In B. Rogowitz and J. Allebach, eds., *Human Vision, Visual Processing, and Digital Display V*, SPIE Proc. Vol. 2179, pp. 109-116, 1994.

De-blocking DCT compressed images

Albert J. Ahumada, Jr. and Rensheng Horng

Human Interface Research Branch
NASA Ames Research Center, Moffett Field, CA 94035

ABSTRACT

Image compression based on quantizing the image in the discrete cosine transform (DCT) domain can generate blocky artifacts in the output image. It is possible to reduce these artifacts and RMS error by correcting DCT domain measures of block edginess and image roughness, while restricting the DCT coefficient values to values that would have been quantized to those of the compressed image.

INTRODUCTION

Lossy image compression in the DCT domain is achieved by the quantization of the DCT coefficients. The quantization of a single coefficient in a single block causes the reconstructed image to differ from the original image by an error image proportional to the associated basis function in that block. When errors are clearly visible, the blockiness of the artifacts distinguishes them from the original image content, suggesting there may be a way of reducing these artifacts. This is an underconstrained vision problem that has no general solution. For example, the original image could be demonstrating DCT compression. Presumably the human visual system identifies the artifacts by making the assumption that the image does not have features in 8 by 8 blocks and that the image is relatively smooth within object boundaries. Our goal here is to try to reduce the blockiness without reducing the accuracy of the reconstruction.

The sum of squared differences between adjacent block edge pixels is a measure of blockiness that tends to increase with the amount of compression. A similar measure taken away from the block edge provides an estimate of the blockiness in the original image. We find that lowering the blockiness value to this estimate, while limiting coefficient change to the quantization level, can reduce both apparent blockiness and RMS (root mean square) error. The spatial frequency weighted contrast energy measures image roughness. Some additional advantage results from restoring the within-block image roughness if the blockiness reduction has increased it.

The goal of this project was the development of algorithms for improving the quality of images that have already been compressed by a JPEG-like scheme in which the image is divided up into blocks, each block is converted to DCT coefficients, and these coefficients are then quantized. In JPEG there follows a stage of lossless encoding that we ignore for the present purpose. We assume that we have the quantized coefficients and the matrix of quantization values. Our problem is to find an image that is more like the original image than the image obtained simply by performing the inverse DCT on the quantized coefficients.

Figure 1 shows five images. The center image is the original image. The upper left and lower left images have been quantized and restored without any de-blocking. The upper right and lower right show the corresponding images after our de-blocking algorithm is applied. Our method can be described as follows:

1) Measure the blockiness of the image and estimate how blocky it should be.

2) Lower the blockiness to the estimate.

3) Ensure that all DCT coefficients quantize to those of the compressed image.

4) If the within-block roughness of the image has increased, restore it to its original value.

5) Ensure that all DCT coefficients quantize to those of the compressed image.

In the rest of the paper, we make this description more precise, show quantitative results for this image and the four other images of Figure 2, and relate our work to that of others. We conclude that if this method is used when the quantization is strong enough to generate significant block artifacts, moderate de-blocking results with no increase in the RMS image error.

THEORY

The DCT image transform

The discrete cosine transform (DCT) has become a standard method of image compression[1,2]. Typically the image is divided into 8×8-pixel blocks, which are each transformed into 64 DCT coefficients. The DCT coefficients $I_{u,v}$, of an $N \times N$ block of image pixels $i_{x,y}$, are given by

$$I_{u,v} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} i_{x,y}\; c_{x,u}\, c_{y,v}\,, \quad u,v = 0,\ldots,N-1, \tag{1a}$$

where

$$c_{x,u} = \alpha_u \cos\left(\frac{\pi u}{2N}[2x+1]\right), \tag{1b}$$

and

$$\alpha_u = \begin{cases} \sqrt{1/N}\,, & u = 0 \\ \sqrt{2/N}\,, & u > 0 \end{cases}. \tag{1c}$$

DCT coefficient quantization

The JPEG compression standard[1,2] requires that uniform quantizers to be used for all of the DCT coefficients, however the quantizer step size used for each coefficient is left to the user. In uniform quantization, a coefficient is quantized by the operation

$$S_{u,v} = \mathrm{Round}\left(\frac{I_{u,v}}{Q_{u,v}}\right). \tag{2}$$

The compressed image contains both the $S_{u,v}$ for all the blocks and the $Q_{u,v}$. To retrieve the image, first the DCT coefficients are restored (with their quantization error) by

$$\hat{I}_{u,v} = S_{u,v}\, Q_{u,v}\,, \tag{3}$$

where $Q_{u,v}$ denotes the quantizer step size used for coefficient $I_{u,v}$. The blocks of image pixels are reconstructed by the inverse transform:

$$\hat{i}_{x,y} = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \hat{I}_{u,v}\, c_{x,u}\, c_{y,v}\,, \tag{4}$$

which for this normalization is the same as the forward transform. Our goal is to find better estimates of these coefficients.

Edge variance: A global measure of blockiness

Suppose i_1 and i_2 are the image values of two pixels that are next to each other in the same row or column, but are in different blocks. We assume that the blockiness of the compressed image is related to the fact that before compression, the values of i_1 and i_2 were usually similar, but they have been made more different by the quantization. We define the edge variance E to be sum of the squared differences for all such pixel pairs.

$$E = \sum (i_1 - i_2)^2, \tag{5}$$

The block edge variance E is our measure of image blockiness.

Edge variance in the DCT domain

Consider two adjacent 8×8 blocks, with block DCT coefficients $I_{1,u,v}$ and $I_{2,u,v}$. Suppose that block 2 is to the right of block 1. Let $i_{1,y}$ be the right edge function of block 1 and $i_{2,y}$ be the left edge of block 2. Then

$$\begin{aligned} i_{1,y} &= \sum_{u,v} I_{1,u,v}\, c_{7,u}\, c_{y,v} \\ &= \sum_{v} c_{y,v} \sum_{u} I_{1,u,v}\, c_{7,u}\,. \end{aligned} \tag{6}$$

Since the DCT is an orthonormal decomposition, the edge variance for these blocks is

$$\begin{aligned} E_{1,2} &= \sum_{y} (i_{1,y} - i_{2,y})^2 \\ &= \sum_{v} (\sum_{u} I_{1,u,v}\, c_{7,u} - \sum_{u} I_{2,u,v}\, c_{0,u})^2 \\ &= \sum_{v} \sum_{u} c_{0,u}{}^2 (I_{1,u,v} - (-1)^u\, I_{2,u,v})^2. \end{aligned} \tag{7}$$

This formula takes advantage of the fact that

$$c_{0,u} = (-1)^u\, c_{7,u}\,. \tag{8}$$

The weighting of the sums or differences by $c_{0,u}$, shows that for a vertical edge, errors in the high vertical frequencies have little effect on the edge variance, because the high frequency DCT basis functions have little amplitude at the block edge.

We estimate the desired value of the edge variance by computing the same measure for the pixels just inside the edge on either side and taking the average. If this estimate is less than the edge variance, we attempt to reduce the edge variance to this value. This reduction is done in the direction of the gradient of edge variance and may not be completely achieved if the minimum reduction in this direction is above the next-to-edge variance.

Adjusting the edge variance in this way only alters the edge pixels. The problem has been reduced at the boundary, but a new problem has been created inside the blocks. We attempt to reduce this problem by monitoring a measure of image roughness in the blocks.

A global measure of intra-block roughness

Our measure of the intra-block roughness is the sum of squares of the DCT coefficients weighted by their spatial frequency,

$$R = \sum_{u,v} (u^2 + v^2)\, I_{u,v}{}^2, \tag{9}$$

summed over all blocks. By weighting each component $I_{u,v}$ by its spatial frequency we obtain a measure closely related to the total edge variance inside the block. If this measure increases after reducing the edge variance, we attempt to return it to its original value by changing it along its gradient.

RESULTS

The five 64×64 images shown in Figure 2 were quantized by constant value quantization matrices, because this appeared to be the most difficult case in previous work. The quantization levels were varied from 5 to 200 in steps of 5. Figure 3 shows the graphs of the results. In the graphs, the images in Figure 2 positions top left, top right, center, bottom left, and bottom right are given lines that are solid, dotted, dashed, dash-dot, and dash-dot-dot, respectively. The compression graph shows the resulting estimated storage required in bits per pixel. The graph of the ratio of edge variance over the edge variance in the original image shows how the edge variance increases with increasing quantization. The center graph shows the ratio of the edge variance to its estimate derived from the next-to-edge pixels. It is much higher at high levels of quantization, because of the strong reduction of within-block variance. At low values of quantization, it tends to the ratio of the two measures in the original image. The lower left graph shows that reduction in edge variance was uniformly achieved. Finally, the RMS error ratio graph shows the ratio of the RMS error in the smoothed picture to that of the compressed image. It shows that the smoothing was usually obtained in conjunction with image accuracy, that the method actually improves the RMS error except in the case that the quantization is very low and the the next-to-edge estimate is also low. Fortunately, in this case, the image will only be slightly changed and there would be no apparent need for de-blocking.

DISCUSSION

The present problem is a special case of the general problem discussed by Wu and Gersho[3], optimal decoding of an image under the assumption of constrained encoding. They formalize the objective of the problem as that of finding the image $\hat{\mathbf{i}}$ that minimizes the average value of some distortion function d,

$$\min_{\hat{\mathbf{i}}} \underset{\mathbf{i}}{E}[\, d(\hat{\mathbf{i}}, \mathbf{i}) \mid \gamma \,].$$

where γ is the encoded image. They point out that if d is the sum of squared differences between pixel values, then the optimal decoder is

$$\hat{\mathbf{i}} = \underset{\mathbf{i}}{E}[\, \mathbf{i} \mid \gamma \,].$$

In an earlier paper[4], they applied this concept to the derivation of optimal additive contribution to the block for each possible level of each DCT coefficient. Their (NLI) decoder gave a 0.7 dB improvement in mean square error on a diverse 23 image training set and about 0.5 dB improvement on new images. They report apparent reduction in blockiness, but since the method was restricted to within blocks, it does not directly attack the problem. We hope to try a generalization of the method, finding the optimal additive contribution to the block and the surrounding block from each coefficient. This method of reconstruction is inherently fast and the large memory requirement can probably be reduced by parametric representations of the tables (which probably will also increase the generalizability). This method is inherently adaptable to local image statistics.

Stevenson[5] has analyzed this problem from the maximum *a posterior* (MAP) point of view. The goal is to find the image $\hat{\mathbf{i}}$ that maximizes the probability of the image given the quantized image $\mathbf{i}'$. That is, the image that maximizes

$$\max_{\hat{\mathbf{i}}} Prob[\, \hat{\mathbf{i}} \mid \mathbf{i}' \,].$$

Using a non-Gaussian Markov random field model for the image distribution, the resulting solution is the minimum of a roughness function similar to ours with the squaring operation replaced by a Huber operation in the space domain. The quantization constraint is also enforced. The method appears to strongly reduce blocking for the Lena picture, JPEG compressed at 30 to 1, but no quantitative measures are reported. The Huber function is reported to allow edges in the original image to persist through the smoothing, but the value of the Huber threshold T is not given and presumably would have to be set higher than Q so that the block edges would be smoothed. It would be interesting to test the relative advantage of using the Huber function in the space domain rather than simple squaring, which can be done as well in either domain. Another question that is raised is the relative value of specialized smoothing at the block boundaries versus a single overall measure of roughness.

Our methods and results are similar to the iterative projection method of Yang, Galatsanos, and Katsaggelos[6], They also use edge variance and the quantization constraint. They compute separate horizontal and vertical edge variances and force them to their correct values in the original image by a weighted averaging of edge pixels. They iterate these two constraints in conjunction with the quantization constraint and range constraints in both the space and DCT domains. Since the constraints are projections onto convex sets, iterating them is guaranteed to terminate, since the original image is a solution. They report a 1 dB improvement in RMS error of reconstruction and strong apparent reduction in the blockiness for the 256×256 Lena image when the PSNR for the original reconstruction was 27.9 dB. Our method differs from their method mainly in the addition of the within block smoothness constraint and the estimation of the edge variance. We find that iterating our constraints can lead to to better performance at high levels of quantization, but can degrade performance at lower levels.

SUMMARY

We have presented a method of estimating DCT coefficients from their quantized values and the quantization matrix, which are both included in the JPEG[2] standard compressed image file. This method of image reconstruction can reduce blockiness and RMS error in DCT quantized images. Although *ad hoc* in nature, it is similar to methods derived by more principled methods that make specific assumptions about image statistics.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance and support of A. B. Watson and the comments of J. Solomon. This work was supported in part by NASA RTOP 506-59-65 and NASA Cooperative Agreement NAC 2-307 with Stanford University.

REFERENCES

1. G. Wallace, "The JPEG still picture compression standard", *Communications of the ACM*, vol. 34, no. 4, pp. 31-44, 1991.

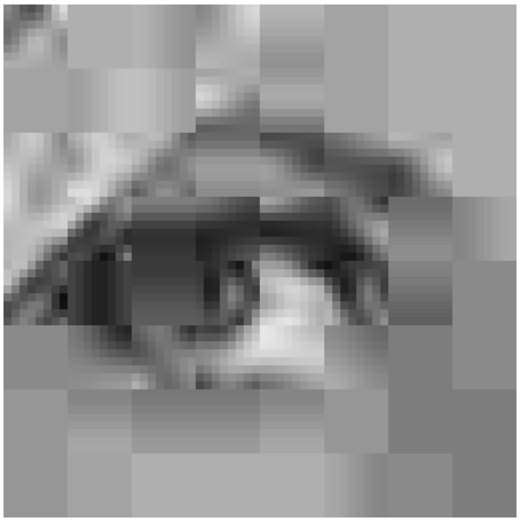
2. W. B. Pennebaker, J. L. Mitchell, *JPEG Still Image Data Compression Standard*, van Nostrand Reinhold, New York, 1993.

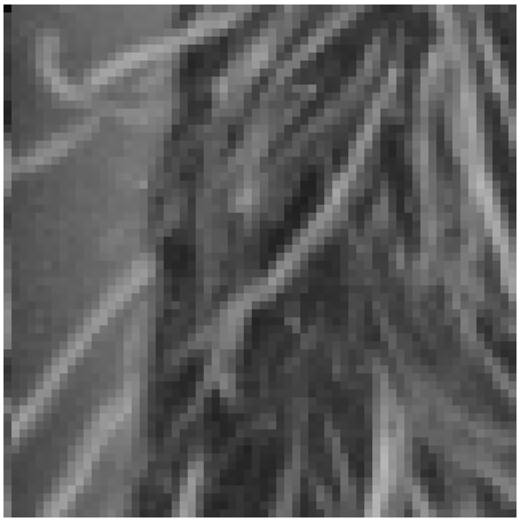
3. S. Wu, A. Gersho, "Enhanced video compression with standardized bit stream syntax", ISCASSP ’93, vol. I, pp. 103-106, 1993.

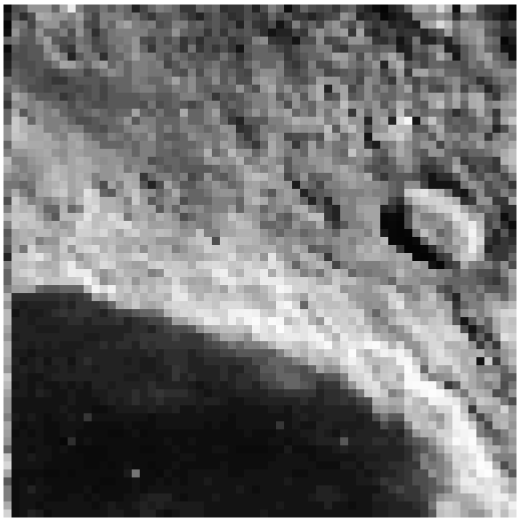
4. S. Wu, A. Gersho, "Enhancement of transform coding by nonlinear interpolation", Visual Communications and Image Processing ’91: Visual Communication vol. 1605, SPIE, Bellingham, WA, pp. 487-498, 1991.

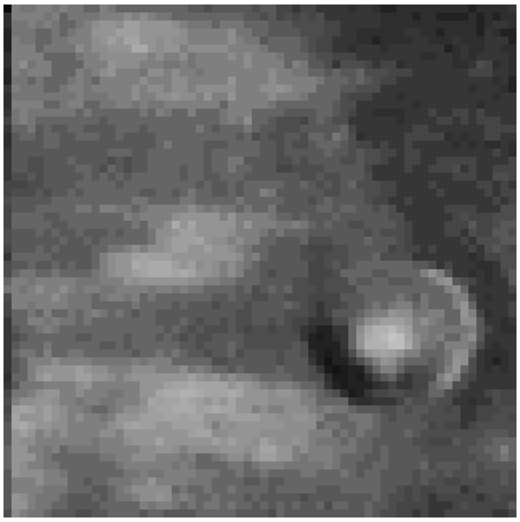
5. Robert L. Stevenson, "Reduction of coding artifacts in transform image coding", ISCASSP ’93, vol. V, pp. 401-404, 1993.

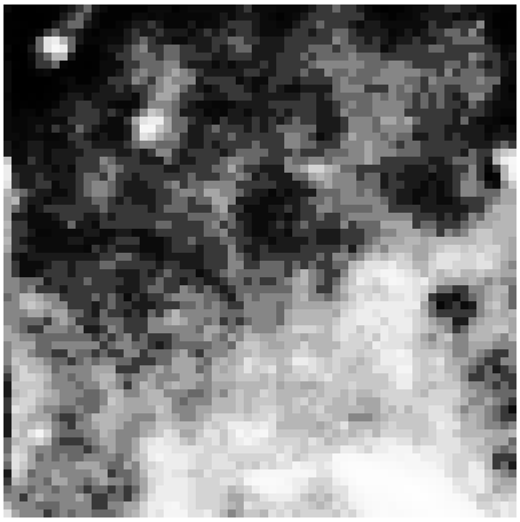
6. Y. Yang, N. Galatsanos, A. Katsaggelos, "Iterative Projection algorithms for removing the blocking artifacts of block-DCT compressed images", ISCASSP ’93, vol. V, pp. 405-408, 1993.



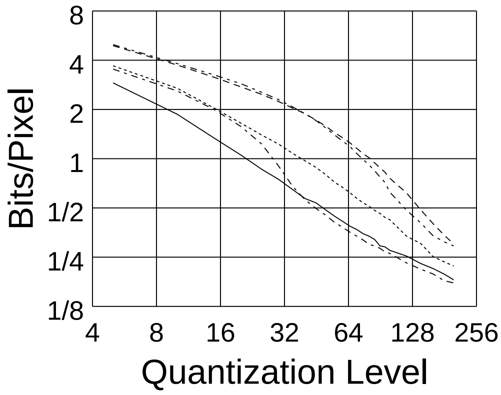




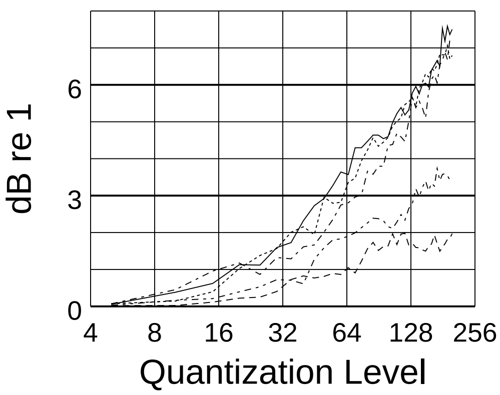




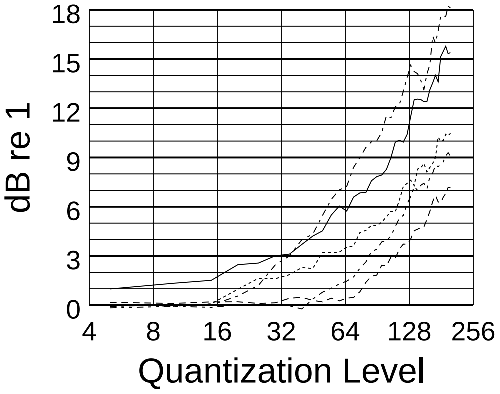
Compression



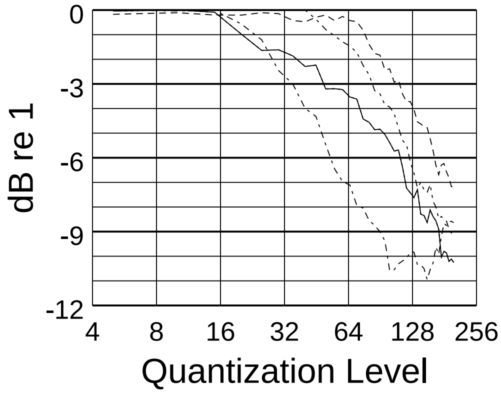
Edge Variance / Original Edge Variance



Edge Variance / Next-to-Edge Variance



Smoothed Edge Variance / Edge Variance



RMS Error Ratio