情報理工学特別研究報告書

題目

画像超解像と標本化定理に基づく画像拡大の 比較に関する研究

学生証番号 954462 氏名 脇益朋也

提出日 令和5年1月26日 指導教員 蚊野 浩

京都産業大学情報理工学部

要約

デジタル画像の拡大はさまざまに利用されている. 典型的な利用例は, 画素数が 1920×1080 画素である地上波デジタル放送の画像を, 3840×2160 画素の 4 K テレビや 7680×4320 画素の 8 K テレビに表示する場合である.

よく知られている画像の拡大手法は最近隣補間,バイリニア補間,バイキュービック補間である.これらは標本化定理に基づく拡大手法と考えられる.これらの中で最も画質の良いバイキュービック補間であっても,拡大画像は平滑化されたような感じになる.これらの手法に比べて,高画質な拡大画像を生成することができる手法として画像超解像がある.画像超解像による拡大画像が高画質になる理由は,画素数を増やす処理の過程で,低解像度画像には存在しない高周波成分が付け加えられるからである.しかし,従来研究では,付加された成分の性質について,十分に明確にされていない.

本研究は、画像超解像において付加される成分の性質を明確に示すことを目標とした。そのために、バイキュービック補間による拡大画像の PSNR と、画像超解像による拡大画像の PSNR を、4×4の小領域ごとに比較した。比較する小領域を一画素ずつずらせることで PSNR の差画像を生成し、画質評価に差が出ている部分と差がない部分を明確にした。その後、差画像から明確になった画像超解像とバイキュービック補間で大きな差が生じている部分と差が生じていない部分について、その画素値をグラフ化した。そのグラフを見ると、画像超解像を用いることで、濃淡変化が大きく、エッジのある部分で特に高画質化されていることがわかった。

結論として、画像超解像による拡大は、エッジが存在する部分について、エッジを 急峻にすることで高画質化していることがわかった.しかし、どのようにそれを行い、 高画質化しているのかはわからなかった.

目次

1章 序論	• • • 1
2章 標本化定理に基づく画素の補間と超解像技術	3
2.1 標本化定理に基づく画素の補間	3
2.2 最近隣補間,バイリニア補間,バイキュービック補間	• • • 5
2.2 合志の超解像	6
2.3 深層ニューラルネットを用いた画像超解像	• • • 7
3章 拡大画像の定量的評価	8
3.1 原画像から縮小画像を生成する方法	8
3.2 原画像と拡大画像の PNSR の比較	8
3.3 高周波成分の比較	• • • 10
4章 バイキュービック補間と画像超解像の比較	• • • 13
4.1 小領域 PSNR の比較による画質の確認	• • • 13
5章 結論	• • • 17
参考文献	• • • 18
謝辞	• • • 18
付録	• • • 19

1章 序論

デジタル画像は画素が縦横に規則的に並んだものである。画像を拡大する処理は、画素と画素の間に新たな画素を補間することにより、画素数が増えたデジタル画像を生成することである。このようなデジタル画像の拡大は、さまざまに利用されている。 典型的な利用例は、画素数が 1920×1080 画素である地上波デジタル放送の画像を、 3840×2160 画素の 4 K テレビや 7680×4320 画素の 8 K テレビに表示する場合である。

よく知られている画像の拡大手法に、最近隣補間、バイリニア補間、バイキュービック補間がある。最近隣補間は補間する画素の値として、最も近くに存在する原画像の画素値を用いる方法である。バイリニア補間とバイキュービック補間は補間する画素を囲む複数の画素値を加重平均する方法である。これらの方法の詳細は2章で説明するが、要するに、標本化定理に基づく補間の計算手順を簡単化したものであると考えることができる。これらの方法の中で、拡大画像の画質が最も良いものはバイキュービック補間である。しかし、バイキュービック補間であっても、拡大画像は平滑化されて劣化したような感じになる。

従来の画像拡大手法は標本化定理に基づいている.従って、画像を拡大して画素数を増やしても、その処理が標本化定理に基づいて忠実に処理されていれば、原画像に含まれる周波数成分は変化しない.そのような拡大画像を表示する時に、画素密度を調整して原画像と同じ大きさになるように表示すれば、同じ画像に見える.一方、拡大画像を、原画像と同じ画素密度で表示すれば見かけの画像は拡大されるが、平滑化され、少しぼやけた感じに見える.

標本化定理に基づく手法に比べて,高画質な拡大画像を生成することができる手法として画像超解像がある.画像超解像には,一枚の画像だけを用いてそれを拡大する単眼超解像と,同じ部分を撮影した複数の画像を合成して一枚の拡大画像を生成する手法がある.本研究では単眼超解像と標本化定理に基づく画像の拡大を比較する.

単眼超解像は、深層ニューラルネットワークによって実現される機械学習に基づくものが多い。この方法では、高解像度画像とそれに対応した低解像度画像のペアを多数準備する。そして、低解像度画像を入力して高解像度画像を出力するように構成した深層ニューラルネットにおいて、正解の高解像度画像を出力するように学習させる。このような学習に成功した深層ニューラルネットに低解像画像を入力すると、正解の高解像度画像に近い画像を出力することができる。

画像超解像による拡大画像が高画質であることを示す方法として、低解像度画像と高解像度画像のペアを例示することで主観的に確認する方法と、PSNR や SSIM などの客観的指標を用いることで全体として画質が改善されていることを示す方法がある.

画像超解像による拡大画像が高画質になる理由は、画素数を増やす処理の過程で、低解像度画像には存在せず、高解像度画像に存在する成分が付け加えられるからである.しかし、従来研究では、付加された成分の性質について、十分に明確にされていない.本研究は、画像超解像において付加される成分の性質を明確に示すことを目標にした.以下、2章では標本化定理に基づく補間、合志と waifu2x の超解像技術について述べる.3章では拡大画像の定量的評価と実験内容について述べる.4章で具体的な実験方法と実験結果を書き、標本化定理に基づく補間と画像超解像を比較し、違いが出る原因や画像超解像の性質を考察する.5章では本研究での成果と今後の課題を述べる.

2章 標本化定理に基づく画素の補間と超解像技術

2.1 標本化定理に基づく画素の補間

画像の拡大は、結局、画素と画素の間に新しい画素を追加することで、縦横に画素数を増加させることである。これを行うための基礎となる考え方は標本化定理である。画像拡大において、しばしば利用される最近傍補間、バイリニア補間、バイキュービック補間などの従来的な手法は、標本化定理を基礎としている。これらの手法は標本化定理に基づく補間を近似する方法であると考えることができる。標本化定理に基づいて、忠実に補間することで、標本化前の元のアナログ信号を完全に再現することができる。この節では、標本化定理に基づいて忠実に補間する方法を、1次元信号の波形を用いて説明する。

まず,元波形(図2.1,図2.2,図2.3の例は振幅が0から1の正弦波)とそれを等間隔で標本化した標本点を用意する.次に,それぞれの標本点に振幅,位相,周波数を合わせたsinc関数を重ねる.最後に,全てのsinc関数を足し合わせる.そうすることで,標本点から元波形を再構成し,再現することができる.標本化定理に基づいて忠実に補間することで,図2.3のように元の信号波形を再現することができる.ただし,図2.3の場合,有限個の標本点を用いて再現したので完全な再現になっていない.

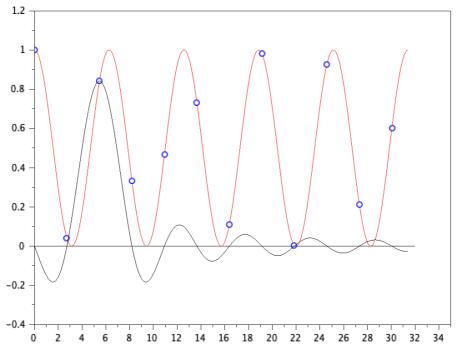


図 2.1 元波形 (赤), その標本点 (青) とそれに合わせた 1 つの sinc 関数 (黒)

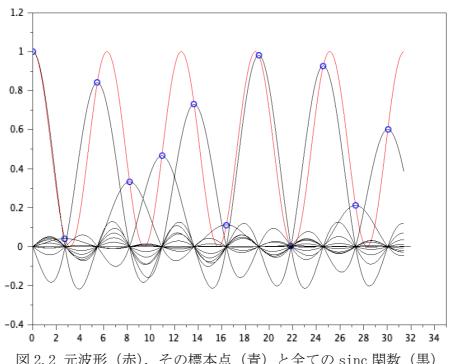


図 2.2 元波形 (赤), その標本点 (青) と全ての sinc 関数 (黒)

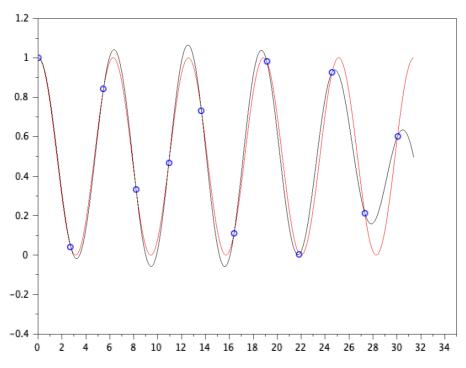


図2.3 元波形(赤), その標本点(青) および 全ての sinc 関数を足し合わせて再現した元波形(黒)

2.2 最近隣補間、バイリニア補間、バイキュービック補間

次に、図 2.4 を用いて、画像を 2 倍に拡大する場合の最近隣補間、バイリニア補間、バイキュービック補間を説明する((注)図 2.4 の領域の拡大率は、厳密に言えば 7/4=1.75 倍の拡大であるが、広い画像の一部と考えれば、全体での拡大率はほぼ 2 倍である)。図 2.4 の灰色の部分は拡大前の画像の画素値が決まっている画素とする。その他の部分は拡大した時に画素値を決める必要がある画素である。青色の画素値を決める場合、最近傍補間は青色画素から 1 番近い画素(この場合、1,2,3,4 のいずれでも良い)を青色の画素値とする方法である。バイリニア補間は青色画素を囲む 2×2 の 4 画素である 1,2,3,4 の画素値の平均を青色の画素値とする方法である。バイキュービック補間は青色画素を囲む 4×4 の 16 画素(この図の灰色の画素全て)の画素値を 用いて加重平均を求め、それを青色の画素値とする方法である。

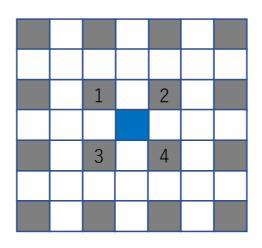
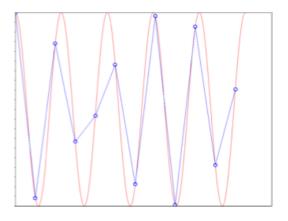


図 2.4 画像を 2 倍に拡大する場合の画像の一部

これらの補間法は、標本化定理に基づく補間を簡単化したものであるから、画像を 拡大しても、原画像に存在しない高周波成分を生成することができない. 従って、元 の画像以上の画質の画像を作り出すことはできない.

バイリニア補間,バイキュービック補間について,1次元信号(正弦波)を使い,信号がどのように補間されるかを確認した.その結果を図2.5に示す.(これらの手法の1次元信号バージョンの名称は正確にはリニア補間,キュービック補間である)



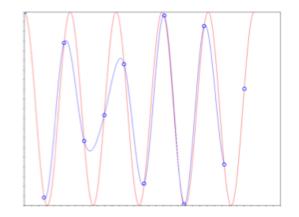


図 2.5 1次元信号での補間シミュレーション結果 (左:バイリニア補間,右:バイキュービック補間)

バイリニア補間は、前後の1画素しか参考にしない. そのため、補間した信号は標本点と標本点を線分で繋ぐ形になる. バイキュービック補間は、前後の2画素の加重平均を用いる. その結果、補間した信号は標本点と標本点を少し丸みが帯びた形で繋ぐ形になる.

2.3 合志の超解像

画像超解像も,拡大前の画像から画素数が増加した拡大画像を生成する方法である. ただし,標本化定理に基づく方法と異なり,元の画像に存在しない高周波成分を付加することで高画質化することができる.そのような画像超解像の一つに合志の方法がある[1].

図 2.6 を用いて合志の方法を説明する. 合志の超解像はまず,入力画像をバイキュービック補間で拡大する. ついで,拡大画像にハイパスフィルタをかけることで高周波成分を抽出する. この高周波成分を3乗し,リミッタをかけた後,拡大画像と足し合わせる. 高周波成分を3乗する過程で,原画像に含まれる周波数の3倍の周波数成分が発生する.



図 2.6 合志の超解像手法



図 2.7 アンシャープマスキングの処理方法

合志の超解像は、画像の鮮鋭化に用いるアンシャープマスキング処理方法と似た処理になっている。図 2.7 にアンシャープマスキングの処理方法を示す。アンシャープマスキングは入力画像に対してハイパスフィルタをかけることで高周波成分を抽出し、入力画像と足し合わせる処理である。これによって画像内の高周波成分を強調することができる。合志超解像とアンシャープマスキングの違いとしては、3乗し、リミッタをかける部分である。3乗をすることによって、入力画像にはない高周波成分を生成することができる。合志の超解像は基本的に3乗をするだけであるため、他の手法に比べて、簡単な方法で高解像度の画像を生成することができる点が特徴である。

2.4 深層ニューラルネットを用いた画像超解像

深層ニューラルネットワークを用いた画像超解像が多数提案されている[2]. SRCNN[3]は 2014 年に発表された3層の畳み込み層で構成されたモデルである. SRCNN をもとにした waifu2x という超解像ソフトウエアが広く利用されており、本研究でも waifu2x を用いた.

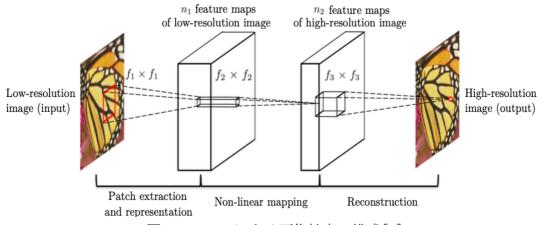


図 2.8 SRCNN による画像拡大の構成[4]

SRCNN による画像拡大の処理の流れを図 2.8 に示す.まず,低解像度画像を入力する.最初の畳み込み層で低解像画像の特徴マップを計算する.2 つ目の層で特徴マップを高解像度の特徴マップに非線形マッピングする.これを入力画像の小領域ごとに

行う. そして, 最後の層で組み合わせることで, 最終的な高解像度画像を生成する.

SRCNN の内部で行われているこれらの処理は、事前に膨大な量の低解像度画像とそれに対応する高解像度画像が準備し、それらを用いて、低解像度画像と高解像度画像の対応関係をネットワークに学習させておくことで実現される.

3章 拡大画像の画質評価

画像拡大技術を定量的に評価するために 512×512 画素の原画像を用意した. 原画像を 128×128 画素の画像に縮小し、これをそれぞれの方法で拡大した後、拡大画像 と原画像を比較した. 比較の方法は PSNR と 2 次元フーリエ変換を用いた.

3.1 原画像から縮小画像を生成する方法

 512×512 画素の原画像を 128×128 画素の画像に縮小するために平均画素法を用いた. 平均画素法では, 画像を一つの面積の四角形に対応させ, それに変換前と変換後の画素数の画素グリッドをあてはめる. 画像縮小の場合, 変換後の一画素に対して, 変換前の複数の画素が対応するので, それらの画素値を面積比で加重平均したものを, 変換後の画素値とする. 図 3.1 は, 7×7 画素の画像を 4×4 画素に縮小する場合の説明図である. 赤いグリッドで示す画像の左上隅の画素は, 黒いグリッドで示す画像の4つの画素値 p_1 , p_2 , p_3 , p_4 を使って p_1 +0.75 p_2 +0.75 p_3 +0.5625 p_4 /3.0625 となる.

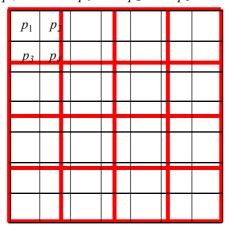


図 3.1 平均画素法による画像縮小を説明する図

 512×512 画素の原画像を 128×128 画素に縮小する場合,縮小後の一画素は原画像では正確に 4×4 画素に対応するので,これら 16 個の画素値を平均すればよい.

3.2 原画像と拡大画像の PNSR の比較

原画像と拡大画像の画質の違いを PSNR (Peak Signal Noise Ratio) で評価した. PSRN は同じサイズの2つの画像が、どの程度近いかを客観的に評価する指標である. 比較対象の拡大手法は最近隣法、バイリニア、バイキュービック、合志の方法、SRCNNを実装したプログラムである waif2x の5種類である. 図3.2の5つの画像と5種類の手法で拡大した場合の PSNR を表3.1 に示す.

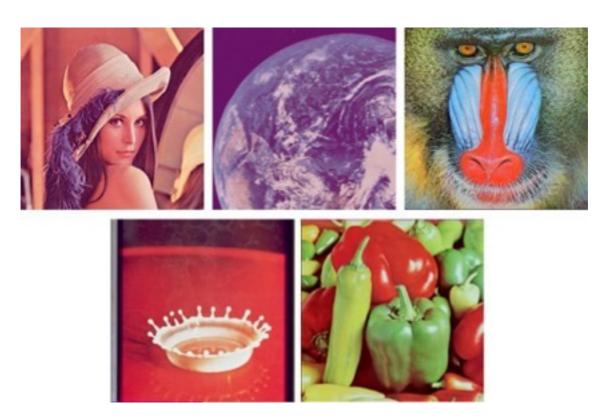


図 3.2 原画像 (左上から lena, earth, mandrill, milkdrop, pepper)

最近傍 バイリニア バイキュービック 合志 waifu2x earth 27.3 28.1 29.1 29.4 29.7 20.7 20.8 mandrill 20.4 20.5 20.7 27.8 29.0 29.8 30.0 30.7 milkdrop 25.7 26.8 27.5 27.7 28.4 pepper 27.3 28.3 29.3 29.5 31.0 lena

表 3.1 PSNR による画質評価の結果

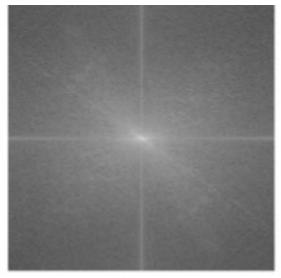
表 3.1 から、mandrill の評価値が、どの拡大方法でも低くなっていることがわかる. Mandrill の画像は毛並みの部分が細かく変化しており、原画像を縮小する際に、毛並みの変化が失われた. そのため、どの拡大方法でも評価値が低くなった. mandrill 以外の評価値を見ると、どの画像に対しても、waifu2x の評価値が最も高くなっている. 全体の画質評価の結果としては、最近傍補間、バイリニア補間、バイキュービック補間、合志の超解像、waifu2x の超解像の順に PSNR の評価値が高くなっている.

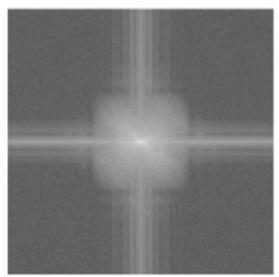
それぞれの拡大方法ごとの評価値の違いを見ると、最近傍とバイリニア補間では

 $1.0\sim1.2$, バイリニア補間とバイキュービック補間では $0.7\sim1.0$ となっており,標本化定理に基づく補間の間でも,補間に用いる画素数の違いから,評価結果に違いが出ていることがわかる.また,バイキュービック補間と合志の超解像では,評価値の違いが $0.2\sim0.4$ になっており,評価値にあまり大きな差が出ていない.バイキュービック補間と waifu2x の超解像では評価値の違いが $0.6\sim0.9$ になっており,画像超解像でも wiafu2x 方がより評価値が高くなっている.

3.3 高周波成分の比較

2章で述べたように、標本化定理に基づく補間では高周波成分が生成されず、画像 超解像では高周波成分が生成される. そこで実際に高周波成分にも違いがあることを 示すために、1ena 画像に2次元 FFT (高速フーリエ変換) を用いて周波成分を表示し、 比較した. 図 3.3 に、原画像とバイキュービック補間、waifu2x の超解像で拡大した 画像の周波数スペクトラムを示す.





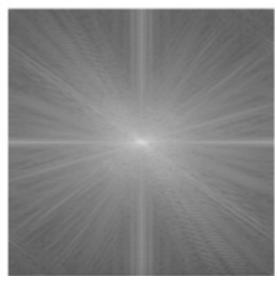


図 3.3 2次元 FFT のよる周波数成分(上:原画像,左下:バイキュービック補間,右下:waifu2xの超解像)

これらの2次元 FFT の画像は、画像の中心に近い位置では低周波成分が出ていることを表している。そして、画像の周辺部に近づくほど高周波成分が出ていることを表している。

まず,図 3.3 上の原画像は、周波数成分が全体に広がっている.バイキュービック 補間で拡大した画像は、中心周りにだけ周波数成分が出ており、全体に広がっていない.原画像と比較すると高い周波数が出ておらず、周波数成分だけで見るとバイキュービック補間は、原画像を再現できるほどの高周波成分は生成されていないことがわかる.また、waifu2x で拡大した画像は、周波数成分が画像の角まで広がっている.原画像と比較すると、原画像よりも高い周波数成分まで出ている.そのため、周波数

成分だけで見ると waifu2x は原画像を再現できるだけの高周波成分を生成できていることがわかる.

4章 バイキュービック補間と画像超解像の比較

3章の結果から、waifu2xによる拡大画像は、バイキュービック補間による拡大画像には存在しない高周波成分が生成されており、その結果、高画質化されている、と考えられる.しかし、3章の結果だけでは、waifu2xによって、画像のどの部分がどのように高画質化されたか、まではわからない.4章ではこれらを明らかにする.

4.1 小領域 PSNR の比較による画質の確認

画像超解像によって高画質化される領域を特定するために、拡大画像内の小領域ごとに PSNR を求め、それを waifu2x とバイキュービック補間で比較する。図 4.1 に示すように、waifu2x による拡大画像について、 4×4 画素の小領域ごとに原画像との PSNR を求める、PSNR を求める小領域は左上隅から 1 画素ごとにずらすことで、拡大画像全体についての PSNR 値の画像を求める、バイキュービックによる拡大画像についても同様の処理を行う。このようにして求めた 2 つの PSNR 値の画像の差分を求めることで、waifu2x とバイキュービック補間の違いを明らかにした。



図 4.1 比較手順

lena, mandrill, pepper の3画像について、PSNR値の差画像を図4.2に示す.これらの図で、灰色(画素値128)になっているところは、waifu2xとバイキュービック補間で PSNR に違いがなかったところである。waifu2xの方が PSNR値が大きい箇所は明るく、小さいところは暗く表示されている。図4.2からわかることは、waifu2xがバイキュービックよりも高画質化できている領域は、概ね、原画像中のエッジに対応していることである。また、原画像中で変化が少ない平坦な領域とmandrillにおける毛並みのように細かい絵柄の箇所は顕著な違いがない。

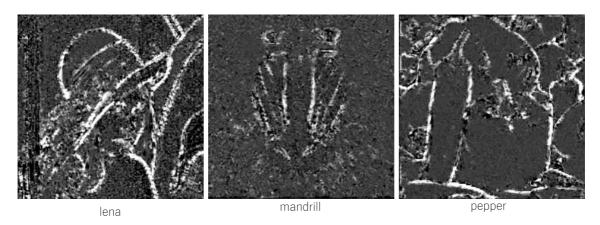


図 4.2 waifu2x とバイキュービック補間の小領域 PSNR の差画像

図 4.2 の PSNR の差画像を観察し、差が大きい箇所を図 4.3 のカラー画像に黄色のマークで示した。また、差が小さい箇所を青色のマークで示した。そして、これらの場所の画素値を、原画像、縮小画像、バイキュービック補間による拡大画像、waifu2xによる拡大画像のそれぞれで取得し、図 4.4 と図 4.5 のグラフに重ねて表示した。これらを観察して、それぞれの方法で拡大した画像の画素値の違いを詳細に分析した。



図 4.3 画素値の表示位置 (黄色:評価値の差がある部分,青:差がない部分)

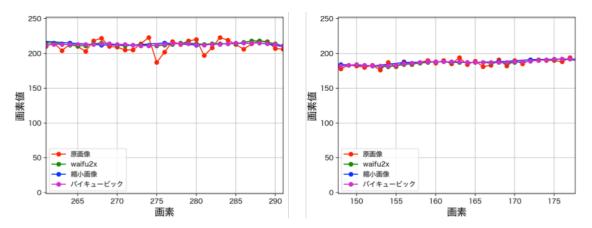
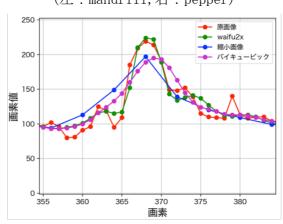


図 4.4 画質評価の差がない部分の画素値

(左:mandrill,右:pepper)



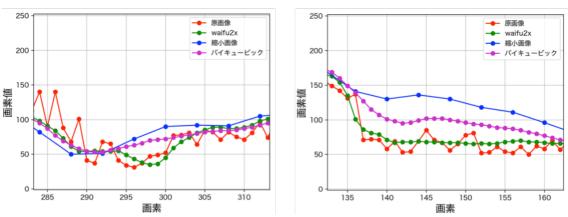


図 4.5 画質評価の差がある部分の画素値

(上:lena,左下:mandrill,右下:pepper)

図 4.4,図 4.5 のグラフは図 4.3 でマークをつけた部分の画素値の表示したものである. 赤が原画像,青が原画像の縮小画像,紫がバイキュービック補間で拡大した画像,緑が waifu2x で拡大した画像のそれぞれの画素値である. x 軸が画像の横方向の画素,y 軸が画素値を表している.

まず、画質評価の差がない部分に注目する、図 4.4 左のグラフは mandrill の髭の部分で、原画像では画素値が小刻みに何度も変化している. このような部分では、原画像の画素値をどちらの拡大方法でも再現できていないことがわかる. また、図 4.4 右のグラフは pepper の色の変化がほとんどない部分である. このような部分では、変化が小さすぎるため、どちらの拡大方法でも差がほとんどでないことがわかる.

次に、画質評価の差が大きい部分について注目する。図 4.5 上のグラフは lena の目の横の髪の部分で縦方向にエッジが続いている部分である。原画像では小さい濃淡変化の後に、大きな濃淡変化があり、その後に小さい濃淡変化がある。縮小画像の画素値の変化を見るとこれらいくつかの濃淡変化が平滑化されている。バイキュービック補間は縮小画像に沿う形で補間されているため、原画像の画素値の変化に対応できていない。waifu2x では中央部分の大きな濃淡変化が原画像に忠実な形になっている。また、その前後の小さい濃淡変化についても、原画像の画素値の変化にある程度対応できている。

図 4.5 下左のグラフは、原画像では小刻みな濃淡変化の後に、少し大きな濃淡変化があり、その後に小刻みな濃淡変化がある。縮小画像ではグラフ全体の濃淡変化が平滑化されている。バイキュービック補間は小刻みな濃淡変化と少し大きな濃淡変化のどちらも縮小画像に沿う形で補間されているため、原画像の画素値の変化に対応できていない。しかし、waifu2x は小刻みな濃淡変化は縮小画像に沿っているが、少し大きな濃淡変化では原画像に沿う形になっている。

図 4.5 下右のグラフは、原画像では小刻みな濃淡変化はあるが、滑らかに画素値が小さくなるような濃淡変化が続いている。縮小画像では全体的に原画像よりも画素値が高く、あまり変化のない濃淡変化になっている。その理由は原画像のエッジ部分であり、その上下の画素は画素値が高い画素と画素値が低い画素がある。このグラフの場合は、縮小時に画素値が高い画素を参照しているため、原画像よりも画素値が高くなっている。バイキュービック補間はその影響を受け、縮小画像と原画像の間の画素値を取るような濃淡変化で原画像とは離れているため、原画像の濃淡変化に対応できていない。しかし、waifu2x は小刻みな濃淡変化には対応できていないが、全体的な濃淡変化を見ると原画像に沿う形になっている。これは waifu2x が何らかの方法で画像のエッジ部分を特定しているからだと考えられる。この結果からも lena 画像のエッジのある部分と同じような結果になっていることが確認できた。

5章 結論

本研究では、画像超解像による拡大画像と、バイキュービック補間による拡大画像の画質評価値の差を可視化し、差がある部分と差がない部分の画素値をグラフで明示した。それによって、画像超解像と標本化定理に基づく画像拡大がどのような違いがあり、画像超解像にどのような性質があるかを明らかにすることを目的として行った。

その結果,バイキュービック補間では縮小画像の画素値に沿うような形で補間されていることがわかった.しかし waifu2x の超解像ではエッジがある部分で,拡大前の画素値に沿わない形で補間されており,バイキュービック補間よりも原画像を再現できていることがわかった.このことから,深層ニューラルネットワークを用いたwaifu2x の超解像による拡大方法には,エッジ部分でより高精度な補間をすることができる性質があることがわかった.

また、waifu2x がエッジ部分でのより詳細な性質を分析するために、どのような方法で画素値の補間を行っているかを明確にすることが今後の課題である.

参考文献

- [1] Seiichi Gohshi, "A new processing method for video: reproduce the frequency spectrum exceeding the Nyquist frequency," MMSys'12, Proc. of 3rd Multimedia Systems Conference, pp.47-52, 2012.
- [2] 内田 奏,「フカヨミ超解像 分野の最前線にズームイン!!」, コンピュータビジョン 最前線 Spring 2022, 共立出版, pp.35-48, 2022 年 3 月.
- [3] C. Dong, C.C. Loy and X. Tang, "Accelerating the super-resolution convolutional neural network," ECCV 2016, pp.391-407.
- [4] Chao Dong, Chen Change Loy, Xiaoou Tang, "Image Super-Resolution Using Deep Convolutional Networks"

謝辞

本論文を作成するにあたり、学部長という忙しい立場の中、丁寧な御指導を賜りました蚊野浩教授に感謝いたします.

付録 本研究で開発したプログラム

開発環境:

プログラム名: image_pixelvalue.py

内容: 2 枚の拡大画像と原画像を読み込み、拡大画像内の 4×4 の小領域ごとに1 画素 ずつずらし、画像全体の PSNR を求める。その後、2 枚の拡大画像の PSNR の差を小領域ごとに求め、その値を画素値とした画像を生成するプログラム。