N94-32467

A VISUAL DETECTION MODEL FOR DCT COEFFICIENT QUANTIZATION

Albert J. Ahumada, Jr.
Andrew B. Watson
NASA Ames Research Center
Moffett Field, California 94035-1000
ahumada@vision.arc.nasa.gov
beau@vision.arc.nasa.gov

ABSTRACT

The discrete cosine transform (DCT) is widely used in image compression, and is part of the JPEG and MPEG compression standards. The degree of compression, and the amount of distortion in the decompressed image are controlled by the quantization of the transform coefficients. The standards do not specify how the DCT coefficients should be quantized. Our approach is to set the quantization level for each coefficient so that the quantization error is near the threshold of visibility. Here we combine results from our previous work to form our current best detection model for DCT coefficient quantization noise. This model predicts sensitivity as a function of display parameters, enabling quantization matrices to be designed for display situations varying in luminance, veiling light, and spatial frequency related conditions (pixel size, viewing distance, and aspect ratio). It also allows arbitrary color space directions for the representation of color. In a further development, we have developed a model-based method of optimizing the quantization matrix for an individual image. The model described above provides visual thresholds for each DCT frequency. These thresholds are adjusted within each block for visual light adaptation and contrast masking. For a given quantization matrix, the DCT quantization errors are scaled by the adjusted thresholds to yield perceptual errors. These errors are pooled non-linearly over the image to yield total perceptual error. With this model we may estimate the quantization matrix for a particular image that yields minimum bit rate for a given total perceptual error, or minimum perceptual error for a given bit rate. Custom matrices for a number of images show clear improvement over image-independent matrices. Custom matrices are compatible with the JPEG standard, which requires transmission of the quantization matrix.

1. INTRODUCTION

1.1 DCT image compression

The discrete cosine transform (DCT) has become an image compression standard (ref. 1, 2, 3) Typically the image is divided into 8x8-pixel blocks, which are each transformed into 64 DCT coefficients. The DCT transform coefficients $I_{m,n}$, of an N x N block of image pixels $i_{j,k}$, are given by

$$I_{m,n} = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} i_{j,k}\, c_{j,m}\, c_{k,n} \quad m,n = 0 \ldots N-1 \tag{1a}$$

where

$$c_{j,m} = \alpha_m \cos\left(\frac{\pi m}{2N}[2j+1]\right), \tag{1b}$$

and

$$\begin{aligned} \alpha_m &= \sqrt{1/N} \quad m = 0 \\ &= \sqrt{2/N} \quad m > 0 \end{aligned} \tag{1c}$$

REFERENCES

1. Rocken, Christian, and Thomas M. Kelecy. 1992. "High-Accuracy GPS Marine Positioning for Scientific Applications." *GPS World*, June, pp. 42-47.

2. Gilbert, Chuck. 1993. "Portable GPS Systems for Mapping: Features Versus Benefits." *Earth Observation Magazine*, October, pp. 43-48.

3. Birk, Ronald J., and Bruce Spiering. 1992. "Commercial Applications Multispectral Sensor System." *Small Satellite Technologies & Applications II*, SPIE Vol. 1691, 49-59.

4. Olsen, Norman T. 1993. "Update Your Database - The North American Continent Has Moved." *GIS World*, September, pp. 40-41.

5. Strange, William E., and John D. Love. 1991. "High Accuracy Reference Networks - A National Perspective." Presented at ASCE Specialty Conference - Transportation Applications of GPS Positioning Strategy, Sacramento, CA, September 18-21.

6. Grunthal, Captain Melvyn C. 1993. *INSTRUCTIONS: Mississippi HARN, 1993.*

7. Dangermond, Jack. 1992. Presentation to the First GIS-in-Business Conference, Denver, CO.

8. Center for Mapping, Ohio State University. 1991. *GPS/Imaging/GIS Project*. Columbus, OH.

9. Hill, C. L., R. J. Birk, E. Christensen, and T. Alexander. 1991. "Airborne Instrument Test System (AITS) Program." *Proceedings of American Society for Photogrammetry and Remote Sensing.*

10. Petersen, Carolyn. 1992. "Precisely San Diego." *GPS World*, April, pp. 25-29.

11. Lewis, Robert. 1992. NAVSTAR Mapping Corporation, personal communication.

The block of image pixels is reconstructed by the inverse transform:

$$i_{j,k} = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} I_{m,n}\, c_{j,m}\, c_{k,n} \quad j,k = 0 \ldots N-1 \tag{2}$$

which for this normalization is the same as the forward transform. Quantization of the DCT coefficients achieves image compression, but it also generates distortion in the decompressed image. If a single coefficient is quantized and its block is reconstructed, the difference between the original image block and the reconstructed block is the error image. This error image has the form of the associated basis function, and its amplitude is proportional to the quantization error of the coefficient. Since the inverse transform is linear, the error image resulting from quantizing multiple coefficients is a sum of such images.

1.2 The Quantization Matrix

The JPEG compression standard (ref. 1, 2) requires that uniform quantization be used for the DCT coefficients, but the quantizer step size to be used for each coefficient is left to the user. The step size used for coefficient $I_{m,n}$ is denoted by $Q_{m,n}$. A coefficient is quantized by the operation

$$S_{m,n} = \text{Round}\left[I_{m,n}/Q_{m,n}\right] \tag{3a}$$

and restored (with the quantization error) by

$$\hat{I}_{m,n} = S_{m,n}\, Q_{m,n}. \tag{3b}$$

Two example quantization matrices can be found in the JPEG standard (ref. 2). These matrices appear in Table 1 following the references. These matrices were designed for a particular viewing situation. No suggestions were provided for how they should be changed to accommodate different viewing conditions, or for compression in a different color space. Our research was initiated to provide quantization matrices suitable for compression in the RGB color representation (ref. 4). Subsequently, a theoretical framework was constructed and additional measurements have been done (ref. 5, 6, 7, 8). Here we summarize the quantization matrix design technique. It can be applied under a wide variety of conditions: different display luminances, veiling luminances, spatial frequencies, and color spaces. The basic idea of the technique is to develop a detection model that predicts the detectability of the artifacts in a perceptual space representation. This step is described in Section 2. A quantizer step size is then determined from the sensitivity of the perceptual space representation to the quantization distortion. This step is described in Section 3.

2. THE DETECTION MODEL

2.1 The Luminance Detection Model

The luminance detection model predicts the threshold of detection of the luminance error image generated by quantization of a single DCT coefficient $I_{m,n}$. We use the subscript Y for luminance, since we assume that it is defined by the 1931 CIE standard (ref. 9). This error image is assumed to be below the threshold of visibility if its zero-to-peak luminance is less than a threshold $T_{m,n}$ given by

$$\begin{aligned} \log T_{Y,m,n} &= P(f_{m,n}; b_Y, k_Y, f_Y) \\ &= \log b_Y \qquad\qquad \text{if } f_{m,n} <= f_Y. \\ &= \log b_Y + k_Y(\log f_{m,n} - \log f_Y)^2 \quad \text{if } f_{m,n} > f_Y \end{aligned} \tag{4}$$

This function P represents a low-pass contrast sensitivity function of spatial frequency. Although luminance contrast sensitivity is more correctly modeled as band-pass, we choose a low-pass function for this application. This ensures that no new artifacts become visible as viewing distance increases. A low-pass function is also convenient because purely chromatic channels are low-pass in this spatial frequency range. We will use P for the luminance and chrominance channels of our model.

The spatial frequency, $f_{m,n}$, associated with the m,nth basis function, is given by

$$f_{m,n} = \frac{1}{2N}\sqrt{(m/W_x)^2 + (n/W_y)^2} \tag{5}$$

where W_X and W_Y are the horizontal and vertical pixel spacings in degrees of visual angle. The term b_Y has three components:

$$b_Y = \frac{s\,T_Y}{\theta_{m,n}}. \tag{6}$$

The parameter s is a fraction, $0<s\leq 1$, to account for spatial summation of quantization errors over blocks. We set it to unity to model detection experiments with only one block (ref. 5) . Our summation results suggest that it should be equal to the inverse of the fourth root of the number of blocks contributing to detection (ref. 6). We suggest the value s=0.25, corresponding to 16x16 blocks. The factor T_Y gives the dependence of the threshold on the image average luminance $\overline{Y}$.

$$\begin{aligned} T_Y &= \frac{\overline{Y}^{a_T} Y_T^{\,1-a_T}}{S_0} \qquad \overline{Y} \leq Y_T \\ &= \frac{\overline{Y}}{S_0} \qquad\qquad \overline{Y} > Y_T \end{aligned} \tag{7}$$

where suggested parameter values are Y_T=15 cd/m^2, S_0= 40, and a_T=0.65.

The product of a cosine in the x with a cosine in the y direction can be expressed as the sum of two cosines of the same radial spatial frequency but differing in orientation. The factor

$$\theta_{m,n} = r + (1-r)\left(1 - \left[\frac{2\,f_{m,0}\,f_{0,n}}{f_{m,n}^2}\right]^2\right) \tag{8}$$

accounts for the imperfect summation of two such frequency components as a function of the angle between them. Based on the fourth power summation rule for the two components when they are orthogonal, r is set to 0.6. An additional oblique effect can be included by decreasing the value of r.

The parameters f_Y and k_Y determine the shape of P and depend on the average luminance $\overline{Y}$.

$$f_Y = f_0 \bar{Y}^{a_f} Y_f^{-a_f}, \quad \bar{Y} \le Y_f$$
$$= f_0, \quad \bar{Y} > Y_f \tag{9}$$

and

$$k_Y = k_0 \bar{Y}^{a_k} Y_k^{-a_k} \quad \bar{Y} \le Y_k$$
$$= k_0 \quad \bar{Y} > Y_k \tag{10}$$

where

$f_0 = 6.8$ cycles / deg
$a_f = 0.182$
$Y_f = 300$ cd / m^2
$k_0 = 2$
$a_k = 0.0706$, and
$Y_k = 300$ cd / m^2.

2.2 The Chrominance Detection Model

We now add two chromatic channels to the luminance-only model. From the large number of color spaces that have been proposed for chromatic discriminations, we have selected one close to that suggested by Boynton (ref. 9): a red-green opponent channel and a blue channel. Our channels are defined in terms of the CIE 1931 XYZ color space. The blue channel is just Z, and the opponent red-green channel O is given by

$$O = 0.47\,X - 0.37\,Y - 0.10\,Z. \tag{11}$$

This opponent channel is approximately the Boynton (ref. 9) (Red-cone)-2(Green-cone) channel. Our model now needs the threshold for quantization noise in the O and Z channels. For simplicity, we model the chromatic thresholds by

$$\log T_{O,m,n} = P(f_{m,n}; \frac{0.36\,s\,T_Y}{\theta_{m,n}}, k_Y, \frac{f_Y}{4}), \tag{12}$$

and

$$\log T_{Z,m,n} = P(f_{m,n}; \frac{3\,s\,T_Y}{\theta_{mn}}, k_Y, \frac{f_Y}{4}), \tag{13}$$

These shapes of these are in agreement with experimental results of Mullen (ref. 10), except that the slopes for the chromatic channels are found to be steeper than that of the luminance channel. The reason for keeping them the same is to prevent strong quantization of purely chromatic channels, since there is a fair amount of individual variability in the exact direction of isoluminance.

Although we previously used $\overline{Z}$ to set the level of the Z threshold (ref. 7), we are using $\overline{Y}$ here under the assumption that the average image color is close to white and hence that they are roughly equal.

Finally, we say that the errors from the quantization of a coefficient are visible, if the error in any of the three channels is visible.

3. QUANTIZATION MATRIX DESIGN

Suppose that one color dimension D in a color space linearly related to our YOZ color space is to be quantized. Let D_Y, D_O, and D_Z be the amplitudes of the errors in YOZ space generated by a unit error in D. An error generated in the D image by quantizing the m,nth DCT coefficient is then below threshold if it is less than

$$T_{D,m,n} = \min\left(\frac{T_{Y,m,n}}{D_Y}, \frac{T_{O,m,n}}{D_O}, \frac{T_{Z,m,n}}{D_Z}\right). \tag{14}$$

The D quantization matrix entries are obtained by dividing the thresholds above by the DCT normalization constants (α_m in Equation (1c)):

$$Q_{D,m,n} = \frac{2T_{D,m,n}}{\alpha_m \alpha_n}, \tag{15}$$

The factor 2 results from the maximum quantization error being half the quantizer step size.

3.1 Quantization in YC_rC_b Space

In an attempt to put all luminance information in a single channel, color images are often represented in the YC_rC_b color space for image compression. (ref. 2) give the transformation from RGB to YC_rC_b as

$$\begin{aligned} Y' &= 0.3\,R + 0.6\,G + 0.1\,B, \\ C_r &= (R - Y')/1.6 + 0.5, \\ C_b &= (B - Y')/2 + 0.5. \end{aligned} \tag{16}$$

Suppose that the viewing conditions are set so that the average image luminance is 40 cd/m^2, the pixel spacings are 0.028 deg, and the monitor calibration of the XYZ outputs for unit RGB inputs are given by the matrix

	X	Y	Z
R	26.1	13.3	2.3
G	25.2	48.9	10.2
B	9.3	4.7	35.7

(17)

The values of D_Y, D_O, and D_Z for each dimension turn out to be:

$$\begin{array}{cccc} & D_Y & D_O & D_Z \\ Y & 66.9 & -1.1 & 48.2 \\ C_r & -17.8 & 17.1 & -4.5 \\ C_b & -7.0 & 0.6 & 67.9 \end{array} \tag{18}$$

The quantization matrices appear in Table 2 following the references.

4. SUMMARY

We have presented a model for predicting visibility thresholds for DCT coefficient quantization error, from which quantization matrices for use in DCT-based compression can be designed. We regard this as preliminary results of work in progress. The quantization matrices computed by the techniques described above take no account of image content. We now show how an extension of this model may be used to optimize quantization matrices for individual images or a class of images.

5. LIMITATIONS OF THE IMAGE-INDEPENDENT APPROACH

The preceding approach constructs a quantization matrix independent of the image. While a great advance over the *ad hoc* matrices that preceded it, the image-independent approach has several shortcomings. The fundamental drawback is that visual thresholds for artifacts are dependent on the image upon which they are superimposed.

First, visual thresholds increase with background luminance., and variations in local mean luminance within the image will in fact produce substantial variations in DCT threshold. We call this *luminance masking*. Second, threshold for a visual pattern is typically reduced in the presence of other patterns, particularly those of similar spatial frequency and orientation, a phenomenon usually called *contrast masking*. This means that threshold error in a particular DCT coefficient in a particular block of the image will be a function of the value of that coefficient in the original image. Third, the image-independent approach ensures that any single error is below threshold. But in a typical image there are many errors, of varying magnitudes. The visibility of this error ensemble is not generally equal to the visibility of the largest error, but reflects a pooling of errors, over both frequencies and blocks of the image. We call this *error pooling*. Fourth, when all errors are kept below a perceptual threshold a certain bit rate will result. The image-independent method gives no guidance on what to do when a lower bit rate is desired. The ad hoc "quality factors" employed in some JPEG implementations, which usually do no more than multiply the quantization matrix by a scalar, will allow an arbitrary bit rate, but do not guarantee (or even suggest) optimum quality at that bit rate. We call this the problem of *selectable quality*.

6. IMAGE-DEPENDENT APPROACH

Here we present a general method of designing a custom quantization matrix tailored to a particular image. This *image-dependent* method incorporates solutions to each of the problems described above. The strategy is to develop a very simple model of perceptual error, based upon DCT coefficients, and to iteratively estimate the quantization matrix which yields a designated perceptual error or bit-rate. We call this the *DCTune* algorithm, because it tunes the DCT quantization matrix to the individual image(ref. 11, 12).

6.1. JPEG DCT Quantization

In the JPEG image compression standard, the image is first divided into blocks of size {8,8}. Each block is transformed into its DCT, which we write $I_{m,n,b}$, where m,n indexes the DCT frequency (or basis function), and b indexes a block of the image. Each block is then quantized by dividing it, coefficient by coefficient, by a quantization matrix (QM) $Q_{m,n}$, and rounding to the nearest integer

$$S_{m,n,b} = \text{Round}\left[I_{m,n,b}/Q_{m,n}\right] \tag{19}$$

The quantization error in the DCT domain is then

$$E_{m,n,b} = I_{m,n,b} - S_{m,n,b} Q_{m,n} \tag{20}$$

6.2 Luminance Masking

Detection threshold for a luminance pattern typically depends upon the mean luminance of the local image region: the brighter the background, the higher the luminance threshold (ref. 13, 14). This is usually called "light adaptation," but here we call it "luminance masking" to emphasize the similarity to contrast masking, discussed in the next section. We can compute a luminance-masked threshold matrix for each block in either of two ways. The first is to make use of a formula such as that supplied by Peterson *et al.* (ref. 7)

$$T_{m,n,b} = \mathrm{apw}[m,n,\bar{Y} I_{00b}/\bar{I}_{00}] \tag{21}$$

where I_{00b} is the DC coefficient of the DCT for block b, $\bar{Y}$ is the mean luminance of the display, and $\bar{I}_{00}$ is the DC coefficient corresponding to $\bar{Y}$ (1024 for an 8 bit image).

A second, simpler solution is to approximate the dependence of $T_{m,n,b}$ upon I_{00b} with a power function:

$$T_{m,n,b} = T_{m,n}\left(I_{00b}/\bar{I}_{00}\right)^{a_T} \tag{22}$$

The initial calculation of $T_{m,n}$ should be made assuming a display luminance of $\bar{Y}$. The parameter a_T takes its name from the corresponding parameter in the formula of Ahumada and Peterson, wherein they suggest a value of 0.65. Note that luminance masking may be suppressed by setting a_T=0. More generally, a_T controls the degree to which this masking occurs. Note also that the power function makes it easy to incorporate a non-unity display Gamma, by multiplying a_T by the Gamma exponent.

6.3 Contrast Masking

Contrast masking refers to the reduction in the visibility of one image component by the presence of another. Here we consider only masking within a block and a particular DCT coefficient. We employ a model of visual masking that has been widely used in vision models,(ref. 15, 16). Given a DCT coefficient $I_{m,n,b}$ and a corresponding absolute threshold $T_{m,n,b}$ our masking rule states that the masked threshold $M_{m,n,b}$ will be

$$M_{m,n,b} = T_{m,n,b} \max\left[1 , \left| I_{m,n,b}/T_{m,n,b} \right|^{w_{m,n}} \right] \tag{23}$$

where $w_{m,n}$ is an exponent that lies between 0 and 1. Because the exponent may differ for each frequency, we allow a matrix of exponents equal in size to the DCT. Note that when $w_{m,n}$ =0, no masking occurs, and the threshold is constant at $T_{m,n,b}$. When $w_{m,n} = 1$, we have what is usually called "Weber Law" behavior, and threshold is constant in log or percentage terms (for $I_{m,n,b} > T_{m,n,b}$). Because the effect of the DC coefficient upon thresholds has already been expressed by luminance masking, we specifically exclude the DC term from the contrast masking, by setting the value of $w_{00} = 0$.

6.4 Perceptual Error and Just-Noticeable-Differences

In vision science, we often express the magnitude of a signal in multiples of the threshold for that signal. These threshold units are often called "just-noticeable differences," or *jnd*'s. Having computed a masked threshold $M_{m,n,b}$, the error DCT may therefore be expressed in jnd's as

$$D_{m,n,b} = E_{m,n,b}/M_{m,n,b} \tag{24}$$

6.5 Spatial Error Pooling

To pool the errors in the jnd DCT we employ another standard feature of current vision models: the so-called Minkowski metric. It often arises from an attempt to combine the separate probabilities that individual errors will be seen, in the scheme known as "probability summation" (ref. 17) . We pool the jnds for a particular frequency m,n over all blocks b as

$$\Psi_{m,n} = \left(\sum_b \left| D_{m,n,b} \right|^{\beta_s} \right)^{1/\beta_s} \tag{25}$$

In psychophysical experiments that examine summation over space a β_s of about 4 has been observed (ref. 17). The exponent β_s is given here as a scalar, but may be made a matrix equal in size to the QM to allow differing pooling behavior for different DCT frequencies. This matrix $\Psi_{m,n}$ is now a simple measure of the visibility of artifacts within each of the frequency bands defined by the DCT basis functions. We call it the "perceptual error matrix."

6.6 Frequency Error Pooling

This perceptual error matrix $\Psi_{m,n}$ may itself be of value in revealing the frequencies that result in the greatest pooled error for a particular image and quantization matrix. But to optimize the matrix we would like a single-valued perceptual error metric. We obtain this by combining the elements in the perceptual error matrix, using a Minkowski metric with a possibly different exponent , β_f

$$\Psi = \left(\sum_{m,n} \Psi_{m,n}^{\beta_f} \right)^{1/\beta_f} \tag{26}$$

It is now straightforward, at least conceptually, to optimize the quantization matrix to obtain minimum bit-rate for a given Ψ, or minimum Ψ for a given bit rate. In practice, however, a solution may be difficult to compute. But if $\beta_f = \infty$, then Ψ is given by the maximum of the $\Psi_{m,n}$. Under this condition minimum bit-rate for a given $\hat{\Psi}$ is achieved when all $\Psi_{m,n} = \hat{\Psi}$.

6.7 Optimization Method

Under the assumption $\beta_f = \infty$, the joint optimization of the quantization matrix reduces to the vastly simpler separate optimization of the individual elements of the matrix. Each entry of the perceptual error matrix $\Psi_{m,n}$ may be considered an independent monotonically increasing function of the corresponding entry $Q_{m,n}$ of the quantization matrix.

6.8 Optimizing QM for a given bit-rate

To obtain a quantization matrix that yields a given bit rate with minimum perceptual error Ψ we note that the bit rate is a decreasing function of Ψ, and its is a simple matter to estimate the requisite perceptual error.

7. APPLICATION TO SPACE IMAGERY

Image compression will play a vital role in the distribution of preview images of science data to scientists at distributed sites, especially in programs such as EOS and the Mission to Planet Earth(ref. 18). Due to the generally high performance and wide availability of the JPEG still image compression standard, we expect it to play an important role in this area. Since the JPEG standard includes the quantization matrix as part of the file, DCTune technology provides a method of optimizing the bit-rate/quality trade-off for each science image.

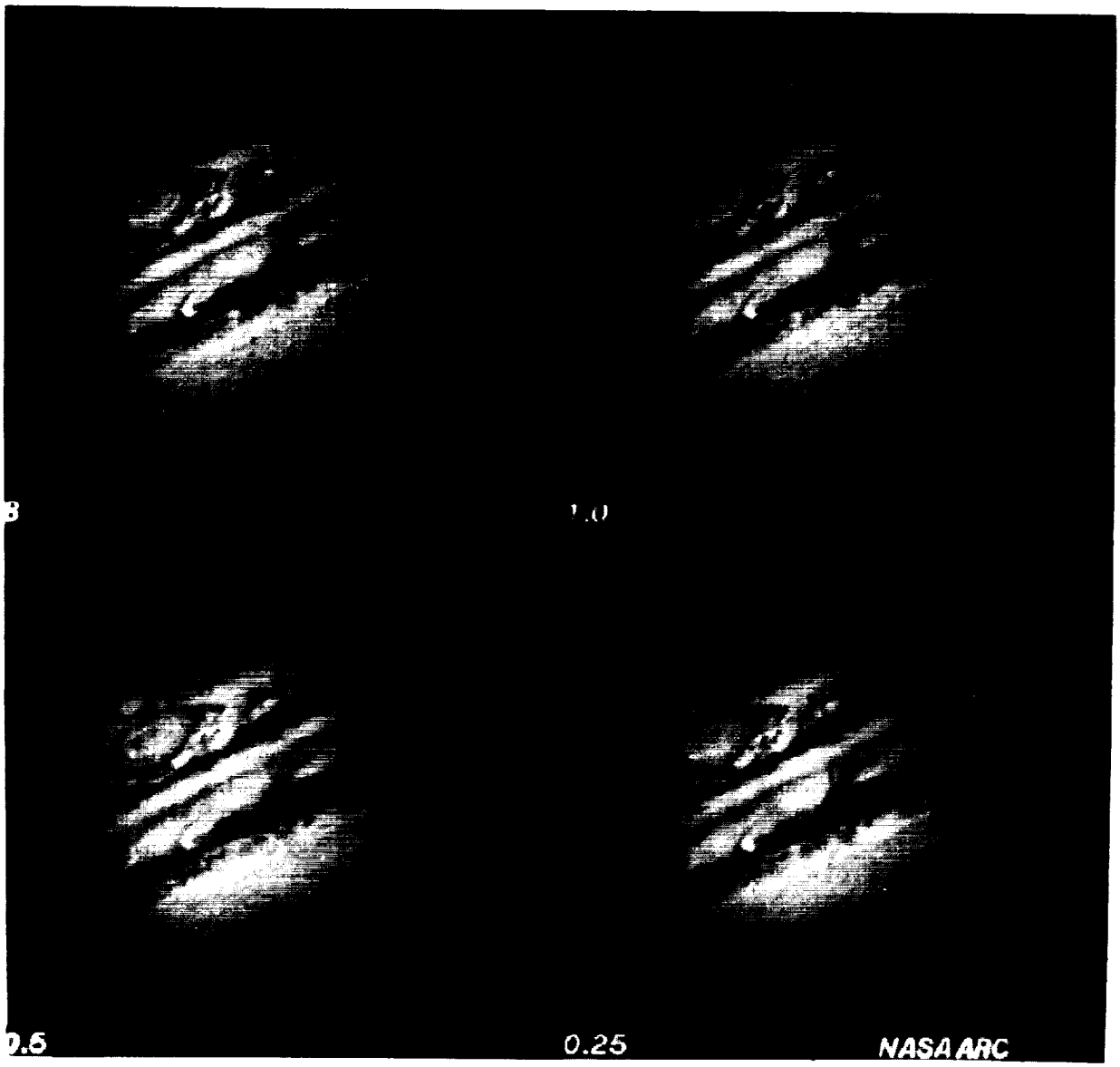

Figure 1. Voyager image of Jupiter compressed to 1.0, 0.5, and 0.25 bits/pixel, using optimal DCTune quantization matrices.

Lossy image compression based on the DCT may also play a role in the recovery of scientific imagery from spacecraft. The Galileo orbiter spacecraft is now on its way to Jupiter. Due to a malfunction of the main antenna, image data will be sent to earth over an auxiliary antenna with approximately 15,000 times lower bandwidth. Image compression will be used to partially compensate for the loss of bandwidth(ref. 19). In support of this effort, we have designed quantization matrices for use in the Galileo mission, based on application of DCTune technology to existing Voyager and Galileo images (ref. 20, 21) . An example of DCTune algorithm applied to an image of Jupiter obtained by the previous Voyager mission is shown in Fig. 1. It shows the original and three levels of optimized compression: 1.0, 0.5, and 0.25 bits/pixel. In this example, the parameter values used were a_T = 0.65, β = 4, $w_{m,n}$= 0.7 , display mean luminance $\bar{Y}$ = 65 cd m^{-2} , image greylevels = 256, $\bar{I}_{00}$ = 1024. The viewing distance was assumed to yield 32 pixels/degree.

8. SUMMARY

We have shown how to compute a visually optimal quantization matrix for a given image. These image-dependent quantization matrices produce better results than image independent matrices. The DCTune algorithm can be easily incorporated into JPEG-compliant applications.

In a practical sense, the DCTune method proposed here solves two problems. The first is to provide maximum visual quality for a given bit rate. The second problem it solves is to provide the user with a sensible and meaningful quality scale for JPEG (or other DCT-based) compression. Without such a scale, each image must be repeatedly compressed, reconstructed, and evaluated by eye to find the desired level of visual quality.

9. ACKNOWLEDGMENTS

We appreciate the help of Heidi A. Peterson and Jeffrey B. Mulligan. This work was supported in part by the IBM Independent Research and Development Program and by NASA RTOP Nos. 506-59-65 and 505-64-53.

10. REFERENCES

1. Wallace, G. The JPEG still picture compression standard. Communications of the ACM. **34**(4): 30-44, 1991.

2. Pennebaker, W. B. and J. L. Mitchell. "JPEG Still image data compression standard." 1993 Van Nostrand Reinhold. New York.

3. LeGall, D. MPEG: A video compression standard for multimedia applications. Communications of the ACM. **34**(4): 46-58, 1991.

4. Peterson, H. A., H. Peng, J. H. Morgan and W. B. Pennebaker. Quantization of color image components in the DCT domain. Human Vision, Visual Processing, and Digital Display. Proc. SPIE. **1453**: 210-222, 1991.

5. Peterson, H. A. "DCT basis function visibility in RGB space." Society for Information Display Digest of Technical Papers. Morreale ed. 1992 Society for Information Display. Playa del Rey, CA.

6. Peterson, H. A., A. J. Ahumada Jr. and A. B. Watson. The Visibility of DCT Quantization Noise. SID Digest of Technical Papers. **XXIV**: 942-945, 1993.

7. Peterson, H., A. Ahumada and A. Watson. "An Improved Detection Model for DCT Coefficient Quantization." Human Vision, Visual Processing, and Digital Display IV. Allebach ed. 1993 SPIE. Bellingham, WA.

8. Ahumada, A. J., Jr. and H. A. Peterson. "Luminance-Model-Based DCT Quantization for Color Image Compression." Human Vision, Visual Processing, and Digital Display III. Rogowitz ed. 1992 Proceedings of the SPIE.

9. Boynton, R. M. "Human Color Vision." 1979 Holt, Rinehart and Winston. New York.

10. Mullen, K. T. The contrast sensitivity of human color vision to red/green and blue/yellow chromatic gratings. Journal of Physiology, Lond. **359**(381-400): 1985.

11. Watson, A. B. "DCT quantization matrices visually optimized for individual images." Human Vision, Visual Processing, and Digital Display IV. Rogowitz ed. 1993 SPIE. Bellingham, WA.

12. Watson, A. B. DCTune: A technique for visual optimization of DCT quantization matrices for individual images. Society for Information Display Digest of Technical Papers. **XXIV**: 946-949, 1993.

13. van Nes, F. L. and M. A. Bouman. Spatial modulation transfer in the human eye. Journal of the Optical Society of America. **57**: 401-406, 1967.

14. Barlow, H. B. "Dark and light adaptation: Psychophysics." Handbook of Sensory Physiology. Hurvich and Jameson ed. 1972 Springer-Verlag. New York.

15. Legge, G. E. and J. M. Foley. Contrast masking in human vision. Journal of the Optical Society of America. **70**(12): 1458-1471, 1980.

16. Legge, G. E. A power law for contrast discrimination. Vision Research. **21**: 457-467, 1981.

17. Robson, J. G. and N. Graham. Probability summation and regional variation in contrast sensitivity across the visual field. Vision Research. **21**: 409-418, 1981.

18. Jaworski, A. Earth Observing System (EOS) Data and Information System (DIS) software interface standards. AIAA/NASA Second International Symposium on Space Information Systems. **AIAA-90-5075**: 1990.

19. Cheung, K.-M. and K. Tong. Proposed data compression schemes for the Galileo S-band contingency mission. 1993 Space and Earth Science Data Compression Workshop. **3191**: 99-109, 1993.

20. Watson, A. B. and A. J. Ahumada Jr. Preservation of photometric accuracy in ICT-compressed imagery. 1993.

21. Watson, A. B., A. J. Ahumada Jr. and M. J. Young. ICT quantization matrix design for the Galileo S-Band Mission. 1993.

11. APPENDIX

	16	11	10	16	24	40	51	61
	12	12	14	19	26	58	60	55
luminance	14	13	16	24	40	57	69	56
quantization	14	17	22	29	51	87	80	62
matrix	18	22	37	56	68	109	103	77
	24	35	55	64	81	104	113	92
	49	64	78	87	103	121	120	101
	72	92	95	98	112	100	103	99
	17	18	24	47	99	99	99	99
	18	21	26	66	99	99	99	99
chrominance	24	26	56	99	99	99	99	99
quantization	47	66	99	99	99	99	99	99
matrix	99	99	99	99	99	99	99	99
	99	99	99	99	99	99	99	99
	99	99	99	99	99	99	99	99
	99	99	99	99	99	99	99	99

Table 1. The default quantization matrices. The Q_{00} value is located in the upper left corner of each matrix.

Y' quantization matrix	15	11	11	12	15	19	25	32
	11	13	10	10	12	15	19	24
	11	10	14	14	16	18	22	27
	12	10	14	18	21	24	28	33
	15	12	16	21	26	31	36	42
	19	15	18	24	31	38	45	53
	25	19	22	28	36	45	55	65
	32	24	27	33	42	53	65	77

C_r quantization matrix	21	21	41	45	55	71	92	120
	21	37	39	38	44	55	70	89
	41	39	51	54	59	69	83	103
	45	38	54	69	80	91	106	126
	55	44	59	80	100	117	136	158
	71	55	69	91	117	144	170	198
	92	70	83	106	136	170	206	243
	120	89	103	126	158	198	243	290

C_b quantization matrix	45	43	103	114	141	181	236	306
	43	78	99	97	113	140	178	228
	103	99	130	138	150	175	212	262
	114	97	138	176	203	232	270	321
	141	113	150	203	254	299	347	403
	181	140	175	232	299	367	434	505
	236	178	212	270	347	434	525	619
	306	228	262	321	403	505	619	739

Table 2. YCrCb quantization matrices. The values in these matrices are obtained following the proceduredescribed in Section 3. The $Q_{0,0}$ value is located in the upper left corner of each quantization matrix. As specified in the JPEG standard, the values have been rounded to the nearest integer. JPEG also requires that values in the quantization matrix be ≤ 255.

S/3-32
253/
p. 10

N94-32468

VOICE AND VIDEO TRANSMISSION USING XTP AND FDDI

John Drummond
Naval Command, Control and Ocean Surveillance Center RDT&E Division (NRaD)
San Diego, CA 92152-5000

Edwin Cheng
Naval Command, Control and Ocean Surveillance Center RDT&E Division (NRaD)
San Diego, CA 92152-5000

Will Gex
Naval Command, Control and Ocean Surveillance Center RDT&E Division (NRaD)
San Diego, CA 92152-5000

ABSTRACT

The use of XTP and FDDI provides a high speed and high performance network solution to multimedia transmission that requires high bandwidth. FDDI is an ANSI and ISO standard for a MAC and Physical layer protocol that provides a signaling rate of 100 Mbits/sec and fault tolerance. XTP is a Transport and Network layer protocol designed for high performance and efficiency and is the heart of the SAFENET Lightweight Suite for systems that require high performance or realtime communications. Our testbed consists of several commercially available Intel based i486 PCs containing off-the-shelf FDDI cards, audio analog-digital converter cards, video interface cards, and XTP software. Unicast, multicast, and duplex audio transmission experiments have been performed using XTP and FDDI. We are working on unicast and multicast video transmission. Several potential commercial applications are described.

INTRODUCTION

Multimedia (voice, video, data, text, and graphics) distribution over high speed networks has many commercial applications which will revolutionize the way we use computers and networks. Several big corporations have already formed strategic alliances to explore new opportunities in this area.

We have been researching and experimenting for several years with high speed networks which utilize Fiber Distributed Data Interface (FDDI), and a high performance network protocol called Xpress Transfer Protocol (XTP). As multimedia increased in popularity, in both military and commercial world, we started to look at the possibility of using XTP and FDDI in voice and video transmission. We have performed many voice transmission experiments using XTP with several PC i486 machines connected via FDDI network. The results indicate voice transmission using XTP and FDDI have many advantages over traditional methods of voice transmission such as fault tolerance, high bandwidth, and data integration. Right now, we are performing video transmission experiments using XTP and FDDI.

XTP

The Xpress Transfer Protocol (XTP)[1] has been developed over the past seven years from a consortium of private industry, academia, and government to address many high performance and realtime issues that were lacking in previously developed transport and network protocols. Certain concepts from existing protocols (e.g.VMTP, GAM-T-103, Delta-t, NETBLT) were modified and combined with new ideas to form the basis for XTP. Experience and other ideas have added to its development to produce the current specification[2].

It is a protocol that spans the Network and Transport layers (layers 3 and 4) of the OSI 7 layer model and therefore has some interesting features due to the coupling of an end-to-end protocol with an intermediate, network protocol (e.g. bandwidth reservation of an intermediate resource by an end host or packet priority assigned by an end host and used by an intermediate router). Because the protocol does not specify policy but