

DCT 係数のスパース化による 画像品質評価法に適した画像圧縮方式

Transform Domain Sparsification for Quality-optimized Encoding of JPEG Images

石田 淳一¹
Junichi Ishida

チョン ジーン²
Gene Cheung

久保田 彰¹
Akira Kubota

オルテガ アントニオ³
Antonio Ortega

¹ 中央大学大学院 理工学研究科
Graduate School of Science and Engineering, Chuo University
³ 南カリフォルニア大学

² 国立情報学研究所
National Institute of Informatics

Department of Electrical Engineering, University of Southern California

1 まえがき

画像品質評価では、平均二乗誤差 (Mean Square Error: MSE) がよく使われる。MSE の評価は人間の印象と必ずしも一致しない問題がある。そこで、人間視覚システム (Human Visual System: HVS) 固有の特性へ近づけるために、様々な画像品質評価法が提案されている。

本稿では、各画像品質評価法による歪み特性と変換係数のスパース性のトレードオフの最適化を行う方式を提案する。提案方式では、非零の変換係数の個数と歪み特性を同時に最小化する最適化問題を解く。JPEG と同様に、直交変換には離散コサイン変換 (DCT) を用い、 8×8 ブロック単位での処理を行う。今回、歪み特性の評価として MSE, PSNR-HVS-M[1], SSIM[2] の 3 つの画像品質評価法を扱った。実験により、全ての場合において、JPEG よりもレート歪み特性が改善されることを示す [3]。

2 提案手法

2.1 目的関数の設定

提案手法では、 8×8 のブロック毎に、非零の DCT 係数の個数と各画像品質評価法による歪み評価の和が最小となる DCT 係数を求める。スパース化した DCT 係数のベクトル Y から、輝度信号 y は $y = \Phi^{-1}Y$ から得ることができる。ここに Φ^{-1} は DCT を行う行列である。非零の DCT 係数の個数は、 Y の 0-ノルム $\|Y\|_0$ を用いる。歪み特性は原信号 x とスパース化した信号 $\Phi^{-1}Y$ を用いて $d(x, \Phi^{-1}Y)$ で表現する。これらを用いて、解くべき最小化問題は次式のように記述される：

$$\arg \min_Y \|Y\|_0 + \lambda \cdot d(x, \Phi^{-1}Y) \quad (1)$$

ここで、 λ は二つの評価量の相対的な重要度を制御するパラメータである。歪み特性 $d(\cdot)$ の評価は、MSE, PSNR-HVS-M[1], SSIM[2] によって算出した。MSE, SSIM は直交変換より、変換領域内で生じた歪みは輝度領域での歪みと等価である。よって、変換領域で歪みの算出ができる。PSNR-HVS-M は計算過程で、輝度領域の情報も用いるため、 $y = \Phi^{-1}Y$ から求めた輝度情報を使用した。(1) 式の最小化は、変換領域内で歪み特性の許容領域 “Don't Care Region” (DCR) と “Bounding Box” (BB) を構築し、効率的に解く。

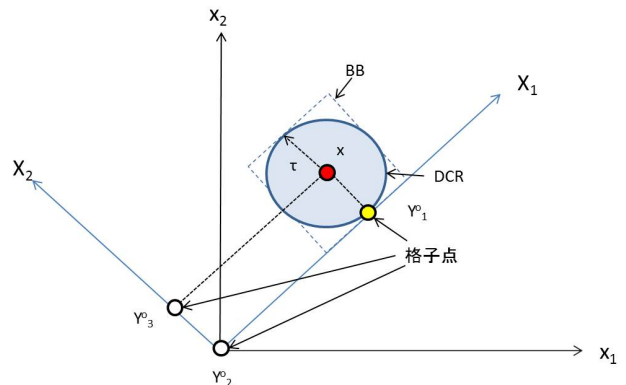


図 1 DCR による BB の構築

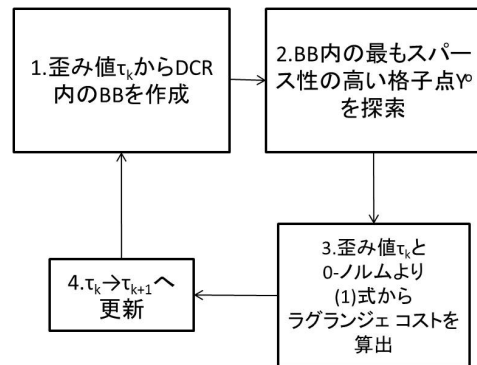


図 2 ラグランジェ コスト算出過程

2.2 許容歪みの設定

最初に、DCR を構築するためのパラメータである許容歪みの算出過程を以下に示す。原信号の DCT 係数のベクトルを X 、スパース化する DCT 係数のベクトルを Y とする。非零の DCT 係数 $Y_i = X_i$ としておく。次に、 j 番目の非零の DCT 係数を $Y_j = 0$ とし、この時に生じる歪み値を算出する。全ての非零の DCT 係数の場合で上記の計算をし、各場合の歪み値を得る。これを元に、歪み値の昇順リスト $\tau = [\tau_1, \tau_2, \tau_3, \dots, \tau_N]$ を作成する。ここで算出した許容歪み τ_k を用いて、DCR を構築する。

2.3 Bounding Box

歪み τ_k をパラメータとして構築した MSE による DCR と BB の二次元での例を図 1 に示す. 原信号が指す点を x とする. スパース性の高い点は格子点 Y_1^o, Y_2^o, Y_3^o である. この 3 点の内, 最適な点を決定する要素として, DCR を用いる. DCR 上にある格子点は一定のスパース性と歪み τ_k 以下を持ち, (1) 式の最適解としてみなすことができる. 一方, DCR 外にある格子点はスパース性はあるものの歪みについては保証することができない. よって図 1 の例では格子点 Y_1^o が最適解となる. PSNR-HVS-M, SSIM においても同様の手順だが, 歪みの評価方法が MSE とは異なり, 各周波数で DCR は変化する. これにより, DCR のみでは, 最適解が求まらない場合がある. そこで, BB を用いて DCR より大きな許容領域を構築することによって, この問題を緩和する.

以上の最適化を用いて, 各 τ_k における (1) 式のラグランジェコストを算出する過程を図 2 に示す. τ によって構成する BB 内で最もスパース性の高い格子点 Y を探索する. この時の τ_k と BB によって得られた 0-ノルムから, (1) 式よりラグランジェコストを算出する. 次に τ_k を τ_{k+1} へ更新する. 1~4 を繰り返す. この反復により, 各歪み値に応じたスパース性を持つ DCT 係数を得る. 得られたラグランジェコストの中で, 最小のものを最適解として得る.

3 実験

画像 “dancers” (画像サイズ: 618×453) の輝度成分を用いて圧縮性能を調べた. JPEG と各画像品質評価法による歪み評価で最適化した提案手法による結果を図 3 から図 5 に示す.

MSE による歪み評価をした結果を図 3 に示す (縦軸: PSNR, 横軸: 符号化量). 図 3 より全ての場合で JPEG を上回り, 最大で 1.0dB の利得を得た.

次に MSE-HVS-M による歪み評価をした結果を図 4 に示す. 図 4 による結果の比較より, PSNR-HVS-M が 53dB 付近で JPEG よりも品質が低下する傾向があった. それ以外の部分では, JPEG を上回り, 最大で 2.0dB の利得を得た.

最後に, SSIM による歪み評価をした結果を図 5 に示す. わずかではあるが, 0.005 程度の利得を得た.

4 むすび

各画像品質評価法の歪みの最小化と DCT 係数のスパース化を同時に実現するための画像圧縮方式について提案した. 提案手法により, JPEG よりも各画像品質評価法による歪み評価で MSE で最大 1.0dB, PSNR-HVS-M で 2.0dB, SSIM で 0.005 程度の品質改善を実現できた.

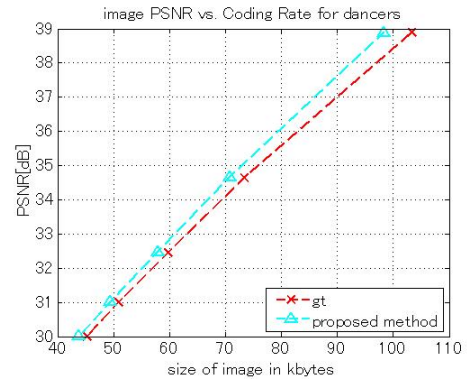


図 3 dancers(MSE)

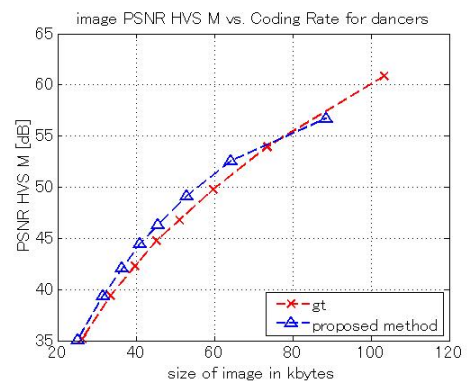


図 4 dancers(PSNR-HVS-M)

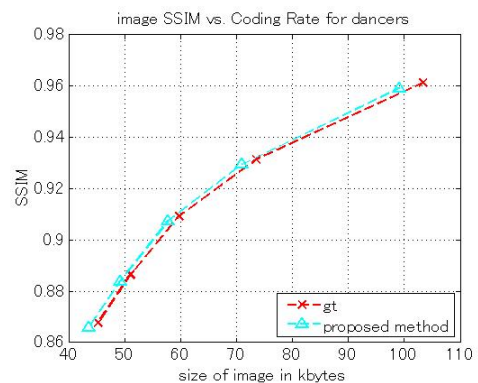


図 5 dancers(SSIM)

参考文献

- [1] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, “On between-coefficient contrast masking of DCT basis functions,” in Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, January 2007.
- [2] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: From error visibility to structural similarity” in IEEE Transactions on Image Processing, vol. 13, no.4, August 2005, pp. 600- 612.
- [3] J. Ishida, G. Cheung, A. Kubota, and A. Ortega, “Quality-optimized Encoding of JPEG Images using Transform Domain Sparsification,” MMSP 2012, Sep 17-19, 2012(to appear).