

Contrast Modulation—How Much is Enough?

Jon Leachtenauer
J/M Leachtenauer, Associates Inc., Annandale, Virginia

Andrew S. Biache and Geoff Garney
Autometric Inc., Springfield, Virginia

Abstract

The effective resolution of a CRT is defined by two metrics, addressability or pixel density, and contrast modulation. Pixel density indicates the size of the smallest potentially resolvable detail. CRTs with pixel densities of 170 pixels per inch (ppi) are becoming increasingly common and are theoretically capable of showing detail on the order of 0.005 inches on the display. Unless the pixel can be distinguished from surrounding pixels, however, the detail will not be seen. The ability to distinguish adjacent pixels is typically expressed in terms of contrast modulation.

Current monitors typically show Cm values of 0.2-0.6 with highest values in the center of the screen and lower values toward the edges. A few monitors show even higher performance. To assess the effects of differences in Cm performance, two monitors were evaluated. One had Cm values of 0.5 to 0.87, the other had values of 0.2-0.4. Both had pixel densities of 100 ppi and were calibrated to a range of 0.1 to 35 fL. Experienced imagery analysts performed the evaluation. They provided National Imagery Interpretability ratings of visible and radar imagery as well as Briggs ratings. Briggs ratings assess the ability to discriminate checkerboard patterns of varying sizes and contrast. Results were analyzed to determine if performance was affected by differences in Cm.

Introduction

Contrast modulation is defined as the ratio of a luminance difference divided by the sum of the luminance values:

$$C_m = (L_B - L_D)/(L_B + L_D) \tag{1}$$

where C_m = contrast modulation, L_B = bright luminance, and L_D = dark luminance. Contrast modulation is measured using an aperture grill of one, two, or three pixels wide commanded to full on and full off values (0 and 255 for an 8-bit display). A C_m value of 0.25 or 25% is considered adequate for imagery, 50% C_m is required for text.[1] Current monitors typically show C_m values of 20-40% at screen center and lower values toward the corners of the display. C_m values for color monitors are generally less than for monochrome monitors due to aperture grill or shadow mask effects, as well as color misconvergence.

The tendency in CRT design over the past several years has been to improve both addressability and C_m. As was shown in an earlier study[2] , increased addressability may not be usable because of human perceptual limitations. The goal of the current study was to assess the effect of improved C_m.

Background

The ability to resolve small detail on a display is a function of both size and contrast. As size decreases, required contrast must increase. The so-called "J" curve defines the threshold.[3] The curve is shown as Figure 1 with the size dimension expressed as addressability (pixels per inch) at an 18 inch viewing distance. A combination of C_m and addressability falling below the line will generally not be resolved. C_m is best at ~17 ppi is an order of magnitude worse at 100 ppi.

From Figure 1, it is apparent that the ability to detect contrast can be improved with magnification. This can be demonstrated by comparing Cm measurements with different aperture grill sizes (one to three pixels). The effect is shown in Figure 2 for five different aperture grill color monitors. The effect of magnification is not fully predictable because of alignment, convergence, and halation effects.

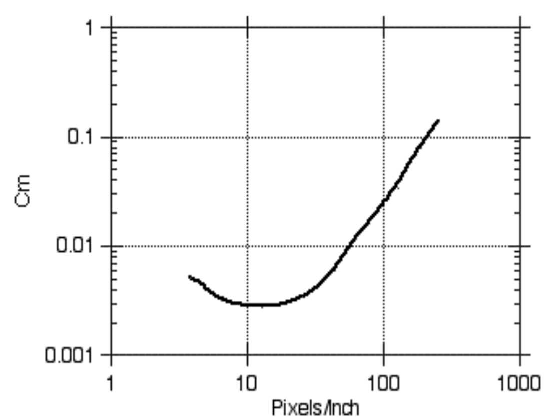

Figure 1. "J" Curve

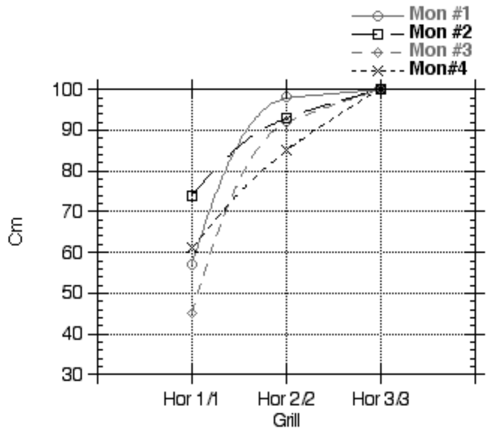

Figure 2. Effect of grill size on horizontal C_m

Method

Imagery was displayed on two monochrome monitors with different levels of C_m. Experienced IAs provided Briggs ratings for both monitors at both 1x and 2x magnification. NIIRS difference ratings (delta-NIIRS) ratings were made on images displayed on one monitor relative to the same images displayed on the other. Subjective quality ratings were also made for each magnification level.

Monitors

Two monochrome monitors were used for the current study. Both had pixel densities of 100 ppi and were calibrated to a dynamic range of 0.1 to 35 fL. The NEMA/DICOM[4] perceptual linearization look up table (LUT) was applied to both displays. Table 1 summarizes monitor characteristics. A comparison of modulation values is shown in Figure 3.

Table 1. Monitor Characteristics

Measure	Monitor 1	Monitor 2
Lmax	35fL	35fL
Lmin	0.1fL	0.1fL
L non-unif. (max)	29%	25%
Vertical Refresh	74Hz	60Hz
C_m Center (V/H)	40/34%	76/58%
C_m Minimum (V/H)	22/28%	53/58%
Pixel Density	100ppi	100ppi

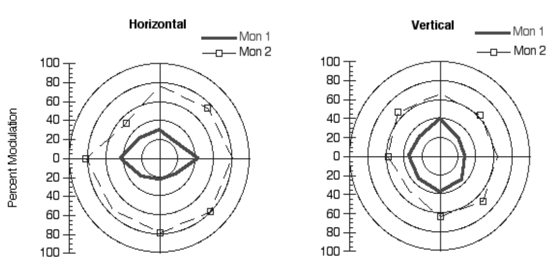

Figure 3. Comparison of horizontal and vertical modulation values as a function of screen position.

Imagery

Ten visible and ten radar images were used in the study. They were chipped to 600 pixels square and displayed at both 1x and 2x magnification. Bi-linear interpolation was used for magnification. The images were centered on the display.

Briggs targets[5] (checkerboards of varying size, contrast and luminance) were also used. Figure 4 shows a target set. Each set consists of eight groups of checkerboards equally spaced across the command level range. Each group contains 17 checkerboard patterns varying in size. Three target sets with differences between dark and light squares of 1 (C-1), 3 (C-3), and 7(C-7) command levels were used at both 1x and 2x magnification. Pixel replication was used for magnification. The images were centered on the display. In addition, the C-7 target was also shown in the upper right and lower left of the display.

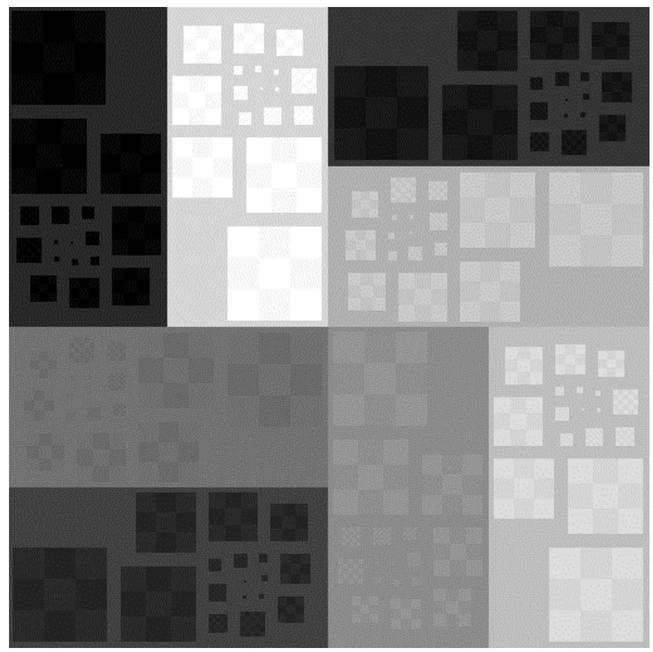

Figure 4. Briggs C-7 Target.

In this case, each of the eight checkerboard sets making up a full Briggs target was displayed separately so as to remain in the corner of the display.

Rating Scales

Briggs ratings require the observer to define the smallest checkerboard (in each set of 17) where the squares can be separately distinguished. The smallest target receives a score of 91, the largest a score of 6. The smallest board is then rated in terms of square definition with ratings ranging from 1 (well-defined squares) to 5 ("blobs"). These ratings are subtracted from the size scores so that the range of scores is 1 to 90. If the largest pattern in each set can not be resolved, a score of 0 is given.

NIIRS ratings are made on a 0 to 9 scale. Each level is defined by six criteria or interpretation tasks; e.g., identify individual rail cars by type.[6] The tasks are of increasing difficulty from 0 to 9 and thus require increasingly better image quality. Separate NIIRS exist for each image type (visible, IR, radar).

The subjective quality ratings asked the IAs to rate relative (Monitor 2 to Monitor 1) display contrast, detail, noise, and sharpness in terms of the impact of the difference on image interpretability

Evaluation Procedure

Eight NIIRS-certified imagery analysts (IAs) participated in the study. The displays were placed in a controlled lighting environment with an ambient light level of ~ 2fc. Each analyst began with a set of Briggs ratings at 1x magnification followed by NIIRS ratings.

NIIRS ratings were made in terms of Monitor 2 relative to monitor 1. These are termed delta-NIIRS ratings. The same procedure was followed at 2x magnification. The order of monitor presentation was counterbalanced for the Briggs ratings. The subjective quality ratings were made at the completion of Briggs and NIIRS ratings for each magnification level.

Data Analysis

Analysis began with outlier analysis. Analyses of variance were performed on all of the rating data. Variables for the Briggs data included magnification, monitor, target contrast, and target brightness. For the Briggs C-7 data, target location was also a variable. For the delta-NIIRS ratings, variables included magnification, monitor and image type. For the quality rating data, variables were monitor, quality attribute, and magnification.

Results

Data from one IA were dropped after outlier analysis. Rater/group correlations for the Briggs data ranged from 0.74 to 0.91. The average rating standard deviation for the Briggs C-7 target was 9.1 (scale of 0 to 90). Only three of the IAs reported NIIRS differences.

Briggs Rating Data

Briggs scores for targets in the center of the displays are summarized in Figure 5. The overall monitor difference was not statistically significant at either magnification level. There was a statistically significant interaction with monitor and Briggs target. Monitor 2 showed significantly higher scores with the C-3 target. There was no significant difference with the C-1 or C-7 target.

The differences between the monitor C-1 scores were largely the result of differences with the sixth brightest target. At 1x magnification, the score for Monitor 1 was 17 points higher than for Monitor 2. There was no apparent reason for this difference, although it is possible that it was due to a quantization loss. The perceptual linearization LUT was applied separately to each monitor based on the monitor's command level/luminance function. For the C-3 target, the Monitor 2 scores were higher than the Monitor 1 scores for all but the sixth brightest target. Again, no obvious explanation was apparent

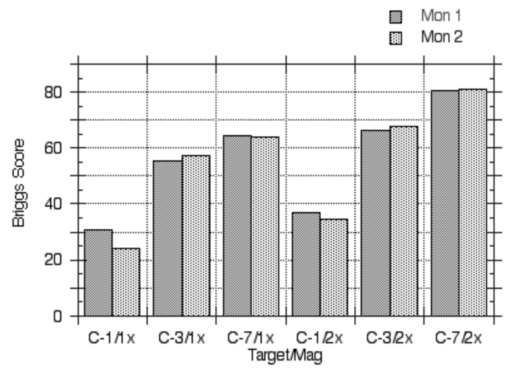

Figure 5. Briggs rating data.

The effect of screen position is shown in Figure 6. Overall, Monitor 2 scores were higher than Monitor 1 by a statistically significant degree. The difference was due to the performance in the corners of the display and was largely due to the results of a single IA. Monitor 1 had Cm values of ~ 20%, whereas Monitor 2 had values of 50% or greater.

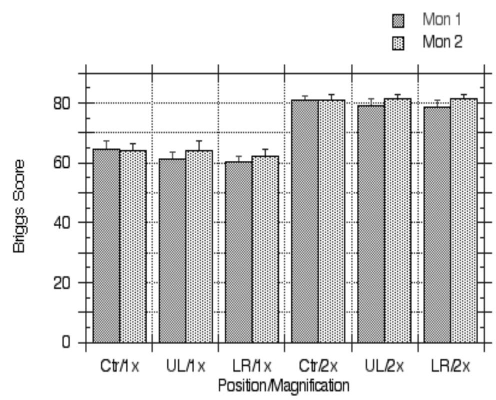

Figure 6. Effect of screen position (center, upper left, lower right).

NIIRS Ratings

As noted earlier, only three of the IAs reported NIIRS differences between the displays. Averaged across seven IAs, the differences ranged from 0.016 to 0.023 in favor of Monitor 1. Although the differences were statistically significant, none were practically significant

Quality Ratings

Subjective quality ratings provided at the completion of each magnification level are shown in Figure 7. Ratings are made on Monitor 2 relative to Monitor 1. Bars marked with an asterisk show statistically significant differences from 0.

Only contrast at 1x and sharpness at 2x show statistically significant differences. Ratings are on a 10-point scale, where a 5 denotes moderate impact on interpretability and a 10 denotes significant impact. The strongest rating observed was a –2.

Discussion and Conclusions

Briggs ratings showed no overall differences between the two monitors. Ratings were higher for Monitor 2 on the C-3 target. Based on the C-7 target, ratings were significantly higher for Monitor 2 in the corners of the display. Monitor 1 shows Cm values of ~ 20% in the corners, Monitor 2 shows values of > 50%. Briggs scores vary with both the size of the squares and the number of squares in a checkerboard. Results of the current study indicate that for the C-1 target, IAs resolved 4 to 8 pixel squares at 1x magnification and 2-4 pixel squares at 2x.

Based on the "J" curve, they should not have been able to resolve 1 pixel squares even at 2x with no modulation loss. For the C-3 target, they could resolve 2-4 pixel squares at 1x and 1-2 pixels at 2x. They should not have been able to resolve one pixel squares at 1x with no modulation loss. They should have been able to resolve one pixel squares at 2x and they did. There clearly were modulation losses and they affected some, but not all of, the Briggs scores. The losses were not sufficient to significantly affect all of the Briggs scores or to distinguish between the two monitors. Results appeared consistent with the "J" curve thresholds.

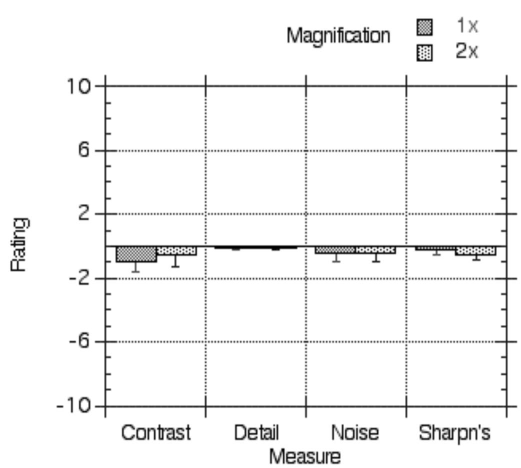

Figure 7. Quality Ratings

Only three of the IAs reported NIIRS differences. Although the difference between the two monitors (in favor of Monitor 1) was statistically significant, the difference was small and mostly the result of ratings from a single IA.

The subjective quality comments favored Monitor 1 in terms of contrast at 1x magnification and in terms of sharpness at 2x magnification. In both cases, the mean was less than one on a ten-point scale.

In their reports on monitor measurements, NIDL [1] uses a Cm threshold of 25% for imagery and 50% for text. Both monitors exceeded the 25% criterion over most of the display, Only Monitor 1 fell below the 25% criterion in the corners of the display. This may explain the C-7 results favoring Monitor 2 based on the corner scores.

Some of the analysts reported a color between the two monitors. Monitor 1 was judged more "black and white" than Monitor 2. Color coordinate measurements were made on both monitors, results indicated that Monitor 2 was "warmer" than Monitor 1.

It appears that some of the IAs may have responded to the color coordinate difference in making their NIIRS ratings and quality attribute ratings. Luminance measurements showed no differences between the two monitors. There is thus no reason to believe that contrast differences existed. It is considered unlikely that the color coordinate difference would have produced a perceptual contrast difference. This conclusion is supported in part by the Briggs ratings. However, Briggs ratings are also affected by contrast modulation and Monitor 2 showed higher Cm values than Monitor 1 (and thus should have been preferred based on physical quality).

The lack of an overall Briggs difference, coupled with the fact that most IAs saw no NIIRS differences between the two monitors, suggests that the NIIRS difference was spurious. The side-by-side comparisons may have tended to force rating differences. Similarly, there is no logical reason to believe that the Monitor 1 was sharper at 2x magnification than Monitor 2.

It is therefore concluded that, in the current study, there was no practical difference between the two monitors, except possibly in the corners of the display. It appears that once some threshold value of Cm is achieved, improving Cm beyond that value offers little benefit. For the tasks evaluated in the current study, the threshold value of Cm appears to be on the order of 25%. Only when values fell below 25% was there a discernable difference. It is possible that other tasks might show higher thresholds and thus larger differences between the two monitors. With aerial imagery, however, one seldom encounters single pixel objects of interest with very small contrast changes. One and two pixel autocorrelation functions tend to be high. Thus, the ability to detect single pixel, low contrast changes is seldom an issue and may help explain results of the current study.

Acknowledgements

The study reported here was funded and carried out at the National Imagery and Mapping Agency. The authors would like to thank Mr. Michael Parcell for his support, as well as the imagery analysts who participated in this study.

References

1. National Information Display Laboratory, *Display Evaluation report, A Survey of Five 2-MegaPixel*

Monochrome CRT Monitors for Medical and Government Imaging Applications, Publication No. 021797-058, Princeton NJ, August 1997
2. Leachtenauer, J., Biache, A., and Garney, G. “Effects of Pixel Density on Softcopy Image Interpretability” *PICS Conference, Final Program and Proceedings*, Society for Imaging Science and Technology, Savannah GA, April 25-28, 1999, pp. 184-188.
3. Farrell, R. and Booth, J. *Design Handbook for Imagery Interpretation Equipment,* Boeing Company, Seattle WA, 1984
4. National Electrical Manufacturers Association (NEMA), *Digital Imaging and Communications in Medicine (DICOM), Part 14 “Grayscale Display Function Standard, Supplement 28,* 28 Jan 1998
5. Briggs, S.J. *Manual Digital Test Target BTP#4*, D180-25066-1, Boeing Aerospace Co., Seattle WA, January 1979
6. Leachtenauer, J. “National Imagery Interpretability Rating Scales: Overview and Product Description”, *ASPRS/ACSM Annual Convention and Exhibition, Technical Papers, Volume 1, Remote Sensing and Photogrammetry,* American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, Baltimore MD, April 1996, pp. 272-281.

Biography

Jon Leachtenauer received his BA and MS degrees from Syracuse University. For the past 40 years, he has been engaged in human factors research on image and display quality. He recently established a consulting firm (J/M Leachtenauer Associates Inc.) and is a consultant within the Office of Technology, National Imagery and Mapping Agency. He consults in the areas of evaluation design, image and display quality and utility, and scales development.