Perceptually lossy compression of documents

Giordano Beretta, Computer Peripherals Laboratory
Vasudev Bhaskaran, Computer Systems Laboratory
Konstantinos Konstantinides, Computer Peripherals Lab
Balas Natarajan, Computer Systems Laboratory

HPL–97–23
January, 1997

facsimile, document compression, image transmission, reading performance, JPEG, CIELAB

We present a method for designing JPEG quantization tables that are tailored for color facsimile transmission. The proposed technique combines results on visual acuity and peak contrast sensitivity of the human visual system with results from rate distortion theory. The new tables yield higher compression ratios for a given visual quality, resulting in shorter transmission times.

1 Introduction

We report on our experience with the implementation of a color facsimile machine according to the recent International Telecommunication Union (ITU) standard (ITU-T T.42 Addendum [10]). With careful consideration of the human visual system's (HVS) properties as they relate to electronic imaging (EI) for compound documents, we have achieved our self-imposed goal of transmitting full color documents at the same cost as a for a binary rendition. With diligent software engineering we have realized an implementation that runs on an inexpensive machine powered by a 66MHz i486 class processor with a minimal amount of dynamic random access (RAM) memory, avoiding the high cost of digital signal processors (DSP) and static RAM. We believe our main contributions are in the imaging processing portion of the system and this will be the main focus this paper. Similar considerations can be applied for publishing on the W^3; we will report on our efforts at a future occasion.

1.1 The discrete nature of perception

From 1860, when the physicist Fechner introduced the discipline called psychophysics, it is known that human observers are able to specify the weakest detectable sensations in terms of the stimulus energy necessary to produce them. The smallest amount of stimulus energy necessary to produce a stimulus is called *stimulus threshold*. When a stimulus is above this threshold, the intensity of this stimulus must be changed by some critical amount before an observer can detect a change in sensation. The amount of change in a stimulus that is necessary to produce a *just noticeable difference* (jnd) in the sensation is called the *difference threshold*.

A more modern model dating to 1966 by Green and Sweets and known as theory of signal detection (TSD), states that signals (or stimuli) are always detected by electronic devices (or human observers) against a background of random activity called *noise*. A signal can only be detected if it is above the noise level, and modulation in a signal above noise can be detected only if the modulation is above the noise level. In either case of psychophysics or TSD, the stochastic nature of detection is important and results are always obtained by averaging the data from a number of observers.

Although stimulus intensity is usually represented as a real number, the sensation is discrete. Unfortunately, due to the stochastic nature of the detection process, which can present large variability from observer to observer and also for a single observer in different experiments, it is very difficult to specify discrete scales for percepts. Mastering this difficulty is the key for the compact representation of electronic images.

For example, there is evidence that the HVS can distinguish approximately 40 levels in each of the three channels for luminance and chromaticity. If we know exactly which stimulus energies correspond to these levels, we can communicate every perceptible color using only a 7-bit datum. However, if we use a 7-bit representation and there is a poor correspondence with the HVS levels, the image will appear degraded or posterized. If we can find good correspondences we can get by with compact representations, using a coarse granularity, but if the correspondence is poor, we are forced to use a conservative fine granularity, entailing a wasteful representation.

In conclusion, the quality of an image does not depend only on the number of bits used for its representation, but mostly in the representation itself. Our quest has been for the best representations of the two dimensions of images: the color information and the spatial information.

1.2 Encoding color information

The representation of color values is called color space or color model operator. In color science, spaces that have the above property of a good correspondence with HVS levels are called *perceptually uniform*. Color facsimile is based on CIELAB [6] which has a number of advantages over other uniform color spaces: it is based on the CIE Standard Colorimetric Observer, it has good perceptual uniformity, it is easy to compute, it is widely used in the printing industry and can be read directly with measurement instruments.

The only critical decision left once the color space has been selected is to decide on the best range for the chromatic channels a^* and b^*. Since the number of bits available to represent each channel is fixed, the size of the allowed color gamut directly determines the accuracy of the color representation. This problem has been studied by A.H. Mutz [16].

It might be argued that CIELAB is too complicated for practical use. However, during the past decade applications such as color copiers have encouraged the development of algorithms to convert between CIELAB and device color spaces that are very efficient and that can easily be implemented in relatively simple hardware. Even for our software implementation, that does not use any hardware assist, the color space conversion time is only slightly longer than the scanning time [1]. It is a surprising testimony of the compartmentalization of engineering when each new technology has to rediscover this fact. For example, one of the limitations currently afflicting digital electronic still cameras (DESC), is that they store only a small number of images. This limitation is because the images can be compressed no more than 1:10, as Kinoshita recently related in his review [11]. One reason for this limitation is that the images are compressed in the sensor's WCYG color space (in the case of Hitachi's color filter array (CFA)) or in its RGB color space (in the case of the other CFAs) [14].

1.3 Encoding spatial information

The other problem for the low compression in DESCs is that the spatial compression is performed using the JPEG method with the DQTs in Annex K of the JPEG standard [11]. This misunderstanding that the example tables (K.1 for luminance and K.2 for chrominance channels) are a recommendation is more detrimental for the color facsimile application or for compound documents in general. The example tables were derived for pictorial images represented in the YUV color space and viewed on a color monitor. In color facsimile a large fraction of the image is typed text, the color space is CIELAB, and the output device is a high resolution printer.

With the K.1 DQT, at 300 dpi the 4CP01 facsimile test chart compresses to 1:36. Using the DQT design methodology first introduced by Lohscheller and then widely reported in the literature [1, 17, 19, 23, 24] the compression ratio can easily be doubled. However, such a compression ratio is insufficient for an acceptable transmission time; to achieve a compression ratio of 1:100 or more it is necessary to operate at suprathreshold levels, i.e., to resort to perceptually lossy compression. The design problem is then to place the artifacts in regions of the image where they do not affect reading efficiency.

2 Reading efficiency

What is important for comfortable reading is the information available at higher stages in the visual cognitive process, most notably post awareness. Reading performance depends not only on the available physical information delivered by a visual stimulus, but also on the ability to process this information cognitively, that is, to utilize it. The reading efficiency of text can easily be assessed experimentally. Common parameters for psychometric functions for reading efficiency are reading rate and letter identification:

- reading speed (words/minute) in drifting text
- reading speed in flashed text
- better than chance (50%) in a character identification task
- discrimination performance (telling character pairs apart)

We limit ourselves to the effects due to visual stimulation. The retinal spatial frequency of a character (the object) is determined by the font size and the reading distance. The object frequency is determined by the frequency band of the bandpass filtering induced by the lossy JPEG compression. In other words, the retinal frequency is an indication of the number of cones equally excited by a sharp stimulus, while the object frequency refers to relative response of cones excited by a blurred stimulus.

For letter discriminability, it has been determined that retinal spatial frequency has no effect, only object spatial frequency. Parish and Sperling [17] found that subjects can best extract upper-case letter information from spatial frequencies of 1.5 cycles per object height, and they can extract it with equal efficiency over a range from 0.074 to 2.3 cycles per degree of visual angle (retinal spatial frequency).

It is useful for our subsequent discussion to note that the characters they used in their experiments were quite large compared to those found in typical body text matter. Also, the typeface used was very crude: a 5×7 pixel computer terminal upper-case font. Typically in these experiments the characters are formed by a large number of pixels and are viewed from a distance or through a telescope.

Other results found in the more recent literature indicate that letter identification is mediated by an octave-wide bandpass filter centered at 3 cycles/character [18] or at 2.5 cycles/character [2]. From these results we learn the font size is not important, but the characters should have distinct features of the size 1/3 to 1/6 of the character size.

Parish and Sperling determined that subjects perform better when the frequency band is higher, and subjects require the smallest signal-to-noise ratio in the highest frequency band they used in their bandpass filter. In other words, the character information should be concentrated in a relatively small number of pixels, i.e., the characters should have sharp edges. As mentioned earlier, these results are for very crude typefaces. For typefaces typically used in printed matter, is has long been known that the more letters are differentiated from each other, the easier is the reading. This differentiation is achieved by setting words in lower case and by using serif letters.

On a superficial reading, these constraints on character contrast might seem to be easily fulfilled. One might think, that if one comes close this might just require some more careful reading in some cases, but there would not be a functional impairment. How important for readability is the contrast of characters? Legge [15] has shown that high contrast is particularly critical for people with low vision. People with normal vision are quite tolerant to contrast loss; a tenfold reduction from maximum contrast results in only a slight decrease in reading speed (less than a factor of two). However, the reading speed of people with low vision declines with any reduction from maximum contrast.

So far, our emphasis has been to show how different and independent retinal and object spatial frequency are. To exploit the results from the work on the contrast sensitivity of the human visual system, we have to identify the similarities and dependencies between the two. As reported by Legge [15] it is possible to define a contrast sensitivity function for reading. At each of a number of letter sizes the contrast required to produce a threshold reading rate can be determined. If these threshold contrasts are plotted as a function of the reciprocal of character size of in degrees, a plot qualitatively similar in shape to that of contrast sensitivity functions for sine-wave gratings is obtained. Recently this assertion has been confirmed also quantitatively. Experiments by Alexander et al. [2] have shown that restricting the spatial-frequency contents of letters by spatial-bandpass filtering results in contrast sensitivity functions that are quite similar in shape to those for D6 patterns.*

Although the discrete cosine transform is a sum of spatially filtered blocks, we will nevertheless assume that contrast sensitivity data for sinusoidal distributions as found in the literature can be applied to the characterization of letter identification for color facsimile.

3 Designing JPEG DQT matrices

Let us first clarify the terms of lossy and lossless. *Lossless* compression means that no data is lost when an image is compressed and then decompressed, *i.e.*, the original and the processed image are physically identical. *Perceptually lossless* compression means that no visible data is lost when an image is compressed and the decompressed, *i.e.*, the original and the processed image are perceptually identical. In *lossy* compression, information is lost. In *perceptually lossy* compression the loss of information is visible. The perceptual alterations to the image are called *compression artifacts* or, in the case of JPEG, *quantization errors*. In Lohscheller's psychophysical method based on the HVS, the elements in the DQT are chosen so that the quantization errors are just below the visibility threshold, *i.e.*, the compression is lossy but perceptually lossless.

The usual method to increase the compression ratio of the JPEG method is to scale the DQTs until the quantization errors can no longer be accepted by the user. This simple method can produce acceptable results in applications such as tele-conferencing, where the images are sequences of talking heads. As noted by Hoshino et al. this method is not applicable for compound documents such as typical in color facsimile, where the images are static, observed for a relatively long time, and the preservation of detail in fine structures such as characters is important. Hoshino et al. note that text must be compressed more conservatively than pictorial images to preserve the sharp edges important for reading efficiency. Their proposed solution is to sharpen the image after decompression, a step that requires additional computation. Our proposed solution is to modify the DQT on an element by element basis to preserve the features important for text.

In a sequence of increasing sophistication, we design the DQT matrices in four steps, that for reasons that will become clear later, we call the physical world, two perceptual worlds, and the semantic world. The four design iterations are carried out in a certain order; as a completely different framework is used in each step, we call them "worlds."

---

* D6 patterns are defined by a sixth spatial derivative of a Gaussian in one direction and by a Gaussian in the orthogonal direction. They are similar in spatial frequency content to a Gaussian-windowed sinusoidal grating or Gabor patch.

3.1 Physical world

In the physical world we examine the statistical properties of typical images used in color facsimile communication. We use a well-known technique called *bit-rate control*, where the key idea is to allocate more bits to the coefficients in which the images contain more energy. A common correlate for the energy is the statistical variance. For our prototype color facsimile machine we followed the example in reference [4].

Let $Y_i[k, l]$ be the (k, l) (for $k, l \in [0, 7]$) DCT element of the i-th 8×8 block in an image. Let B be the number of blocks in the image. The variance $V_y[k, l]$ of the (k, l) frequency component in an image is then computed as

$$V_y[k, l] = \operatorname{var}(Y[k, l]) = \frac{1}{B}\sum_{i=1}^{B}(Y_i[k, l] - M_y[k,l])^2 \tag{1}$$

where $M_y[k, l]$ denotes the mean and

$$M_y[k,l] = \frac{1}{B} \cdot \sum_{i=1}^{B} Y_i[k, l]. \tag{2}$$

Let $N_{k, l}$ be the number of bits allocated for the (k, l) DCT element. For logarithm in base 2,

$$N_{k, l} = \frac{1}{2} \cdot \log\frac{V_y[k, l]}{D} \tag{3}$$

Given an overall bits-per-pixel (bpp) rate,* the value of D is defined by solving for

$$64 \cdot \text{bpp} = \sum_{k=0}^{7}\sum_{l=0}^{7} N_{k, l} \tag{4}$$

Finally, the elements $Q[k, l]$ of the DQT are defined as

$$Q[k, l] = \frac{2048}{2^{N_{k, l}}}, \tag{5}$$

where 2048 represents the maximal range that the output of the discrete cosine transform may have, namely [-1024, 1024].

The bits-per-pixel rate does not represent a real bit-rate. In practice one has to experiment with different values and pick a value for which the visual quality is acceptable. For example, in color facsimile images bpp = 4 yields images that are perceptually lossless. With bit-per-pixels rates lower than 2.5 we can start seeing artifacts.

* This is the bit rate *before* the Huffman encoder.

The above technique yields discrete quantization tables that result in good compression ratios and acceptable visual quality. The compression can be further increased by taking into consideration the human visual system.

3.2 Perceptual world (contrast sensitivity)

In this step we note that image quality depends on the contrast visible at a given spatial resolution. We weight the DQT elements by the contrast sensitivity function (CSF) of the HVS [5]. The result is much better than just increasing the a global scaling factor like the *q*-factor, which is equivalent to a constant CSF. De Queiroz and Rao [20] have described in detail how to weigh the DQT by the CSF and we refer to their paper for further details.

Up to this point the resulting compression method is lossy, but we cannot make any assertion on the presence and magnitude of artifacts, because we have not applied the visibility thresholds to the image or the discrete cosine transform (DCT) kernels.

3.3 Perceptual world (visual masking)

The CSF is determined with a sine-wave grating. When a new structure (texture) is added to an existing structure, the new structure may mask the old structure or vice-versa. Quantization noise is a structure that is added to the original image; noise that is masked by the image is perfectly acceptable, allowing for higher compression ratios than the CFS method alone. This method is used mostly for image-dependent (adaptive) compression, but it turns out that there is only little variation in the DQT elements for images of similar contents, which is the case in color facsimile communication.

The main researcher active in this field is Watson [23] of NASA Ames and in the past couple of years some excellent work has been performed by Westen [24] at the TU Delft in the Netherlands. Watson developed an iterative method where the DQT matrix elements are changed until the quantization errors in each DCT kernel element are at a predetermined multiple of visual threshold (jnd—just noticeable difference). Watson’s method is based on a model of the HVS, which is much more convenient than having to perform psychophysical experiments.

Although Watson’s method works also by starting with a random DQT matrix, by performing the bit-rate control step and weighting it by the CSF, Watson’s method converges much more rapidly. After this step we have a compressed image where for each DCT kernel element we know the quantization error in terms of jnds.

3.4 Semantic world

The critical image element with respect to data compression for color facsimile communication is text, which consists mostly of high frequency information. For the quality of color facsimile, the reading efficiency is a critical factor. When compressing, we can discard information that does not impact reading performance. We identify the parts of characters in fonts that affect reading performance (*e.g.*, serifs) and discard prevalently spatial information not related to these character parts (see Fig. 1). This is possible because within a window covering three adjacent components, coefficient energy in the DCT does not depend on the phase angle [7].

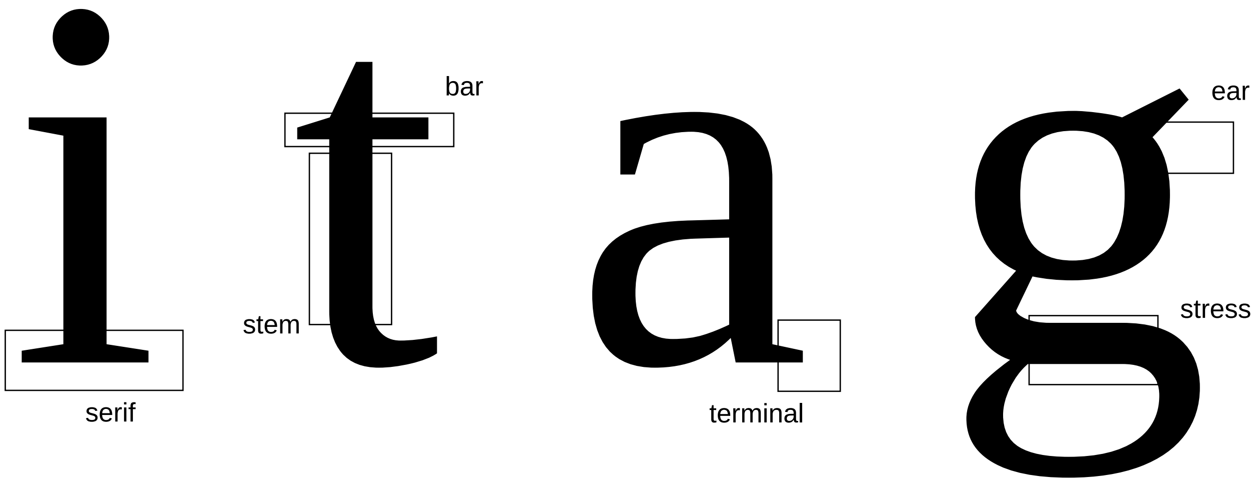

Figure 1. Some parts of the Times New Roman typeface. *Serif*: the short finishing stroke at the end of stems. *Stem* or *main stroke*: straight vertical or full-length diagonal stroke. *Bar*: horizontal stroke in A, H, e, t, etc. *Terminal*: end of a stroke not terminated with a serif. *Stress*: direction of thickening in a curved stroke. *Ear*: Small stroke springing from the top of the g.

Note that the situation is different in printed text than in text displayed on a television monitor. In the first case, the output device operates at a relatively high resolution and the user must be able to read long text efficiently; thus we aim to preserve features like serifs that differentiate characters and bind words. In television application text is mostly used for display, titles, and captions, for which sans-serif typefaces are more suited and sharp character edges are more important. Yanagihara [25] has studied the design of DQTs for television applications.

So far, these considerations have applied only to the luminance channel, because the chrominance channels are usually considered to have the same CSF as the luminance channel, just shifted towards the lower frequencies, which is where the subsampling of the chrominance channels for $a*$ and $b*$ comes from. In the case of color facsimile, the chrominance channels are easier to handle than in the general case and we can quantize more aggressively. As Vivienne Smith [22] of the University of Chicago has noted recently, the DQTs for the red/green channel can be weighted reciprocally to the luminance channel DQT because in practice little or no text information is in this channel alone (document creators tend to avoid designs that are hard to read and there is substantial evidence of masking interactions among the chrominance and the luminance channels in the HVS). As for the blue/yellow channel, it is sufficient to consider prominently the DC element, since all spacial information relating to text is already in the luminance channel versus the blue/yellow channel.

In summary, we start with a DQT matrix optimized based on bit-rate control techniques, which takes into account the energy of typical images used in color facsimile communication. Then, instead of applying a global q-factor to achieve a higher compression rate we weigh the DQT elements according the contrast sensitivity of the HVS. After considering the visual masking occurring in average images to produce a DQT matrix with uniform and well-defined quantization error visibility, a second weighting preserving the character parts that impact reading performance of text is applied to increase the compression ratio. This leads us to modify the bit allocation equation (3) with

$$N_{k,\,l} = \frac{1}{2} \cdot \log\left(w[k,\,l]\frac{V_y[k,\,l]}{D}\right) \tag{6}$$

where $w[k, l]$ denotes the weight of each element in the DQT based on visual quality. Table 1 shows an example for a weight table

1	1	1	3/4	1	3/4	1	1
1	3/4	1/2	1/4	1/2	1/4	1/2	1/2
1	1/2	1/4	1/8	1/4	1/8	1/4	1/4
3/4	1/4	1/8	1/8	1/8	1/8	1/8	1/8
1	1/2	1/4	1/8	1/8	1/8	1/8	1/8
3/4	1/4	1/8	1/8	1/8	1/8	1/8	1/8
1	1/2	1/4	1/8	1/8	1/8	1/8	1/8
1	1/2	1/4	1/8	1/8	1/8	1/8	1/8

Table 1. Weight table for color facsimile. An example of a conceptional weight table for the luminance channel optimized for color facsimile transmission at 300 dpi.

4 JPEG Huffman tables

The example Huffman tables (HT) in the JPEG standard have been determined for the perceptually lossless compression of YC_bC_r images. In color facsimile, when DQT matrices designed according to the method above are used, compression is visually lossy. Furthermore, the perceptually uniform CIELAB color space is used, which is different from the YC_bC_r color space. In the following we briefly summarize the results presented in detail at a previous meeting [12].

Our experimental results show that we can design good constrained-length Huffman codes using the simple ad hoc technique of setting to 2^{-16} all symbol probabilities less than 2^{-16} and then proceeding as if there were no constraints on the codeword lengths. For each of our test images, first we compute the probability distribution of all possible symbols for which Huffman code words are needed, and then we design image-dependent custom tables. We have found that using custom Huffman tables we can improve the compression ratio by a factor between 8% and 14% compared to using the example tables from the JPEG standard. We have also found that if we use image-independent tables obtained by averaging the statistics of all test images, we only lose about 0.5% compression ratio. Hence we are presently using image-independent custom Huffman tables, achieving an average improvement of 11% in the compression ratio of typical images used in color facsimile communication.

5 Experimental results

We have achieved error-free color facsimile communication at speeds up to 9600 bit/s, the limiting factor being the speed at which we can execute the modem functions in software [3]. With our custom JPEG table we have been able to achieve 1:98 or better compression ratios, achieving transmission times of less than three minutes in a software-only implementation. With a V.34 facsimile modem it should be possible to transmit a full-color facsimile page in under a minute, which is an acceptable transmission time.

Although compression artifacts are visible in the images, the reading efficiency is not impaired. Qualitatively the images perform similar to images that have been scanned and printed on a desktop computer, which can hardly be compared with the poor visual quality of conventional binary facsimile images.

References

[1] A.J. Ahumada, Jr. and H.A. Peterson, “Luminance-Model-Based DCT Quantization for Color Image Compression,” *Human Vision, Visual Processing, and Digital Display III*, **1666**, 365-374, SPIE, San Jose, 1992.

[2] K.R. Alexander, W. Xie, and D.J. Derlacki, “Spatial-frequency characteristics of letter identification,” *J. Opt. Soc. Am. A*, **11**, 9, 2375–2382, 1994.

[3] G. Beretta, K. Konstantinides, D.T. Lee, H.J. Lee, A.H. Mutz, “Implementing a color facsimile machine,” Proc. *IS&T's 11th International Congress on Advances in Non-Impact Printing Technologies,* 475–478, IS&T, Hilton Head, 1995.

[4] V. Bhaskaran and K. Konstantinides, Image *and Video Compression Standards — Algorithms and Architectures,* Kluwer Academic Publishers, Boston, 1995.

[5] W.-H. Chen and C.H. Smith, “Adaptive Coding of Monochrome and Color Images,” *IEEE Transactions on Communications,* **COM-25**, 1285-1292, 1977.

[6] CIE, *Colorimetry*, 2nd Edition, 15.2, Vienna, Austria, 1986.

[7] R.J. Clarke, “Spectral Response of the Discrete Cosine and Walsh-Hadamard Transforms,” *IEE Proc.*, **130**, Part F, 309-313, 1983.

[8] Y. Hoshino, D. Ando, S. Higashino, and M. Tsukamoto, “Applicability of a standardized discrete cosine transform coding method to character images,” *J. Electronic Imaging,* **1**, 3, 322–327, 1992.

[9] ITU-T Rec. T.81 | ISO/IEC 10918-1 (JPEG), *Information technology —- Digital compression and coding of continuous-tone images,* 1992.

[10] ITU (Q.5/8 Rapporteur), *Amendments to ITU-T Rec. T.4 to enable continuous-tone colour and grayscale mode for Group 3*, **COM 8-44-E**, June 1994.

[11] H. Kinoshita, “Image Compression and Quality,” Proc. *SPSTJ Symposium on Digital Photography '96,* 21–24, IS&T Tokyo Chapter, Tokyo, 1996.

[12] K. Konstantinides, “Effects of Custom Huffman Tables on the Compression of CIELAB Images Using JPEG, *SID Digest of Technical Papers*, **XXVII**, 951–954, San Diego, 1996.

[13] K. Konstantinides, H.J. Lee, G.B. Beretta, and A.H. Mutz, “A Testbed for Color Imaging and Compression,” Proc. *Intern. Conf. on Digital Signal Processing*, **1**, 29–34, Limassol, Cyprus, 1995.

[14] M.A. Kriss, “Image Quality and Color Filter Arrays in Digital Still Camera Systems,” Proc. *SPSTJ Symposium on Digital Photography '96,* 29–33, IS&T Tokyo Chapter, Tokyo, 1996.

[15] G.E. Legge, “Reading: effects of contrast and spatial frequency,” *Applied Vision*, OSA Technical Digest Series, **16**, 90–93, 1989.

[16] A.H. Mutz, “Effect of Gamut Range Upon Color Quantization During JPEG Coding,” Proc. *SID'94 Digest,* San Jose, California, **XXV**, 891–893, 1994.

[17] D.H. Parish and G. Sperling, "Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination," *Vision Research*, **31**, 7/8, 1399–1415, 1991.

[18] D.G. Pelli, C.W. Burns, M. Raghavan, and B. Farell, "Visual Factors in Letter Identification," *IS&T's 47th Annual Conference/ICPS*, 411, 1994.

[19] H.A. Peterson, A.J. Ahumada, Jr., and A.B. Watson, "An Improved Detection Model for DCT Coefficient Quantization," *Human Vision, Visual Processing, and Digital Display IV*, **1913**, 191-201, SPIE, San Jose, 1993.

[20] R.L. de Queiroz and K.R. Rao, "Human visual sensitivity-weighted progressive image transmission using the lapped orthogonal transform," *J. Electronic Imaging*, 1, 3, 328–338, 1992.

[21] R.J. Safranek and J.D. Johnson, "A Perceptually Tuned Sub-band Image Coder With Image Dependent Quantization and Post-quantization Data Compression," Proc. *ICASSP 89*, **3**, 1945-1948, 1989.

[22] V. Smith, personal communication, April 1995.

[23] A.B. Watson, "DCT Quantization Matrices Visually Optimized for Individual Images," *Conference on Human Vision, Visual Processing, and Digital Display IV*, **1913**, 202–216, SPIE, San Jose, 1993.

[24] S.J.P. Westen, R.L. Lagendijk, and J. Biemond, "Optimization of JPEG Color Image Coding Using a Human Visual System Model," *Conference on Human Vision, Visual Processing, and Digital Display VII*, **2657**, 370–381, SPIE, San Jose, 1996.

[25] H. Yanagihara, *Image compression based on pattern fineness and edge presence*, United States Patent 5,374,958, 1994.