A Lossless Compression Method of Multi-scale JPEG Images via DCT Coefficients Prediction

Haohan Li[1], Zhaoyi Sun[1] and Jie Sun[1]

[1]Theory Lab Hong Kong R&D Center, Huawei Technology Co. Ltd, Hong Kong, China.

Contributing authors: li.haohan@huawei.com; sun.zhaoyi@huawei.com; j.sun@huawei.com;

Abstract

JPEG

Keywords: JPEG, DCT, DCT prediction, Lossless

1 Introduction

2 DCT

Discrete cosine transform(DCT) which is the key to the JPEG baseline algorithm has become a standard image compression method. With this technique the signal is converted from image color domain to the frequency domain losslessly. The image is divided into two-dimensinal 8x8 blocks of pixels. Each block is transformed into 64 DCT coefficients. The DCT coefficients are weights associated with the DCT basis functions, from which the block of image RGB pixels can be reconstructed.

2.1 2-D DCT and IDCT

The DCT transforms an input signal from the time domain into the frequency domain. Given an image, let $S \in \mathbb{R}^{8\times 8}$ denote the block of DCT coefficients for a given 8×8 image block, in the spatial domain, the pixel at coordinates

(x, y) is denoted S_{yx}. To transform S into an image in the frequency domain, D, we can use the following:

$$D_{vu} = \frac{1}{4}\alpha(u)\alpha(v)\sum_{y=0}^{N-1}\sum_{x=0}^{N-1} S_{yx} cos\left(v\pi\frac{2y+1}{2N}\right) cos\left(u\pi\frac{2x+1}{2N}\right)$$

where $\alpha(k) = \begin{cases} \frac{1}{\sqrt{8}}, & k = 0, \\ \frac{1}{2}, & k > 0. \end{cases}$

To rebuild an image in the spatial domain from the frequencies obtained above, we use the IDCT:

$$S_{yx} = \frac{1}{4}\sum_{v=0}^{N-1}\sum_{u=0}^{N-1} \alpha(u)\alpha(v)D_{vu} cos\left(v\pi\frac{2y+1}{2N}\right) cos\left(u\pi\frac{2x+1}{2N}\right) \tag{1}$$

Mathematically, the DCT is perfectly reversable and we are not lossing any information on pixel domain. However, the JPEG baseline algorithm will throw away the higher frequency coefficients in order to save image storage without sacrificing too much image quality. JPEG does this by dividing the DCT coefficients by a quantization matrix in order to get long runs of zeros.

2.2 Quantization Process and Reconstruction Error

Let $Q \in \mathbb{R}_{+}^{8\times 8}$ denote the quantization matrix. For the coefficient at position (v, u) we write D_{vu} with corresponding quantization step size Q_{vu}. The quantization operation is defined as

$$\tilde{D}_{vu} = \mathrm{round}\left(\frac{D_{vu}}{Q_{vu}}\right),$$

where $\mathrm{round}(\cdot)$ denotes rounding to the nearest integer. The reconstructed coefficient after inverse quantization is then

$$\hat{D}_{vu} = \tilde{D}_{vu}\, Q_{vu}.$$

The quantization error in the transform domain is defined as the difference between the reconstructed coefficient and the original coefficient:

$$E_{vu} \triangleq \hat{D}_{vu} - D_{vu} = \left(\mathrm{round}\left(\frac{D_{vu}}{Q_{vu}}\right) - \frac{D_{vu}}{Q_{vu}}\right) Q_{vu}.$$

Defining the normalized quantization residue as

$$\varepsilon_{vu} \triangleq \mathrm{round}\left(\frac{D_{vu}}{Q_{vu}}\right) - \frac{D_{vu}}{Q_{vu}},$$

we obtain

$$E_{vu} = Q_{vu}\,\varepsilon_{vu}, \qquad \varepsilon_{vu} \in \left(-\tfrac{1}{2}, \tfrac{1}{2}\right].$$

Thus, the quantization error is bounded within $\pm Q_{vu}/2$ and is determined jointly by the step size Q_{vu} and the coefficient value D_{vu}.

3 Modeling Spatial-Domain Error from Quantization Noise

Quantization introduces an error in the DCT domain that propagates into the reconstructed image through the inverse transform. Assume that each E_{vu} follows an independent Gaussian distribution with zero mean and variance σ_{vu}^2, i.e.,

$$E_{vu} \sim \mathcal{N}(0, \sigma_{vu}^2).$$

The spatial-domain reconstruction of an 8×8 block at position (x, y) is obtained via the inverse DCT (IDCT) Eq.1:

$$\hat{S}_{yx} = \sum_{u=0}^{7}\sum_{v=0}^{7} \alpha(u)\alpha(v)\hat{D}_{vu} \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right],$$

where $\alpha(k) = \begin{cases} \frac{1}{\sqrt{8}}, & k = 0, \\ \frac{1}{2}, & k > 0. \end{cases}$

The reconstruction error in the spatial domain is defined as

$$\Delta S_{yx} = \hat{S}_{yx} - S_{yx}.$$

Substituting $\hat{D}_{vu} = D_{vu} + E_{vu}$, we obtain

$$\Delta S_{yx} = \sum_{u=0}^{7}\sum_{v=0}^{7} \alpha(u)\alpha(v)E_{vu} \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right].$$

Thus, the spatial error ΔS_{yx} is a linear combination of the quantization errors $\{E_{vu}\}$, where the coefficients are determined entirely by the IDCT basis functions.

Since the quantization errors are assumed independent Gaussians, ΔS_{yx} is also Gaussian, with zero mean and variance

$$\mathrm{Var}[\Delta S_{yx}] = \sum_{u=0}^{7}\sum_{v=0}^{7}\left(\alpha(u)\alpha(v)\cos\left[\frac{(2x+1)u\pi}{16}\right]\cos\left[\frac{(2y+1)v\pi}{16}\right]\right)^2 \sigma_{uv}^2. \tag{2}$$

4 Analysis of Uniform Error Variance in the Spatial Domain

We consider the quantization error model in JPEG, where frequency-domain quantization errors E_{uv} are modeled as zero-mean independent Gaussian random variables.

Define

$$\phi_{uv}(x,y) = \alpha(u)\alpha(v)\cos\left(\frac{(2x+1)u\pi}{16}\right)\cos\left(\frac{(2y+1)v\pi}{16}\right).$$

Then from Eq.2

$$\mathrm{Var}[\Delta S_{yx}] = \sum_{u=0}^{7}\sum_{v=0}^{7}\phi_{uv}(x,y)^2\,\sigma_{uv}^2.$$

Uniform Error Variance Condition

One practical desideratum for perceptual quality in block transform is that the reconstruction error induced by quantization be spatially *uniform* within each transform block. Localized peaks of reconstruction error (large pixel-wise deviations) are salient and readily detected by human eyes, whereas spatially homogeneous, low-amplitude noise is less perceptible. Therefore we want the variance in the spatial domain to be constant across all (x,y):

$$\mathrm{Var}[\Delta S_{yx}] = C, \quad \forall (x,y).$$

This is equivalent to requiring

$$\sum_{u=0}^{7}\sum_{v=0}^{7}\phi_{uv}(x,y)^2\,\sigma_{uv}^2 = C, \quad \forall (x,y).$$

In matrix form, let $\mathbf{w}$ be the vector of length 64 containing all σ_{uv}^2, and let

$$A_{(x,y),(u,v)} = \phi_{uv}(x,y)^2.$$

Then the condition is

$$A\mathbf{w} = C \cdot \mathbf{1}_{64},$$

where $\mathbf{1}_{64}$ is the all-ones vector of length 64.

Trivial Solution

Since the 2D DCT transform is orthogonal, if we set

$$\sigma_{uv}^2 = \sigma^2, \quad \forall (u, v),$$

then each spatial position (x, y) receives the same total variance, yielding

$$\mathrm{Var}[\Delta S_{yx}] = \sigma^2 \sum_{u,v} \phi_{uv}(x, y)^2 = C.$$

Thus, the uniform solution is guaranteed when all σ_{uv}^2 are equal. This is the trivial solution.

Non-Trivial Solutions

Recall the notation 64×64 matrix A by

$$A_{(x,y),(u,v)} \;=\; \phi_{uv}(x, y)^2,$$

where rows are indexed by spatial locations (x, y) (lexicographically) and columns by frequency indices (u, v).

We performed a singular value decomposition (SVD) of A. The computed singular values (nonincreasing order) are

$$1,\; \underbrace{0.353553391}_{\times 6},\; \underbrace{0.125}_{\times 9}, \underbrace{0, \ldots, 0}_{\times 48}.$$

From these singular values we obtain

$$\mathrm{rank}(A) = 16, \qquad \dim\big(\ker(A)\big) = 64 - 16 = 48.$$

Since $\mathrm{rank}(A) = 16 < 64$, A is rank-deficient and thus has a nontrivial null space of dimension 48. A particular solution is the constant vector $\sigma = C\mathbf{1}_{64}$ (the trivial solution). The general solution set is

$$\boxed{\sigma \;=\; C\mathbf{1}_{64} \;+\; n, \qquad n \in \ker(A) \text{ (any vector in the 48-dim null space)}}$$

i.e. any null-space perturbation n added to the constant solution yields another valid solution that still produces spatially uniform variance C.

Thus if we want the spatial-domain variance at every pixel to be constant, then the condition is $\sigma_{uv}^2 \in \mathcal{S} \;=\; \big\{\sigma \in \mathbb{R}^{64} \;\big|\; \sigma = C\mathbf{1}_{64} + n,\; n \in \ker(A)\big\}$.. A straightforward calculation shows that the quantization error variance is proportional to the square of the quantization step, i.e., $\sigma_{uv}^2 = O(Q_{uv}^2)$. Next, we summerize the above conditions in a lemma and give an example to verify it.

Lemma 1 Let $A \in \mathbb{R}^{64\times 64}$ denote the linear transform mapping the quantization-domain variances $\sigma^2 = (\sigma_{uv}^2)_{u,v}$ to the spatial-domain pixel variances. Then the set of admissible variance vectors that yield constant spatial-domain variance is the affine subspace

$$\mathcal{S} \;=\; \Big\{\sigma \in \mathbb{R}^{64} \;\mid\; \sigma = C\mathbf{1}_{64} + n,\; n \in \ker(A)\Big\}.$$

Furthermore, the quantization error variance is asymptotically quadratic in the quantization step size:

$$\sigma_{uv}^2 \;\asymp\; Q_{uv}^2, \qquad (u,v) \in \{0,\dots,7\}^2.$$

We test lemma1 on 100 random images using standard JPEG quantization table and calculate the spatial domain pixel variances. The result is shown in Fig.1. The result variance is within $[3.9, 4.3]$ which is approximately uniform across the spatial domain.

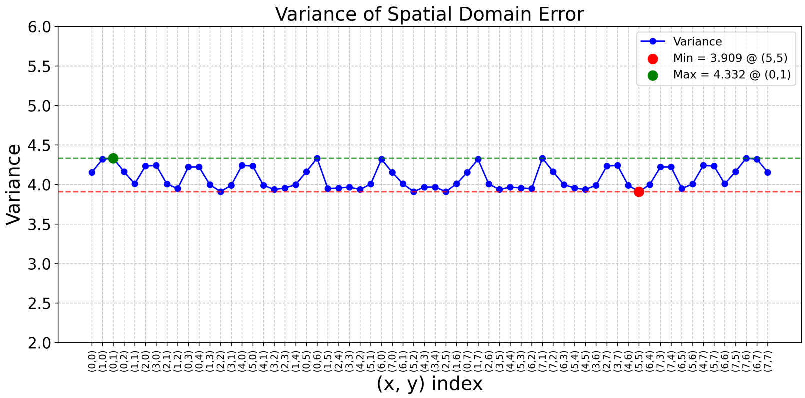

Fig. 1 Spatial domain pixel variances

Lemma 2 Let

$$\mathcal{S} \;=\; \Big\{\sigma \in \mathbb{R}^{64} \;\mid\; \sigma = C\mathbf{1}_{64} + n,\; n \in \ker(A)\Big\}$$

be the affine subspace of admissible variance vectors (here $C \in \mathbb{R}$ and $\mathbf{1}_{64}$ is the all-ones vector). Define the monotone cone

$$\mathcal{M} \;=\; \{\, t \in \mathbb{R}^{64} \mid t_1 \le t_2 \le \cdots \le t_{64} \,\}.$$

Let $M \in \mathbb{R}^{64\times d}$ be any matrix whose columns form a basis of $\ker(A)$ and set

$$B := [\,\mathbf{1}_{64}\;\; M\,] \in \mathbb{R}^{64\times(1+d)},$$

$D \in \mathbb{R}^{63\times 64}$ the forward difference matrix, $(D\sigma)_i = \sigma_{i+1} - \sigma_i,\; i = 1,\dots,63.$

Then the minimum Euclidean distance between the two sets is strictly positive and is attained. More precisely,

$$\operatorname{dist}(\mathcal{S},\mathcal{M}) \;=\; \min_{\substack{\sigma\in\mathcal{S}\\ t\in\mathcal{M}}} \|\sigma - t\|_2 \;=\; \min_{p\in\mathbb{R}^{1+d},\; t\in\mathbb{R}^{64}} \|Bp - t\|_2 \quad \text{s.t.} \quad Dt \ge 0,$$

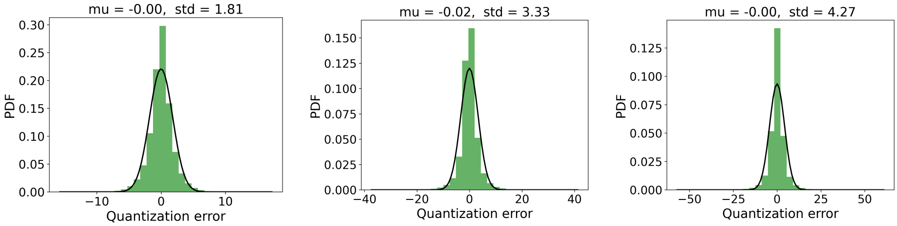

Fig. 2 Distribution of color space quantization error with different quantization step size. From left to right the quantization steps are 3, 10, 16. As the quantization step increase, the variance of quantization error will also increase.

References