

低圧縮画像に着目した二値評価でのデータセットの作成 Dataset of Bi-Selection For Low-Compression Images

本田 宗一郎¹⁾ 福嶋 慶繁¹⁾
Soichiro Honda Norishige Fukushima

1 はじめに

インターネット上で画像を効率よく扱うには、画像の圧縮が重要な役割を持ち、圧縮の大きさを決定するには画質評価を行うことが必要となる。人の感覚は、主観評価で測られるべきであるが、人間が一つ一つ判断するとコストが高すぎる。よって通常は、主観評価データセットを作成し、それに基づいた客観評価指標を用いて機械的に画質評価を行う。

代表的な主観評価データセットには、LIVE [1], TID2008/2013 [2, 3], KADID-10K [4] などが挙げられる。これらのデータセットでは汎用的な劣化を計測しており、何か特別な劣化に特化したものではない。圧縮の劣化に限定すると、JPEG [5] と JPEG2000 [6] のみであり、しかも JPEG2000 はデファクトスタンダードな圧縮フォーマットではない。また Mean Opinion Score (MOS) が利用されており、これは多段階で画像の評価をする方法である。多段階で行うためには、明らかな劣化から原画像に近い劣化まで幅広く判定をする必要がある。ただし、他の劣化と違い符号化の劣化はできるだけ原画像に近いくようにと実装されているため、明らかな劣化以外はずべて良い画像のように思えてしまう。よって高精度圧縮の計測が正確にできていない

圧縮の高精度画像のように人の目から見た際の劣化の差が小さい場合は、Just Noticeable Difference (JND) が利用される。これは、2枚の画像を同じか異なるかの2値評価で判定する方法である。そこで本実験では、JNDに基づいてニアロスレス圧縮による符号化劣化の度合いを実験によって測り、主観評価データセットを作成する。

2 提案手法

JNDに基づいた本実験では、様々な圧縮画像の劣化の度合いを計測した。被験者には、原画像と圧縮画像を比較し、識別できるかの判断をさせた。実験では10枚のグレイスケール画像を使用しており、すべて Kodak の画像から選択をしている。Kodak の画像は 768 × 512 か 512 × 768 のサイズであるため、大きさを揃えるために 512 × 512 に切り取りを行っている (図 1)。また、10枚の画像を周波数成分によって、高周波成分の多い画像集合 (a, b, c, d, e) と低周波成分の多い画像集合 (f, g, h, i, j) に分けている。

それぞれの画像は、JPEG [5], WebP [7], HEIF [8] の圧縮フォーマットで圧縮をしている。また WebP では、デブロッキングフィルタがある場合とない場合の両方で圧縮をしているため、圧縮は合計で4種類行っている。また 512 × 512 の画像を 1024 × 1024 に最近傍法で拡大をすることで dpi が 1/2 の比較画像を作成した。ただし、圧縮をした後に拡大をしていることに注意してもらいたい。また、512 × 512 と 1024 × 1024 の画像でクオリティパラメータの大きさは異なる。4種の劣化を5つの劣化の段階で10枚の画像を使用しているため、解像度

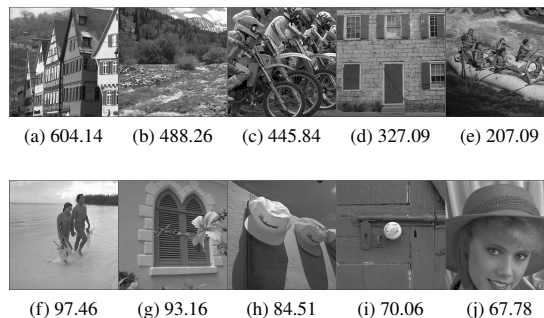


図 1: 各画素を中心とした 3 × 3 のパッチごとの分散の平均で (a) ~ (e) 高周波成分の多い画像と (f) ~ (j) 低周波成分の多い画像に分類

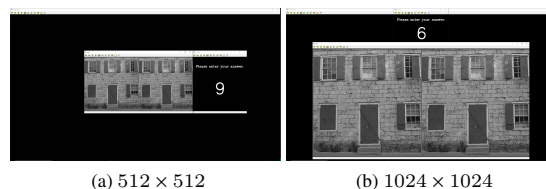


図 2: 実験環境 (デスクトップのスクリーンショット)

の違いも含め、合計で $4 \times 5 \times 10 \times 2 = 400$ の比較画像を作成した。被験者は、名古屋工業大学の掲示板を中心に募集をした 10代, 20代の非画像処理研究者とし、研究室内の学生は含めていない。それぞれの比較画像に対し 30人が判断をしているため、判断回数の合計は $400 \times 30 = 12000$ 回となる。

次に実験方法を説明する。まず、被験者に原画像と圧縮画像を並べて提示し、2値評価で比較をしてもらった。実験のインターフェースは図 3 のようになる。被験者には 2つの画像を同じだと判断した場合は“1”を、異なると判断した場合は、“0”を入力するように指示した。1回の判断の制限時間は12秒である。また、仮に同じだと感じた場合でも6秒以上かけて判断をするように依頼した。そして12秒以上経過した場合は同じと判断したものとした。これを1人あたり400回繰り返した。また、使用したディスプレイはEIZO CS 270 (27インチ, 2560 × 1440) で、視距離は視力の影響を受けないように50cmとした。実験のインターフェースは図 2 に示す。なお、実験の最初と最後での慣れによる結果の差を極力無くすために、実験が始まる前にブロックノイズやぼかしが発生しやすいポイントを被験者に教えた。

3 実験結果

図 3 から図 5 は、PSNR, GMSD [9] で評価した結果である。y 軸は識別率、x 軸は客観評価指標の値で、high のラベルは高周波成分の多い画像で、low は低周波成分の多い画像であることを示している。ただし図 5 の (a) のみクオリティパラメータで色分けがされている。

PSNR (図 3) を見ると周波数成分によって 2 グループ

1) 名古屋工業大学 Nagoya Institute of Technology

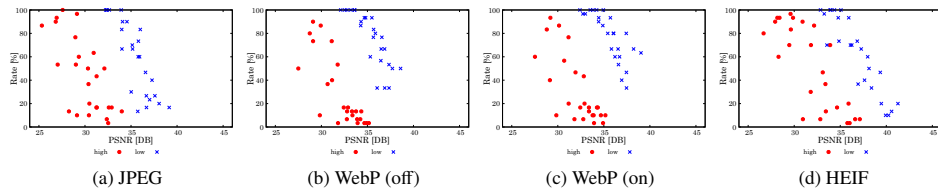


図 3: PSNR (512 × 512)

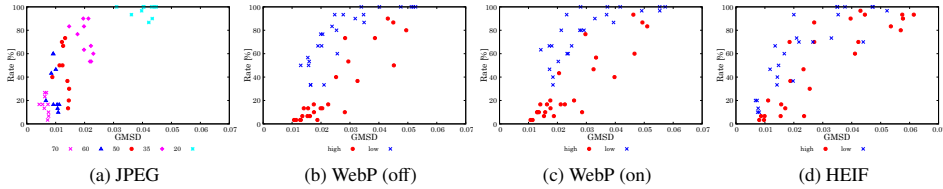


図 4: GMSD (512 × 512)

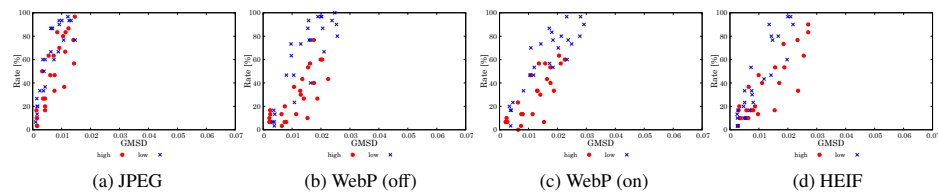


図 5: GMSD (1024 × 1024)

表 1: スピアマンの順位相関係数
512 × 512

	JPEG	WebP (off)	WebP (on)	HEIF
PSNR	-0.1839	-0.0454	0.0565	-0.4583
GMSD	0.8084	0.7815	0.7683	0.8072

	JPEG	WebP (off)	WebP (on)	HEIF
PSNR	-0.4090	-0.3483	-0.3238	-0.6167
GMSD	0.8107	0.7935	0.8378	0.8307

に分かれている。PSNR では入力画像の周波数成分の影響を大きく受けるため、ニアロスレス圧縮画像の評価を行うことが困難である。GMSD (図 4) は PSNR よりも高い相関がある。解像度を 1/2 にすると、グループ間の差は小さくなっている (図 5)。図 4 の (a) では、クオリティパラメータが 20 と 70 の値がエッジクラスターとなっているが、既存のデータセット [1, 2, 3, 4] では、この 20 と 70 の間の点が限られている。

次にスピアマンの順位相関係数を計算すると表 1 ととなる。PSNR では、JPEG と WebP は相関がほとんどなく、HEIF でも相関が弱い。これに対し GMSD は強い相関を持つ。またデブロッキングフィルタは画像が大きいと効果があり、小さいと逆効果となっている。デブロッキングフィルタは圧縮時に生じるブロックノイズを平滑化し、人が見た時の画像品質を向上させるものだが、平滑化によって模様も消えてしまっている。画像の模様は、本実験での判定において重要な要素となっている。

4 まとめ

本論文では、JND に基づいた実験により高画質圧縮の主観評価データセットを作成した。圧縮は JPEG, WebP, HEIF で dpi は 2 種類用意し、デブロッキングフィルタの有無での比較も行った。またデータセットの評価は PSNR, GMSD で行っている。

謝辞 科研費(21H03465, 21K17768)および(独)環境再生保全機構の環境研究総合推進費(JPMEERF2022M01)

により実施した。

参考文献

- [1] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
- [2] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "Tid2008 - a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, pp. 30–45, 01 2009.
- [3] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. J. Kuo, "Image database tid2013: Peculiarities, results and perspectives," *Signal Processing: Image Communication*, vol. 30, p. 57–77, 2015.
- [4] H. Lin, V. Hosu, and D. Saupe, "Kadid-10k: A large-scale artificially distorted iqa database," in *Proceedings of International Conference on Quality of Multimedia Experience (QoMEX)*, 2019, pp. 1–3.
- [5] G. K. Wallace, "The jpeg still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. 18–34, 1992.
- [6] M. Rabbani, "JPEG2000: Image Compression Fundamentals, Standards and Practice," *Journal of Electronic Imaging*, vol. 11, no. 2, p. 286, 2002.
- [7] WebP homepage, <http://code.google.com/speed/webp/>, (15 Feb 2023).
- [8] M. M. Hannuksela, J. Lainema, and V. K. M. Vadakital, "The high efficiency image file format standard [standards in a nutshell]," *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 150–156, 2015.
- [9] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, 2014.