

2019 年度

早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻 修士論文

学習型超解像の特性を活かした高解像度化技術

High-Resolution Technology Utilizing the Characteristics of Example-based Super-Resolution

矢野 仁愛

(5118F104-4)

提出日 2020.01.29

指導教員 渡辺裕教授 (印)

研究指導名：オーディオビジュアル情報処理研究

目次

第1章	序論	1
1.1	研究の背景.....	1
1.2	本研究の目的.....	2
1.3	本論文の構成.....	2
第2章	関連研究	4
2.1	まえがき.....	4
2.2	補間技術を用いた高解像度化技術.....	4
2.3	超解像技術の概要および分類.....	5
2.3.1	超解像技術.....	5
2.3.2	超解像技術の分類.....	6
2.3.2.1	再構成型超解像.....	7
2.3.2.2	学習型超解像.....	8
1.)	自己合同性を利用した超解像技術.....	9
2.)	事前学習による辞書を利用した超解像技術.....	10
2.4	提案手法に使用する従来手法.....	13
2.4.1	Super-Resolution Convolutional Neural Network (SRCNN).....	13
2.4.2	Sparse-coding Super-Resolution (ScSR).....	15
2.4.3	Adjusted anchored neighbor regression (A+).....	17
2.5	画質評価指標.....	18
2.5.1	PSNR.....	18
2.5.2	SSIM.....	19
2.5.3	BRISQUE.....	20
2.5.4	各種画質評価指標の比較.....	22
2.5.5	画質評価指標と評価実験の関係.....	25
2.6	むすび.....	26
第3章	提案手法	27
3.1	まえがき.....	27
3.2	提案手法の概要と構成技術.....	27
3.2.1	マルチステップ超解像技術 (多段超解像技術).....	28

3.2.2	方向性を考慮した超解像技術.....	31
3.3	提案手法の構成.....	33
3.4	むすび.....	35
第4章	評価実験.....	36
4.1	まえがき.....	36
4.2	評価実験.....	36
4.2.1	実験方法.....	36
4.2.2	実験結果.....	36
4.2.3	まとめと考察.....	50
4.3	むすび.....	51
第5章	結論.....	52
	謝辞.....	53
	参考文献.....	54
	図一覧.....	61
	表一覧.....	62
	研究業績.....	63

第1章 序論

1.1 研究の背景

近年, テレビやスマートフォンをはじめとする出力デバイスの解像度が飛躍的に向上している. 2018年12月から, BS・東経110度CSによる4K・8K実用放送である「新4K8K実用放送」も開始され, 現在, 4K・8K放送がより一層推進されている [1][2]. この推進に付随する形で, 4K(対応)テレビの出荷台数も年々向上しており, 2019年11月期における4K(対応)テレビの出荷台数は約23.6万台, 薄型テレビに占める割合は57.6%となっている [3][4]. 4K・8Kテレビの出荷台数は, 今後さらに拡大する見込みであり, 一般社団法人電子情報技術産業協会(JEITA)の見通しでは, 2022年に64%まで拡大するとされている [2][4].

また, スマートフォンについても同様の傾向がみられる. 2017年に発売されたiPhone 8やiPhone 8 Plusでは, 解像度がそれぞれ $1,334 \times 750$, $1,920 \times 1,080$ [5]であったが, 2019年発売のiPhone 11 ProやiPhone 11 Pro Maxでは, $2,436 \times 1,125$ および $2,688 \times 1,242$ [6]になるなど, 年々解像度が拡大している. Androidにおいても解像度の拡大がみられ, 2018年8月にはXperia XZ2 Premium SOV38で4Kサイズ($3,840 \times 2,160$)のスマートフォンも発売され, 2019年モデルではFHD+以上のサイズが大半を占めている [7].

このような出力機器の解像度の向上に伴い, 低解像度画像を高解像度ディスプレイで出力する機会がますます増加している. その一方で, 低解像度画像を従来の補間技術によりそのまま拡大すると, ぼやけやジャギーといった画質劣化が生じる問題があり, これらを防ぐためには, 事前の信号処理によって画素数を増やす高解像度化技術が必要となる [8][9]. 従来は, この高解像度化技術として, 補間技術が用いられてきた. しかしながら, 補間技術では, 不足している画素を, 参照画像中に存在する複数画素の平均値などから算出するため, 高周波成分が生成されず, 依然として出力画像がぼやけていた [8][10]. そこで, 高周波成分も復元可能な高解像度化技術である, 超解像技術が考案され, 様々な手法が提案されている [19]. 中でも, 機械学習により生成した辞書を利用して超解像処理を行う「学習型超解像 (Example-based Super-Resolution)」は, その能力の高さから注目を浴びている [11].

従来、超解像技術は、ピーク信号対雑音比：Peak Signal-to Noise Ratio (PSNR) や構造的類似性：Structural similarity (SSIM) [12]といった、再構成精度による客観的画質評価値を向上させることを目的に行われてきた。一方で、再構成精度による画質評価と知覚品質とが一致しないことが示されている [13][14]。この研究報告によると、再構成誤差を最小にするモデルは視覚的に不快な結果を生む傾向があり、反対に、優れた知覚品質を持つモデルは PSNR や SSIM のような歪み測定では評価されない傾向がある [13]。したがって、高解像度化技術の考案において、再構成精度、および、知覚品質の、双方を改善することが最大の課題である。

以上のような背景から、本研究では、学習型超解像の特性を活かし、再構成精度をできる限り保ちながら、知覚品質を改善することを目指す。

1.2 本研究の目的

超解像技術分野において驚異的な能力を発揮している学習型超解像は、超解像処理前に学習済みの辞書や情報（以下単に辞書）を使用して、低解像度画像から高解像度画像を取得する技術である。学習型超解像は、高拡大倍率においても高精細な画像を再現することが可能であり、再構成精度（復元精度）や知覚品質の高さから注目されている [11]。一方で、前述のように、再構成精度の向上を図ると知覚品質が低下してしまうことが知られており、双方の向上が課題である。

そこで、本研究では、従来手法と同程度の再構成精度のもとで、知覚品質を大幅に向上させることを目的としている。本研究では、学習型超解像である Super-Resolution Convolutional Neural Network (SRCNN) [15]、Adjusted anchored neighbor regression (A+) [16]、および Sparse-coding Super-Resolution (ScSR) [17]について、その特性を活かしたマルチステップ超解像技術 (Multi-Step Super-Resolution : MSSR) により、知覚品質の向上を図る。さらに、過去に提案した「方向性を考慮した超解像技術 (回転・反転を伴う超解像)」[18]をマルチステップ超解像技術に適用することにより、従来手法に対する再構成精度の低下を抑制する。

1.3 本論文の構成

以下に本論文の構成を示す。

- 第1章 本章であり，研究の背景，および目的について述べている．
- 第2章 関連研究について述べる．まず，超解像技術の基礎技術である補間技術について簡潔に述べ，超解像技術の概要と分類について記述する．続いて，学習型超解像の代表手法について記述し，本研究で用いる SRCNN, A+, および ScSR について述べる．最後に，画質評価指標である PSNR, SSIM, および非参照画質評価モデルによる評価方法について記述する．
- 第3章 提案手法の概要と構成について述べる．まず，提案手法の基礎技術であるマルチステップ超解像技術と，方向性を考慮した超解像技術について説明する．続いて，これらの手法を合成した提案手法について説明する．
- 第4章 提案手法の評価実験の概要，結果，および考察について述べる．
- 第5章 本論文のまとめを行う．

第2章 関連研究

2.1 まえがき

本章では、本研究の関連研究について述べる。

本研究では、学習型超解像の原点的手法である Super-Resolution Convolutional Neural Network (SRCNN) [15]と、スパースコーディングベースの Adjusted anchored neighbor regression (A+) [16], および Sparse-coding Super-Resolution (ScSR) [17]を利用し、これらの特性を活かすことで、従来手法 [15][16][17]の再構成精度を保持しつつ、知覚品質を向上させる手法の提案を目的としている。

そこで、本章ではまず、超解像処理の基礎技術である補間技術について簡潔に述べ、次に、超解像技術の概要および分類について記述する。続いて、学習型超解像の代表手法について記述し、提案手法に用いる SRCNN, A+, および ScSR について述べる。最後に、再構成精度の測定手法である PSNR, SSIM と、非参照画質評価モデルによる評価手法について記述する。

2.2 補間技術を用いた高解像度化技術

本節では、超解像技術の基礎技術である、補間技術について述べる。

補間技術は、デジタル画像処理において、画像の拡大、縮小や変形を行う場合や、低解像度の画像を高解像度で出力する場合に、空間的歪みを補正する用途で広く利用されてきた [19][20]。この補間技術の手法としては、バイキュービック法（三次畳み込み内挿法 [20]）、バイリニア法（双一次内挿法）、ニアレストネイバー、Lanczos 法などの補間フィルタが利用されている。

これらの補間技術は、計算コストが低いという利点も持つが、その一方で復元性能が非常に限定されたものとなる [21]。特に、高周波成分の不足により、出力画像にぼやけが生じる問題がある [10]。さらに、先鋭化を行う上で重要となる、画像の境界領域

においても、本来の輝度変化を再現できずジャギーが生じる問題があった [8]。そのため、従来の補間技術のみでは視覚的に良好な結果を取得することは困難であった。

この問題を解決するための高解像度化手法として、低解像画像を補間により拡大した後、エッジ部分を強調することで見かけ上の画質を向上させる手法も用いられてきた [22]。しかしながら、この技術のみでは原理上、低解像画像中で欠損している高周波に関する情報を復元することは不可能であるため、依然としてぼやけが解消されなかった [19][22]。加えて、エッジだけでなく画像上のノイズまで強調してしまうという問題も生じていた [22]。

そこで、「補間技術の性能限界を超えること」を目指した超解像技術が考案され、近年、さまざまな手法が考案されている [19]。

2.3 超解像技術の概要および分類

2.3.1 超解像技術

超解像技術とは、一枚、または複数枚の入力低解像度画像から、高解像度の出力画像を生成する技術である。超解像技術では、低解像度画像中で欠損している高周波成分を付加することにより、高解像度画像の取得が可能である [23][24]。

一般に、低解像度の画像を用いて高解像度の画像を推定する問題は、多数の解が、あらゆる位置の低解像度画素に対して存在するため、不良設定問題となる [15][22][25]。言い換えると、特定の解を持たない未決定性逆問題として与えられる [15]。そこで、この解法と言える数々の超解像手法が提案されてきた。

超解像技術の過去の用途は、顕微鏡、医療画像、衛星画像、天体観測、および監視映像など [17][26]であり、光学撮像システムの分解能を超えた物体の認識や解析を行うことを目的として利用されてきた [22][26]。これに加え、近年では超解像技術を符号化処理に導入する手法も検討され始めている [27][28]。これらの技術では、超解像技術を符号化技術に適用することで、情報量の削減によるストレージの削減、および高周波成分の復元による高精細な画像の取得を可能にしている。

近年の出力デバイスの解像度の拡大に伴い、今後さらに超解像技術の需要や用途が増加することが予測される。

2.3.2 超解像技術の分類

超解像技術は、複数の入力画像を利用して高解像度の画像を取得するマルチフレーム超解像技術と、1枚のみの入力画像から高解像度の画像を取得する単一画像超解像とに分類される [29]。超解像技術の分類を図 2.1 に示す。

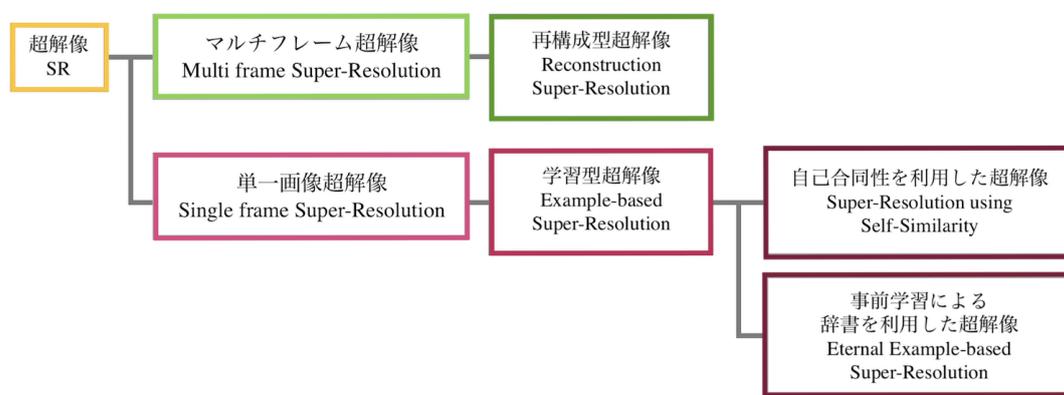


図 2.1 超解像技術の分類

マルチフレーム超解像の主な手法は再構成型超解像（Reconstruction Super-Resolution）であり，単一画像超解像の主な手法は学習型超解像（Example based Super-Resolution） [11]である。学習型超解像はさらに，自己合同性を利用した超解像技術と，事前学習による辞書を利用した超解像技術とに分類される [15]。

ここで，学習型超解像と一般にいう場合，後者の事前学習による辞書を利用した超解像技術を指すが，本章では正確な分類に基づき，自己合同性を利用した超解像技術も学習型超解像に分類している。（なお，本論文第 2 章以外で事前学習による辞書を利用した超解像技術について記述する場合は，学習型超解像と呼ぶ。）

これらの超解像処理の手法のうち，事前学習により取得した辞書を利用して超解像処理を行う学習型超解像は，特に復元精度が高く，高速な手法として知られている。

以下に再構成型超解像，および学習型超解像に関する説明と，それらに属する手法について述べる。

2.3.2.1 再構成型超解像

再構成型超解像は、複数の低解像度画像を利用して高解像度の画像を取得する超解像技術である。この手法は、カメラの分解能を超える高解像度の画像を推定、生成できる技術であり、幅広い製品に応用されてきた [26].

再構成型超解像ではまず、同じ被写体が存在する複数の低解像度画像を取得し、次に、これらの画像間におけるピクセル単位の位置合わせを行う [26]. この位置ずれには平行移動だけでなく回転や変形や大きさの変化なども含まれているため、被写体の移動は、オプティカルフローを利用した Lucas-Kanade 法やブロックマッチング法を用いて検出される [23]. 続いて、位置合わせ後の参照画像の画素値を、補間情報として処理画像に付加し、標本点の密度を増やすことで画素量を増加させ、テクスチャの細部を再現する [10][23][21].

この再構成型超解像は、動画像フレームの高解像度化技術としても適用されており [10], その概念は図 2.2 の通りである。なお、図中の画像は参考文献 [30][31] の標準画像を使用しており、以後、標準画像と言う場合はこの画像を示す。

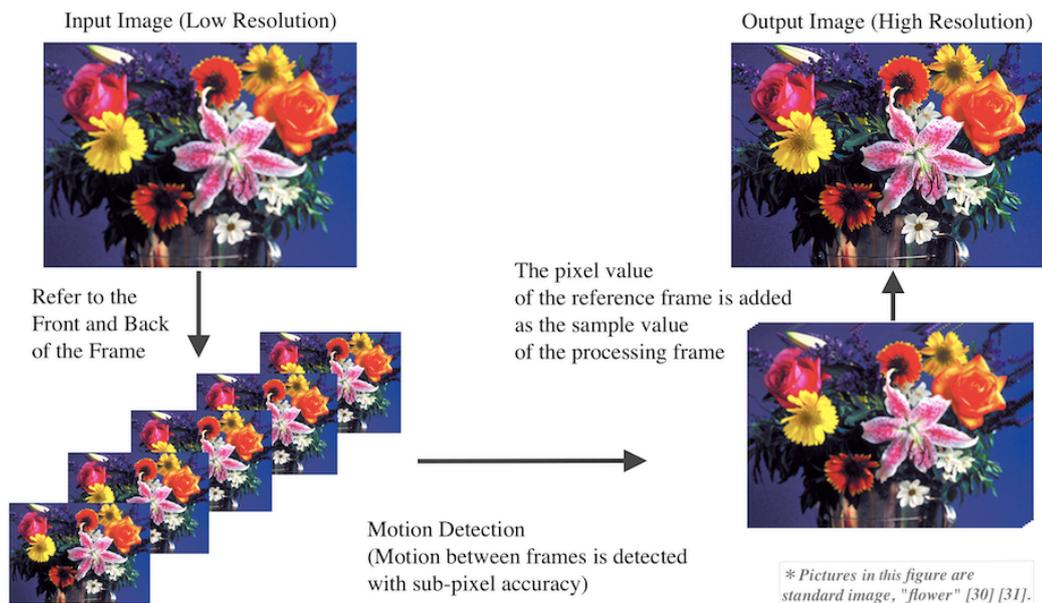


図 2.2 再構成型超解像の動画像への適用 (概念図)

図 2.2 の通り，再構成型超解像を動画像に適用する場合は，入力フレームの前後フレームを参照し，サブピクセル精度でフレーム間の動きを検出して，標本値を付加する [10][23]．この手法では，カメラか被写体のいずれかに位置ずれが存在するときに解像度の向上を図ることができる [10]．これは，処理フレーム中の被写体における輝度変化と同様の輝度変化が，参照フレームの異なる位置で標本化されることに起因する [10]．したがって，取得した複数のフレームについて，フレーム中の被写体の動作を検出し，その動作に合わせて標本点を増加させることで高解像度の画像を取得することが可能となる [10][21]．

再構成処理の手法は多数考案されている．代表的な手法には，先見情報を利用して事後確率を最大化し高解像度の画像を取得する MAP (Maximum a posteriori) 法 [32][33][34]や，最尤推定に基づく ML (Maximum likelihood) 法 [35]，元の画像の部分的な情報と事前に知られている特性から欠落したデータを復元することができる POCS (Projection onto convex sets) 法 [36][37]や，投影から画像を再構成する手法である Back-Projection 法 [38]などがある．

再構成型超解像は出力画像の破綻が少なく，解像度の向上が行いやすい [26] という利点がある．さらに，複数画像間における位置ずれを正確に検出できれば鮮鋭な画像を取得できる [10] という利点がある．特に小さな全体的な動きがある場合には，高い忠実性を補償するという特徴を持つ [21]．

一方で，位置ずれが存在する複数の入力画像が必要であり，その数が充分でない場合，このアルゴリズムの性能が急速に低下する [25]．また，複数の画像間で画素毎のサブピクセル精度の位置合わせを行う処理が必要となる．さらに，複数の画像データにアクセスするための演算処理性能やメモリー容量も必要となる [26]．また，大規模な要素や大きな動きを処理することは困難である [21]．

2.3.2.2 学習型超解像

本項では，学習型超解像の概要，および，学習型超解像に属する自己合同性を利用した超解像技術と，事前学習による辞書を利用した超解像技術について述べる．

低解像度画像を利用して高解像度画像を推定する問題は，多数の解が，あらゆる位置の低解像度画素に対して存在するため，特定の解を持たない未決定性逆問題となる [15][22][25]．しかし，この問題は一般に，「事前情報」を用いて解に制約を付

加することで軽減される [15][22][25]. この事前情報を得るために, 最先端の手法のほとんどは学習型超解像を利用している [15].

単一画像超解像の主な手法は, 学習型超解像であり, 学習型超解像はさらに, 自己合同性を利用した超解像技術と, 事前学習による辞書を利用した超解像技術とに分類される [15]. 以下に, 各技術の説明を述べる.

1.) 自己合同性を利用した超解像技術

自己合同性を利用した超解像技術とは内部例ベースの超解像技術 (**internal example-based methods**) であり, 入力画像中のパッチの類似性を利用して高解像度の画像を取得する [15]. 自己合同性を利用した超解像技術の概念図を, 図 2.3 に示す.

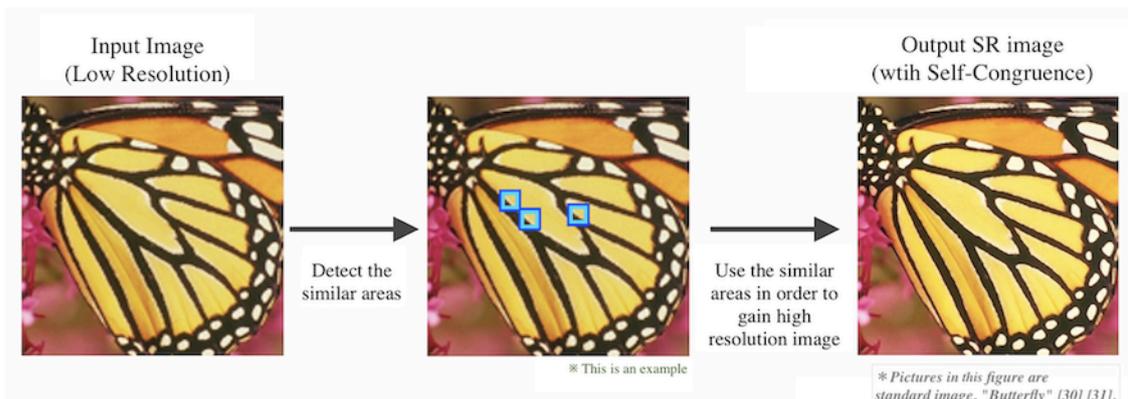


図 2.3 自己合同性を利用した超解像技術

自己合同性を利用した超解像技術では, 類似点を重畳することで画像の解像度を向上させる [15]. 自己合同性 (自己類似性, または自己相似性とも呼ばれる) とは, 「被写体輪郭などの境界領域において, エッジに沿って輝度の変化が類似したパターンが並ぶことが多い」という自然画像の性質を指す [19]. この自己合同性により, 単一自然画像内のパッチは, 同一画像中で何度も繰り返される傾向がある [40]. さらに, 自己合同性は同じスケール内だけでなく, 異なるスケールにおいても同様に観測されることが示されている [40]. したがって, これら

の反復した類似点を各領域において重畳することで、反復領域の解像度を向上させることが可能となる。

図 2.3 はパッチ選択とその重畳に関する大まかな概念のみを示したものであり、実際には入力画像とその低周波成分を利用して低解像度-高解像度のパッチペアを作成する処理が必要となる。パッチペアの作成は前処理段階で行われ、入力画像から異なる種類のソースを生成することで作成される [41]。超解像処理段階ではまず、エッジ・輪郭・テクスチャ領域などの詳細な特徴領域の検出が行われる。続いて、検出された各領域は他の領域と照合され、類似する領域が重畳される。また、他の領域においては類似する低周波成分が検出され、その差分が局所的に埋め込まれる。

自己合同性を利用した超解像技術では、入力低解像画像のみを利用して超解像処理を行うことができ、外部データベースを必要としないという利点がある。さらに、再構成型超解像よりも計算時間を大幅に短縮することができるという利点もある [41]。

その一方で、復元精度が向上する領域は自己合同性のある領域に限定され、画像中に十分な反復テクスチャが存在する領域でしか良好な性能が得られないという特徴を持つ [19][21]。

2.) 事前学習による辞書を利用した超解像技術

事前学習による辞書を利用した超解像技術は、機械学習を用いて作成した辞書や情報を利用して、高解像度の画像を取得する技術である。辞書作成時には、低解像度から高解像度への非線形マッピングが学習される [21]。この手法では、事前にマッピングが学習されているため、超解像処理時には、高解像度画像へ写像を行うだけで高解像度画像を取得できる [21][25]。

第 2.3.2 項で述べたとおり、一般的に「学習型超解像」という場合は、事前学習による辞書を利用した学習型超解像を指す[8]。(以下、学習型超解像と記述する場合、事前学習による辞書を利用した学習型超解像を指すこととする。)

学習型超解像の概念図を、図 2.4 に示す。

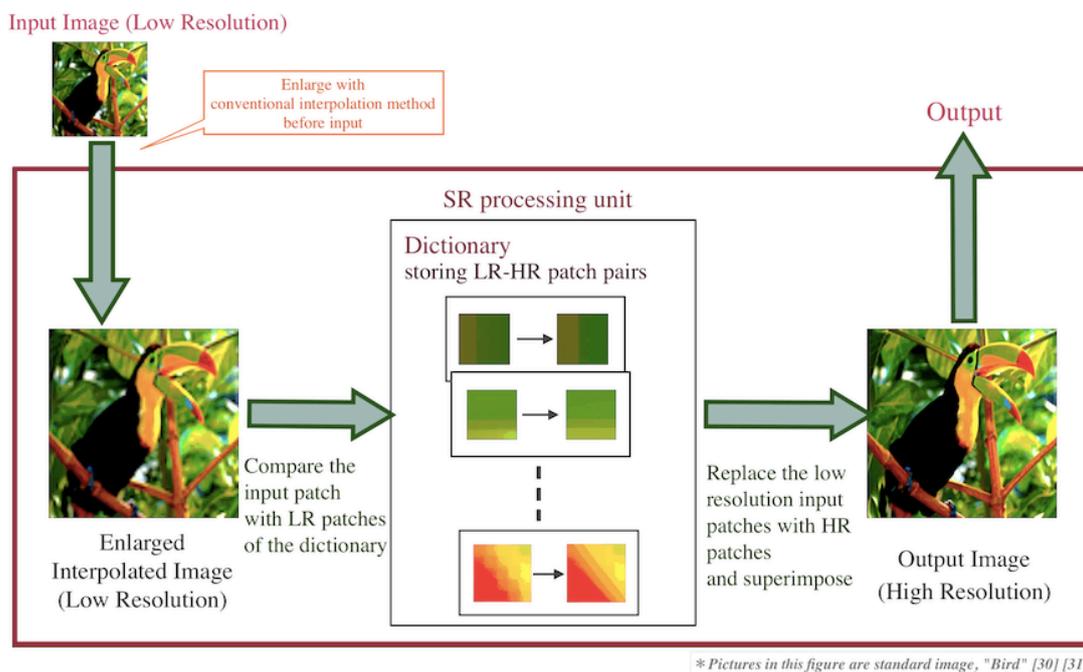


図 2.4 学習型超解像の概念図

図 2.4 に示すとおり，学習型超解像では，まず入力された低解像度画像を既存の補間技術により拡大し，この補間拡大画像をパッチに分割する [8]．超解像処理部には，事前に学習した高解像度および低解像度のパッチペアを保存した辞書を持ち，分割した各パッチについて，辞書中の低解像度パッチと照合する [26]．さらに，対応する高解像度のパッチに置き換えて，それらを重畳することによって，低解像度画像中で欠損していた高周波成分を推定する [8][25][26]．

過去の学習型超解像手法には，Locally linear embedding (LLE) [42]や，Neighbor Embedding (NE) [43][44]，スパースコーディングに基づく超解像 [17][45][46] や，Simple Function (SF) [36]，Anchored Neighborhood Regression (ANR) [48]や，Adjusted anchored neighbor regression (A+) [16]，および，ランダムフォレストを用いた超解像技術などがある．

さらに，Deep Learning を超解像技術に初めて適用した，Super-Resolution Convolutional Neural Network (SRCNN) [15]が提案されて以降，Convolutional Neural Network や Deep Neural Network を利用した，さまざまな手法が考案された．その例として，Fast Super-Resolution Convolutional Neural Networks (FSRCNN) [49]，Enhanced Deep Super-Resolution Network (EDSR) [50]や，Dual-State Recurrent Network (DSRN) [51]，Residual Channel Attention Networks (RCAN) [52]など

が挙げられる。これらの手法では、Neural Network の利用により、知覚的に尤もらしい結果を生み出す傾向がある。この他にも、近年多数の超解像技術が提案されており、今後さらに性能が向上することが見込まれる。

学習型超解像は、単一の入力画像のみで超解像処理を行うことができ、他の超解像手法と比較して、PSNR や SSIM などの再構成精度（復元精度）や知覚品質が非常に高いという特徴がある [8]。

一方で、従来の辞書利用の学習型超解像は、高い性能を得るには多様なデータベースを保存しておく必要があり、辞書サイズが増大すると処理時間も増加するという特徴があった [8][26]。しかし、この問題点は年々改善されている。例えばスパース性を利用した超解像技術では辞書サイズが小さくなるため、再構成型超解像などの複数枚の画像を扱う手法と異なり、必要メモリーを削減でき、処理速度も高速化できるという特徴がある。

ここから、これまでに述べた再構成型超解像、自己合同性を利用した超解像技術と学習型超解像とを比較する。

まず、学習型超解像と再構成型超解像とを比較する。再構成型超解像は複数の画像間に微小な動きが存在する場合においては前者を上回る可能性がある。一方、学習型超解像は単一の入力低解像度画像から高解像度の画像を予測・生成することが可能である。したがって、複数画像を利用するための演算処理性能やメモリー容量を必要としない。さらに、拡大倍率が高い場合においても高精細な画像を取得することができる。加えて、再構成型超解像では画質が低下する「画像間で大きな動きが存在する場合」においても、学習型超解像では高精度な解像度拡大を行うことが可能である。よって、総合的にみて、事前学習による辞書を利用した学習型超解像の方が有効であると考えられる。

また、学習型超解像と自己合同性を利用した超解像技術とを比較すると、後者は外部データベースを必要とせず、十分な反復テクスチャがある領域においてのみ前者よりも性能が良い可能性がある。しかし、学習型超解像は、画像の一部のみならず、画像全体において画質の向上を図ることが可能であるため、より自然な画像を得ることができ、再構成精度や知覚品質も高くなる。また、自己合同性を利用した超解像は、外部データベースは必要としないものの、低解像度画像「入力後」に、入力画像とその低周波成分からパッチペア辞書を作成する処理が必要となる。そのため、事前に用意した辞書を利用する後者の処理の方が、効率的かつ実用的であると言える。

以上をとおして、超解像技術のうち学習型超解像は、効率的であり、画像中で

画質の変化が生じず，特に PSNR や SSIM 等の再現精度が高いことがわかる．さらに，学習型超解像では，失われた高周波成分の細部を人間的に尤もらしい表現で再現できるため，客観的評価指標だけでなく，視覚的にもはっきりとした良好な画像を取得することが可能である．

このように，学習型超解像は，これまでの超解像技術の能力を圧倒するほどの性能がゆえ，近年盛んに研究されている．

2.4 提案手法に使用する従来手法

前節の通り，超解像技術の中で最も注目を浴びているのは，学習型超解像である．

本研究では，学習型超解像のうち，畳み込みニューラルネットワークを利用した Super-Resolution Convolutional Neural Network (SRCNN) [15]と，スパースコーディングを利用した A+(Adjusted anchored neighbor regression) [16], および ScSR (Sparse-coding Super-Resolution) [17]を使用する．これらの手法はどれも，再構成精度が高く，高速な処理速度で実行可能な手法として知られている．

本節では，これらの手法について説明する．

2.4.1 Super-Resolution Convolutional Neural Network (SRCNN)

本項では参考文献 [15]の手法についてまとめる．

SRCNN は畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) を利用した超解像技術であり，CNN は低解像度-高解像度間のエンドエンド間のマッピングを学習している．この手法は，学習時に，部分空間のモデル化を行うための辞書や多面性を，CNN の中間層を通して学習する．

SRCNN では，超解像時に，入力低解像度画像を事前処理としてバイキュービック補間で所望のサイズに拡大し，それ以降の処理は全て超解像処理部の CNN を通して行われる．この CNN 内で，パッチ抽出と重畳を含む，低解像度から高解像度画像へのマッピングが行われる．

C. Dong らは実験から，超解像処理に利用する CNN においては，層数が深いほど

性能が良くなるわけではないことを示しており、4層や5層まで進むと、収束を保証する適切な学習率を設定することが困難となるため、3層のネットワークを利用している。3層のCNNの構成を図2.5に示す。

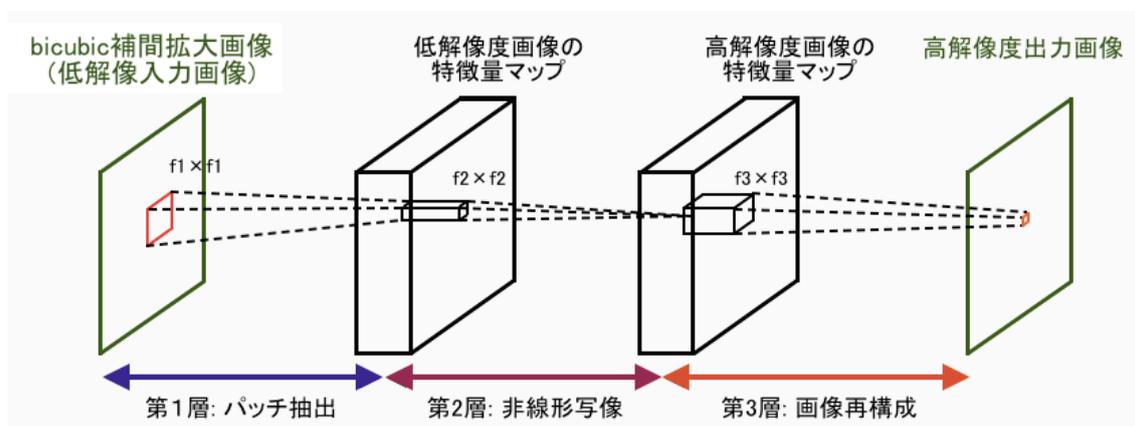


図 2.5 SRCNN のネットワークの構成図

なお、 f_1 , f_2 , f_3 はフィルタの空間サイズを示し、標準設定値として、 $f_1 = 9$, $f_2 = 1$, $f_3 = 5$ を利用している。この CNN における全体的な処理について述べる。

補間拡大画像は、CNN の第 1 層に入力され、オーバーラップしたパッチが抽出される。各パッチは高次元ベクトルとして表され、一連の特徴マップが抽出される。フィルタ応答としては Rectified Linear Unit (ReLU) [53] を利用する。続いて、第 2 層では、第 1 層で出力された n_1 次元ベクトルから n_2 次元ベクトルへの非線形写像が行われ、高解像度のパッチへのマッピングが行われる。 n_1 , n_2 は実験から、増加させるほど優れた性能を示すが、修復時間が長くなるため、64, 32 と設定されている。第 2 層で出力された n_2 次元ベクトルのそれぞれが、再構成処理に使用される高解像度パッチを示し、第 3 層ではこれらのパッチの再構成処理を行うことで最終的な高解像度画像が出力される。ここで、第 2 層にて写像された高解像度のパッチは異なるセットの特徴マップを含むため、第 3 層において、オーバーラップした領域が平均化され、隣接する高解像度予測パッチが合成される。

SRCNN は、CNN 訓練後、超解像処理はフィードフォワード処理で行われる [54]。したがって、使用時には最適化問題を解く必要がなく、従来の超解像手法よりもはるかに高速に動作する。また、SRCNN はシンプルな構造で設計されているにも関わらず、優れた性能を持つ。さらに、フィルタ数および層数が適切に設定されているため、オンラインでも利用可能な実用的な処理速度を示す。

2.4.2 Sparse-coding Super-Resolution (ScSR)

本項では参考文献 [17]の手法についてまとめる。(J. Yang らは一般画像用と顔画像用とで別のアプローチを取っているが、ここでは前者の一般画像の ScSR について説明する.)

ScSR はスパースコーディングに基づいた超解像技術であり、「高解像度の信号における基底の線形結合の関係が低解像度の射影から正確に復元される」というスパース信号表現の研究結果 [55][56]に基づいた手法である。ここで、基底とは画像構造を抽象化した成分であり、スパースコーディングとは、入力された低解像度パッチを辞書中に保存された複数の基底の線形結合で近似するためのスパース係数を決定する問題である [8][57].

従来の辞書利用の学習型超解像では、高解像度画像と低解像度画像からサンプリングしたパッチペア (LR-HR パッチペア) についての直接的なマッピングを学習するところ、ScSR では、スパースコーディングを利用する。これにより、ScSR では直接的な写像の学習を行う必要がなく、超解像処理に必要なメモリー容量、およびアルゴリズムの速度が大幅に改善される。

ScSR は、高解像度パッチ用の辞書 D_h と低解像度パッチ用の辞書 D_l 、という同じスパース表現を持つ連結した2つの辞書を使用しており、これらの辞書は、自然画像からランダムにサンプリングされた10万種類の画像のパッチペアから訓練されている。この2つの辞書の訓練は共同で行われ、これによって高解像度-低解像度パッチペア間のスパース表現の類似性をそれぞれの辞書に強制している。ScSR の辞書のサイズは1024であり、J. Yang らは256から2048の4種の辞書サイズを用いた実験結果から、辞書サイズ1024において、処理時間と精度の双方で良好な結果が得られることを導いている。

ScSR の構成図を、図 2.6 に示す。

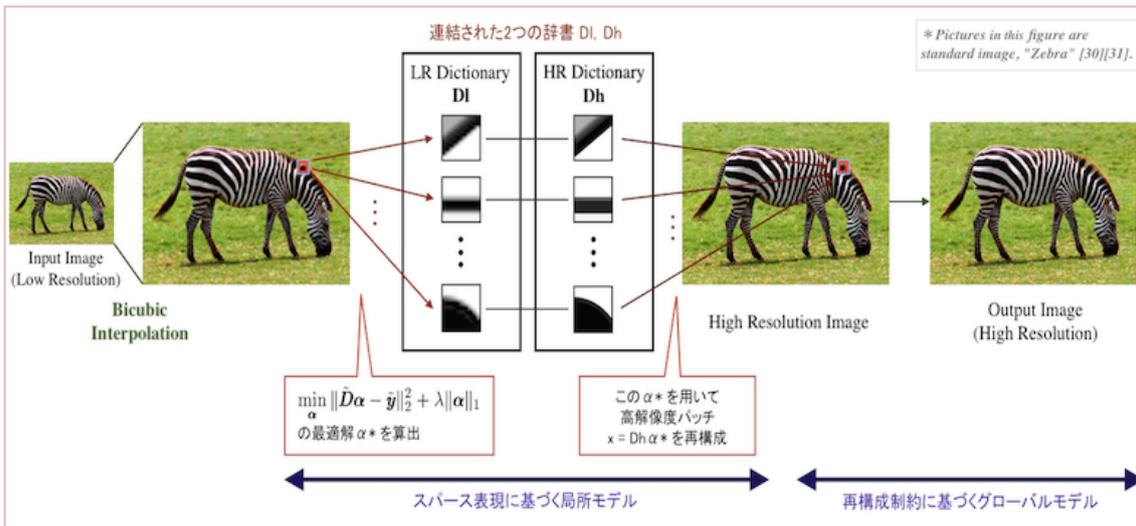


図 2.6 ScSR の構成図 (概念図)

ScSR の超解像処理段階ではまず、バイキュービック補間を利用して取得した補間拡大画像が、1 ピクセルのオーバーラップを含む 3×3 のパッチに分割される。辞書には、画素値の絶対強度ではなく、各パッチの平均ピクセル値を減算した画像のテクスチャが登録されているため、始めに低解像度パッチの平均画素値が算出される。続いて、各パッチについて、隣接パッチ間の空間的互換性を考慮したスパース表現が求められる。このとき、高周波成分を回復するための局所的な細部情報を表現するために、スパース事前分布の局所モデル (local model from the sparsity prior) が使用される。スパース事前分布の局所モデルとは、「高解像度パッチが、適切に選択された過完備辞書においてスパースに表現され、そのスパース表現が低解像度パッチから回復できる」と仮定するモデルのことである。このモデルに基づいて求めたスパース係数は、高解像度パッチを生成するために使用される。

したがって、低解像度パッチがスパース表現によって基底の線形結合で示された後、各基底は連結された 2 つの辞書によって、対応する高解像度の基底に置き換えられる。次に、先で算出したスパース係数を直接使用して基底が結合され、高解像度出力パッチが生成される。続いて、出力された高解像度パッチは、上で計算した平均画素値と加算されて、テクスチャ成分から画素値の絶対強度へと変換される。これによって、局所的な各高解像度パッチの生成が完了する。

ここで、低解像度パッチと辞書 D_l の基底の線形結合との差異やノイズが原因で、再構成制約条件を厳密に満たさない場合がある。再構成制約とは、「高解像度画像に

ぼかしフィルタとダウンサンプリング関数を掛けた劣化画像と、入力低解像度画像が同等であること」を要求する制約である。そこで、ScSR では最後に再構成制約のグローバルモデルを適用することで、アーティファクトが除去され、画像が全体的に自然になる。この計算は、最急降下法を用いて効率的に行われる。

一般に、事前学習辞書利用の学習型超解像は、辞書を作成する際の訓練データに実験結果が依存すると考えられており、学習されていない未知データを入力した場合に適合できない状態が生じると考えられている（これを過剰適合という）。しかし、スパースコーディングを利用した超解像技術では、低解像パッチを基底に分解してから超解像処理を行うため、性能が入力画像の種類に依存しにくいという特徴がある。したがって、訓練データ以外の未知データに対しても適合し、訓練データへの過剰適合を減らすことができる [16]。

また、スパース表現は雑音に対してロバストである。そのため、他の多くの超解像手法では雑音除去と SR を同時に行うことができないところ、ScSR のアルゴリズムでは低解像度画像のノイズに対してより頑強に処理することが可能である。さらに、ScSR では、従来の辞書利用の超解像技術と異なり、大きなパッチデータベースを必要とせず、前述のコンパクトな 2 種のパッチペア表現を学習した辞書のみを利用する。したがって、計算コストを大幅に削減できるという利点がある。

2.4.3 Adjusted anchored neighbor regression (A+)

本項では参考文献 [16]の手法についてまとめる。

A+ は、スパース辞書と、近傍埋め込み法：Neighbor Embedding (NE) に基づく手法である。スパース辞書は、2.4.1.2 項で記述した通り、訓練データの過剰適合を減らすことができる。さらに、大きなパッチデータベースを必要とせず、計算コストも少ない手法である。また、近傍埋め込み法は、「全ての低解像パッチと高解像パッチが、局所的に類似した形状を有する、低次元の非線形マニホールド上に存在する」と仮定する。この仮定により、入力パッチを既存のデータベースパッチの補間と近似することができる。よって、1 組の補間係数が低解像度空間、および高解像度空間の双方に適用され、この係数を利用して、対応する HR パッチの線形補間を実行できる。

これらの手法はどちらも辞書ベースの手法であり、パッチの辞書やパッチに基づく atom の辞書 (Dictionary of patch-based atoms) に依存して超解像処理を行う。そのため、辞書の学習段階において、自然なパッチを効率的に表現するように訓練すること

が可能であり，時間計算量を大幅に削減することができる．

また，A+ は，超解像手法としては，ANR (Anchored Neighborhood Regression) [47] と SF (Simple Function) [48] とを合成した手法である．ANR はスパース辞書を学習し，辞書の atom に固定された独立変数を学習する手法である．また，SF は低解像パッチ空間をクラスタ化して各クラスタの低解像から高解像へのマッピングを学習する手法である．これらの手法は，超解像出力画像の品質を犠牲にすることなく，単一画像超解像の時間計算量を大幅に低減させる．

A+は，これらの手法を組み合わせることで，効率的な辞書ベースの超解像処理を行う．A+ は ANR の特徴と固定独立変数をベースにしているが，辞書の独立変数を学習する代わりに SF と同様の完全な訓練資料を使用する．また，ANR の重要な側面である，特徴量やテスト時間計算量などを継承して調整すると同時に，性能を大幅に向上させる．

2.5 画質評価指標

本節では，再構成精度による画質評価指標 (PSNR および SSIM [12]) と，非参照画質評価モデルによる画質評価手法 (Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)) について記述し，超解像技術に対する特徴をまとめる．

2.5.1 PSNR

PSNR は，画像および映像の劣化程度を定量的に表現するための客観的評価尺度の 1 つであり，MSE (Mean Square Error) とピーク値から算出可能である [58]．

以下に，参考文献 [58][59] をもとに，8 bit 画像に対する PSNR の導出式を示す．

輝度信号が 8 bit で表現されるとき，画像信号の最大値は 255 である．評価の基準となる画像 (原画像) の画素を $O(i, j)$ ，評価対象となる画像 (超解像画像) の画素を $R(i, j)$ とし，画像サイズを $M \times N$ とするとき，PSNR は以下のように定義される [58][59]．

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (2.1)$$

$$MSE = \frac{1}{M \times N} \sum_i^M \sum_j^N [O(i,j) - R(i,j)]^2 \quad (2.2)$$

ここで、MSE は、原画像と比較画像（超解像画像）の差分二乗平均を示す。

PSNR が常用対数を用いてデシベル (dB) 単位で表現されるのは、多くの信号においてダイナミックレンジが広いことを考慮したものである [58]。上記の対数化は、人間の視覚は Weber Fechner の法則に従うことから、主観品質と PSNR とがより一致する結果をもたらしている [58]。特に、PSNR の高い画像において、PSNR と主観品質が高い相関関係にあることが知られている [58]。また、PSNR は複雑な計算は不要なため、客観的評価尺度として、広く使用されている [58]。

その一方で、PSNR は参照画像に対する誤差とピーク値のみに基づいて計算され、人間の視覚特性を考慮した画質評価指標ではないため、知覚品質と合致しない場合がある。例えば、全体的にぼやけているが各画素の誤差はわずかである場合と、原画像にほぼ等しくはっきりしているが数画素の画素値が大幅に異なる場合について、PSNR では近い値となり得る。また、PSNR の低い画像や画像中で画質に偏りが存在する場合、知覚品質との間に乖離が生じ得ることが知られている [58]。

2.5.2 SSIM

SSIM [12]は、画像の構造情報の類似度に基づき画質を評価する客観的評価尺度である [58][60]。SSIM 値は、ピアソンの相関係数を用いた構造的類似度項に、輝度比較項、およびコントラスト比較項を乗算することで計算される [12][58]。

以下に、参考文献 [12][58][60]をもとに、SSIM の導出式を示す。

μ_x , σ_x , および σ_{xy} をそれぞれ、 11×11 画素の正規化された円対称ガウス関数により重み付けられた $x[i]$ の 平均値, 標準偏差, および $x[i] y[i]$ の共分散とし, μ_y および σ_y を, 11×11 画素の正規化された円対称ガウス関数により重み付けられた $y[i]$ の 平均値, および標準偏差とすると, 原画像, および比較画像中の小領域 (ウィンド

ウ) $x[i], y[i]$ に対する SSIM は, 式 2.3 から計算される [12][58].

$$SSIM(x, y) = \frac{(2\mu_x \times \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 \times \mu_y^2 + C_1)(\sigma_x^2 \times \sigma_y^2 + C_1)} \quad (2.3)$$

ここで, 定数 K_1, K_2 が 1 より非常に小さく, L を画素値のピーク値とするとき (8bit 画像の場合 $L = 255$), C_1, C_2 は, 式 2.4, および 式 2.5 から得られる [12][58].

$$C_1 = (K_1 L)^2 \quad (2.4)$$

$$C_2 = (K_2 L)^2 \quad (2.5)$$

文献 [12]では, 定数 K_1, K_2 は, $K_1 = 0.01, K_2 = 0.03$ が適するとされている [12].

ここで, 式 2.3 により得た SSIM は, 画像中のウィンドウに対する評価値であり, 画像全体に対する画質評価値は, それらの平均値から得られる (式 2.6) [58].

$$MSSIM(x, y) = \frac{1}{M} \sum_{j=1}^M SSIM(x, y) \quad (2.6)$$

SSIM は, PSNR よりも知覚品質に相関する場合があることが報告されている [58].

その一方で, PSNR と同様, 参照画像 (原画像) が必要であり, SSIM は この原画像との差異に基づいて画質評価を行う. そのため, 比較画像が原画像とは異なる構造情報を持つが, 知覚的には高画質であるとき, SSIM と知覚品質とが一致しない結果となる場合がある.

2.5.3 BRISQUE

BRISQUE [14]とは, 画質評価に参照画像 (原画像) を必要としない, 非参照画質評価 (No-Reference Image Quality Assessment (NRIQA)) モデルである.

BRISQUE は, リンギング, ブラー, ブロッキングといった歪み特徴量を計算せず, 代わりに, 局所的に正規化された輝度係数のシーン統計を使用する [14]. これによ

り、歪みによる画像の自然さの損失が定量化され、品質の全体的な尺度が導かれる [14]. BRISQUE で使用される基本的な特徴量は、空間的な自然シーン統計モデルの下での局所的に正規化された輝度、およびその積の経験的分布から導き出される [14].

以下に、参考文献 [14]による、BRISQUE の算出方法を示す.

BRISQUE では、与えられた強度画像 $I(i, j)$ に対して、はじめに、局所平均減算と除算正規化を用いて、局所正規化輝度を計算する [14][61]. これは、Ruderman [61] の観察結果、「対数コントラストのゼロから局所平均変位を除去して局所分散を正規化するために、対数コントラスト輝度に局所非線形操作を適用すると、非相関効果がある」に基づくものである.

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C} \quad (2.7)$$

ここで、 $I(i, j)$ は画像の行列 (i, j) 成分における輝度値、 $i \in 1, 2, \dots, M$, $j \in 1, 2, \dots, N$ は画像サイズ $M \times N$ に対する空間インデックス、 $C = 1$ は分母がゼロになる傾向がある際の不安定性を防ぐための定数項である. また、 $\mu(i, j)$ および $\sigma(i, j)$ は式 2.8, および 式 2.9 により計算される.

$$\mu(i, j) = \sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} I_{k,l}(i, j) \quad (2.8)$$

$$\sigma(i, j) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} \left(I_{k,l}(i, j) - \mu(i, j) \right)^2} \quad (2.9)$$

$w = \{w_{k,l} \mid k = -K, \dots, K, l = -L, \dots, L\}$ は、局所領域 $(K \times L)$ における円対称ガウス重み関数であり、実装では、 $K = L = 3$ である.

BRISQUE では、上記の計算により得られた $\hat{I}(i, j)$ を平均減算コントラスト正規化係数 (MSCN 係数) と呼び、MSCN 係数は自然画像の単位正規ガウス特性に強く従う [14][61]. そこで、MSCN 係数は人間の初期の視覚におけるコントラストゲインマスキングプロセスをモデル化するために使用される [14][61][62][63]. BRISQUE の考案者である Mittal らは、「MSCN 係数には歪みの存在によって変化する特徴的な統

計特性があるため、これらの変化を定量化することで、画像に影響する歪みの種類、およびその知覚品質を予測できるようになる」としている [61].

次に、「水平 (H), 垂直 (V), 主対角 (D1), 二次対角 (D2)」の 4 つの方向に沿った、隣接 MSCN 係数のペアワイズ積の経験的分布を使用して、隣接するピクセル間の統計的関係をモデル化する. ここで、隣接 MSCN 係数のペアワイズ積は下式により算出される.

$$H(i, j) = \hat{I}(i, j) \hat{I}(i, j + 1) \quad (2.10)$$

$$V(i, j) = \hat{I}(i, j) \hat{I}(i + 1, j) \quad (2.11)$$

$$D1(i, j) = \hat{I}(i, j) \hat{I}(i + 1, j + 1) \quad (2.12)$$

$$D2(i, j) = \hat{I}(i, j) \hat{I}(i + 1, j - 1) \quad (2.13)$$

これらの隣接 MSCN 係数のペア積は、非対称一般化ガウス分布 *Asymmetric Generalized Gaussian Distribution (AGGD)* に適合され、文献 [64] のモーメントマッチングベースの手法により統計モデルのパラメータが推定される [14]. BRISQUE では、MSCN 係数の確率分布から 36 次元 (各スケール 18 次元) の特徴量が作成される [14]. 文献 [14] から、これらの統計的特徴量は人間の知覚品質とよく相関することが示されている. 最後に、特徴から品質空間へのマッピングが学習され、知覚品質の自動非参照測定手法が生成される [14].

BRISQUE は、PSNR および SSIM よりも統計的に優れていることが、文献 [14] の実験結果より示されている. さらに、BRISQUE モデルの学習には、主観的オピニオンスコアが使用されることから、BRISQUE 値は人間の感知する画質と相関するという特徴がある [65].

2.5.4 各種画質評価指標の比較

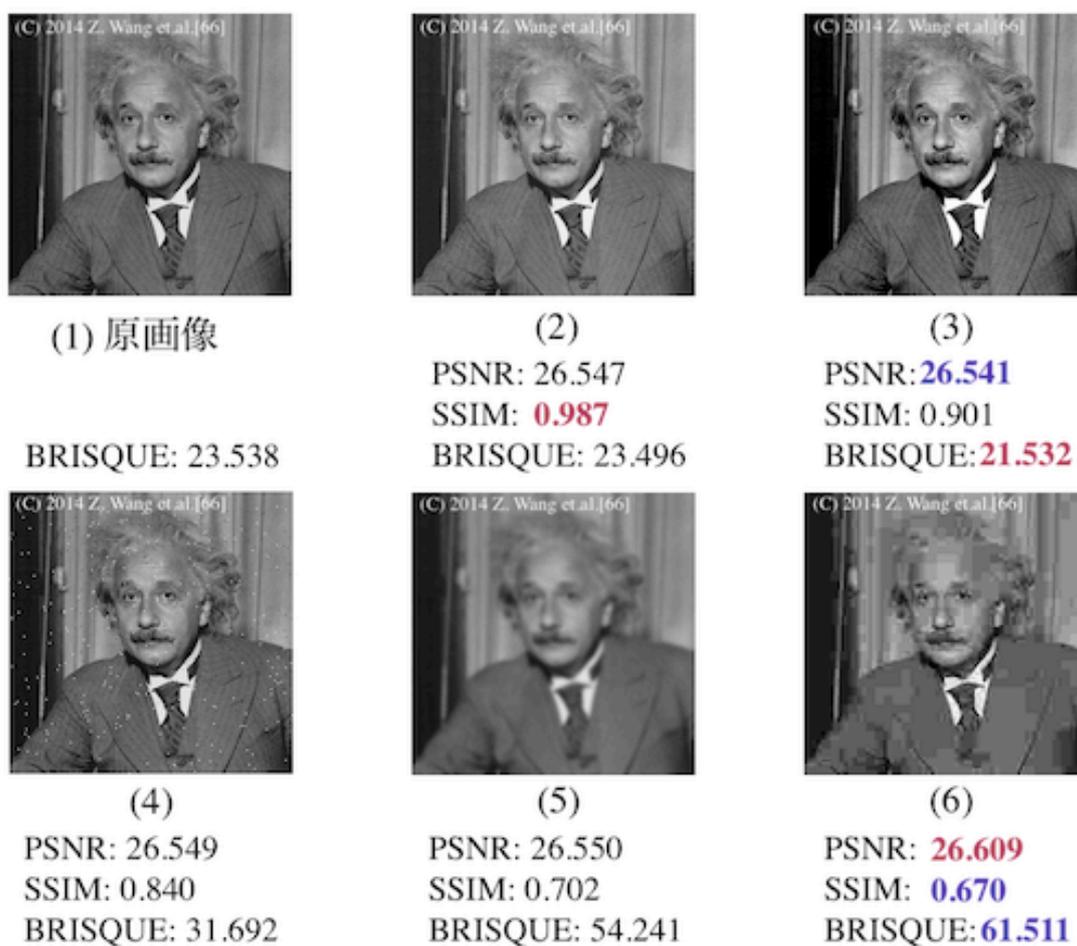
第 1 章で述べたように、従来、超解像技術は、PSNR や SSIM [12] といった再構成精度による客観的画質評価値を向上させることを目的に行われてきた.

一方で、近年の研究結果から、再構成精度による画質評価と知覚品質とが比例しない場合があることが示されている [13][14][60]. この研究報告によると、優れた知覚品質を持つ超解像手法は PSNR や SSIM のような歪み測定では評価されない傾向があ

るとされている [13].

そこで、本項では、文献 [66]から取得した画像について PSNR 値、SSIM 値、および BRISQUE 値を計算し、それぞれの特徴を比較する。

実験結果を図 2.7 に示す。



(C) 2014 Z. Wang et.al. / 文献 [66]より画像引用

図 2.7 画質評価指標 PSNR, SSIM, BRISQUE による画質評価値の比較

ここで、PSNR 値、および SSIM 値は高いほど高画質であることを示すのに対し、BRISQUE 値は低い値であるほど高画質であることを示すことに注意されたい。

実験結果をより比較しやすくするために、PSNR, SSIM, BRISQUE, 主観品質による順位の比較表を表 2.1 に示す.

表 2.1 順位比較

Image \ IQA	(2) Luminous	(3) High contrast	(4) Dots	(5) Blurs	(6) Block Noise
PSNR	4	5	3	2	1
SSIM	1	2	3	4	5
BRISQUE	2	1	3	4	5
主観品質	2	1	3	4	5

実験結果より、PSNR では、視覚的に不快な画像 (5) および (6) を上位 2 位にしている。これらの画像はそれぞれ、(5) : 全体的にぼやけているが画素値の誤差が少ない画像、および (6) : ブロックノイズが生じているが画素値の誤差が少ない画像、である。PSNR は、参照画像に対する誤差およびピーク値 (8bit 画像のため 255) のみを用いて算出されるため、画素値の誤差の少ないこれらの画像において、他の画像よりも高い評価値を算出してしまいう結果となった。さらに、PSNR は、主観品質で最も順位の高かった画像 (2) の順位を最下位に順位づけており、人間の知覚的に明瞭な画像であっても、原画像との画素誤差が大きければ「低画質」と判断されてしまうことがわかる。

このように、PSNR は、元の画像 (原画像) に対する誤差を測るには、計算複雑度も低く、簡便な手法であるが、人間の視覚特性を考慮した手法ではないため、知覚品質とは必ずしも一致しないことが確認できる。

続いて、SSIM の実験結果について確認する。SSIM では、主観品質に比較的近い結果が取得できており、文献 [58] の「SSIM は主観品質との相関の観点で、PSNR より優れる場合がある」ことを確認することができた。一方、SSIM の上位 2 種は、主観品質の上位 2 種とは反対であった。これは、SSIM が PSNR と同様、参照画像に対する比較をもとに画質評価を行うためである。SSIM は、原画像との画像構造上の差異に基づいて画質を評価するため、比較画像が知覚的には高画質であっても、原画像とは異なる構造情報を持つ場合 (今回の場合はコントラストにおける構造情報が原画像と大きく異なったため)、SSIM と知覚品質とが一致しない結果となり得ることが確認できた。

最後に、BRISQUEの実験結果について確認する。BRISQUEによる順位は主観品質の順位と全て一致した。さらに、BRISQUEでは、視覚的に大きな変化のない画像(1)から(3)間ではBRISQUE値の変化の程度が小さく、全体的に不明瞭でぼやけた画像(5)で大幅にBRISQUE値が増加し、さらにブロック歪が多く視覚的に不快な画像(6)でよりBRISQUE値が増加(低画質と評価)する、という知覚的な画質評価に近い結果となった。さらに、BRISQUEでは値のレンジが広いので、画質評価地の変化が直感的によりわかりやすいと感じた。

以上の結果から、PSNRによる画質評価は、「原画像との誤差」という復元精度を測るには十分であるが、誤差値のみで画質が評価されることから、主観品質にそぐわない場合があることが確認できた。また、SSIMの結果は比較的主観品質に近い値となることが確認されたが、原画像に対する画像構造特徴量の類似性に基づいて画質を評価することから、視覚的に明瞭な結果であっても画像構造が異なれば画質を低く評価する場合もあることが確認できた。最後に、BRISQUEによる画質評価は、主観的な画質評価に近く、定量的な知覚品質の評価指標として適した手法であることが確認できた。

2.5.5 画質評価指標と評価実験の関係

本項では、画質評価指標と本研究の評価実験との関係について述べる。

2.5.4項の実験結果を踏まえ、高解像度化技術の考案において「画質の向上のみ」を意識する場合は、再構成精度であるPSNRやSSIM以上に、BRISQUEによる知覚品質を改善することが重要であるといえる。

一方、PSNRやSSIMはどれだけ正しい画像を生成できるかを判断するために簡便な手法であり、正解画像から乖離した結果を生み出さないためには重要な判断基準である。そのため、本研究では、PSNRを超解像処理が正解画像に対する誤差値がどれだけ小さいかを判断するために、SSIMを、画像構造特徴がどれだけ近いかを判断するために使用する。

さらに、BRISQUEは主観品質に比較的忠実に画質評価を行うことから、主観品質に近い客観的評価値として、画質の向上の程度を測定するために本研究に使用する。

本研究では、従来手法であるA+, ScSR, SRCNNに対し、再構成精度の低下をで

きる限り抑制し，知覚品質を大幅に改善することを目指す．

2.6 むすび

本章では，超解像技術について，その基礎技術である補間技術について説明し，超解像技術の概要と分類について述べた．さらに，各分類の手法と特徴について記述し，本研究に使用する SRCNN, A+, および ScSR について述べた．

続いて，画質評価指標として，PSNR, SSIM, および BRISQUE について説明・比較し，超解像技術に対する特徴を述べた．最後に，本研究で上記の画質評価値を使用する目的について述べた．

第3章 提案手法

3.1 まえがき

本章では，提案手法の概要と構成について述べる．

まず，提案手法の概要について述べ，提案手法を構成するマルチステップ超解像技術 [67] と，方向性を考慮した超解像技術 [18] について説明する．続いて，これらの手法を合成した提案手法について述べる．

3.2 提案手法の概要と構成技術

前述のように，PSNR や SSIM 等の再構成精度は，知覚的な画質と必ずしも一致しないことが知られており，優れた知覚品質を持つ手法であっても，これらの歪み測定手法では評価されない傾向がある [13]．そのため，高画質化技術における知覚的な画質改善のみを目標とする場合，主観品質評価に近い BRISQUE 等の IQA 値の向上を測ることが望ましい．その一方で，PSNR や SSIM は，原画像に対する画素値，または構造特徴量に対する正確さを示す指標であることから，原画像からの乖離を防ぐためには重要な指標であるといえる．

そこで，提案手法では，学習型超解像の特性を活かしたマルチステップ超解像技術 (Multi-Step Super-Resolution : MSSR) [67] を用いて，まず知覚品質を向上させ，BRISQUE 値の向上を図る．さらに，「方向性を考慮した超解像技術 (回転・反転を伴う超解像)」 [18] を マルチステップ超解像技術に適用することで，再構成精度 (PSNR および SSIM) を改善し，知覚品質を向上と再構成精度の双方を向上させることを目指す．

従来手法 [18][67] は，我々が過去の研究において提案した手法であり，以下にその説明を述べる．

3.2.1 マルチステップ超解像技術（多段超解像技術）

マルチステップ超解像技術は、学習型超解像における特性を利用することで、超解像処理部を多段化した手法である。学習型超解像の特性とは、「学習型超解像の超解像処理部では、低解像度から高解像度への同サイズのマッピングが行われる」という性質のことである。図 3.1 に、再度、学習型超解像の概念図を示す。

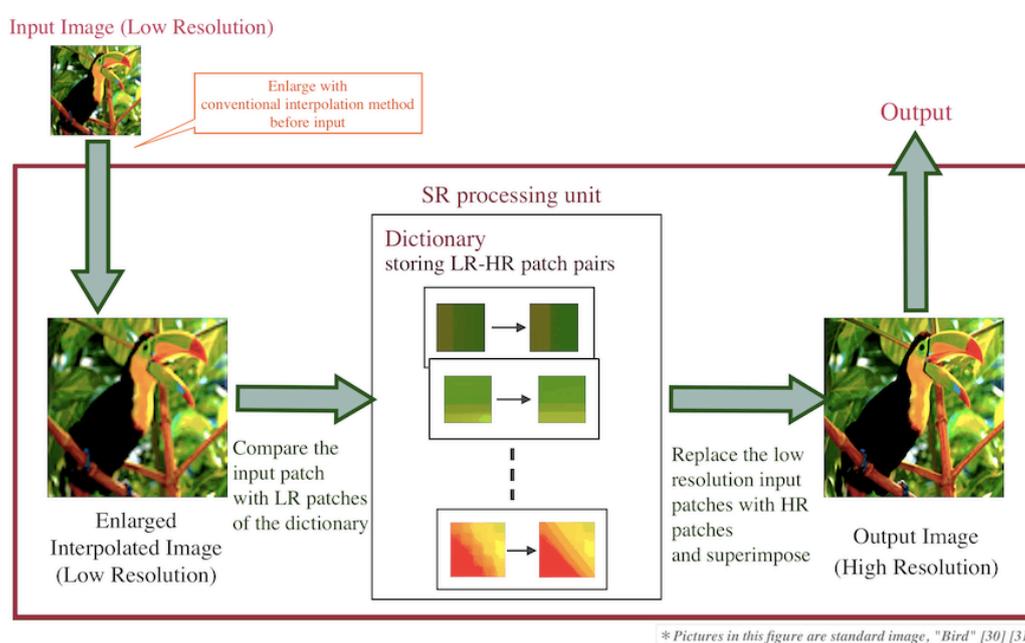
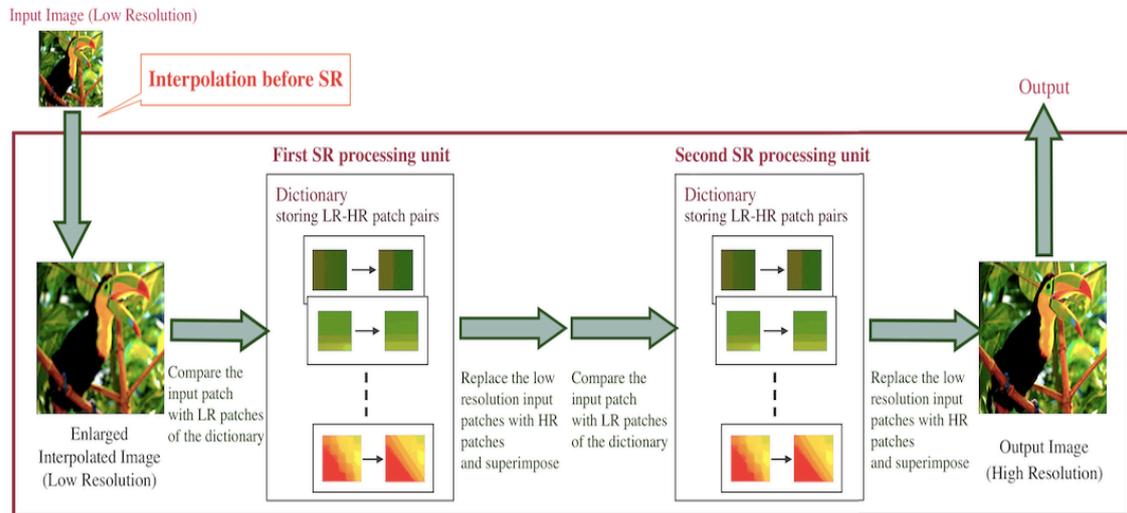


図 3.1 学習型超解像の概念図（再掲）

図 3.1 に示すとおり、一般に、学習型超解像では、事前処理として **Bicubic** 法などを用いた補間拡大処理が行われ、超解像処理部では同サイズ間の変換が実行される。そのため、「学習型超解像は、補間拡大処理を除く超解像処理部のみを見れば、擬似的に高画質化フィルタとみなせる」という仮説を立てることが可能である。マルチステップ超解像技術では、この仮説に基づき、通常の超解像処理後に、補間拡大処理を除く超解像処理部（すなわち高画質化変換処理）を直列的に結合する。

図 3.2 にマルチステップ超解像技術の概念図を示す。

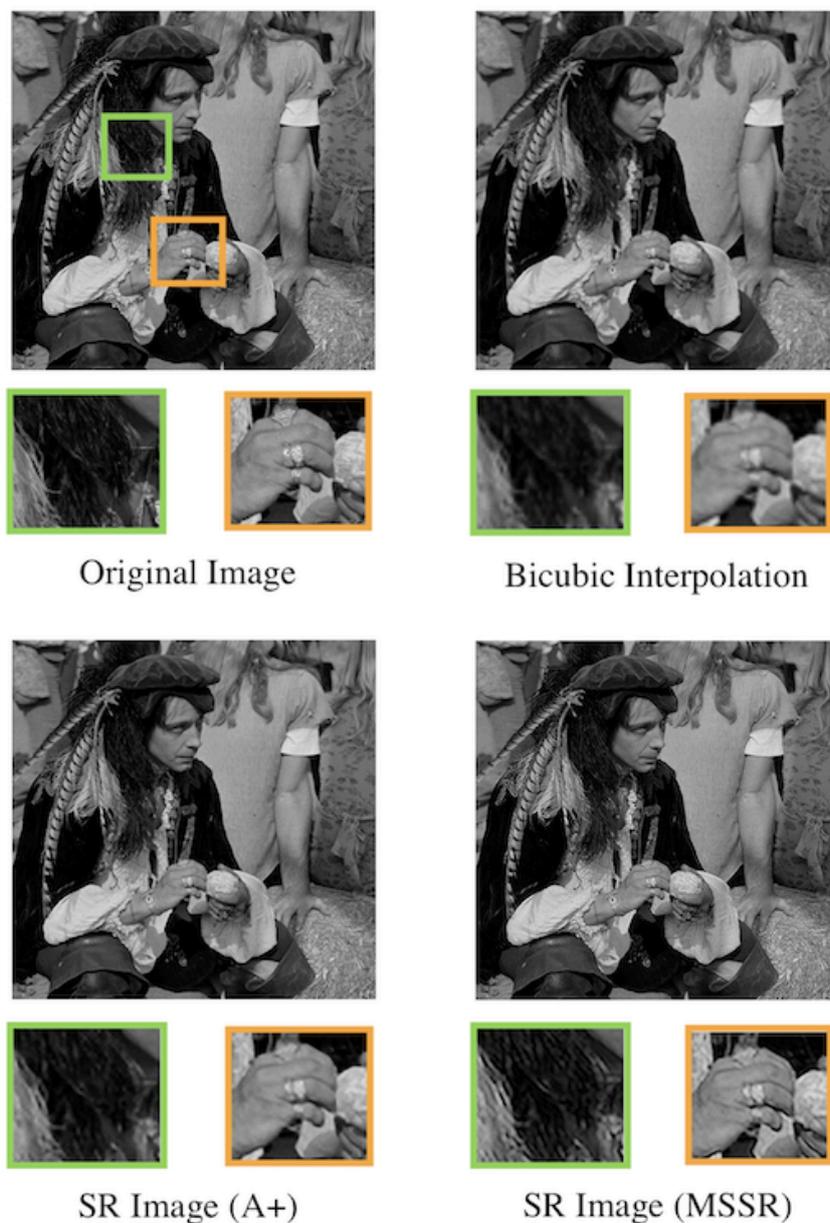


* Pictures in this figure are standard image, "Bird" [30][31].

図 3.2 マルチステップ超解像技術の概念図

図 3.2 に示すとおり，マルチステップ超解像技術では，Bicubic 補間拡大処理を除く 2 段の超解像処理部を連結する．過去の研究では，この 1 段目の超解像処理手法として A+，または ScSR を使用し，2 段目の超解像処理手法として，Bicubic 補間を除く SRCNN を使用している [67]．

通常，SRCNN のプログラム中では，輝度値 $[0, 255]$ を $[0, 1]$ に正規化する処理が実行される．一方，標準画像 Set 5，Set 14 [30][31] を用いた実験結果より，マルチステップ超解像技術の 2 段目に SRCNN を適用する場合には，この正規化を行わないことで再構成精度が向上することを確認した．さらに，マルチステップ超解像技術の評価実験から，本手法では視覚的に大幅に画質が改善されることを確認した．図 3.3 に，マルチステップ超解像技術の 1 段目に A+ を使用し，標準画像 *Man* を使用した場合の超解像結果の比較を示す．



** Pictures in this figure are standard image, "Man" [30][31].*

図 3.3 マルチステップ超解像技術による実験結果 (A+ to SRCNN)

図 3.3 に示すとおり、マルチステップ超解像技術では、従来手法と比較して、視覚的に大幅に画質が改善した。特に、細い線や輪郭部分が、従来手法よりも、さらにはっきりとしていることを確認した。その一方で、PSNR や SSIM といった再構成精度が向上しないことを確認した [67]。これは、先鋭化する際に画素位置がずれることで、画素値が大きく変化したことが原因であると考えられる。

3.2.2 方向性を考慮した超解像技術

続いて、方向性を考慮した超解像技術（回転・反転を伴う超解像技術） [18]について説明する。

辞書利用の学習型超解像では、訓練画像として ImageNet [68]で提供されるような一般画像を使用している。そのため、学習に使用した画像に応じて、特徴量の方向性の偏りが少なからず生じてしまう問題がある。訓練画像中の被写体が人工物や人である場合を考えると、それらは重力に従って垂直の構造を取ることが多く、その次に水平、斜め、弧の構造を取ると考えられる。よって、これらの訓練画像では、必然的に垂直方向のエッジの量が多くなる。

したがって、学習型超解像では、上記の「偏りを持った訓練画像」を用いて作成した辞書を利用することから、辞書のパッチ中に存在するテクスチャの方向性が原因で、入力画像の方向性に依存して画質向上度に差が生じるという問題があった [18][69][70]。

そこで、方向性を考慮した超解像技術では、この「辞書のパッチ中の方向性」を逆に利用し、複数の方向に対して超解像処理を行った結果を重畳することにより、再構成精度を向上させた。

図 3.4 に方向性を考慮した超解像技術の構成図を示す。

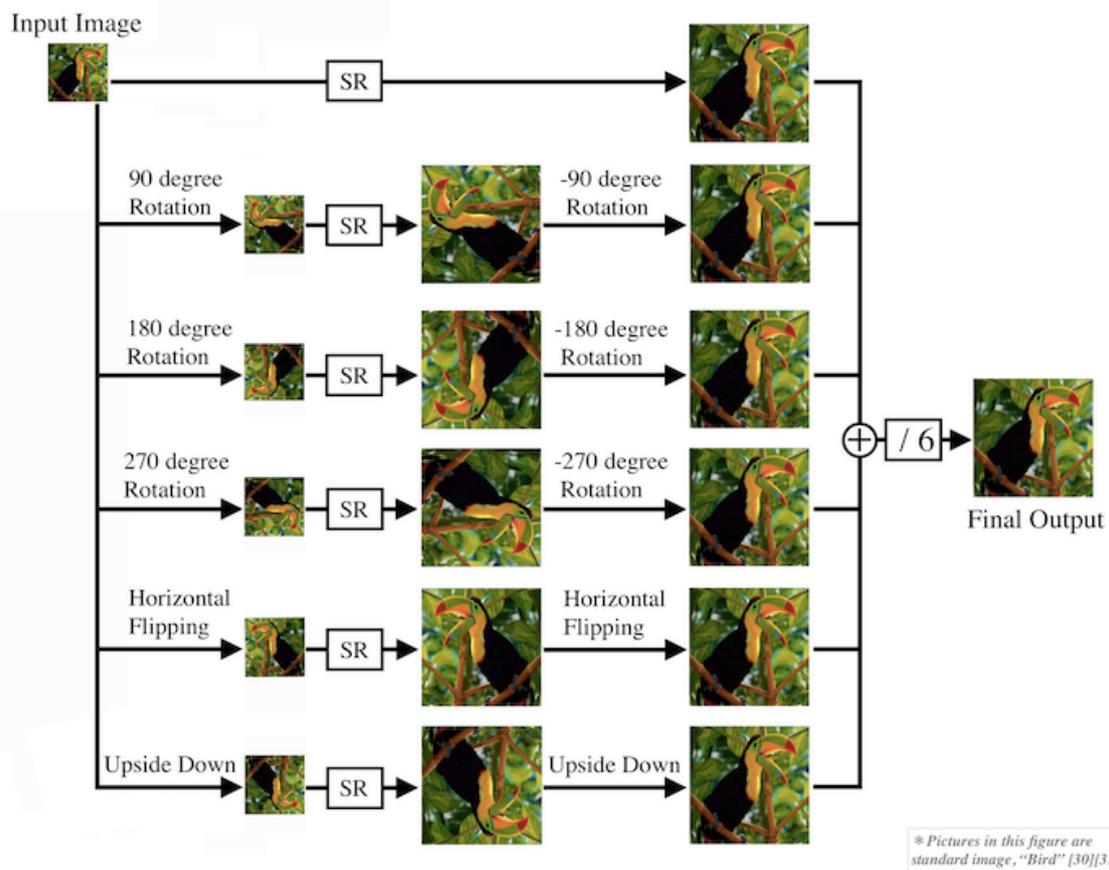


図 3.4 方向性を考慮した超解像技術の構成図

図 3.4 に示すとおり，方向性を考慮した超解像技術では，超解像処理の前処理として入力画像を回転・反転させる処理を，後処理として元の回転状態への復元処理を付加することで，入力パッチ中の方向性を考慮する．さらに，それらの操作により取得した複数枚の超解像結果を重畳（平均化）することで，再構成精度を向上させる．

方向性を考慮した超解像技術の評価実験結果から，本手法では，従来手法よりも PSNR や SSIM が向上することを確認した．さらに，超解像画像を目視により定性的に比較した結果，本手法ではリングングや擬輪郭といった画質劣化が低減されており，より自然な画像が取得できることを確認した．

さらに，我々は過去の研究において，方向性を考慮した超解像技術が再構成精度を向上させる要因についても研究を行った [69]．その結果，本手法による画質向上は，回転・反転処理，Bicubic 補間拡大処理後の超解像処理部において，6 枚の画像の画素値が変化したことが原因であることが確認された．我々はこの結果から，再構成精度

の向上は、以下のように生じると考察した [69].

SRCNN は辞書中で、入力低解像度パッチが高次元ベクトルとして表現され、非線形写像により高解像度パッチへのマッピングが行われる。また、A+ や ScSR ではスパース信号表現に基づいた超解像処理が行われ、基底の高解像度への置換処理、およびそれらの線形和により高解像度画像を取得する。したがって、上記の学習型超解像手法では、画像を回転、反転させることにより、超解像処理時に選択される低解像度パッチや、低解像度の基底の組み合わせが変化することになる。そのため、これらに対応する高解像度パッチに置換すると、必ず画素値の異なるパッチへ変換される。

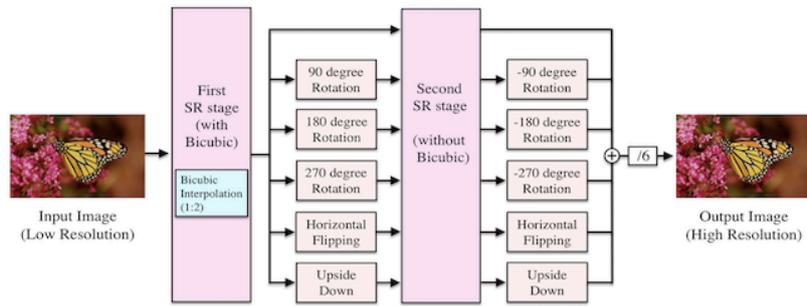
ここで、辞書内では、入力された低解像度画像パッチに最も類似したパッチ、または基底の線形和が選択されるが、この選択には誤差が生じ得る。加えて、これを高解像度パッチ、または高解像度の基底に置換する際にも誤差が生じる可能性がある。例えば、エッジの過度な強調が行われたり、余計な高周波成分が付加されたりする可能性がある。方向性を考慮した超解像技術では、複数の回転状態に対する出力結果を足し合わせることで、上記の「低解像度パッチ選択の誤差」および「高解像度変換時の誤差」が緩和され、より滑らかな出力画像を取得することができたと考えられる。

さらに、上記の結論に基づけば、SRCNN, A+ や ScSR に限らず、一般的な学習型超解像手法ではパッチや基底の選択における誤差、およびマッピングによる誤差が生じるため、他の学習型超解像手法にも適用可能であると示唆した [69].

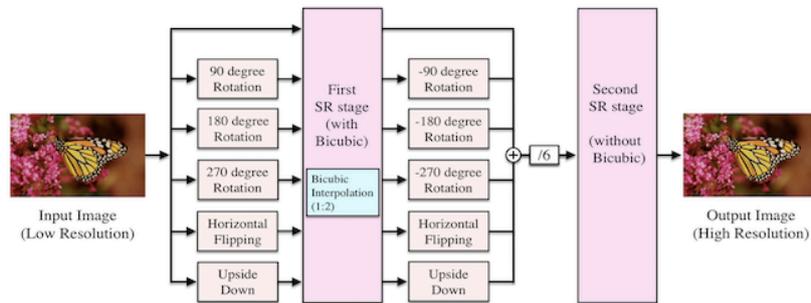
3.3 提案手法の構成

提案手法では、視覚的に画質を大幅に向上させるマルチステップ超解像技術と、方向性を考慮した超解像技術とを合成することで、従来手法である A+, ScSR, および SRCNN よりも知覚品質を向上させ、再構成精度をできる限り従来手法に近づける。

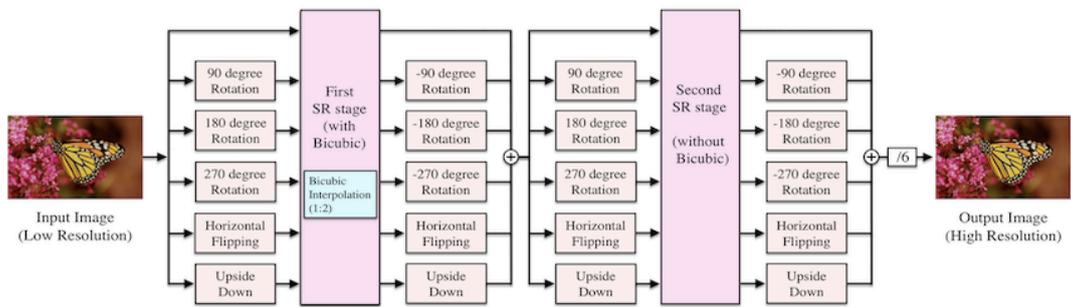
本研究では、4 種類の手法を提案する。図 3.5 に提案手法の構成図を示す。



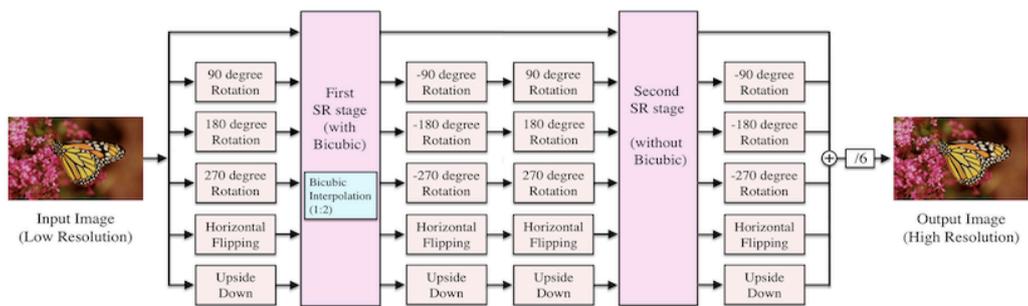
(1) MSSR-2



(2) MSSR-3



(3) MSSR+



(4) MSSR+2

* Pictures in this figure are standard image, "Monarch" [30][31].

図 3.5 提案手法の構成図

各種提案手法において、1 段目の超解像処理部には、A+, または ScSR を使用し、2 段目には SRCNN を使用した。

まず、MSSR-2 では、マルチステップの超解像技術の 2 段目に方向性を伴う超解像技術を適用する。これにより、MSSR-2 では、2 段目の超解像処理部において、より正確な（再構成精度の高い）超解像処理を実行することを図る。

次に、MSSR-3 では、方向性を伴う超解像技術を 1 段目に適用する。そのため、MSSR-3 では、1 段目の超解像処理部で出力される超解像結果の再構成精度が向上することが予測される。したがって、原画像に対するエラーがより少ない画像が 2 段目の超解像処理部に渡されるため、2 段目の超解像処理部において、より正確な高画質化の実行を期待できる。

続いて、MSSR+では、1 段目と 2 段目の双方に、方向性を伴う超解像技術を適用する。この手法では、マルチステップ超解像の各段の超解像処理部で、再構成精度が向上することが予測される。したがって、MSSR-2 や MSSR-3 よりも PSNR や SSIM の向上が見込まれる。

最後に、MSSR+2 は、方向性を伴う超解像技術をそのまま適用するのではなく、1 段目の超解像処理後に平均化処理を行わない。これにより、MSSR+2 では、2 段目の超解像処理部において、1 段目と同じ回転・反転状態に対して超解像処理を実行することになる。

3.4 むすび

本章では、はじめに、提案手法の概要と、提案手法を構成するマルチステップ超解像技術、および方向性を考慮した超解像技術について説明した。続いて、これらの手法を合成した提案手法について述べた。

第4章 評価実験

4.1 まえがき

本章では，提案手法の評価実験の概要，結果，および考察について述べる．

4.2 評価実験

本研究では，提案手法に対する評価実験として，MSSR-2，MSSR-3，MSSR+，および MSSR+2 について，従来の超解像手法と画質を比較し，性能を評価する．

4.2.1 実験方法

本研究では，画質評価指標を用いた定量的比較，および目視による定性的比較により，提案手法の評価を行った．評価実験には，標準画像 Set 5 および Set 14 [30][31] を使用した．また，再構成精度の超解像処理部で入力画像が 1 : 2 に拡大されることを考慮し，テスト画像に 2 : 1 ダウンサンプリングを行ったものを入力画像とした．

画質評価指標には，PSNR，SSIM，および BRISQUE を使用した．それぞれ，PSNR は，超解像処理が正解画像に対する誤差がどれだけ小さいかを判断する目的で，SSIM は，画像構造特徴の類似度を判断する目的で使用した．さらに，主観品質に比較的忠実に画質評価を行う BRISQUE を用いて，画質の向上の程度を定量的に測定した．

4.2.2 実験結果

表 4.1 から表 4.7 に，1 段目の超解像処理部に A+ を使用した場合における，画質評価値の比較を示す．まず，BRISQUE 値の比較結果を，表 4.1 および表 4.2 に示す．

なお，表中の Original は，事前処理においてダウンサンプリングを行う前の原画像

を示し, Bicubic はダウンサンプリング後に Bicubic 法を用いて補間拡大した結果を示す. また, A+ (T1) は, 従来手法 A+ [16]のみを使用した場合の超解像結果を示し, A+ (6) は, 方向性を考慮した超解像 [18]を A+に適用した場合の超解像結果を示す. MSSR は マルチステップ超解像技術 [67]を使用した場合の超解像結果を示し, MSSR-2 から MSSR+2 は, 第3章で記述した提案手法による超解像結果を示す.

表中の赤字は各画像における最高画質を示し, 下線は原画像よりも高画質である場合を示す. また, 薄紫の塗りつぶしは BRISQUE 値が 40 を上回る場合を示す.

表 4.1 BRISQUE の比較 (A+ 使用時)

Method Image	Original	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	33.4766	42.1806	42.7065	43.4581	43.4582	43.4581	43.4582	43.4582
Baby	28.2013	<u>22.3921</u>	<u>22.7796</u>	<u>21.6847</u>	<u>19.9471</u>	<u>21.8369</u>	<u>20.0209</u>	19.4451
Barbara	24.7818	<u>24.7417</u>	24.5671	32.4029	33.2722	31.8460	32.8394	32.8920
Bird	35.8869	37.5129	37.6817	25.6677	<u>26.5028</u>	<u>25.6929</u>	<u>26.5360</u>	<u>26.8230</u>
Bridge	9.4536	40.3840	41.3550	43.2547	43.4523	43.2756	43.4540	43.4545
Butterfly	42.8245	<u>42.1281</u>	42.0706	43.4471	43.4582	43.4413	43.4582	43.4582
Coastguard	43.4556	43.4582	43.4582	43.4582	43.4582	43.4582	43.4582	43.4582
Comic	39.8219	43.2271	43.3257	43.4582	43.4582	43.4582	43.4582	43.4582
Face	36.6044	34.4924	34.7425	<u>20.2218</u>	<u>22.0198</u>	19.9257	<u>22.0388</u>	<u>22.0556</u>
Flowers	21.8328	31.4709	32.8270	40.5720	42.4991	40.5208	42.4098	42.2359
Foreman	30.2200	31.4919	31.5207	10.4051	<u>13.9840</u>	<u>10.5622</u>	<u>13.8558</u>	<u>14.2869</u>
Head	36.6680	34.5984	34.8333	20.2911	<u>21.9957</u>	<u>20.0984</u>	<u>22.0370</u>	<u>21.9999</u>
Lena	16.2192	32.2501	33.1067	25.8766	27.3020	26.1714	27.2734	27.3970
Man	16.7750	40.0472	40.3377	43.3242	43.4303	43.2998	43.4338	43.4338
Monarch	2.7060	23.0019	23.4029	8.9264	15.3439	9.0176	16.5046	15.1100
Pepper	28.0928	38.0494	38.5068	29.1836	29.8055	29.0986	29.7944	30.0030
Ppt3	40.7031	41.3573	41.3290	37.0277	<u>38.5330</u>	<u>37.1402</u>	<u>38.6429</u>	<u>38.9978</u>
Woman	27.6664	<u>26.8437</u>	<u>26.9976</u>	12.2865	<u>15.6849</u>	<u>12.5584</u>	<u>15.7189</u>	<u>16.2873</u>
Zebra	21.8629	42.7485	43.1335	43.4432	43.4582	43.4463	43.4582	43.4582

BRISQUE 値の比較から, BRISQUE が 40 以下の場合, Barbara を除く全ての画像で, 提案手法の BRISQUE 値が, 従来手法 A+ よりも減少した(すなわち高画質となった). また, BRISQUE が 40 を超える場合, BRISQUE 値は各手法間であまり変化しないことが確認できた. また, 原画像 (Original) との BRISQUE 値の比較から, マルチステップ超解像技術を使用した超解像技術 (表中, 右 5 種) では, 原画像よりも高画質となる確率 (表下線) が従来手法よりも圧倒的に高いことがわかる.

続いて, 上記の結果について平均値を比較したものを表 4.2 に示す.

表 4.2 BRISQUE の平均値の比較 (A+ 使用時)

Method	Original	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Average	28.2765	35.3882	35.7201	30.9679	32.1612	30.9635	32.2027	32.1954
A+(T1)との差	-7.1118	-	+0.3319	-4.4203	-3.2270	-4.4247	-3.1856	-3.1928

BRISQUE の平均値の比較から, マルチステップ超解像技術を使用した超解像手法 (MSSR から MSSR+2) では, 従来手法である A+よりも BRISQUE 値が改善された. また, 表 4.2 の結果から, 提案手法 MSSR-3 において BRISQUE 値が平均的に最も削減されることを確認した.

次に, PSNR について比較した結果を, 表 4.3 から表 4.5 に示す.

表 4.3 PSNR の比較 1 [dB] (A+ 使用時)

Method Image	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	25.6530	25.6783	23.9428	24.2715	23.9661	24.2827	24.2917
Baby	38.5225	38.5680	35.0262	35.7270	35.0611	35.7554	35.7670
Barbara	28.6971	28.6251	26.5154	27.0379	26.5244	26.9610	26.9612
Bird	41.1493	41.3187	34.5658	35.2894	34.5854	35.3197	35.3593
Bridge	29.0955	29.1342	26.8749	27.3250	26.8861	27.3238	27.3369
Butterfly	32.0313	32.1484	24.8041	25.6578	24.8176	25.6684	25.6927
Coastguard	30.5984	30.7104	28.8937	29.1091	28.8932	29.1226	29.1554

表 4.4 PSNR の比較 2 [dB] (A+ 使用時)

Method Image	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Comic	28.3257	28.3568	24.9632	25.4680	24.9878	25.4794	25.4897
Face	35.7472	35.7749	33.9085	34.2801	33.9222	34.2825	34.2997
Flowers	33.0812	33.1368	28.9177	29.4689	28.9280	29.4824	29.5061
Foreman	37.0230	37.2137	31.6228	32.4391	31.6555	32.4903	32.4945
Head	35.7716	35.7989	33.9229	34.2992	33.9387	34.3045	34.3197
Lena	36.5989	36.6371	33.0002	33.4156	33.0187	33.4242	33.4398
Man	30.8865	30.9113	28.4740	28.8706	28.4962	28.8815	28.8972
Monarch	37.0284	37.1384	30.6127	31.4092	30.6341	31.4326	31.4573
Pepper	37.0875	37.1077	33.9191	34.2874	33.9076	34.2856	34.3072
Ppt3	30.4563	30.5541	26.9299	27.3988	26.9364	27.4123	27.4404
Woman	35.3197	35.4175	30.0617	30.7329	30.0548	30.7179	30.7317
Zebra	33.7001	33.8540	28.6549	29.4951	28.6590	29.5038	29.5421

表 4.5 PSNR の平均値の比較 [dB] (A+ 使用時)

Method Image	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	33.5144	33.5834	29.7690	30.3149	29.7828	30.3227	30.3416

PSNR の比較から、我々が過去に提案した「方向性を伴う超解像技術」を用いた超解像結果 (A+(6)) で、最も高い PSNR 値となった。さらに、この手法をマルチステップ超解像技術に適用することで、提案手法の 4 種 (MSSR-2 から MSSR+2) の PSNR 値が向上することが確認できた。特に、提案手法 MSSR+2 では、マルチステップ超解像技術を用いた超解像手法の中で、最高の PSNR 値をとることを確認した。

続いて SSIM による画質評価値について比較する。

表 4.6 SSIM の比較 (A+ 使用時)

Method Image	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	0.8630	0.8637	0.8383	0.8456	0.8388	0.8458	0.8461
Baby	0.9855	0.9857	0.9747	0.9779	0.9749	0.9781	0.9782
Barbara	0.9075	0.9057	0.8810	0.8875	0.8807	0.8859	0.8861
Bird	0.9969	0.9970	0.9857	0.9878	0.9858	0.9878	0.9879
Bridge	0.8480	0.8490	0.8129	0.8239	0.8135	0.8242	0.8244
Butterfly	0.9836	0.9840	0.9273	0.9377	0.9276	0.9379	0.9381
Coastguard	0.8479	0.8507	0.8288	0.8365	0.8300	0.8374	0.8378
Comic	0.9345	0.9350	0.8958	0.9045	0.8963	0.9048	0.9049
Face	0.9564	0.9566	0.9418	0.9450	0.9419	0.9450	0.9452
Flowers	0.9684	0.9688	0.9360	0.9419	0.9363	0.9422	0.9424
Foreman	0.9834	0.9837	0.9615	0.9674	0.9617	0.9676	0.9677
Head	0.9564	0.9567	0.9416	0.9449	0.9418	0.9449	0.9452
Lena	0.9915	0.9916	0.9836	0.9848	0.9836	0.9849	0.9849
Man	0.8753	0.8761	0.8336	0.8431	0.8340	0.8433	0.8438
Monarch	0.9942	0.9943	0.9792	0.9820	0.9793	0.9821	0.9822
Pepper	0.9932	0.9933	0.9879	0.9886	0.9879	0.9886	0.9886
Ppt3	0.9831	0.9835	0.9412	0.9462	0.9412	0.9462	0.9465
Woman	0.9876	0.9878	0.9690	0.9735	0.9691	0.9735	0.9736
Zebra	0.9847	0.9852	0.9536	0.9600	0.9537	0.9600	0.9603

表 4.7 SSIM の平均値の比較 (A+ 使用時)

Method Image	A+ (T1)	A+(6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	0.9495	0.9499	0.9249	0.9305	0.9252	0.9305	0.9307

SSIM の結果は PSNR による結果と同様の結果となった.

提案手法（4種）では，マルチステップ超解像技術に，方向性を考慮した超解像技術を適用することで，MSSR よりも SSIM が向上することが確認できた．さらに，提案手法 MSSR+2 は，マルチステップ超解像技術を用いた超解像手法の中で最も高い SSIM 値となった．

以上の結果をまとめると，BRISQUE 値の比較から，提案手法では，4 種類のいずれの手法においても，従来手法である A+より平均 BRISQUE 値が改善されることが示された．さらに，再構成精度の比較から，提案手法では，MSSR の再構成精度を改善できていることを確認した．中でも，MSSR+2 では，マルチステップ超解像技術を用いた超解像手法の中で，Barbara を除く全ての画像において，最高 SSIM 値をとることが確認できた．

続いて，超解像画像の目視による比較結果について述べる．

図 4.1，図 4.2 に従来手法，および提案手法を超解像画像の比較を示す．

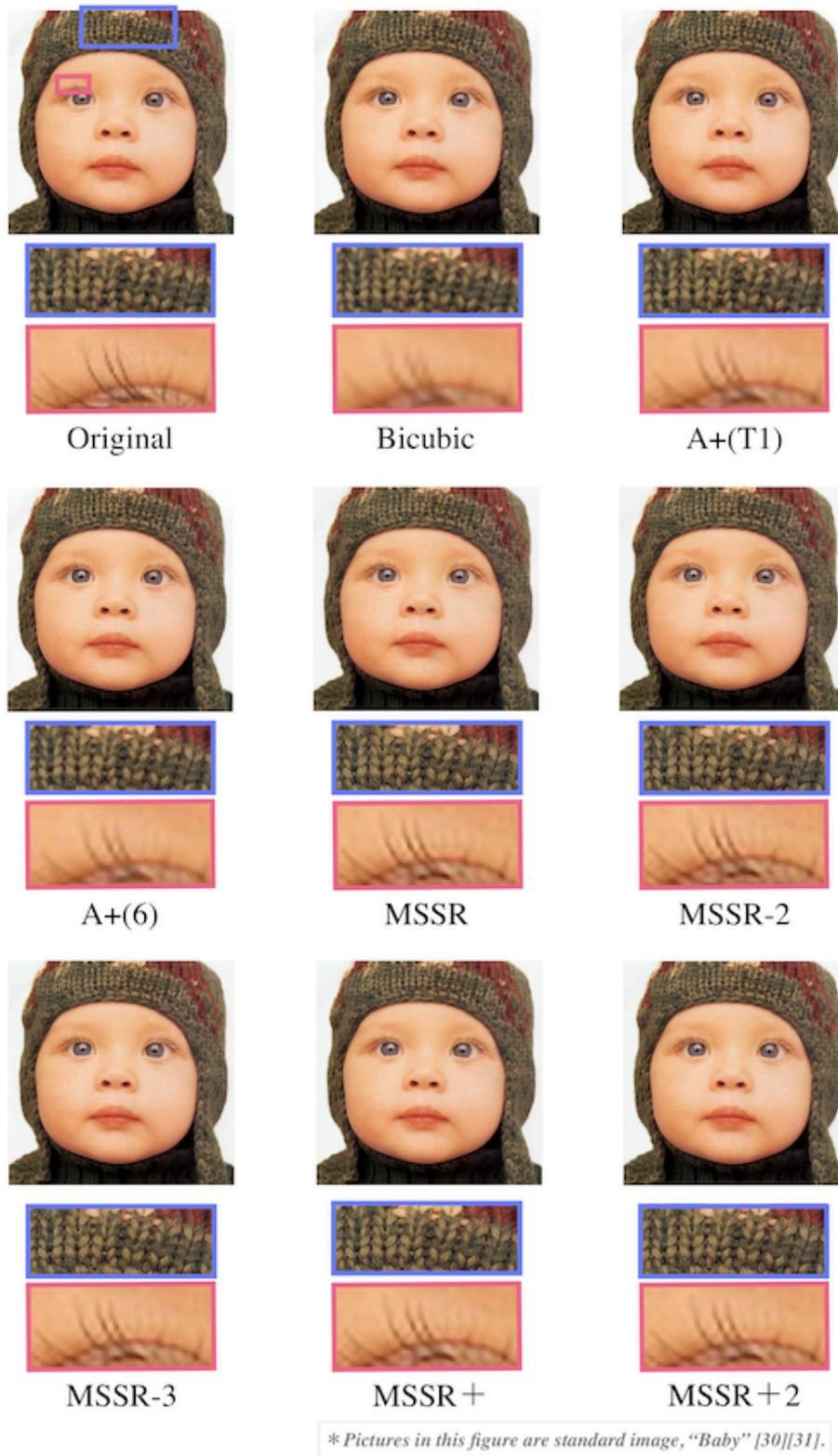


図 4.1 超解像画像の比較 1 (A+ to SRCNN)



* Pictures in this figure are standard image, "Flower" [30][31].

図 4.2 超解像画像の比較 2 (A+ to SRCNN)

目視による比較から、MSSR-3、および MSSR+2 による超解像画像が、最も高画質であった。特に、MSSR+2 の超解像結果では、全体的に滑らかに表現されており、自然な画像が生成できていることを確認した。

次に、1 段目の超解像処理部に ScSR を使用した場合の画質評価値の比較を、表 4.8 から表 4.13 に示す。

表中の ScSR (T1) は、従来手法 ScSR [17] を使用した場合の超解像結果を示し、ScSR (6) は、方向性を考慮した超解像 [18] を ScSR に適用した場合の超解像結果を示す。

表 4.8 BRISQUE の比較 (ScSR 使用時)

Method Image	Original	ScSR (T1)	ScSR (6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	33.4766	37.7566	39.5840	43.4580	43.4582	43.4582	43.4582	43.4582
Baby	28.2013	<u>24.4594</u>	<u>23.5438</u>	<u>22.6502</u>	<u>21.7406</u>	<u>20.8240</u>	20.3594	<u>21.1064</u>
Barbara	24.7818	22.2219	<u>23.8721</u>	30.1308	31.2079	30.2582	31.2264	31.1874
Bird	35.8869	36.3723	36.4348	20.4878	<u>22.4352</u>	<u>21.4015</u>	<u>23.0078</u>	<u>23.3624</u>
Bridge	9.4536	35.7344	37.2890	43.0232	43.4493	43.0622	43.4502	43.4483
Butterfly	42.8245	42.1765	<u>42.7201</u>	43.4476	43.4582	43.4423	43.4582	43.4582
Coastguard	43.4556	43.4582	43.4582	43.4582	43.4582	43.4582	43.4582	43.4582
Comic	39.8219	43.0158	43.3030	43.4582	43.4582	43.4582	43.4582	43.4582
Face	36.6044	<u>33.2562</u>	<u>33.9187</u>	<u>21.6840</u>	<u>23.3197</u>	21.3224	<u>22.8609</u>	<u>23.1289</u>
Flowers	21.8328	31.6142	28.7702	38.3521	40.7993	37.7168	39.7094	40.8621
Foreman	30.2200	29.2621	29.8455	<u>4.3167</u>	<u>5.3526</u>	4.2987	<u>5.5776</u>	<u>5.7248</u>
Head	36.6680	32.9902	33.7576	<u>21.4465</u>	<u>23.2994</u>	21.1412	<u>22.6668</u>	<u>23.0740</u>
Lena	16.2192	23.1903	28.3615	21.6248	23.3676	22.4952	23.6278	23.7735
Man	16.7750	30.0121	35.6014	43.2570	43.4248	43.2979	43.4343	43.4298
Monarch	2.7060	21.5916	22.4841	9.6356	17.4792	8.1124	14.7985	15.5914
Pepper	28.0928	34.0615	37.1655	23.6718	<u>26.9591</u>	<u>24.2364</u>	<u>26.2863</u>	<u>26.5436</u>
Ppt3	40.7031	38.5083	39.8533	36.6618	<u>38.4107</u>	<u>37.0405</u>	<u>38.6545</u>	<u>38.5023</u>
Woman	27.6664	25.9771	25.9095	<u>13.3676</u>	<u>12.8931</u>	10.3846	<u>12.1360</u>	<u>12.9506</u>
Zebra	21.8629	31.6534	36.1479	43.3569	43.4577	43.3963	43.4578	43.4576

表 4.9 BRISQUE の平均値の比較 (ScSR 使用時)

Method	Original	ScSR (T1)	ScSR (6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Average	28.2765	35.3882	33.7905	29.8678	31.1278	29.6213	30.7940	31.0514
A+(T1)との差	-4.2137	-	+1.3004	-2.6223	-1.3623	-2.8688	-1.6961	-1.4388

BRISQUE 値の比較から, 1 段目の超解像処理部に ScSR を使用する場合においても, A+を使用する場合と同様の結果が示された.

BRISQUE 値が 40 以下の場合, Barbara を除く全ての画像において, 提案手法の BRISQUE 値が従来手法 ScSR の BRISQUE 値よりも減少した. また, 原画像の BRISQUE 値との比較から, マルチステップ超解像技術を使用した超解像 (表中, 右 5 種) では, 原画像より画質が向上する確率 (表下線) が, 従来手法よりも圧倒的に高いことがわかる.

さらに, BRISQUE の平均値の比較から, マルチステップ超解像技術を使用した超解像手法 (MSSR から MSSR+2) では, 従来手法である ScSR よりも BRISQUE 値が改善されることが確認できた. また, 表 4.9 の結果から, 提案手法 MSSR-3 において BRISQUE 値が平均的に最も削減できることを確認した.

続いて, PSNR, および SSIM について比較した結果を, 表 4.10 から表 4.13 に示す.

表 4.1 0 PSNR の比較 (ScSR 使用時)

Method Image	ScSR (T1)	ScSR (6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	25.3219	25.3894	23.8232	24.1263	23.8587	24.1539	24.1508
Baby	38.0480	38.4129	34.8309	35.5244	34.9249	35.6146	35.6267
Barbara	28.5358	28.6336	26.5905	27.0484	26.6270	27.0754	27.0756
Bird	39.6128	40.3103	34.3915	35.0699	34.4890	35.1627	35.2105
Bridge	27.4844	27.6558	25.2050	25.7049	25.2380	25.7393	25.7454
Butterfly	31.0481	31.4528	24.7206	25.5567	24.8005	25.6408	25.6431
Coastguard	30.4613	30.6732	28.6087	28.8644	28.6801	28.9509	29.0098
Comic	27.8819	28.0758	24.9045	25.4192	24.9387	25.4529	25.4658
Face	35.5296	35.6914	33.7985	34.1623	33.8500	34.2062	34.2177
Flowers	32.5339	32.8264	28.8123	29.3695	28.8576	29.4146	29.4358
Foreman	34.5896	34.7385	30.9890	31.5446	31.2857	31.7673	31.7209
Head	35.5460	35.7064	33.8077	34.1770	33.8588	34.2180	34.2321
Lena	36.2219	36.4913	32.9636	33.4121	33.0153	33.4441	33.4624
Man	30.5785	30.7469	28.3688	28.7845	28.4193	28.8211	28.8377
Monarch	36.1726	36.5608	30.4994	31.3019	30.5834	31.3884	31.3905
Pepper	34.1698	34.3819	32.3959	33.1025	32.5461	33.1501	33.0477
Ppt3	29.4705	29.6887	26.5736	27.1300	26.6070	27.1636	27.1451
Woman	34.6160	34.9803	30.0382	30.7988	30.0919	30.8223	30.8325
Zebra	33.0742	33.5607	28.4996	29.3503	28.5735	29.4021	29.4208

表 4.1 1 PSNR の平均値の比較 [dB] (ScSR 使用時)

Method Image	ScSR (T1)	ScSR (6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	32.6788	32.9462	29.4643	30.0236	29.5392	30.0836	30.0879

PSNR の比較から、1 段目の超解像処理部に ScSR を使用する場合においても、方

向性を伴う超解像技術を用いた超解像結果 (ScSR(6)) で、最も高い PSNR 値となった。さらに、この手法をマルチステップ超解像技術に適用した提案手法 (4 種) では、マルチステップ超解像技術よりも PSNR が向上することを確認した。特に、提案手法のうち、MSSR+2 は、実験に使用した全てのデータセットにおいて、マルチステップ超解像を利用した超解像手法の中で、最も高い PSNR 値をとることを確認した。

表 4.1 2 SSIM の比較 (ScSR 使用時)

Method Image	ScSR (T1)	ScSR (6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	0.8285	0.8304	0.8073	0.8140	0.8080	0.8146	0.8147
Baby	0.9821	0.9835	0.9716	0.9750	0.9722	0.9754	0.9756
Barbara	0.8996	0.9009	0.8755	0.8810	0.8765	0.8819	0.8819
Bird	0.9929	0.9938	0.9821	0.9841	0.9824	0.9843	0.9845
Bridge	0.8427	0.8473	0.8030	0.8157	0.8047	0.8171	0.8174
Butterfly	0.9788	0.9806	0.9234	0.9337	0.9246	0.9348	0.9348
Coastguard	0.8481	0.8527	0.8245	0.8306	0.8256	0.8325	0.8335
Comic	0.9252	0.9281	0.8878	0.8967	0.8888	0.8975	0.8977
Face	0.9305	0.9321	0.9172	0.9203	0.9178	0.9209	0.9212
Flowers	0.9564	0.9585	0.9245	0.9305	0.9253	0.9312	0.9314
Foreman	0.9778	0.9788	0.9566	0.9623	0.9580	0.9633	0.9632
Head	0.9302	0.9318	0.9169	0.9201	0.9175	0.9206	0.9210
Lena	0.9871	0.9876	0.9791	0.9805	0.9793	0.9806	0.9807
Man	0.8692	0.8733	0.8273	0.8376	0.8290	0.8387	0.8392
Monarch	0.9922	0.9929	0.9775	0.9804	0.9779	0.9807	0.9807
Pepper	0.9840	0.9845	0.9783	0.9794	0.9787	0.9796	0.9796
Ppt3	0.9725	0.9739	0.9326	0.9383	0.9336	0.9392	0.9388
Woman	0.9844	0.9855	0.9662	0.9712	0.9666	0.9715	0.9716
Zebra	0.9819	0.9835	0.9508	0.9574	0.9516	0.9579	0.9580

表 4.1 3 SSIM の平均値の比較 [dB] (ScSR 使用時)

Method Image	ScSR (T1)	ScSR (6)	MSSR	MSSR-2	MSSR-3	MSSR+	MSSR+2
Baboon	0.9402	0.9421	0.9159	0.9215	0.9167	0.9222	0.9224

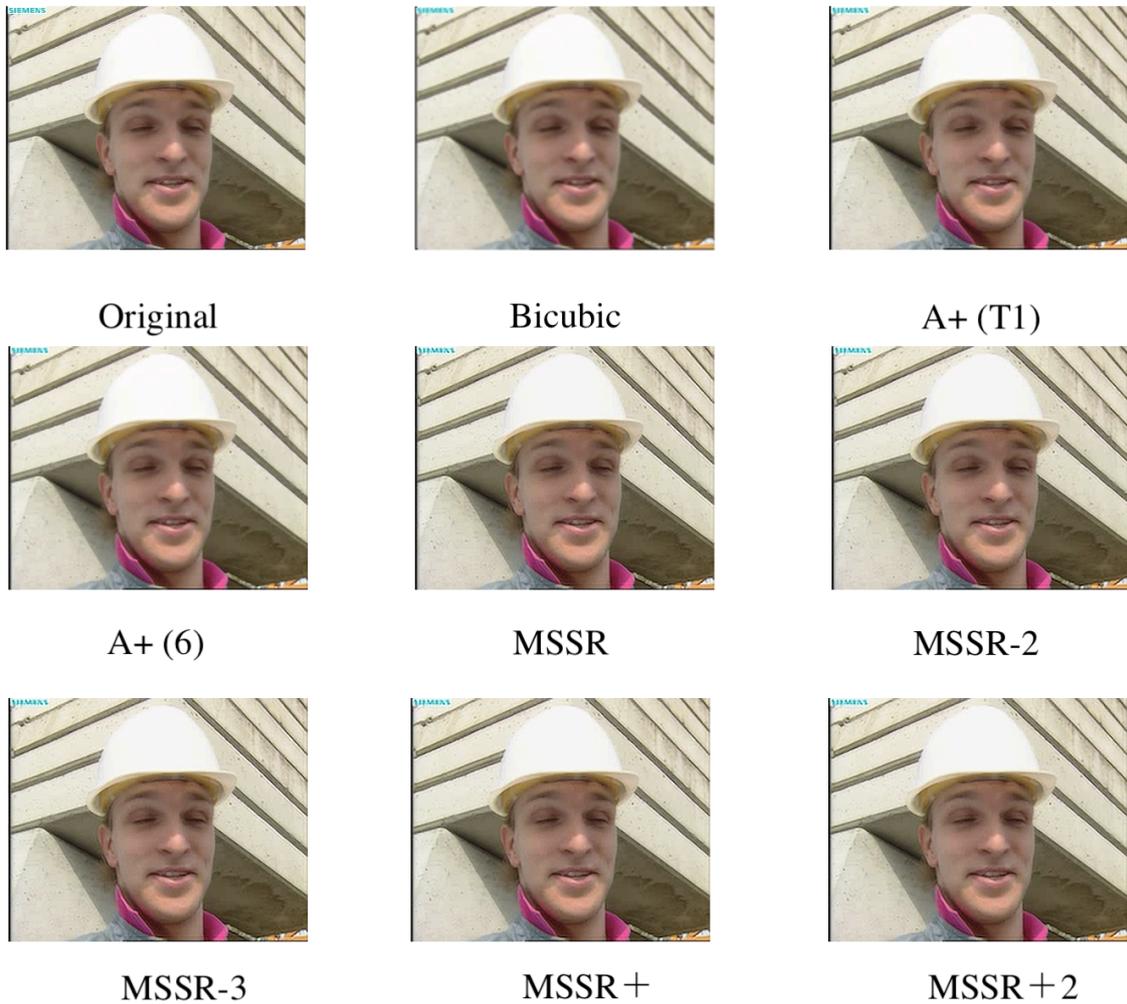
SSIM の比較結果より，提案手法（4 種）では，マルチステップ超解像技術に，方向性を考慮した超解像技術を適用することで，MSSR よりも SSIM が向上することを確認した．さらに，提案手法 MSSR+2 は，マルチステップ超解像技術を用いた超解像手法の中で最も高い SSIM 値をとった．

ScSR を使用した場合の以上の結果をまとめる．

BRISQUE 値の比較から，提案手法では，4 種類のいずれの手法においても，従来手法である ScSR より平均 BRISQUE 値が改善されることを確認した．さらに，再構成精度である PSNR および SSIM の比較から，提案手法では，MSSR の再構成精度を改善できていることが示された．特に，MSSR+2 では，全ての画像において，マルチステップ超解像技術を用いた超解像手法中で，最も高い SSIM 値をとることを確認した．

最後に，超解像画像の目視による比較結果について述べる．

図 4.3 に，超解像画像の比較を示す．



** Pictures in this figure are standard image, "Foreman" [30][31].*

図 4.3 超解像画像の比較 (ScSR to SRCNN)

目視による比較から、マルチステップ超解像技術を使用した 5 種類の超解像手法では、全ての手法において、従来手法 ScSR よりも視覚的な改善が見られた。中でも特に、MSSR+2 の超解像結果では、全体的に先鋭かつ自然に画像が生成できていることを確認した。

4.2.3 まとめと考察

本研究では，提案手法について，BRISQUE，PSNR，およびSSIMを用いて評価した．BRISQUE 値の比較から，BRISQUE 値が 40 以下の場合，Barbara を除く全ての画像において，提案手法は従来手法よりも画質が向上することが示された．また，BRISQUE の平均値の比較から，マルチステップ超解像技術を使用した超解像手法，5 種類全てにおいて，従来手法である A+，および ScSR よりも BRISQUE 値が改善されることを確認した．中でも，MSSR-3 において，最も平均 BRISQUE 値が削減される（すなわち最も高画質となる）ことを確認した．さらに，マルチステップ超解像技術を使用した超解像では，原画像よりも画質が向上し得ることが示され，この確率が従来手法よりも圧倒的に高いことが示された．

続いて，PSNR，および SSIM の比較から，提案手法では，マルチステップ超解像技術に，方向性を考慮した超解像技術を適用することで，マルチステップ超解像技術による再構成精度の低下を改善することが確認できた．特に，提案手法のうち，MSSR+2 は，マルチステップ超解像を利用した超解像手法のうち，最も高い PSNR 値および SSIM 値をとることを確認した．

さらに，目視による超解像画像の比較から，マルチステップ超解像技術を使用した 5 種類の超解像手法では，全ての手法において，従来手法 ScSR よりも視覚的な改善が見られた．その中でも特に，MSSR+2 の超解像結果では，全体的に先鋭かつ自然に画像が生成できていることを確認した．

以上の結果をもとに総合的に判断すると，MSSR+2 を使用することで，従来手法である A+，ScSR，および SRCNN よりも，知覚品質が大幅に向上し，さらに，再構成精度の低下も最小限に抑制できると考えられる．

ここで，MSSR-3 で最も平均 BRISQUE 値が削減される（すなわち最も高画質となる）結果となった原因と，MSSR+2 による知覚品質，および再構成精度の向上について考察する．

MSSR-3 は，マルチステップ超解像の 1 段目の超解像処理部に方向性を考慮した超解像技術を適用する手法である．MSSR-3 では，この構成により，2 段目の超解像処理部に，より画素値のエラーの少ない画像が渡される．したがって 2 段目の超解像処理段階において，より正確な超解像処理が実行できたと考えられる．

また，MSSR+2 による再構成精度の向上について考察する．MSSR+2 では，方向性を伴う超解像技術をそのまま適用するのではなく，1 段目の超解像処理後に平均化

処理を行わずに、直接 2 段目と回転・反転処理に連結する。したがって MSSR+2 では、1 段目と 2 段目の超解像処理部で同じ方向性に対して超解像処理が実行され、これにより再構成精度が向上したと考えられる。1 段目と 2 段目の超解像処理部で同じ方向性を持つことで再構成精度が向上した原因には、以下が考えられる。

第 3.2.2 項で述べたとおり、辞書利用の学習型超解像では、訓練画像として ImageNet [68] で提供されるような一般画像を使用している。この、訓練画像中の被写体が人工物、あるいは人である場合を考えると、それらは重力に従って垂直の構造を取ることが多く、その次に水平、斜め、弧の構造を取ると予想される。したがって、学習型超解像の辞書中には、必然的に垂直方向のエッジの量が多くなると考えられる。そのため、1 段目と 2 段目の超解像処理部で、同じ方向性を持つ画像について超解像処理を行うことで、辞書中の低解像度パッチ選択時、および、高解像度変換時に同じ方向性を持つパッチが適用され、これらの重畳により、方向性に基づくエラーが削減されたと考えられる。

4.3 むすび

本章では、提案手法の評価実験の概要を述べ、実験結果を比較した。実験結果から、提案手法 MSSR-2 を使用することで、従来手法の再構成精度を保ちながら、大幅に主観品質を向上させることができることを示した。

最後に実験結果のまとめ、および考察について述べた。

第5章 結論

本研究では、学習型超解像技術の従来手法である A+, ScSR, および SRCNN の再構成精度をできる限り保ちながら、知覚品質を大幅に向上させることを目的とし、学習型超解像の特性を活かした超解像技術を提案した。

提案手法では、マルチステップ超解像技術を利用することで知覚品質の向上を図り、さらに、方向性を考慮した超解像技術を適用することで、再構成精度の低下を抑制した。

評価実験の結果から、全ての提案手法において、平均 BRISQUE 値が従来手法よりも改善されることを確認した。続いて、提案手法のうち、MSSR+2 で再構成精度 (PSNR, および SSIM) が最高値をとることを確認した。さらに、目視による定性的な比較から、全ての提案手法において、従来手法である A+ や ScSR よりも主観品質が大幅に向上することを確認した。中でも、MSSR+2 の超解像結果では、全体的に先鋭かつ自然に画像が生成できていることを確認した。

評価実験結果に対する総合的な判断より、MSSR+2 を使用することで、従来手法よりも大幅に知覚品質が向上し、再構成精度の低下も最小限に抑えられることを確認した。

謝辞

本研究に際して、丁寧かつ素晴らしいご指導をしてくださり、実験環境および快適な研究環境を与えてくださった渡辺教授に、心より感謝いたします。

日頃から貴重な意見をくださり、研究室における温かい環境を提供してくださった、渡辺研究室の皆様に感謝いたします。

最後に、常に心を支えてくださり、私をここまで育ててくださり、生活を支えてくださっている家族に深く感謝いたします。

参考文献

- [1] 総務省：“放送政策の展開”，令和元年版 情報通信白書 ICT 白書 進化するデジタル経済とその先にある Society 5.0，第4節，2019.
- [2] 経済産業省：“テレビジョン受信機の現状について”，第1回 総合資源エネルギー調査会 省エネルギー・新エネルギー分科会 省エネルギー小委員会 テレビジョン受信機判断基準ワーキンググループ，資料2，Jan. 2019.
- [3] 一般社団法人 電子情報技術産業協会（JEITA）：“2019年 民生用電子機器国内出荷統計”，2019.
- [4] 一般社団法人 電子情報技術産業協会（JEITA）：“JEITA 調査統計ガイドブック 2018-2019 ～Executive Summary～”，Jul. 2017.
- [5] Apple（日本）：“iPhone 8 - 仕様 -”，<https://www.apple.com/jp/iphone-8/specs/>，[1/2020 閲覧].
- [6] Apple（日本）：“iPhone 11 Pro - 仕様 -”，<https://www.apple.com/jp/iphone-11-pro/specs/>，[1/2020 閲覧].
- [7] au：“ディスプレイ 仕様（スペック） | Android(TM) 技術情報”，<https://www.au.com/developer/android/kishu/display/>，[1/2020 閲覧].
- [8] 本橋直樹，中村聡史，鈴木俊博：“複数の縮小基底辞書を用いたスパースコーディングに基づく学習型超解像の高速化”，Ricoh Technical Report, No.42, pp.16-25, Feb. 2017.
- [9] 井田孝，松本信幸，五十川賢造，“画像の自己合同性を利用した再構成型超解像”，情報処理学会研究報告，Vol.59, No.23, pp.135-140, Dec. 2007.
- [10] 松本信幸，井田孝：“画像のエッジ部の自己合同性を利用した再構成型超解像”，電子情報通信学会論文誌，Vol. J93-D, No. 2, pp.118-126, Feb. 2010.
- [11] W. T. Freeman, T. R. Jones, E. C. Pasztor: “Example-based super-resolution”, IEEE Computer Graphics and Applications, Vol.22, Issue.2, pp.56-65, Mar/Apr. 2002.
- [12] 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, Apr. 2004.
- [13] Y. Blau, R. Mechrez, R. Timofte, T. Michaeli, and L. Zelnik-Manor: “2018 PIRM Challenge on Perceptual Image Super-resolution”, Proceedings of the European Conference on Computer Vision (ECCV) Workshops 2018, pp.1-22, Oct. 2018.
- [14] A. Mittal, A. K. Moorthy, and A. C. Bovik: “No-Reference Image Quality Assessment in the Spatial Domain”, IEEE Transactions on Image Processing, Vol.21, No.12, pp.4695-4708, Dec. 2012.
- [15] C. Dong, C. C. Loy, K. He, and X. Tang: “Image Super-Resolution Using Deep Convolutional Networks”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.38, No.2, pp.1-14, Jul. 2015.
- [16] R. Timofte, V. De Smet, and L. Van Gool: “A+: Adjusted anchored neighborhood regression for fast super-resolution”, IEEE Asian Conference on Computer Vision, pp. 111-126, Nov. 2014.
- [17] J. Yang, J. Wright, T. Huang, and Y. Ma: “Image super-resolution via sparse representation”, IEEE Transactions on Image Processing, Vol.19, No.11, pp.2861-2873, May. 2010.
- [18] 矢野仁愛, 梅田聖也, 渡辺裕, 猪飼知宏, 中條健, 伊藤典男: “画像の方向性を考慮した超解像処理技術 (Super-Resolution Technology Considering Directionality of Image)”, 2017年映像情報メディア学会冬季大会, 14B-2, Dec. 2017.
- [19] 竹島秀則, 加藤宣弘, 金子敏充: “サブピクセルシフト画像を用いた複数フレーム超解像技術”, 東芝レビュー, Vol.64, No.9, Jun. 2009.
- [20] R.G. Keys: “Cubic convolution interpolation for digital image processing”, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 29, No.6, pp.1153-1160, Dec. 1981.
- [21] Dingyi Li, Zengfu Wang: “Video Super-Resolution via Motion Compensation and Deep Residual Learning”, IEEE Transaction on Computational Imaging, Vol. PP, No.99, pp.1-15, Dec. 2017.
- [22] 池谷彰彦, 広明敏彦: “超解像ソリューション”, NEC 技報, Vol.60, No.2, pp.24-26, Apr.

2007.

- [23] 厚木七恵, 永田明德: “超解像技術による画像拡大の検討”, マルチメディア分散協調とモバイルシンポジウム, 2011 論文集, pp.1043-1048, Jun. 2011.
- [24] S. C. Park, M. K. Park, M. G. King: “Super-Resolution. Image Reconstruction: A Technical Overview”, IEEE Signal Processing Magazine, pp.21-36, May. 2003.
- [25] C. Lai, F. Li, B. Li, S. Jin: “Image super-resolution based on segmentation and classification with sparsity”, 2016 2nd IEEE International Conference on Computer and Communications, pp.563-567, Oct. 2016.
- [26] 小森秀樹: “超解像技術”, 映像情報メディア学会誌, Vol.63, No.10, pp.1400-1402, Oct. 2009.
- [27] 松尾康孝, 岩村俊輔, 井口和久, 境田慎一: “超解像技術を用いた実時間超高精細映像符号化システムの開発”, 映像情報メディア学会誌, Vol.70, No.1, pp.J22-J28, Dec. 2015.
- [28] S. Umeda, N. Yano, H. Watanabe, T. Ikai, T. Chujoh, N. Ito: “Deep Learning Based Super-Resolution for Future Video Coding”, International Workshop on Advanced Image Technology (IWAIT2018), No.89, pp.1-4, Jan. 2018.
- [29] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang: “Image super-resolution: The techniques, applications, and future”, Signal Process, Vol.128, pp.389-408, Nov. 2016.
- [30] C. Dong, C. C. Loy, K. He, and X. Tang: “Image Super-Resolution Using Deep Convolutional Networks”, <http://mmlab.ie.cuhk.edu.hk/projects/SRCNN.html>, Jan. 2015.
- [31] R. Timofte, V. De Smet, and L. Van Gool: “A+: Adjusted anchored neighborhood regression for fast super-resolution”, http://www.vision.ee.ethz.ch/~timofte/ACCV2014_ID820_SUPPLEMENTARY/index.html.
- [32] E. Levitan, G. T. Herman: “A Maximum a Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography”, IEEE Trans. Medical Imaging, Vol.6, No.3, pp.185-192, Sept. 1987.
- [33] R. R. Schulz and R. L. Stevenson: “Extraction of high-resolution frames from video sequences”, IEEE Transaction on Image Process, Vol.5, No.6, pp.996-1011, Jun. 1996.

- [34] 田中正行, 奥富正敏: “周波数領域最適化法による MAP 型超解像処理の高速化”, 情報処理学会論文誌, Vol.47, No. SIG 10 (CVIM15), pp.12-22, Jul. 2006.
- [35] A. J. Rockmore, A. Macovski: “A Maximum Likelihood Approach to Emission Image Reconstruction from Projections”, IEEE Nuclear and Plasma Sciences Society, pp.1428-1432, Aug. 1976.
- [36] M. Sezan, H. Stark: “Image restoration in CT by the method of projection onto convex sets”, ICASSP 1983 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol.8, pp.139-142, Apr. 1983.
- [37] M.I. Sezan, H. Stark: “Image Restoration by the Method of Convex Projections : Part2 - Applications and Numerical Results”, IEEE Transaction on Medical Imaging, Vol. MI-1, No.2, pp.95-101, Oct. 1982.
- [38] M. Irani, S. Peleg: “Improving resolution by image registration”, CVGIP : Graphical Models and Image Processing, Vol.53, No.3, pp.231-239, May. 1991.
- [39] C.J. Thompson, T.M. Peters: “A Fractional Address Accumulator for Fast Back-Projection”, IEEE Transactions on Nuclear Science, Vol.NS-28, No.4, pp. 3648-3650, Aug. 1981.
- [40] D. Glasner, S. Bagon, M. Irani: “Super-Resolution from a Single Image”, 2009 IEEE 12th International Conference on Computer Vision, pp.349-356, Sept. 2009.
- [41] S.L. Lo, C.T. Chiu: “Single Image Super-Resolution Using Hybrid Patch Search and Local Self-Similarity”, IEEE International Symposium on Circuits and Systems (ISCAS), pp.1-4, May. 2017.
- [42] S. T. Roweis, L. K. Saul: “Nonlinear dimensionality reduction by locally linear embedding”, Science, Vol. 290, No.5500, pp.2323-2326, Dec. 2000.
- [43] H. Chang, D.Y. Yeung, Y. Xiong: “Super-resolution through neighbor embedding”, the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol.1, pp.275-282, Feb. 2004.
- [44] M. Bevilacqua, A. Roumy, C. Guillemot, M. A. Morel: “Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding”, BMVC, pp.1-10, 2012.
- [45] J. Yang, J. Wright, T. Huang, Y. Ma: “Image super-resolution as sparse representation of

- raw image patches”, the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp.1-8, Jun. 2008.
- [46] R. Zeyde, M. Elad, M. Protter: “On Single Image Scale-Up Using Sparse-Representations”, *Curves and Surfaces 2010*, pp.711-730, Jan. 2010.
- [47] R. Timofte, V. D. Smet, L. V. Gool: “Anchored neighborhood regression for fast example-based super-resolution”, *IEEE International Conference on Computer Vision*, pp.1920-1927, Dec. 2013.
- [48] C. Y. Yang, M. H. Yang: “Fast Direct Super-Resolution by Simple Functions”, *2013 IEEE International Conference on Computer Vision (ICCV)*, pp.561-568, Dec. 2013.
- [49] C. Dong, C. C. Loy, and X. Tang: “Accelerating the Super-Resolution Convolutional Neural Network”, *European Conference on Computer Vision (ECCV) 2016*, pp.1-17, Aug. 2016.
- [50] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee: “Enhanced Deep Residual Networks for Single Image Super-Resolution”, the *2017 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-9, Jul. 2017.
- [51] W. Han, S. Chang, D. Liu, M. Yu, M. Witbrock, and T. S. Huang: “Image Super-Resolution via Dual-State Recurrent Networks”, the *2018 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-10, May. 2018.
- [52] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu: “Image Super-Resolution Using Very Deep Residual Channel Attention Networks”, *European Conference on Computer Vision (ECCV) 2018*, pp.1-16, Jul. 2018.
- [53] V. Nair, G.E. Hinton: “Rectified linear units improve restricted Boltzmann machines”, *International Conference on Machine Learning*, pp.807-814, Jun. 2010.
- [54] A. Kappeler, S. Yoo, Q. Dai, A. K. Katsaggelos: “Video Super-Resolution with Convolutional Neural Networks”, *IEEE Transactions on Computational Imaging*, Vol.2, No.2, pp.109-122, Mar. 2016.
- [55] E. J. Candes: “Compressive Sampling”, *Proceedings of the international congress of mathematicians*, Vol.3, pp.1433-1452, Aug. 2006.

- [56] D. L. Donoho: “Compressed sensing”, IEEE Transactions on Information Theory, Vol.52, No.4, pp.1289-1306, Apr. 2006.
- [57] 日野英逸, 村田昇: “スパース表現の数理とその応用”, 研究報告コンピュータビジョンとイメージメディア, Vol.2012-CVIM-183, No.20, pp.1-10, Sept. 2012.
- [58] 杉本修: “9章 符号化画質”, 電気情報通信学会, 知識の森, 2群 画像・音・言語 5編 画像符号化, pp.1-12, Apr. 2013.
- [59] 高橋玲: “1章 QoE: アプリケーションの品質”, 電気情報通信学会 知識の森, 3群 コンピュータネットワーク 5編 通信品質, pp.1-27, Jan. 2011.
- [60] 中易洋介, 前田純治, 鈴木幸司: “顕著性マップを用いた画質評価の実験的検討”, 第77回全国大会講演論文集 2015, No.1, pp.559-560, Mar. 2015.
- [61] J D. L. Ruderman: “The statistics of natural images”, Network: Computation in Neural Systems, Vol. 5, No.4, pp.517-548, Jul. 1994.
- [62] M. Carandini, D. J. Heeger, and J. A. Movshon: “Linearity and normalization in simple cells of the macaque primary visual cortex”, Journal of Neuroscience, Vol.17, No.21, pp.8621-8644, Nov. 1997.
- [63] M. J. Wainwright, O. Schwartz, and E. P. Simoncelli: “Natural Image Statistics and Divisive Normalization: Modeling Nonlinearities and Adaptation in Cortical Neurons”, Statistical Theories of the Brain, pp.203-222, Jan. 2001.
- [64] N. E. Lasmari, Y. Stitou, and Y. Berthoumieu: “Multiscale skewed heavy tailed model for texture analysis”, 2009 16th IEEE International Conference on Image Processing (ICIP), pp.2281-2284, Nov. 2009.
- [65] MATLAB: “非参照画質評価モデルの学習と使用”, https://jp.mathworks.com/help/images/train-and-use-a-no-reference-quality-assessment-model.html#mw_aef2dae4-669b-4ea9-a2c7-10a2b4a6d0c7 [1/2020 閲覧].
- [66] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli: “The SSIM Index for Image Quality Assessment”, <https://ece.uwaterloo.ca/~z70wang/research/ssim/>, Nov. 2014. [1/2020 閲覧].
- [67] 矢野仁愛, 渡辺裕: “多段超解像処理による視覚的な画質改善 (Visual Improvement

of Image Quality by Multiple-Step Super-Resolution)”, 2018 年映像情報メディア学会
冬季大会, 14D-5, Dec. 2018.

[68] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, L. F. Fei: “ImageNet: A large-scale
hierarchical image database”, IEEE Conference on Computer Vision and Pattern
Recognition, pp. 248-255, Jun. 2009.

[69] 矢野仁愛, 梅田聖也, 渡辺裕, 猪飼知宏, 中條健, 伊藤典男: “画像の方向性を考
慮した超解像処理技術 (Super-Resolution Technology Considering Directionality of
Image)”, 2017 年映像情報メディア学会冬季大会, 14B-2, Dec. 2017.

[70] 矢野仁愛, 渡辺裕: “複数の出力画像を用いた機械学習に基づく超解像処理技術”,
2018 年画像符号化シンポジウム・2018 年映像メディア処理シンポジウム
(PCSJ/IMPS 2018), P-4-8, Nov. 2018.

図一覧

図 2.1	超解像技術の分類.....	6
図 2.2	再構成型超解像の動画像への適用 (概念図).....	7
図 2.3	自己合同性を利用した超解像技術.....	9
図 2.4	学習型超解像の概念図.....	11
図 2.5	SRCNN のネットワークの構成図.....	14
図 2.6	ScSR の構成図 (概念図)	16
図 2.7	画質評価指標 PSNR, SSIM, BRISQUE による画質評価値の比較	23
図 3.1	学習型超解像の概念図 (再掲)	28
図 3.2	マルチステップ超解像技術の概念図.....	29
図 3.3	マルチステップ超解像技術による実験結果 (A+ to SRCNN).....	30
図 3.4	方向性を考慮した超解像技術の構成図.....	32
図 3.5	提案手法の構成図.....	34
図 4.1	超解像画像の比較 1 (A+ to SRCNN)	42
図 4.2	超解像画像の比較 2 (A+ to SRCNN)	43
図 4.3	超解像画像の比較 (ScSR to SRCNN)	49

表一覧

表 2.1 順位比較.....	24
表 4.1 BRISQUE の比較 (A+ 使用時)	37
表 4.2 BRISQUE の平均値の比較 (A+ 使用時)	38
表 4.3 PSNR の比較 1 [dB] (A+ 使用時)	38
表 4.4 PSNR の比較 2 [dB] (A+ 使用時)	39
表 4.5 PSNR の平均値の比較 [dB] (A+ 使用時)	39
表 4.6 SSIM の比較 (A+ 使用時)	40
表 4.7 SSIM の平均値の比較 (A+ 使用時)	40
表 4.8 BRISQUE の比較 (ScSR 使用時)	44
表 4.9 BRISQUE の平均値の比較 (ScSR 使用時)	45
表 4.1 0 PSNR の比較 (ScSR 使用時)	46
表 4.1 1 PSNR の平均値の比較 [dB] (ScSR 使用時)	46
表 4.1 2 SSIM の比較 (ScSR 使用時)	47
表 4.1 3 SSIM の平均値の比較 [dB] (ScSR 使用時)	48

研究業績

- [1] 矢野仁愛, 梅田聖也, 渡辺裕, 猪飼知宏, 中條健, 伊藤典男: “画像の方向性を考慮した超解像処理技術 (Super-Resolution Technology Considering Directionality of Image)”, 2017 年映像情報メディア学会冬季大会, 14B-2, Dec. 2017.
- [2] S. Umeda, N. Yano, H. Watanabe, T. Ikai, T. Chujoh, N. Ito: “Deep Learning Based Super-Resolution for Future Video Coding”, International Workshop on Advanced Image Technology (IWAIT2018), No.89, pp.1-4, Jan. 2018 (Best Paper Award).
- [3] 矢野仁愛: “画像の方向性を考慮した超解像処理技術 (Super-Resolution Technology Considering Directionality of Images)”, 卒業論文, Feb. 2018.
- [4] 矢野仁愛, 梅田聖也, 渡辺裕, 猪飼知宏, 中條健, 伊藤典男: “方向性を考慮した超解像処理技術による FVC 符号化画像の画質改善 (Quality Improvement of FVC Encoded Image Using Super-Resolution Considering Feature Directionality)”, 情報処理学会第 80 回全国大会, 2U-02, May. 2018.
- [5] 梅田聖也, 矢野仁愛, 渡辺裕, 猪飼知宏, 中條健, 伊藤典男: “FVC 量子化パラメータ制御による超解像の特性 (Quantize parameter dependent super resolution characteristics in Future Video Coding)”, 情報処理学会第 80 回全国大会, 2U-03, Mar. 2018.
- [6] 梅田聖也, 矢野仁愛, 渡辺裕: “映像符号化における超解像処理画像の画質改善 (Quality improvement of Super-Resolution Image for Video Coding)”, 2018 年画像符号化シンポジウム・2018 年映像メディア処理シンポジウム (PCSJ / IMPS 2018), P-1-8, Nov. 2018.
- [7] 矢野仁愛, 渡辺裕: “複数の出力画像を用いた機械学習に基づく超解像処理技術”, 2018 年画像符号化シンポジウム・2018 年映像メディア処理シンポジウム (PCSJ/IMPS 2018), P-4-8, Nov. 2018.

- [8] 梅田聖也, 矢野仁愛, 渡辺裕: “符号化画像に対する超解像処理の特性改善について (On Quality Improvement of Super Resolution Processing for Coded Images)”, 2018 年映像情報メディア学会冬季大会, 14D-4, Dec. 2018.
- [9] 矢野仁愛, 渡辺裕: “多段超解像処理による視覚的な画質改善 (Visual Improvement of Image Quality by Multiple-Step Super-Resolution)”, 2018 年映像情報メディア学会冬季大会, 14D-5, Dec. 2018.
- [10] 矢野仁愛, 渡辺裕: “マルチステップ超解像技術による高画質化処理 (Image Quality Enhancing by Multi-Step Super-Resolution)”, 電子情報通信学会総合大会, D-11-9, Mar. 2019.
- [11] N. Yano, H. Watanabe: “Image Quality Enhancement with Machine Learning Based Multi-Step Super-Resolution”, International Conference on Artificial Intelligence in Information and Communication (ICAIIIC) 2020, Feb. 2020 (発表予定).
- [12] 矢野仁愛, 渡辺裕: “機械学習に基づくマルチステップ超解像を応用した高解像度化処理手法 (High-Resolution Processing Applying Machine Learning Based Multi-Step Super-Resolution)”, 電子情報通信学会総合大会, Mar. 2020 (発表予定).

