

〈一般研究課題〉 劣化画像の学習型リアルタイム
復元システムの開発

助成研究者 名古屋工業大学 後藤 富朗



劣化画像の学習型リアルタイム復元システムの開発

後藤 富朗
(名古屋工業大学)

Development of Learning-based Real-time Restoration System for Degraded Images

Tomio Goto
(Nagoya Institute of Technology)

Abstract :

In recent years, with the spread of smartphones, web services such as SNS have become more and more familiar. In recent years, with the spread of smartphones, web services such as social networking services have become familiar, and the opportunities to take pictures with cameras have increased. In particular, in the case of endoscope images, the images are taken in dark places, and the degradation components are more severe than in brightly lit rooms or outdoors. In this application, we introduce a learning restoration method that learns between sharp images without motion blur and images containing motion blur, and propose a high-performance restoration method specialized for endoscope images. The effectiveness of this method is confirmed by experiments.

1. はじめに

近年、スマートフォンの普及に伴い、SNSなどのWebサービスが身近なものとなっており、以前よりカメラで撮影をする機会が増加している。しかし、撮影された写真は周囲の明るさなどの環境や撮影する際の手ぶれにより画質が大きく劣化してしまう場合がある。また、近年では4Kや8Kといった高解像度ディスプレイなどの表示デバイスが、より身近なものとなってきているため、以前よりも高画質な画像や動画を表示させる機会が増加している。そのため、以前の低画質なデバイスでは視認できていなかった劣化成分が視認できてしまう場合がある。画質劣化の例としては、カメラの動きや焦点ずれに起因するぼけやぶれ、また撮影する際に被写体が動くことにより生じる被

写体ぶれなどが挙げられるが、これらの劣化成分は視覚的に不快であるため、画質を改善することが要求される。この劣化成分を含む画像から理想的な画像を得る場合、劣化画像1枚から復元を行わなければならないため、処理が複雑になり復元が困難となる。そのため、様々なアプローチによる研究が行われている[1,2,3].

近年の劣化画像復元手法の研究においては、幅広い種類の画像において破綻することなく復元することができる、ロバストな復元手法[4,5]が取り上げられることが多い。特に近年では、人工知能の研究課題の一つである機械学習による劣化画像の復元技術が注目を集めている。その中でも近年では深層学習による技術が盛んに行われており、画像分野以外でも様々な場面での応用が期待されている。特に、内視鏡などの医療分野では学習法による良好な結果が期待できる。内視鏡動画画像の場合、体内の奥に進むことにより生じる左右独立したぶれや臓器の収縮や弛緩により生じる部位ごとに独立した動きなど、デジタルカメラの撮影時とは違った特殊な状況に起因するぶれ劣化が考えられる。深層学習を用いた復元処理を行えば、そういった劣化動画画像に対しても良好な結果を得ることが期待できる。

劣化画像復元の分野では復元する際に深層学習によりPSFを推定して、それを劣化画像に畳みこむことによって復元を行う手法と、直接ぶれのない画像を推定する2種類の方針で研究が行われている。さらに、使用する深層学習のアーキテクチャにも様々な種類のものが存在する。しかし、どちらの手法でも復元に失敗はしないもののぼけが残ってしまう場合が多いことや学習に時間がかかること、上記の反復処理手法に比べてパラメータが非常に多く、調整が困難な点など依然として課題が多い。

そこで本稿では、Taoらの検討[6]を参考にし、学習法を用いた劣化画像復元の手法を提案する。最近の研究では、画像[7]や動画[8,9]に対して高性能劣化画像復元技術が開発されており、その中でもマルチスケールCNNを用いた、Nahらの手法[7]では、最先端の結果を出すことに成功している。彼らの手法は、小さいスケールの画像から始まり、原画像と等しいサイズまでの拡大過程の間に鮮明な画像を徐々に推定していく手法である。本稿では、非常に高い性能を実現しているこの手法を応用することにより、内視鏡画像に特化した手法を提案する。

2. 学習型ぶれ画像復元手法の概要

Taoらの手法[6]を基にディープラーニングを用いた内視鏡ぶれ画像の復元手法を提案する。

2.1 ネットワーク構造

マルチスケール画像におけるぶれ除去の新たなネットワーク構造であるScale-Recurrent Network(以下、SRN) 構造としている。SRNの提案された全体的な構造を図1に示す。入力画像からいくつかのスケールでダウンサンプリングすることにより、対応する鮮明な画像を得ることができる。本手法では、入力画像と等しいスケールの画像が、最終出力の画像として得られる。

各スケールでは、それぞれのスケールにおける理想画像を推定し、アップサンプリングの際の初期値として理想画像が使用され、理想画像はスケールごとに以下のように推定される。

$$I^i, h^i = \text{Net}_{SR}(B^i, I^{i+1}, h^{i+1}; \theta_{SR}) \quad (1)$$

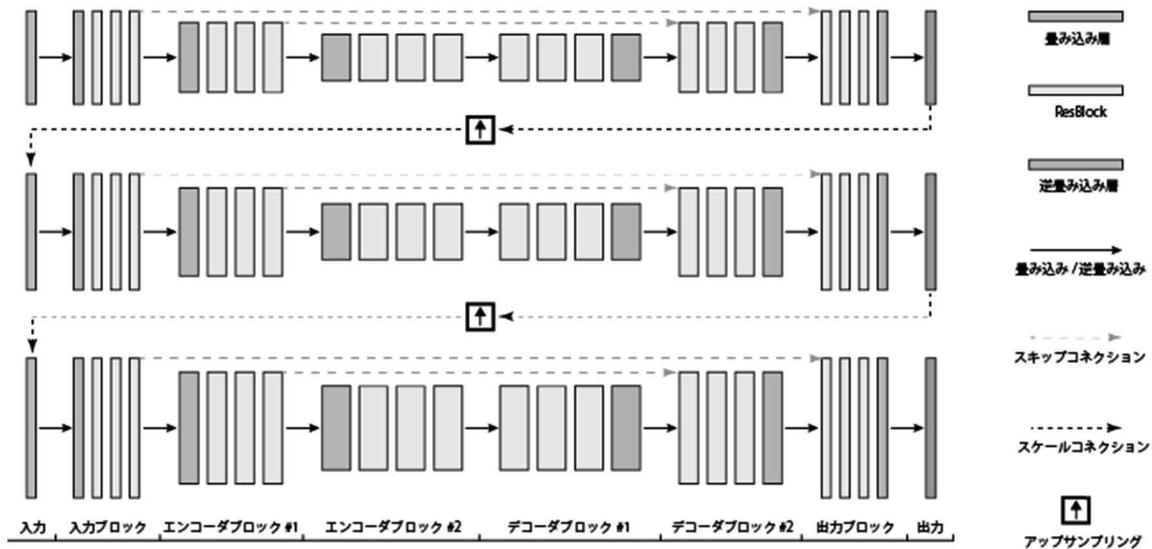


図1. SRNのネットワーク構造

ここで、 i はスケールであり $i=1$ は最大スケールを表す。 B^i 、 I^i はそれぞれ、 i 番目のスケールにおける劣化画像および推定理想画像である。 Net_{SR} は、パラメータ θ_{SR} によって定義づけられる SRN を表す。 h^i はネットワークにおいて再発性が用いられていることに起因するスケール間でのフローを表す。 h^i は前のスケールより、理想画像およびカーネル情報を取得する。(.) \uparrow は $(i-1)$ 番目から i 番目のスケールに移る際の演算子である。

2.2 エンコーダ・デコーダ

エンコーダ・デコーダネットワーク[10, 11] は、小さいサイズの画像とチャンネルを用いて、入力データを徐々に特徴マップに変換し(エンコーダ)、その後入力画像のサイズに戻す対称CNN 構造を示す。エンコーダ・デコーダに対応する特徴マップ間のスキップコネクションは、様々なスケールにおける情報を組み合わせるために広く使用されている。また、勾配伝搬の機能も持つため、処理の収束において有効である。このネットワークは、多くの処理において有効である一方、ぶれ除去処理を行う場合、各スケールにおいてエンコーダ・デコーダネットワークを直接使用することとなり、ネットワークの収束が遅くなるといった問題が生じる。この問題を改善するために導入するネットワークが、ResBlockである。使用するエンコーダ・デコーダResBlockでは、1つの畳み込み層にいくつかのResBlockが含まれており、エンコーダ・デコーダResBlockはそれぞれ対称になっている。逆畳み込み層は、特徴マップのサイズを2倍にし、チャンネルを半分にするために使用される。Taoらの手法[6]では42層で処理を行っているが、提案法として18層に変更する。また、内視鏡画像に特化した復元を実現するために、SRNではGoProテストセット画像2013枚のみを使用していたが、これに加えて5078枚の医療画像を学習画像とする。

3. 実験結果

提案手法による有効性を確認するため、実験を行った。本稿では、PSF推定による復元処理[12]を従来法とし、復元性能の比較を行う。本実験では、学習用画像としてTaoらにより提案されたGoProテストセットおよび鮮明な内視鏡画像を用いる。復元対象画像としては、内視鏡画像のみを使用する。実験結果の評価には、PSNRを用いる。実験条件を表1に示す。

客観評価の結果を図2に、復元結果の中で提案法が良好な画質の復元画像を図3, 4に示す。図2より、提案法では多くの画像で0dB以上のPSNRとなっており、劣化画像から画質が改善していることが確認できる。また、一部の画像で従来法のPSNRを下回っているものの、大半の画像では従来法を上回っていることが確認できる。次に図3, 4より、従来法では画像全体にリングングが発生しているのに対し、提案法ではリングングが発生しておらず、客観評価においても改善されていることが確認できる。また、18層で処理を行いbais処理を行うことにより、復元性能を維持しながら、処理コストを削減することに成功した。

表1. SRNと提案法における実験条件

	SRN	提案法
層数	42	18
カーネルサイズ	5×5	5×5
ストライド	2	2
活性化関数	ReLU	ReLU
バッチサイズ	256×256	256×256
学習画像	2013 枚	7091 枚

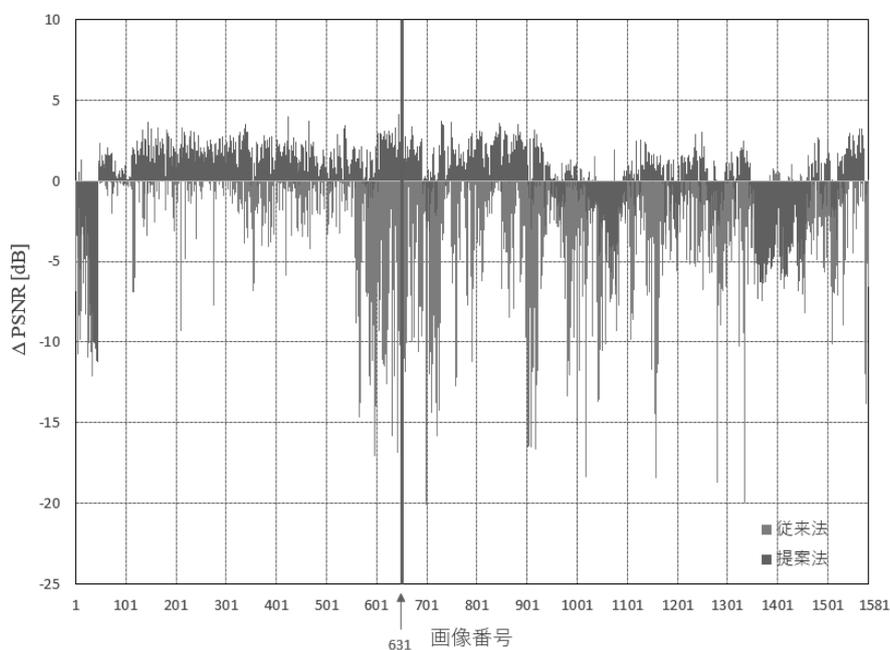


図2. 復元結果の比較

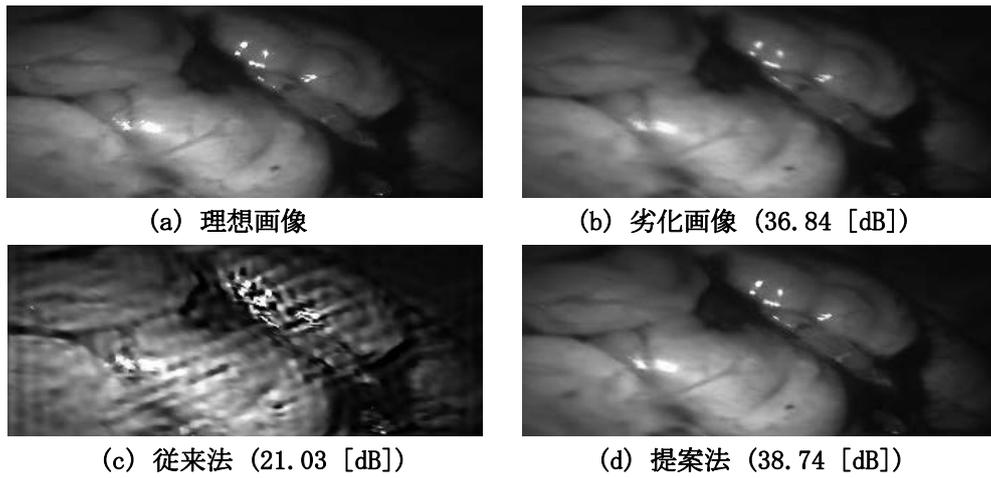


図3. 復元画像の比較

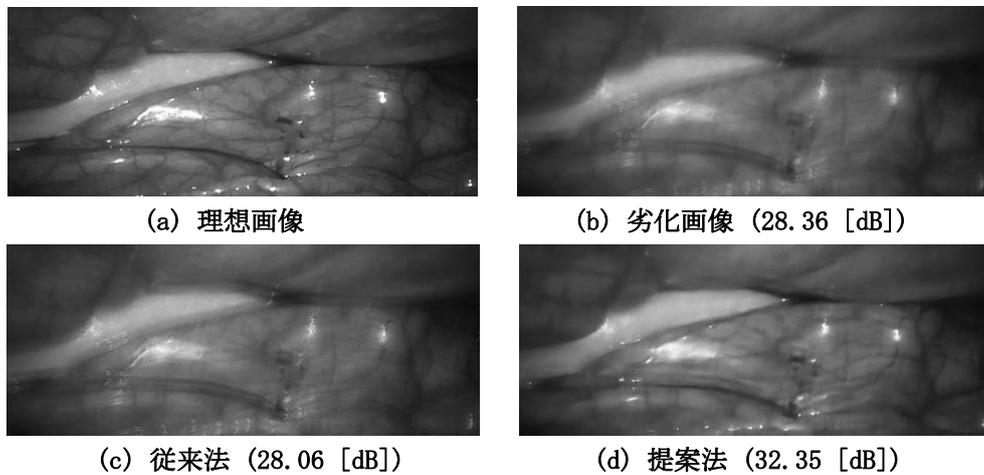


図4. 復元画像の比較

4. むすび

本稿では、内視鏡画像に対するブラインドぶれ画像復元手法の高性能化を目的として、学習法を用いた処理の提案を行い、提案法による実験を行い、客観評価手法を用いることで総合的な復元性能を評価するとともに従来法との復元性能と比較して総合的に良好な結果が得られた。従来法では、復元に成功した場合、客観評価値が非常に高く、復元画像も細部まで復元される良好な結果が得られた。しかし、復元に失敗した場合、リングングなどの劣化成分が生じてしまい、客観評価値も大きく低下してしまうという結果となった。内視鏡画像を実験の対象としているため、もともと人の体内に存在しないような画像の劣化成分の発生は、医療現場での使用を想定すると、致命的な欠陥であり、従来法の復元結果は医療画像復元技術として適切ではない。一方、今回提案した学習法を用いた手法を使用した場合、復元対象画像全体で総合的に復元性能が高く、従来法と違い安定した処理が可能である。また、客観評価値が低い画像に関しても劣化成分は部分的であり、そのほとんどが医療器具などの臓器とは関係がない部分であった。この結果より、従来と比較してロバスト性の高い、医療画像に特化した復元性能を実現した。

今後の課題として、復元画像全体でのロバスト性の向上および復元性能の更なる向上が挙げられる。現在の学習画像に加え、ぼけに特化した画像を増やすことにより、様々な画像に対応した復元

が可能になると考えられる。さらに、本論文では劣化画像として実画像ではなく数種類の動画像より作成した劣化画像を再現した画像を使用しているため、実際の内視鏡カメラ画像を使用した実験を検討する必要がある。

参考文献

- [1] 石原信人, 小松進一: "ニューラルネットワークモデルを用いた劣化像のブラインド回復法(特集最新の超解像・復元技術)", *Optics plus Ellectrics*, Vol.28, No.3, pp.258-262, 2008.
- [2] 和泉大佑, 小川貴弘, 長谷川美紀: "ぶれにより劣化した低解像度動画像の高解像度化に関する考察: Psf を用いた事後確率の高精度化に関する検討", *電子情報通信学会技術研究報告. IE, 画像工学*, Vol.110, No.421, pp.159-164, 2011.
- [3] 浅井晴香, 小山田雄仁, 斎藤英雄, 太田垣康二, 江口満男: "ケプストラム解析を用いたブレ画像のblind deconvolution", *情報処理学会研究報告, CVIM*, Vol.2009, No.26, pp.1-8, 2009.
- [4] J. Pan, Zhouchen Lin, Zhixun Su, Ming-Hsuan Yang: "Robust Kernel Estimation with Outliers Handling for Image Deblurring", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2800-2808, 2016.
- [5] T. Yue, Sunghyun Cho, J. Wang, Q. Dai: "Hybrid Image Deblurring by Fusing Edge and Power Spectrum Information", *European Conference on Computer Vision (ECCV)*, pp.79-93, 2014.
- [6] X. Tao, H. Gao, X. Shen, J. Wang, J. Jia: "Scalerecurrent Network for Deep Image Deblurring", *Proc. of CVPR*, pp.8174-8182, 2018.
- [7] S. Nah, T. H. Kim, K. M. Lee: "Deep multi-scale convolutional neural network for dynamic scene deblurring", *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.3883-3891, 2017.
- [8] T. Hyun Kim, K. Mu Lee, B. Scholkopf, M. Hirsch: "Online video deblurring via dynamic temporal blending network", *International Conference on Computer Vision (ICCV)*, pp.4038-4047, 2017.
- [9] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, Z. Wang: "Real-time single image and video super-resolution using an efficient subpixelconvolutional neural network", *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1874-1883, 2016.
- [10] X. Mao, C. Shen, Y.-B. Yang: "Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections", *NIPS*, pp.2802-2810, 2016.
- [11] O. Ronneberger, P. Fischer, T. Brox: "Convolutional Networks for Biomedical Image Segmentation", *MICCAI*, pp.234-241, 2015.
- [12] T. Nagata, S. Motohashi, T. Goto, R. Aoki and H. Chen: "A Study on Blind Image Restoration of Blurred Images Using R-map", *International Workshop on Advanced Image Technology (IWAIT)*, No.35, 2018.