The Perceived Image Quality of Reduced Color Depth Images

Cathleen M. Daniels and Douglas W. Christoffel
Imaging Research and Advanced Development
Eastman Kodak Company, Rochester, New York

Abstract

A psychometric experiment was conducted with twenty-four observers to determine the perceived image quality of reduced color depth images compared to a 24-bit reference image. The reduced color depth images were created by processing 24-bit images to 8-bit with 5 different algorithms and with palettes containing equal and unequal numbers of colors per channel. The images were viewed with a variety of different display conditions. The 24-bit images and images processed with the sequential scalar quantization algorithm had significantly higher relative image quality ratings than all other algorithms. The relative image quality ratings for the unequal palette were significantly higher than the equal palette. The effects of algorithm and palette on relative image quality ratings depended on display gamma.

Introduction

There has been a large of amount of research devoted to developing improved color-depth reduction (quantization) algorithms for hardcopy and softcopy display systems.[1] But, very little of this research has provided data comparing the perceptual image quality of these algorithms. The goal of the current research was to determine the relative differences in perceived image quality between 24-bit images and 8-bit images generated with various types of color quantization algorithms. It was also of interest to examine the quality of 8-bit images viewed under some typical, but varied display conditions.

Perceived Image Quality Experiment

Observers

Twenty-four Eastman Kodak employees participated in this study. The group was balanced by image evaluation experience and involvement with digital imaging. Eight of the observers were expert print judges. All observers had normal color vision and normal or corrected-to-normal visual acuity. If necessary, they wore corrective lenses.

Experimental Design

To determine relative perceived image quality, 8-bit images were compared to a 24-bit reference image. The experimental variables comprised a fully within-subjects factorial design. There were five algorithms, two fixed palettes, two display gammas, and two display addressabilities. The color quantization algorithms were: multilevel halftone dither patterns (Bayer and blue noise), error diffusion (simple and Floyd-Steinberg), and the sequential scalar quantization algorithm (SSQ). The palettes were: unequal numbers of colors per channel (R = G = 8, B = 4) and equal numbers of colors per channel (R = G = B = 6). Images generated with the SSQ algorithm did not have a fixed palette. The display gammas were original (2.45) and adjusted (1.36). Finally, display addressabilities used in this study were: 832 pixels x 624 lines and 1024 pixels x 768 lines.

The image presentations were blocked by display addressability and gamma (4 blocks). Block presentation order and scene order within block were counterbalanced with a Greco-Latin square design. Image order within scene was randomized (algorithms and palettes). The dependent measure was ratio-scaled image quality.

Scenes

Four different scenes were used for this experiment: Wedpose, Cake, Parkbench, and Pine (Figure 1). The scenes varied in illumination type, camera-to-subject distance, and amount of fleshtones and smoothly varying tonal surfaces. The displayed image size was 286 lines x 429 pixels and the images were horizontal in orientation.

Image Processing

Color Processing. For all algorithms except the SSQ algorithm, images in XYZ tristimulus values were transformed to monitor RGB code values (original and adjusted gammas) with a white point of 9300K (Figure 2). The phosphor matrix was measured using peak R, G, and B code values. To describe monitor gamma, the luminance to code value relationship was measured by displaying a series of seventeen equal code value triplets. The original gamma was approximately 2.45 and the adjusted gamma was approximately 1.36. The gammas were calculated by fitting a power function to the plotted values in Figure 2. The images were then processed through the color quantization algorithms. The SSQ algorithm was specified to accept images in SMPTE RGB code values (CCIR 709 primaries with a D65 white point and 2.2 gamma). Therefore, the images in XYZ tristimulus values were transformed to SMPTE RGB code values and then quantized with the SSQ algorithm. After the images were processed with the SSQ algorithm, they were transformed to XYZ tristimulus values with a phosphor matrix based on the CCIR 709 primaries and the reverse 2.2 gamma look-up table. Then, as for the other algorithms, the images were transformed to monitor RGB code values for original and adjusted gammas.

a) Wedpose

b) Cake

c) Parkbench

d) Pine

Figure 1. Study scenes.

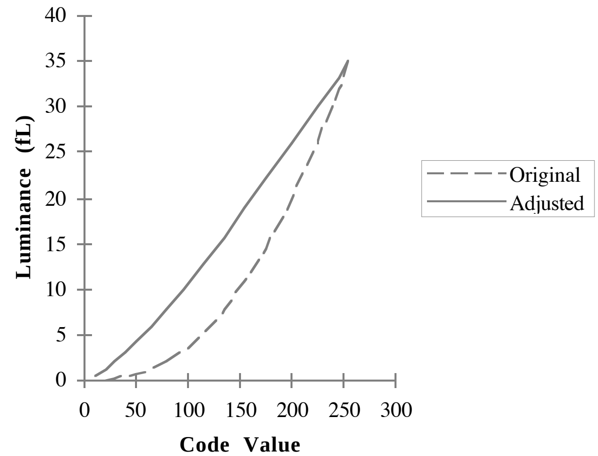

Figure 2. Relationship of code value to luminance for adjusted and original gammas measured on the Sony se II monitor.

Color Quantization Algorithms. The images were quantized from 24-bit to 8-bit by digitally processing the images through five different algorithms: Bayer and blue noise multilevel halftone dither, simple and Floyd-Steinberg error diffusion, and the SSQ algorithm. The algorithms used here are commonly used versions of those cited in other published literature.[1-12] Therefore, the details of the algorithms will not be described in this paper. Dither algorithms use a halftone pattern to simulate the appearance of additional colors that are not present in the palette. A 16 x 16 dither matrix was used for both multilevel halftone dither algorithms. This dither matrix was used to quantize the image code values to create a halftoned image. Because the dither matrix pixel dimensions were smaller than the image pixel dimensions used in this study, the halftone pattern was tiled across the image. The Bayer dither algorithm has been described in more detail elsewhere.[2] The blue noise dither matrix attempts to maximize the high frequency content of the halftone patterns thereby making them less visible to human observers.[3-7] This blue noise halftone pattern has less obvious texture than the Bayer pattern. The Bayer pattern tends to exhibit texture contours in areas of smoothly varying tonal structure (e.g., sky). The error diffusion algorithms also quantize image code values; however, the quantization error is “diffused” to the unprocessed surrounding pixel(s) to preserve the mean pixel values of the image.[9,10] This error can be propagated to the next pixel as in the case of simple error diffusion or the error can be weighted differentially and diffused to surrounding pixels as in the case of Floyd-Steinberg error diffusion.[10] Images generated with simple error diffusion have more “worm-like” artifacts that appear on smooth regions of the image. All of the previous algorithms employ pre-specified fixed palettes, which will be described below. The SSQ algorithm provides image-dependent quantization through calculating a set of palette colors with luminance-chrominance weighting.[11,12] This algorithm also allows for a spatial activity measure, which was not used to process images in the current study.

Palettes. The SSQ algorithm calculated an adaptive palette while the remaining algorithms were processed with two different palettes. These palettes contained equal (R = G = B = 6) and unequal (R = G = 8, B = 4) numbers of colors per channel. The palette consists of an RGB factorial of these code values. The unequal palette contained 256 colors, thus system colors have to be chosen from the palette. The equal palette contained 216 colors with 40 additional system colors that can chosen by the system designer. The ith palette color is given by the nearest integer of i * [255.0 / (n-1)] where, n = number of colors in the channel.

Viewing Environment and Softcopy Display

The study was conducted in a darkened room and observers adapted to the ambient light level during a practice session. The monitors were placed on desks that were covered with non-reflective gray paper. All images were viewed at a constant distance of sixteen inches with a chin rest (Figure 3). The chin rest was covered with non-reflective gray felt and the monitor front encasement was masked with gray cardboard.

Two identical Sony se II monitors were digitally calibrated to the white point for the 9300K Sony software specification. The monitors may be driven at both low and high addressabilities with 24-bit color video cards installed in Macintosh PowerPCs. The reference monitor was configured with the original gamma and an addressability of 832 x 624. The test monitor gamma (adjusted and original) and addressability varied according to the experimental conditions. System software was used to select the appropriate condition. Note that the image size varies with display addressability. Therefore, the images were smaller in the 1024 x 768 condition than the 832 x 624 condition. Images were displayed in random order within experimental blocks with locally written "C" code.

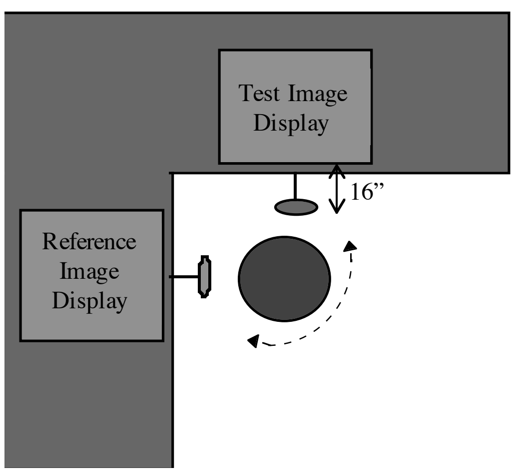

Figure 3. Viewing environment.

Procedure

To rate image quality on a ratio scale, observers compared the reduced color depth images to a 24-bit reference image using fixed modulus magnitude estimation. The reference image modulus was assigned a value of 100. Observers practiced the ratio technique using a set of images processed with all of the image manipulations. They were asked to continue to refer to the reference image when making the ratings. Observers were also asked to judge quality as they normally would when viewing their own images.

Results

Data Analysis. These data were analyzed with analysis of variance. Two separate analyses were performed: with palette as a factor and without palette as a factor. It was necessary to perform the second analysis to assess all algorithms including 24-bit and SSQ within the full experimental design. Therefore, the effect of palette as well as interactions with palette will be reported from the analysis, which includes palette as a factor. Post-hoc tests were conducted for significant main effects. Simple effects F-tests and post-hoc tests were conducted to analyze interactions. Duncan's multiple range test was used for all post-hoc tests. The critical p-value was set at 0.05.

Image Quality Rating Results. There was a significant effect of algorithm on relative image quality ratings [$F(5,115) = 185.73$, $p < 0.0001$]. The 24-bit images and images processed with the SSQ algorithm were rated higher than all of the other images. However, the SSQ images were not rated significantly different than the 24-bit images. There was also a significant effect of scene on relative image quality ratings [$F(3,69) = 15.35$, $p < 0.0001$]. The original gamma was rated significantly higher than the adjusted gamma [$F(1,23) = 48.80$, $p < 0.0001$] and the unequal palette was rated significantly higher than the equal palette [$F(1,23) = 166.53$, $p < 0.0001$]. Although the 1024 x 768 addressability resulted in a smaller image, it was rated significantly higher than the 832 x 624 addressability [$F(1,23) = 50.95$, $p < 0.0001$].

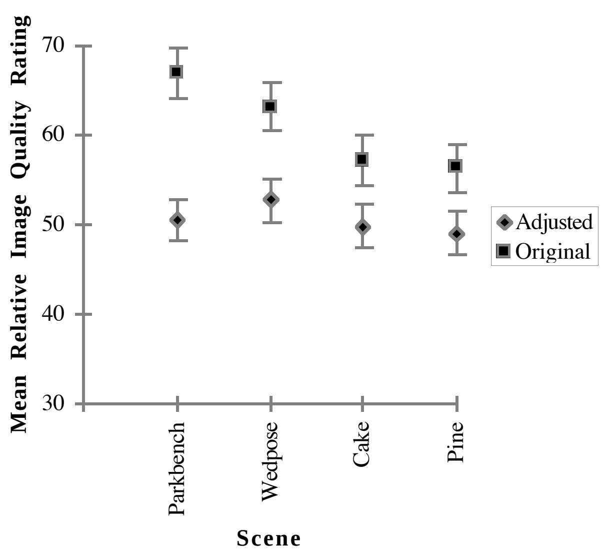

Figure 4. Mean relative image quality and standard error bars as a function of scene and gamma.

There were a few significant two-way interactions, but no significant three-way interactions. There was a significant interaction of scene and gamma [F(3,69) = 10.66, p < 0.0001] (Figure 4) and a significant interaction of algorithm and scene [F(9,207) = 7.91, p < 0.0001] (Figure 5). The Parkbench scene was more sensitive to gamma settings than the other scenes. For the scene and algorithm interaction, the observers' sensitivity to algorithm was only slightly differentially affected by scene. The SSQ and 24-bit images were rated consistently higher than other algorithms for each of the four scenes. There was a significant interaction of algorithm and gamma [F(5,115) = 33.95, p < 0.0001] (Figure 6), as well as a significant interaction of palette and gamma [F(1,23) = 7.66, p < 0.02] (Figure 7). Again, the ordering of ratings by algorithm was not largely different by gamma. The unequal palette for the original gamma was consistently the highest rated condition.

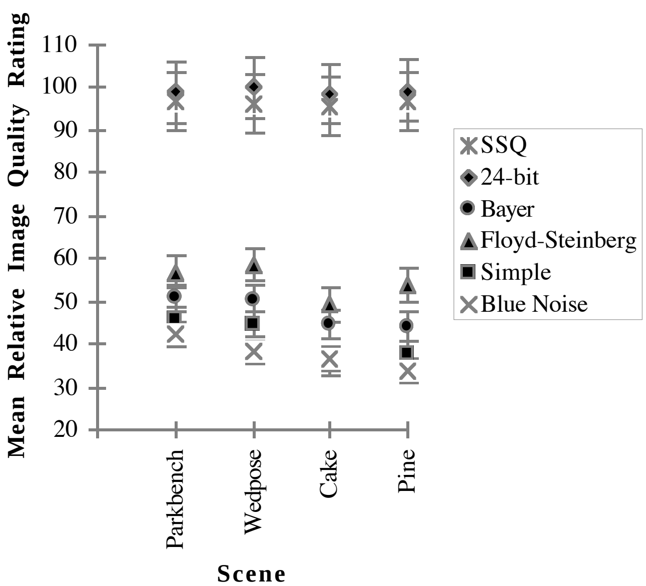

Figure 5. Mean relative image quality and standard error bars as a function of scene and algorithm.

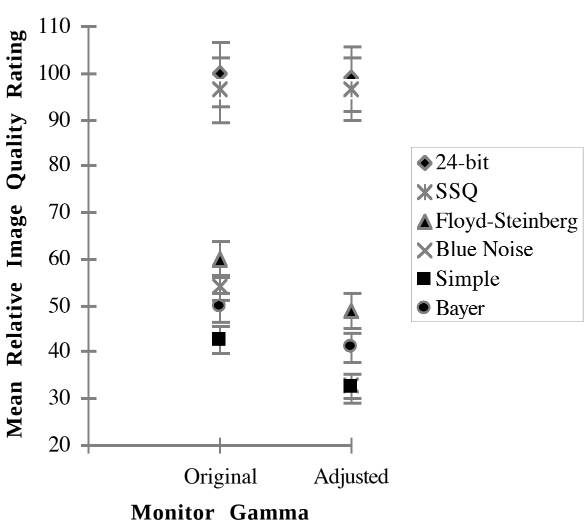

Figure 6. Mean relative image quality and standard error bars as a function of algorithm and gamma.

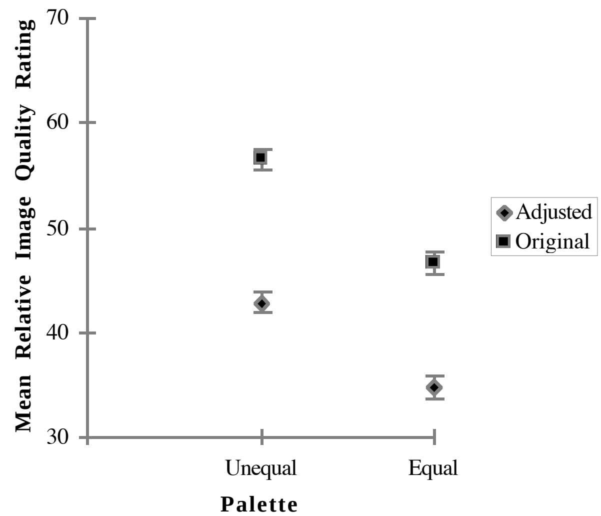

Figure 7. Mean relative image quality and standard error bars as a function of palette and gamma.

Colorimetric Error Analysis

Two of the scenes from the visual experiment were used to gain an understanding of the colorimetric errors introduced by the algorithms. These scenes were chosen based on the number of unique colors or code value triplets that are represented in the images. The Wedpose scene contains the largest number of colors and the Parkbench scene contains the least number of colors. Over one-hundred-fifty 48 pixel x 48 line and 9 pixel x 9 line patches were digitally placed into the center of each of the image combinations described above. The patches included a set of 100 randomly chosen RGB triplets and a set of "typical" scene colors, which consisted of neutrals, fleshtones, and memory colors. This montage was processed through the color quantization algorithms. The images were then transformed to XYZ tristimulus values and average XYZ tristimulus values were calculated for each patch. Colorimetric calculations were made with the appropriate white points for each experimental condition. To determine colorimetric error, CIELAB ΔE^*ab values were determined relative to the 24-bit values for each experimental condition. Based on experience, the monitor characterization is expected to have an error of $\pm$ 1 ΔE^*ab. For this analysis, the error is assumed to be applied equally across all algorithms and monitor conditions.

The mean ΔE^*ab values were very similar for Wedpose and Parkbench scenes for all algorithms except for the SSQ algorithm. Therefore, this difference is only shown for the SSQ algorithm (Figure 8). For the larger patch, the colorimetric error was lower for the SSQ algorithms than the other algorithms. But, for the smaller patch, the colorimetric error was much higher for the SSQ algorithm than the other algorithms. This result is not unexpected because the SSQ algorithm uses a histogram process to calculate an image-dependent palette and the pixel values of the larger patches will have a higher frequency in the histogram.

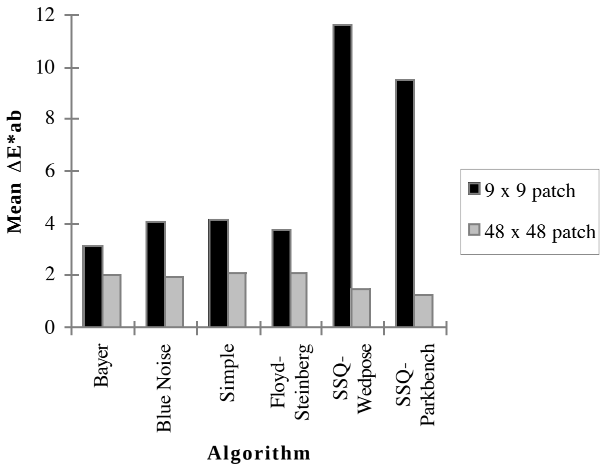

Figure 8. Mean ΔEab as a function of algorithm and patch size.*

Comparison of Results

The image quality results should be compared to the colorimetric errors to better understand the underlying reasons for perceived image quality assessments. As shown in Figure 9, colorimetric accuracy is inversely proportional to mean relative image quality ratings. For the 9 x 9 patch, the SSQ algorithm has slightly higher colorimetric error (lower color accuracy) than the other algorithms, but much higher relative image quality ratings. This result implies that spatial artifacts may be a strong driver of image quality. Indeed, many observers commented that they found the halftone dither patterns as well as error diffusion artifacts to be objectionable.

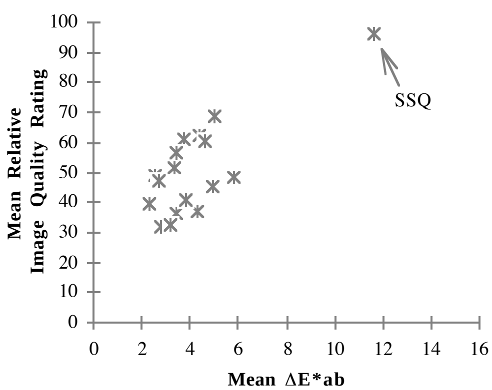

Figure 9 . Relationship of mean ΔEab and mean relative image quality rating for the 9 x 9 patch.*

Conclusions

The adaptive palette SSQ algorithm appears to be perceptually superior to non-adaptive palette algorithms for softcopy applications. We cannot conclude, however, whether this is due to the use of image-specific palette selection or the SSQ algorithm. Colorimetric error was a poor predictor of image quality for images containing visible spatial artifacts. Images processed with the unequal palette appear to be preferred in terms of image quality over images processed with the equal palette. Finally, it should be noted that display variables are quite important to perceived image quality. The algorithms described in this paper provide color quantization in monitor RGB space. Therefore, display gamma and palette may have a combined impact on visible image quality. These factors should be considered when designing web-based and low-cost imaging systems.

Acknowledgments

The authors acknowledge the study observers who patiently provided data. Laurie Schaefer collected and summarized the data. Kevin Spaulding, Peter Burns, Geoff Woolfe, Mike Miller, and Doug Beaudet are gratefully acknowledged for their technical expertise.

References

1. R. Ulichney, *Digital Halftoning*, MIT Press, Cambridge, (1990).
2. B. E Bayer, An optimum method for two-level rendition of continuous-tone pictures, *Proc. IEEE Int. Conf. Commun.,* **26**, 11-15 (1973).
3. K. E. Spaulding, R. L. Miller, and J. Schildkraut, Methods for generating blue-noise dither matrices for digital halftoning, *J. Elect. Imaging*, **6**, 208-230 (1997).
4. T. Mitsa, and K. J. Parker, Digital halftoning technique using a blue-noise mask, *J. Opt. Soc. Am. A*, **9**, 1920-1929 (1992).
5. J. Sullivan and R. L. Miller., Digital halftoning with minimum visual modulation patterns, U.S. Patent 4,920,501 assigned to Eastman Kodak Company (1990).
6. J. R. Sullivan and L. A. Ray, Digital halftoning with correlated minimum visual modulation patterns, U.S. Patent 5,214,517 assigned to Eastman Kodak Company (1993).
7. K. E. Spaulding, and L. A. Ray, Method and apparatus for generating a halftone pattern for a multilevel output device, U.S. Patent 5,586,203 assigned to Eastman Kodak Co. (1996).
8. K. E. Spaulding, D. W. Couwenhoven, and R. L. Miller, Improved error diffusion incorporating a visual model, Presented at Electronic Imaging '98, San Jose', CA (1998).
9. R. S. Gentile, E. Walowit, and J. P. Allebach, Quantization and multilevel halftoning of color images for near original image quality, *J. Opt. Soc. Am. A*, **7**, 1019-1026 (1990).
10. R. W. Floyd and L. Steinberg, An adaptive palette for spatial grayscale, *Proc. SID*, **17**, 75-77 (1976).
11. R. Balasubramanian, C. A. Bouman, and J. P. Allebach, Sequential scalar quantization of color images, *J. Elect. Imaging*, **3**, 45-59 (1994).
12. J. P. Allebach, C. A. Bouman, and R. Balasubramanian, Sequential product code quantization of digital color image, U.S. Patent 5,544,284 assigned to Eastman Kodak Company (1996).