

符号化ノイズを含む画像への PSNR 適応超解像

PSNR Adaptive Super-resolution for Images with Coding Noise

堀隼也[†] 龚子臣[†] 梅田聖也[†] 渡辺裕[†]
Toshiya Hori[†] Zichen Gong[†] Seiya Umeda[†] Hiroshi Watanabe[†]

中條 健[‡] 猪飼 知宏[‡] 佐々木 瑛一[‡] 伊藤 典男[‡]
Takeshi Chujo[‡] Tomonori Ikai[‡] Eiichi Sasaki[‡] Norio Ito[‡]

[†] 早稲田大学
Waseda University

[‡] シャープ株式会社
Sharp Corporation

1. はじめに

地上放送高度化に向け地上放送での 4K8K 放送の研究開発が進み始めている。地上放送の伝送容量は小さく、8K 放送実現には更なる符号化効率の向上が必要となる。これに対し、低解像画像を高解像画像に復元する超解像技術をコーデックに含める案を我々は検討している。8K 解像度の動画を 4K 解像度に縮小してから伝送し、デコーダ側で超解像を用いて元のサイズに復元する。この際、8K 映像が単純な 4K 映像の線形補間となることを防ぐため、追加の情報を送り 8K 映像を復元する。

2. 符号化画像に対する超解像の問題点

提案システムの実現には符号化画像に対応した超解像が必要となる。符号化画像は量子化パラメータ qP により品質が変化する。そのため CNN-base の超解像を用いる場合、 qP を考慮しネットワークを学習させる必要がある。しかし、 qP ごとに CNN を学習させると、超解像効果が十分に発揮できないという問題がある。データセット DIV2K[1] を $x.264$ にて $qP=36-45$ でそれぞれ符号化し、 qP ごとのデータセットを作成する。データセットごとに SRCNN を学習させ、各 SRCNN に対してそれぞれ $qP=40$ の符号化画像 10 枚を入力する。その際の結果を表 1 に示す。各画像内で PSNR が最も向上した結果を黄色で示した。入力画像の $qP(=40)$ と最も効果のあった SRCNN の学習時の qP は異なり、画像によりばらつきが見られた。

表 1 qP ごとに学習した SRCNN による画質改善効果

PSNR[dB]	input image (coded with $QP=40$)									
	im_801	im_802	im_803	im_804	im_805	im_806	im_807	im_808	im_809	im_810
INPUT	28.360	32.140	35.007	27.882	28.416	29.904	23.672	28.453	29.919	27.997
36	+0.146	+0.047	+0.144	+0.148	+0.102	+0.067	-0.349	+0.030	+0.072	+0.187
37	+0.127	+0.033	+0.111	+0.117	+0.093	+0.054	-0.369	+0.020	+0.061	+0.168
38	+0.180	+0.058	+0.108	+0.144	+0.080	-0.133	-0.850	-0.182	+0.095	+0.111
39	+0.150	-0.003	-0.051	+0.102	+0.038	-0.160	-0.872	-0.206	+0.046	+0.085
40	+0.195	+0.076	+0.157	+0.149	+0.102	-0.099	-0.830	-0.170	+0.113	+0.127
41	+0.187	+0.065	+0.144	+0.131	+0.110	-0.030	-0.641	-0.064	+0.108	+0.159
42	+0.234	+0.078	-0.004	+0.000	+0.033	-0.559	-1.549	-0.556	+0.198	-0.196
43	+0.215	+0.059	+0.110	+0.090	+0.098	-0.203	-0.975	-0.220	+0.135	+0.055
44	+0.242	+0.086	+0.124	+0.086	+0.096	-0.286	-1.176	-0.327	+0.162	+0.002
45	+0.180	+0.043	-0.058	-0.097	-0.019	-0.695	-1.708	-0.672	+0.166	-0.318
most effective training QP	44	44	40	40	41	36	-	36	42	36

3. 入力画像 qP と学習画像の qP の関係

$qP=40$ で符号化した画像でも、異なる qP の符号化画像を学習した SRCNN を用いる場合に超解像効果が高まる理由として画像の構成周波数による影響が考えられる。符号化画像は DCT 係数を量子化している。そのため、画像を構成する、値が 0 でない有意な DCT 係数が増えれば、その分誤

差も大きくなる。符号化画像ごとに SRCNN の学習を行うと、画像ごとに構成周波数の数が異なり、劣化が少ない画像、多い画像を混ぜて学習することになる。そのため、SRCNN は各 qP のデータセットに対して平均的に PSNR が向上するように学習される。これにより、同一 qP 内の全画像に対して平均的に超解像効果を持つことになるが、各画像に対して超解像を十分に発揮することができないと考えられる。またこれにより、入力画像と異なる qP で学習をした SRCNN が入力画像に対し親和性が高ければ入力画像と同じ qP で学習をした SRCNN よりも超解像効果が高くなると考えられる。よって、符号化画像に対し最適な超解像の学習方法を考える必要がある。

4. 提案手法

qP ごとのデータセットでは学習する画面内容にばらつきがあった。そこで、データセットを劣化度により分類し SRCNN を学習する。さらに、画像をパッチ分割し、画面の小領域ごとに劣化度に応じた超解像を適用する。これにより、画面内容に配慮した超解像を実現する。 $qP=40$ で符号化したデータセットを 96×96 で分割、パッチの PSNR を測定し、PSNR によりパッチを分類。PSNR ごとのデータセットを作成する。本稿では 26dB から 3dB 刻みで分類した。各データセットで SRCNN をそれぞれ学習する。そして、 $qP=40$ の符号化画像を入力し、パッチごとに PSNR を測定し、PSNR に応じたデータセットで学習した SRCNN を適用していく。表 1 と同じ画像に対し提案手法を施した結果を表 2 に示す。また、表 1 中の黄色で示した最も向上した PSNR と比較をした。多くの画像において提案手法が qP ごとの学習よりも PSNR が向上した。

表 2 PSNR ごとに学習した SRCNN による画質改善効果

PSNR[dB]	input coded image($QP=40$)									
	im_801	im_802	im_803	im_804	im_805	im_806	im_807	im_808	im_809	im_810
INPUT	28.360	32.140	35.007	27.882	28.416	29.904	23.672	28.453	29.919	27.997
PSNR-adaptive/Proposed	-0.274	+0.059	-0.034	-0.225	+0.174	+0.214	-0.025	+0.117	+0.192	+0.282
training with $QP=40$	+0.195	+0.076	+0.157	+0.149	+0.102	-0.099	-0.830	-0.170	+0.113	+0.127
training with $QP=$ BEST	+0.242	+0.086	+0.157	+0.149	+0.110	+0.067	+0.181	+0.030	+0.198	+0.187

5. まとめ

本手法により、学習画像の条件を符号化劣化と結びつけることにより、超解像効果を高めることに成功した。一方で、符号化劣化の特性と超解像効果との関連性を掘り下げ根拠あるものしていく必要がある。

参考文献

[1] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In CVPR Workshops, July 2017. 4