A New Metric for Color Halftone Visibility

Qing Yu and Kevin J. Parker, Robert Buckley and Victor Klassen**
Dept. of Electrical Engineering, University of Rochester , Rochester, NY
**Corporate Research & Technology, Xerox Corporation, Webster, NY*

Abstract

Color halftoning with stochastic processes such as error diffusion and stochastic screening eliminates the Moiré problems since the halftone dots created by these processes are relatively unstructured. This removes the constraints of the rotation angle that apply to conventional color halftoning with ordered dither. In this paper, a series of psychophysical experiments address the visibility of color halftone textures from different schemes of stochastic halftoning as well as Bayer's dithering. Specifically, color halftone patches are displayed on a monitor screen, and the distance at which an observer can confidently see the texture is identified. This distance is used as a measurement of the texture visibility of the corresponding halftone patch. The experimental design and set-up will be outlined first, then the experiment results will be reported. A new metric for color halftone texture visibility is also proposed and it is tested with the experimental results.

Introduction

Color imaging normally requires mixing of three additive primary colors (RGB) for CRT display or subtractive primary colors (CMY) for print. Printing technology can also utilize a fourth primary (K) to provide a better black hue, enlarge the color gamut and improve image quality. Additional colors can be added to further enlarge the color gamut. Color halftoning is the process of generating halftone images for the different color planes for a printing or display device. Color image halftoning is significantly more complicated than halftoning for a gray scale image. All the qualities required of black-and-white halftone images apply to color halftone images which are composed of multiple color planes, and where, the interactions between color planes must be precisely controlled.

In conventional halftoning, the same clustered-dot screen can be used to halftone the C, M, Y, K planes separately to obtain four halftone images, which are then used to control the placing of color on paper. To reduce Moiré patterns which are caused by the low-frequency components of the interference of different color planes, the screens are typically oriented at different angles (usually 30° apart), thereby adding more constraints to the halftone process. Stochastic halftoning (or blue-noise halftoning), such as stochastic screening and error diffusion,[1,2] eliminate the Moiré concern since the color halftone dots created by these processes are relatively unstructured, thus removing the constraints of rotation angle.

To better understand the behaviour of the human visual system (HVS) perception of color halftone patterns, a series of psychophysical experiments have been developed. The experimental results have been used to develop new quality metric for color halftone texture visibility. In the following sections, the outline of the experiment is given first, then experimental results, observations and data analysis are reported.

Experiment Outline

Stimulus

In this experiment, we used displayed halftone patches with different colors and textures as experimental stimuli. These halftone patterns were chosen so that the samples contained common colors (such as skin tone and sky) in color printing as well as halftone textures from different blue-noise halftoning schemes. These halftone patches were created from six uniform contone images (skin color, sky color, green color, 25% gray, 37% gray, and 50% gray respectively) using five different color halftone algorithms: four blue-noise schemes (dot-on-dot, dot-off-dot (or four mask), the adaptive scheme and error diffusion)[3,4] and Bayer's dithering.[5]

The halftone patches were displayed in random sequence on a 21-inch SGI monitor that has a pitch of 0.26 mm with 24-bit-color display ability. Since monitors normally have difficulty reproducing images at the highest spatial frequency, i.e., on-off-on pixel sequence that is very common in blue-noise halftone images, each pixel of the halftone patches (256 by 256 in size originally) was replicated in both horizontal and vertical directions to make a 2 by 2 square. The final patches displayed on the monitor were all 512 pixels by 512 pixels (or about 13.3 cm by 13.3 cm) in size. The test site had conventional office lighting with fluorescent illumination, with no glare visible on the display.

patch	X	Y	Z
50%gray	27.26	29.08	40.27
37%gray	35.34	37.82	51.90
25%gray	43.05	45.98	63.32
green	6.57	12.25	5.28
skin	47.47	46.67	56.81
sky	30.44	33.92	66.09

Table 1: Average tristimulus values of the color patches.

Table 1 lists the average tristimulus values (measured with a colorimeter for display) of the color halftone patches used in this experiment.

Subjects

Eight University of Rochester graduate students between the ages of 20 and 30 were used as subjects. All of these subjects had some imaging background. Before each experiment, all subjects were tested to have normal color vision based on correct reading of a set of isochromatic plates (Ishihara Plates) and approximately 20/20 vision (some with correction lens). At the end of each experiment, a subject was also tested to identify his (or her) actual 20/20 vision distance, which was the longest distance at which a subject could accurately read at least 7 out of 8 characters in the 20/20 line on a Snellen adult eye chart. This distance was used as a normalizing factor in later analysis.

Procedures

Before each test patch was displayed, the subject was asked to stand at a maximum distance (approximately 7 meters) from the monitor where no texture could be perceived for all test patches. Then, a test patch was displayed on a monitor and the subject was asked to walk slowly toward the monitor and stop where he/she could confidently see the texture in the patch. No back-and-forth movement was allowed in this process. The distance between the stop position and the monitor was recorded. To expand the distance range as well as to keep subjects from learning a typical distance, three multi-level (three levels) halftoned patches and the six original contone patches were inserted randomly into the test sequence. Therefore, for each subject, a total of 39 test patches were displayed and the corresponding distances were measured. Subjects were allowed to take a break if they felt fatigued.

Experimental Results

Observations

Before the results are reported, several observations are worth mentioning. To study the repeatability of the experiment, in a preliminary experiment, two test patches were randomly picked and repeated during a display sequence. As the results show, the difference between those two measurements for each repeated pattern is small for most of the observers, as shown in Table 2, where the difference between the two measurements is given as the percentage of the average distance for the repeated halftone patches. Therefore, intra-observer error was not considered a serious issue for this experiment.

Another observation is that by using the normalizing factor of the actual 20/20 vision distance for each observer, the inter-observer variability is greatly reduced, to a final average value of 11.2% over all observers and patches.

Subject	1	2	3	4	5	6	7	8
Difference(%)	5.5	3.5	6	2.4	4	7.9	3.6	2.9

Table 2: Intra-observer variation.

Experimental Data

Table 3 lists the mean value of the distance distribution of the 8 observers for each halftone patch. We also determined that the median value of the same distribution was very close to the mean value for each patch.

original	dod	4m	adap	ed	bayer	multi
50%gray	377	330	316	542	X	X
37%gray	356	291	355	224	282	X
25%gray	324	275	279	213	X	X
green	341	335	367	276	193	233
skin	265	261	296	244	282	204
sky	309	293	369	277	181	198

Table 3: Mean value (across the observers) of the normalized visibility distance (cm) for each halftone patch (dod stands for dot-on-dot, 4m for dot-off-dot, adap for the adaptive scheme, ed for error diffusion, bayer for Bayer's dithering and multi for multi-level dot-on-dot).

The "X" in Table 3 means that data for that cell is either not available (multi-level patches for 25%, 37% and 50% gray) or not reliable (most subjects claimed that the bayer patches for 25% and 50% gray were contone). From Table 3, we can come up with the following observation. For most of the observers, the error diffusion scheme performed well (with small average distance) for all of the six color patches used in this experiment except the 50% gray patch (due to typical error diffusion artifacts). Also, for most of the observers, the dot-off-dot (or four mask) scheme outperformed the dot-on-dot and adaptive scheme, which might suggest that the HVS is less sensitive to texture in chrominance channel than in luminance channel.[6] This agrees with the general vision theory.[7] Finally, for certain Bayer's dithered patches, the texture visibility is extremely low (very short distances); and for some other Bayer's dithered patches, the texture visibility is relatively high. This again agrees well with our experience with the texture variations that exist using Bayer's dithered halftone patterns.

Experimental Data Analysis

Analysis Approach

Because of the nature of the experimental set-up and procedure, this experiment examines the image quality of halftone textures from different color halftone schemes. Silverstein and Farrell[8] had investigated the relationship between the perceived image fidelity and image quality of halftone textures, and according to their report, there is "at most a weak correlation between image fidelity and image quality". Therefore, objective metrics such as frequency-weighted-mean-square-error (FWMSE), for which a corresponding contone image is used as a reference, are not applicable here since the observers had no such reference images. Theoretically, the visibility of halftone texture or the rank order of color halftone rendering by the human visual system (HVS), should be directly related to the distinct characteristics[9] of individual halftone patches, as well as the properties of the HVS. Three issues should be considered before developing a halftone texture metric.

First, in color halftoning, we have to deal with not only texture that is related to luminance channel as in black-and-

white halftoning, but also texture related to chrominance channel and their weightings. It has been widely accepted that the HVS is more sensitive to texture in the luminance channel than in the chrominance channel.[7] Therefore, the first step of a possible approach could be to find out if there is a strong correlation between the texture in luminance channel and the experimental results, in this case, the distance.

The second issue pertains to the nature of this experiment. When an observer is far away from the monitor, the corresponding HVS function will be an extreme narrow-band lowpass filter in the frequency domain, such that higher frequency content is not detected. As he/she moves toward the monitor, the bandwidth of the corresponding HVS filter is gradually expanded so that more and more frequency contents at those sample low frequencies would be captured, which would ultimately contribute to the halftone texture as detected by eyes. Therefore, intuitively, there should be a positive correlation between the low frequency noise and the texture visibility of any halftone pattern.

The first two issues are directly related to the properties of the halftone patterns, which is the objective part of this experiment. The third issue pertains to the subjective part of this experiment, the texture detection model of the HVS. The detection model we propose here is based on the Weber's Law.

A New Texture Visibility Metric Based On Weber's Law

As early as the last century, working mainly with the discrimination of lifted weights, the German physiologist E. H. Weber (1834) discovered that when a stimulus is applied to the human sense organs, there appears to be a lawful relationship between the size of difference threshold and the stimulus intensity level. This relationship is known as Weber's law: the change in stimulus intensity $\Delta\phi$ that can just be discriminated is a constant fraction (c) of the starting intensity of the stimulus ϕ:

$$\Delta\phi / \phi = c; \qquad (1)$$

For human vision,[10] the constant c is found to be around 1/30. This prediction is typically confirmed for a fairly wide range of stimulus intensities, except that c tends to increase greatly at low intensities.

Extending Weber's law to this experiment, we propose a new metric for color halftone texture visibility. This metric is outlined as following:

1. Define the displayed halftone patch as h(x,y), which has three color planes RGB.
2. Make the RGB to XYZ conversion with the following matrix transform to get the XYZ representation of the image:

$$XYZ = M * RGB \qquad (2)$$

where M is a 3 by 3 matrix for the monitor used in this experiment:

$$M = \begin{vmatrix} 0.3944 & 0.3235 & 0.2161 \\ 0.2050 & 0.7089 & 0.0861 \\ 0.0217 & 0.1355 & 1.2091 \end{vmatrix} \qquad (3)$$

Now the Y plane y(x,y) has the luminance image.

3. Define a new variable as WebLimit, which is given by the following equation:

$$WebLimit = AVG * Webfactor. \qquad (4)$$

where AVG is the average (AVG) of the luminance image, the Webfactor is set as 1/30 for all the color patches.

4. Take a Fourier transform (FT) to find the corresponding Y(K,L), the luminance image in frequency domain.
5. Define a variable F_{cf} as cut-off radial frequency, and set the starting value of F_{cf} to 1.
6. Define a new image C(K,L) in frequency domain according to the following equation:

$$C(K,L)= Y(K,L), \text{ if } R(K,L) \le F_{cf} = 0, \quad \text{otherwise} \qquad (5)$$

where $R(K,L) = \text{sqrt}(K^2+L^2)$, the corresponding radial frequency of location (K, L) in the frequency domain.

7. Take the inverse Fourier transform (IFT) of C(K,L), and find the corresponding image domain pattern c(x,y).
8. Calculate the standard deviation (SD) of c(x,y).
9. Compare the value of SD with WebLimit defined in [3]. If SD is less than WebLimit, then increase F_{cf} by 1 and go back to [6]. Otherwise, report the current F_{cf} value and exit.

In this process, the iterations of increasing the cut-off radial frequency mimic the HVS contrast sensitivity adjustment during the physical process of walking slowing toward the monitor, and the minimum value of F_{cf} is searched such that:

$$\left\{ \frac{SD\{Y(x,y) * LPF(x,y \mid F_{cf})\}}{\bar{Y}} \right\} \ge Webfactor \qquad (6)$$

where SD{} is the standard deviation function, LPF() is a lowpass filter with cut-off radial frequency F_{cf}.

One advantage of this model is that, as Weber's law defines a relative measurement of intensity discrimination (in this case, the luminance), we are able to study the correlation between luminance texture and visual response for all halftones, no matter what the original color is or what halftone algorithm is used. Simply put, the current study can be narrowed down to a 2-D space. In one dimension we can rank the halftone patterns by the distances at which the observers first reported seeing the texture; in another dimension we rank the halftone patterns by the cut-off frequency when WebLimit is exceeded for the corresponding luminance image.

Table 4 lists the cut-off radial frequency for each halftone patch. A regression is carried out to find out the linear relationship between 1/cut-off-frequency and the mean distance across the observers for each color patch, and a R^2 value of 0.57 is obtained. However, when all the green patches are excluded, a R^2 value of 0.86 is obtained. As given by Table 1, all the green patches have relatively low (around 10) luminance intensity compared with other patches. According to Weber's law,[10] the Webfactor tends to

increase greatly at low intensities, therefore, we believe that a larger value of Webfactor should be used for the green halftone patches. The cut-off frequencies for all the green patches are recalculated with a Webfactor of 1/15, and the results are given in Table 5.

original	dod	4m	adap	ed	bayer	multi
50%gray	30	41	43	16	181	X
37%gray	38	47	40	56	45	X
25%gray	44	49	49	56	128	X
green	24	26	24	21	45	33
skin	50	51	47	51	45	67
sky	40	45	38	50	91	61

Table 4: Threshold frequency for each halftone patch when WebLimit is exceeded. The number is relative to a 256x256 discrete Fourier transform of 256x256 luminance images.

	dot	4m	adap	ed	bayer	multi
green	38	38	38	33	64	52

Table 5: Cut-off frequencies for green patches when Webfactor equals to 1/15.

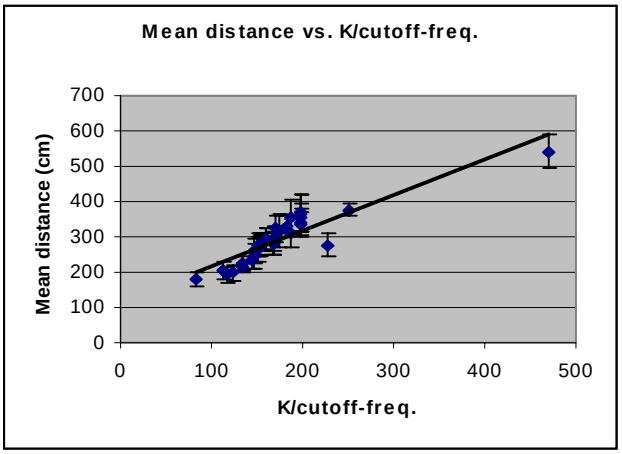

Figure 1: Mean distance of texture visibility vs. K/cutoff-frequency (K=7527) for each halftone patch.

Figure 1 shows mean distance of texture visibility vs. K/cutoff-frequency (K=7527) for all the 31 halftone/ multitone patches after Table 4 is updated with Table 5. As it shows, the new metric gives a consistent ranking of the texture visibility for different halftone patches (in terms of the distance) based on the corresponding cut-off radial frequencies. A linear fitting is further carried out and a R^2 value of 0.81 is obtained.

Discussion

Zhang et al.[11] recently proposed a color image quality metric S-CIELAB to predict texture visibility of printed halftone patterns (illustrated in Figure 2).

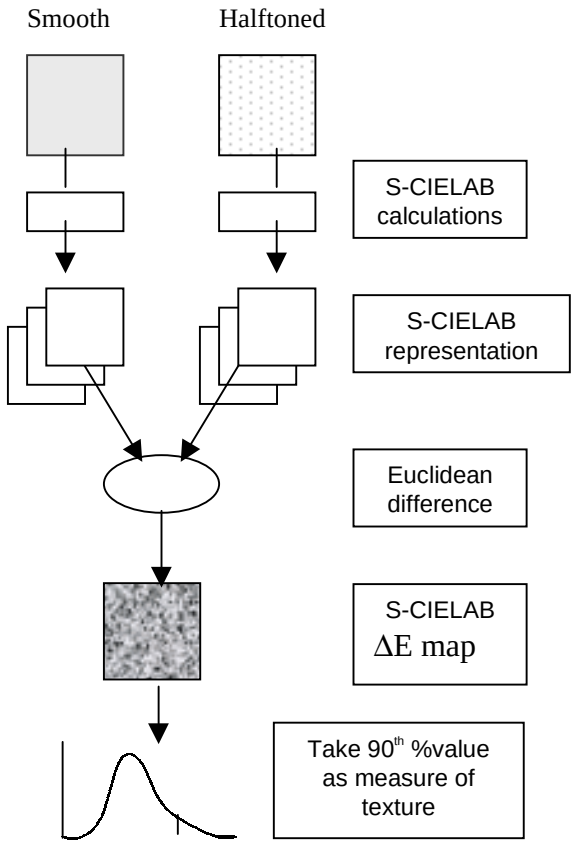

Figure 2: S-CIELAB metric for halftone texture visibility

In the S-CIELAB calculation, the original image is first converted into one luminance and two chrominance color components. Each component image is then passed through a spatial filter that is selected according to the spatial sensitivity of the human eye for that color component. The filtered images are transformed into XYZ color space and finally into CIELAB space. The S-CIELAB calculation also requires targeted printer or display resolution and viewing distance for the design of the spatial filters.

As Figure 2 shows, the S-CIELAB calculations require a reference image to calculate the ΔE map (the 90th percentile value of ΔE map is used as the metric value). For this experiment, since no such reference images were shown to the observers, the average value of each halftone patch was used as the reference image, and the 90% ΔE was calculated using CIE94 color difference formula.[12] The default settings for the S-CIELAB calculation (72dpi resolution and 10 inch viewing distance) have been used.

Figure 3 shows the mean distance vs. 90% ΔE for each halftone patch used in the experiment. Ideally, we were expecting that the S-CIELAB metric values should correlate with the visual distance measurement. However, as Figure 3 shows, this is not the case. One typical observation from the S-CIELAB metric values is that it consistently predicts lower halftone texture visibility for Bayer's dithered patches, i.e., the 38% gray and skin patches from Bayer's dithering have smaller 90% ΔE values compared to those of patches from error diffusion. This doesn't agree with our

experimental results, which were more accurately predicted by our new metric that is based on Weber's law.

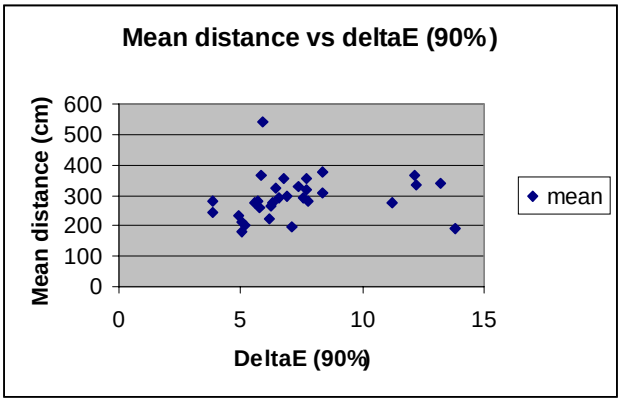

Figure 3: Mean distance vs. 90% ΔE (predicted by S-CIELAB metric) for each patch.

We have demonstrated that our simple model performs significantly better at ranking halftones than S-CIELAB with the default settings of 72 dpi and 10 inches viewing distance. This raises some interesting questions. It is easy to observe that no halftone scheme looks particularly good when displayed at 72 dpi and viewed at 10 inches.

Zhang et al.[11] varied the contrast of halftones in order to derive a quality metric, which provided a better correlation with S-CIELAB. The most notable difference between their data and calculations and ours was that the print resolution was 150 dpi for dispersed dither, while the resolution in our calculation was 72, about half theirs. Both Zhang et al. and we used a fixed distance and resolution in our S-CIELAB calculations: the key differences are in the choice of resolution and the psychophysical experiment.

Because actual viewers will evaluate halftones based on viewing them at varying distances (changing their contrast is not feasible), our experiment may provide a more meaningful ranking. We hypothesize that the fundamental difference in our results derives from the difference in experimental techniques: distance vs. contrast variation. This raises three questions. First, would the S-CIELAB predictions be near constant had we used the distances in Table 4 in the S-CIELAB calculations? Second, by varying the viewing distance used for S-CIELAB to find a distance with a constant S-CIELAB prediction, would we end up with distances that correlate well with our measurements? And finally, how well would S-CIELAB perform at ranking the halftones given a fixed viewing distance and resolution typical of print media? Answers to these questions remain as future work.

Some similarities exist between our new texture visibility model and S-CIELAB. For example, S-CIELAB and our model both distinguish the luminance channel (we ignore the remaining channels, they filter them more heavily); both models filter the luminance channel, removing high spatial frequencies: S-CIELAB uses a Gaussian-like filter, while we use an ideal lowpass filter with a varying cut-off. Further experiments will explore this key issue.

Conclusion

In this paper, we reported a visual experiment for studying visibility of color halftone texture. A new metric, which combines low frequency noise measurement with a detection model based on Weber's law, is further proposed. We have shown that this new metric for color halftone texture visibility correlates well with the experimental results. This enables an automatic and objective metric for ranking the qualities of different halftone patterns, where quality is defined by the distance where the HVS perceives halftone texture.

Reference

1. T. Mitsa and K. J. Parker, "Digital halftoning using a blue noise mask," Journal of the Optical Society of America A, vol. 9, pp. 1920 – 1929, Nov. 1992.
2. R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial greyscale," Proceedings of the Society for Information Display, vol. 17, no. 2, pp. 75 – 77, 1976
3. Q. Yu and K. J. Parker, "Adaptive color halftoning for minimum perceived error using the blue noise mask," in Proceedings, SPIE Electronic Imaging Conference: Color Imaging: Device – Independent color, Color Hard Copy, and Graphic Arts II, vol. 3018, pp. 272 - 277, 1997.
4. Q. Yu and K. J. Parker, "Stochastic screen halftoning for electronic devices," in printing, Journal of Visual Communication and Image Representation, vol. 8, December, 1997.
5. B. E. Bayer, "An optimum method for two-level rendition of continuous-tone pictures," in Proceedings, IEEE International Conference on Communication, pp. 2611 – 2615, 1973.
6. Q. Yu, K. J. Parker, and M. Yao, "Color halftoning with blue noise mask," in Proceedings, Fourth Color Imaging Conference: Color Science, Systems, and Applications, pp. 77 – 80, IS&T/SID, 1996.
7. T. Cornsweet, Visual Perception. New York, NY: Academic Press, 1971.
8. D. A. Silverstein and J. E. Farrell, "The relationship between image fidelity and image quality," in Proceedings, IEEE International Conference on Image Processing, vol. I, pp. 881 – 884, 1996.
9. Q. Yu and K. J. Parker, "Quality issues in blue noise halftoning," in Proceedings, SPIE Electronic Imaging Conference: Color Imaging: Device – Independent color, Color Hard Copy, and Graphic Arts III, vol. 3300, pp. 379 - 385, 1998.
10. G. Gescheider, Psychophysics Method, Theory, and Application. Hillsdale, NJ: Lawrence Erlbaum Associates, 1985.
11. X. Zhang, D. A. Silverstein, J. E. Farrell, and B. A. Wandell, "Color image quality metric S-CIELAB and its application on halftone texture visibility," IEEE Comp. Confer., 1997.
12. R. Berns, "Colorimetry I," Optics and Photonics News, pp. 23-27, September, 1995.