

A Patch-Based Neural Network Super-Resolution for Low-Delay Real-Time Processing

メタデータ	言語: jpn 出版者: 公開日: 2020-01-09 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	http://hdl.handle.net/2297/00056487

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.



学位論文要旨

低遅延リアルタイム処理に適した
パッチ型ニューラルネットに基づく
超解像に関する研究

A Patch-Based Neural Network Super-Resolution
for Low-Delay Real-Time Processing

金沢大学大学院 自然科学研究科
電子情報科学専攻

学 籍 番 号 : 1624042006

氏 名 : 青木 玲央

主任指導教員名 : 今村 幸祐

Abstract

In recent years, getting higher resolution of images and videos has become a trend of the times, and particularly in the industrial field such as endoscopic surgery and automatic driving, low delay real time super-resolution is often required. On the other hand, in the field of recent super resolution research, many of high-performance deep learning methods are reported. However, due to the most of them are composed of convolutional neural networks (CNN), there are problems about cost and process-delay to realize low-delay real-time processing in industry.

In the present paper, we propose a novel super resolution method based on patch based deep neural networks (PDNN). Specifically, to accomplish the real-time low-delay processing, patch-based fully connected neural networks are designed for converting low-resolution quality images into high-resolution quality image. In addition, I introduce also a unique scheme to match the positions of each neuron to interpolation pixel positions for improving the performance.

In the numerical simulations, it was confirmed that the cost performance of proposed method is better than that of conventional CNN based method such as SRCNN (Super-Resolution via Convolutional Neural Network) and VDSR (Very Deep Super Resolution). And it was also confirmed that the matching the positions makes better performance. In addition, the potential performance of the proposed method and cost trade off was also discussed.

As a conclusion, the proposed method can be said that it is a suitable super-resolution system for realizing low-delay and real-time processing. And when the reconstruction performance level is almost same as SRCNN 9-1-5, the calculation cost can be reduced to about 2%, and the memory cost can be also reduced to about 11% against the SRCNN.

1. 序論

2000年代前半から始まった映像表示機器産業における高解像度化は時代の潮流であり、現在では4K、8Kといったヒトの視覚限界に迫る高画質化まで到達してきている。また、精密医療を追求する内視鏡手術やAIを用いた映像解析など、産業応用における高精細画像、或いは、画像の高精細化に関する需要は益々高まる傾向にあり、撮像系だけでなく表示系を含めたシステム全体での取り組みが行われている。特に、内視鏡手術や自動運転などの産業応用の現場ではリアルタイム性が重視され、遅延なく画像を高精細化する超解像技術が求められている。一方で、学術分野における超解像研究では、Deep Learningの発展によって機械学習型超解像の復元性能が近年急速に高まってきており、1枚型超解像における中心的な立ち位置に変わってきている。しかしながら、それらの多くは畳み込みニューラルネットワーク(CNN)によって達成されているため、4K / 8Kなどの高解像度映像表示システムにおいて、低遅延・リアルタイム処理を実現するには、コストや遅延等の面で課題がある。

本研究は高解像にも対応可能でかつ低遅延リアルタイム処理が実現可能なニューラルネットワーク学習型超解像システムの実現が目的である。具体的には、パッチ別処理に対応した全結合型ニューラルネットワーク構造による超解像システムを構成することで、CNNと異なり受容野を限定しながら低コストでDeepな推定を可能にするシステムを提案した。提案法では、補間画素位置を考慮した学習条件を設定することで、学習時のばらつきを抑え、高い復元性能を達成した。

2. 低遅延リアルタイム処理の課題と関連研究

まず、本研究で目指す超解像システムの枠組みについて説明する。図1は、一般的な映像表示機器を想定した映像表示システムにおける画像処理部の挿入位置を示している。図の通り、画像処理部の望ましい配置はフォーマット変換後となる。これは様々な信号フォーマットへの対応が不要になることや、顧客のニーズに応じた処理の切り替え、或いは取り外しが容易となる利便性が得られるためである。また、超解像処理の観点からは、拡大処理は予めScalerで行われるため、画質の精細化としての補正が超解像処理として画像処理部で行われることになる。これは、映像ソース側で予め拡大された画像にも適用できる

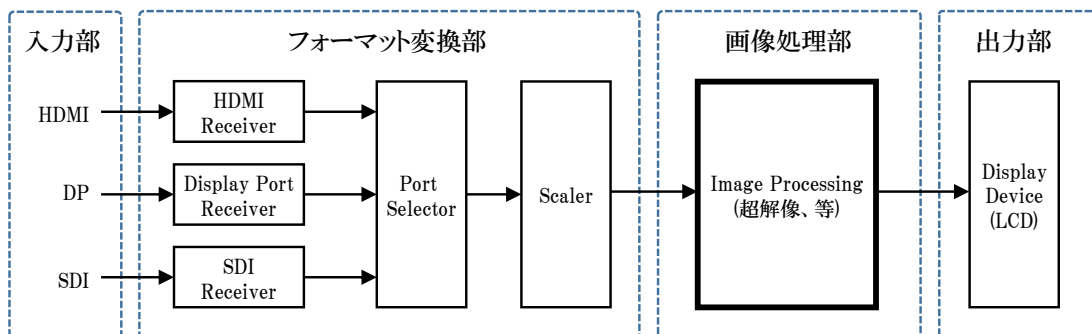


図1 映像表示システムの内部構成

という利点にもつながり、処理前後で解像度が変わらない処理が超解像アルゴリズムの必要要件ともなる。また、遅延に関する概念も表示機器では重要で、図2で示したようにラインオーダーの遅延しか許容されないのが一般的である。以上の要件は超解像システムが低遅延リアルタイム処理可能かどうかを判断する上で重要であり、アルゴリズム構築の視点では、以下2点が満たされる必要がある。

低遅延リアルタイム処理の実現条件：

- ① 出力画素値の計算が一定時間（ラインオーダーの遅延）で完了できること。
- ② 出力画素値の計算に必要な事前情報がラインオーダーの範囲で収まっていること。

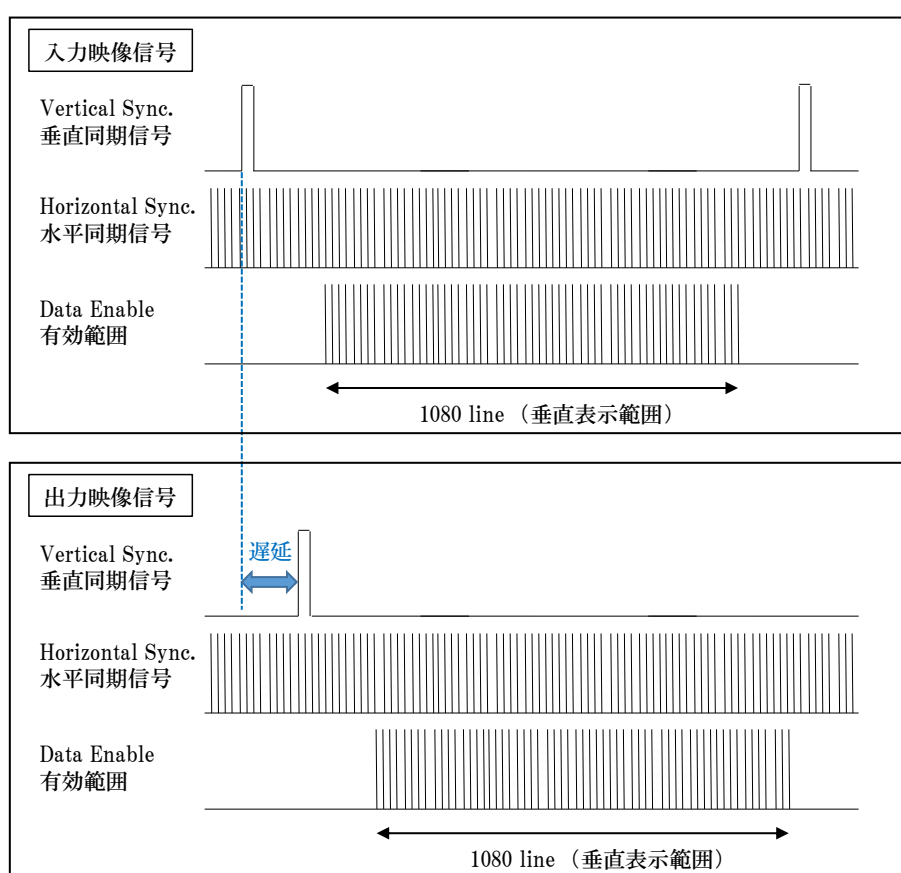


図2：低遅延リアルタイム処理（制御信号のみ表示）

なお、4K / 8K といった高解像度映像に現実的に対応していくには、並列演算可能なデバイスを利用する必要がある。本研究では産業分野で一般的に使用される ASIC (Application Specific Integrated) や FPGA (Field-Programmable Gate Array) といった自由度が高いハードウェア素子を前提にアルゴリズムの構築を検討している。

次に、上記、低遅延・リアルタイム処理の実現という観点で、近年の発展が目覚ましい機械学習に基づく超解像手法について俯瞰し、現状の課題を述べる。

任意画像に対する学習型超解像の枠組みを確立させた **Example-based Super Resolution** [1] では、対となる低解像度 (LR) 画像と高解像度 (HR) 画像を事例として蓄積したデータベースを構築し、復元時には類似する LR 画像を参照して HR 画像を推定するといった手法が提案された。この事例参照型超解像を低遅延リアルタイム処理の実現という本研究の観点で見ると、処理遅延が短いパッチベースでかつ、固定時間で処理を終えられる **one-pass** 処理という利点を持つ一方で、高性能を得るには膨大な事例が必要であり、かつ、その事例の探索コストもまた膨大になるといった課題がある。

その後、**Sparse coding Super Resolution (ScSR)** [2] では、LR 画像と HR 画像のそれぞれをスパース基底の重ね合わせで表現し、LR 画像から HR 画像への変換は LR 画像に対応したスパース基底から HR 画像に対応したスパース基底への変換で実現することを提案した。これにより、事例参照型超解像に比べ小規模なデータベース (基底ベクトル群) でも幅広い画像パターンに対応することが可能となった。一方で、LR 画像をスパース基底の和で表現するためには **OMP** 法 [3] などの反復演算によってスパース係数を算出する必要があり、そのコストが課題である。

そして、一般画像認識の分野で **Deep Learning** の発展が盛んになった頃、超解像としての機能を **End-to-End** 学習によって実現したのが **SRCNN** [4] であった。SRCNN は復元に使用する辞書ではなく、LR 画像から HR 画像への変換そのものを学習しているという点が異なっており、ネットワーク構造を決定した段階で復元処理に必要な計算量が確定される利点を持つ。しかしながら、多層の畳み込み演算は SRCNN を 4K で実現するには 28 TFLOPS と計算量が非常に多く、産業応用上の課題がある。

3. パッチベース型 DNN による超解像システム

図 3 は本論文の提案手法である SR-PDNN の全体像を示している。システム全体の入力としては、既にバイキュービック補間で拡大された低解像度品質の入力画像 (input LR

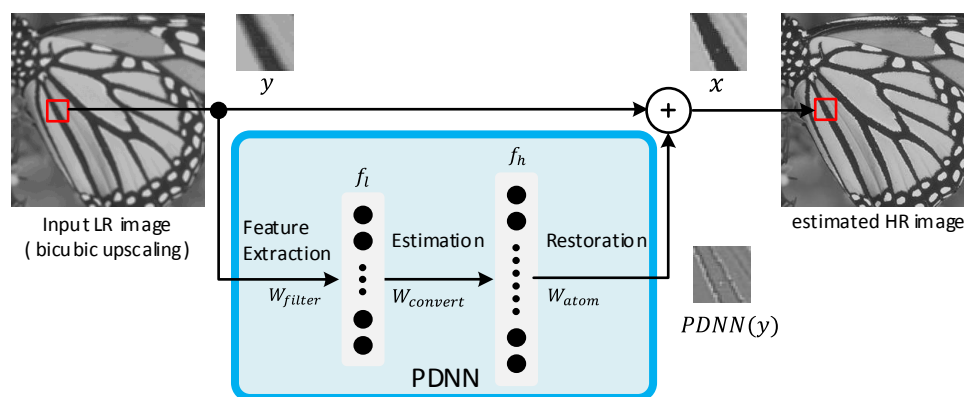


図 3 : SR-PDNN の構成

image) があり, 出力は推定された HR 画像 (estimated HR image) となる. ここで, PDNN は画像をパッチ分解された結果を入出力する全結合型のニューラルネットであり, 次の式 (1) から式 (3) によって定義される 3 層構造を持つ.

第 1 層 (特徴抽出部) :

$$Features(y) = soft\ sign(W_{filter} \times y) \quad (1)$$

第 2 層 (推定部) :

$$p(W_{atom}|y) = soft\ sign(W_{convert} \times Features(y)) \quad (2)$$

第 3 層 (復元部) :

$$PDNN(y) = W_{atom} \times p(W_{atom}|y) \quad (3)$$

ただし, y は入力である LR 画像パッチ画像を現す. また, $\theta \triangleq \{W_{filter}, W_{convert}, W_{atom}\}$ は PDNN の入出力特性を決定する重み行列であり, 以下の式 (4) で表される損失関数によって最適化される.

$$loss(\theta) = \frac{1}{N_s} \sum_{n=1}^{N_s} \|PDNN(y_n; \theta) - (x_n - y_n)\|^2 \quad (4)$$

ただし, y_n は n 番目の学習サンプルにおける LR 画像, x_n は n 番目の学習サンプルにおける HR 画像であり, N_s は学習用サンプル画像セットの総数を表す. また, 学習に使用されるパッチ画像は拡大率に応じたステップ幅で切り出しを行い, 補間画素位置と各ニューロンとの対応関係を一致させる. これは, 図 4 に示したように拡大時に生じるジャギーは拡大率の周期で発生し, かつ, 拡大補間時の補間パターンによって定まるためである.

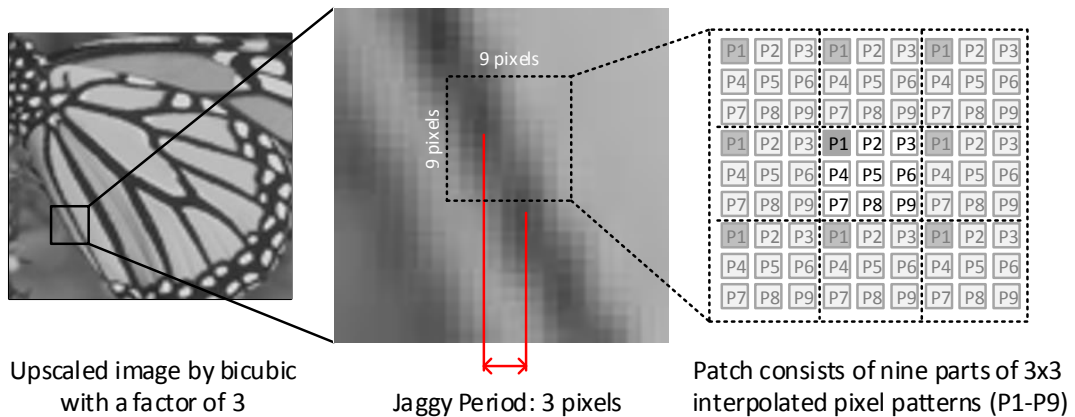


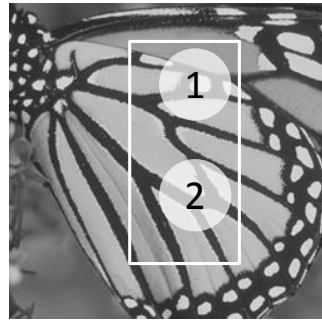
図 4: 拡大画像に含まれるジャギーと補間画素位置の関係 (拡大率:3)

4. シミュレーションによる性能評価と解析

図5は学習データを91 images [2]にした場合における復元結果画像の画質比較である。図中 (a) は Set5 [5] に含まれる butterfly, (b) は Set14 [6] に含まれる comic, (c) は拡大率 4 の butterfly の各復元結果の一部を切り取ったものであり, (d) は拡大率 3 の comic の各復元結果の一部を切り取ったものである。注目点として, 円領域 1 ではリングングが, 円領域 2, 3 ではジャギーが現れやすい領域である。これらを比較すると, ScSR, SRCNN, Ours (SR-PDNN) はいずれも LR 画像のボケを軽減する効果は見られるが, ScSR にはかなり多くのジャギーやリングングが様々な輪郭線の部分を中心に残っていることがわかる。また, SRCNN の画質も ScSR に比べてジャギーの発生エリアは抑えられているものの, 未だ多くの場所に残っている。対して, 提案手法ではわずかにジャギーは残っているものの, 殆どの箇所でジャギーやリングングが軽減し, 改善していることがわかる。

図6はCNNによる復元性能と提案手法であるPDNNによる復元性能について, コストパフォーマンスを比較している。横軸は積和演算量を表し, 縦軸は推定された出力画像と入力画像のPSNRの差分を示している。このとき, 図中で左上に位置するほどコストパフォーマンスが高いことになる。PDNN手法は明らかにCNN手法と比べて低いコストで同等の性能を実現できており, コストパフォーマンスが高い手法であるということがわかる。より具体的にはSRCNN(9-1-5)相当の復元性能であれば, SR-PDNNの場合, 約2%の積和演算量と約11%のメモリコストで到達することが確認されている。

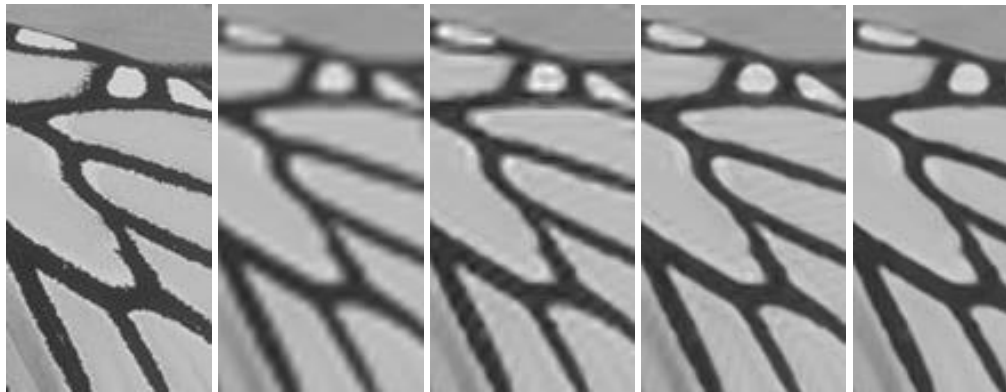
図7は, 学習時のパッチ画像切り出しステップ幅を3に固定した場合において, 復元時のパッチ切り出しステップ幅を1~9まで変化させた場合における復元性能の比較を示している。最も性能が高いのは復元時のステップ幅を3にした時であり, 次に, 6, 9と続く。これらのステップ幅はいずれも補間画素位置が一致する場合を示しており, 補間画素位置を一致させる優位性が確認できる。なお, ステップ幅3が最も高い結果となるのは, パッチの重ねあわせによる平均化効果が高いためと考えられる。



(a)



(b)



Original

Bicubic

ScSR

SRCNN

Ours

(c)



Original

Bicubic

ScSR

SRCNN

Ours

(d)

図 5：復元結果画像の画質比較

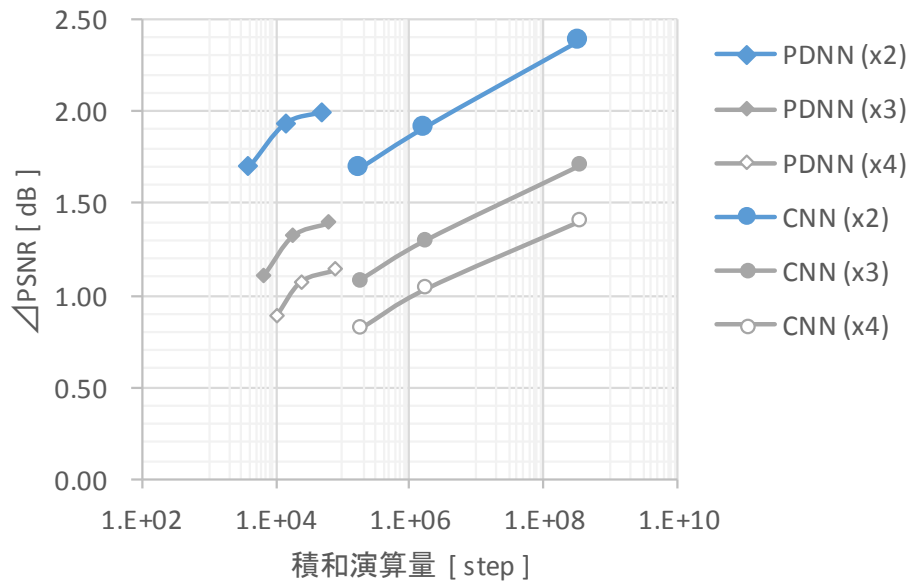


図 6：積和演算量に対する性能比較

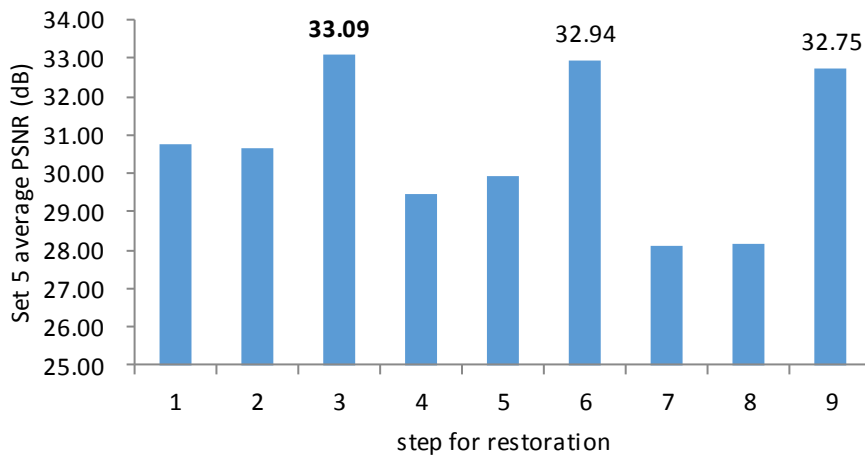


図 7：Set5 での PSNR 比較結果

結論

本研究では、4K/8K などの高解像度にも対応でき、かつ、低遅延リアルタイム処理が実現可能な超解像システムの構築を目的とした。提案手法である SR-PDNN では、パッチ処理に基づいた全結合型ニューラルネット構造をとることで、低遅延性とリアルタイム性を確保した。また、補間画素位置に応じたパッチ切り出しによる学習ばらつきの低減を図ること推定精度の向上を行い、結果として、従来の畳み込みニューラルネット型超解像手法と比較して、優れたコストパフォーマンスを実現した。具体的には、SRCNN 9-1-5 程度の性能であれば、約 2% の積和演算量と、約 11% のメモリコストでの実現が可能となった。

参考文献

- [1] W.T. Freeman, T.R. Jones, and E.C. Pasztor, “Example-based super-resolution,” *IEEE Computer graphics and Applications*, vol.22, no.2, pp.56–65, 2002.
- [2] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image super-resolution as sparse representation of raw image patches,” *Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, pp.1–8, IEEE, 2008.
- [3] Y.C. Pati, R. Rezaifar, and P.S. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers (ACSSC 1993)*, pp.40–44, 1993.
- [4] C. Dong, C.C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” *European Conference on Computer Vision (ECCV 2014)*, *Lecture Notes in Computer Science*, vol.8692, pp.184–199, Springer, 2014.
- [5] M. Bevilacqua, A. Roumy, C. Guillemot, M. Alber. “Low-complexity single-image super-resolution based on nonnegative neighbor embedding,” *Proceedings of British Machine Vision Conference 2012 (BMVC 2012)*, 2012.
- [6] R. Zeyde, M. Elad, M. Protter, “On single image scale-up using sparse-representations,” *Proceedings of 7th International Conference on Curves and Surfaces 2010*, pp. 711-730. Springer, 2010.

学位論文審査報告書（甲）

1. 学位論文題目（外国語の場合は和訳を付けること。）

低遅延リアルタイム処理に適したパッチ型ニューラルネットに基づく超解像に関する研究

(A Patch-Based Neural Network Super-Resolution for Low-Delay Real-Time Processing)

2. 論文提出者 (1) 所属 電子情報科学 専攻

(2) 氏名 青木 玲央 (あおき れお)

3. 審査結果の要旨（600～650字）

当該学位論文に関し、令和元年7月31日に第1回学位論文審査委員会を開催した。同日に口頭発表を実施し、その後第2回審査委員会を開催した。慎重審議の結果、以下の通り判定した。なお口頭発表における質疑を最終試験に代えるものとした。

本論文は、低遅延でかつリアルタイム処理に適したパッチ型ニューラルネットワークによる超解像システムに関する研究をまとめたものである。当該システムにおける超解像では、低解像度品質画像から高解像度品質画像への変換処理を、パッチベースの全結合型ニューラルネットワークにより構築している。その際、対象画像のパッチ分解時に、拡大率に応じた位置からの切り出しを行うことにより、各パッチにおける補間画素位置と処理ニューロンの対応関係を一致させ、超解像性能の向上を実現している。結果として、従来のSRCNN (Super-Resolution via Convolutional Neural Network) と呼ばれる超解像手法の9-1-5層と同等の復元性能とした場合、約2%の演算量と約11%のメモリコストによる超解像度処理の実現を達成している。

以上のように、高速化・高解像度化が求められる現在のメディア処理において、実用性と汎用性の高い超解像システムを提案し、学術的な知見を得ている。よって、本論文は博士（工学）に値すると判定した。

4. 審査結果 (1) 判定 (いずれかに○印) 合格 ・ 不合格

(2) 授与学位 博士 (工学)

5. 学位論文及び参考論文に不適切な引用や剽窃が無いことの確認

確認済み (確認方法: iThenticate による)

未確認 (理由:)