



Title	単板式デジタルカメラのための画質改善法に関する研究
Author(s)	河野, 克也
Citation	北海道大学. 博士(情報科学) 甲第13075号
Issue Date	2018-03-22
DOI	10.14943/doctoral.k13075
Doc URL	http://hdl.handle.net/2115/70201
Type	theses (doctoral)
File Information	Katsuya_Kono.pdf



[Instructions for use](#)

単板式デジタルカメラのための
画質改善法に関する研究

北海道大学 大学院情報科学研究科
情報理工学専攻
数理科学講座 情報数理学研究室
河野 克也

目次

第 1 章	序論	4
1.1	本研究の背景	4
1.2	本研究の目的	7
1.3	本論文の構成	8
第 2 章	単板式デジタルカメラによる写真撮影	11
2.1	はじめに	11
2.2	撮像部による画像取得	11
2.3	画像処理エンジンによる現像処理	13
2.4	現像画像の保存	14
第 3 章	ブロックベース線形回帰によるデモザイキング	19
3.1	はじめに	19
3.2	既存のデモザイキング手法	19
3.3	ブロックベース線形回帰	25
3.4	線形回帰を利用した欠損データ推定手法	26

3.5	線形回帰手法のデモザイキングへの適用	28
3.6	欠損要素に対する複数の推定値候補の扱い	32
3.7	計算効率の改善	36
第 4 章	ブロックベース線形回帰によるデモザイキング手法に関する数値実験	42
4.1	実験用画像	42
4.2	提案手法に関する実験条件	43
4.3	OBS 法, MV 法による推定に関する実験	46
4.4	既存手法との性能比較実験	49
第 5 章	JPEG 再適用法の改良法による画質改善手法	55
5.1	JPEG 圧縮アルゴリズム	55
5.2	既存の画質改善手法	61
5.3	画像データベースを利用した JPEG 再適用法	66
第 6 章	JPEG 再適用法の改良法に関する数値実験	73
6.1	実験用画像	73
6.2	画像特徴量に関する実験	74
6.3	画像特徴量とカーネル関数の組み合わせに関する実験	82
6.4	既存手法との性能比較実験	85
第 7 章	総括	100
7.1	本論文による成果	100

7.2	今後の課題	102
	謝辞	103
	参考文献	104

第 1 章

序論

1.1 本研究の背景

1988年に写真画像をデジタル形式で記録するカメラが初めて発表されて以来、企業競争による性能向上やコンピュータの普及によりデジタルカメラの市場は拡大し、2002年には出荷数が従来の銀塩カメラを上回り [1]、現在ではデジタルカメラは銀塩カメラに代わる重要な写真撮影デバイスとなっている。

デジタルカメラと銀塩カメラの違いは撮影画像の記録、および現像方法にある。銀塩カメラでは、フィルム上にレンズから集めた光を結像させ、フィルムに塗布された銀塩感光材料を化学変化させることで画像を記録し、フィルムに対して薬品処理を施すことにより記録した画像を可視化する。一方でデジタルカメラは、フィルムではなく撮像素子に結像させることによって、電気信号に変換したのち、デジタルデータとして画像情報を取得する。取得されたデータは基本的にカメラ内で人間の目で見られる形式に変換され、メモリーカード等の記録媒体に保存される。

このような仕組みの違いにおいて、デジタルカメラには様々な利点が存在する。例え

ば、銀塩カメラで用いられるフィルムは使い切りであり、フィルムの最後まで記録が行われたら新たなフィルムを購入する必要があるが、デジタルカメラでは記録媒体内のデータが一杯になっても、画像データを他に移したり不要な画像データを削除することで、繰り返し撮影を行なうことが可能である。また、フィルムに記録された画像は薬品処理による現像を行なうまで確認できないが、デジタルカメラでは液晶パネルが搭載されており、撮影したその場で液晶画面で撮影画像を確認することができ、撮影に失敗しても気軽に撮り直しを行なうことが可能である。さらに、インターネットを通じて撮影画像を電子メールで送信したり、Web で発信することも容易である。

デジタルカメラによる写真の撮影方式には、大きく分けて単板方式と多板方式の二つが存在する。単板方式は、単一の撮像素子を用いて各色成分の取得を行なう。撮像素子中のそれぞれのフォトセンサに対し、取得したい色成分に対応するカラーフィルタを規則的に並べ、1画素につき異なる1色の情報を持つ画像 (RAW 画像) を取得したのち、補間処理によりカラー画像を生成する。一方、多板方式では、取得する色成分の数だけ撮像素子を用意し、レンズから集めた光をダイクロイックプリズムにより取得させたい色成分に分解し、それぞれの色成分に対応する素子に感光させることにより、カラー画像を生成する。

多板方式では各画素に対して全ての色成分を取得するため補間処理を行なう必要がなく、高画質、高感度な写真撮影が可能である。しかしながら、カメラ筐体の小型化や製造コスト削減の要請により、現在普及しているデジタルカメラは単板方式を採用しているものが大多数である。

単板方式による写真撮影において、上記の補間処理はデモザイキングと呼ばれている。デモザイキングの性能が撮影画像の画質に大きく影響することから、優れたデモザイキング手法の構築は画像処理分野における重要な研究課題の一つとなっている。自然画像は各色成分の間において類似した構造を持つことが多いなどといった、自然画像の特徴を利用した様々なデモザイキング手法が現在までに提案されているが、それらの手法の多くは、補間処理において数学的最適性が不明瞭な過程が存在するという問題を含んでいる。

デジタルカメラにより撮影された画像を記録媒体に保存する際、画像データのサイズの削減のため、撮影画像に対しデータ圧縮処理が施される場合がある。現在は JPEG (Joint Photographic Experts Group)[2] 方式で圧縮を行なうのが一般的であるが、JPEG 方式はデータを完全に復元することが不可能な非可逆圧縮方式であるため、圧縮画像には、ブロックノイズと呼ばれる、画像中の平坦領域に発生するモザイク状のノイズや、モスキートノイズと呼ばれる、主にエッジ領域で発生する霽がかったノイズが混入し、画質の劣化が生じてしまう。圧縮率を上げるに従い、これらのノイズによる画質劣化が顕著となるため、特に高圧縮率で圧縮された JPEG 画像に対して画質改善の需要が存在する。

JPEG 画像の画質改善手法の 1 つに、JPEG 再適用法 [3] と呼ばれる画質改善法が存在する。JPEG 再適用法は、入力画像に対する様々な方向への画像シフト操作と JPEG の再適用により、ノイズの発生位置を移動させた複数の画像を生成し、それらの加重平均を計算することで、ノイズを拡散させ画質改善を図る手法であり、単純な方法でありながら高い画質改善効果を示している。しかしながら、どの入力画像に対しても一定の重み係数を用いているため、入力画像によっては十分な画質改善が行われない可能性がある。

1.2 本研究の目的

本研究は，単板式デジタルカメラにより撮影された画像の画質改善を実現することを目的とし，高性能なデモザイキング手法による高画質カラー画像の生成法の構築，および JPEG 圧縮画像に対する新たな画質改善法の構築を行なう。

デモザイキングにおいては，画像をブロックに分割し，各々の画像ブロックに対して線形回帰を適用することにより補間を行なう，ブロックベース線形回帰 [4, 5] と呼ばれる手法を RAW 画像に対して適用することで，効果的な補間を行なう手法を提案する [6]。前述の数学的最適性に関する問題について，当該手法は各々のブロック画像データは同一の確率変数ベクトルから得られた実現値であるという仮定のもと，画像ブロックの期待二乗誤差の最小化という明確な最適化基準の元で補間を行っている点が特徴である。

また，JPEG 画像の画質改善手法については，JPEG 再適用法における加重平均の重み係数の設定法に関する問題に対し，入力画像から算出される画像特徴量を定義した上で，画像データベースの情報を利用することで最適な重み係数を求め，更なる画質改善を試みる 2 種の手法を提案する [7, 8, 9]。1 つは，入力画像から得られる画像特徴量が最も近い画像をデータベースから選択し，当該データベース画像に対し最適な画質改善効果を与える重み係数を利用して画質の改善を行なう方法である。もう 1 つは，画像特徴量から最適な重み係数を与える関数の存在を仮定し，それを推定した上で，当該関数から得られる重み係数を利用してさらなる画質改善を行なう方法である。JPEG 再適用法の画質改善過程より，推定する関数は非線形であると予想されることから，画像特徴量と最適な改善効

果を与える重み係数からなる訓練データをデータベース画像から作成し，非線形なデータに対する解析手法として知られるカーネル線形回帰 [10] を利用して推定を行なう。

1.3 本論文の構成

本論文は，本章を含め全 7 章から構成される。

第 2 章では，単板式デジタルカメラでの写真撮影によりデジタル画像が得られるまでの流れを，デジタルカメラの構成を踏まえて説明する。

第 3 章では，単板式デジタルカメラでの写真撮影におけるデモザイキングに関して，既存の研究を総括したのち，ブロックベース線形回帰手法を用いて補間を行なう手法について述べる．具体的には，線形回帰によるデータ欠損推定手法について述べ，所与の RAW 画像から，一定の大きさで重複を許して切り出した画像ブロック群に対して当該手法を利用した補間を行なうことで，デモザイキング処理を実現する方法を示す．また，重複して画像ブロック切り出す操作に起因する，補間されるべき画素に対して得られる複数の推定値候補を統合するための指針を示す．さらに，補間の過程で必要となる一般逆行列の計算，および前述の複数の推定値候補に対する計算効率を改善する方法についても述べる。

第 4 章では，第 3 章において提案したブロックベース線形回帰手法によるデモザイキング手法について数値実験を実施し，その有効性を検証する．具体的には，補間処理の過程における，切り出すブロックの重複により得られる複数の推定値候補の利用方法について，各々の方法による最終的な推定結果の比較を行う．また，提案手法と既存のデモザイキン

グ手法について、客観的画質評価指標の一つである CPSNR(Color Peak Signal-to-Noise Ratio), および実際のデモザイキング画像の例を用いて性能比較を行なう。

第 5 章では, JPEG 圧縮画像の画質改善手法について述べる。最初に JPEG 圧縮アルゴリズムについて述べたのち, 提案手法の元となった JPEG 再適用法をはじめ, JPEG 画像に対する既存の画質改善手法を総括する。JPEG 再適用法における加重平均の重み係数について, どの画像に対しても一定の値を用いるが故に画質改善性能に限界があるという従来法の問題点を改善すべく, JPEG 画像から算出される画像特徴量と画像データベースを導入し, 入力画像に対して, 画像特徴量が最も近いデータベース画像に対する最適な重み係数を利用する JPEG 再適用法の改良法や, 入力画像の画像特徴量から最適な重み係数を与える関数を, データベース画像を訓練データとして利用したカーネル線形回帰により推定し, 当該関数を利用して算出した重み係数を利用する JPEG 再適用法の改良法について述べる。

第 6 章では, 第 5 章で提案した, 画像データベースを利用した JPEG 再適用法について数値実験を実施し, その有効性を検証する。具体的には, 提案した画像特徴量の有効性を検証すべく, 実験画像と画像データベースの画像に対して, 最適な重み係数同士の距離と各々の特徴量の距離との相関や, 実験画像に対しデータベース画像に対する最適な重み係数を用いて画質改善を行なった際の PSNR 改善値と特徴量の距離との相関の調査を行なう。また, カーネル線形回帰を利用した JPEG 再適用法における, 提案した 4 種のカーネル関数の有効性について, 画質改善画像の PSNR 改善値による評価を行なう。さらに, 提案手法と既存の画質改善手法について, PSNR, SSIM(Structural Similarity

Index) の改善値，および実際の画質改善画像例を用いて性能比較を行なう．

最後に第 7 章では，本論文の総括を行なう．

第 2 章

単板式デジタルカメラによる写真撮影 [1, 11, 12]

2.1 はじめに

本章では、単板式デジタルカメラの構造と、撮影画像が生成、保存される仕組みを述べる。デジタルカメラはレンズ、撮像部、及び画像処理エンジンを通じて撮影画像を生成する。レンズは被写体から発せられた光を集光し、撮影部に光像を形成する働きを持つ。撮像部ではアナログ画像である光像を電気信号に変換することでデジタル画像を取得する。画像処理エンジンでは、得られたデジタル画像に対する補間処理や、劣化や歪み等の補正処理を行ない、人間の目に適した画像に修正する、いわゆる現像処理が行われる。上記の処理により得られた撮影画像は、適宜画像圧縮処理を施したのち記憶媒体に保存される。

2.2 撮像部による画像取得

撮像部は光学フィルタ部と撮像素子部から構成される。

2.2.1 光学フィルタ部

光学フィルタ部は主に光学ローパスフィルタと赤外フィルタから構成され、レンズにより形成された光像から不要な情報の除去を行なう。光学ローパスフィルタは水晶の複屈折性を利用し、光像の高周波成分を減衰させる働きを持つ。この機能により、後述するデモザイキングによるアーティファクトを抑制することが可能となる。一方で、高周波成分の減衰により撮影画像の解像感が失われてしまうという問題も発生するため、鮮明さを優先する理由で、近年では光学ローパスフィルタを搭載していないデジタルカメラも開発されていることを付記しておく。赤外フィルタは人間の目に見えない赤外線を除去する働きを持つ。

2.2.2 撮像素子部

撮像素子部は、光信号を電気信号に変換するフォトセンサを縦横に並べることで作られるフォトセンサアレイに、フォトセンサに入射させる光束を集光させるマイクロレンズアレイ、および色情報を取得するためのカラーフィルタアレイ (CFA) を積層した形で構成され、光学フィルタ部を通した光像からサンプリングを行ない、デジタル画像を生成する。

フォトセンサには CCD (Charge Coupled Device) センサや CMOS (Complementary Metal Oxide Semiconductor) センサが利用されている。CCD センサは主にコンパクトデジタルカメラで採用されており、CMOS センサは主に携帯電話向けのカメラモジュールや一眼レフタイプのデジタルカメラで採用されている。

フォトセンサは光の強さしか感知できないため、カラーフィルタアレイを利用し光像を原色に分解した上で、1つのフォトセンサに対し1色の情報を取得させる。カラーフィルタアレイには、光の三原色である赤、緑、青を採用する RGB 原色フィルタや、シアン、マゼンタ、イエローを採用する CMY 補色フィルタなどの種類が存在し、個別の微小カラーフィルタを一定の規則で配列させた構造を持っている。現在では RGB 原色フィルタの1つである、図 2.1 に示す Bayer CFA[13] が主流となっている。

フォトセンサから得られたアナログ電子信号に対し A/D(Analog/Digital) 変換を行うことでデジタル画像データを取得する。この画像は RAW 画像と呼ばれている。

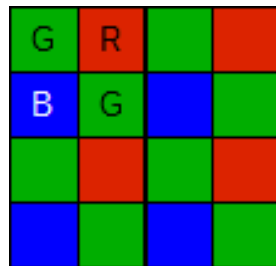


図 2.1: Bayer CFA パターン [13]

2.3 画像処理エンジンによる現像処理

取得された RAW 画像データは 1 画素につき 1 色の情報しか保持していない不完全なものであるため、画像処理エンジンは、最初に補間により RAW 画像を各画素が 3 色すべての色情報を保持するカラー画像に変換する処理を行なう。この処理はデモザイキングと呼ばれている。デモザイキング処理において、本来撮影対象が有していない情報がカラー画像に発生する場合がある。例として、ジッターノイズや偽色が挙げられる。ジッターノ

イズはジャギー状の模様の中で、主に画像のエッジ領域で発生する。ジッパーノイズの例を図 2.2 に示す。被写体の衣服において、黄色の塗料が付着している部分との輪郭に細かい刻み目のような模様が確認できる。偽色は撮影対象が有していない色が発生する現象である。偽色の例を図 2.3 に示す。被写体の灰色のカーペットにおいて、緑色が写り込んでいることが確認できる。このようなアーティファクトの発生を避け高画質なカラー画像を得るために、様々なデモザイキング手法の研究がなされている。

デモザイキングを行なったのち、必要に応じて撮影素子の特性や素子表面に付着したごみ等に起因する画像ノイズの除去、レンズの性質により劣化した光像により発生するシェーディングや歪曲収差、倍率色収差の補正、階調や色の調整、手ブレの補正等の処理を行ない、現像処理が完了する。以上の現像処理は基本的にデジタルカメラの内部で行われるが、より高画質な画像を得るために、RAW 画像データの現像処理を PC などデジタルカメラ外で実行する場合もある。

2.4 現像画像の保存

現像された画像はデジタルカメラ内の記録媒体に保存される。画像データをより多く保存するため、撮影画像に対して圧縮処理を施す場合がある。圧縮形式としては JPEG 方式が用いられるのが一般的であるが、JPEG 方式は非可逆な圧縮方式であるため、保存された画像には圧縮に伴うノイズが混入し、画質が劣化してしまう。JPEG 圧縮により混入するノイズの例を図 2.4 に示す。図 2.4(a) の画像において、背景がモザイク状に見えることが確認でき、このようなノイズはブロックノイズと呼ばれている。また、被写体の

ペンの縁に靄がかかった模様が発生していることが確認でき、これはモスキートノイズと呼ばれている。

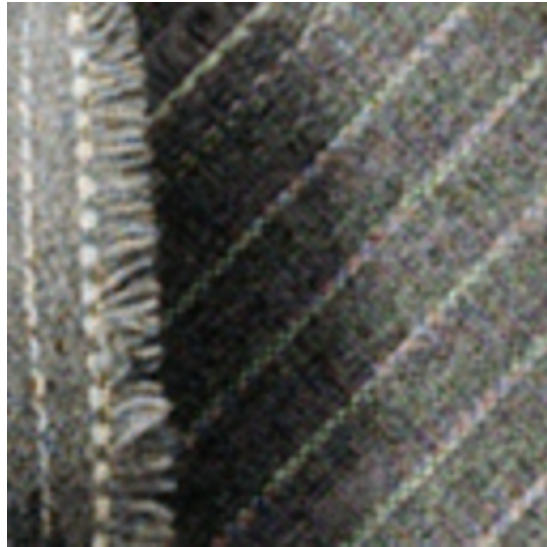


(a) デモザイキング画像

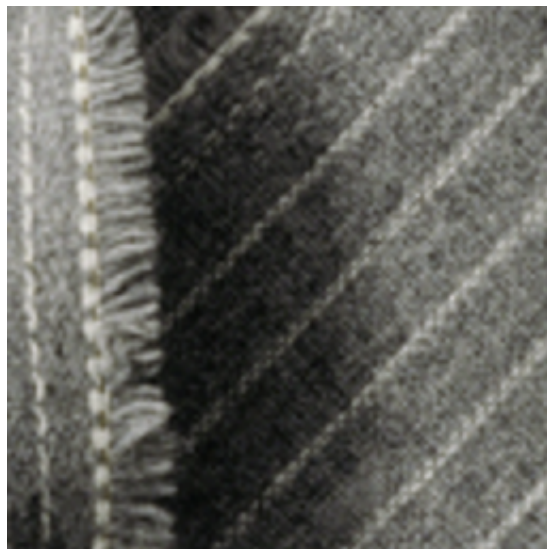


(b) 本来の撮影画像

図 2.2: ジッパーノイズ



(a) デモザイキング画像



(b) 本来の撮影画像

図 2.3: 偽色



(a) JPEG による圧縮画像



(b) 原画像

図 2.4: JPEG 圧縮に起因するノイズ

第 3 章

ブロックベース線形回帰によるデモザイキング [6]

3.1 はじめに

本章では，単板式デジタルカメラによる画像撮影過程の一つであるデモザイキング処理に関して，既存のデモザイキング手法を総括したのち，ブロックベース線形回帰 [4, 5] を利用した画像補間によりデモザイキングを行なう手法を提案する [6].

3.2 既存のデモザイキング手法

第 2 章で述べたように，現在最も普及している CFA パターンは，Bayer CFA [13] であり，現在に至るまで提案されてきたデモザイキング手法の多くは Bayer CFA パターンに対するものである．最も単純な方法は各色成分に対し独立にバイリニア補間 [14] を適用するものである．ほとんどの手法では，自然画像が各色成分間において類似した特徴を持つという性質に着目し，色成分の補間において他色の情報も利用している [15, 16, 17, 18]. とりわけ，緑色と赤色，緑色と青色に対する色差成分に対して補間を行なうことでデモザ

イキングを実現するものが多く、これは色差画像が比較的滑らかなものとなることから、補間画像に発生するアーティファクトを効果的に抑制できるという理由による。最近では色差補間後に画像の持つ非局所的冗長性を利用した処理を行なう手法 [19, 20] や、色差補間が有効に機能するための条件を解析し、画素の特徴に応じて色差補間と色成分内でのバイリニア補間を効果的に組み合わせる手法 [21] が提案されている。また、緑色成分の情報から赤色・青色成分を推定した際の残差に対して補間を行う手法も提案されている [22, 23, 24, 25]。ここでは、文献 [19, 22] の手法について詳細を述べる。

3.2.1 非局所的冗長性を利用した手法 [19]

Zhang らは、補間対象となる画素について、色差と色成分の勾配の情報を利用して補間を行ったのち、補間対象の画素を中心とするブロックと類似したブロックを画像中から集め、それらの情報を利用した後処理を行なうことで、デモザイキングを実現する手法を提案している。

緑色成分の補間

図 3.1 はデモザイキングを行なう RAW 画像の一部を表している。はじめに緑色成分の補間を行なうが、ここでは赤色成分 R_0 を取得した画素に対する補間方法を述べる。まず、 R_0 の画素と隣接する上下左右の画素における緑色と赤色の色差成分を

$$\begin{cases} d_{gr}^n = G_2 - (R_0 + R_{10})/2 \\ d_{gr}^s = G_4 - (R_0 + R_{12})/2 \\ d_{gr}^w = G_1 - (R_0 + R_9)/2 \\ d_{gr}^e = G_3 - (R_0 + R_{11})/2 \end{cases} \quad (3.1)$$



図 3.1: RAW 画像の一部分

と計算する．次に，これらの色差成分に対し，それぞれの方向における勾配により重みづけした加重平均を計算することで， R_0 の画素における緑色と赤色の色差成分を

$$\hat{d}_{gr} = \bar{w}_n d_{gr}^n + \bar{w}_s d_{gr}^s + \bar{w}_w d_{gr}^w + \bar{w}_e d_{gr}^e, \quad (3.2)$$

$$\bar{w}_n = \frac{\nabla_n^{-1}}{C_{gr}}, \bar{w}_s = \frac{\nabla_s^{-1}}{C_{gr}}, \bar{w}_w = \frac{\nabla_w^{-1}}{C_{gr}}, \bar{w}_e = \frac{\nabla_e^{-1}}{C_{gr}}, \quad (3.3)$$

$$\begin{cases} \nabla_n = |G_2 - G_4| + |R_0 - R_{10}| + \frac{1}{2}|G_1 - G_{14}| + \frac{1}{2}|G_3 - G_{15}| + \epsilon \\ \nabla_s = |G_2 - G_4| + |R_0 - R_{12}| + \frac{1}{2}|G_1 - G_{19}| + \frac{1}{2}|G_3 - G_{18}| + \epsilon \\ \nabla_w = |G_1 - G_3| + |R_0 - R_9| + \frac{1}{2}|G_2 - G_{13}| + \frac{1}{2}|G_4 - G_{20}| + \epsilon \\ \nabla_e = |G_1 - G_3| + |R_0 - R_{11}| + \frac{1}{2}|G_2 - G_{16}| + \frac{1}{2}|G_4 - G_{17}| + \epsilon \\ C_{gr} = \nabla_n^{-1} + \nabla_s^{-1} + \nabla_w^{-1} + \nabla_e^{-1} \end{cases} \quad (3.4)$$

と推定する．なお， ϵ は $\nabla_n^{-1}, \dots, \nabla_e^{-1}$ が 0 になるのを避けるための微小な定数である．

以上より， R_0 の画素における緑色成分は

$$\hat{G}_0 = R_0 + \hat{d}_{gr} \quad (3.5)$$

と計算される．青色成分を取得した画素に対する補間は，上記赤色成分の情報を青色成分に置き換えた上で，同様の手順で行なう．

非局所的冗長性を利用した補間精度向上

以上の手順により補間された全ての緑色成分に対し，NLM (NonLocal Means filtering) や NAT (Nonlocal Adaptive Thresholding) と呼ばれる手法を適用することにより，補

間精度を向上させる．着目する補間画素を $u_1 \in \mathbf{R}$, u_1 を中心とする $s \times s$ サイズのブロックをベクトル形式で表したものを $\mathbf{b}_1 \in \mathbf{R}^{s^2}$ とし, u_1 を中心とする $w \times w (w > s)$ の領域から得られる, 画素 u_i を中心とするブロックを $\mathbf{b}_i (i \in \{1, \dots, N\})$ とする．添字 i は $d_i = s^{-2} \|\mathbf{b}_1 - \mathbf{b}_i\|_1$ の値が小さいブロックから順に割り振られるものとし, $\|\cdot\|_p$ はベクトルの p -ノルムを表す．

NLM では, $n (n \leq N)$ 個のブロック $\mathbf{b}_1, \dots, \mathbf{b}_n$ から, u_i の加重平均

$$\hat{u}_1 = \sum_{i=1}^n \frac{\exp(-d_i/\sigma)}{D} u_i, \quad (3.6)$$

$$D = \sum_{i=1}^n \exp(-d_i/\sigma) \quad (3.7)$$

を新たな補間結果として採用する．ここで, σ は重み調整パラメータ, D は重み係数の正規化を行なう係数を表す．

NAT では, n 個のブロック $\mathbf{b}_1, \dots, \mathbf{b}_n$ から, 行列 $P = [\mathbf{b}_1, \dots, \mathbf{b}_n] \in \mathbf{R}^{s^2 \times n}$ を構成する．ここで, 各 \mathbf{b}_i に対応する真のブロック \mathbf{c}_i は, P の共分散行列の固有ベクトルを並べた行列 $\Phi \in \mathbf{R}^{s^2 \times s^2}$ とスパース行列 $A = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n] \in \mathbf{R}^{s^2 \times n}$ を用いて, $Q = [\mathbf{c}_1, \dots, \mathbf{c}_n] = \Phi A$ と表せるものと仮定し,

$$\operatorname{argmin}_A \|A\|_1 \text{ s.t. } \|P - \Phi A\|_F \leq \tau \quad (3.8)$$

を解き, $\hat{\mathbf{c}}_1 = \Phi \boldsymbol{\alpha}_1$ から新たな補間結果 \hat{u}_1 を得る．なお, τ は微小定数であり, $\|\cdot\|_F$ は行列のフロベニウスノルムを表す．

赤色・青色成分の補間

緑色成分の補間を終えた後, 赤色および青色成分の補間を行なう．ここでは, 青色成分の補間について述べる．まず赤色成分を取得した画素に対する補間を行なうが, R_0 の画

素を例に述べる。 R_0 の画素に隣接する斜め上、斜め下の画素それぞれにおける青色と緑色の色差成分を

$$\begin{cases} d_{br}^{nw} = B_5 - \hat{G}_5 \\ d_{br}^{ne} = B_6 - \hat{G}_6 \\ d_{br}^{se} = B_7 - \hat{G}_7 \\ d_{br}^{sw} = B_8 - \hat{G}_8 \end{cases} \quad (3.9)$$

と計算する。次に、色差成分に対し、それぞれの方向における勾配により重みづけした加重平均を計算することで、 R_0 の画素における緑色と青色の色差成分を

$$\hat{d}_{br} = \bar{w}_{nw}d_{br}^{nw} + \bar{w}_{ne}d_{br}^{ne} + \bar{w}_{se}d_{br}^{se} + \bar{w}_{sw}d_{br}^{sw}, \quad (3.10)$$

$$\bar{w}_{nw} = \frac{\nabla_{nw}^{-1}}{C_{br}}, \bar{w}_{ne} = \frac{\nabla_{ne}^{-1}}{C_{br}}, \bar{w}_{se} = \frac{\nabla_{se}^{-1}}{C_{br}}, \bar{w}_{sw} = \frac{\nabla_{sw}^{-1}}{C_{br}}, \quad (3.11)$$

$$\begin{cases} \nabla_{nw} = |B_5 - B_7| + |R_{21} - R_0| + |\hat{G}_5 - \hat{G}_0| + \epsilon \\ \nabla_{ne} = |B_6 - B_8| + |R_{22} - R_0| + |\hat{G}_6 - \hat{G}_0| + \epsilon \\ \nabla_{se} = |B_5 - B_7| + |R_{23} - R_0| + |\hat{G}_7 - \hat{G}_0| + \epsilon \\ \nabla_{sw} = |B_6 - B_8| + |R_{24} - R_0| + |\hat{G}_8 - \hat{G}_0| + \epsilon \\ C_{br} = \nabla_{nw}^{-1} + \nabla_{ne}^{-1} + \nabla_{se}^{-1} + \nabla_{sw}^{-1} \end{cases} \quad (3.12)$$

と推定する。以上より、 R_0 の画素における青色成分は

$$\hat{B}_0 = \hat{G}_0 + \hat{d}_{br} \quad (3.13)$$

と計算される。

最後に、緑色成分を取得した画素に対する青色成分の補間は、 G_1 の画素を例として、

$$\hat{B}_1 = G_1 + \hat{d}_{bg}, \quad (3.14)$$

$$\hat{d}_{bg} = \bar{w}_n d_{bg}^n + \bar{w}_s d_{bg}^s + \bar{w}_w d_{bg}^w + \bar{w}_e d_{bg}^e, \quad (3.15)$$

$$\begin{cases} d_{bg}^n = B_5 - \hat{G}_5 \\ d_{bg}^s = B_8 - \hat{G}_8 \\ d_{bg}^w = \hat{B}_9 - \hat{G}_9 \\ d_{bg}^e = \hat{B}_0 - \hat{G}_0 \end{cases}, \quad (3.16)$$

$$\bar{w}_n = \frac{\nabla_n^{-1}}{C_{bg}}, \bar{w}_s = \frac{\nabla_s^{-1}}{C_{bg}}, \bar{w}_w = \frac{\nabla_w^{-1}}{C_{bg}}, \bar{w}_e = \frac{\nabla_e^{-1}}{C_{bg}}, \quad (3.17)$$

$$\begin{cases} \nabla_n = |G_{14} - G_1| + |B_5 - B_8| + \frac{1}{2}|R_{21} - R_9| + \frac{1}{2}|R_{10} - R_0| + \epsilon \\ \nabla_s = |G_{19} - G_1| + |B_5 - B_8| + \frac{1}{2}|R_{24} - R_9| + \frac{1}{2}|R_{12} - R_0| + \epsilon \\ \nabla_w = |G_1 - G_{26}| + |R_0 - R_9| + \frac{1}{2}|B_{27} - B_5| + \frac{1}{2}|B_{25} - B_8| + \epsilon \\ \nabla_e = |G_1 - G_3| + |R_0 - R_9| + \frac{1}{2}|B_5 - B_6| + \frac{1}{2}|B_8 - B_7| + \epsilon \\ C_{bg} = \nabla_n^{-1} + \nabla_s^{-1} + \nabla_w^{-1} + \nabla_e^{-1} \end{cases} \quad (3.18)$$

の形で行われる。赤色成分も同様の手順で補間を行なった後、赤色・青色成分それぞれに対し NLM または NAT を適用し、デモザイキングを完了する。

3.2.2 残差補間を利用した手法 [22]

禧久らは、RAW 画像に対しはじめに緑色成分の補間を行なった後、緑色成分に対して Guided フィルタ [26] を適用することにより赤色・青色成分の仮の補間を行なった上で、これらの仮の色成分と撮影により取得している赤色・青色成分との残差について改めて補間を行なうことで、デモザイキングを実現する手法を提案している。

Guided フィルタによる仮補間

補間を終えた緑色成分からなる画像から、 $\alpha \times \beta$ の大きさの画像ブロックを切り出し、それぞれの画像ブロックに対して、赤色成分を

$$\hat{R}_{i,j} = aG_{i,j} + b \quad (i \in \{1, \dots, \alpha\}, j \in \{1, \dots, \beta\}) \quad (3.19)$$

と推定する。ここで、 $G_{i,j}$ は画像ブロック中の (i, j) 要素の緑色成分を表し、 a, b は

$$J(a, b) = \sum_{(i,j) \in S} \left(\hat{R}_{i,j} - R_{i,j} \right)^2 \quad (3.20)$$

を最小とするように決定する。 $R_{i,j}$ は画像ブロック中の (i, j) 要素における赤色成分の真の値であり、 S は撮影により $R_{i,j}$ が既知の (i, j) からなる集合である。全てのブロックに

対して推定を行なうことで、RAW 画像全体における赤色成分の仮補間が完了する。青色成分も同様の手順で仮補間が行われる。

残差補間

撮影で得られている赤色・青色成分それぞれに対し、仮補間画像との残差について、バイリニア補間 [14] を適用する。その後、得られた残差成分画像に対し、仮補間画像の画素値を加えることで、それぞれの色成分の補間を完了する。

3.3 ブロックベース線形回帰 [4, 5]

前述の既存手法はいずれも高いデモザイキング性能を示しているが、色差補間や残差補間、勾配の利用といった、画素の補間において数学的な最適性が不明瞭な過程が存在する。ブロックベース線形回帰は観測画像から一定の大きさのブロックを切り出し、各々のブロック画像データは同一の確率変数ベクトルから得られた実現値であるという仮定を置いた上で、推定画像ブロックと未知の真の画像ブロックの期待二乗誤差の最小化という明確な最適化基準の元で画素の補間を行なうものである。

ブロックベース線形回帰による補間は固有空間 BPLP (Back Projection for Lost Pixels) 法 [27] に端を発している。固有空間 BPLP 法は切り出したブロックデータのうち、欠損を含まないデータの主成分分析により、いくつかの主要な主成分が張る部分固有空間を構成し、当該部分固有空間の基底の線形結合により欠損を含むブロックを推定する方法である。田中らは固有空間 BPLP 法において、推定精度の改善を実現し [28]、主成分分析を適用する分散共分散行列と欠損ブロックに対応する分散共分散行列が等しいとい

う条件のもとでは、全ての主成分を用いた改良固有空間 BPLP 法に対応する解が期待二乗誤差最小の意味での最適解であり、当該最適解は固有空間を利用せずとも、上記の線形回帰手法により得られることを明らかにした [29].

自然画像が持つ特徴とブロックベース線形回帰の有効性について、自然画像は局所的に似通った構造を持ち、画像をブロックに分割した際にブロック同士に高い相関があることが多い。本手法での当該画像ブロック群への確率構造の導入はこの性質に沿っていると考え、自然画像の補間問題の 1 つであるデモザイキングにおいて有効に機能すると考えられる。

3.4 線形回帰を利用した欠損データ推定手法 [29]

$\mathcal{I} = \{1, \dots, n\}$ を添字集合とし、 $\mathcal{Y} = \{\mathbf{y}_i \in \mathbf{R}^d \mid i \in \mathcal{I}\}$ を d 次元ベクトルで表される観測データの集合とする。観測データには欠損が含まれているものもあり、欠損要素の値は 0 として扱う。 \mathcal{J} を \mathcal{Y} のうち欠損を含むものの添字から構成される \mathcal{I} の部分集合とする。また、 $\mathcal{X} = \{\mathbf{x}_i \in \mathbf{R}^d \mid i \in \mathcal{I}\}$ を \mathbf{y}_i に対応する真のデータの集合とする。このとき、 \mathbf{y}_i の非欠損部分に対応する位置の値が 1、欠損部分に対応する位置の値が 0 であるような対角行列 $A_i \in \mathbf{R}^{d \times d}$ を用いて、

$$\mathbf{y}_i = A_i \mathbf{x}_i \quad (3.21)$$

と表現できる。目的は、欠損を含む観測データ $\mathbf{y}_j (j \in \mathcal{J})$ の欠損位置、すなわち A_j が既知であるという条件のもとで \mathbf{y}_j と A_j から \mathbf{x}_j を推定することである。

ここで、欠損の推定のために線形回帰モデルの導入を行なう。 I_d を d 次元単位行列、 $\text{rank}(A_j) = r_j < d$ とし、 $M_j \in \mathbf{R}^{r_j \times d}$ を A_j の対角成分が 0 となる行を除いて得られる

行列, $N_j \in \mathbf{R}^{(d-r_j) \times d}$ を $I_d - A_j$ の対角成分が 0 となる行を除いて得られる行列と定める. 例えば,

$$A_j = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

のとき,

$$M_j = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad N_j = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

である. M_j, N_j はそれぞれ \mathbf{x}_j の非欠損部分に対応する値, \mathbf{x}_j の欠損部分に対応する値を抽出する作用を持ち,

$$M_j' M_j + N_j' N_j = I_d, \quad N_j M_j' = O$$

を満たす. ' は転置作用素を, O は零行列をそれぞれ表す. また, \mathcal{X} の要素はすべて同一の確率変数ベクトル \mathbf{x} から得られた実現値であると仮定する.

以上の準備のもとで, 以下の目的関数 $J(B)$ を最小化する行列 $B_j \in \mathbf{R}^{(d-r_j) \times r_j}$ を求める問題として \mathbf{x}_j の推定を定式化する.

$$B_j = \underset{B}{\operatorname{argmin}} J(B), \quad (3.22)$$

$$J(B) = E_{\mathbf{x}_j} \|B M_j \mathbf{y}_j - N_j \mathbf{x}_j\|_2^2. \quad (3.23)$$

$E_{\mathbf{x}_j}$ は \mathbf{x}_j に関する期待値演算を表す. \mathbf{x}_j を \mathbf{x} と同じ確率変数ベクトルとみなし,

$M_j \mathbf{y}_j = M_j \mathbf{x}_j$ であることを利用すると, 式 (3.23) は

$$\begin{aligned} J(B) &= E_{\mathbf{x}_j} \|B M_j \mathbf{y}_j - N_j \mathbf{x}_j\|_2^2 \\ &= E_{\mathbf{x}} \|B M_j \mathbf{x} - N_j \mathbf{x}\|_2^2 \\ &= \operatorname{tr}[(B M_j - N_j) R (B M_j - N_j)'] \end{aligned} \quad (3.24)$$

と変形できる. $\text{tr}[\cdot]$ は行列のトレース, $R \in \mathbf{R}^{d \times d}$ は \mathbf{x} の相関行列

$$R = E_{\mathbf{x}}[\mathbf{x}\mathbf{x}'] \quad (3.25)$$

を表すが, R は未知の情報であるため, 以下のようにして欠損を受けていないデータから事前に推定を行なう.

$$\hat{R} = \frac{1}{|\mathcal{I} - \mathcal{J}|} \sum_{i \in \mathcal{I}, i \notin \mathcal{J}} \mathbf{x}_i \mathbf{x}_i'. \quad (3.26)$$

$|\cdot|$ は集合の要素数を表す. 式 (3.22) の解は

$$B_j = N_j \hat{R} M_j' (M_j \hat{R} M_j')^+ \quad (3.27)$$

で与えられる. ここで, $+$ は行列の MP(Moore-Penrose) 一般逆行列 [30] を表す. これにより, 欠損要素の推定値は B_j を用いて,

$$\mathbf{z}_j = B_j M_j \mathbf{y}_j \quad (3.28)$$

と表され, \mathbf{x}_j の推定値は

$$\hat{\mathbf{x}}_j = M_j' M_j \mathbf{y}_j + N_j' \mathbf{z}_j \quad (3.29)$$

と表される.

なお, 以上の手順で \mathcal{X} の全てのデータの推定を行ったのち, 相関行列 R の再推定

$$\hat{R} = \frac{1}{|\mathcal{I}|} \left(\sum_{i \in \mathcal{I}, i \notin \mathcal{J}} \mathbf{x}_i \mathbf{x}_i' + \sum_{j \in \mathcal{J}} \hat{\mathbf{x}}_j \hat{\mathbf{x}}_j' \right) \quad (3.30)$$

を行ない, 得られた新たな \hat{R} を用いての \mathbf{x}_j の更新を繰り返すことで推定精度の向上を図ることが可能である.

3.5 線形回帰手法のデモザイキングへの適用

前節で述べた線形回帰による欠損要素推定手法の, デモザイキングへの適用方法について述べる.

撮影により得られた $p \times q$ 画素の RAW 画像を $H \in \mathbf{R}^{p \times q \times 3}$ とする。ここで、 H は使用した CFA パターンにより全ての画素において色情報の欠損を含むことを付記しておく。 H から 1 画素ずつ位置を移動させながら大きさ $\alpha \times \beta$ の画像ブロックを切り出し、それらを Y_j ($j \in \mathcal{J} = \mathcal{I}$) とする。添字 j は画像ブロックを切り出す位置に対応しているものとする。この操作により切り出される画像ブロックの個数は

$$n = (p - \alpha + 1)(q - \beta + 1)$$

である。

Y_j の赤色成分、緑色成分、青色成分をそれぞれ $Y_j^R, Y_j^G, Y_j^B \in \mathbf{R}^{\alpha \times \beta}$ とし、また、 Y_j に対し、CFA パターンに従い色情報を取得できている画素に対応する要素が 1、その他の要素が 0 であるようなマスク画像を $S_j \in \mathbf{R}^{\alpha \times \beta \times 3}$ 、 S_j の赤色成分、緑色成分、青色成分をそれぞれそれぞれ $S_j^R, S_j^G, S_j^B \in \mathbf{R}^{\alpha \times \beta}$ とする。さらに、 $\text{vec}(\cdot)$ を行列に対するベクトル化操作 [31]、 $\text{diag}(\cdot)$ をベクトルに対する対角行列化操作とする。すなわち、

$$P = [\mathbf{p}_1 \cdots \mathbf{p}_\gamma] \in \mathbf{R}^{\delta \times \gamma}, \quad \mathbf{q} = [q_1, \dots, q_\delta]' \in \mathbf{R}^\delta$$

に対し、

$$\text{vec}(P) = [\mathbf{p}'_1, \dots, \mathbf{p}'_\gamma]' \in \mathbf{R}^{\delta \gamma}, \quad (3.31)$$

$$\text{diag}(\mathbf{q}) = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & q_\delta \end{bmatrix} \in \mathbf{R}^{\delta \times \delta} \quad (3.32)$$

である。

以上の準備のもとで、

$$\mathbf{y}_j = [\text{vec}(Y_j^R)', \text{vec}(Y_j^G)', \text{vec}(Y_j^B)']', \quad (3.33)$$

$$A_j = \text{diag}([\text{vec}(S_j^R)', \text{vec}(S_j^G)', \text{vec}(S_j^B)']') \quad (3.34)$$

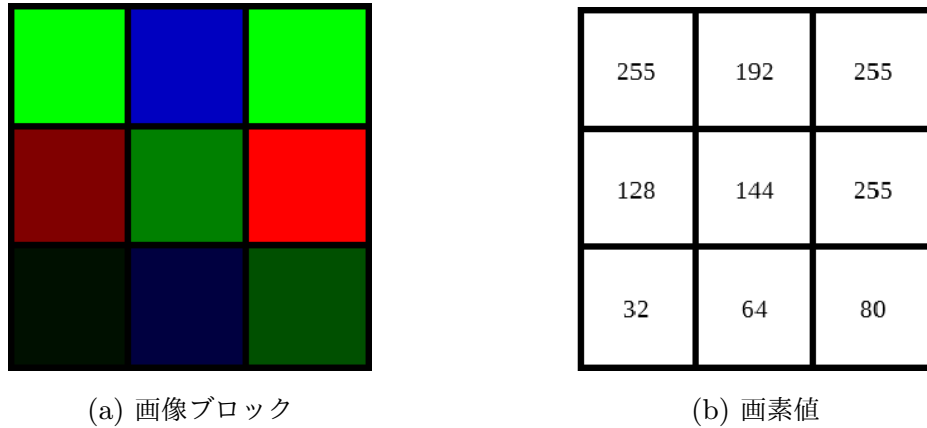


図 3.2: RAW 画像から切り出された 3×3 画像ブロックの例

と対応付ける．図 3.2 は， $\alpha = \beta = 3$ として，RAW 画像から切り出した画像ブロック Y_j の例である．この画像ブロックについて，

$$Y_j^R = \begin{bmatrix} 0 & 0 & 0 \\ 128 & 0 & 255 \\ 0 & 0 & 0 \end{bmatrix}, Y_j^G = \begin{bmatrix} 255 & 0 & 255 \\ 0 & 144 & 0 \\ 32 & 0 & 80 \end{bmatrix}, Y_j^B = \begin{bmatrix} 0 & 192 & 0 \\ 0 & 0 & 0 \\ 0 & 64 & 0 \end{bmatrix},$$

$$S_j^R = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, S_j^G = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, S_j^B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

であり，

$$\mathbf{y}_j = [0 \ 128 \ 0 \ 0 \ 0 \ 0 \ 0 \ 255 \ 0 \ 255 \ 0 \ 32 \ 0 \ 144 \ 0 \ 255 \ 0 \ 80 \ 0 \ 0 \ 0 \ 192 \ 0 \ 64 \ 0 \ 0 \ 0]',$$

$$A_j = \text{diag}([0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 0]')$$

と表される．切り出された各画像ブロック Y_j に対して，前節による線形回帰を用いた補間を通じて所与の RAW 画像 H のデモザイキングを実現する．

なお， $\mathcal{J} = \mathcal{I}$ であるので，式 (3.26) において $i \in \mathcal{I}$, $i \notin \mathcal{J}$ なる \mathbf{x}_i が存在せず，提案手法の枠組みのみでは初回の相関行列 R の推定は不可能である．そのため，提案手法を適用する前に既存の手法によるデモザイキングを行ない，得られたフルカラー画像から R

の推定を行なう必要がある。また、 H 上を 1 画素ずつ移動させながら画像ブロックを切り出すことから、 H の四隅を除く全ての画素は複数の画像ブロックに含まれることになる。そのため、提案手法により全ての画像ブロックの補間を終えた時、1 つの欠損要素に対し当該要素を含むブロックと同数の推定値の候補を得ることができる。これらの複数の推定値候補に対して、次節で紹介する手法を適用することにより、最終的な推定値を決定する。

提案手法を用いて実際にデモザイキングを行なう際の具体的手順を以下に示す。

1. 式 (3.30) による推定精度向上のための相関係数 R の再推定回数を定める。
2. RAW 画像 H に対し、既存の手法によるデモザイキングを適用する。
3. 2. あるいは 5. により得られたカラー画像から画像ブロックを切り出し、相関行列 R の推定を行なう。
4. H から画像ブロックを切り出し、3. で推定した R を用いた線形回帰による補間を行なう。
5. 次節に示す方法により、 H の四隅を除く画素に生じている欠損に対して、得られた複数の推定値候補から最終的な推定値を決定し、カラー画像を取得する。
6. 3.~5. を 1. で定めた回数だけ繰り返し行なう。

3.6 欠損要素に対する複数の推定値候補の扱い

着目欠損要素を含む K 個の画像ブロックデータに対応する、真のデータからなる集合を $\{\mathbf{x}_{j_k} \in \mathbf{R}^{3\alpha\beta} \mid k \in \{1, \dots, K\}, j_k \in \mathcal{J}\}$ とし、着目する欠損は \mathbf{x}_{j_k} の s_k 番目の要素に生じているものとする。ここで、着目欠損要素の真値を u とすると、

$$u = \mathbf{e}'_{s_1} \mathbf{x}_{j_1} = \dots = \mathbf{e}'_{s_K} \mathbf{x}_{j_K} \quad (3.35)$$

となり、各画像ブロックから得られる u の推定値は

$$\hat{u}_k = \mathbf{e}'_{s_k} \hat{\mathbf{x}}_{j_k} \quad (3.36)$$

と表せる。 $\mathbf{e}_{s_k} \in \mathbf{R}^{3\alpha\beta}$ は s_k 番目の要素が 1 の単位ベクトルである。本論文では、各 \hat{u}_k から最終的な欠損要素の推定値を決定する方法について、OBS (Optimal Block Selection) 法、MV (Minimum Variance) 法の 2 つを述べる。

3.6.1 OBS 法 [4]

OBS 法では、各 \hat{u}_k の推定誤差

$$e_k = u - \hat{u}_k \quad (3.37)$$

を平均 0、分散 σ_k^2 に従う確率変数と仮定する。この仮定のもとで、推定誤差の分散 σ_k^2 の推定値が最小である \hat{u}_k を最終的な推定値として採用する。すなわち

$$\hat{u}_{(OBS)} = \hat{u}_{k_{min}}, \quad k_{min} = \underset{k}{\operatorname{argmin}} \sigma_k^2 \quad (3.38)$$

であり、 σ_k^2 については、

$$\hat{\sigma}_k^2 = E_{\mathbf{x}_{j_k}} \|\mathbf{u} - \hat{u}_k\|_2^2 \quad (3.39)$$

と推定される. 式 (3.35) と (3.36) から,

$$u - \hat{u}_k = \mathbf{e}'_{s_k} (\mathbf{x}_{j_k} - \hat{\mathbf{x}}_{j_k}) \quad (3.40)$$

と変形でき, 式 (3.29) により,

$$\mathbf{e}'_{s_k} (\mathbf{x}_{j_k} - \hat{\mathbf{x}}_{j_k}) = \mathbf{e}'_{s_k} (\mathbf{x}_{j_k} - M'_{j_k} M_{j_k} \mathbf{y}_{j_k} - N'_{j_k} \mathbf{z}_{j_k}) \quad (3.41)$$

となる. ここで, $\mathbf{e}'_{s_k} M'_{j_k} M_{j_k} = \mathbf{0}'$, $M_{j_k} \mathbf{y}_{j_k} = M_{j_k} \mathbf{x}_{j_k}$ であることと式 (3.28) から,

$$\begin{aligned} \mathbf{e}'_{s_k} (\mathbf{x}_{j_k} - M'_{j_k} M_{j_k} \mathbf{y}_{j_k} - N'_{j_k} \mathbf{z}_{j_k}) &= \mathbf{e}'_{s_k} (\mathbf{x}_{j_k} - N'_{j_k} B_{j_k} M_{j_k} \mathbf{y}_{j_k}) \\ &= \mathbf{e}'_{s_k} (I_{3\alpha\beta} - N'_{j_k} B_{j_k} M_{j_k}) \mathbf{x}_{j_k} \end{aligned} \quad (3.42)$$

と変形できる. 以上より, 式 (3.39) から,

$$\begin{aligned} \hat{\sigma}_k^2 &= E \mathbf{x}_{j_k} \| \mathbf{e}'_{s_k} (I_{3\alpha\beta} - N'_{j_k} B_{j_k} M_{j_k}) \mathbf{x}_{j_k} \|_2^2 \\ &= E \mathbf{x} \| \mathbf{e}'_{s_k} (I_{3\alpha\beta} - N'_{j_k} B_{j_k} M_{j_k}) \mathbf{x} \|_2^2 \\ &= \mathbf{e}'_{s_k} (I_{3\alpha\beta} - N'_{j_k} B_{j_k} M_{j_k}) R (I_{3\alpha\beta} - N'_{j_k} B_{j_k} M_{j_k})' \mathbf{e}_{s_k} \end{aligned} \quad (3.43)$$

を得る.

3.6.2 MV 法 [5, 32]

MV 法では, 各推定誤差 e_k について互いに無相関であると仮定する. すなわち, $k, \ell \in \{1, \dots, K\}$ ($k \neq \ell$) に対し,

$$E \mathbf{x}_{j_k} \mathbf{x}_{j_\ell} [e_k e_\ell] = 0 \quad (3.44)$$

が成立するものとし, この仮定のもとで誤差の分散が最小となるような \hat{u}_k の凸結合

$$\hat{u}_{(MV)} = \sum_{k=1}^K \alpha_k \hat{u}_k, \quad \sum_{k=1}^K \alpha_k = 1, \quad \alpha_k \geq 0 \quad (3.45)$$

を構成し, 最終的な推定値として採用する. 真値との誤差 $e_{(MV)}$ の分散 $\sigma_{(MV)}^2$ は

$$\hat{\sigma}_{(MV)}^2 = E \mathbf{x}_{j_1} \dots \mathbf{x}_{j_K} \| u - \hat{u}_{(MV)} \|_2^2 \quad (3.46)$$

と推定される。式 (3.45) から、

$$\begin{aligned}
 u - \hat{u}_{(MV)} &= u - \sum_{k=1}^K \alpha_k \hat{u}_k \\
 &= u - \sum_{k=1}^K \alpha_k (u - e_k) \\
 &= \sum_{k=1}^K \alpha_k e_k
 \end{aligned} \tag{3.47}$$

と変形でき、無相関性の仮定と合わせて、式 (3.46) は、

$$\begin{aligned}
 \hat{\sigma}_{(MV)}^2 &= E_{\mathbf{x}_{j_1} \dots \mathbf{x}_{j_K}} \left\| \sum_{k=1}^K \alpha_k e_k \right\|_2^2 \\
 &= \sum_{k=1}^K E_{\mathbf{x}_{j_1} \dots \mathbf{x}_{j_K}} \|\alpha_k e_k\|_2^2 \\
 &= \sum_{k=1}^K \alpha_k^2 \hat{\sigma}_k^2
 \end{aligned} \tag{3.48}$$

と変形できる。

ここで、式 (3.48) の値を最小化する $\alpha_1, \dots, \alpha_K$ の値を求めることを考える。式 (3.45)

より、

$$\alpha_1 = 1 - \sum_{k=2}^K \alpha_k \tag{3.49}$$

であるから、式 (3.48) は、

$$\hat{\sigma}_{(MV)}^2 = \sum_{k=2}^K \alpha_k^2 \hat{\sigma}_k^2 + \left(1 - \sum_{k=2}^K \alpha_k \right)^2 \hat{\sigma}_1^2 \tag{3.50}$$

と変形できる。 α_ℓ ($\ell \in \{2, \dots, K\}$) による偏微分を行ない、

$$\frac{\partial}{\partial \alpha_\ell} \hat{\sigma}_{(MV)}^2 = 2\alpha_\ell \hat{\sigma}_\ell^2 + 2 \left(\sum_{k=2}^K \alpha_k - 1 \right) \hat{\sigma}_1^2 = 0 \tag{3.51}$$

を解くことで $\hat{\sigma}_{(MV)}^2$ の最小値を得ることが可能である．式 (3.51) は，

$$\alpha_\ell \frac{\hat{\sigma}_\ell^2}{\hat{\sigma}_1^2} = 1 - \sum_{k=2}^K \alpha_k \quad (3.52)$$

と変形でき，式 (3.49) を用いて，

$$\alpha_\ell = \alpha_1 \frac{\hat{\sigma}_1^2}{\hat{\sigma}_\ell^2} \quad (3.53)$$

が得られる．また，式 (3.53) を式 (3.45) に代入すると，

$$\alpha_1 + \sum_{k=2}^K \alpha_1 \frac{\hat{\sigma}_1^2}{\hat{\sigma}_k^2} = 1 \quad (3.54)$$

となり，

$$\alpha_1 = \frac{1/\hat{\sigma}_1^2}{\sum_{k=1}^K 1/\hat{\sigma}_k^2} \quad (3.55)$$

を得る． $\alpha_2, \dots, \alpha_K$ も同様の手順で解を求めることができ，

$$\alpha_k = \frac{1/\hat{\sigma}_k^2}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} \quad (3.56)$$

と表される．以上より，最終的な推定値は

$$\hat{u}_{(MV)} = \sum_{k=1}^K \frac{1/\hat{\sigma}_k^2}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} \hat{u}_k \quad (3.57)$$

と表される．また，このとき， $\hat{\sigma}_{(MV)}^2$ の値は

$$\begin{aligned} \hat{\sigma}_{(MV)}^2 &= \sum_{k=1}^K \alpha_k^2 \hat{\sigma}_k^2 \\ &= \sum_{k=1}^K \left(\frac{1/\hat{\sigma}_k^2}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} \right)^2 \hat{\sigma}_k^2 \\ &= \frac{1}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} \end{aligned} \quad (3.58)$$

となる．なお，MV 法による推定誤差の分散と OBS 法による推定誤差の分散に関して，

$$\hat{\sigma}_{(MV)}^2 = \frac{1}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} = \frac{1/\hat{\sigma}_{k_{min}}^2}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} \hat{\sigma}_{k_{min}}^2 < \hat{\sigma}_{k_{min}}^2 \quad (3.59)$$

が成立する． k_{min} は K 個の推定値候補のうち，誤差の分散が最小であるもののインデックスを表す．これは，各 e_k における無相関性の仮定が成立する状況においては，MV 法は OBS 法よりも最終的な推定値における誤差の分散が小さくなり，より精度の高い補間が可能であることを意味している．

3.7 計算効率の改善

3.7.1 MP 一般逆行列の計算回数の削減

ブロックベース線形回帰によるデモザイキングでは，RAW 画像から切り出された全ての画像ブロックに対し，式 (3.27) で B_j を求めるにあたり，MP 一般逆の計算を行なっている．一般に，ブロックベース線形回帰における画像補間を現実的な大きさの画像に対して適用する場合，この大量の MP 一般逆行列の計算がボトルネックとなり，処理を完了するまでに膨大な時間を要してしまう．しかし，デモザイキングにおいては欠損が規則的に生じている点を考慮することで計算回数を抑え，処理時間を短縮することが可能である．

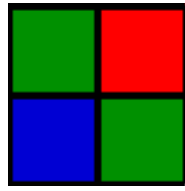
式 (3.27) に示したように， B_j は

$$B_j = N_j \hat{R} M_j' (M_j \hat{R} M_j')^+$$

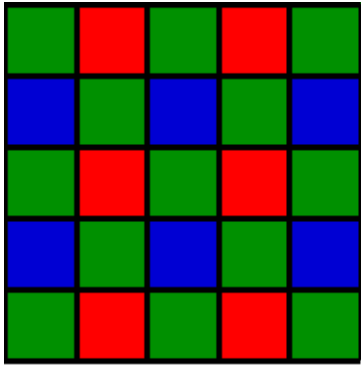
の形式で与えられる．ここで， \hat{R} は相関行列， M_j および N_j は S_j ，すなわち撮影に用いた CFA パターンから得られるマスク画像の情報から構成されるものであり， B_j は各画像ブロック \mathbf{y}_j の情報には依存していないことがわかる．ここで，各 S_j のうち相異なる

るものの個数は CFA パターンの繰り返しにおける最小のサイズ以下となる。例えば，Bayer CFA パターンの繰り返しの最小サイズは，図 3.3(a) に示すように $2 \times 2 = 4$ であり，Bayer CFA で撮影された RAW 画像から切り出された画像ブロックに対するマスク画像 S_j の数も 4 通りとなる。 $\alpha = \beta = 5$ で切り出した画像ブロックに対応するマスク画像を図 3.3(b) ~ 図 3.3(e) に示す。

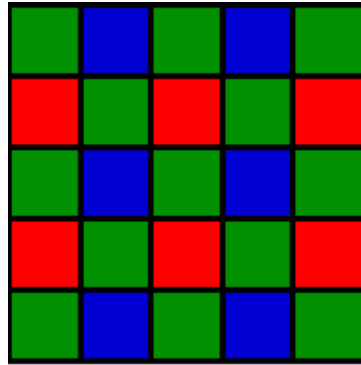
以上から，あらかじめ相異なる S_j に対して式 (3.27) を計算し，それぞれ得られた B_j の値を保持しておくことで， B_j を逐一計算することなく全画像ブロックに対する補間が可能である。



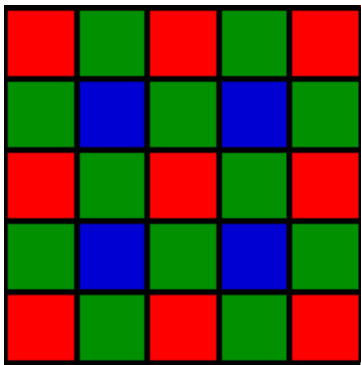
(a)



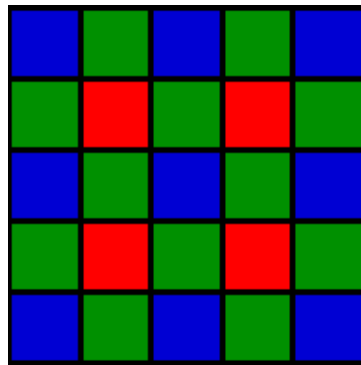
(b)



(c)



(d)



(e)

図 3.3: (a) Bayer CFA パターンにおける繰り返しの最小サイズ (2×2),
 (b) ~ (e) 5×5 画像ブロックを切り出した時の, Bayer CFA パターンに対応するマスク
 画像 S_j

3.7.2 複数の推定値候補に対する効率的な計算方法

前節では、重複を許して画像ブロックを切り出すことにより得られる、欠損要素に対する複数の推定値候補の扱い方について述べたが、最終的な推定結果を得るために、1つの欠損要素に対して、各々の画像ブロックから得られる推定値や誤差の分散を全て取得した後に処理を行なうのは非効率的である。ここでは、着目する欠損画素に対し、画像ブロックから新たな推定情報が得られるごとに値の更新を行ない、全ての推定情報を保持することなく最終的な推定結果を求めるための方法を述べる。

OBS 法 [4]

OBS 法では、はじめに得られた推定値とその誤差の分散の値を保持しておき、新たな推定値が得られた際、誤差の分散について保持している推定値との比較を行ない、より小さい値だったときにのみ保持する情報の更新を行なう。この操作の繰り返しにより、最終的に $\hat{u}_{(OBS)}$ を求めることができる。

MV 法 [5, 32]

着目欠損要素を含む画像ブロックが K 個存在するものとする。推定値 $\hat{u}_1, \dots, \hat{u}_k$ ($k \in \{1, \dots, K\}$) と、それに対応する誤差の分散 $\hat{\sigma}_1^2, \dots, \hat{\sigma}_k^2$ の値が判明しているとき、各推定値に対し最小分散凸結合を適用したものを、

$$\hat{u}_{(MV)_{1\dots k}} = \sum_{\ell=1}^k \frac{1/\hat{\sigma}_\ell^2}{\sum_{m=1}^k 1/\hat{\sigma}_m^2} \hat{u}_\ell \quad (3.60)$$

と定める.

$$\hat{u}_{(MV)_{1\dots 1}} = \hat{u}_1, \quad (3.61)$$

$$\hat{u}_{(MV)_{1\dots K}} = \sum_{\ell=1}^K \frac{1/\hat{\sigma}_\ell^2}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} \hat{u}_\ell = \hat{u}_{(MV)} \quad (3.62)$$

である. また, $\hat{u}_{(MV)_{1\dots k}}$ の誤差の分散は, 式 (3.58) より,

$$\hat{\sigma}_{(MV)_{1\dots k}}^2 = \frac{1}{\sum_{m=1}^k 1/\hat{\sigma}_m^2} \quad (3.63)$$

で与えられ,

$$\hat{\sigma}_{(MV)_{1\dots 1}}^2 = \hat{\sigma}_1^2, \quad (3.64)$$

$$\hat{\sigma}_{(MV)_{1\dots K}}^2 = \frac{1}{\sum_{m=1}^K 1/\hat{\sigma}_m^2} = \hat{\sigma}_{(MV)}^2 \quad (3.65)$$

である. ここで, $\hat{\sigma}_{(MV)_{1\dots k+1}}^2$ と $\hat{\sigma}_{(MV)_{1\dots k}}^2$ について,

$$\begin{aligned} \hat{\sigma}_{(MV)_{1\dots k+1}}^2 &= \frac{1}{\sum_{m=1}^{k+1} 1/\hat{\sigma}_m^2} \\ &= \frac{1}{\sum_{m=1}^k 1/\hat{\sigma}_m^2 + 1/\hat{\sigma}_{k+1}^2} \\ &= \frac{1}{1/\hat{\sigma}_{(MV)_{1\dots k}}^2 + 1/\hat{\sigma}_{k+1}^2} \end{aligned} \quad (3.66)$$

が成立し, $\hat{u}_{(MV)_{1\dots k+1}}$ と $\hat{u}_{(MV)_{1\dots k}}$ について,

$$\begin{aligned} \hat{u}_{(MV)_{1\dots k+1}} &= \sum_{\ell=1}^{k+1} \frac{1/\hat{\sigma}_\ell^2}{\sum_{m=1}^{k+1} 1/\hat{\sigma}_m^2} \hat{u}_\ell \\ &= \hat{\sigma}_{(MV)_{1\dots k+1}}^2 \left(\sum_{\ell=1}^k \frac{\hat{u}_\ell}{\hat{\sigma}_\ell^2} + \frac{\hat{u}_{k+1}}{\hat{\sigma}_{k+1}^2} \right) \\ &= \hat{\sigma}_{(MV)_{1\dots k+1}}^2 \left(\frac{1}{\hat{\sigma}_{(MV)_{1\dots k}}^2} \sum_{\ell=1}^k \frac{1/\hat{\sigma}_\ell^2}{\sum_{m=1}^k 1/\hat{\sigma}_m^2} \hat{u}_\ell + \frac{\hat{u}_{k+1}}{\hat{\sigma}_{k+1}^2} \right) \\ &= \hat{\sigma}_{(MV)_{1\dots k+1}}^2 \left(\frac{\hat{u}_{(MV)_{1\dots k}}}{\hat{\sigma}_{(MV)_{1\dots k}}^2} + \frac{\hat{u}_{k+1}}{\hat{\sigma}_{k+1}^2} \right) \end{aligned} \quad (3.67)$$

が成立する. 以上より, 新たな推定値が得られるごとに式 (3.66), 式 (3.67) を繰り返し適用することで, $\hat{u}_1, \dots, \hat{u}_k$ や $\hat{\sigma}_1, \dots, \hat{\sigma}_k$ の値をすべて保持せずとも, $\hat{u}_{(MV)}$ を求めることができる.

第 4 章

ブロックベース線形回帰によるデモザイキング手法に関する数値実験

本章では、第 3 章において提案したブロックベース線形回帰によるデモザイキング手法に関して、その有効性を検証すべく数値実験を行なう。まず、提案手法における、重複を許可した画像ブロックの切り出しにより得られる複数の推定値候補の扱いに関して、OBS 法、MV 法の性能を調査する。その後、提案手法と既存のデモザイキング手法について性能の比較を行なう。

4.1 実験用画像

Kodak dataset 24 枚のフルカラー画像 (512×768 画素 18 枚, 768×512 画素 6 枚)[33], および McMaster dataset 18 枚のフルカラー画像 (500×500 画素)[19] に対し, Bayer CFA [13] パターンに従い色情報を欠損させることにより, RAW 画像に相当する画像を作成し, 実験画像として使用する。Kodak dataset は画像処理における様々な分野の研究において, 実験用標準画像として広く利用されている。また, McMaster dataset は,

デモザイキング手法の研究に伴い McMaster 大学により作成されたものであり, Kodak dataset と比較して色成分間の相関が小さく, また急激に色が変わる領域を多く持つ画像から構成されている点が特徴である. 図 4.1 に Kodak dataset の画像を, 図 4.2 に McMaster dataset の画像を示す. なお, 本論文では, 比較実験に用いた手法の処理の都合上, 実験画像の外側 10 画素を除外した上でデモザイキングを適用していることを付記しておく.

4.2 提案手法に関する実験条件

RAW 画像から切り出すブロックサイズのパラメータを $\alpha = \beta = 8$ とし, 初回の相関行列 R の推定には文献 [15] の手法 (HILI) を用いる. また, 推定精度向上のための補間の繰り返し回数を 5 回に設定する. 以降, OBS 法, MV 法を用いた提案手法をそれぞれ BBLR(Block-Based Linear Regression)-OBS, BBLR-MV と呼ぶこととする.



(a) No.1 - No.6



(b) No.7 - No.12



(c) No.13 - No.18

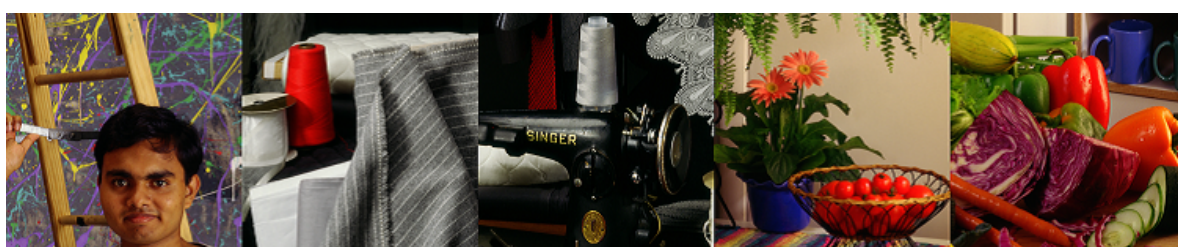


(d) No.19 - No.24

図 4.1: Kodak dataset の画像



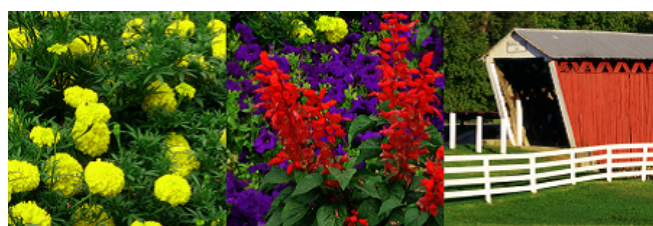
(a) No.1 - No.5



(b) No.6 - No.10



(c) No.11 - No.15



(d) No.16 - No.18

図 4.2: McMaster dataset の画像

4.3 OBS 法, MV 法による推定に関する実験

BBLR-OBS, BBLR-MV それぞれを用いたデモザイキングにより得られた画像について, 画像の客観的画質評価基準である CPSNR を比較することで, OBS 法と MV 法の性能を確認する. CPSNR は評価対象となる画像と, それに対応する真の画像に対する, 赤・緑・青色成分それぞれにおける平均二乗誤差 $MSE(R)$, $MSE(G)$, $MSE(B)$ を用いて,

$$CPSNR(\text{dB}) = 10 \log_{10} \frac{3 \cdot 255^2}{MSE(R) + MSE(G) + MSE(B)} \quad (4.1)$$

と定義される. 表 4.1 は Kodak dataset による結果であり, 24 枚全ての画像に対して BBLR-MV が BBLR-OBS より高い CPSNR を示している. また, 表 4.2 は McMaster dataset による結果であり, こちらも 18 枚全ての画像に対して BBLR-MV が BBLR-OBS を上回る CPSNR を示していることを確認できる. 以上の結果より, MV 法による推定値の凸結合が有効に機能しているものと考えられ, 提案手法において各ブロックから得られる複数の推定値に対して MV 法を採用すべきであると結論付ける.

表 4.1: BBLR-OBS と BBLR-MV によるデモザイキング画像の CPSNR (dB) の比較 (Kodak dataset)

No.	BBLR-OBS	BBLR-MV
1	37.468	38.207
2	40.231	41.117
3	42.452	42.857
4	41.518	41.910
5	37.170	37.781
6	39.155	40.714
7	42.095	42.575
8	33.559	35.193
9	40.657	41.464
10	41.681	42.103
11	38.322	39.607
12	42.165	42.904
13	33.992	36.596
14	36.512	36.988
15	39.865	40.384
16	42.884	44.017
17	40.476	41.660
18	36.769	37.625
19	38.231	39.840
20	39.979	40.722
21	38.166	39.521
22	37.311	37.956
23	42.886	43.389
24	34.941	35.492
average	39.104	40.026

表 4.2: BBLR-OBS と BBLR-MV によるデモザイキング画像の CPSNR (dB) の比較 (McMaster dataset)

No.	BBLR-OBS	BBLR-MV
1	28.382	28.832
2	33.908	34.230
3	32.450	32.697
4	36.690	37.301
5	33.135	33.939
6	38.261	39.058
7	37.892	38.861
8	37.366	37.479
9	36.635	36.932
10	38.109	38.504
11	38.939	39.304
12	37.496	38.047
13	39.863	40.242
14	37.840	38.238
15	38.171	38.516
16	33.207	33.962
17	33.268	34.235
18	34.280	34.757
average	35.883	36.396

4.4 既存手法との性能比較実験

提案手法の BBLR-MV との比較対象として、既存のデモザイキング手法のうち、HILI, および文献 [19](LDI-NLM, LDI-NAT), [21](EICC), [22](RI), [24](ARI), [25](WMLRI) の手法を採用する。なお、これらの比較手法について、HILI を除き、著者が公開している MATLAB 用のプログラムコードを利用している。

表 4.3, 表 4.4 はそれぞれのデータセットから生成した RAW 画像から、各手法によって得られたカラー画像の CPSNR を示している。表 4.3 は Kodak dataset による結果であり、BBLR-MV が 24 枚中 14 枚の画像において最も高い CPSNR を示し、24 枚の画像の CPSNR 平均値に関しても、最も優れた結果を残していることが確認できる。一方、表 4.4 は McMaster dataset による結果であるが、これらの画像に対しては良好な結果は得られず、特に ARI をはじめとする残差補間を利用した手法に対して劣る結果となった。McMaster dataset の画像は Kodak dataset の画像と比較して、画像中の領域間における色の変化が顕著である。そのため、McMaster dataset の画像に対するデモザイキングにおいて、切り出したブロック群が同一の確率変数ベクトルから生成された実現値であるという、提案手法における仮定が成立しておらず、正確な推定が実現しなかったものと考えられる。

次に、各手法により生成されたカラー画像の例を挙げ補間効果の検証を行なう。Kodak dataset No.13 の画像、McMaster dataset No.1 の画像から作成された RAW 画像に対し、デモザイキングを適用して得られたカラー画像の一部 (64×64 画素) を拡大したもの

を、それぞれ図 4.3, 図 4.4 に示す。図 4.3 において、画像上部の雪山の領域に着目すると、多くの既存手法では偽色が目立っているが、BBLR-MV では効果的に偽色の発生を抑制できていることが確認できる。一方、図 4.4 では、BBLR-MV は補間における誤差により、BBLR-MV よりも高い CPSNR を示した手法と比較して、画像中のエッジ領域にジッターノイズが若干現れている事が確認でき、やや画質が劣っている事が分かる。以上より、Kodak dataset のような画像内で色の変化が比較的少ない画像に対して、提案手法は有効であると結論付ける。

表 4.3: Kodak dataset から得られたデモザイキング画像の CPSNR (dB)

No.	BBLR-MV	HILI	LDI-NLM	LDI-NAT	EICC	RI	ARI	WMLRI
1	38.207	32.062	35.279	34.448	35.575	35.553	38.809	36.319
2	41.117	37.958	38.521	39.780	40.331	39.457	39.666	40.471
3	42.857	39.823	41.488	41.936	42.012	41.219	42.673	42.566
4	41.910	38.830	38.790	39.653	40.210	40.136	40.690	40.807
5	37.781	33.511	36.844	35.812	36.726	36.661	38.263	37.291
6	40.714	33.380	35.906	35.794	38.487	38.438	40.509	38.736
7	42.575	39.553	41.919	41.493	42.423	41.995	42.666	42.570
8	35.193	29.140	33.350	32.596	34.035	34.009	34.975	34.486
9	41.464	38.185	40.470	40.522	41.405	41.106	41.557	41.810
10	42.103	38.737	40.473	40.828	41.533	41.497	41.550	41.919
11	39.607	34.763	37.001	36.934	38.273	38.181	39.623	38.839
12	42.904	38.759	39.948	41.007	42.520	42.309	43.178	42.930
13	36.596	29.616	31.712	30.726	31.961	32.001	35.201	32.522
14	36.988	34.601	36.259	36.311	36.662	36.308	37.542	37.434
15	40.384	38.056	38.288	38.483	39.056	38.770	38.612	39.142
16	44.017	36.512	38.468	39.024	42.302	42.207	43.242	42.448
17	41.660	37.806	39.161	39.050	40.060	40.072	41.131	40.659
18	37.625	33.967	35.343	34.825	35.380	35.400	36.863	35.888
19	39.840	33.728	37.790	37.504	39.126	39.122	40.459	39.421
20	40.722	37.342	39.443	39.408	39.982	39.933	40.852	40.281
21	39.521	34.146	36.762	35.910	37.276	37.234	39.237	37.698
22	37.956	35.425	37.125	37.272	37.753	37.667	38.048	38.229
23	43.389	41.004	41.874	42.271	43.127	42.223	43.157	43.081
24	35.492	32.253	33.605	33.357	34.175	34.326	35.361	34.430
average	40.026	35.798	37.742	37.706	38.766	38.576	39.744	39.166

表 4.4: McMaster dataset から得られたデモザイキング画像の CPSNR (dB)

No.	BBLR-MV	HILI	LDI-NLM	LDI-NAT	EICC	RI	ARI	WMLRI
1	28.832	27.413	28.897	29.180	29.296	29.119	29.629	29.405
2	34.230	33.204	34.855	35.032	35.172	35.003	35.212	35.334
3	32.697	31.251	33.128	32.593	33.445	33.750	34.634	34.032
4	37.301	33.901	36.473	35.929	36.433	37.861	37.891	37.967
5	33.939	32.007	33.754	34.083	34.413	33.915	35.418	34.411
6	39.058	35.209	37.102	37.841	38.467	38.289	39.752	38.799
7	38.861	35.039	36.247	35.948	36.982	36.946	39.704	37.012
8	37.479	36.157	37.827	37.490	38.660	36.963	39.506	37.275
9	36.932	35.282	37.038	37.000	37.155	35.917	37.882	36.819
10	38.504	36.923	38.371	38.771	38.547	38.148	39.242	39.079
11	39.304	37.873	39.252	39.579	39.910	39.427	40.199	40.170
12	38.047	36.199	38.587	38.883	39.762	39.606	40.024	39.801
13	40.242	38.721	40.799	40.733	41.099	40.268	40.592	40.614
14	38.238	37.134	38.479	38.696	38.973	38.921	39.023	39.074
15	38.516	37.607	38.911	38.954	39.308	38.367	39.380	39.221
16	33.962	31.545	33.317	33.855	35.139	35.165	35.642	35.419
17	34.235	31.286	32.570	32.874	33.368	32.440	34.687	33.176
18	34.757	33.183	35.381	35.123	35.986	36.518	36.431	36.410
average	36.396	34.441	36.166	36.254	36.784	36.479	37.492	36.890

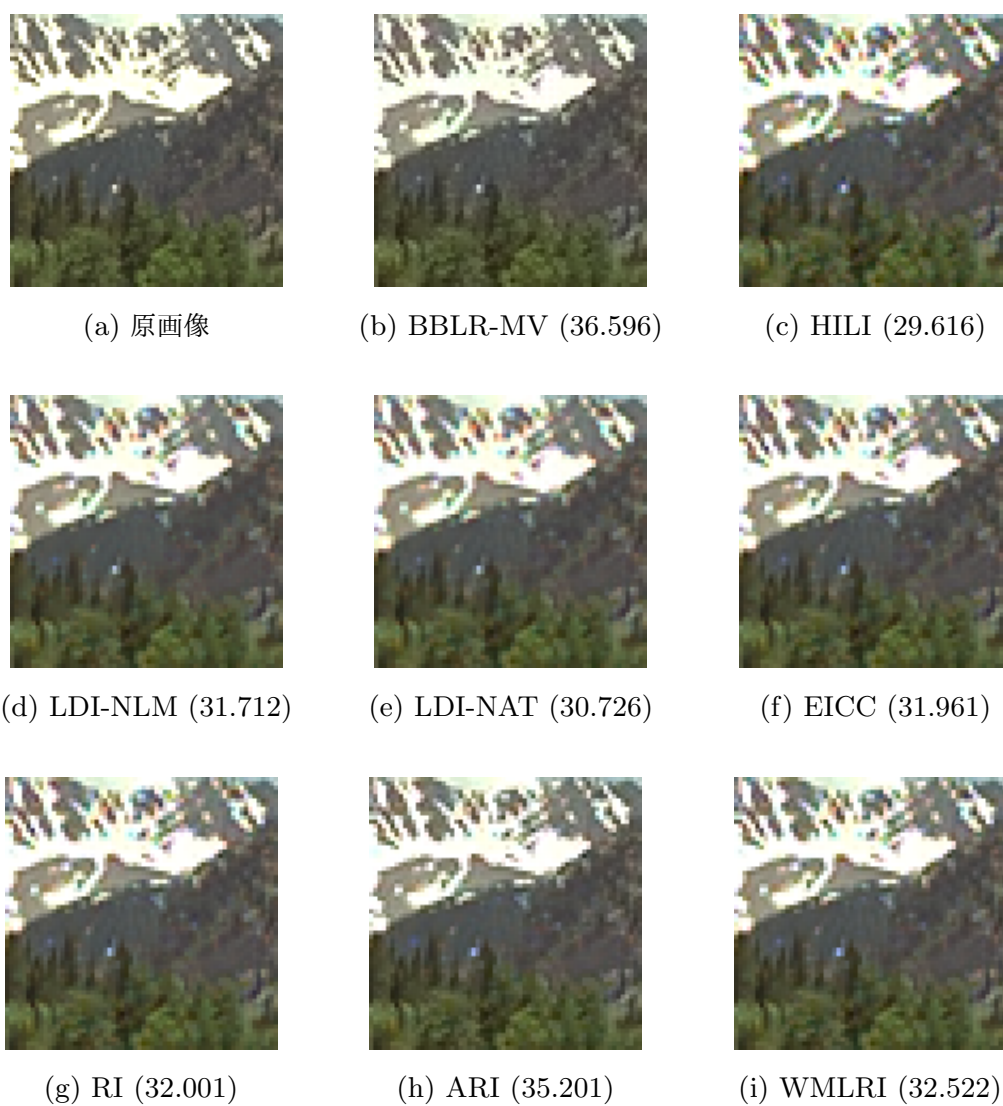


図 4.3: Kodak dataset No.13 に対するデモザイキング結果の拡大画像と CPSNR(dB)

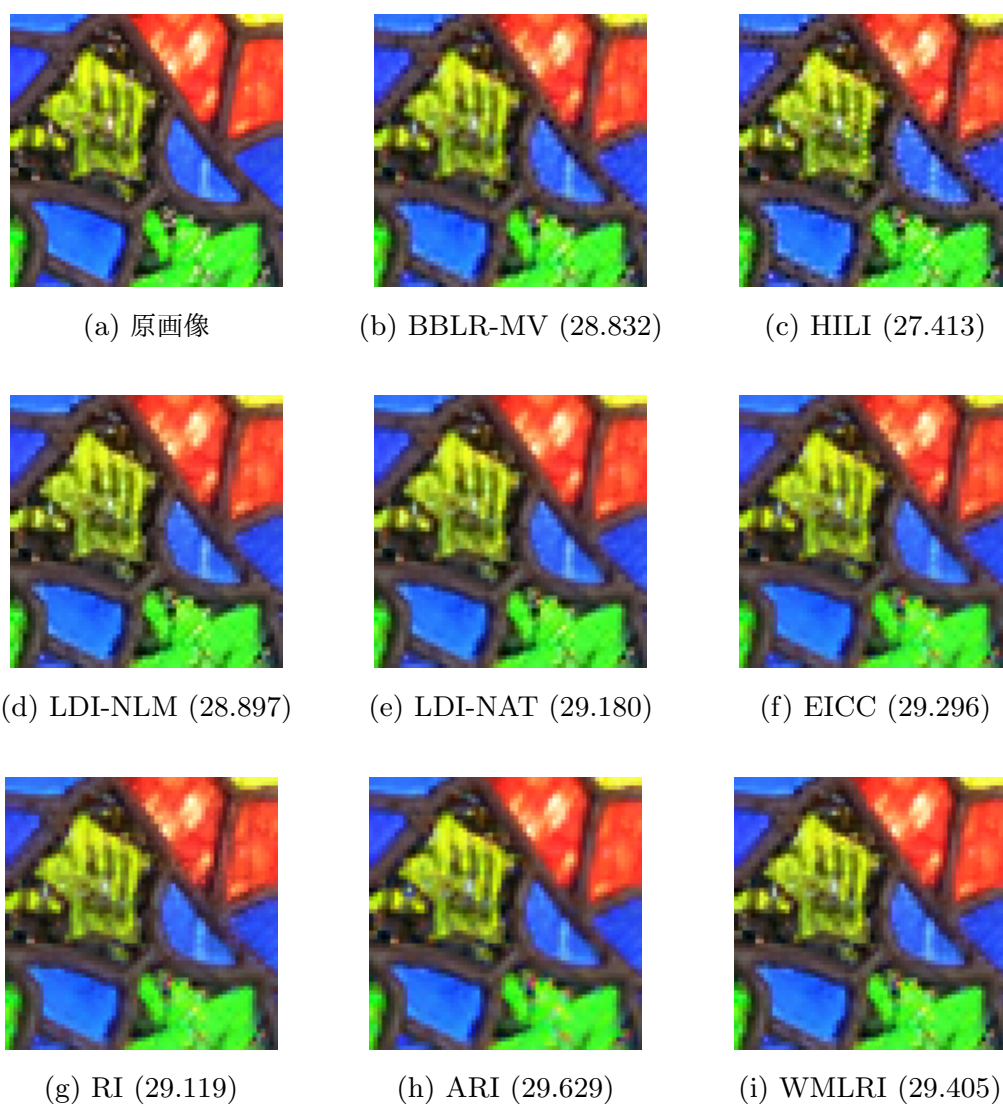


図 4.4: McMaster dataset No.1 に対するデモザイキング結果の拡大画像と CPSNR(dB)

第 5 章

JPEG 再適用法の改良法による画質改善手法

本章では、JPEG 圧縮アルゴリズムと既存の JPEG 画像の画質改善手法の例を紹介したのち、JPEG 再適用法に対して、画像データベースを導入した手法を提案する。

5.1 JPEG 圧縮アルゴリズム

JPEG では、以下の手順により画像データの圧縮を行っている。

1. 対象画像の RGB 色空間から $YCbCr$ 色空間への変換を行なう
2. 色差 (C_b, C_r) 成分のダウンサンプリングを行なう
3. 輝度，色差成分に対し， 8×8 ブロック分割し，2次元 DCT 変換を行なう
4. DCT 係数の量子化を行なう
5. エントロピー符号化を行なう

上記の手順において、2,4,5 の操作により画像データの圧縮が実現されている。なお、圧縮対象がグレースケール画像の場合、上記手順 1,2 は省略される。

5.1.1 RGB 色空間から YC_bC_r 色空間への変換

RGB 色空間で表現されている画像を、輝度と色差からなる YC_bC_r 色空間での表現に変換する。変換式は以下の形式で表される [34]。

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}. \quad (5.1)$$

ここで、 R, G, B, Y, C_b, C_r 各成分は 0 から 255 までの値を取りうる。なお、 YC_bC_r 色空間から RGB 色空間への変換式は、

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} Y \\ C_b - 128 \\ C_r - 128 \end{bmatrix} \quad (5.2)$$

と表される。

5.1.2 色差成分のダウンサンプリング

人間は輝度成分の変化には敏感であるが、色差成分に関しては比較的感度が低いという視覚特性がある。この特性に基づき、 C_b, C_r 成分に対してダウンサンプリングを行なう。水平、垂直方向に対してサンプリング間隔の設定ができ、ダウンサンプリングを行わないことも可能である。 C_b, C_r を水平、垂直方向にそれぞれ 1/2 ずつ間引く方式が一般的である。

5.1.3 ブロック DCT 変換

輝度成分, およびダウンサンプリングされた色差成分に対して, それぞれ 8×8 サイズのブロックに分割し, 2次元 DCT 変換

$$B_{p,q} = c_p c_q \sum_{m=1}^8 \sum_{n=1}^8 A_{m,n} \cos \frac{(2m-1)(p-1)\pi}{16} \cos \frac{(2n-1)(q-1)\pi}{16} \quad (5.3)$$

$$(1 \leq p \leq 8, 1 \leq q \leq 8),$$

$$c_p = \begin{cases} 1/2\sqrt{2} & (p=1) \\ 1/2 & (\text{otherwise}) \end{cases}, \quad c_q = \begin{cases} 1/2\sqrt{2} & (q=1) \\ 1/2 & (\text{otherwise}) \end{cases}$$

を施すことにより, 周波数領域への変換 (ブロック DCT 変換) を行なう. ここで, $A_{m,n}$ は分割されたブロックの画素値からなる行列の (m,n) 要素であり, $B_{p,q}$ は変換後の行列の (p,q) 要素である. 一般に, 自然画像は低周波成分が支配的であるため, その 2次元 DCT 変換においては, 低周波成分の係数に画像のエネルギーが集中する. よって, 後に述べる量子化によるデータの削減を効果的に行なうことが可能となる. なお, 2次元逆 DCT 変換は,

$$A_{m,n} = \sum_{p=1}^8 \sum_{q=1}^8 c_p c_q B_{p,q} \cos \frac{(2m-1)(p-1)\pi}{16} \cos \frac{(2n-1)(q-1)\pi}{16} \quad (5.4)$$

$$(1 \leq m \leq 8, 1 \leq n \leq 8),$$

$$c_p = \begin{cases} 1/2\sqrt{2} & (p=1) \\ 1/2 & (\text{otherwise}) \end{cases}, \quad c_q = \begin{cases} 1/2\sqrt{2} & (q=1) \\ 1/2 & (\text{otherwise}) \end{cases}$$

と表される. 周波数領域に変換されたそれぞれのブロックに対して, 式 (5.4) により空間領域に戻す操作をブロック IDCT 変換と呼ぶことにする.

5.1.4 周波数領域における量子化

2次元 DCT 変換を施されたブロックに対し、各要素の値の精度を落とす処理を行なう。

量子化操作は

$$C_{p,q} = \text{round} \left(\frac{B_{p,q}}{Q_{p,q}} \right) \quad (5.5)$$

の形式で表され、 $Q_{p,q}$ は $B_{p,q}$ に対する量子化精度を定める定数であり、 $Q_{p,q}$ の値が大きくなるほど、粗く量子化されることになる。round(\cdot) は小数点以下の値に対する四捨五入操作を表す。 $Q_{p,q}$ の値を並べた行列は量子化テーブルと呼ばれている。量子化テーブルは任意に設定することが可能であり、通常、情報が集中する低周波成分に対しては小さな値が設定され、情報量が少なく、また人間には知覚しにくい高周波成分には大きな値が設定される。これにより、画質に与える影響を抑えながら圧縮効率を上げることが可能となる。なお、逆量子化は $C_{p,q}$ に対し $Q_{p,q}$ を乗じる事により行われる。JPEG 標準 [2] によるガイドラインでは、表 5.1 に示すテーブルが提案されている。

5.1.5 エントロピー符号化

量子化が行われたブロック群に対し、直流成分 (ブロック中の (1, 1) 要素) と交流成分 ((1, 1) 以外の要素) に分けてエントロピー符号化を行なう。直流成分にはハフマン符号化が用いられ、交流成分にはハフマン符号化に加えランレングス符号化も用いられる。

直流成分に関しては、対象の 1 つ前のブロックの直流成分との差分をハフマン符号化する。交流成分に関しては、はじめにジグザグスキャンと呼ばれる、図 5.1 に示す順番でブロック内のデータを 1 次元に並べる操作を行なう。前述の量子化操作により、ブロックの

表 5.1: 標準量子化テーブル

(a) 輝度成分用量子化テーブル

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(b) 色差成分用量子化テーブル

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

高周波成分に対応する要素は 0 となっている場合が多く、ジグザグスキャンにより、データ列には 0 が連続して現れることが期待できる。この手順により連続して現れた 0 に対してランレングス符号化を施した上で、ハフマン符号化を適用する。

5.1.6 JPEG の復号とノイズの混入

JPEG 画像の復号は、上記の圧縮手順の逆を辿ることで実行される。

1. エントロピー符号化に対する復号を行ない、得られたデータを 8×8 に並べる

*	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

図 5.1: ジグザグスキャン

2. 並べられたデータに対し逆量子化を行なう
3. 8×8 ブロック IDCT 変換を行なう
4. 色差成分のアップサンプリングを行なう
5. $YCbCr$ 色空間から RGB 色空間への変換を行なう

ここで、データの削減が行われる操作であった前述の圧縮手順 2,4,5 に対応する、復号手順 1,2,4 について、1. における復号は完全に圧縮前の情報を再現可能である一方で、2. では圧縮時の四捨五入操作により、量子化前の値との間に誤差が生じ、4. でのアップサンプリング操作においても、元の色差成分との間に誤差が生じてしまう。これらの誤差が復号画像内にノイズとして現れる。図 2.4 で示したブロックノイズやモスキートノイズは、いずれも圧縮手順 4. に起因するものであり、前者は主に、隣接するブロック同士における低周波成分の差が大きくなり、輝度・色差成分の連続性が失われる事により発生し、後者はブロック内の高周波成分が除去される事により発生する。

5.2 既存の画質改善手法

現在までに数多くの JPEG 画像の画質改善手法が提案されており，例えば画素の特性に応じたフィルタ処理を用いる手法 [35, 36, 37, 38]，原画像に関する事前情報に確率モデルを利用する手法 [39, 40]，ウェーブレット変換を利用する手法 [41, 42]，TV(Total-Variation) 正則化法 [43] を利用する手法 [44, 45, 46]，冗長 DCT 変換とソフトシュリンケージを利用する手法 [47]，JPEG の再適用による手法 [3] など，アプローチも様々である．ここでは，TV 正則化法を利用した文献 [44] の手法と，文献 [47] の手法，そして本論文で提案する手法の元となった文献 [3] の手法について述べる．

なお，以降の説明において，改善を行なう JPEG 画像 (入力画像) はグレースケール画像を対象としている．カラー画像に対しては，入力画像の復号時に得られる輝度成分・色差成分それぞれに対し，グレースケール画像とみなした上で画質改善手法を適用する必要がある．

5.2.1 TV 正則化法を利用した画質改善手法 [44]

TV 正則化法を利用した画質改善手法のうち，代表的なものに Alter らにより提案された手法がある． y を入力画像， T, T^{-1} をそれぞれブロック DCT 変換，ブロック IDCT 変換， $J(Q)$ を入力画像から得られる量子化テーブル Q による JPEG 圧縮・復号の操作とする．このとき， y に対応する真の画像 x に対し，

$$y = J(Q)x \tag{5.6}$$

が成立する．以上の準備のもとで，Alter らは JPEG 画像の画質改善問題に対し，次の制約付き最適化問題とみなして定式化を行なっている．

$$\inf_u \left\{ \int_{\Omega} |\nabla u| d\Omega \right\} \text{ subject to } u \in C, \quad (5.7)$$

$$C = \{u \mid T\{J(Q)u\} = T\{y\}\}. \quad (5.8)$$

C による制約は，画質改善画像にブロック DCT 変換を適用して得られる DCT 係数が，入力画像の DCT 係数が存在していた量子化区間内に属することを意味する．さらに，式 (5.7) の TV 正則化項における各勾配に対し重みを持たせることで，ブロックノイズのさらなる抑制も行っている．式 (5.7) の解は，以下に示す最急降下法により求められる．

$$u^{(0)} = y, \quad (5.9)$$

$$u^{(\ell+1)} = P_C\{u^{(\ell)} - t^{(\ell)}g(u^{(\ell)})\}, \quad (5.10)$$

$$\hat{x} = u^{(\ell_{max})}. \quad (5.11)$$

ここで， ℓ は反復計算のカウンタ数， ℓ_{max} は画質改善を行なうにあたって設定する反復回数を表す． $t^{(\ell)}$ は最急降下法のステップ幅 $1/(\ell + 1)$ ， $g(u^{(\ell)})$ は $\int_{\Omega} |\nabla u^{(\ell)}| d\Omega$ の劣勾配を表している．また， $P_C(\cdot)$ は C への凸射影を表す．具体的には，ブロック DCT 変換を行なった際の DCT 係数が，入力画像の DCT 係数が存在していた量子化区間から外れていた場合，区間の上限，もしくは下限の値に修正する操作である．

5.2.2 冗長 DCT 変換とソフトシュリンケージを利用した画質改善手法 [47]

所与の画像から，様々な方向へのシフト操作により得られる複数枚の画像に対し DCT 変換を行なう操作は冗長 DCT 変換と呼ばれている．小松らは，画質改善画像の DCT 係

数に対して Alter らの手法と同様の制約を課した上で、TV 正則化法の代わりに、入力画像に対して冗長ブロック DCT 変換を施し、ソフトシュリンケージを用いて DCT 係数を縮退させる操作を繰り返すことにより画質改善を行なう手法を提案している。

所与の画像に対し (i, j) ($i, j \in \{-3, -2, \dots, 4\}$) だけシフトを行なう操作を S_k ($k \in \{1, \dots, 64\}$) とする。 k は各シフト方向に対応するインデックスを表す。 また、 S_k に対する逆操作、すなわち $(-i, -j)$ だけシフトを行なう操作を S_k^{-1} とし、反復更新式

$$u^{(0)} = y, \quad (5.12)$$

$$u^{(\ell+1)} = P_C \left\{ \frac{1}{64} \sum_{k=1}^{64} S_k^{-1} T^{-1} H_{\tau^{(\ell)}}(T S_k u^{(\ell)}) \right\}, \quad (5.13)$$

$$\hat{x} = u^{(\ell_{max})} \quad (5.14)$$

の計算により画質の改善を行なう。ここで、 $H_{\tau^{(\ell)}}(z)$ は DCT 係数の交流成分 z に対するソフトシュリンケージ操作

$$H_{\tau^{(\ell)}}(z) = \begin{cases} \text{sgn}(z) \cdot (|z| - \tau^{(\ell)}) & (|z| > \tau^{(\ell)}) \\ 0 & (\text{otherwise}) \end{cases}, \quad (5.15)$$

$$\text{sgn}(z) = \begin{cases} 1 & (z > 0) \\ -1 & (\text{otherwise}) \end{cases} \quad (5.16)$$

であり、パラメータ $\tau^{(\ell)}$ は、

$$\tau^{(\ell+1)} = \tau^{(0)} \left(1 - \frac{\ell}{\ell_{max}} \right) \quad (5.17)$$

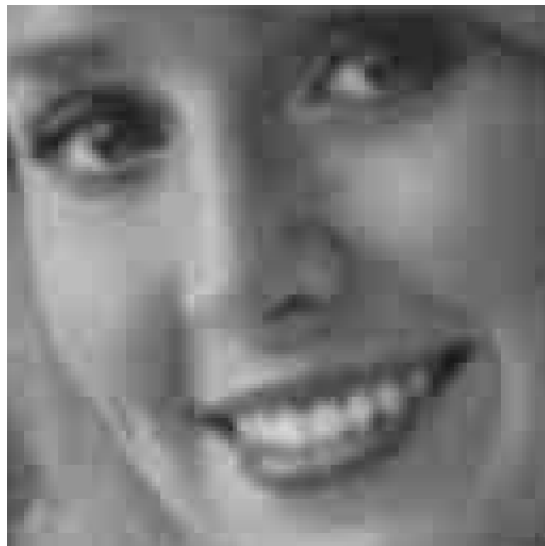
により求められる。初期値 $\tau^{(0)}$ は手動で設定する。

5.2.3 JPEG 再適用法による画質改善手法 [3]

Nosratinia によって提案された JPEG 再適用法は、入力画像に対するシフト操作から得られる複数の画像に対し、JPEG の再適用、すなわち JPEG による圧縮・復号操作を



(a) 入力画像



(b) (4,4) シフトと JPEG 再適用により得られた画像

図 5.2: シフト操作と JPEG 再適用を適用した画像の例

行ない、それらを逆シフトすることにより入力画像と位置合わせを行なった上で、加重平均を適用することで画質を改善する手法である。JPEG の再適用と位置合わせが行われた画像の例を図 5.2 に示す。このように、シフト操作と JPEG の再適用により、入力画像に混入したノイズが除去されると同時に、新たなノイズが入力画像と異なる位置に発生す

る。こうして得られる複数枚の画像の加重平均を計算することで、入力画像中に存在していたブロックノイズやモスキートノイズを拡散させることができ、結果として画質の改善が期待できる。

JPEG 再適用法における画質改善は、

$$\hat{x} = \sum_{k=1}^{64} \alpha_k S_k^{-1} J(Q) S_k y \quad (5.18)$$

と定式化される。 α_k は加重平均における重み係数を表す。なお、式 (5.18) は、 $\hat{\mathbf{x}} = \text{vec}(\hat{x})$ として、

$$\hat{\mathbf{x}} = X(Q, y) \boldsymbol{\alpha}, \quad (5.19)$$

$$X(Q, y) = [\text{vec}(S_1^{-1} J(Q) S_1 y), \dots, \text{vec}(D_{64}^{-1} J(Q) D_{64} y)], \quad (5.20)$$

$$\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_{64}]'$$

とも記述できる。

重み係数の設定について、文献 [3] では 2 種類の方法が提案されている。1 つは各画像に対し重みをつけることはせず、単純平均

$$\alpha_1 = \dots = \alpha_{64} = \frac{1}{64} \quad (5.21)$$

を利用する方法である。もう 1 つは最小二乗誤差 (MMSE) 推定による方法であり、あらかじめ何らかの画像集合 $\{x_1, \dots, x_n\}$ と、当該集合の要素に対し $J(Q)$ を適用することにより得られる JPEG 復号画像集合 $\{y_1, \dots, y_n\}$ を用いて、

$$\boldsymbol{\alpha}^{(MMSE)} = \underset{\boldsymbol{\alpha}}{\text{argmin}} \sum_{m=1}^n \|X(Q, y_m) \boldsymbol{\alpha} - \mathbf{x}_m\|_2^2 \quad (5.22)$$

$$= \frac{\sum_{m=1}^n X(Q, y_m)' \mathbf{x}_m}{\sum_{m=1}^n X(Q, y_m)' X(Q, y_m)} \quad (5.23)$$

により得られる係数ベクトルを利用する方法である。文献 [3] では、単純平均を利用し画質改善を行なうことを推奨している。

5.3 画像データベースを利用した JPEG 再適用法

JPEG 再適用法では反復計算が不要で、必要なパラメータが加重平均の重み係数ベクトルの値のみであり、極めて単純な方法でありながら、効果的な画質改善を実現する。しかしながら、どのような画像に対しても同じ重み係数ベクトルを用いているため、画像によっては改善が十分でない可能性がある。

この問題に対して、入力画像から算出される画像特徴量と画像データベースを導入することにより、入力画像に対する最適な重み係数ベクトル (最適係数ベクトル) を推定し、それを利用することでさらなる画質改善を行なう方法を構築する。ここでは、画像特徴量が類似している画像同士では最適係数ベクトルも類似しているという仮定のもと、入力画像と最も近い特徴量を持つデータベース画像に対する最適係数ベクトルを利用して JPEG 再適用法を行なう手法 [7, 8]、および画像特徴量から入力画像に対する最適係数ベクトルを与える関数の存在を仮定した上で、データベース画像から画像特徴量と最適係数ベクトルからなる訓練データセットを構成し、カーネル線形回帰を実行することにより当該関数の推定を行ない、得られた関数から算出される重み係数ベクトルを利用して JPEG 再適用法を行なう手法 [9] を提案する。以降、前者による手法を DB 法、後者による手法を KR 法と呼ぶこととする。

5.3.1 最適係数ベクトル

まず, JPEG 再適用法 [3] における, 入力画像に対する最適係数ベクトルを, 原画像と画質改善画像の二乗誤差を最小化するものと定める. これは $\mathbf{x} = \text{vec}(x)$ と y を用いて,

$$\boldsymbol{\alpha}^{(OPT)} = \underset{\boldsymbol{\alpha}}{\text{argmin}} \|X(Q, y)\boldsymbol{\alpha} - \mathbf{x}\|_2^2 \quad (5.24)$$

$$= X(Q, y)^+ \mathbf{x} \quad (5.25)$$

と表される. なお, 入力画像 y に対する原画像 x は当然のことながら未知である. そのため, 実際に $\boldsymbol{\alpha}^{(OPT)}$ そのものを利用して画質改善を行なうことは不可能であることを付記しておく.

5.3.2 画像特徴量

JPEG 画像から計算される画像特徴量について, 以下の 2 種を提案する.

画像特徴量 Σ [7]

Σ は, 所与の入力画像 y を 8×8 の排反なブロックに分割し, 個々のブロックを 64 次列ベクトル $\mathbf{b}_1, \dots, \mathbf{b}_n$ と見なした上で算出された 64 次の共分散行列であり, 以下のよう

に表される.

$$\Sigma = \frac{1}{n} \sum_{i=1}^n (\mathbf{b}_i - \boldsymbol{\mu})(\mathbf{b}_i - \boldsymbol{\mu})', \quad (5.26)$$

$$\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^n \mathbf{b}_i. \quad (5.27)$$

n は分割したブロックの個数を表す.

画像特徴量 $X'X$ [8]

$X'X$ は式 (5.20) の行列 $X(Q, y)$ から計算され,

$$X(Q, y)'X(Q, y) \quad (5.28)$$

という形式で与えられる.

以降, 画像特徴量 $\Sigma, X'X$ をまとめて, \mathbf{Y} と呼ぶこととする. なお, 提案手法の過程において, これらの画像特徴量は $\text{tr}[\mathbf{Y}]$ で割ることにより正規化を行った上で利用する.

5.3.3 DB 法 [7, 8]

画像データベース, および当該データベースに所属する画像を

$$\text{DB}_x = \{x_1^{(DB)}, \dots, x_n^{(DB)}\}$$

と表す. DB 法では, 以下の手順により入力画像の画質改善を実現する.

1. DB_x に属するそれぞれの画像に対し, $J(Q)$ を適用して得られる JPEG 画像データベース $\text{DB}_y = \{y_1^{(DB)}, \dots, y_n^{(DB)}\}$ を作成する. 量子化テーブル Q は入力画像 y の圧縮に用いられたものと同じのものを利用する.
2. 入力画像 y と DB_y に属するすべての画像について, それぞれ画像特徴量 $\mathbf{Y}^{(INPUT)}, \mathbf{Y}_i^{(DB)}$ ($i \in \{1, \dots, n\}$) を算出する.
3. y と各画像特徴量が最も近いデータベース画像のインデックス

$$i_{opt} = \underset{i}{\operatorname{argmin}} \|\mathbf{Y}^{(INPUT)} - \mathbf{Y}_i^{(DB)}\|_2^2 \quad (5.29)$$

を求める.

4. データベース画像 $y_{i_{opt}}^{(DB)}$ と $x_{i_{opt}}^{(DB)}$ から, 式 (5.25) を用いて, $y_{i_{opt}}^{(DB)}$ に対する最適係数ベクトル

$$\alpha_{i_{opt}}^{(DB)} = X(Q, y_{i_{opt}}^{(DB)}) + \mathbf{x}_{i_{opt}}^{(DB)} \quad (5.30)$$

を算出する.

5. 入力画像 y に対し, $\alpha_{i_{opt}}^{(DB)}$ を利用した JPEG 再適用法により, 画質改善画像を得る.

5.3.4 KR 法 [9]

KR 法では, 以下の手順により入力画像の画質改善を実行する.

1. DB 法と同様にして, DB_x に属するそれぞれの画像に対し, $J(Q)$ を適用して得られる JPEG 画像データベース DB_y を作成する.
2. DB_y に属するすべての画像に対し, 画像特徴量 $\mathbf{Y}_i^{(DB)}$, および最適係数ベクトル $\alpha_i^{(DB)} = X(Q, y_i^{(DB)}) + \mathbf{x}_i^{(DB)}$ を算出し, 後述するカーネル線形回帰に利用する訓練データセット $\{(\mathbf{Y}_i^{(DB)}, \alpha_i^{(DB)})\}$ を作成する. 以降, 便宜上このデータセットを $\{(\mathbf{Y}^{(i)}, \alpha_k^{(i)}) \mid (k \in \{1, \dots, 64\}, i \in \{1, \dots, n\})\}$ と表記する.
3. 得られたデータセットを利用したカーネル線形回帰により, 画像特徴量から最適係数ベクトル与える関数を推定する.
4. 入力画像 y から算出される画像特徴量 $\mathbf{Y}^{(INPUT)}$ を当該関数に入力し, 得られた重み係数ベクトルを利用した JPEG 再適用法により, 画質改善画像を得る.

カーネル線形回帰による関数の推定

画像特徴量から最適係数ベクトルを与える関数を $\mathbf{f} = [f_1(\cdot), \dots, f_{64}(\cdot)]'$ とする. \mathbf{f} と KR 法の手順 2. から得られる訓練データセット $\{(\mathbf{Y}^{(i)}, \alpha_k^{(i)})\}$ の間に,

$$\alpha_k^{(i)} = f_k(\mathbf{Y}^{(i)}) \quad (5.31)$$

が成立するものと仮定する.

次に, $f_k(\cdot)$ を

$$\hat{f}_k(\cdot) = \sum_{i=1}^n \beta_k^{(i)} K(\cdot, \mathbf{Y}^{(i)}) \quad (5.32)$$

とモデル化する. $K(\cdot, \mathbf{Y}^{(i)})$ はカーネル関数 [10] を表す. このモデルにおける回帰係数 $\beta_k^{(i)}$ は目的関数

$$J(\beta_k^{(1)}, \dots, \beta_k^{(n)}) = \sum_{j=1}^n \left(\alpha_k^{(j)} - \hat{f}_k(\mathbf{Y}^{(j)}) \right)^2 \quad (5.33)$$

が最小となるように決定する. $\beta_k = [\beta_k^{(1)}, \dots, \beta_k^{(n)}]'$ とすると, 最適な係数 β_k は

$$\begin{aligned} \beta_k^{(OPT)} &= \operatorname{argmin}_{\beta_k} J(\beta_k) \\ &= \operatorname{argmin}_{\beta_k} \|\alpha_k^{(TD)} - G\beta_k\|_2^2 \\ &= G^+ \alpha_k^{(TD)} \end{aligned} \quad (5.34)$$

と与えられる. ここで,

$$\alpha_k^{(TD)} = [\alpha_k^{(1)}, \dots, \alpha_k^{(n)}]'$$

であり, G はグラム行列

$$G = \begin{bmatrix} K(\mathbf{Y}^{(1)}, \mathbf{Y}^{(1)}) & K(\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}) & \dots & K(\mathbf{Y}^{(1)}, \mathbf{Y}^{(n)}) \\ K(\mathbf{Y}^{(2)}, \mathbf{Y}^{(1)}) & K(\mathbf{Y}^{(2)}, \mathbf{Y}^{(2)}) & & \vdots \\ \vdots & & \ddots & \vdots \\ K(\mathbf{Y}^{(n)}, \mathbf{Y}^{(1)}) & \dots & \dots & K(\mathbf{Y}^{(n)}, \mathbf{Y}^{(n)}) \end{bmatrix}$$

である。

以上により，求める関数は次のように推定される。

$$\begin{aligned}\hat{f}_k(\cdot) &= \sum_{i=1}^n (G^+ \alpha_k^{(TD)})_i K(\cdot, \mathbf{Y}^{(i)}) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_k^{(i)} (G^+)_{ij} K(\cdot, \mathbf{Y}^{(j)}).\end{aligned}\quad (5.35)$$

得られた関数 $\hat{f}_k(\cdot)$ と入力画像 $\hat{\mathbf{x}}$ より算出される $\mathbf{Y}^{(INPUT)}$ から，

$$\hat{\alpha}^{(OPT)} = \frac{[\hat{f}_1(\mathbf{Y}^{(INPUT)}), \dots, \hat{f}_{64}(\mathbf{Y}^{(INPUT)})]'}{\sum_{k=1}^{64} \hat{f}_k(\mathbf{Y}^{(INPUT)})}\quad (5.36)$$

を算出し， $\hat{\alpha}^{(OPT)}$ を用いて JPEG 再適用法を実行することで画質改善画像を得る。なお，式 (5.36) の分母は正規化のための係数である。

画像特徴量の次元削減

画像特徴量 \mathbf{Y} は 64×64 対称行列であり，その次元は $64 \times 63/2 = 2016$ となる。カーネル線形回帰における推定精度を高めるために，画像特徴量の情報をできるだけ保持しつつ \mathbf{Y} の次元を削減することを考える。本論文では \mathbf{Y} に対する固有値分解により得られる固有値及び固有ベクトルを利用することで新たな低次元の画像特徴量を作成する。

$$\lambda_i (i \in \{1, \dots, 64\}) \text{ を} \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{64}\quad (5.37)$$

を満たす \mathbf{Y} の固有値， \mathbf{v}_i を λ_i に対応する固有ベクトルとする。以上の準備のもとで，

新たな 3 種類の画像特徴量

$$\mathbf{Y}_M = [\lambda_1, \dots, \lambda_{64}]',\quad (5.38)$$

$$\mathbf{Y}_v = \mathbf{v}_1,\quad (5.39)$$

$$\mathbf{Y}_{li} = \sum_{i=1}^{64} \lambda_i \mathbf{v}_i\quad (5.40)$$

を提案し, \mathbf{Y}_o を

$$\mathbf{Y}_o = \text{vec}(X(Q, y)'X(Q, y)) \quad (5.41)$$

と定める. \mathbf{Y}_o は式 (5.28) を列ベクトルの形式で表したものである. 以降, $\mathbf{Y}_M, \mathbf{Y}_v, \mathbf{Y}_{li}, \mathbf{Y}_o$ を総称して \mathbf{Y} と呼ぶことにする.

カーネル関数の選択

カーネル線形回帰では採用するカーネル関数により推定結果が変化するが, 訓練データに対する適切なカーネル関数の構成方法についての明確な指針は存在していない. 本論文では以下に示す 4 種類のカーネル関数を回帰に用いる候補として採用する.

$$K_\sigma(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)}) = \exp\left(-\frac{1}{\sigma}\|\mathbf{Y}^{(i)} - \mathbf{Y}^{(j)}\|_2\right), \quad (5.42)$$

$$K_{\sigma_g}^{(G)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)}) = \exp\left(-\frac{1}{\sigma_g^2}\|\mathbf{Y}^{(i)} - \mathbf{Y}^{(j)}\|_2^2\right), \quad (5.43)$$

$$K_{\sigma_l}^{(L)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)}) = \exp\left(-\frac{1}{\sigma_l}\|\mathbf{Y}^{(i)} - \mathbf{Y}^{(j)}\|_1\right), \quad (5.44)$$

$$K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)}) = (\mathbf{Y}^{(i)'}\mathbf{Y}^{(j)} + 1)^n. \quad (5.45)$$

また, 各カーネル関数におけるパラメータ $\sigma, \sigma_g, \sigma_l, n$ は訓練データに対し leave-one-out cross-varidation を実施して決定する.

第 6 章

JPEG 再適用法の改良法に関する数値実験

本章では、第 5 章で提案した、画像データベースを利用した JPEG 画像の画質改善手法である KR 法, DB 法の有効性を検証すべく、数値実験を行なう。まず、提案した 2 種の画像特徴量 $\Sigma, X'X$ について、どちらがより画質改善に有効であるか検証を行なう。次に、KR 法において、次元を削減した画像特徴量とカーネル関数の最適な組み合わせを調査する。最後に、既存手法との性能比較実験を行なう。

6.1 実験用画像

画質改善の対象とする実験画像として、図 6.1 に示す、標準画像データベース SIDBA に含まれる 12 枚の 256×256 画素、256 階調のグレースケール画像 (Airplane, Barbara, Boat, Building, Cameraman, Goldhill, Lax, Lenna, Lighthouse, Mandrill, Milkcrow, Pepper) を採用し、これらの画像に JPEG 圧縮・復号を適用して入力画像を作成する。なお、圧縮の際に利用する量子化テーブルは表 5.1 に示した輝度成分用標準量子化テー

ブルに対し，画質調整のための定数を乗じたものを用いる．具体的には， Q_r を表 5.1 に示す量子化テーブル， p を品質調整パラメータ ($p = 1, 2, \dots, 100$) とし，

$$Q(p) = \begin{cases} \text{round} \left(\frac{50}{p} Q_r \right) & (1 \leq p < 50) \\ \text{round} \left(\frac{100-p}{50} Q_r \right) & (50 \leq p < 100) \\ \mathbf{1} & (p = 100) \end{cases} \quad (6.1)$$

と定める．ここで， $\mathbf{1}$ は全ての成分が 1 である 8×8 行列であり， $Q(p)$ の要素について，255 を超えた要素はその値を 255 へと修正するものとする． p が小さいほど画像の圧縮率が高くなることを付記しておく．本論文では $Q(50), Q(45), \dots, Q(5)$ の 10 種類の量子化テーブルを用いる．画像 Lenna を $Q(50), Q(45), \dots, Q(5)$ により圧縮して得られた入力画像を拡大して表示したものを図 6.2, 図 6.3 に示す．また，画像データベースとしては，入力画像に採用したものとは別の 252 枚，256 階調のグレースケール自然画像を用いた．

6.2 画像特徴量に関する実験

12 種類の入力画像に対して，JPEG 復号画像データベースに属する 252 枚の画像との画像特徴量の距離

$$D_Y = \|Y^{(INPUT)} - Y_i^{(DB)}\|_F \quad (6.2)$$

と最適係数ベクトルの距離

$$D_c = \|\alpha^{(OPT)} - \alpha_i^{(DB)}\|_2 \quad (6.3)$$

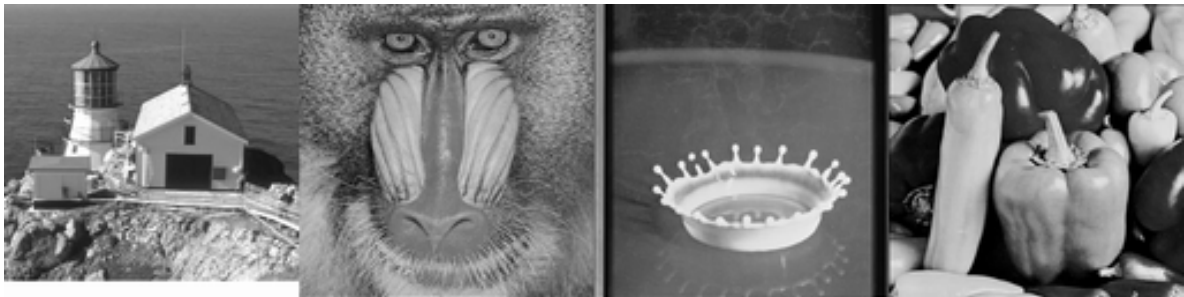
との相関，および， D_Y と $\alpha^{(OPT)}$ を用いた画質改善での PSNR 改善値 (IPSNR) との相関を調査する．PSNR は評価対象の画像と，それに対応する真の画像に対する平均二乗



(a) Airplane, Barbara, Boat, Buildings



(b) Cameraman, Goldhill, Lax, Lenna



(c) Lighthouse, Mandrill, Milkcrown, Pepper

図 6.1: 実験用原画像

誤差 MSE を用いて,

$$\text{PSNR (dB)} = 10 \log_{10} \frac{255^2}{\text{MSE}} \quad (6.4)$$

と定義される. D_Y と D_c が強い正の相関を持てば, 画像特徴量により DB 法において最適係数ベクトルに近い係数ベクトル α が得られることが期待でき, 各評価基準と IPSNR



(a) $p = 50$



(b) $p = 45$



(c) $p = 40$



(d) $p = 35$



(e) $p = 30$

図 6.2: 入力画像 Lenna ($p = 50, \dots, 30$)



(a) $p = 25$



(b) $p = 20$



(c) $p = 15$



(d) $p = 10$



(e) $p = 5$

図 6.3: 入力画像 Lenna ($p = 25, \dots, 5$)

値に強い負の相関を持てば、当該評価基準により良好な画質改善結果が得られることを意味する。なお、最適係数ベクトルによって最適な画質改善結果は得られるものの、最適係数ベクトルの近傍のベクトルが良好な画質改善結果を与える保証はない。したがって、後者の相関がより重要な意味を持つ。

上記の相関係数の一覧を表 6.1～表 6.3 に示す。 D_c については、どちらの特徴量も、どの画質においても大きな相関は見られないことが確認できる。一方で IPSNR については、全ての画質においてどちらの特徴量においても負の相関が確認できる。これは、画像特徴量の値が近いほど高い画質改善効果が期待できることを意味する。さらに、 $X'X$ が Σ に対してより強い相関を持つことから、画像特徴量として $X'X$ を用いるべきであると結論付ける。以後の実験では、 Y として $X'X$ を採用するものとする。

表 6.1: $D_{\mathbf{Y}}$ と D_c , $D_{\mathbf{Y}}$ と IPSNR の相関係数 ($p = 5, \dots, 20$)

Image	$p = 5$				$p = 10$			
	D_c		IPSNR		D_c		IPSNR	
	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ
Airplane	-0.104	-0.041	-0.642	-0.555	0.063	0.063	-0.738	-0.659
Barbara	0.118	-0.068	-0.607	-0.324	0.195	0.057	-0.441	-0.257
Boat	-0.040	-0.135	-0.678	-0.448	0.115	-0.006	-0.705	-0.516
Building	-0.072	-0.144	-0.619	-0.427	0.099	0.012	-0.580	-0.445
Cameraman	-0.033	-0.085	-0.700	-0.516	0.153	0.061	-0.714	-0.576
Goldhill	-0.033	-0.098	-0.660	-0.468	0.142	0.048	-0.636	-0.499
Lax	0.030	0.102	-0.585	-0.522	-0.007	0.040	-0.300	-0.316
Lenna	0.064	0.170	-0.644	-0.598	0.168	0.181	-0.670	-0.632
Lighthouse	-0.037	-0.080	-0.636	-0.450	-0.033	-0.050	-0.598	-0.485
Mandrill	-0.078	-0.096	-0.688	-0.527	0.009	-0.028	-0.558	-0.462
Milkcrow	0.026	0.071	-0.672	-0.624	0.192	0.142	-0.781	-0.709
Pepper	-0.026	0.118	-0.648	-0.616	0.129	0.167	-0.714	-0.665
Image	$p = 15$				$p = 20$			
	D_c		IPSNR		D_c		IPSNR	
	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ
Airplane	0.182	0.098	-0.861	-0.736	0.101	0.026	-0.881	-0.746
Barbara	0.164	0.041	-0.658	-0.453	0.113	-0.002	-0.691	-0.485
Boat	0.132	-0.017	-0.808	-0.590	0.070	-0.064	-0.860	-0.657
Building	0.098	0.003	-0.835	-0.665	0.067	-0.028	-0.861	-0.690
Cameraman	0.167	0.054	-0.838	-0.669	0.110	0.006	-0.861	-0.695
Goldhill	0.161	0.041	-0.822	-0.651	0.081	-0.022	-0.853	-0.685
Lax	-0.010	0.044	-0.459	-0.438	-0.024	0.016	-0.518	-0.458
Lenna	0.133	0.128	-0.788	-0.723	0.092	0.078	-0.849	-0.732
Lighthouse	-0.010	-0.019	-0.836	-0.681	0.003	-0.041	-0.840	-0.680
Mandrill	-0.051	-0.045	-0.717	-0.607	-0.113	-0.081	-0.777	-0.658
Milkcrow	0.227	0.127	-0.887	-0.763	0.180	0.069	-0.882	-0.749
Pepper	0.152	0.125	-0.839	-0.742	0.137	0.068	-0.881	-0.737

表 6.2: $D_{\mathbf{Y}}$ と D_c , $D_{\mathbf{Y}}$ と IPSNR の相関係数 ($p = 25, \dots, 40$)

Image	$p = 25$				$p = 30$			
	D_c		IPSNR		D_c		IPSNR	
	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ
Airplane	-0.028	-0.082	-0.887	-0.760	-0.149	-0.147	-0.833	-0.716
Barbara	0.003	-0.112	-0.805	-0.602	-0.093	-0.197	-0.745	-0.559
Boat	-0.004	-0.135	-0.883	-0.689	-0.105	-0.221	-0.833	-0.654
Building	-0.028	-0.108	-0.878	-0.718	-0.131	-0.180	-0.828	-0.683
Cameraman	-0.051	-0.120	-0.870	-0.713	-0.127	-0.179	-0.818	-0.674
Goldhill	-0.079	-0.127	-0.882	-0.723	-0.137	-0.173	-0.817	-0.675
Lax	0.041	0.032	-0.694	-0.569	0.104	0.052	-0.631	-0.499
Lenna	-0.034	-0.014	-0.871	-0.757	-0.109	-0.051	-0.798	-0.686
Lighthouse	-0.088	-0.115	-0.881	-0.724	-0.109	-0.073	-0.818	-0.677
Mandrill	-0.103	-0.020	-0.857	-0.736	-0.018	0.067	-0.800	-0.683
Milkcrow	0.004	-0.063	-0.889	-0.781	-0.106	-0.113	-0.848	-0.737
Pepper	-0.002	-0.024	-0.882	-0.755	-0.117	-0.073	-0.814	-0.697
Image	$p = 35$				$p = 40$			
	D_c		IPSNR		D_c		IPSNR	
	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ
Airplane	-0.144	-0.125	-0.835	-0.736	-0.084	-0.043	-0.826	-0.738
Barbara	-0.128	-0.220	-0.774	-0.608	-0.125	-0.211	-0.732	-0.582
Boat	-0.141	-0.248	-0.844	-0.680	-0.139	-0.237	-0.817	-0.667
Building	-0.138	-0.156	-0.837	-0.711	-0.099	-0.117	-0.827	-0.713
Cameraman	-0.136	-0.163	-0.829	-0.700	-0.091	-0.118	-0.806	-0.691
Goldhill	-0.125	-0.141	-0.831	-0.707	-0.057	-0.069	-0.820	-0.709
Lax	0.085	0.028	-0.634	-0.529	0.038	-0.016	-0.524	-0.469
Lenna	-0.135	-0.055	-0.820	-0.722	-0.099	-0.008	-0.803	-0.714
Lighthouse	-0.064	-0.002	-0.845	-0.716	0.004	0.041	-0.833	-0.715
Mandrill	-0.003	0.070	-0.826	-0.718	0.065	0.101	-0.805	-0.707
Milkcrow	-0.135	-0.100	-0.841	-0.758	-0.093	-0.040	-0.824	-0.754
Pepper	-0.147	-0.072	-0.836	-0.728	-0.136	-0.029	-0.808	-0.720

表 6.3: $D_{\mathbf{Y}}$ と D_c , $D_{\mathbf{Y}}$ と IPSNR の相関係数 ($p = 45, \dots, 50$)

Image	$p = 45$				$p = 50$			
	D_c		IPSNR		D_c		IPSNR	
	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ	$X'X$	Σ
Airplane	-0.099	-0.043	-0.828	-0.758	-0.080	-0.005	-0.830	-0.761
Barbara	-0.174	-0.217	-0.718	-0.595	-0.157	-0.184	-0.699	-0.585
Boat	-0.168	-0.231	-0.816	-0.683	-0.131	-0.186	-0.821	-0.690
Building	-0.042	-0.034	-0.841	-0.741	-0.015	-0.004	-0.824	-0.730
Cameraman	-0.085	-0.105	-0.820	-0.717	-0.077	-0.086	-0.811	-0.711
Goldhill	0.017	0.020	-0.837	-0.740	0.018	0.026	-0.827	-0.735
Lax	0.037	-0.002	-0.465	-0.443	0.089	0.082	-0.351	-0.380
Lenna	-0.088	0.049	-0.802	-0.738	-0.056	0.105	-0.790	-0.732
Lighthouse	0.130	0.143	-0.853	-0.742	0.107	0.107	-0.821	-0.714
Mandrill	0.145	0.141	-0.816	-0.726	0.168	0.151	-0.794	-0.699
Milkcrow	-0.060	0.042	-0.832	-0.783	-0.042	0.077	-0.825	-0.780
Pepper	-0.101	0.065	-0.822	-0.753	-0.128	0.058	-0.797	-0.741

6.3 画像特徴量とカーネル関数の組み合わせに関する実験

提案したカーネル関数と \mathbf{Y} に対して次元削減を行ない得られた画像特徴量に関して、入力画像に対し KR 法を適用して画質改善を行ない、改善画像における IPSNR によって最適な組み合わせを評価する。まず、各画像特徴量 $\mathbf{Y}_M, \mathbf{Y}_v, \mathbf{Y}_{li}, \mathbf{Y}_o$ に対し最も有効なカーネル関数を調査する。表 6.4 ~ 表 6.7 は各々の \mathbf{Y} に対し $K_\sigma(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$, $K_{\sigma_g}^{(G)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$, $K_{\sigma_l}^{(L)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$, $K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$ をそれぞれ用いて得られた 12 枚の画質改善画像の IPSNR 平均値を入力画像の品質 p ごとに掲載したものである。これらの表から、 $\mathbf{Y}_v, \mathbf{Y}_{li}, \mathbf{Y}_o$ に対しては、ほぼ全ての画質において最大の IPSNR 平均値を与えた $K_\sigma(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$ を採用し、 \mathbf{Y}_M に対しては $K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$ を採用するのが適切であると判断できる。

次に、これらの組み合わせ同士で同様に比較を行なうと、 \mathbf{Y}_v が 4 種類、 \mathbf{Y}_{li} が 3 種類、 \mathbf{Y}_o と \mathbf{Y}_M が 2 種類の画質において最大の IPSNR 平均値を与えていることが分かる。以上の結果より、KR 法として用いる画像特徴量とカーネル関数の組み合わせは \mathbf{Y}_v と $K_\sigma(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$ が最適であると結論付ける。

表 6.4: Y_M を用いた場合の 12 枚の画質改善画像における IPSNR 平均値

p	$K_{\sigma}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_g}^{(G)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_l}^{(L)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$
5	0.767	0.752	0.759	0.755
10	0.698	0.663	0.698	0.697
15	0.696	0.680	0.699	0.704
20	0.693	0.675	0.690	0.702
25	0.688	0.658	0.687	0.698
30	0.705	0.643	0.703	0.706
35	0.711	0.675	0.715	0.721
40	0.731	0.665	0.728	0.738
45	0.732	0.666	0.725	0.731
50	0.722	0.697	0.714	0.736

表 6.5: Y_v を用いた場合の 12 枚の画質改善画像における IPSNR 平均値

p	$K_{\sigma}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_g}^{(G)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_l}^{(L)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$
5	0.759	0.626	0.757	0.678
10	0.707	0.616	0.695	0.644
15	0.709	0.641	0.691	0.675
20	0.697	0.631	0.686	0.655
25	0.696	0.644	0.691	0.657
30	0.708	0.633	0.704	0.698
35	0.725	0.683	0.724	0.705
40	0.732	0.637	0.730	0.712
45	0.730	0.647	0.727	0.717
50	0.724	0.662	0.718	0.699

表 6.6: \mathbf{Y}_{li} を用いた場合の 12 枚の画質改善画像における IPSNR 平均値

p	$K_{\sigma}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_g}^{(G)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_l}^{(L)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$
5	0.758	0.705	0.756	0.665
10	0.698	0.618	0.692	0.686
15	0.703	0.641	0.689	0.670
20	0.694	0.635	0.696	0.681
25	0.697	0.642	0.688	0.680
30	0.707	0.659	0.700	0.699
35	0.724	0.691	0.724	0.702
40	0.743	0.661	0.735	0.733
45	0.740	0.699	0.732	0.716
50	0.737	0.666	0.732	0.723

表 6.7: \mathbf{Y}_o を用いた場合の 12 枚の画質改善画像における IPSNR 平均値

p	$K_{\sigma}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_g}^{(G)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_{\sigma_l}^{(L)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$	$K_n^{(P)}(\mathbf{Y}^{(i)}, \mathbf{Y}^{(j)})$
5	0.762	0.725	0.762	0.647
10	0.704	0.657	0.703	0.649
15	0.706	0.676	0.702	0.680
20	0.694	0.655	0.694	0.669
25	0.694	0.646	0.693	0.670
30	0.708	0.677	0.706	0.693
35	0.724	0.704	0.724	0.706
40	0.739	0.707	0.738	0.722
45	0.739	0.707	0.738	0.721
50	0.731	0.647	0.724	0.708

6.4 既存手法との性能比較実験

提案手法である DB 法, KR 法と, 従来の JPEG 再適用法 (RJ 法)[3], TV 正則化法を応用した手法 (ADF 法)[44], および 冗長 DCT 領域においてソフトシュリンケージを適用し画質改善を行なう手法 (RDCTS 法) [47] との性能比較実験を行なう.

DB 法に関する条件として, KR 法と同様の画像データベースを使用し, 画像特徴量は次元削減を適用しない \mathbf{Y}_o を採用する. ADF 法では, 最急降下法のステップ幅 $t^{(\ell)}$ に関して, 文献 [44] では実際の画質改善問題においては ℓ に依存しない定数 $1/2$ を用いており, 本論文でもこの値を採用する. また, 求解における更新式計算の反復回数に関しての明確な指針が与えられていないため, ここでは, 各データベース画像に対し JPEG 圧縮を行なって得られた画像に対し, ADF 法を適用した場合の最適な反復更新回数の平均値を l_{max} として設定する. 表 6.8 に, 具体的な l_{max} の値を示す. なお, TV 項の勾配に対する重みづけについて, 今回の実験においてより優れた結果を示した, 重みづけを行わない従来の TV 項を用いた場合の改善結果を掲載している.

RDCTS 法では, 更新式計算の反復回数 l_{max} を文献 [47] と同様の 20 回とし, シュリンケージパラメータの初期値 $\tau^{(0)}$ については RDCTS 法が最大限の性能を発揮できるよう, 各入力画像に対し 0.1 刻みで変化させながら IPSNR を調べ, 最も良い IPSNR を与

表 6.8: ADF 法による更新式の反復回数

p	5	10	15	20	25	30	35	40	45	50
l_{max}	13	16	11	9	8	7	7	6	6	5

えた値を採用している。なお、現実問題において画質改善を行なう際は入力画像に対する RDCTS 法の最適なパラメータを求めることは不可能であるので、現実問題においては今回の実験の結果より性能が劣化することを付記しておく。

6.4.1 客観的評価

それぞれの手法により得られた画質改善画像について、IPSNR、および画質改善評価基準 SSIM[48] の改善値 (ISSIM) の比較を行なうことにより、性能評価を行なう。SSIM は、評価対象の画像を y 、 y に対応する真の画像を x とし、 x, y の同じ位置に存在する画素に対して、

$$\text{SSIM} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (6.5)$$

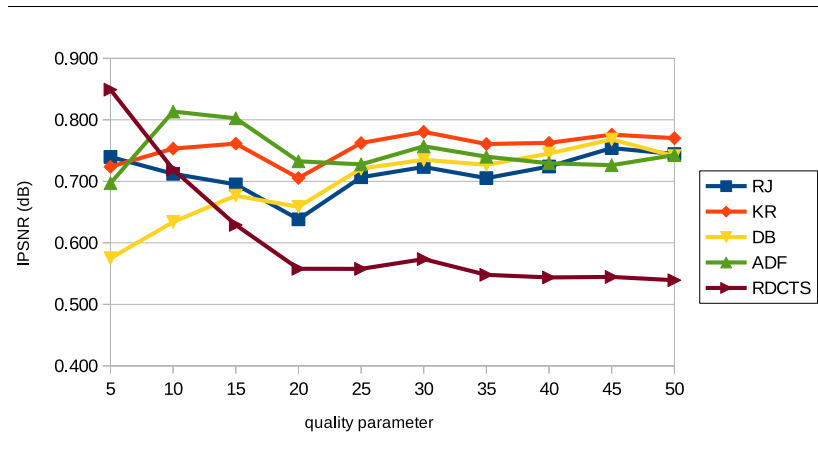
と定義される。ここで、 μ_x, σ_x^2 はそれぞれ、 x における着目画素を中心とする領域 W_x に対し、回転対称なガウスフィルタを適用したうえで計算される平均と不偏分散であり、 μ_y, σ_y^2 は y における着目画素を中心とする領域 W_y から同様にして計算される平均と不偏分散である。また、 σ_{xy} は W_x, W_y から計算される共分散である。 C_1, C_2 は分母の値が小さくなる事により SSIM の値が不安定になるのを防ぐための定数であり、文献 [48] では、 W_x, W_y を 11×11 画素の正方形、

$$C_1 = (0.01 \cdot 255)^2, C_2 = (0.03 \cdot 255)^2$$

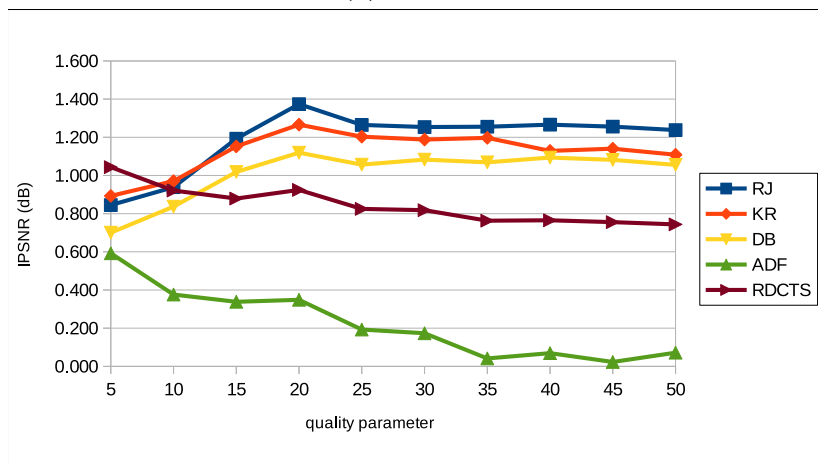
と定めている。式 (6.5) による計算が可能な全ての画素から得られる SSIM の平均値を、画像全体に対する SSIM 値としている。

それぞれの画質改善画像の IPSNR を図 6.4 ~ 図 6.7 に、各品質における 12 枚の画質改善画像の IPSNR の平均値、および ISSIM の平均値を表 6.9, 表 6.10 に示す。

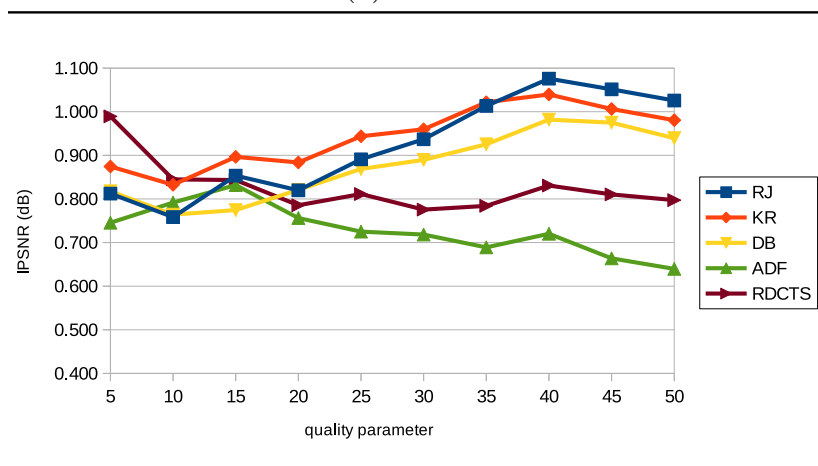
図 6.4 ~ 図 6.7 より, JPEG 再適用法による画質改善手法同士の比較を行なうと, DB 法による画質改善については, RJ 法よりも IPSNR が低くなっている結果が散見される. 一方で, KR 法については, Barbara において RJ 法 に劣る結果となったものの, 他の画像については多くの場合において IPSNR が上回っていることが確認できる. また, DB 法の IPSNR をほぼ全ての場合で上回っており, カーネル線形回帰による関数の推定が成功したものと考えられる. KR 法とその他の既存手法に関しては, $p = 5$ では RDCTS 法に及ばなかったが, 広い範囲の品質において, 多くの画像で KR 法が最も優れた IPSNR を示していることが確認できる. Cameraman, Lax, Milkcrow, Pepper で ADF 法を下回ったことに関しては, これらの画像は強いエッジ領域を多く含み, JPEG 圧縮・復号時に発生した多量のモスキートノイズを除去しきれなかったためと考えられる. 表 6.9, 表 6.10 からも, KR 法が広い範囲の品質において他手法よりも優れた IPSNR, ISSIM を与えていることが確認できる.



(a) Airplane

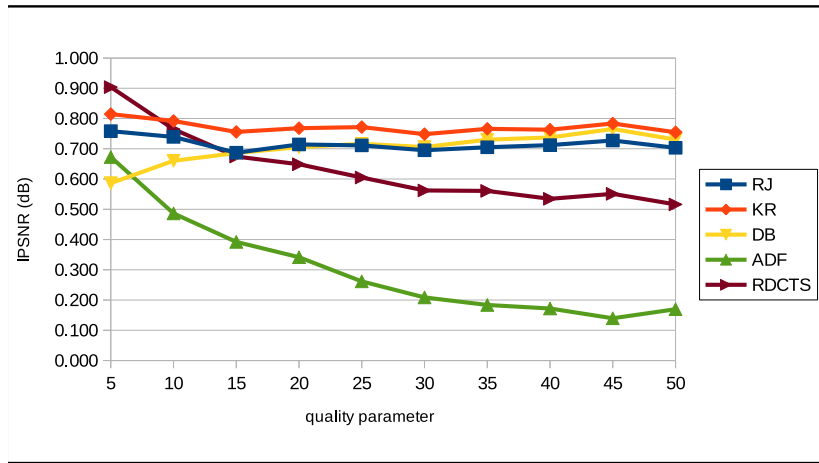


(b) Barbara

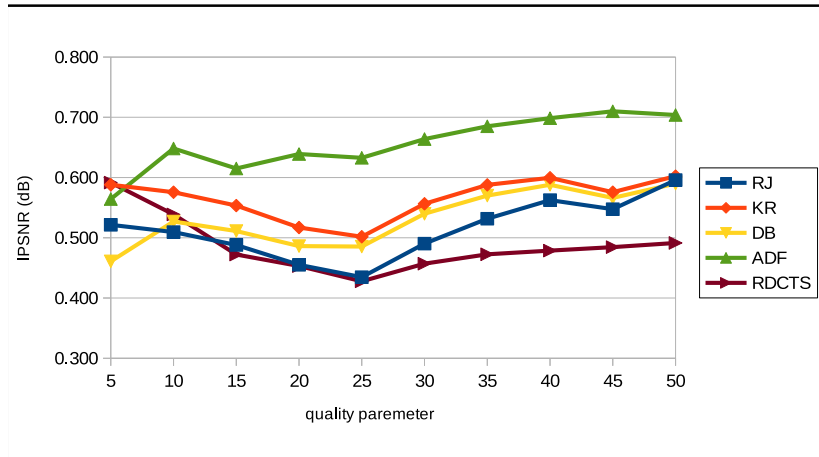


(c) Boat

図 6.4: 画質改善画像の IPSNR (Airplane, Barbara, Boat)



(a) Buildings

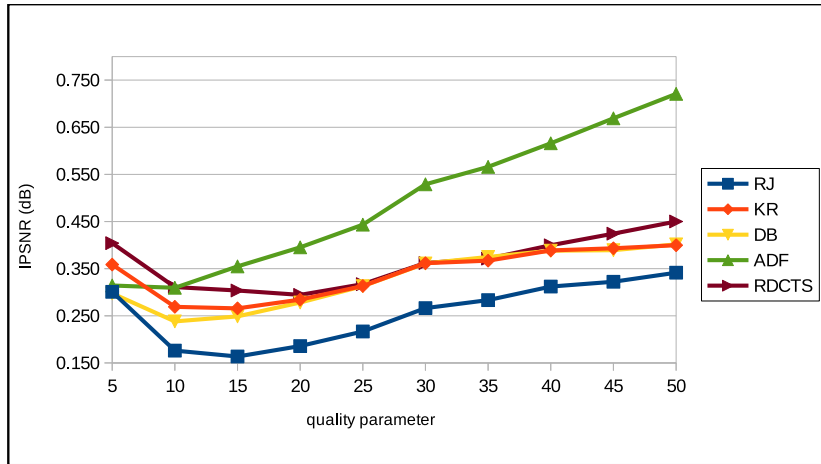


(b) Cameraman

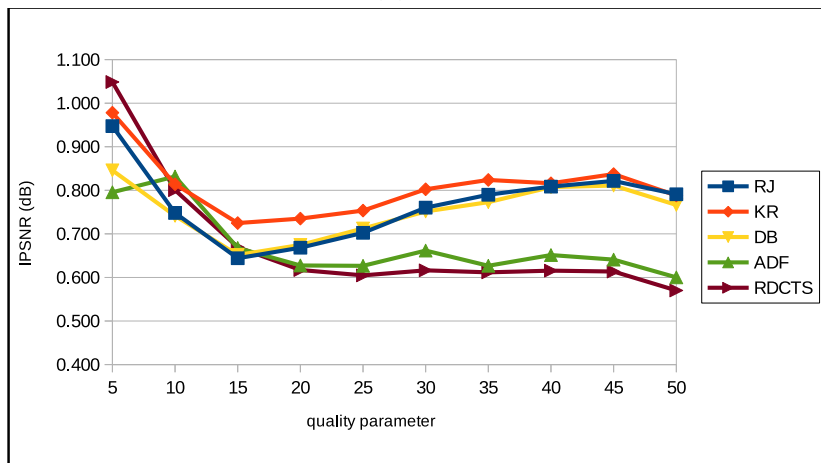


(c) Goldhill

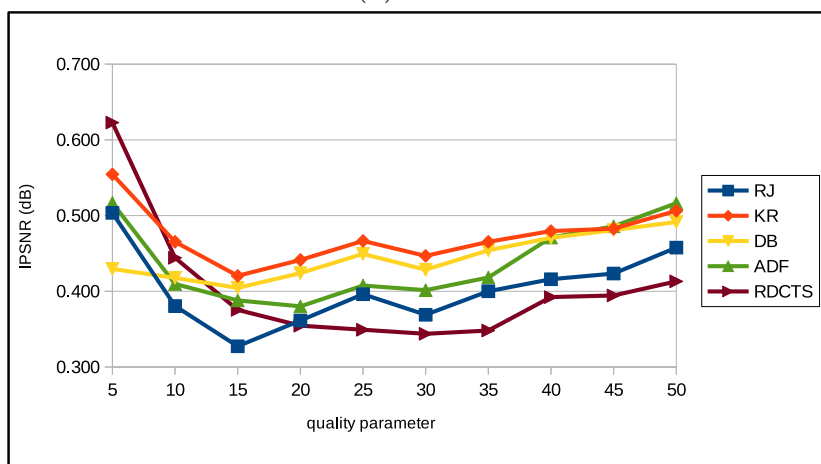
図 6.5: 画質改善画像の IPSNR (Buildings, Cameraman, Goldhill)



(a) Lax

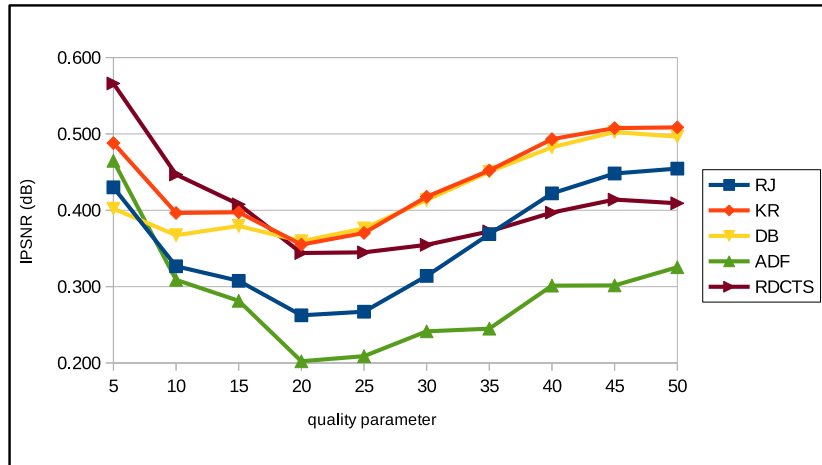


(b) Lenna

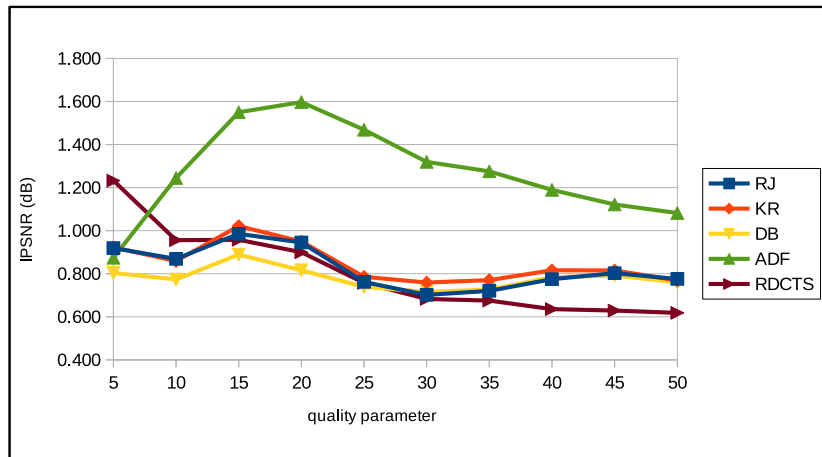


(c) Lighthouse

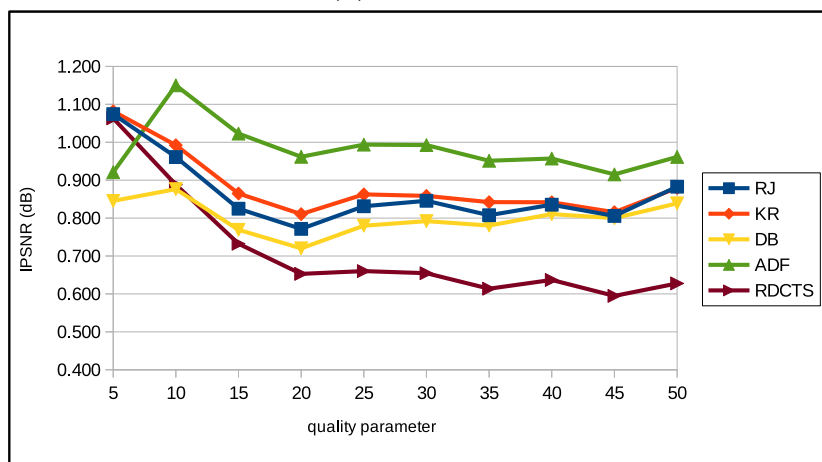
図 6.6: 画質改善画像の IPSNR (Lax, Lenna, Lighthouse)



(a) Mandrill



(b) Milkcrown



(c) Pepper

図 6.7: 画質改善画像の IPSNR (Mandrill, Milkcrown, Pepper)

表 6.9: 各改善手法での 12 枚の画質改善画像における IPSNR 平均値

p	RJ	KR	DB	ADF	RDCTS
5	0.724	0.759	0.619	0.653	0.854
10	0.654	0.707	0.628	0.668	0.704
15	0.652	0.709	0.636	0.646	0.634
20	0.649	0.697	0.638	0.620	0.595
25	0.645	0.696	0.650	0.587	0.566
30	0.660	0.708	0.667	0.586	0.561
35	0.682	0.725	0.684	0.560	0.553
40	0.709	0.732	0.710	0.577	0.562
45	0.712	0.730	0.713	0.560	0.559
50	0.716	0.724	0.703	0.574	0.556

表 6.10: 各改善手法での 12 枚の画質改善画像における ISSIM 平均値

p	RJ	KR	DB	ADF	RDCTS
5	0.0289	0.0353	0.0288	0.0255	0.0382
10	0.0217	0.0251	0.0225	0.0146	0.0240
15	0.0175	0.0202	0.0183	0.0104	0.0172
20	0.0146	0.0168	0.0154	0.0072	0.0132
25	0.0131	0.0150	0.0140	0.0054	0.0108
30	0.0121	0.0138	0.0130	0.0051	0.0091
35	0.0113	0.0127	0.0122	0.0035	0.0079
40	0.0106	0.0117	0.0114	0.0038	0.0071
45	0.0097	0.0107	0.0106	0.0029	0.0062
50	0.0091	0.0100	0.0098	0.0034	0.0056

6.4.2 画質改善画像の例

品質 $p = 5, 25, 50$ の入力画像 Lenna, Cameraman から得られる改善画像を図 6.8~ 図 6.13 に示す. Lenna に関して, $p = 5$ では, KR 法が DB 法や RJ 法と比較して僅かにブロックノイズを抑制できており, RDCTS 法と比較しても同程度の画質改善効果が確認できる. また, ADF 法は他の 3 手法と比べてあまりよい改善効果が得られていないことがわかる. $p = 25, 50$ においても KR 法が他手法よりも良い改善効果を与えていることが被写体の女性の右目周辺から判断できる. 一方 Cameraman に関しては, $p = 5$ においては Lenna と同様の判断を下すことができるが, $p = 25, 50$ と画質が高くなると, KR 法では被写体のカメラや男性と背景の境界に発生したモスキートノイズの除去性能が ADF 法と比較して若干弱いことが確認される.



図 6.8: 入力画像 Lenna ($p = 5$) から得られる画質改善画像



(a) 入力画像



(b) RJ



(c) DB



(d) KR



(e) ADF



(f) RDCTS

図 6.9: 入力画像 Lenna ($p = 25$) から得られる画質改善画像



(a) 入力画像



(b) RJ



(c) DB



(d) KR



(e) ADF



(f) RDCTS

図 6.10: 入力画像 Lenna ($p = 50$) から得られる画質改善画像



(a) 入力画像



(b) RJ



(c) DB



(d) KR



(e) ADF



(f) RDCTS

図 6.11: 入力画像 Cameraman ($p = 5$) から得られる画質改善画像



(a) 入力画像



(b) RJ



(c) DB



(d) KR



(e) ADF



(f) RDCTS

図 6.12: 入力画像 Cameraman ($p = 25$) から得られる画質改善画像



(a) 入力画像



(b) RJ



(c) DB



(d) KR



(e) ADF



(f) RDCTS

図 6.13: 入力画像 Cameraman ($p = 50$) から得られる画質改善画像

第 7 章

総括

7.1 本論文による成果

本論文第 3 章では，単板式デジタルカメラの現像過程で適用されるデモザイキング処理について，ブロック分割した画像に対し線形回帰による欠損データ補間手法を適用する，ブロックベース線形回帰によりデモザイキングを実現する手法を提案した．また，本提案手法において，重複を許してブロック分割を行なった際に得られる，補間画素に対する複数の推定値について，推定誤差の分散を最小とする線形凸結合を構成する手法を提案し，補間精度を向上させた．さらに，写真撮影に使用する CFA パターンの繰り返しに着目した，補間処理における計算量削減方法，および，上記の最小分散凸結合による補間精度向上法について，新たに得られる推定候補値の情報を逐次反映させ計算する方法を提案し，計算の効率化を実現した．

本論文第 4 章では，第 3 章で提案したブロックベース線形回帰によるデモザイキング手法について，数値実験を実施することにより有効性を確認した．具体的には，最小分散凸結合を利用した補間精度の向上法がデモザイキングにおいて有効に機能することや，既存

のデモザイキング手法との性能比較実験を行なうことにより，提案手法が偽色の抑制に効果的であること，および，色情報の変化が比較的小さい撮影画像に対して優れたデモザイキング性能を有することを確認した．なお，本論文では Bayer CFA パターンを利用して得られた RAW 画像に対してのデモザイキングを実施したが，提案手法は CFA パターンによらずデモザイキングを適用することが可能である．そのため，デモザイキングによりカラー画像を得るための手段としてだけではなく，CFA パターン同士における性能比較としての用途で本提案手法を利用することも可能であると考えられる．

本論文第 5 章では，従来の JPEG 再適用法における加重平均の重み係数について，どの画像に対しても一定の重み係数を用いるが故に画質改善性能に限界があるという問題に対して，JPEG 画像から算出される画像特徴量の定義と画像データベースの導入を行ない JPEG 再適用法の改良を行なった．入力画像に対して，画像特徴量が最も近いデータベース画像に対する最適な重み係数を利用して JPEG 再適用法を適用する手法，および，データベース画像を訓練データとして利用したカーネル線形回帰により推定し，当該関数を利用して算出した重み係数を利用して JPEG 再適用法を適用する手法を提案した．

本論文第 6 章では，第 5 章で提案した JPEG 再適用法の改良法について，数値実験を実施することにより有効性を確認した．画像特徴量の有効性や関数推定に適した画像特徴量やカーネル関数の検証，および既存の JPEG 画質改善手法との性能比較実験を行なうことにより，多くの画像において，カーネル線形回帰を利用した JPEG 再適用法が従来の JPEG 再適用法に勝る結果を示し，その他の既存手法に対しても優れた画質改善性能を有することを確認した．

7.2 今後の課題

ブロックベース線形回帰に関して、計算の効率化を実現したとはいえ、デモザイキング処理の完了までに Intel(R) Core(TM) i7-4600U CPU 2.10GHz, 8GB RAM の環境で 1 枚 2 分弱の時間を要するため、デジタルカメラ内部での現像において提案手法を用いることは困難であり、現状では PC 等でのオフライン手法での利用が必要となる。補間性能を維持しながらの計算コストのさらなる削減が今後の課題となる。

画像データベースを用いた JPEG 再適用法については、カーネル線形回帰による関数推定の頑健性に関する調査や、従来の JPEG 再適用が画質改善において良好に機能する機序の理論的解明が挙げられる。

謝辞

本研究を行うにあたって，多忙な中，指導教員として多くの御指導，御助言を頂きました，情報数理学研究室教授 田中 章先生，副指導教員として貴重な御助言を頂きました情報認識学研究室教授 工藤 峰一先生，知能情報学研究室教授 杉本 雅則先生，情報解析学研究室教授 今井 英幸先生に深く感謝いたします。また，研究，論文作成を行うにあたり御協力頂いた情報数理学研究室の皆様にも感謝いたします。

参考文献

- [1] 神崎洋治, 西井美鷹, 体系的に学ぶ デジタルカメラのしくみ 第2版, 日経BP ソフトプレス, 2009.
- [2] ISO/IEC IS 10918-1. ITU-T Rec. T.81, “Digital Compression Coding of Continuous Tone Still Images - Part 1; Requirements and Guidelines,” Feb. 1994.
- [3] A. Nosratinia, “Enhancement of JPEG-Compressed Images by Reapplication of JPEG,” *The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, vol.27, pp.69–79, Feb. 2001.
- [4] A. Tanaka, T. Ogawa, and M. Haseyama, “Image Inpainting by Block-Based Linear Regression with Optimal Block Selection,” *Proceedings of APSIPA Annual Summit & Conference 2012*, Paper ID:8, Dec. 2012.
- [5] K. Kohno and A. Tanaka, “Image Inpainting Based on Block-Based Linear Regression and Minimum Variance Convex Combinations,” *Proceedings of the 15th IASTED International Conference on Signal and Image Processing (SIP 2013)*, pp.368–373, 2013.
- [6] 河野克也, 田中章, “ブロックベース線形回帰を利用したデモザイキング,” *電子情報*

-
- 通信学会論文誌 D, vol.J100-D, no.5, pp.605–612, 2017.
- [7] 河野 克也, 田中 章, “画像データベースと JPEG 再適用を利用した JPEG 画像の画質改善,” 電子情報通信学会技術研究報告, vol.111, no.465, pp.13–18, March 2012.
- [8] K. Kohno, A. Tanaka and H. Imai, “A Novel Criterion for Quality Improvement of JPEG Images Based on Image Database and Re-application of JPEG,” Proceedings of APSIPA Annual Summit & Conference 2012, Paper ID:112, Dec. 2012.
- [9] 河野克也, 田中章, “カーネル回帰を用いた JPEG 再適用法による JPEG 画像の画質改善,” 第 28 回信号処理シンポジウム講演論文集, pp.446–447, Nov. 2013.
- [10] 福水健次, カーネル法入門 : 正定値カーネルによるデータ解析, 朝倉書店, 2010.
- [11] 映像情報メディア学会, デジカメの画像処理, オーム社, 2015.
- [12] 津軽海渡, 木村誠聡, 図解雑学 デジタルカメラ, ナツメ社, 2002.
- [13] B.E. Bayer, “Color imaging array,” U.S. Patent 3971065, 1975.
- [14] P. Longre, X. Zhang, P.B. Delahunt, and D.H. Brainard, “Perceptual Assessment of Demosaicing Algorithm Performance,” Proceedings of the IEEE, vol.90, pp.123–132, Jan. 2002.
- [15] H. Malvar, L. He, and R. Cutler, “High-Quality Linear Interpolation for Demosaicing of Bayer-Patterned Color Images,” Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004), pp.485–488, 2004.

-
- [16] L. Zhang and X. Wu, “Color Demosaicking Via Directional Linear Minimum Mean Square-Error Estimation,” *IEEE Transactions on Image Processing*, vol.14, no.12, pp.2167–2178, 2005.
- [17] D. Paliy, V. Katkovnik, R. Bilcu, S. Alenius, and K. Egiazarain, “Spatially Adaptive Color Filter Array Inter-Poalition for Noiseless and Noisy Data,” *International Journal of Imaging Systems and Technology*, vol.17, no.3, pp.105–122, 2007.
- [18] I. Pekkucuksen and Y. Altunbasak, “Gradient Based Threshold Free Color Filter Array Interpolation,” *Proceedings of IEEE International Conference on Image Processing (ICIP2010)*, pp.137–140, 2010.
- [19] L. Zhang, X. Wu, A. Buades, and X. Li, “Color Demosaicking by Local Directional Interpolation and Non-Local Adaptive Thresholding,” *Journal of Electronic Imaging*, vol.20, no.2, 2011.
- [20] G.-G. Wang, X.-C. Zhu, and Z.-L. Ga, “Image Demosaicing by Non-Local Similarity and Local Correlation,” *Proceedings of IEEE International Conference on Signal Processing (ICSP 2012)*, vol.2, pp.806–810, 2012.
- [21] S.P. Jaiswal, O.C. Au, V. Jakhetiya, Y. Yuan, and H. Yang, “Exploitation of Inter-Color Correlation for Color Image Demosaicking,” *Proceedings of IEEE International Conference on Image Processing (ICIP2014)*, pp.1812–1816, Oct. 2014.

-
- [22] D. Kiku, Y. Monno, M. Tanaka and M. Okutomi, “Residual Interpolation for Color Image Demosaicking,” Proceedings of IEEE International Conference on Image Processing (ICIP2013), pp.2304–2308, Sept. 2013.
- [23] D. Kiku, Y. Monno, M. Tanaka and M. Okutomi, “Minimized-Laplacian Residual Interpolation for Color Image Demosaicking,” Proceedings of IS&T/SPIE Electronic Imaging (EI2014), Digital Photography X, vol.9023, pp.90230L–1–8, Feb. 2014.
- [24] Y. Monno, D. Kiku, M. Tanaka and M. Okutomi, “Adaptive Residual Interpolation for Color Image Demosaicking,” Proceedings of IEEE International Conference on Image Processing (ICIP2015), pp.3861–3865, Sept. 2015.
- [25] D. Kiku, Y. Monno, M. Tanaka and M. Okutomi, “Beyond Color Difference: Residual Interpolation for Color Image Demosaicking,” IEEE Transactions on Image Processing (TIP), vol.25, pp.1288–1300, March 2016.
- [26] K. He, J. Sun, and X. Tang, “Guided Image Filtering,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.35, no.6, pp.1397–1409, 2013.
- [27] 天野敏之, 佐藤幸男, “固有空間法を用いた BPLP による画像補間,” 電子情報通信学会論文誌 D, vol.J85-D2, no.3, pp.457–465, 2002.
- [28] 田中章, 宮腰政明, “線形推定理論に基づく固有空間 BPLP 法の改善,” 電子情報通信学会論文誌 D, vol.J90-D, no.10, pp.2840–2847, 2007.
- [29] 田中章, 小川貴弘, 長谷山美紀, 宮腰政明, “固有空間 BPLP 法の補間精度に関する

-
- 解析,” 電子情報通信学会論文誌 A, vol.J94-A, no.2, pp.116–126, 2011.
- [30] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and Its Applications*, John Wiley & Sons, 1971.
- [31] J.R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, 1988.
- [32] A. Tanaka and K. Kohno, “Multi-frame Image Denoising Based on Minimum Noise Variance Convex Combination with Difference-Based Noise Variance Estimation,” *IEICE Transactions on Fundamentals*, vol.E96-A, no.10, pp.2066–2070, 2013.
- [33] “Kodak Image Dataset,” <http://r0k.us/graphics/kodak/>.
- [34] “JPEG File Interchange Format,” <https://www.w3.org/Graphics/JPEG/jfif3.pdf>, 1992.
- [35] H.C. Leeves and J.S. Lim, “Reduction of Blocking Effect in Image Coding,” *SPIE Journal of Optical Engineering*, vol.23, no.1, pp.34–37, Feb. 1984.
- [36] Draft ITU-T Rec. H.263 Version 2, “Video Coding for Low Bit Rate Communication,” 1998.
- [37] V.K. Nath, D. Hazarika and A. Mahanta, “Blocking Artifacts Reduction Using Adaptive Bilateral Filtering,” *Proceedings of IEEE International Conference on Signal Processing and Communications (SPCOM2010)*, pp. 243–250, July 2010.
- [38] Y. Monobe, T. Watanabe, A. Kojima and T. Kurosawa, “Enhancement of JPEG

-
- Coded Images Using a Region-Based Algorithm,” *The Journal of the Institute of Image Information and Television Engineers*, vol.56, no.8, pp.1291–1298, Aug. 2002.
- [39] T. Oezcelik and J.C. Brailean and A.K. Katsaggelos, “Image and Video Compression Algorithms Based on Recovery Techniques Using Mean Field Annealing,” *Proceedings of the IEEE*, vol.83, no.2, pp.304–316, Feb. 1995.
- [40] J. Li and C.-C.J. Kuo, “Coding Artifact Removal with Multiscale Postprocessing,” *Proceedings of IEEE International Conference on Image Processing (ICIP1997)*, pp.45–48, Oct. 1997.
- [41] R.A. Gopinath, M. Lang, H. Guo, and J.E. Odegard, “Wavelet-Based Post Processing of Low Bit Rate Transform Coded Images,” *Proceedings of IEEE International Conference on Image Processing (ICIP1994)*, pp.913–917, Dec. 1994.
- [42] H. Choi, T. Kim, “Blocking-Artifact Reduction in Block-Coded Images Using Wavelet-Based Subband Decomposition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.10, no.5, pp.801–815, Aug. 2000.
- [43] L. Rudin, S. Osher and E. Fatemi, “Nonlinear Total Variation Based Noise Removal Algorithms,” *Journal of Physica D: Nonlinear Phenomena*, vol.60, pp.259–268, Nov. 1992.
- [44] F. Alter, S. Durand and J. Froment, “Adapted Total Variation for Artifact Free Decompression of JPEG Images,” *Journal of Mathematical Imaging and Vision*,

-
- vol.23, no.2, pp.199–211, Sept. 2005.
- [45] 小松隆, 齊藤隆弘, “TV 正則化法を用いた JPEG 圧縮画像の超解像デコーディング,” 電子情報通信学会論文誌 D, vol.J90-D, no.7, pp.1671–1674, July. 2007.
- [46] 後藤富朗, 大野永貴, 平野智, 桜井優, “Total Variation 正則化を用いた JPEG 圧縮画像に対するノイズ除去法,” 映像情報メディア学会誌, vol.64, no.11, pp.1647–1654, Nov. 2010.
- [47] T. Komatsu, Y. Ueda and T. Saito, “Super-Resolution Decoding of JPEG-compressed Image Data with the Shrinkage in the Redundant DCT Domain,” Proceedings of Picture Coding Symposium 2010, pp.114–117, Dec. 2010.
- [48] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE Transactions on Image Processing, vol.13, no.4, pp.600–612, April 2004.