135.56	10.02	96.66	486.76	7.15

4.6.5 実行時間の評価

表 4.5 に Blind deconvolution アルゴリズムの処理時間の比較結果を示す。提案手法は高速な手法である Cho の手法より処理時間を短縮することができており、比較対象の手法の中で最も高速に動作する。これは、4.3.1 節の前処理フィルタの改善による処理時間の短縮と式 (4.13) の周波数ベースの逆畳み込みによる DFT 計算の削減が主な要因である。

4.7 まとめと今後の展望

本章ではノンパラメトリックな Blind deconvolution である Cho の手法をベースにカーネル推定部と逆畳み込み部分の改善を行った。

カーネル推定部では、Guided フィルタによる前処理と周波数軸上で特徴抽出フィルタを用いたカーネル推定問題を提案した。Guided フィルタを用いた前処理では、ぶれ画像や逆畳み込みによって生じるノイズやリングングを除去しつつ、従来法の前処理フィルタよりも高速な処理を実現している。また、特徴抽出フィルタを用いることで、従来法のしきい値処理の計算コストを省きつつ、しきい値処理では取れない細かな画像のエッジ・テクスチャ部を最適化問題に適用することでカーネル推定精度を向上させていている。

逆畳み込みの過程では、周波数軸上でガイド画像を用いない差分逆畳み込みの手法を提案した。この手法は画像自身の代わりに画像とガイド画像の差分を逆畳み込みすることで、リングング成分が増幅されるのを防ぐ。また最適化問題の誤差項・制約項をそれぞれ周波数軸の値とし、共役勾配法を用いて解くことで反復法の各反復における冗長な DFT/IDFT 算出の計算コストを削減した。

提案手法は従来の Blind Deconvolution 手法と比較して高速な処理を実現しつつ、よ

り高いカーネル推定精度とテクスチャ部分の復元性にすぐれる逆畳み込みを実現した。

しかしながら、提案手法の逆畳み込みは、スパース制約を用いない主に平坦な部分にリンクングが発生しやすい問題点があり、Non-blind, Blind 問題ともに出力画像の見た目に悪影響を与えていた問題点が見受けられ、その改善が望まれる。

第 5 章

結 論

本論文では、辞書ベースの超解像および、ぶれ推定・ぶれ除去問題に関して既存手法の問題点を明らかにするとともに、その解決策を提案した。本章ではそれらを統括するとともに、提案法における問題点や将来の展望について述べる。

辞書ベースの超解像手法

本論文の第3章において、一枚の画像に対する辞書ベースの超解像手法を取り扱った。この手法は、入力画像を予め Bicubic 法等で拡大したものを低解像度画像とし、低解像度画像から切り出した低解像度パッチを低解像度辞書と係数ベクトルの線形結合で近似する。従来手法では、 L_1 再構成を用いた係数ベクトル算出の計算コストがかかる点、低解像度・及び高解像度の辞書を予め用意しておく必要がある点に問題であった。提案手法では係数ベクトル算出の L_1 再構成問題を L_2 再構成に変更した超解像手法を提案した。 L_2 再構成を用いた場合、係数 α の導出問題を行列の積のみで計算が可能となるため、最適化問題が必要な L_1 再構成問題と比較して計算コストを削減することができる。また、入力低解像度画像からパッチをサンプルすることによって再構成問題に用いる低解像度辞書・高解像度辞書のペアを動的に生成する手法を提案した。これによって辞書サイズ・パッチサイズ等のパラメータが異なる辞書を予め用意しておく必要がなくなり、柔軟に辞書を構築できる。標準画像を用いた評価の結果、高解像度パッチの再構成において、 L_2 再構成を用いた場合でも L_1 再構成を用いる従来法から PSNR を大きく落とすことなしに画像再構成を可能であることを示した。また、事前辞書を用いた再構成と、自画像からサンプルした辞書を用いた再構成の比較から、自画像からサンプルした辞書を用いても事前辞書と同等の PSNR をもつ画像再構成が可能であることを示した。

提案手法の問題点として、画像中の方向成分の偏りなどによって、辞書中にサンプルされるパッチの種類に偏りが生じ、画像の再構成精度を悪化させる問題が発生した。また、自画像からのパッチのサンプルにかかる時間が画像の画素数に比例して大きくなるため、特に大きなサイズの画像を処理する際の計算コスト改善が望まれる。

画像のぶれ除去手法

本論文の第4章では、モーションぶれと焦点ぼけを含む劣化画像の補正としてぶれ除去問題を取り扱った。ぶれ除去には既知のぶれ軌跡を取り除く Non-blind Deconvolution と、ぶれ軌跡が未知の Blind Deconvolution があるが、本論文では Blind Deconvolution を取り扱った。また、Blind Deconvolution 手法にはぶれ軌跡に直線・折れ線などの想定を加える手法が存在するが、本手法では、これらの想定を行わないため、モーションぶれ・焦点ぼけの両方に対応できる。

従来手法では、前処理フィルタをかけた画像からエッジ検出を行い、これに対してカーネル推定を行う。また、これによって得られたカーネルをもとに画像のぶれ補正を行い、次のカーネル推定に用いる補正画像を得る。

提案手法ではカーネル推定問題の前処理過程を高速化・精度向上させると共に、特徴抽出フィルタを導入することで、従来手法のエッジ検出処理の計算コストを省きつつ、画像のエッジ・テクスチャ部をカーネル推定問題に適用することで、カーネル推定精度を向上させた。逆畳込み問題では、周波数軸上でガイド画像を用いない差分逆畳み込みの手法を提案した。この手法は、画像とガイド画像の差分値を逆畳み込みすることで、リングング成分の増幅を抑えることができる。また、誤差関数・制約項を周波数空間の値とすることで冗長なDFT計算を削減し、処理時間の削減を実現した。テスト画像を用いた性能評価では、提案手法は従来の手法と比較して高速な処理を実現しつつ、より高いカーネル推定精度とテクスチャ部分の復元性に優れる逆畳み込みを実現した。

本論文では、画像の劣化過程のうち、画像のぼけ行列、ダウンサンプル、モーションぶれに着目した。画像のぼけ行列とダウンサンプルを劣化過程に含む画像には超解像問題、ぼけ行列とモーションぶれを含む画像にはぶれ除去問題が適用できる。画像の劣化復元手法の改良とともに、本論文では、アルゴリズムの処理時間短縮を主眼に入れた改善を行った。画像の超解像手法では、ノルム制約の緩和によって閉形式の解を得ることで、画像のぶれ除去手法では冗長な周波数変換計算の削減によって高速化を実現した。近年、撮像機器・表示機器の高解像度化は著しく、これら高解像度の画像を実用的な時

間で処理できるかどうかはアルゴリズムの実用化の観点から重要な課題となっている。本論文のアルゴリズムは画像の超解像・ぶれ除去の手法を実用化に一步近づけるとともに、他の最適化ベースの画像再構成手法を高速化する方法の一つを提案できたものと考えている。

画像に含まれる劣化過程の要因には多岐に渡り、それぞれの劣化に対する復元の過程はそれぞれ異なったものとなる。これらの劣化過程を取り除き、劣化のない画像情報を提供することは、画像信号の利用者によりよい経験をもたらす他、商業的・工業的な利用価値を高めることにもつながる。本研究で得られた知見がそのような画像の劣化画像の復元技術、ひいては画像処理技術のさらなる発展につながることを願い、本論文を結ぶ。

参考文献

- [1] H. Hou and H. Andrews, “Cubic splines for image interpolation and digital filtering,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.26, no.6, pp.508–517, Dec. 1978.
- [2] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, United States Governm. Press Office Los Angeles, CA, 1950.
- [3] C.E. Duchon, “Lanczos filtering in one and two dimensions,” Journal of Applied Meteorology, vol.18, no.8, pp.1016–1022, 1979.
[https://doi.org/10.1175/1520-0450\(1979\)018<1016:LFI0AT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<1016:LFI0AT>2.0.CO;2)
- [4] F. Zhou, W. Yang, and Q. Liao, “Interpolation-based image super-resolution using multisurface fitting,” IEEE Transactions on Image Processing, vol.21, no.7, pp.3312–3318, July 2012.
- [5] R. Timofte, V. De, and L.V. Gool, “Anchored neighborhood regression for fast example-based super-resolution,” 2013 IEEE International Conference on Computer Vision, pp.1920–1927, Dec. 2013.
- [6] B.K. Gunturk and M. Gevrekci, “High-resolution image reconstruction from multiple differently exposed images,” IEEE Signal Processing Letters, vol.13, no.4, pp.197–200, April 2006.
- [7] Y. Tian and K.H. Yap, “Joint image registration and super-resolution from low-resolution images with zooming motion,” IEEE Transactions on Circuits and Systems for Video Technology, vol.23, no.7, pp.1224–1234, July 2013.
- [8] S.J. Kang, D.G. Yoo, S.K. Lee, and Y.H. Kim, “Multiframe-based bilateral motion estimation with emphasis on stationary caption processing for frame rate up-conversion,” IEEE Transactions on Consumer Electronics, vol.54,

- no.4, pp.1830–1838, Nov. 2008.
- [9] W. Bao, X. Zhang, L. Chen, L. Ding, and Z. Gao, “High-order model and dynamic filtering for frame rate up-conversion,” IEEE Transactions on Image Processing, vol.27, no.8, pp.3813–3826, Aug. 2018.
 - [10] S.P. Kim and W.Y. Su, “Recursive high-resolution reconstruction of blurred multiframe images,” IEEE Transactions on Image Processing, vol.2, no.4, pp.534–539, Oct. 1993.
 - [11] A. Marquina and S.J. Osher, “Image super-resolution by tv-regularization and bregman iteration,” Journal of Scientific Computing, vol.37, no.3, pp.367–382, Dec. 2008.
<https://doi.org/10.1007/s10915-008-9214-8>
 - [12] S. Farsiu, M.D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super resolution,” IEEE Transactions on Image Processing, vol.13, no.10, pp.1327–1344, Oct. 2004.
 - [13] Y.R. Li, D.Q. Dai, and L. Shen, “Multiframe super-resolution reconstruction using sparse directional regularization,” IEEE Transactions on Circuits and Systems for Video Technology, vol.20, no.7, pp.945–956, July 2010.
 - [14] Q. Yuan, L. Zhang, and H. Shen, “Multiframe super-resolution employing a spatially weighted total variation model,” IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.3, pp.379–392, March 2012.
 - [15] S. Baker and T. Kanade, “Limits on super-resolution and how to break them,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.9, pp.1167–1183, Sept. 2002.
 - [16] D. Zhou, X. Shen, and W. Dong, “Image zooming using directional cubic convolution interpolation,” IET Image Processing, vol.6, no.6, pp.627–634, Aug. 2012.
 - [17] 井上恵太, 村平宏太, 田口 亮, “ジャギーが生じない修正双線形補間法,” 電気学会論文誌C (電子・情報・システム部門誌), vol.131, no.6, pp.1273–1274, 2011.
 - [18] X. Zhang and X. Wu, “Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation,” IEEE Transactions on Image Processing, vol.17, no.6, pp.887–896, June 2008.
 - [19] K.W. Hung and W.C. Siu, “Robust soft-decision interpolation using weighted least squares,” IEEE Transactions on Image Processing, vol.21, no.3, pp.1061–

1069, March 2012.

- [20] K.W. Hung and W.C. Siu, “Fast image interpolation using the bilateral filter,” IET Image Processing, vol.6, no.7, pp.877–890, Oct. 2012.
- [21] W. Freeman, E. Pasztor, and O. Carmichael, “Learning low-level vision,” International journal of computer vision, vol.40, no.1, pp.25–47, 2000.
<http://link.springer.com/article/10.1023/A:1026501619075>
- [22] J. Sun, N.-N. Zheng, H. Tao, and H.-Y. Shum, “Image hallucination with primal sketch priors,” 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., vol.2, pp.II–729, June 2003.
- [23] H. Chang, D. Yeung, and Y. Xiong, “Super-resolution through neighbor embedding,” Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2004.
<http://ieeexplore.ieee.org/document/1315043>
- [24] R. Timofte, V. De Smet, and L. Van Gool, “A+: Adjusted anchored neighborhood regression for fast super-resolution,” Asian Conference on Computer Vision Springer, pp.111–126 2014.
- [25] Y. Romano, J. Isidoro, and P. Milanfar, “RAISR: Rapid and accurate image super resolution,” IEEE Transactions on Computational Imaging, vol.3, no.1, pp.110–125, March 2017.
- [26] C. Dong, C.C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” CoRR, vol.abs/1501.00092, 2015.
<http://arxiv.org/abs/1501.00092>
- [27] J. Kim, J.K. Lee, and K.M. Lee, “Accurate image super-resolution using very deep convolutional networks,” CoRR, vol.abs/1511.04587, 2015.
<http://arxiv.org/abs/1511.04587>
- [28] A. Kappeler, S. Yoo, Q. Dai, and A.K. Katsaggelos, “Video super-resolution with convolutional neural networks,” IEEE Transactions on Computational Imaging, vol.2, no.2, pp.109–122, June 2016.
- [29] M. Haris, G. Shakhnarovich, and N. Ukita, “Deep back-projection networks for super-resolution,” CoRR, vol.abs/1803.02735, 2018.
<http://arxiv.org/abs/1803.02735>
- [30] W.T. Freeman, T.R. Jones, and E.C. Pasztor, “Example-based super-

resolution,” IEEE Computer Graphics and Applications, vol.22, no.2, pp.56–65, March 2002.

- [31] A. Choudhury, P. vanBeek, and A. Segall, “Image detail enhancement using a dictionary technique,” 2012 19th IEEE International Conference on Image Processing, pp.977–980, Sept. 2012.
<https://ieeexplore.ieee.org/document/6467025/>
- [32] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” IEEE Transactions on Image Processing, vol.15, no.12, pp.3736–3745, Dec. 2006.
- [33] R. Khan, R. Sablatnig, A. Bais, and Y.M. Khawaja, “Comparison of reconstruction and example-based super-resolution,” 2011 7th International Conference on Emerging Technologies, pp.1–6, Sept. 2011.
- [34] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image super-resolution as sparse representation of raw image patches,” IEEE Conference on Computer Vision and Pattern Recognition, pp.1–8, June 2008.
<http://ieeexplore.ieee.org/document/4587647>
- [35] E.J. Candes and M.B. Wakin, “An introduction to compressive sampling,” IEEE Signal Processing Magazine, vol.25, no.2, pp.21–30, March 2008.
<https://ieeexplore.ieee.org/document/4472240>
- [36] D.L. Donoho, “Compressed sensing,” IEEE Transactions on Information Theory, vol.52, no.4, pp.1289–1306, April 2006.
<https://ieeexplore.ieee.org/document/1614066>
- [37] J. Yang, J. Wright, T.S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” IEEE Transactions on Image Processing, vol.19, no.11, pp.2861–2873, Nov. 2010.
<https://ieeexplore.ieee.org/document/5466111/>
- [38] Y. Yitzhaky and N.S. Kopeika, “Identification of blur parameters from motion blurred images,” Graphical Models and Image Processing, vol.59, no.5, pp.310–320, 1997.
<http://www.sciencedirect.com/science/article/pii/S1077316997904354>
- [39] R. Lokhande, K.V. Arya, and P. Gupta, “Identification of parameters and restoration of motion blurred images,” Proceedings of the 2006 ACM Symposium on Applied Computing, pp.301–305, ACM, New York, NY, USA, 2006.

<http://doi.acm.org/10.1145/1141277.1141347>

- [40] Q. Li and Y. Yoshida, “Parameter estimation and restoration for motion blurred images,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E80-A, no.8, pp.1430–1437, 1997.
- [41] M.E. Moghaddam and M. Jamzad, “Motion blur identification in noisy images using fuzzy sets,” Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology, 2005., pp.862–866, Dec. 2005.
<https://ieeexplore.ieee.org/document/1577212/>
- [42] 米司健一, 田中正行, 奥富正敏, “直線的手ぶれ画像復元のための PSF パラメータ推定手法,” 情報処理学会論文誌コンピュータビジョンとイメージメディア(CVIM), vol.47, no.9, pp.107–110, June 2006.
- [43] B. Scholkopf, J. Platt, and T. Hofmann, “Blind motion deblurring using image statistics,” pp.841–848, MIT Press, 2007.
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6287478>
- [44] N. Joshi, R. Szeliski, and D. Kriegman, “PSF estimation using sharp edge prediction,” IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008., pp.1–8, June 2008.
- [45] K. Patanukhom and A. Nishihara, “Identification of piecewise linear uniform motion blur,” TENCON 2007 - 2007 IEEE Region 10 Conference, pp.1–4, Oct. 2007.
- [46] Q. Shan, W. Xiong, and J. Jia, “Rotational motion deblurring of a rigid object from a single image,” IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007., pp.1–8, Oct. 2007.
- [47] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, and W.T. Freeman, “Removing camera shake from a single photograph,” ACM Trans. Graph., vol.25, no.3, pp.787–794, July 2006.
<http://doi.acm.org/10.1145/1141911.1141956>
- [48] W.H. Richardson, “Bayesian-based iterative method of image restoration,” JOSA, vol.62, no.1, pp.55–59, 1972.
- [49] L.B. Lucy, “An iterative technique for the rectification of observed distributions,” The astronomical journal, vol.79, p.745, 1974.
- [50] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, “Efficient marginal likelihood optimization in blind deconvolution,” CVPR 2011, pp.2657–2664, June

2011.

- [51] L. Xu and J. Jia, “Two-phase kernel estimation for robust motion deblurring,” Proceedings of the 11th European Conference on Computer Vision: Part I, pp.157–170, ECCV’10, Springer-Verlag, Berlin, Heidelberg, 2010.
<http://dl.acm.org/citation.cfm?id=1886063.1886077>
- [52] L. Xu, S. Zheng, and J. Jia, “Unnatural l0 sparse representation for natural image deblurring,” 2013 IEEE Conference on Computer Vision and Pattern Recognition, pp.1107–1114, June 2013.
- [53] S. Cho and S. Lee, “Fast motion deblurring,” ACM Trans. Graph., vol.28, no.5, pp.145:1–145:8, Dec. 2009.
<http://doi.acm.org/10.1145/1618452.1618491>
- [54] K. Rank, M. Lendl, and R. Unbehauen, “Estimation of image noise variance,” IEE Proceedings - Vision, Image and Signal Processing, vol.146, no.2, pp.80–84, Aug. 1999.
<https://ieeexplore.ieee.org/document/788764/>
- [55] S.-C. Tai and S.-M. Yang, “A fast method for image noise estimation using laplacian operator and adaptive edge detection,” 2008 3rd International Symposium on Communications, Control and Signal Processing, pp.1077–1081, March 2008.
<https://ieeexplore.ieee.org/document/4537384/>
- [56] X. Liu, M. Tanaka, and M. Okutomi, “Single-image noise level estimation for blind denoising,” IEEE Transactions on Image Processing, vol.22, no.12, pp.5226–5237, Dec. 2013.
<https://ieeexplore.ieee.org/document/6607209/>
- [57] R.C. Gonzalez and R.E. Woods, Digital Image Processing, 2nd edition, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.
- [58] A. Buades, B. Coll, and J.M. Morel, “A non-local algorithm for image denoising,” 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), vol.2, pp.60–65vol.2, June 2005.
- [59] J. Darbon, A. Cunha, T.F. Chan, S. Osher, and G.J. Jensen, “Fast nonlocal filtering applied to electron cryomicroscopy,” 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp.1331–1334, May 2008.

- [60] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *IEEE Transactions on Image Processing*, vol.16, no.8, pp.2080–2095, Aug. 2007.
- [61] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space,” *2007 IEEE International Conference on Image Processing*, vol.1, pp.I-313–I-316, Sept. 2007.
- [62] K. Dabov, A. Foi, and K. Egiazarian, “Video denoising by sparse 3d transform-domain collaborative filtering,” *2007 15th European Signal Processing Conference*, pp.145–149, Sept. 2007.
- [63] Y. Xiao, T. Zeng, J. Yu, and M.K. Ng, “Restoration of images corrupted by mixed gaussian-impulse noise via l_1-l_0 minimization,” *Pattern Recognition*, vol.44, no.8, pp.1708–1720, 2011.
<http://www.sciencedirect.com/science/article/pii/S0031320311000495>
- [64] W.K. Pratt, “Median filtering,” *Semiannual Report*, Univ. of Southern California, 1975.
<https://ci.nii.ac.jp/naid/10024701835/>
- [65] D.R.K. Brownrigg, “The weighted median filter,” *Commun. ACM*, vol.27, no.8, pp.807–818, Aug. 1984.
<http://doi.acm.org/10.1145/358198.358222>
- [66] V.R. Vijaykumar, D. Ebenezer, and P.T. Vanathi, “Detail preserving median based filter for impulse noise removal in digital images,” *2008 9th International Conference on Signal Processing*, pp.793–796, Oct. 2008.
- [67] Y. Dong and S. Xu, “A new directional weighted median filter for removal of random-valued impulse noise,” *IEEE Signal Processing Letters*, vol.14, no.3, pp.193–196, March 2007.
- [68] B. Xiong and Z. Yin, “A universal denoising framework with a new impulse detector and nonlocal means,” *IEEE Transactions on Image Processing*, vol.21, no.4, pp.1663–1675, April 2012.
- [69] J. Liu, X.C. Tai, H. Huang, and Z. Huan, “A weighted dictionary learning model for denoising images corrupted by mixed noise,” *IEEE Transactions on Image Processing*, vol.22, no.3, pp.1108–1120, March 2013.
- [70] J. Jiang, L. Zhang, and J. Yang, “Mixed noise removal by weighted encoding

with sparse nonlocal regularization,” IEEE Transactions on Image Processing, vol.23, no.6, pp.2651–2662, June 2014.

- [71] T. Huang, W. Dong, X. Xie, G. Shi, and X. Bai, “Mixed noise removal via laplacian scale mixture modeling and nonlocal low-rank approximation,” IEEE Transactions on Image Processing, vol.26, no.7, pp.3171–3186, July 2017.
- [72] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, “Least angle regression,” Ann. Statist., vol.32, no.2, pp.407–499, 04 2004.
<https://doi.org/10.1214/009053604000000067>
- [73] M.A.T. Figueiredo, R.D. Nowak, and S.J. Wright, “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,” IEEE Journal of Selected Topics in Signal Processing, vol.1, no.4, pp.586–597, Dec. 2007.
<https://ieeexplore.ieee.org/document/4407762/>
- [74] A. Chambolle, “An algorithm for total variation minimization and applications,” Journal of Mathematical Imaging and Vision, vol.20, no.1, pp.89–97, Jan. 2004.
- [75] V. Caselles, A. Chambolle, and M. Novaga, “The discontinuity set of solutions of the TV denoising problem and some extensions,” Multiscale modeling & simulation, vol.6, no.3, pp.879–894, 2007.
- [76] J.L. Starck, M. Elad, and D.L. Donoho, “Image decomposition via the combination of sparse representations and a variational approach,” IEEE Transactions on Image Processing, vol.14, no.10, pp.1570–1582, Oct. 2005.
<https://ieeexplore.ieee.org/document/1510691/>
- [77] M.V. Afonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo, “An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems,” IEEE Transactions on Image Processing, vol.20, no.3, pp.681–695, March 2011.
<https://ieeexplore.ieee.org/document/5570998/>
- [78] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, “Non-local sparse models for image restoration,” 2009 IEEE 12th International Conference on Computer Vision, pp.2272–2279, Sept. 2009.
- [79] Z. Lu, H. Li, and W. Li, “Image deblocking via group sparsity optimiza-

- tion,” 2015 IEEE International Symposium on Circuits and Systems (ISCAS), pp.1582–1585, May 2015.
- [80] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Weighted nuclear norm minimization with application to image denoising,” 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp.2862–2869, June 2014.
- [81] S. Ono and I. Yamada, “A convex regularizer for reducing color artifact in color image recovery,” 2013 IEEE Conference on Computer Vision and Pattern Recognition, pp.1775–1781, June 2013.
- [82] P.S. Huang, S.D. Chen, P. Smaragdis, and M. Hasegawa-Johnson, “Singing-voice separation from binaural recordings using robust principal component analysis,” 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.57–60, March 2012.
- [83] I.-Y. Jeong and K. Lee, “Vocal separation using extended robust principal component analysis with schatten p/lp -norm and scale compression,” 2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), pp.1–6, Sept. 2014.
- [84] Y. Ikemiya, K. Itoyama, and K. Yoshii, “Singing voice separation and vocal F0 estimation based on mutual combination of robust principal component analysis and subharmonic summation,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol.24, no.11, pp.2084–2095, Nov. 2016.
- [85] C. O’Brien and M.D. Plumbley, “Automatic music transcription using low rank non-negative matrix decomposition,” 2017 25th European Signal Processing Conference (EUSIPCO), pp.1848–1852, Aug. 2017.
- [86] S.G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” IEEE Transactions on Signal Processing, vol.41, no.12, pp.3397–3415, Dec. 1993.
<https://ieeexplore.ieee.org/document/258082/>
- [87] Y.C. Pati, R. Rezaifar, and P.S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, pp.40–44vol.1, Nov. 1993.
<https://ieeexplore.ieee.org/document/342465/>
- [88] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli, “Image denoising

using scale mixtures of gaussians in the wavelet domain,” IEEE Transactions on Image processing, vol.12, no.11, pp.1338–1351, 2003.

- [89] S. Roth and M.J. Black, “Fields of experts: A framework for learning image priors,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005., vol.2IEEE, pp.860–867 2005.
<https://ieeexplore.ieee.org/document/1467533/>
- [90] A. Levin, A. Zomet, and Y. Weiss, “Learning how to inpaint from global image statistics,” Proceedings Ninth IEEE International Conference on Computer Vision, p.305, 2003.
- [91] N. Parikh and S. Boyd, Proximal Algorithms, vol.1, Foundations and Trends in Optimization, 2014.
- [92] M.R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, vol.49, NBS Washington, DC, 1952.
- [93] C.C. Paige and M.A. Saunders, “LSQR: An algorithm for sparse linear equations and sparse least squares,” ACM Trans. Math. Softw., vol.8, no.1, pp.43–71, March 1982.
<http://doi.acm.org/10.1145/355984.355989>
- [94] J. Sun, N.-n. Zheng, and H.-y. Shum, “Image hallucination with primal sketch priors,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.2, no.c, pp.II–729–36, 2003.
[http://ieeexplore.ieee.org/document/1211539](https://ieeexplore.ieee.org/document/1211539)
- [95] D. Donoho, “For most large underdetermined systems of linear equations the minimal L1-norm solution is also the sparsest solution,” Communications on pure and applied mathematics, 2006.
<http://onlinelibrary.wiley.com/doi/10.1002/cpa.20132/abstract>
- [96] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” Sixth International Conference on Computer Vision, 1998., pp.839–846, Jan. 1998.
- [97] S. Paris, P. Kornprobst, and J. Tumblin, Bilateral Filtering, Now Publishers Inc., Hanover, MA, USA, 2009.
- [98] C. Ludusan, O. Lavialle, S. Pop, R. Terebes, and M. Borda, “Image enhancement using a new shock filter formalism,” Acta Technica Napocensis, Electronics and Telecommunications, vol.50, no.3, 2009.

- [99] 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.
<http://doi.acm.org/10.1145/1360612.1360672>
- [100] Y. Wang and W. Yin, “Compressed sensing via iterative support detection,” Rice University CAAM Technical Report, 2009.
- [101] K. He, J. Sun, and X. Tang, “Guided image filtering,” IEEE transactions on pattern analysis & machine intelligence, vol.35, no.6, pp.1397–1409, 2013.
- [102] J.S. Lim, Two-dimensional Signal and Image Processing, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1990.
- [103] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, “Image deblurring with blurred/noisy image pairs,” ACM Trans. Graph., vol.26, no.3, July 2007.
<http://doi.acm.org/10.1145/1276377.1276379>
- [104] N. Asuni and A. Giachetti, “TESTIMAGES: a large-scale archive for testing visual devices and basic image processing algorithms,” Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference, ed. by A. Giachetti, pp.63–70, The Eurographics Association, 2014.
<http://dx.doi.org/10.2312/stag.20141242>
- [105] D. Krishnan and R. Fergus, “Fast image deconvolution using hyper-laplacian priors,” Proceedings of the 22Nd International Conference on Neural Information Processing Systems, pp.1033–1041, NIPS’09, Curran Associates Inc., USA, 2009.
<http://dl.acm.org/citation.cfm?id=2984093.2984210>

謝辞

本研究は、著者が慶應義塾大学 大学院 理工学研究科 後期博士課程在学中に行ったものである。本論文をまとめるにあたり、ご指導、ご助言を厚く賜りました指導教員および本論文の主査である慶應義塾大学 理工学部 池原 雅章 教授に心から感謝申し上げます。

またご多忙な中、本論文の副査を快く引き受けてくださった上、多くの貴重なご意見を賜りました慶應義塾大学 理工学部 岡田 英史 教授、斎藤 英雄 教授、湯川 正裕 准教授に厚く御礼申し上げます。

最後になりますが、同期である博士課程 山口 拓郎 氏を始めとする、様々な面で私を支えていただいた池原研究室のすべての諸兄に深謝し、また陰ながら心の支えとなってくれた家族という貴重な存在に対してここに感謝の意を示します。