27.1: Determining the Appropriate Color Bit Depth for a Small Portable Display

Jennifer Gille, Alan Lewis, Scot Hastings and Clarence Chui
QUALCOMM MEMS Technologies, San Jose, CA, USA

Abstract
Appropriate color bit depth for user acceptability of small portable displays depends on contrast ratio, gamut, gamma, pixel pitch, dither, size and uses. We determined target bit depth values for an interference-color reflective display, testing acceptability and artifact visibility using diverse observers and images in a display-simulation optical system.

1. Introduction
Displays differ widely in the principles employed to produce color images, but all developers share the goals of establishing a color tonescale that is without unacceptable banding or dither noise, and of doing so at the lowest possible manufacturing cost. If bit depth is a factor in the performance/cost trade-off function, then determining the right bit depth for a display will be an important element of the development process. The small size and intermittent use of small portable displays suggest that while well-established threshold functions for visibility of artifacts will apply, we still need to test potential display configurations for acceptability to viewers.

QUALCOMM's reflective IMOD (*i*nterferometric *mod*ulator) display uses interference colors generated by a reflective membrane suspended over a thin film stack in a two-plate MEMS configuration [1]. The gamut is determined by the plate gaps for the red, green and blue channels. Grayscale is produced spatially by subpixel area, so the gamma (digital input to luminance transfer function) is linear. Pixel pitch is based on MEMS feature size, and contrast ratio is a complex function of the gamut and feature size. Our goal was to determine the most appropriate bit depth for an IMOD display in a 1.8" diagonal format suitable for a small portable device.

2. Contrast and Tonescale
There is considerable guidance in the vision science literature concerning the visibility of color differences for adjacent light-emitting or reflective areas. We know that the visibility of color tonescale steps, and therefore appropriate bit depth, depends on maximum white/black contrast ratio, color gamut, gamma, and pixel pitch if dither (halftoning) is used [2].

Contrast ratio and color gamut depend not only on the choice of primaries, their brightness and relative strengths, but also on viewing environment. Figures 1 and 2 show actual measurements of two displays in current commercial products and illustrate this point. The maximum white luminances, measured in the dark, of the CSTN and OLED displays are 118 and 59.5 cd/m^2 respectively.

At ADEAC 2006 [3], we showed, using the above data, that in office-level lighting for the two commercial devices, 24-bit color is extremely redundant. Our method was to calculate $\Delta E_{ab}*$ for each actual color point with its 26 nearest neighbors and to examine the histogram of the results for values less than or greater

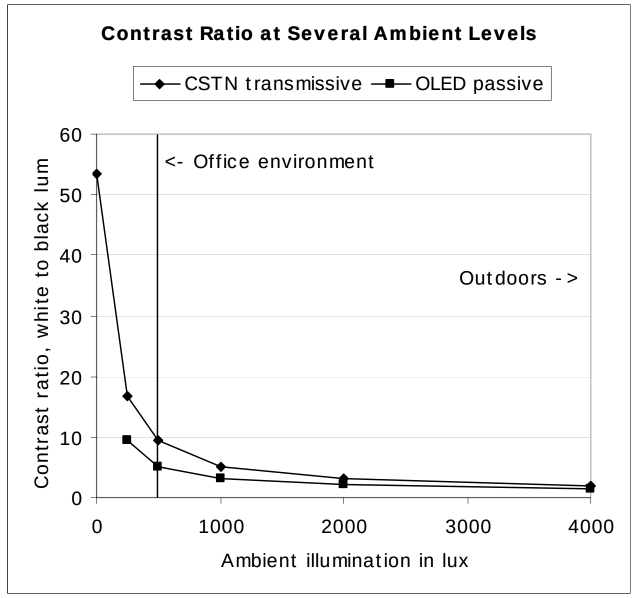

Figure 1. White/black contrast ratio reduction for two commercial displays in ambient light.

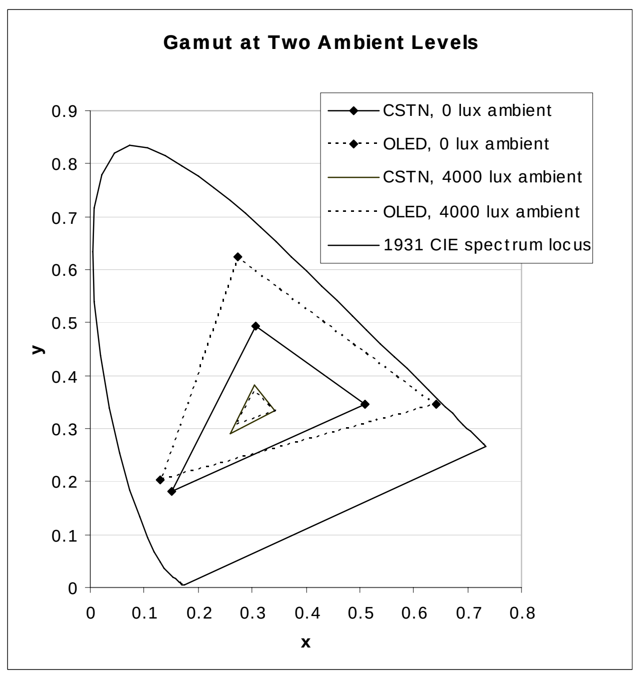

Figure 2. Gamut shrinkage for two commercial displays in ambient light. (CIE 1931 x,y diagram)

than a nominal Just Noticeable Difference (JND) of $\Delta E_{ab}^* = 2.3$ [4, 5]. The same computational technique can be used to estimate the values that should be tested in determining a target bit depth.

There is yet a good deal of myth surrounding the digital input to the luminance transfer function, or gamma, of a display; "gamma" being shorthand for the exponent in a power-law relationship. In the analog TV world, gamma is manipulated to enhance the apparent contrast of an image. In the digital image world, it represents both a digital coding scheme and a display parameter. It has been shown [6] that a power-law function with gamma = 2.2 is not necessarily most like the human visual response, and that the growth function curve that is the result of applying the more justifiable Weber fraction varies in shape depending on the white/black contrast ratio. The smaller the contrast ratio, the shallower the curve, as shown in Figure 3.

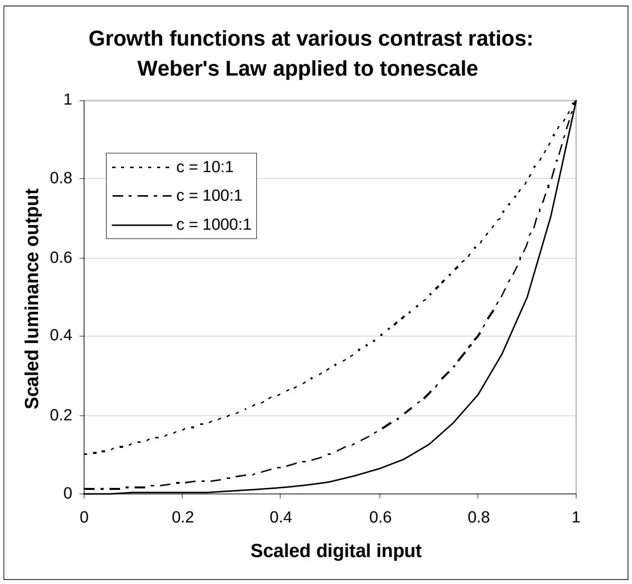

Figure 3. Constant Weber fraction curves for three black/white contrast ratios.

Images encoded in sRGB (and many other coding schemes) use the equivalent of a gamma of 2.2. If the encoding digitization produced steps that were approximately Just Noticeably Different, then a display with a linear gamma parameter (gamma = 1, for instance with spatial or temporal grayscale) such as the IMOD would need extra bits for the levels required to successfully display the low end of the tone curve, thus wasting levels at the high end. When the contrast ratio is low, however, the 2.2 gamma encoding scheme generates more levels than needed in the low end of the tone scale, and the disadvantage of the mismatch between the linear display gamma and the 2.2 encoding scheme is greatly mitigated. Color processing may be applied to recover apparent contrast, but extra bits are not necessarily needed. Of course, even if a display does have a gamma of 2.2, reduced contrast reduces the number of needed bits, as in the ADEAC results cited.

3. Dither, Spot Size and Tonescale

The threshold Weber fraction of 1-2% is derived from relatively large-area tests of spot visibility. As spot size decreases, the contrast needed to discriminate the spot from background increases. Vision models will take spot size into account when analyzing discriminability, and an accessible tool of this type is available in SCIELAB, a spatial extension to the CIE L*a*b* deltaE color difference metric [7-9]. Input to SCIELAB is two images that are to be compared, and output is a point-by-point image map of dE*ab values.

In order to derive a plot of threshold Weber fraction against decreasing spot size for grayscale stimuli, we generated a large series of images of a single pixel on a 20x20 pixel background at various pixel sizes, contrasts, background levels, and both polarities, and compared them to images of their respective backgrounds alone. We defined contrast as the difference between the single-pixel luminance and the background luminance divided by the background luminance. We defined threshold as max(dE*ab)=2.3 in the output map [4, 5]. The results, showing increased contrast needed for smaller size, are plotted in Figure 4. Separate curves are shown for the separate background and polarity conditions because the measure of contrast we used is not the same as the dE*ab measure.

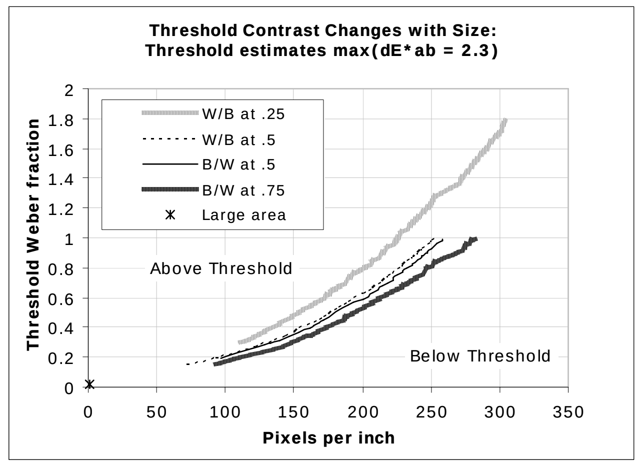

Figure 4. Threshold Weber fraction curves derived using SCIELAB. In the legend, B/W refers to dark spots on a lighter background, and W/B the opposite; .25, .5, and .75 refer to the background luminance as a proportion of maximum screen luminance. The large area point is the ordinary 1-2% Weber fraction.

SCIELAB has been calibrated against human performance, but we wanted to validate our particular experimental set-up described below. Using the curves above, we selected 16 combinations of pixels/in., Weber fraction, background level and polarity, such that six of the combinations were at predicted threshold, five above, and five below. Ten split-field images were prepared for each combination with a .25% spatter of dots on either the top or bottom half and background only on the other half. Four observers were tested in a method of constant stimuli, and the results agreed well with the predictions.

It has been shown that there is a tonescale/pixel-size trade-off for visible quantization artifacts only if a dither or halftoning scheme is used [10-12]. Specific results from these trade-off studies vary, however, because of the different technologies on which the phenomenon was examined, from printers to CRTs to LCDs. The visibility of dither artifacts is also dependent on the image content, and for some kinds of graphics, dither may not be desirable. It would seem that the effects of the use of dither particularly require an empirical approach to assessing the acceptability of a bit depth / resolution / contrast ratio / gamma / gamut combination for a small portable display. On the other hand, we were interested in establishing the usefulness of a computational vision model such as SCIELAB in our display parameter specification process.

4. Bit Depth for a Particular Specification

We chose a particular set of IMOD display specifications (contrast ratio, color gamut, gamma=1, pixel size) for which to determine the target bit depth. Based on previous visualizations and the Weber fraction curves above, we predicted that 12 bit grayscale, 4 bits per channel (444), would show some dither and banding artifacts but still be acceptable to viewers, and that 15 bits (555) would show little dither noise or banding. We also predicted that a mixed bit depth (453) would perform midway between (444) and (555) because it can exploit the lower resolution requirements of blue cones and increase the resolution in the green primary channel, which carries much of the luminance information on the display and therefore has a critical need for sufficient bit depth [13]. We further predicted that the observers would be most critical of artifacts in pictures of people.

Performing the actual tests, both psychophysical and computational, was important because of an essential feature of IMOD displays: at the mirror level, they are one-bit devices, although sub-subpixel binary coding may be used to create higher bit depths when calculated at the pixel level. In the apparatus described below, we are able to simulate pixels via a cartoon, a simplified rendering of the geometry of a pixel where sub-subpixel segments are either on or off, as in an actual display.

4.1 Acceptability Study

Our testing apparatus is a 12’ black tunnel, about 4’x4’, with a large LCD display (in portrait mode) on an optical rail down the center and optics at the open, observer seat, end that create an image of the display that seems smaller and closer. The apparatus is similar to the one described by Silverstein et al. [14], with the exception that it is monocular. The optics and LCD display were positioned such that the image of the display was at 16" viewing distance, with a 1.17" portion of the 1.8" area being shown. The observer is placed using a head- and chinrest, and holds a keyboard that allows him or her to control the rate at which images are presented.

The stimuli were images made from 13 snapshots of faces, 4 snapshots of nature scenes, and 6 graphics images with text and icons. Color processing was applied, and images at (222), (333), (444), (453) and (555) bits per channel were generated, as was a 24-bit color version to be used in the computational study. Stimuli were presented in random order for the acceptability judgments.

Fourteen observers participated in the study, both male and female, varying in age from early twenties to late fifties, and of various backgrounds. They were instructed to give a binary judgment to each image of “acceptable” (1) or “unacceptable” (0), based on small portable display usage. Responses were given orally and recorded by the experimenter.

The judgments of all observers for a given image/bit depth combination were averaged over observers and images for an overall acceptability score for a particular bit depth, shown as the solid black line in Figure 5. Acceptability scores for each type of image across bit depths were also calculated, and are plotted using dotted lines.

The data support a (444) bit depth as the target value for an IMOD display with the given specifications. Interestingly, the (555) value did not surpass the (444), and the (453) images, if anything, were less acceptable. Also, the graphics images were judged as having high acceptability even at the lowest bit depth, despite the fact that they were not created for the specific palette of the IMOD specification, and did show dither noise.

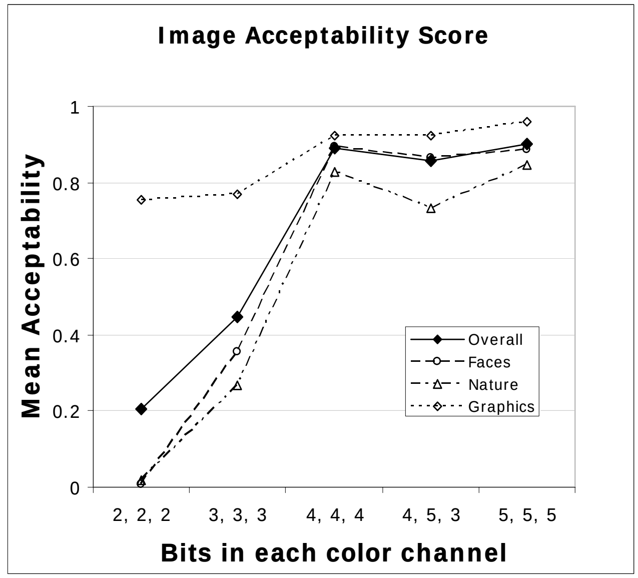

Figure 5. Average image acceptability scores by bit depth, for all images and for images by type.

4.2 Computational Image Quality study

We compared each image as generated for the acceptability study to the 24-bit version generated at the same time and with the same contrast, gamut, gamma, and pixel size specifications, using SCIELAB. The output of SCIELAB is a map of dE*ab values that can be viewed for evaluation of a particular comparison. A summary, the histogram of dE*ab values, can also be examined to advantage. In this case, however, we wanted to be able to combine the dE*ab information into a single number per image, so that we might further combine over images in order to compare across bit depths.

For the 1994 Gille et al. paper [10], the output of what was then known as the ViDEOS/Sarnoff Human Vision Model was a JND map, similar to the dE*ab map of SCIELAB. In that study, thresholds in a dual-random-staircase psychophysical study were compared to threshold contours (JND = 1) derived from the vision model by calculating the RMS JND value for the image comparisons of that study. There, the common measure was JNDs, and the measured thresholds could be graphed together with the calculated contours.

In the present study, we decided that a similar RMS(dE*ab) summary measure would be a reasonable figure-of-merit for each image, where a high number should correspond to low acceptability and vice versa. We calculated the average RMS(dE*ab) for all images and for each type of image, across bit depths. In order to allow a visual comparison of the results of the computational study to those of the psychophysical study, we simply inverted the average RMS values to produce an “image quality” score. The results are shown in Figure 6.

The computational results captured many of the features of the acceptability results, including supporting (444) as the target bit depth for this display, the non-superiority of the (453) bit depth, and even the greater acceptability of the graphics images generally. This last result was somewhat surprising, as we had postulated a purely "psychological" reason for expecting a lower acceptability score for spotty face images.

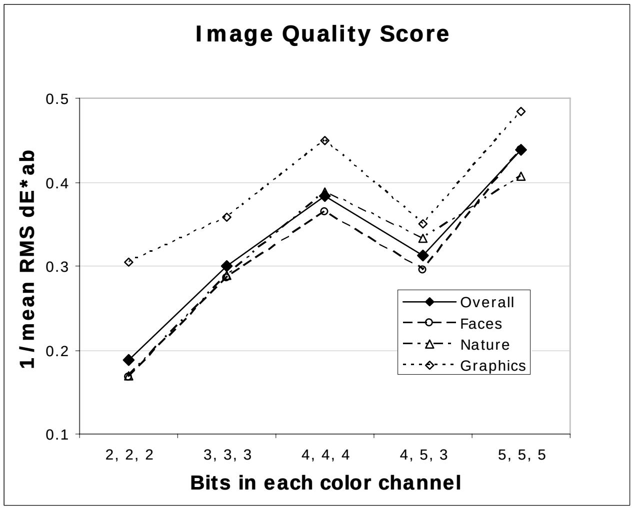

Figure 6. Average image quality scores by bit depth, for all images and for images by type.

5. Discussion

Small portable displays, which typically have low contrast ratios and small gamuts under their ordinary conditions of use, do not require a large bit depth in order to produce good, useable images, especially if the pixel size is small enough to take advantage of dither. Even a linear gamma is not a disadvantage under these conditions, but all the specifications of a display need to be taken into account when determining a target bit depth.

We used data from a detailed simulation acceptability study to establish the target bit depth for a particular reflective interference-color IMOD display. We further used a computational model to make the same judgment, and found good agreement between the two methods.

6. Acknowledgements

Measurement data from commercial displays were provided by Siva Chapalamadugu. Specifications for IMOD primaries used in the display simulations were provided by Kostadin Djordjev.

7. References

[1] http://www.qualcomm.com/technology/imod/media/pdf/IMOD_Technology_Overview_WP.pdf (accessed 3/12/07)

[2] J. Larimer, J. Gille, "Visual performance depends upon signal resolution: Frame rate, dot pitch and bit depth guidelines" Proceedings of the America's Display Engineering and Applications Conference, Vol. **1**, pp. 43-45, 2004.

[3] J. Gille, S. Chapalamadugu, A. Lewis, M. Kirby, "Numbers of colors for small portable displays" Paper presented at the Third Americas Display Engineering and Applications Conference, October 24-26, 2006, Atlanta, GA, 2006.

[4] M. Mahy, L. Van Eycken, A. Oosterlinck, "Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV" Color Research and Application, Vol. **19**(2), pp. 105-121, 1994.

[5] M. Aldaba, J. Linhares, P. Pinto, S. Nascimento, K. Amano, D. Foster, "Visual sensitivity to color errors in images of natural scenes" Visual Neuroscience, Vol. **23**, pp. 555-559, 2006.

[6] J. Larimer, J. Gille, J. Luszcz, "Display resolution trade-offs: Grayscale spacing" Conference Record of the 1997 International Display Research Conference, p. 262, 1997.

[7] http://white.stanford.edu/~brian/scielab/scielab.html (accessed 3/1/07)

[8] X. Zhang, B. Wandell, "A spatial extension of CIELAB for digital color image reproduction" Society for Information Display Symposium Technical Digest, Vol. **27**, pp. 731-734, 1996

[9] X. Zhang, D. Silverstein, J. Farrell, B. Wandell, "Color image quality metric S-CIELAB and its application on halftone texture visibility" Proceedings of COMPCON 97, p. 44, 1997.

[10] J. Gille, R. Martin, J. Larimer, "Spatial resolution, grayscale, and error diffusion trade-offs:Impact on display system design" Conference Record of the 1994 International Display Research Conference, pp. 381-385, 1994.

[11] J. Farrell, "Grayscale/resolution tradeoffs in image quality" Proceedings of the Conference on Raster Imaging and Digital Typography, Boston, 1991.

[12] M. Rodriduez, "Graphic arts perspective on digital halftoning" Human Vision, Visual Processing, and Digital Display V, B. E. Rogowitz and J. P. Allebach, Editors, Proc. SPIE 2179, pp. 47-59, 1994.

[13] R. Martin, J. Gille, J. Larimer, "Detectability of reduced blue pixel count in projection displays" Proceedings of the Society for Information Display, Vol. **24**, pp. 606-609, 1993.

[14] L. Silverstein, J. H. Krantz, F. E. Gomer, Y. -Y. Yeh, and R. W. Monty, "Effects of spatial sampling and luminance quantization on the image quality of color matrix displays," J. Opt. Soc. Am. A **7**, pp. 1955-68 , 1990.