DCTune perceptual optimization of compressed dental X-Rays

Andrew B. Watson[a], **Mathias Taylor**[b], **& Robert Borthwick**[b]

[a]NASA Ames Research Center, Moffett Field, CA 94035-1000[-]
[b]Foothill College, Los Altos Hills, CA

ABSTRACT

This is a brief report on research on the subject of DCTune optimization of JPEG compression of dental x-rays. DCTune is a technology for optimizing DCT quantization matrices to yield maximum perceptual quality for a given bit-rate, or minimum bit-rate for a given perceptual quality. In addition, the technology provides a means of setting the perceptual quality of compressed imagery in a systematic way. We optimized matrices for a total of 20 images at two resolutions (150 and 300 dpi) and four bit-rates (0.25, 0.5, 0.75, 1.0 bits/pixel), and examined structural regularities in the resulting matrices. We also conducted some brief psychophysical studies to validate the DCTune quality metric and to demonstrate the visual advantage of DCTune compression over standard JPEG.

Keywords: jpeg, compression, quantization, adaptive, image quality

1. INTRODUCTION

1.1. Dental Teleradiography

In current dental practice, x-rays of completed dental work are typically sent to the insurer for verification. It is economically advantageous to transmit instead digital scans of the x-rays. Further economies will be realized if the images are sent in compressed form. This is an example of a class of medical images for which quality standards are arguably lower than in general diagnostic imaging, and thus in which image compression may be less controversial. There is nonetheless a desire in this context for 1) reliable specification of image quality, and 2) optimal quality for a given bit-rate.

1.2. DCTune

DCTune is a method of optimizing the quantization matrix for DCT-based image compression, notably the JPEG standard. The basis of the algorithm is a perceptual error metric which transforms the quantization error into a perceptually meaningful unit, the just-noticeable-difference. We define our perceptual quality measure as the inverse of the perceptual error. The perceptual error model incorporates the human contrast sensitivity function, light adaptation, contrast masking, and error pooling [1, 2, 3, 4]. For a given image, the algorithm will compute an optimal quantization matrix at either a specified bit-rate or a specified quality. The user must provide viewing parameters, such as pixels/degree and average luminance.

In this study, we sought to explore the behavior and value of the DCTune algorithm when applied to digital scans of dental x-rays. Some previous work on application of DCTune to medical imaging is reported elsewhere [5, 6].

2. SOURCE MATERIAL

A total of 60 images were provided by Tau Corporation of Los Gatos, California. These consisted of a representative set of 20 dental x-rays, each scanned at 3 resolutions: 150, 300, and 600 dpi. A tableau of the complete set at 150 dpi is shown in Figure 1.

[-] Further author information - ABW: email: beau@vision.arc.nasa.gov ; WWW: http://vision.arc.nasa.gov

3. IMAGE CROPPING

Our optimization software at present can only deal with images composed of integral DCT blocks. To simplify processing of the images, all were cropped to the nearest multiple of 8 in both height and width, so that each would be composed of an integer number of 8x8 pixel DCT blocks.

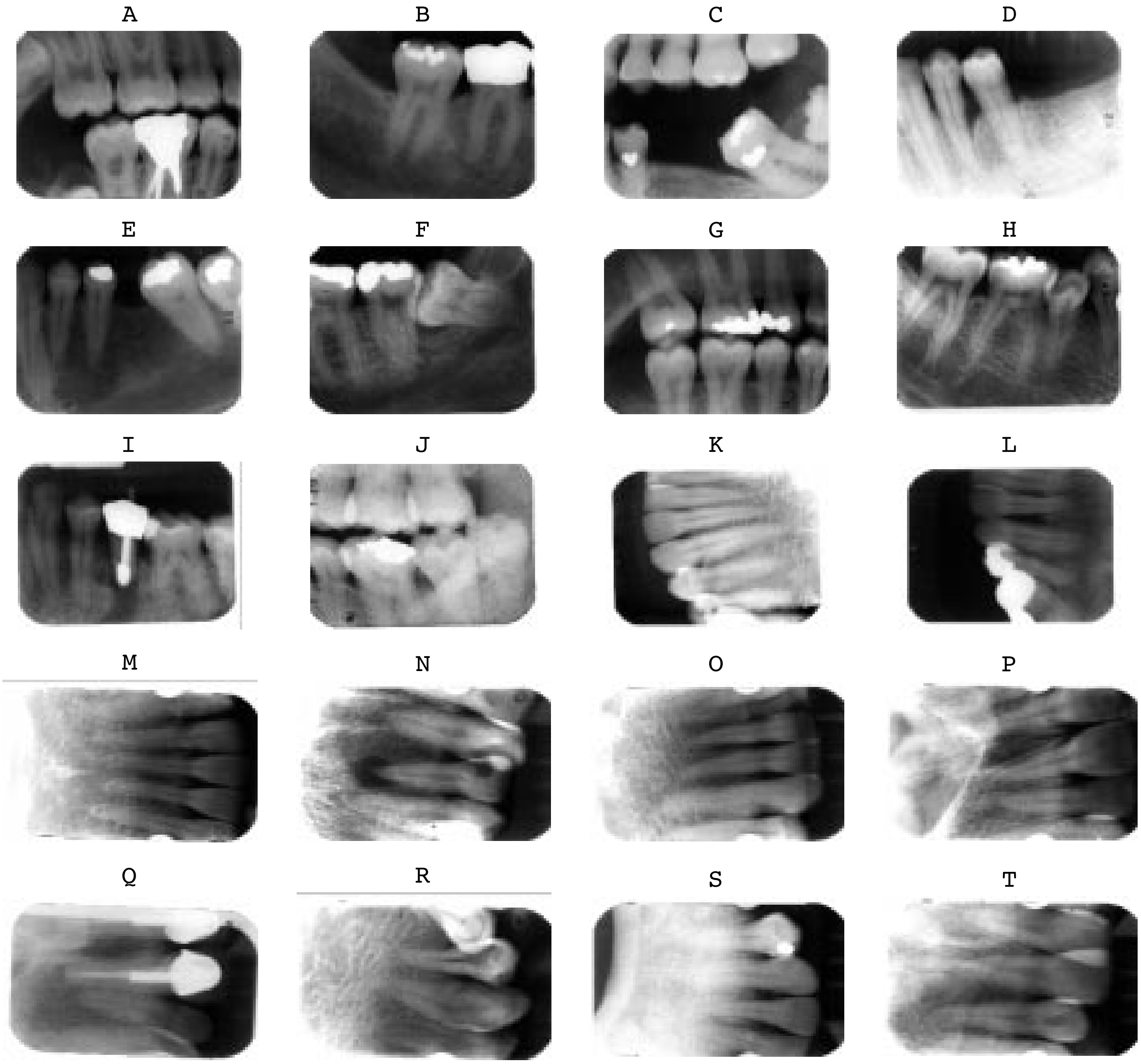


Figure 1. The twenty images used in this study.

4. REGION OF INTEREST

As provided, all of the images consist of a grayscale image upon a white background (see Figure 1). The transition from the typically gray image to a white background is a very large signal which may result in large quantization errors, but which is of no interest to the examiner. Consequently we defined a “region of interest” (ROI) within which quantization errors would be considered significant. While manual construction of such areas might be useful, we automated the process by defining a block to be outside the region of interest if it contained 8 or more white (255) pixels. An example of an image and its ROI are shown in

Figure 2. The excluded region consists primarily of border, but includes occasional portions of the true image (eg fillings). It is unclear whether exclusion of these areas will cause a problem, but a better method for constructing the ROI

could doubtless be found. Because the ROI is defined in terms of blocks, its extent will depend somewhat on the resolution of the image. The example shown is for 300 dpi.

Figure 2. An image and its region of interest (ROI). The ROI is shown in white.

5. OPTIMIZATIONS

The DCTune algorithm requires that the user specify some aspects of the intended display. Here we assumed the following attributes: display resolution = 32 pixels/degree, display mean luminance = 33.5 cd m^{-2}. For each image at each resolution, we computed optimized quantization matrices for four bit-rates: 0.25, 0.5, 0.75, and 1.0 bits/pixel. It was assumed that these would span the range of interest in the present application. The optimization process also generates a set of samples from the function relating bit-rate and perceptual error. We define perceptual quality (often abbreviated here to "quality") as the inverse of perceptual error. In theory, a perceptual error or quality of 1 corresponds to just detectable artifacts. The quantization matrices computed for one image are shown in Table 1. The same four matrices are also shown graphically in Figure 3.

49	94	169	255	255	255	255	255	21	20	21	31	59	255	255	255
88	255	255	255	255	255	255	255	19	28	24	33	85	255	255	255
255	255	255	255	255	255	255	255	20	24	89	255	255	255	255	255
255	255	255	255	255	255	255	255	34	40	255	255	255	255	255	255
255	255	255	255	255	255	255	255	142	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255

12	9	10	14	25	43	117	153	8	6	6	9	13	24	76	117
9	14	11	13	20	46	255	255	6	9	7	8	12	20	255	255
11	12	23	43	108	255	255	255	6	7	12	16	25	72	255	255
15	14	44	153	255	255	255	255	9	8	16	41	97	255	255	255
27	32	82	255	255	255	255	255	16	13	33	63	131	255	255	255
70	53	255	255	255	255	255	255	36	31	65	139	255	255	255	255
114	116	255	255	255	255	255	255	60	58	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255

Table 1. Quantization matrices for image A at 150 dpi at 0.25, 0.5, 0.75 and 1 bits/pixel.

Figure 4 shows the DCTune versions of image A at 150 dpi at the target bit-rates. We defer any precise statement about visibility of artifacts, but note that artifacts are clearly visible at 0.25 bits/pixel, and nearly invisible at 1.0 bits/pixel. The computed qualities of these four images were: 0.197, 0.452, 0.819, 1.26. In theory, then, artifacts should become visible somewhere between 0.75 and 1.0 bits/pixel for this image.

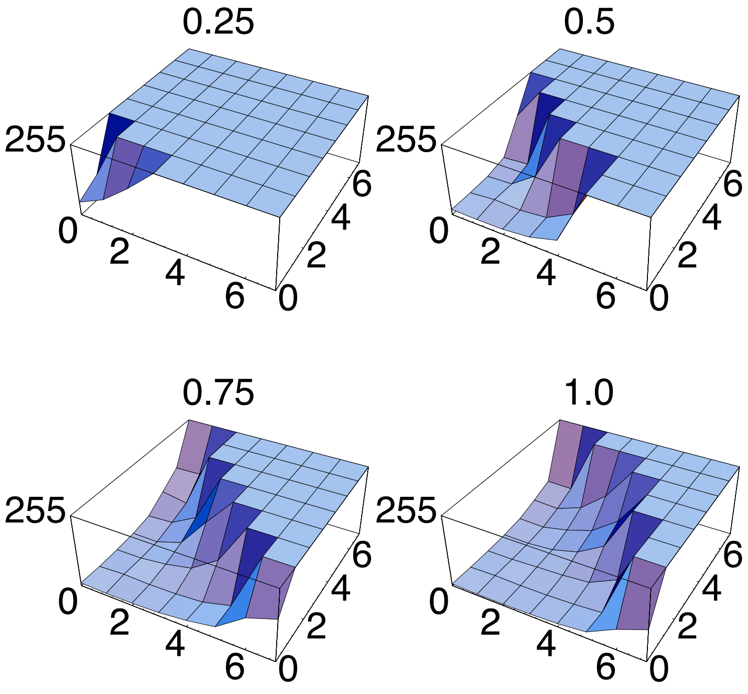
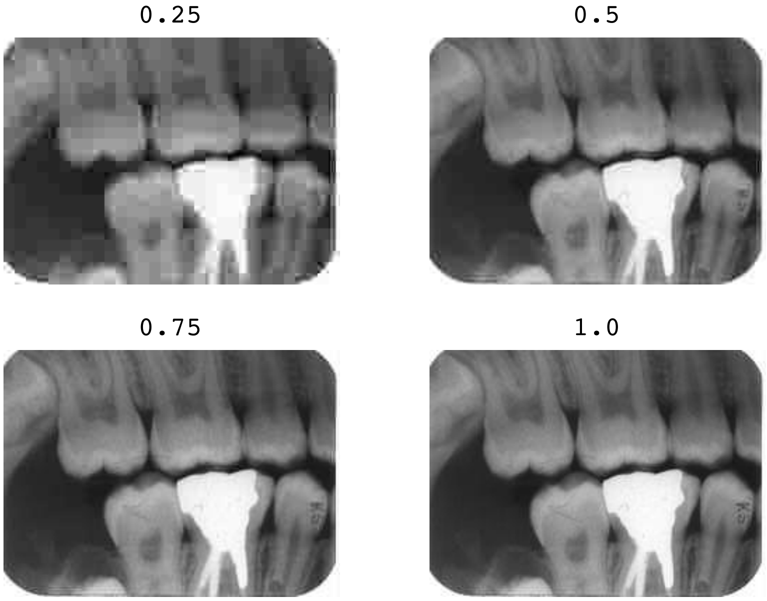


Figure 3. Quantization matrices for image A at 150 dpi at bit-rates of 0.25, 0.5, 0.75, and 1 bits/pixel, shown as surfaces.


0.25

0.5

0.75

1.0

Figure 4. Image A at 150 dpi compressed using DCTune to 0.25, 0.5, 0.75, and 1.0 bits/pixel, shown at approximately actual size.

6. ANALYSES

6.1. *Quality vs Bit-rate*

The first relationship we examine is quality vs bit-rate. This is shown for all twenty 150 dpi images in Figure 5. Recall that perceptual quality is defined as the inverse of perceptual error. For each image, quality is very nearly a linear function of bit-rate. A dashed line has been drawn at quality=1, theoretically the point of perceptually lossless compression. For these images, this level of quality requires between 0.6 and 1.0 bits/pixel, depending upon the image.

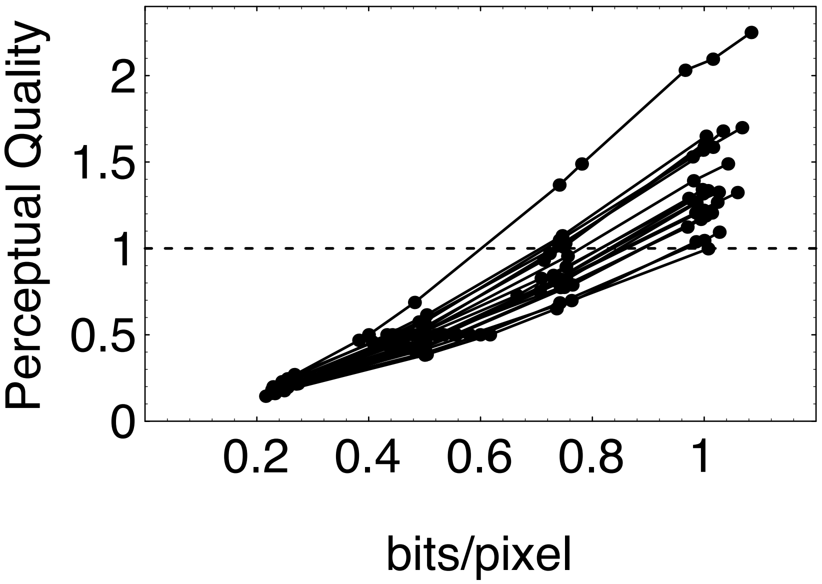


Figure 5. Quality vs bit-rate for twenty dental x-rays at 150 dpi resolution.

6.2. *Gaussian Model for Quantization Matrix*

In earlier research [7] we observed that DCTune quantization matrices are often well-modeled by a two-parameter inverse Gaussian quantization model (IGQM). Specifically, quantization matrix entries Q(u,v) can be approximated by

$$Q(u,v) = a \exp\left(\frac{u^2 + v^2}{w^2}\right) \qquad (1$$

where a and w are the amplitude and width of the Gaussian, respectively, where $\sqrt{u^2 + v^2}$ is the radial frequency, and where u and v are indexed from 0. The amplitude is the quantization value at the lowest frequencies, while the width corresponds to the frequency at which the quantization value has increased by 1/e, or about 37%. We have fit this model to the four matrices computed for each image, corresponding to bit-rates of 0.25, 0.5, 0.75, and 1.0 bits/pixel. The fit minimized the squared error in log Q. The resulting fits for images A-J are shown in Figure 6, which shows log Q vs radial frequency. The quality of fit is remarkably good. The fits for images K-T were equally good.

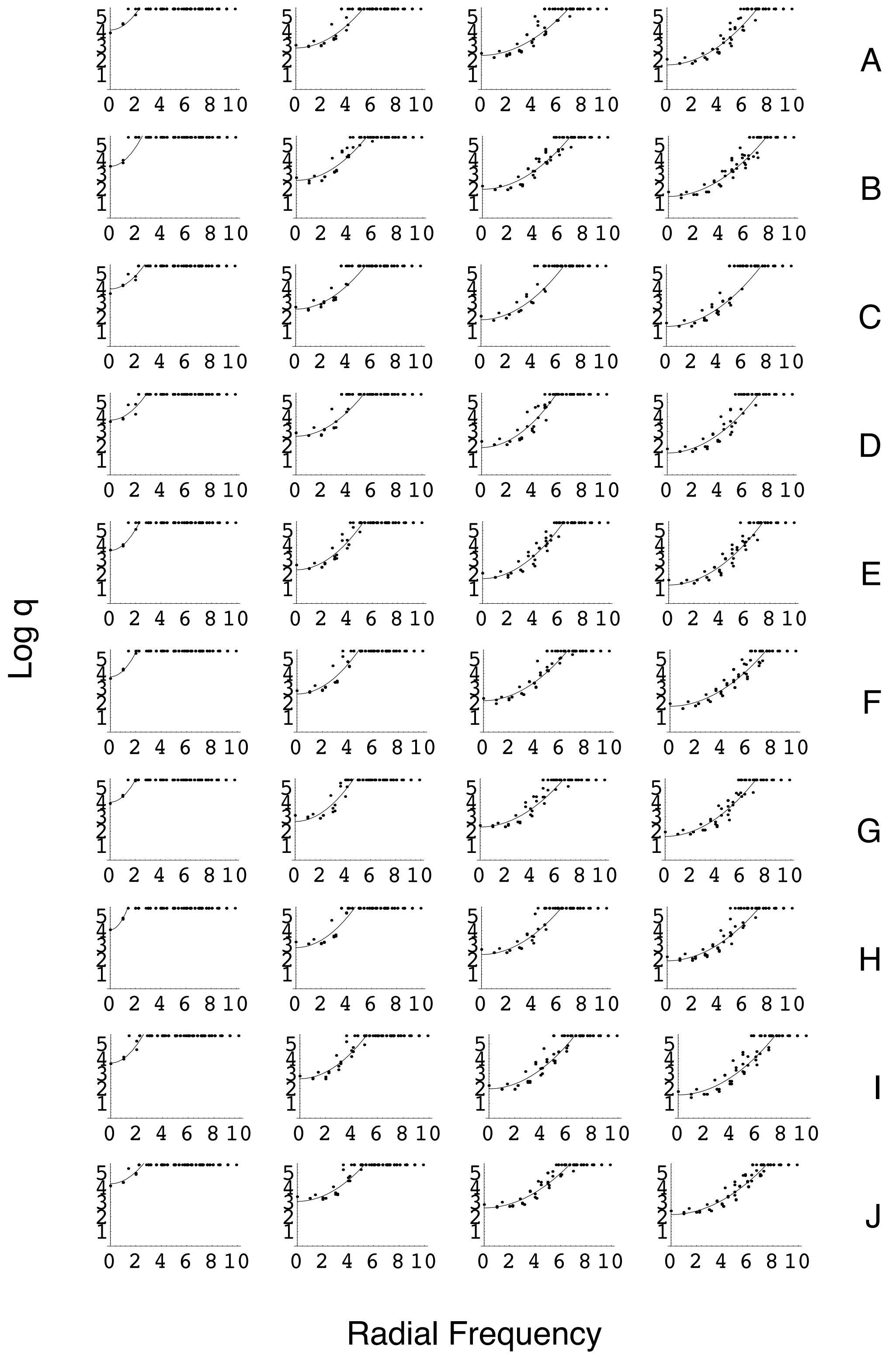


Figure 6. Fit of the Gaussian model to quantization matrices for images A-J.

6.3. Parameter Trajectories

Each fit yields a pair of parameters, amplitude and width. The amplitude is the quantization value at the lowest frequencies, while the width corresponds to the frequency at which the quantization value has increased by 1/e, or about 37%. It is of interest to see how these parameters behave as the quality or bit-rate is varied. One way to visualize this behavior is to plot the trajectory through this two parameter space that is taken by the optimized matrices as the bit-rate is varied. These are shown in Figure 7. There is a remarkable homogeneity to the behavior. Decreases in bit-rate produce changes in amplitude and width that are nearly proportional to one another. This means that tuning the matrix for different bit-rates requires varying both width and amplitude. It should be noted that this differs from the conventional practice with JPEG of varying quality by varying a scalar multiplier of the whole matrix, analogous to varying only the amplitude (i.e. a horizontal trajectory).

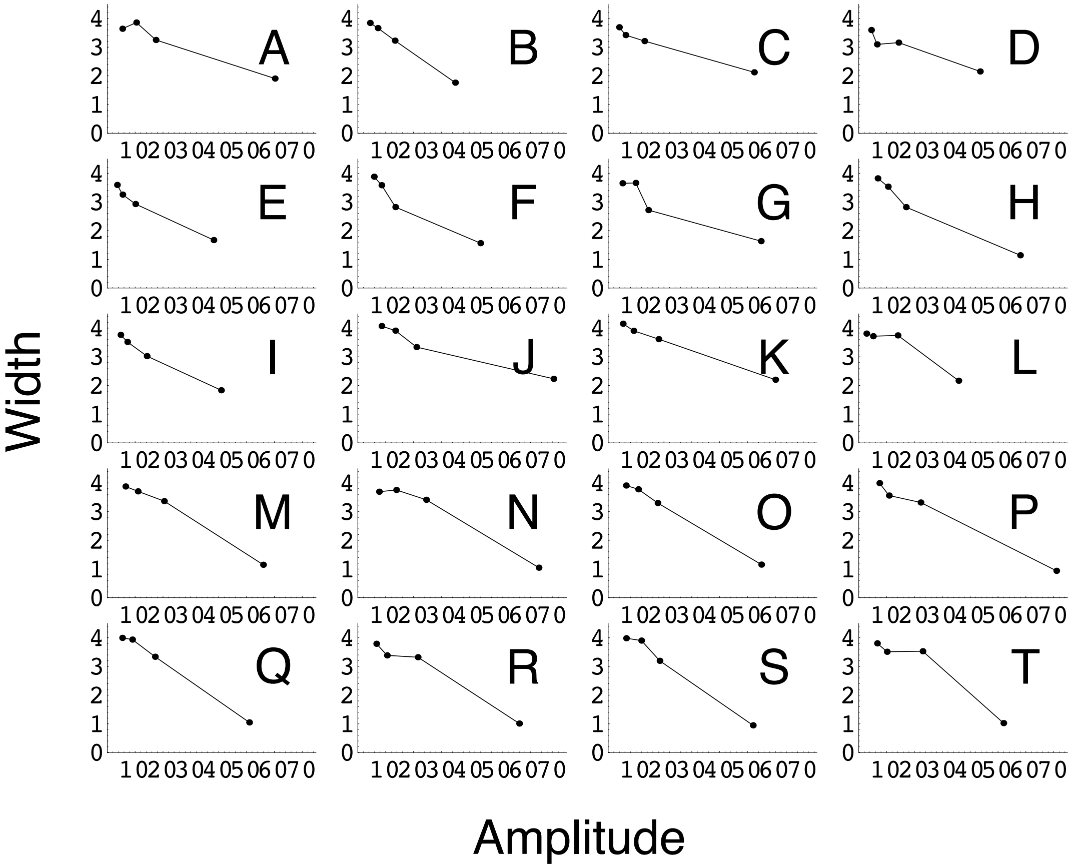


Figure 7. Parameter trajectories for twenty images at 150 dpi. In each panel, the four points correspond to the four bit-rates: 0.25, 0.5, 0.75, and 1 bits/pixel.

6.4. Amplitude vs Bit-rate

A further insight into the model behavior is obtained by plotting the amplitude parameter as a function of bit-rate for all images. This plot is shown in Figure 8. As bit-rate increases from 0.25 bits/pixel, there is a rapid but saturating decline in the amplitude, from a value of about 50 to an asymptotic value of about 5.

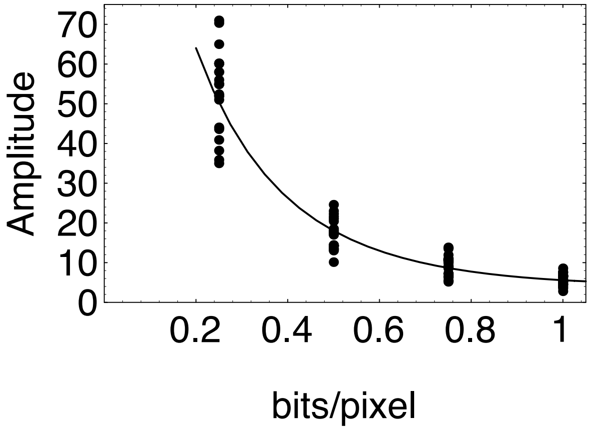


Figure 8. Amplitude parameter as a function of bit-rate for all twenty images at 150 dpi. The curve is Equation 2.

The systematic shape of the data in Figure 8 suggests a functional form. The data are fit reasonably well by an inverse Gaussian function of bit-rate, given by the following equation where *b* is bit-rate in bits/pixel:

$$a = 4.82 \exp\left[\left(\frac{b - 1.25}{0.652} \right)^2 \right] \qquad (2$$

The peak of the inverse Gaussian was placed rather arbitrarily at 1.25 bits/pixel, while the other parameters were obtained by least squares fit. These parameters correspond to the minimum value of a (4.82) and the distance in bit-rate one must go from 1.25 to obtain a 37% increase in amplitude (0.652 bits/pixel). This equation is of course valid only below 1.25 bits/pixel, above that point we assume *a* to remain constant.

6.5. Width vs bit-rate

The corresponding plot of width as a function of bit-rate is shown in Figure 9. It appears very much as the complement of Figure 8. Indeed, we observed earlier that the parameters amplitude and width are linearly related as bit-rate varies, thus we expect that a good describing function would be a linear function of equation 2. Such a function is shown by the curve in Figure 9 and in equation 3 below.

$$w = 3.42 - 0.204 \exp\left[\left(\frac{b - 1.25}{0.652} \right)^2 \right] \qquad (3$$

One feature of the results in Figure 8 and Figure 9 is the evidently greater variability at the lowest bit-rate. This is perhaps not surprising, given the asymptotic nature of the curves, but reminds us that variability may be quite low at practical levels of compression, in this case, probably above 0.5 bits/pixel.

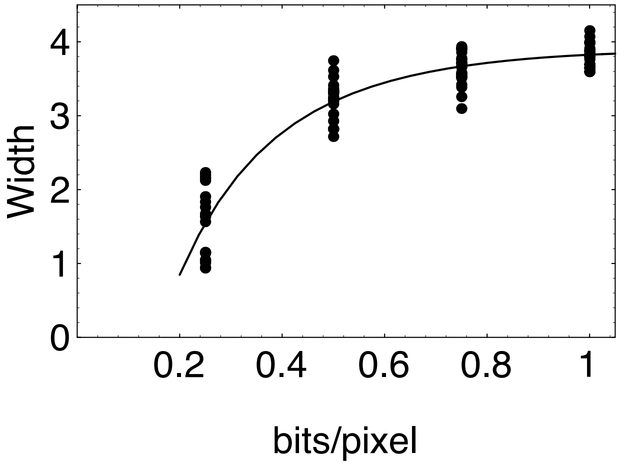


Figure 9. Width as a function of bit-rate for all twenty 150 dpi images. Curve is equation 3.

6.6. Amplitude vs Quality

One advantage offered by the DCTune method is the possibility of adjustable quality. To achieve this it is of interest to know the relation between the amplitude of the inverse Gaussian quantization model (IGQM) and the perceptual quality measure yielded by DCTune. This is shown in Figure 10. We also provide a purely empirical function fit to the data, shown in equation (4) and drawn in Figure 10.

$$a = \exp\left(87.0 - 207.4\, q + 168.2\, q^2 - 43.09\, q^3\right) \qquad (4$$

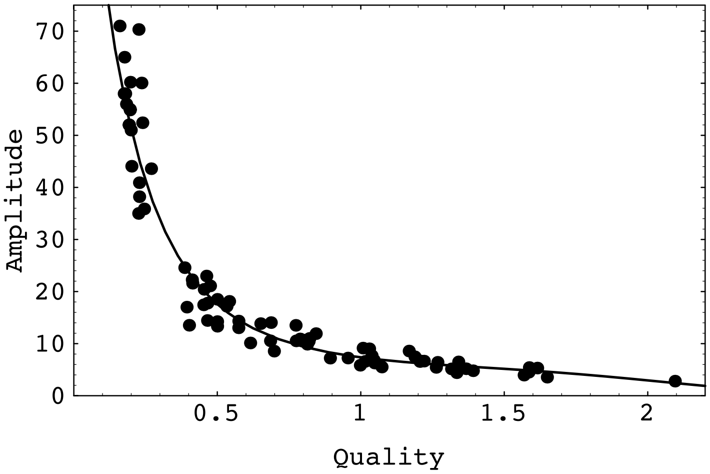


Figure 10. Amplitude of the Gaussian quantization model as a function of perceptual quality. The curve is equation 4.

6.7. Width vs quality

The comparable relationship between width and quality is shown in Figure 11. Again relying on the linear relationship between width and amplitude, we found the best linear function of equation (4), which turns out to be:

$$w = 4.128 - 0.05146\, \exp\left(4.974 - 5.935\, q + 3.923\, q^2 - 0.9645\, q^3\right) \qquad (5$$

This function is shown in Figure 11 and it is evidently a good approximation.

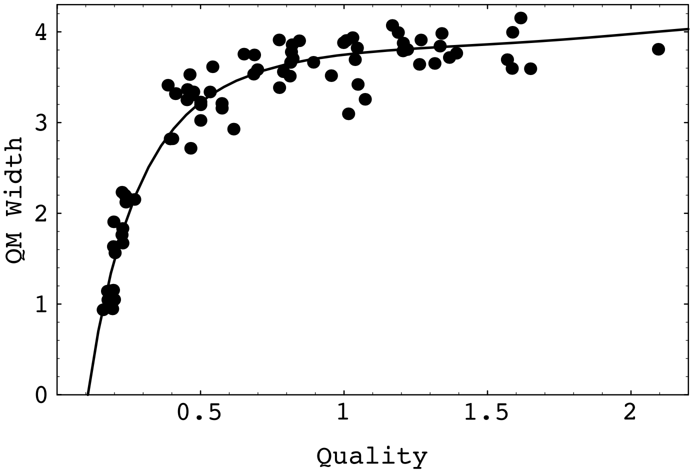


Figure 11. Width vs quality for twenty images at 150 dpi. The curve is equation 5.

6.8. 300 vs 150 dpi

All of the previous analyses were presented for the set of images scanned at 150 dpi. At 300 dpi, the parameters trajectories are similar, but somewhat less orderly. The data relating amplitude to bit-rate show a marked change from 150 to 300 dpi; the points are lower, especially at the lowest bit-rates. This is reflected in the fitted function, analogous to equation (2), which has corresponding parameters of 3.012 and 2.073. This may in part be due to the reduced fraction of the total compressed file that must be devoted to overhead. Likewise, the function relating width to bit-rate is higher, primarily at the lowest bit-rates. In summary, at 300 dpi, quantization matrices are somewhat gentler and broader, especially at the bit-rate of 0.25 bits/pixel. The data and function relating amplitude to quality are, as might be expected, also lower at 300 than at 150 dpi. The corresponding fitted function parameters are: 4.424, -5.777 , 3.8326, -0.977. The data and function relating width to quality are not markedly different at 300 dpi, but show more scatter, and somewhat larger values at the lowest bit-rate.

7. SOME STANDARD MATRICES

For convenience, we use the formulas developed above to compute a set of eight "standard" matrices, for quality values of 0.25, 0.5, 0.75, and 1, at each of the scanning resolutions of 150 and 300 dpi.

41	53	112	255	255	255	255	255
53	68	143	255	255	255	255	255
112	143	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

18	19	26	42	82	195	255	255
19	21	28	46	90	214	255	255
26	28	38	61	120	255	255	255
42	46	61	99	195	255	255	255
82	90	120	195	255	255	255	255
195	214	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

10	11	14	20	35	70	163	255
11	12	15	22	38	76	176	255
14	15	19	28	48	95	222	255
20	22	28	41	70	140	255	255
35	38	48	70	120	240	255	255
70	76	95	140	240	255	255	255
163	176	222	255	255	255	255	255
255	255	255	255	255	255	255	255

7	8	10	14	23	44	95	241
8	8	11	15	25	47	102	255
10	11	13	19	31	58	127	255
14	15	19	27	44	83	181	255
23	25	31	44	72	136	255	255
44	47	58	83	136	255	255	255
95	102	127	181	255	255	255	255
241	255	255	255	255	255	255	255

Table 2. Standard matrices for 150 dpi at quality levels of 0.25 0.5, 0.75, and 1.0.

25	29	46	99	255	255	255	255
29	34	53	115	255	255	255	255
46	53	85	183	255	255	255	255
99	115	183	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

11	12	16	26	52	127	255	255
12	13	18	29	57	140	255	255
16	18	24	39	77	188	255	255
26	29	39	63	127	255	255	255
52	57	77	127	253	255	255	255
127	140	188	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

6	7	9	14	25	56	145	255
7	7	10	15	28	61	158	255
9	10	13	19	36	79	206	255
14	15	19	30	56	122	255	255
25	28	36	56	102	224	255	255
56	61	79	122	224	255	255	255
145	158	206	255	255	255	255	255
255	255	255	255	255	255	255	255

4	5	6	10	17	36	90	255
5	5	7	10	18	39	98	255
6	7	9	13	24	50	125	255
10	10	13	20	36	76	190	255
17	18	24	36	64	136	255	255
36	39	50	76	136	255	255	255
90	98	125	190	255	255	255	255
255	255	255	255	255	255	255	255

Table 3. Standard matrices for 300 dpi at quality levels of 0.25 0.5, 0.75, and 1.0.

8. PSYCHOPHYSICAL EVALUATION

8.1. Experiment 1: Thresholds for Visually Lossless Compression

An important function of a perceptual error metric is that it provides a basis for setting the quality of a compressed image. The DCTune perceptual quality measure offers this possibility. To evaluate the accuracy of the metric we conducted a brief psychophysical study. The goal was to measure the ability of observers to detect compression artifacts at various levels of compression, and from these data to estimate a threshold degree of compression at which artifacts are detected with a certain criterion probability (in this case, 75% correct). The threshold degree of compression can be quantified either in terms of bit-rate or DCTune quality.

Five of the images (A-E) optimally compressed to bit-rates of 0.25, 0.5, 0.75, and 1.0 bits/pixel, along with the uncompressed image, were presented in random order to the observer. Each image was presented a total of 16 times, for a total of 400 trials (5 images x 5 bit-rates x 16 trials). Exposure duration of each image was 1 second. After each presentation, the observer reported (via a key press) whether the image appeared to be compressed, or visually lossless. Viewing distance was 65 cm, yielding 32 pixels/cm on an Apple Macintosh 16 inch color display. Calibrated display software was used allowing specification of the display gamma [8, 9, 10].

Figure 12 shows an example of proportion of "lossless" reports as a function of the DCTune quality measure for one observer and one image (image A, observer SFL). The point for the uncompressed case is plotted arbitrarily at quality=3. The proportions rise rapidly between quality values of 0 and 2. A linear psychometric function has been fit to the data, with both slope and position allowed to vary. Threshold, corresponding to the quality at which 75% of trials result in "lossless" judgments, is then estimated from the line. The values of these estimates for both observers on all five images are shown in Figure 13.

The DCTune error metric is calibrated in such a way that the transition from apparently lossy to lossless should occur at quality = 1. The points in Figure 13 are remarkably close to this value, except for image B for observer SFL, in which none of the examined bit-rates yielded perceptually lossless compression. Observer ABW also showed a somewhat higher threshold for this image. This is evidently due to a particular artifact (between the two teeth) that was highly salient and not properly weighted by the DCTune algorithm.

One caveat must be made. The DCTune error metric is calibrated for a duration of around 1 second. The success of the metric evident in Figure 13 is consistent with this duration, which is analogous to a brief glance at a compressed image. Longer scrutiny will render artifacts more visible. We speculate that DCTune perceptual error (and quality) will vary as the 1/4 power of duration. In other words, increasing the observation period from 1 second to 16 seconds would reduce the perceptual quality by a factor of two. This rule is only expected to work from durations of perhaps 100 msec to durations of perhaps 16 seconds. At shorter durations, quality would increase more rapidly, while at longer durations, quality would no longer decline. This apparently paradoxical effect of duration is due to the narrow meaning of quality used here: invisibility of artifacts.

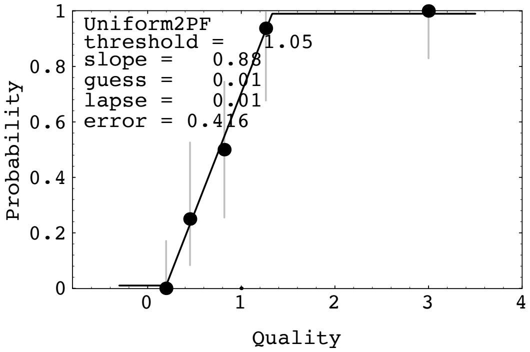


Figure 12. Proportion of "lossless" judgments versus DCTune quality.

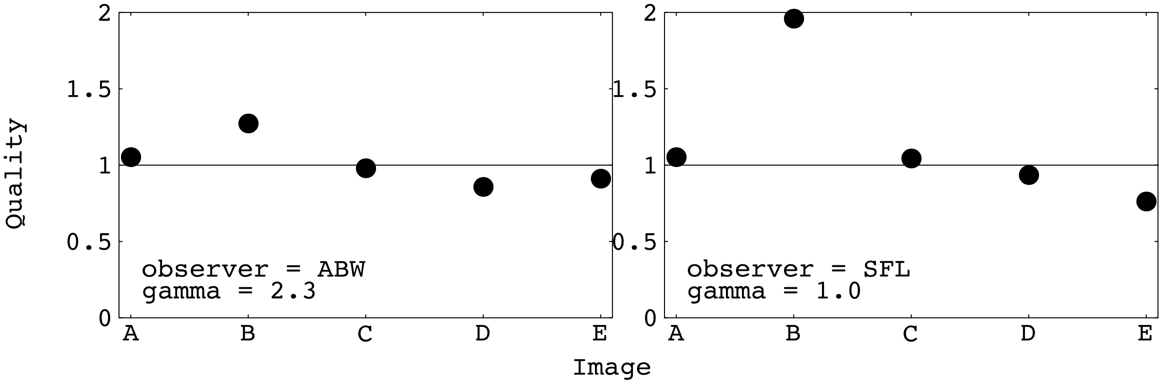


Figure 13. Thresholds for visually lossless compression, in units of DCTune quality, for five images and two observers.

8.2. Experiment 2: Comparative Ratings of JPEG and DCTune

DCTune offers two advantages: controllable quality and optimal quality for a given bit-rate. In the previous section we illustrated the controllable quality aspect. Here we show that, at least in preliminary data, DCTune consistently yields better subjective quality than standard JPEG.

We used methods that closely paralleled those of another recent study illustrating the advantage of DCTune [11]. We first computed eight compressed versions of each of four of the images (A-D), four via DCTune and four via standard JPEG. The four versions for each image and method differed in bit-rate (0.25, 0.5, 0.75, 1.0 bits/pixel). The JPEG versions were produced by scalar multiplication of the example grayscale quantization matrix contained in the JPEG standards document [12].

Each trial consisted of a 9 second presentation of two images, one JPEG and one DCTune, placed vertically above and below the center of the display. The observer responded with a number between -5 and 5, reflecting the relative quality of the two images. The sign indicated whether the upper or lower image was better, and the magnitude, by how much. The complete set of conditions consisted of each possible JPEG version combined with each possible DCTune version, for each source image (4 x 4 x 4 = 64). Four trials were completed for each condition, yielding a total of 256 trials. Display gamma was 2.3, and other experimental conditions were as in experiment 1. Three observers took part in this experiment.

To simplify the results somewhat, in Figure 14 we show only the relative average rating for the two versions when both are at equal bit-rates. When both images are at 0.5 bits/pixel or above, relative ratings are close to zero. This is not surprising, since these images are approaching visually lossless bit-rates (see Figure 12 and Figure 13). In general, differences among compression methods can only be demonstrated when at least one of them is below the visually lossless point. When both images are at 0.25 bits/pixel, the DCTune version shows a sizable and consistent advantage. The advantage is large enough that it can be seen in the printed examples shown in Figure 15. These are the four images used in Experiment 2.

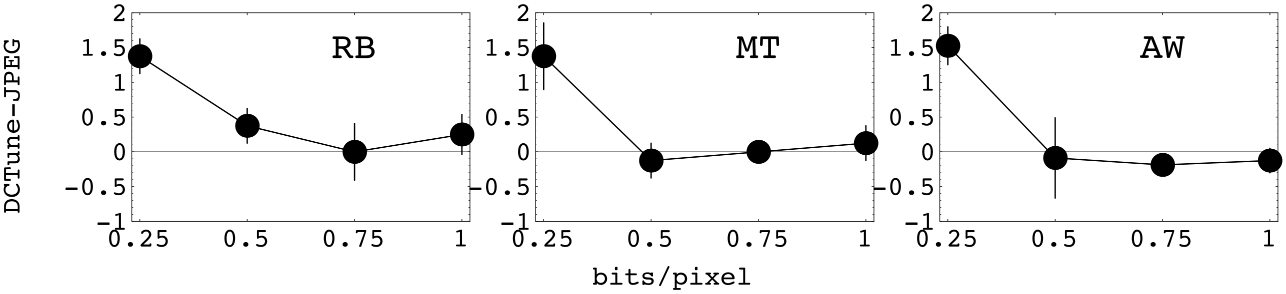


Figure 14. Relative rating of DCTune vs JPEG as a function of bit-rate for three observers.

Figure 15. DCTune (top) and JPEG (bottom) versions of four images at 0.25 bits/pixel.

9. SUMMARY AND CONCLUSIONS

We have performed some preliminary analyses of 20 images of dental x-rays provided by Tau Corporation. DCTune optimized quantization matrices were computed at four bit-rates (0.25, 0.5, 0.75, and 1.0 bits/pixel) for each image at two scanning resolutions (150 and 300 dpi). The following observations were made.

1. DCTune quality is a linear function of bit-rate.
2. The inverse Gaussian model provides a good fit to optimized quantization matrices.
3. At 150 dpi, Gaussian model parameter trajectories are approximately linear and oblique (both parameters must be varied as bit-rate is varied).
4. The amplitude parameter is approximately an inverse Gaussian function of distance in bit-rate from 1.25 bits/pixel.
5. The width parameter as a function of bit-rate is a linear function of the amplitude parameter.
6. The amplitude parameter versus quality is an orderly function, which can be modeled as an exponential of a third order polynomial.
7. The width parameter versus quality is a linear function of the amplitude parameter.
8. In going from 150 to 300 dpi, amplitude parameters are substantially lower and widths larger at corresponding bit-rates or qualities.
9. Psychophysical thresholds for visually lossless compression occur at a DCTune quality value of about 1.
10. At 0.25 bits/pixel, comparative ratings give DCTune a substantial advantage over JPEG. As visually lossless bit-rates are approached, this advantage necessarily diminishes.

10. REFERENCES

1. A.B. Watson,"Perceptual optimization of DCT color quantization matrices," IEEE International Conference on Image Processing, 1, 100-104 (1994).
2. A.B. Watson,"Image data compression having minimum perceptual error," US Patent 5,426,512, (1995).
3. A.B. Watson,"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).
4. A.B. Watson,"DCT quantization matrices visually optimized for individual images," Human Vision, Visual Processing, and Digital Display IV, Proceedings of the SPIE, 1913, 202-216 (1993).
5. M.P. Eckert and D.N. Jones,"Optimized DCT quantization matrices for scanned 12 bit radiographs," Medical Imaging, Proceedings of the SPIE, 2707, (1996).

6. M.P. Eckert and D.N. Jones,"Optimising a quantisation matrix for overlapped transform coding of medical x-ray images," International Picture Coding Symposium, Proceedings, 1, 275-280 (1996).
7. A.B. Watson, A.J. Ahumada, Jr. and M.J. Young, "ICT quantization matrix design for the Galileo S-Band Mission," NASA NASA Technical Memorandum (1993).
8. D.G. Pelli,"VideoToolbox," Spatial Vision, , in press (in press).
9. D.G. Pelli and L. Zhang,"Accurate control of contrast on microcomputer displays," Vision Research, 31(7), 1337-1350 (1991).
10. A.B. Watson and J.A. Solomon,"Psychophysica: Mathematica notebooks for psychophysical experiments," Spatial Vision, in press, (1996).
11. T. van Dijk, J.-B. Martens and A.B. Watson,"Quality assessment of JPEG-coded images using numerical category scaling," European Symposium on Advanced Networks and Services, in press (1995).
12. W.B. Pennebaker and J.L. Mitchell,JPEG Still image data compression standard, Van Nostrand Reinhold, New York (1993).