A Comparison of Still-Image Compression Standards Using Different Image Quality Metrics and Proposed Methods for Improving Lossy Image Quality

Multidimensional DSP – Literature Survey

Eric Heinen

3/21/08

Abstract **–** Several techniques have been developed for modifying the JPEG standard lossy compression algorithm in order to improve image quality, and two of them are discussed in this report. Also, since these techniques have mainly been evaluated in terms of peak signal-to-noise ratio (PSNR), which is an outdated image quality metric (IQM), the topic of image quality assessment is investigated. Finally, the goals and proposed methods of this project, mainly to compare existing lossy compression schemes and investigate novel approaches, are outlined.

INTRODUCTION

The JPEG standard has been around for over a decade and has become the most widely used algorithm for lossy compression of still images. However, since the introduction of JPEG, some advances have been made towards improving the quality of images compressed by lossy techniques. Several of these advances have been made by building upon the JPEG framework, and either optimize or slightly modify certain elements of the JPEG encoder and/or decoder [3, 4, 5, 6]. In addition, JPEG2000 was introduced to try to overcome some of the perceived drawbacks of JPEG, such as blocking artifacts which can occur at high compression ratios [1].

Most of the improved JPEG compression schemes in [3, 4, 5, 6] are only evaluated in reference to JPEG via peak signal-to-noise ratio (PSNR). Even [1] evaluates JPEG2000 using PSNR, which is simply an outdated way of estimating perceptual image quality. Also, it seems that none of the improved JPEG compression schemes have been compared to each other, nor have they been compared to JPEG2000. Therefore, one of the main goals of my project is to evaluate and compare a couple of the improved JPEG compression schemes, JPEG2000 and JPEG using more appropriate measures of image quality.

IMAGE QUALITY ASSESSMENT

In order to evaluate and compare lossy compression techniques, we need to be able to automatically assess image quality degradation. Since we have access to both the distorted (compressed/decompressed) and original images, it is appropriate to consider full-reference image quality metric (IQM) algorithms.

Sheikh et. al. recently performed a statistical evaluation of ten recent full-reference IQMs [2]. First, a database of different types of images was compiled, including "faces, people, animals, closeup shots, wide-angle shots, nature scenes, man-made objects, images with distinct foreground/background configurations, and images without any specific object of interest" [2]. Then, a set of distortions was applied to each image, including but not limited to JPEG compression, white noise addition, and Gaussian blur. Next, a series of human trials was conducted to determine the degradation in mean opinion score (DMOS) between the original and distorted images. Finally, a set of different performance metrics was computed to statistically evaluate how well IQM scores correlated to DMOS scores. These performance metrics were computed for each distortion type and also averaged across the entire database. One of the performance metrics used was "standard deviations of the residuals between different IQMs and subject scores," and these values are listed in Table 1, below.

Table 1. Standard deviations of the residuals between IQMs and subject scores

	JP2K#1	JP2K#2	JPEG#1	JPEG#2	WN	GBlur	FF	All data
PSNR	14.0187	17.7418	16.0393	17.6145	12.5156	14.7832	18.4089	17.4095
JND	12.7536	13.6865	13.4591	12.4137	13.5250	11.0678	17.4220	15.2179
DCTune	16.8368	20.4284	16.1153	15.9592	15.5692	15.9808	22.3555	19.6881
PQS	13.8184	15.4409	14.5328	13.2059	13.9788	11.7365	16.6407	15.3918
NQM	13.3480	14.9805	14.2746	13.1875	12.3555	12.6605	20.2826	16.0460
Fuzzy S7	13.9155	16.2779	15.1627	15.3378	16.6603	17.4870	17.5017	19.0501
BSDM (S4)	14.7145	15.0257	14.9488	13.4766	14.2101	10.5723	15.3048	14.9935
SSIM(MS)	12.5328	13.8039	13.0988	12.3362	13.2394	10.9694	16.6708	14.7366
IFC	13.7056	14.2344	14.9091	13.4828	13.0657	10.3416	15.1184	14.4437
VIF	12.1360	13.4389	13.0252	12.2638	12.4004	10.1183	14.7365	14.0411

The IQM that performed best was VIF, presumably standing for "Visual Information," which uses an information-theoretic framework [7]. Another popular IQM that performed well uses a measure of structural similarity (SSIM) between the original and distorted images [8]. Both of these IQMs outperformed peak signal-to-noise ratio (PSNR) by wide margins. Furthermore, the IQM of [7] was shown to be statistically significantly better or equal to all of the other IQMs tested.

IMPROVED JPEG COMPRESSION SCHEMES

Human Visual System (HVS) Model

Sreelekha and Sathidevi modified the JPEG standard encoding process by integrating an explicit model of the human visual system (HVS) [3]. The block diagram for their encoder is depicted in Figure 1, below.

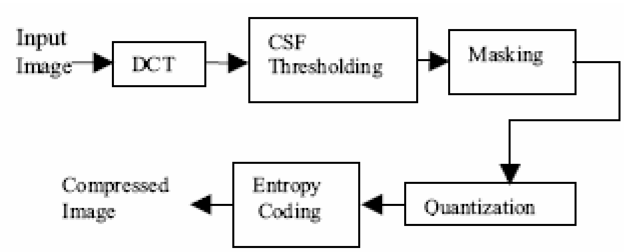

Figure 1. Improved JPEG Compression Scheme Using HVS Model [3]

The first stage of the encoder, after computing DCT coefficients, is so-called contrast sensitivity function (CSF) thresholding. The goal of this stage is to take advantage of an optical limitation of the HVS: sensitivity to spatial frequencies. As shown in Figure 2 on page 5, for any given frequency a certain amount of contrast (the contrast threshold) is needed to elicit a neurological response. Contrast sensitivity, defined as the inverse of contrast threshold, is modeled by the CSF in equation (1), below. The encoding scheme in [3] discards DCT coefficients below their corresponding CSF values, as they are perceptually insignificant.

$$H(f) = 2.6\left(0.0192 + 0.114f\right) e^{-\left(0.114f\right)^{1.1}} \quad where\ f = \sqrt{f_x^2 + f_y^2} \tag{1}$$

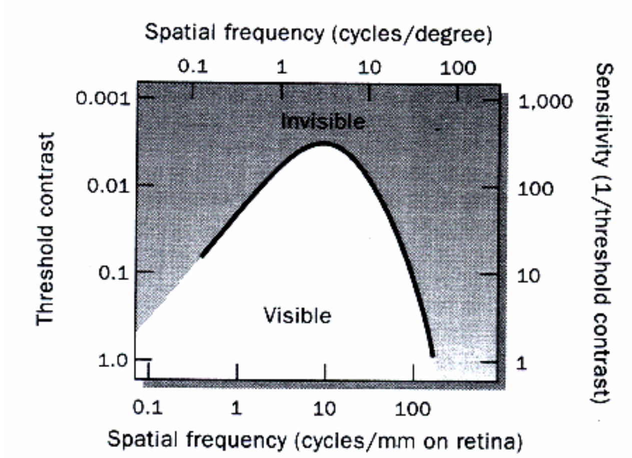

Figure 2. Contrast sensitivity function for an adult human [3]

The next stage of the encoding process takes advantage the masking phenomenon of the visual cortex, "whereby a visual signal can be diminished or hidden in the presence of other visual signals" [3]. There are two different types of masking. First, the HVS is less sensitive to local variations in brighter regions of an image. This is called luminance masking. Second, the sensitivity of the HVS to an image component "is reduced in the presence of other image components of similar frequency and orientation" [3]. This is called contrast masking. Like the CSF thresholding stage of the encoder, the masking stage defines luminance and contrast masking thresholds, and discards DCT coefficients below these thresholds.

Next, the encoder performs quantization of the DCT coefficients using "a set of uniform quantizers such that quantization noise is less than the masking threshold at each frequency" [3]. This is then followed by run-length encoding and Huffman entropy coding. Finally, perceptual quality of this encoder is experimentally shown to beat the JPEG standard on a set of test images, and as measured by both PSNR and MOS.

Joint Thresholding and Quantizer Selection

Crouse and Ramchandran developed an image-adaptive JPEG encoder (shown in Figure 3 on page 7) that performs a joint optimization of the quantization, coefficient thresholding, and Huffman entropy coding processes within a rate-distortion (R-D) framework [4]. This optimization problem is posed formally as equation (2), and reformulated using LaGrange multipliers as equation (3), below. The details of the optimization algorithm and implementation are much too involved for this report. However, it is important to note that this optimization problem is equally applicable to a JPEG2000 encoder.

$$\min_{\mathbf{T},\mathbf{Q},\mathbf{H}} D\left(\mathbf{T},\mathbf{Q}\right) \;\; \textit{subject to} \;\; R\left(\mathbf{T},\mathbf{Q},\mathbf{H}\right) \le R_{budget} \tag{2}$$

$$\min_{\mathbf{T},\mathbf{Q},\mathbf{H}} \left[J(\lambda) = D\left(\mathbf{T},\mathbf{Q}\right) + \lambda R\left(\mathbf{T},\mathbf{Q},\mathbf{H}\right) \right] \tag{3}$$

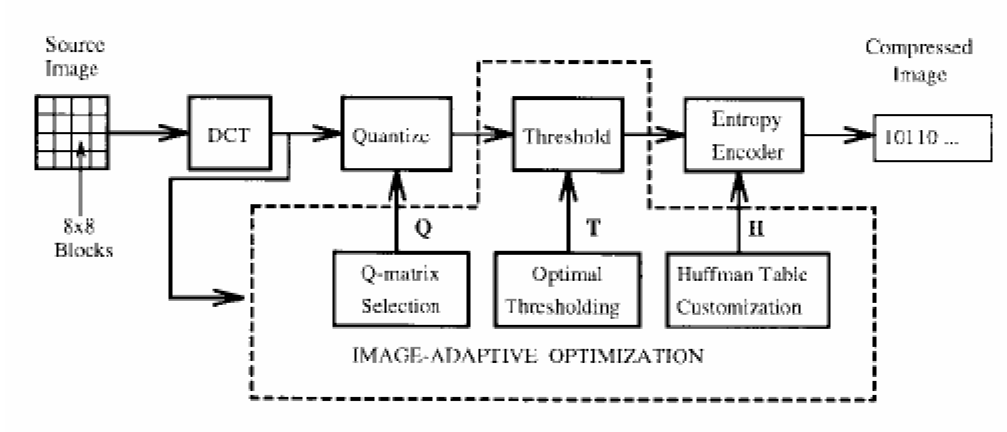

Figure 3. Image-adaptive optimization of JPEG encoder

PROJECT GOALS

The three main goals of my project are to:

- Evaluate and compare the JPEG standard, a couple of improved JPEG compression schemes, and the JPEG2000 standard using one or more IQM other than PSNR.
- Investigate the possibility of applying JPEG improvements to JPEG2000 since the two compression standards have similar frameworks.
- Investigate new methods for improving image quality of JPEG and/or JPEG2000 compression standards.

PROPOSED METHODS

For experimental evaluation I will need to:

- Select a database of different types of images. I will probably select a subset of the images used in [2], which are posted online (live.ece.utexas.edu).

- Compress and decompress each image in the database using each compression scheme being evaluated, and over a range of compression ratios. Note: There is a list of URLs for several open source implementations of JPEG and JPEG2000 encoders/decoders in [1].
- For every combination of IQM, compression (standard with or without algorithm improvement) and compression ratio, compute the average IQM of the database.
- Plot IQM (y) vs. Compression Ratio (x) vs. Compression Scheme (param).

REFERENCES

[1] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 Still Image Compression Standard," *IEEE Signal Processing Magazine*, pp. 36-58, September 2001.

[2] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," *IEEE Trans. on Image Processing*, vol. 15, no. 11, pp. 3441-3452, November 2006.

[3] G. Sreelekha and P. S. Sathidevi, "An Improved JPEG Compression Scheme Using Human Visual System Model," *Proc. IEEE International Workshop on Systems, Signals, and Image Processing*, pp. 98-101, June 2007.

[4] M. Crouse and K. Ramchandran, "Joint Thresholding and Quantizer Selection for Transform Image Coding: Entropy-Constrained Analysis and Applications to Baseline JPEG," *IEEE Trans. on Image Processing*, vol. 6, no. 2, February 1997.

[5] C. Wang, Z. Hou, and A. Yang, "An Improved JPEG Compression Algorithm Based on Sloped-facet Model of Image Segmentation", *Proc. IEEE International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 2893-2896, September 2007.

[6] H. Wang and S. Kwong, "Novel Variance Based Approach to Improving JPEG Decoding," ICIT 2005, *Proc. IEEE International Conference on Industrial Technology*, pp. 475-479, December 2005.

[7] H. R. Sheikh and A. C. Bovik, "Image Information and Visual Quality," *IEEE Trans. on Image Processing*, vol. 15, no. 2, pp. 430-444, February 2006.

[8] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multi-scale Structural Similarity for Image Quality Assessment,” *Proc. IEEE Asilomar Conf. Signals, Systems, and Computers*, November 2003.