Are Medical Image Display Systems Perceptually Optimal? Measurements before and after Perceptual Linearization

Bradley M. Hemminger*, Hartwig R. Blume#

*Department of Radiology, University of North Carolina,
Chapel Hill NC 27599-7510, bmh@rad.unc.edu
#Philips Medical Systems, Shelton CT 06484

This paper was presented as a poster at the SPIE Medical Imaging Conference in February 1996. The paper appears in the proceedings of that conference, and the reference is

Hemminger BM, Blume HR, "Are Medical Image Display Systems Perceptually Optimal? Measurements before and after Perceptual Linearization", Proceedings of Medical Imaging 1996: Image Capture, Formatting and Display, SPIE Vol 2707-58, 1996.

Note: in some copies I have included the figures separately, one per page in the appendix, because of the low quality of the small figures when reproduced by photocopying.

For more information contact:

Brad Hemminger
Senior Research Associate
513 Old Infirmary Building
Department of Radiology
University of North Carolina
Chapel Hill, NC 27599-7510

bmh@rad.unc.edu
(919) 966-2998
(919) 966-5934 (fax)

Are Medical Image Display Systems Perceptually Optimal? Measurements before and after Perceptual Linearization

Bradley M. Hemminger*, Hartwig R. Blume#

*Department of Radiology, University of North Carolina,
Chapel Hill NC 27599-7510, bmh@rad.unc.edu
#Philips Medical Systems, Shelton CT 06484

ABSTRACT

Perceptual linearization has been advocated for medical image presentation, both for the faithful reproduction of images, and for standardizing the appearance across different display devices. It is currently being proposed as the standard display function for medical image presentation by ACR/NEMA working group 11 (display function standard). At this time, studies have not been made to evaluate how close existing display systems are to being perceptually linearized. This paper presents a methodology for quantitatively calculating the perceptual linearity of a display device based on a statistical measure, the *linearization uniformity measure* (LUM), of standard deviation of the ratio of contrast thresholds of the display system versus the contrast thresholds of the human observer. Currently available medical image display systems are analyzed using LUM metric, and their pre-linearization and post-linearization results are compared with that of the desired human observer response curve. We also provide a better description of the achievable dynamic range of a display device, based on the three quantitative measures: the standard deviation of the contrast threshold ratios, the mean of the contrast threshold ratios, and the number of DDL steps used.

1. INTRODUCTION

The perceptual linearization of video display monitors plays a significant role in medical image presentation[1,11,12]. First, it allows the maximum transfer of information to the human observer since each change in digital driving level of the display yields a perceptually equal step in perceived brightness by the human observer. Second, for an image to be perceived as similarly as possible when seen on different displays, the two displays must be standardized, which can be done when they have been perceptually linearized. Third, perceptual linearization allows us to calculate the perceived dynamic range of the display device, which allows comparing the maximum inherent contrast resolution of different devices.

Perceptual linearization was first suggested for medical image presentation by Pizer[1], and in follow-up work[2,3,4,5,6,7,8] at the University of North Carolina at Chapel Hill (UNC). To best visually present an image represented as digital data to the human observer, we would like to maximize the information transferred in mapping the digital driving levels to perceived brightness levels. Perceptually linearizing the mapping from the image data space to the human observer's visual sensory space most faithfully transmits changes in intensities in the image to the human observer.[3,6,9,10] This simply means that to the human observer, equal absolute changes in the input values to the display system should result in equal absolute changes in the perceived visual sensation.

The process of displaying an image on a video display monitor to the human observer is depicted in figure 1. This paradigm applies equally well to the display of images on film. Initially, an object, such as the human body, is scanned and the resulting signal (for instance tissue density) is represented on the computer as a matrix of points, called pixels. This scanning samples the original source data (continuous analog function) into discrete data (set of digital values). Each pixel is represented by a scalar value, usually in the range of 0 to 4096 for medical image data. These are the values referred to as *Recorded Intensities* in Figure 1. The second step is that some set of image processing operations, such as intensity windowing, or contrast enhancement may be performed on the *Recorded Intensities* resulting in the *Displayable Intensities*. These values are then scaled into *Digital Driving Levels* (DDLs), which must be in the range accepted by the *Digital to Analog Converter* (DAC) of the display system. This scaling is done by a table lookup operation, often referred to as a *Lookup Table* (LUT) or colormap table. LUTs are often used to do intensity windowing dynamically, or to implement a linearization LUT (these are sometimes called *gamma correction* curves). The output of the LUT goes to the DAC, which takes the input DDL and converts it to an analog voltage level which is used to drive the monitor at different luminance levels. The luminance generated by the monitor is then recorded and processed by the eye-brain human visual system, resulting in the sensation of brightness by the human observer.

From the standpoint of linearization there are two important relationships in this process, that of the DDLs of the computer's DAC versus the luminance of the monitor, and that of the monitor luminance versus the brightness perceived by the human observer. The first relationship of DAC to luminance will be referred to as *DACLUM*. The second relationship, that of luminance to perceived brightness, is best examined using a *luminance Contrast Sensitivity Function*, abbreviated as CSF in this paper. CSFs measure the change in luminance (ΔL) required for a target, so that it may be detected from the surround luminance (L) as a function of the surround luminance. More specifically, *Contrast Thresholds* (CT) are defined as $\Delta L/L$, while CSFs are defined as its reciprocal, i.e. $L/\Delta L$. CSFs in this paper will refer to $L/\Delta L$ versus L, while in vision literature, CSFs usually refer to $L/\Delta L$ versus spatial frequency of the target.

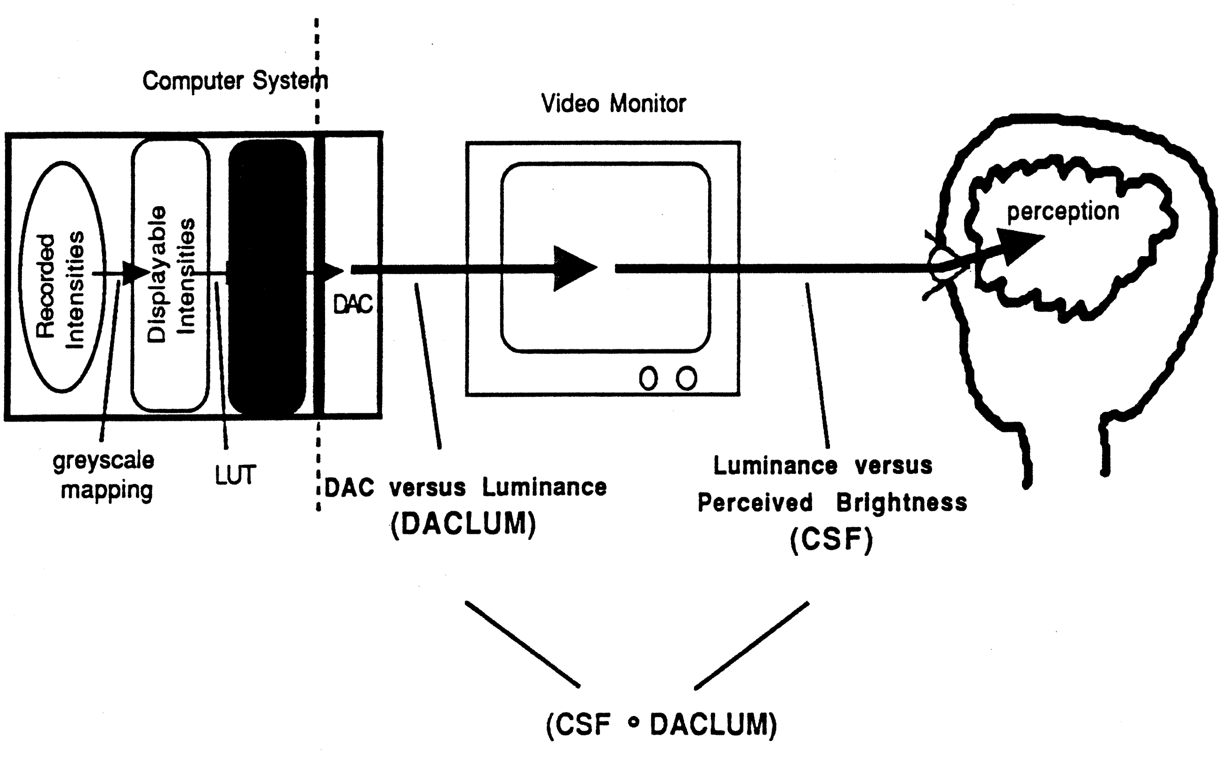


Figure 1. Diagram of components of Perceptual Linearization

If we think of the DACLUM and CSF curves as functions, and compose them on their common variable of luminance, we arrive at a CSF•DACLUM function that defines the overall effect of the DACs, monitors, and human perception in the display system. This is usually referred to as the *characteristic curve* of the display system. The inverse of this function can be determined and used to remap the image values to perceptually linearize the relationship between the grey levels of the image in the computer and the sensation of brightness to the human observer.

Pizer, in his initial description of perceptual linearization gave both an intuitive and a formal analytical approach[1]. In the intuitive approach one calculates

$$L_i = L_{i-1} + (L_{i-1}) * (1 / (CSF(L_{i-1}))$$

until L_i reaches or exceeds the luminance of the maximum DDL. L_i represents the luminance at the ith DDL value, and $(1 / (CSF(L_i))$ the contrast threshold at luminance L_i. Thus in this formulation one simply steps 1 JND in luminance at each step, starting at the minimum luminance, until the maximum luminance is reached. The analytical formulation is given by Pizer[1] and the specifics of implementing the linearization by Cromartie[8]. Also an approximation that further simplifies the analytical solution is given by Ji[35]. In work on linearized color scales other authors have developed methods that supersample in the perceptual scale, and then choose the closest digital driving scale of the monitor[36,37].

Initial work at UNC used experimentally measured CSFs based on simple detection tasks on video monitors. Recently, several investigators have proposed using vision models as more general CSF predictors[9,15,16,24,25,26,35]. Specific parameters to the Barten visual model that match medical image presentations have been proposed as a standard in the literature[12,42] and is currently proposed by ACR/NEMA working group 11 (Display Function Standard) as the method for standardizing and linearizing medical image displays. The Barten model with these parameters is the visual model used throughout this paper.

2. BACKGROUND

2.1 Linearization Calculations

In all of the above linearization approaches, the final step takes a calculated desired luminance level and then selects the DDL that produces the luminance nearest in value to the desired luminance. Because there are limited discrete samples in the DDL range (256), and since they are often not distributed in a fashion matching the CSF function, errors may be introduced during this matching step.

Another important issue is the number of DDLs in the resulting map. Linearization methods attempt to create mappings with each DDL step being an equal fraction of a JND step apart, often with the default implementation creating a table of steps being one JND apart. If the number of distinguishable grey levels, or *perceived dynamic range* (PDR) range of the monitor is significantly less than the number of available DDLs on the DAC then we face the issue of whether to use more DDLs. Choosing not to do so means the contrast resolution must be downsampled to the smaller value of the PDR rather than using the number of DDLs available. For instance, on a display system with a PDR of 80, we would have to downsample the input greyscale range of 4096 levels to just 80 levels. This coarse quantization of the input data may be undesirable. On the other hand, if we resample the linearization curve at a higher resolution, we introduce more matching error, because it becomes more difficult to accurately match desired luminances with the discrete set of luminances available on the display system.

Standardization of video displays, as well as hardcopy displays has assumed an increasingly important role. Currently, ACR/NEMA working group 11 (Display Function Standard) is addressing standardization by perceptual linearization. In order for this to work effectively, we must be able to calculate the accurate linearizations, and have quantitative measures of conformance of a given perceptual linearization. Hemminger[43] has proposed a methodology that allows quantitative comparisons of linearizations using the LUM statistic, and that describes the actual resulting PDR, or *Achievable PDR*. We will use this measure to analyze existing display systems.

II.2 Display Systems

In the past two years, there have been significant advances in display systems. The first has been the advent of commercially available high spatial resolution, high brightness (luminance) monitors, approaching 200 ftL (685.20 cd/m^2). The second has been the increase in DAC resolution from 8 bits (256 levels) to 10 bits (1024 levels). Actually the available higher resolution DACs are only 10 bit DACs on the output side, i.e. they can produce only 256 levels at a time, but from a palette of 1024 possible choices.

The high brightness monitors introduce several effects. First, they increase the contrast range of the display systems. For example, the dynamic range of the four high brightness monitors, at the operating conditions for which they were provided to us by the manufacturers, were measured to be 150:1, 256:1, 274:1, and 538:1. Second, by having more luminance values available at the higher luminances, there are proportionally fewer luminances in the display distribution from the lower end. This second factor is advantageous because it is primarily at the lower luminance values that the display systems characteristic curves are most perceptually non-linear. As most of the luminance range is in the photopic range where the Weber-Fechner Law holds, the video display systems are more linear, and more similar to the film displays. The trade-off of the increasing luminance range, is that the contrast threshold between adjacent DDL steps will increase, thus requiring more DDLs to be available in the display system to provide similar contrast resolution.

The other area of change is the introduction of the 8/10 bit DACs (2^8 levels at any one time out of a palette of 2^{10} possible levels). These are significant in that they allow much better linearizations. Previously, with 8/8 DACs a calculated linearization that used most all of the available 256 levels would either be nearly identical to the default distribution, or it would have to repeat DDL values in order to omit other values, in order to change the distribution. In neither case would the linearization be significantly improved compared to the default characteristic curve. Choosing to use significantly fewer than the 256 levels available, however, causes a corresponding reduction in contrast resolution. On the other hand, an 8/10 bit DAC supporting the choice of 256 out of 1024 levels, makes it much more likely that one can obtain an accurate linearization using all 8bits (256 levels). Thus, the display system can be better linearized with an 8/10 DAC, without the corresponding tradeoff in contrast resolution (caused by a decrease in the number of DDLs used) that occurs with 8/8 bit DACs.

III. METHODS

III.1 Theory

Perceptual linearization is based on maintaining equality between digital driving level steps and perceived brightness steps. Equal changes in digital driving levels should confer equal changes in perceived brightness. The measure of change is usually defined as a contrast threshold, ΔL, where

$$\Delta L = \text{(change in luminance over an interval)/(mean luminance of interval)}$$

Thus, a measure of the perceptual linearity of a display system would be how equal are the contrast thresholds resulting from adjacent DDLs compared to the contrast threshold predicted by vision models. The contrast threshold of the monitor is easily calculated for the interval $[DDL_i, DDL_{i+1}]$ by

$$CT_{display} = (L_{i+1} - L_i)/((L_i)$$

assuming the background to be the L_i value, and the contrast difference to be $(L_{i+1} - L_i)$. We base the contrast threshold for perceived brightness on human visual models, specifically the Barten model with the parameters proposed by Blume[12] and Hemminger[42]. From these models we can calculate the predicted contrast threshold value, CT_{human}, at a specific luminance value. Computing contrast threshold values of the monitor at each interval, as described above, allows comparison with the visual model predicted contrast threshold at the mean luminance of each interval.

Perceptual linearization dictates maintaining equal step sizes across the range of the monitor. A given display system's contrast threshold, as calculated from its step sizes, may not be equal to those predicted by the human visual model. For instance, the step size of the display system may be four times the human observer threshold because there are too few DDLs. Or a very high resolution DAC may have step sizes that are subthreshold, that is, smaller than the human observer contrast threshold. In either case, the important consideration for perceptual linearization is that the ratio between display system contrast threshold and the human observer contrast threshold be the same across the range of the display system. Thus, we want to calculate

$$\text{ratio} = CT_{\text{display}} / CT_{\text{human}}$$

and measure how consistent it is across the range of the display system. Considering these ratio values to be members of a statistical population suggests using the statistical measure of standard deviation to quantitatively define how similar the individual members are to their common mean. Since we have the complete population defined, the standard deviation is simply the square root of the sum of the squares of the distance from each sample point to the mean of the population. Other measures might be considered, for instance, the variance (the square of the standard deviation) which would penalize outliers in a distribution more, and the mean population deviation (simple average of the distances to the mean). We have chosen to work with the standard deviation of the population as our measure. This results in a single value that provides a quantitative measure of the perceptual error of a display system.

While standard deviation provides a good handle on the error of the linearization, two other factors need to be defined to completely describe the linearization remapping. First, because the remapping may only use some of the digital driving levels available in order to produce a different transfer function curve on the display system, the resulting linearization often has fewer DDLs than the default system configuration which uses all DDLs. As a result, in general, the linearized remapping has larger contrast threshold steps because it must cover the same luminance range, but with fewer DDL steps. Because there are adverse effects due to "contouring" when too few DDLs are used to cover a luminance range[34], it is important to quantify this effect by calculating the mean $CT_{\text{display}}/CT_{\text{human}}$ ratio. Finally, the number of DDL steps actually used is an important factor as the number of DDLs is often limited, both at the DAC level, and on some workstations by the window manager application. Thus, a display can be characterized by three values, the standard deviation of the ratio of the display contrast threshold step sizes versus the ideal human observer response, the mean of these ratios, and the total number of DDLs used in the linearization.

III.2 Procedure for Measuring Conformance of Linearization on Display System

Once the linearization has been computed, it is necessary to measure the faithfulness of the perceptual linearization on the display system. This is because errors may occur in the resulting linearization due to display system factors, or to suboptimal linearization calculation techniques. Measuring the resulting luminance values of the display system after perceptual linearization provides a quantitative measure of the quality of the linearization of that display system, and provides the guarantee a specific level of conformance to a standard. Additionally, as part of the conformance measurements, this section describes the calculation of two values, the theoretical realizable dynamic range of the display system, and the achieved dynamic range of the display system. Below is a description of the conformance protocol currently under consideration by ACR/NEMA WG 11. The same protocol can be used for both video and film display systems. In the case of video display, the mappings are achieved by changes to the digital values used to drive the DAC that produces luminances for the monitor. For film display, the mappings are achieved by changes to the digital values used to drive the DAC that drives the laser which exposes the film.

1) Measure the characteristic curve (DDLs versus Luminance) at all DDLs of the display system, using test pattern defined by ACR/NEMA working group 11 (center square in middle of screen, occupying 10% of screen area, driven at test DDL, while remainder of screen is driven at constant DDL = 20% of peak luminance of display system). Note, that for a 8/10 DAC system this means 1024 levels.

2) Calculate the perceptual linearization function.

3) Cause the display system to use the calculated perceptually linear mapping. This will generally be done via the manufacturer's application. However, the end user can easily accomplish this by calculating the linearization LUT and replacing the *Default* display system LUT with the *Linearized* LUT (either via the hardware LUT or via a software LUT depending on display system capabilities).

4) Use formula (1) below, with the luminance versus DDL measurements recorded in step (1), to calculate the ratio of the Contrast Threshold for the linearized display system to that of the desired Contrast Threshold as calculated from the Barten model, for each measured luminance level. Then calculate the standard deviation of the resulting sample population of these ratios. The standard deviation should be small, indicating a uniform matching of luminance step sizes between the display system and those calculated for the human visual system by the Barten model. The values assumed by the ratio of CTs is in the range [0.0,infinity], and centered around 1.0 where the display system ratio is equal to the human visual model predicted ratio. Because the ranges between the two equal sides are not equal, we remap [0.0 , 1.0] to [-infinity, 0.0] and [1.0 to +infinity] to [0.0, +infinity]. This gives us a uniform distribution centered around 0.0, and allows us to calculates the mean and standard deviation. Thus, negative values will now match those ratios where the display ratio was less than the human visual model predicted ratio, and positive ratios those where the human visual model was greater than the human visual model predicted ratio. One last refinement is made for the case where the ratio is equal to zero. This implies that the display CT was 0.0, i.e. there was no luminance increment between adjacent DDLs. In general there is a luminance increment, only it was too small for the measurement technique used, or re-sampling method utilized. Thus, instead of mapping a value of 0.0 to -infinity, which is not desirable, we instead choose to map it to a value corresponding to a CT 10% less than the smallest CT in the dataset (assuming this to be a reasonable lower bound for how small the increment really was). Specifically, for each interval between adjacent DDLs, *i* and *j*, in a display distribution (linearized or not), the Display Distribution versus the Barten model predicted ratio is defined as

Display/Barten ratio = (Display System Contrast Threshold) / (Barten Method Calculated Contrast Threshold) (1)

Display/Barten ratio = ((Luminance[j] - Luminance[i]) / Luminance[i]) / Barten_CT(Luminance[i])

remapping the ratio distribution to be equally distributed around 0.0

(Display/Barten ratio = 0.0) => Modified ratio = *Minimum replacement ratio*
(Display/Barten ratio < 1.0) => Modified ratio = 1.0 - (1.0/(Display/Barten ratio))
(Display/Barten ratio >= 1.0) => Modified ratio = (Display/Barten ratio) - 1.0

where *Minimum replacement ratio* is chosen so that ratio values of zero are mapped to less than the smallest positive ratio value. For our calculations we have chosen Minimum replacement ratio to be 10% less than the smallest ratio value greater than zero.

Linearization Uniformity Measure (*LUM*) = STDDEV {Modified ratios}

ACR/NEMA working group 11 is currently evaluating as a possible conformance measure that a display system is said to conform to the display function standard if the Linearization Uniformity Measure is less than a specified constant, which is currently being determined.

Definition: Display System is Linearized if *LUM* $\leq$ LUM_DEFINED_CONSTANT

III.3 Further Measures of Display System

The quality of the standardization of a display system is determined by the single quantitative measurement of the standard deviation of the collection of Display-Barten ratios, the Linearization Uniformity Measure. An important related factor is the number of discrete output levels of the display system. For instance, the Linearization Uniformity Measure could be improved by using significantly fewer output levels, at the cost of decreasing contrast resolution. While the Linearization Uniformity Measure is influenced by the choice of the number of discrete output levels in the linearized output LUT, the appropriate number of output levels is determined by the clinical application, including possible grey scale image processing that may occur independently of the perceptual linearization. Thus, this perceptual linearization does not proscribe a certain number of grey levels of output. However, in general, the larger number of grey levels available, the higher the possible image quality because the contrast resolution is increased. It is recommended that the number of necessary output driving levels for the Linearized distribution is determined prior to linearization of the display system (based on clinical applications of display system), so that this information can be used when calculating the *Linearized* distribution to avoid using distributions with fewer output levels than needed.

In addition to the LUM value which describes the accuracy of the linearity of the display system, two values are produced during the linearization that provide useful measures for comparing display systems. The first is the number of theoretically achievable JNDs of the display system. The second is the number of realized JNDs of the display system. The number of theoretically achievable JNDs is simply the number of JNDs predicted by the visual model given the luminance range of the display device used. The more useful number of *Achievable* JNDs, describes how many JNDs are actually realized given the specifics of the display system (i.e. number of bits in DAC and distribution of luminance values). This number is calculated by beginning at the minimum luminance of the display system, and then stepping one JND in luminance above the current luminance value, and then choosing the smallest DDL with a luminance value at least as large as this new value. Repeating this through all the available DDLs will produce a sequence of steps, all at least 1 JND apart, and the number of steps in this sequence is the number of achievable JNDs of the display system. Shown below is a pseudocode (based on C programming language) program is given for calculating the number of achievable JNDs for a display system with characteristic curve DPY, and N digital driving levels.

```
i = 0;
L[i] = DPY[0]
while ((NextLum = L[i] + BartenCT ( L[i] )) < L[n] )
 {
     i = i + 1;
     L[i] = find_minimum_DDL_step(NextLum);
     /*  returns value DPY[k],  where DPY [k-1] < NextLum <= DPY [k]   */
 }
NumberAchievableJNDs = i;
```

4. RESULTS

We have analyzed the default and linearized states of common display systems, including the currently available high spatial resolution, high brightness monitors combined with 8/10 bit DAC display adaptors, an example inexpensive workstation, and film displayed on a mammography lightbox. In the appendix are shown two pictures for each display system. On the left side is a plot of DDLs on the horizontal axis versus luminance (log scale) on the vertical axis, for both the default characteristic curve and the perceptually linearized curve. On the right side is the contrast threshold (vertical axis) versus luminance values (horizontal axis) for the same display system. In this plot four curves are pictured: those for the default characteristic curve (labeled *.default.XGRAPH); those for the linearized curve (labeled *.lin.XGRAPH); those for Barten visual model predicted curve at 1 JND step size (labeled Barten Model); and those for Barten visual model at a step size matched to the display system luminance range (labeled *.barten.XGRAPH). The last being that which is used to calculate the perceptual linearization). The plots in the appendix and the rows in table 1 are of the same display systems: first the four high spatial resolution, high brightness display systems, then the lowend workstation, and finally film on a mammography lightbox.

In the DDL versus luminance plots the luminance (vertical) axis is on a log scale to easily distinguish which curves are linear in log(luminance). Except for very low light levels, the entire display system curve should ideally be linear in log scale, to match the human contrast sensitivity response. As seen in these plots most of the display characteristic curves are non-linear, power curves instead. Most are smooth although several have some anomalies, especially at the low end. On the other hand, the perceptually linearized curves, using the 8/10 DAC, clearly achieve linearity on log scale, and faithfully match the human visual response function. In the contrast threshold versus luminance curves, the desired Barten predicted contrast thresholds (at 1 JND step size) are seen near the bottom, shown with the dashed line with the largest spacing. Usually overlying the Barten curve is the characteristic (default) curve of the display. Thus, the contrast thresholds of the 10 bit display are similar to the most sensitive observer. Unluckily, the characteristic curve generally does not have the same curvature as the Barten curve, thus making it not perceptually linear. Above these two curves are the Barten curve, but scaled to the contrast steps achievable by this display system, while being perceptually linear (shown as un-broken line). Overlying this is the linearized curve for this display system. A successful linearization curve should be centered directly on the Barten scaled curve. The amount of vertical deviation of the linearized curve from the Barten scaled curve indicates how much the contrast thresholds differ. Larger deviations imply poorer linearizations. In all cases the linearized curve should track the Barten scaled curve better than the characteristic curve tracks the Barten single JND curve. This shows the better conformance to equally spaced (in human visual terms) steps. Additionally, in most cases, the linearized curve will show smaller amounts of deviation from the Barten scaled curve than does the characteristic curve from the Barten 1 JND curve.

Overall, the high brightness display systems had fairly consistent curves, and were easily linearized with an 8/10 DAC. The lower end workstation, however, had a poorer luminance distribution, and could not easily be linearized with an 8/8 DAC. The film displayed on the mammography lightbox also exhibited larger non-linearities than the high brightness monitors, and was not as easily corrected with an 8/8 DAC as seen in the appendix. A better linearization was achieved with an 8/10 DAC (not shown).

The second point of the study was to examine the effect of DAC resolution on the linearizations. We calculated the linearization uniformity measure for the characteristic (default) curve, the linearized with 256 levels out of 256 levels result, the linearized with 200 levels out of 256 levels result, and the linearized with 256 out of 1024 levels result for all the display systems listed in the appendix. These values are shown below in table 1.

	Default	*256/256*	*200/256*	*256/1024*
Monitor 1	0.81	0.78	0.52	0.16
*Monitor 2**	0.55	0.55	0.35	0.16
Monitor 3	0.92	1.11	1.02	0.47
Monitor 4	1.28	1.02	0.48	0.28
Monitor 5	2.32	1.08	0.98	NA
Film-LightBox	2.10	0.84	0.55	0.19

Table 1. Linearization Uniformity Measure (LUM) values of all display systems evaluated (shown across rows) are charted. For each display system, four LUM values are shown. Columns one, two, three, and four contain, respectively, the LUM values for the characteristic (default) display system, same display system linearized with 256 output levels from 256 possible, the same display system linearized with 200 output levels from 256 possible, and the same display system linearized with 256 output levels from 1024 possible. *Note that Monitor2 was NOT sampled at all 1024 levels, but instead only sampled at 32 levels, and interpolated to fill in the 1024 values. This generally provides a more uniform sampling than is found when actually sampling at all 1024 levels, so the calculated LUM values shown for Monitor2 are likely somewhat better (lower) than they should be.

Assuming LUM is a good metric for measuring the quality of the linearization, three important conclusions can be drawn from table 1. First, that linearization with the same number of input levels as output levels does not significantly improve the perceptual linearity of the system. The values in table 1 show the 256/256 system to have similar values to the original default characteristic curve, and in some cases worse. This result is expected due to the lack of flexibility in choosing DDLs when the output range is the same as the input range[42]. Second, that when the number of DDLs in the output range is smaller than the number of DDLs to chose from (200 out of 256 in column 2 or 256 out of 1024 in column 3), the linearization improves. Using the 8/8 DAC but limiting our number of output levels to 200 shows an improvement over the 256/256 linearization for each display system. Similarly, the 256/1024 shows a substantial improvement over both the 256/256 and the 200/256 cases. Thus, as a rule of thumb, the larger the number of potential DDL choices versus the number utilized, the better the possible linearization. The trade off of decreasing the number of output DDLs to improve linearity, is the loss of contrast resolution. Thus, as discussed earlier, it is important to know the contrast resolution requirements of clinical applications. For instance, we could achieve a LUM value of 0.18 for Monitor 2 if we choose to use a 80/256 mapping (only 80 output levels). Most clinical applications, however, would see a loss of contrast information if only 80 distinct output levels were utilized. The third important point, is that with an 8/10 DAC, and the 256/1024 linearization, the LUM values were significantly improved in all cases, while maintaining the full 256 output levels. Thus 8/10 DACs are superior to the 8/8 DACs in achieving a perceptually linear system while maintaining the full contrast resolution (256 levels) of the display system. Note, however, that the contrast thresholds achieved by the 8/10 DAC are still coarser than the visual capabilities of the human observer as seen in the CT versus luminance plots of the appendix. This supports previous arguments[42,34,11,12] that the appropriate DAC resolution is in the range of 10-12 bits, and suggests the next step would be for manufacturers to produce a 10/12 bit DAC, which may turn out to be sufficient for medical imaging display purposes. In the future, we plan to conduct formal observer studies to evaluate the effect of the LUM metric with respect to observer performance and observer preference.

5. DISCUSSION

To date, the linearization methods developed have been aimed at simply implementing a reasonable $(\mathrm{CSF} \bullet \mathrm{DACLUM})^{-1}$ function. For many monitors the limitation of 8 bit DACs and sub-optimal luminance distributions (not matching the human observer CSF) mean that calculated linearizations may not be very optimal, and in many cases, turn out to be worse than not linearizing. Improvements have occurred with the advent of the 8/10 bit DACs, which allow more accurate perceptual linearizations without decreasing contrast resolution. Developers of display adapters should continue to increase their DAC resolution, with 10 to 12 bit DACs expected to meet the needs of medical image display. While the characteristic (default) curves of most of the high end monitors we evaluated were reasonably close to being perceptually linear, some improvements remain, especially at low luminances. More significant improvements will be required for the lower end display systems, for instance generic workstations, which were not as perceptually linear. We encourage the monitor manufacturers to better match their luminance distributions to CSF distributions, specifically the proposed Barten standard, both to improve the inherent perceptual linearity of their system, and to better allow for after-market perceptual linearization corrections of the display system.

6. ACKNOWLEDGMENTS

This work was supported in part by NIH grants P01-CA47982, R01-CA60193, and R01-CA44060.

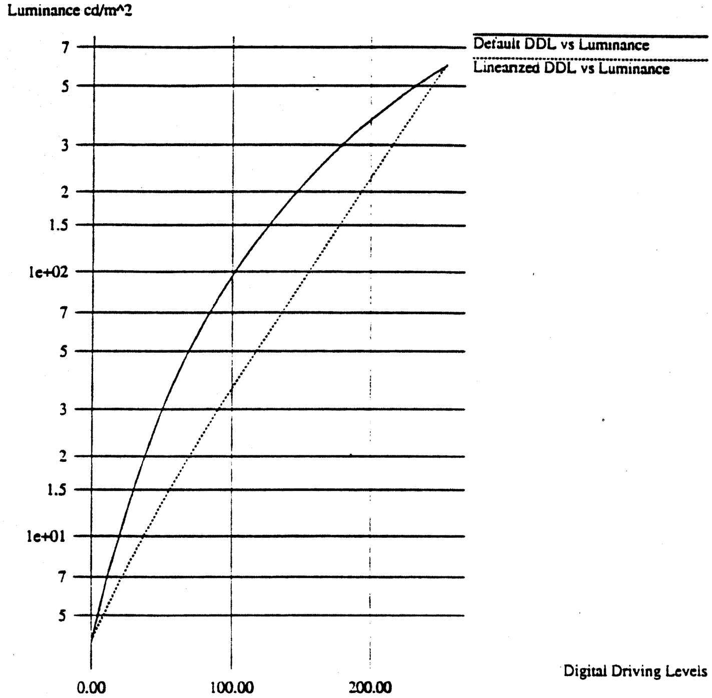
7. FOLLOW-UP WORK

While the LUM metric proposed by Hemminger, and used in this paper, correctly measures of the amount of deviation in the linearization, there are two drawbacks. First, it presents data relative to the Barten curve. This requires the use of the DisplayCT/BartenCT ratios, including rescaling, to map the data so that it is statistically analyzable using simple methods. Second, the BartenCT data used for the comparison is a scaled version of the 1 JND stepsize original Barten curve. To address this, Hemminger has proposed a more elegant metric for this problem, the number of JNDs (vertical axis) versus $DDL_{interval}$ (horizontal axis) which captures more directly the issue of the size of each step of a luminance distribution. This then yields to a standard multiple regression statistical analysis to compute the R^2 values of B_0 (straight line no slope), B_1 (linear), B_2 (quadratic), and B_3 (cubic) fits. The R^2 value of the B_0 fit produces a number from 0.0 (best) to 1.0 (worst) which defines the quality of the linearity of any luminance distribution. Additionally, the intercept, B_0, indicates the mean number of JND steps per interval, which is the best measure of contrast resolution. We expect the results with the new metric to be similar to the previous LUM definition using contrast ratios because it is calculated from essentially the same variances. We are currently in the process of migrating our programs over to the new LUM metric.

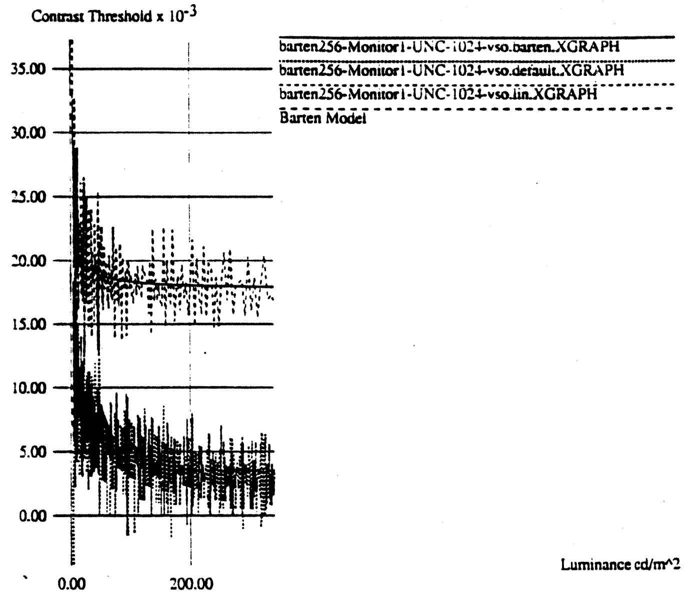
8. APPENDIX

The following graphs are of DDL versus luminance, and luminance versus contrast threshold, for each of the four high brightness monitors, the low-end workstation, and film on a mammography lightbox. A discussion of these graphs is given in section IV RESULTS.

DDLs vs Luminance

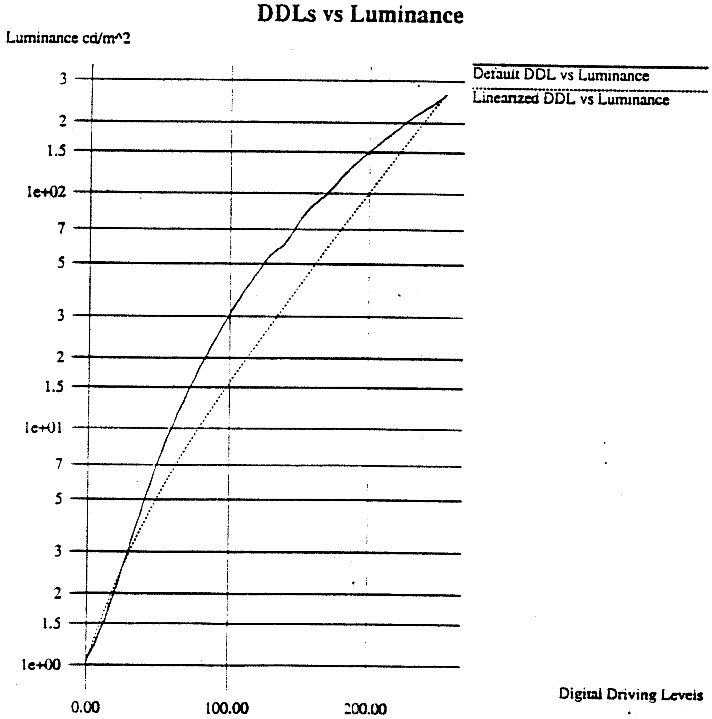


Contrast Thresholds vs Luminance

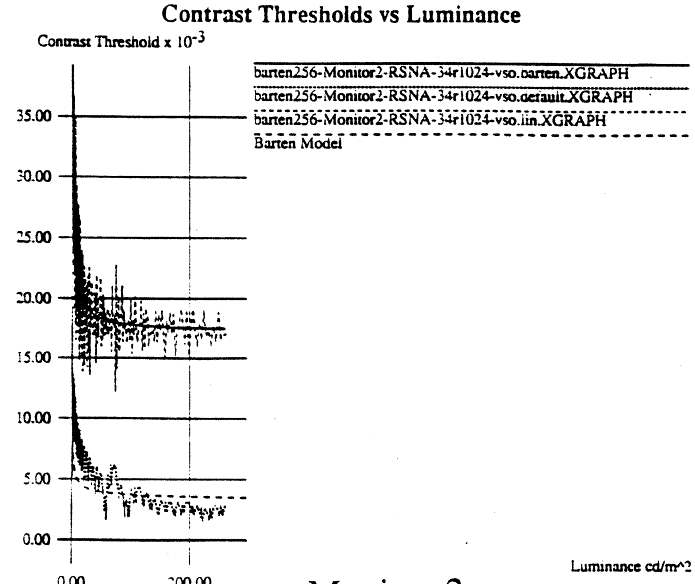


Monitor 1
(High Brightness)

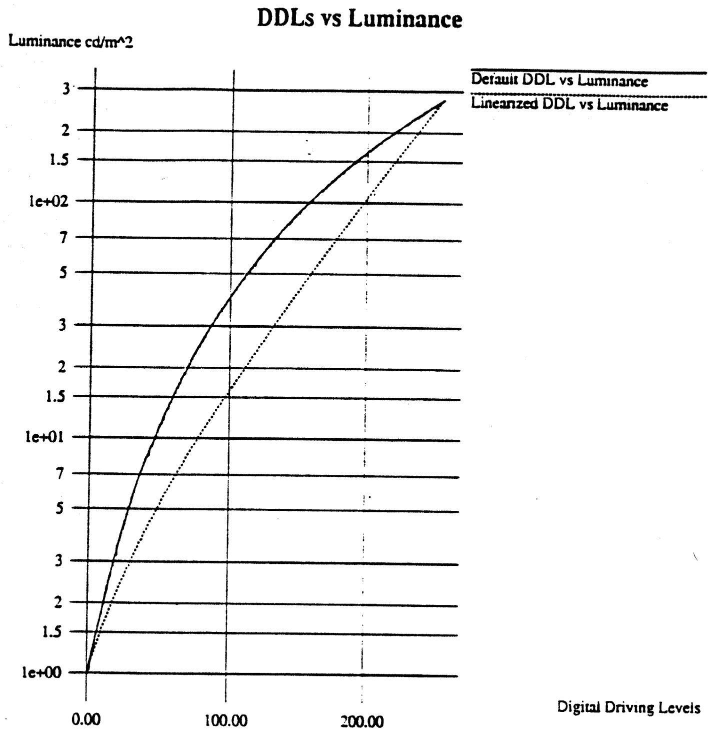
DDLs vs Luminance

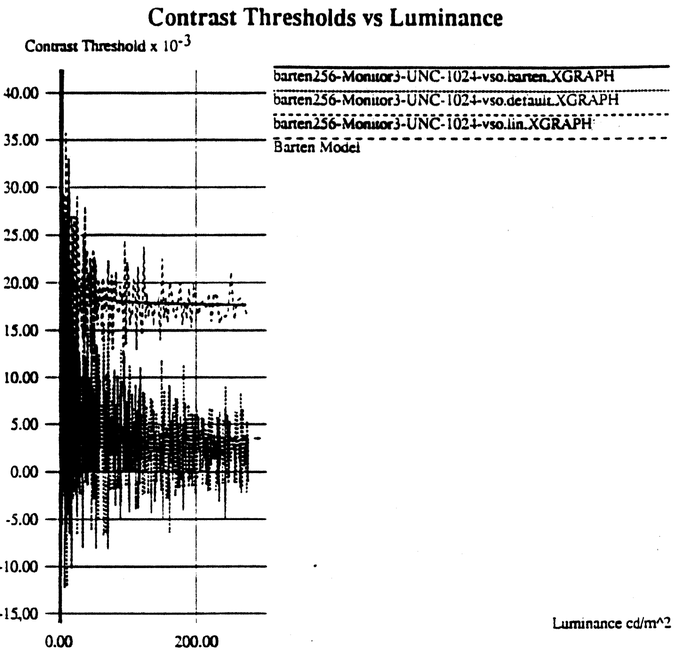


Contrast Thresholds vs Luminance

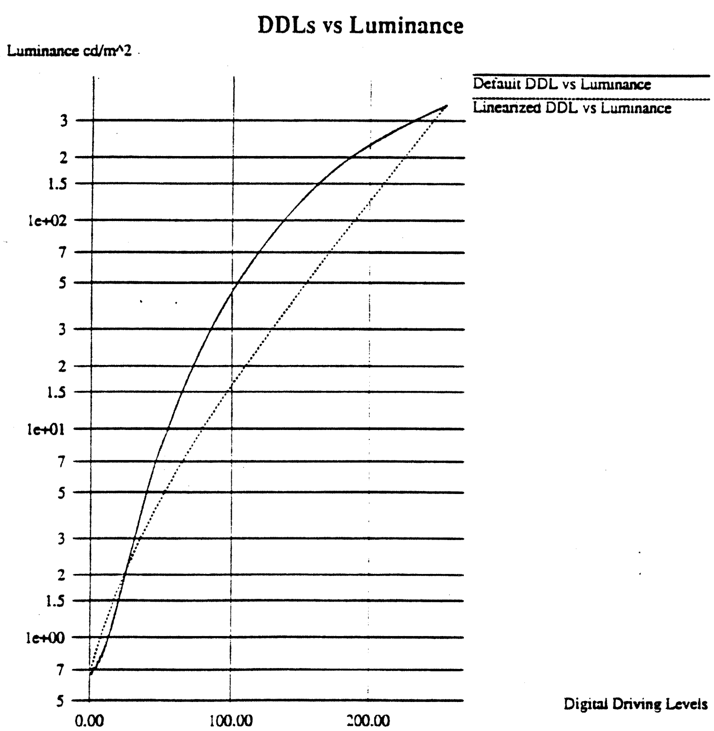


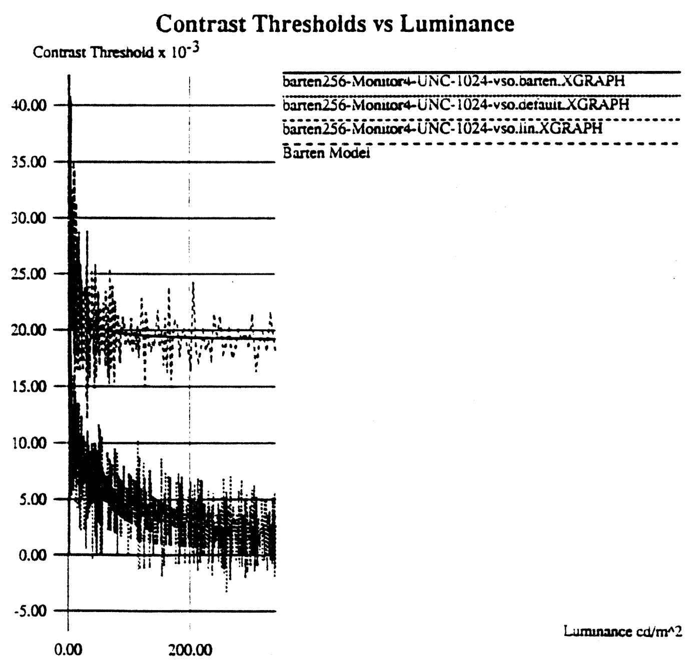
Monitor 2
(High Brightness)



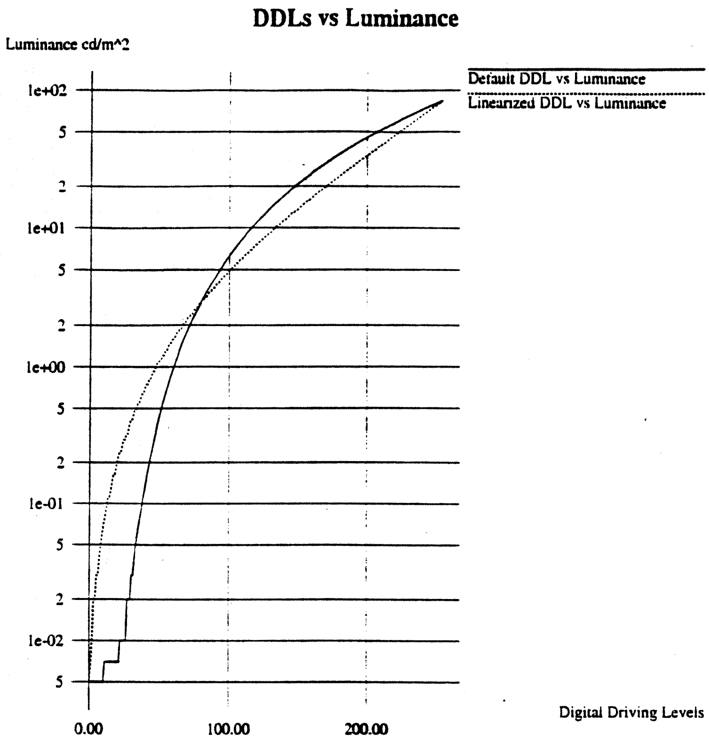


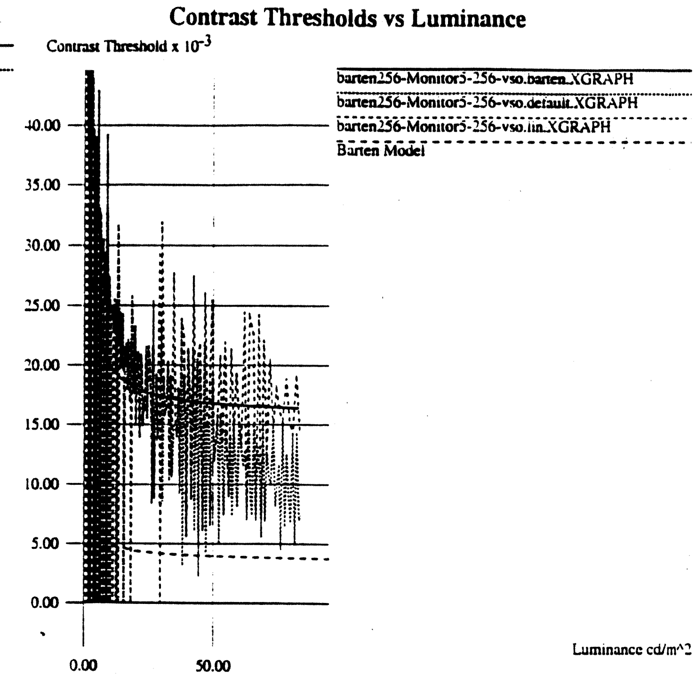
Monitor 3
(High Brightness)



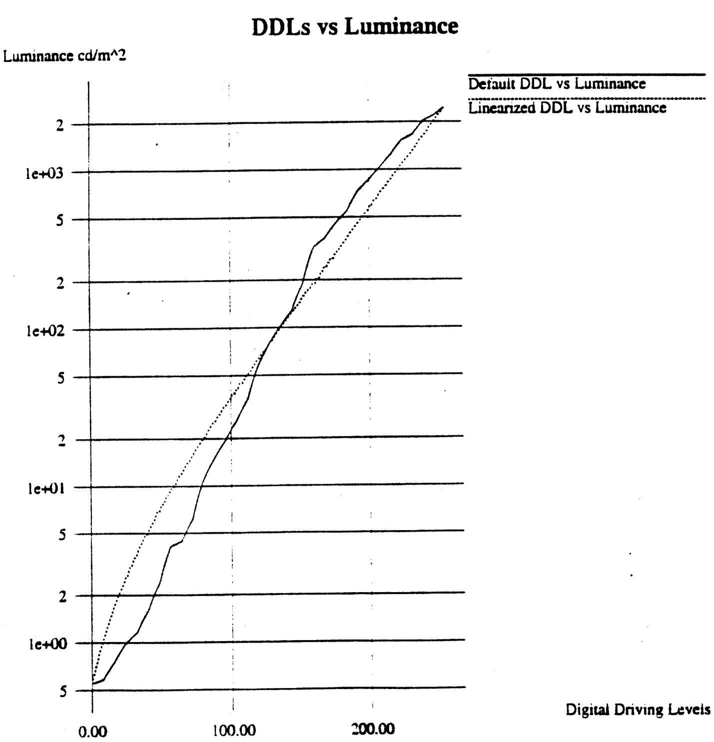


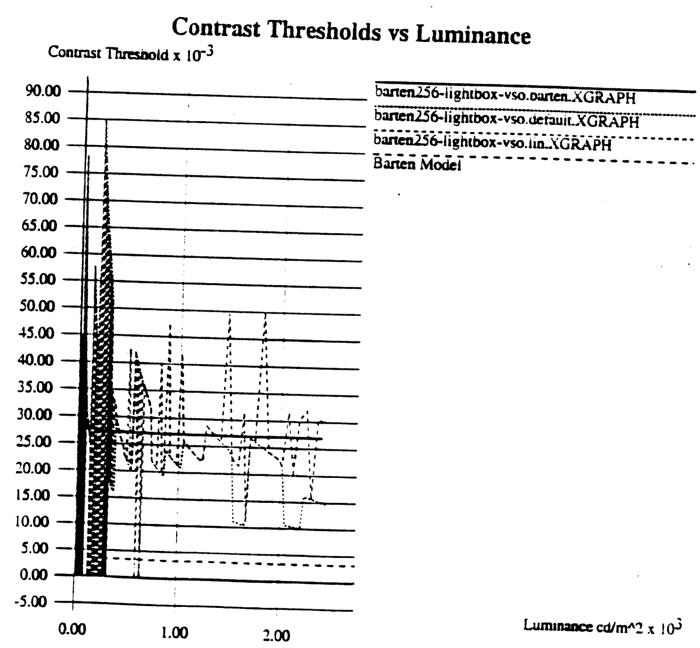
Monitor 4
(High Brightness)





Monitor 5
(low end workstation)





Film on Lightbox

8. REFERENCES

1. Pizer, SM, Chan, FH, "Evaluation of the number of discernible levels produced by a display", *Information Processing in Medical Imaging*, R DiPaola and E. Kahn, Eds., Editions INSERM, Paris, pp. 561-580, 1980.

2. Pizer SM, Intensity mappings: liearization, image-based, user-controlled", SPIE Display Technology II, Vol 271, pp21-27, 1981.

3. Pizer, SM, "Intensity mapping to linearize display devices", *Computer Graphics and Image Processing*, 17, pp 262-268, 1981.

4. Pizer SM, Johnston RE, Zimmerman JB, Chan FH, "Contrast perception with video displays", SPIE Vol 318, PACS for Medical Applications, pp223-230, 1982.

5. Pizer SM, "Psychovisual Issues in the Display of Medical Images", Proc NATO Advanced Studies Institute on Pictorial Information Systems in Medicine, 1985.

6. Johnston RE, Zimmerman JB, Rogers DC, Pizer SM, "Perceptual Standardization", Proceedings of Picture Archiving and Communcations Systems III", SPIE Vol 536, pp. 44-49, 1985.

7. Rogers DC, Johnston RE, Pizer SM, "Effect of Ambient Light o Electronically Displayed Medical Images As Measured by LUmikance Discrimination Thresholds", Journal Optical Society of America, 4, 5, May 1987.

8. Cromartie RC, Johnston RE, Pizer SM, Roger DC, "Standization of Electronic Display Devices Based on Human Perception", University of North Carolina at Chapel Hill, Technical Report 88-002, December 1987.

9. Ji, T, Roehrig, H, Blume, H, Guillen, J, "Optimizing the display function of display devices", *Proceedings Medical Imaging VI: Image Capture, Formatting*, and Display, Vol 1653, pp. 126-139, February 1992.

10. Briggs SJ, "Photometric technique for deriving a "best gamma" for displays", Advances in Display Technology, SPIE 199, pp 134-145, 1979.

11. Blume H, Roehrig H, Browne M, Ji TL, "Comparison of the Physical Performance of High Resolution CRT Displays and Films Recorded by Laser Image Printers and Display on Light-Boxes and the Need for a Display Standard", Proceedings of Medical Imaging: Image Capture, Formatting and Display, SPIE volume 1232, pp97-114, 1990.

12. Blume H, Daly S, Muka E, "Presentation of Medical Images on CRT displays A Renewed Proposal for a Display standard", Proceedings of Medical Imaging: Image Capture, Formatting and Display, SPIE volume 1897, pp215-231, 1993.

13. Naiman AV, Makous W, "Spatial non-linearities of grayscale CRT pixels", Proceedings of Human Vision, Visual Processing, and Digital DIsplay III, SPIE volume 1666, pp41-56, 1992.

14. Roehrig H, Blume H, Ji TL, Sundaresham MK, "Noise of the CRT Display System", Proceedings of Medical Imaging: Image Capture, Formatting and Display, SPIE volume 1897, pp232-245, 1993.

15. Barten PGJ, "Physical Model for the Contrast Sensitivity of the Human Eye", Proceedings of Human Vision, Visual Processing, and Digital Display III, SPIE vol 1666, 1992.

16. Daly S, "The visible differences predictor: an algorithm for the assessment of image fidelity", Proceedings of Human Vision, Visual Processing, and Digital Display III, SPIE vol 1666, 1992.

17. Murchison C (Editor), A Handbook of General Experimental Pschology, p769, Clark University Press, Worcester MA, 1934.

18. Hill CR, The Physics of Medical Imaging, p567-583, Webb S (Ed), IOP Publishing Ltd, Philadephia PA, 1988.

19. Wyszecki, G, Stiles, WS, Color Science: Concepts and Methods, Quantitative Data and Formulae, pp. 567-570, 2nd Edition, John Wiley and Sons, New York, 1982.

20. Cornsweet TN, Visual Perception, Academic Press, New York, 1970.

21. Haber RN, Hershenson M, The Psychology of Visual Perception, Holt Rinehart and Winston, New York NY, 1973.

22. Bijl P, Koenderink JJ, "Visibility of Elliptical Gaussian Blobs", Vision Research, Vol 33, No. 2, pp243-255, 1993.

23. Bijl P, Koenderink JJ, Toet A, "Visibility of Blobs with a Gaussian Luminance Profile", Vision Research, Vol 29, No. 4, pp447-456, 1989

24. Daly, S, "Quantitative Performance Assessment of an Algorithm for Determination of Image Fideltiy", S.I.D. Digetst of Technical Papers, V. 24, pp317-320, 1993.

25. Rogers JG, Carel WL, "Development of design criteria for sensor displays devices", Report HAC Ref. No. C6619, Huges Aircraft Company, Culver City, CA, Dec 1973.

26. Barten P, "Spatio-Temporal Model for the Contrast Sensitivity of the Human Eye and its Temporal Aspects", Proceedings of Human Vision, Visual Processing, and Digital Display IV, SPIE Vol 1913, pp2-14, 1993.

27. Lie I, "Visual Detection and Resolution as a Function of Adaption and Glare", Vision Research Vol 21, pp1793-1797, 1981.

28. Wildt GJ, Waarts RG, "The Influence of the Surround on Contrast Sensitivity", 2nd Intl. Conf. on Visual Psychophysics and Medical Imaging, Brussles, Belgium, July 2-3, 1981. (NY, USA:IEEE 1981).

29. Hood DC, Finkelstein MA, Sensitivity to light Handbook of Perception and Human Performance, Vol 1, eds Boff KR, Kaufman L, and Thomans JP, pp5-28, Wiley, 1986.

30. Kundel HL, "Visual Clues in the Interpretation of Medial Images", Journam of Clinical Neuropysiology, 7(4):472-483, 1990.

31. Burbeck CA, Pizer SM, "Object Representation by Cores", UNC-CS Technical Report 94-004, 1994.

32. DeGroot SG, Gebhard JW, "Pupil Size determined by adapting luminance", Journal of Optical Society America, 42, p492-295, 1952.

33. Kirk RE, Experimental Design: Procedures for Behavioral Sciences, pp261, Second Edition, Wadsworth, Belmont CA, 1982.

34. Sezan MI, Yip KL, Daly SJ, "Uniform Perceptual Quantization: Applications to Digital Radiolography", IEEE Transactions on Man, Machine, and Cybernetics, Vol SMC-17, No 4, pp622-634, 1987.

35. Ji T, Roehrig H, Blume H, Guilllen J, "Optimizing the Display Function of Display Devices", Proceedings of Medical Imaging VI: Image Capture, Formatting, and Display, eds Kim Y, SPIE vol 1653, 1992.

36. Robertson PK, "Color Image Display: Computational Framework Based on a Uniform Color Space", PhD thesis, Australian National University, April 1985, (CSIRO Tech Report No. 27)

37. Levkowitz, H, Herman GT, "Color Scales for Image Data", *IEEE Computer Graphics and Applications*, Vol 12, No 1, pp72-80, Jan 1992.

38. Jain AN, Fundamentals of Digital Image Processing, pp99-123, Prentice Hall, Englewood Cliffs NJ, 1989.

39. Max J, "Quantizing for Minimum Distortion", IRE Trans. on Info. Th., Vol IT-6, No. 1, pp7-12, March 1960.

40. Daly S, personal communication, Jan 1994.

41. Barrett HH, "Objective Assessment of Image Quality: Effect of Quantum Noise and Object Variability", J. Opt. Soc. Am., 7, pp1266-1278, 1990.

42. Hemminger BM, Johnston RE, Rolland JP, Muller K, "Perceptual Linearization of video display monitors for medical image presentation", Proceedings of Medical Imaging 1994: Image Capture, Formatting, and Display, eds Kim Y, SPIE vol 2264, 1994.

43. Hemminger BM, "Minimum Perceptual Error Calculation for Perceptual Linearization", UNC Chapel Hill Dept of Computer Science Technical Report, TR94-032, 1994.