NASA-CR-201011

SPATIAL RESOLUTION, GRAYSCALE, and ERROR DIFFUSION TRADE-OFFS: IMPACT ON DISPLAY SYSTEM DESIGN

FINAL REPORT

Cooperative Agreement NCC2-788

Western Aerospace Laboratories, Inc

Los Gatos, CA

Jennifer L. Gille, Ph.D.

Principal Investigator

Submitted to:

NASA Ames Research Center

Moffett Field, CA

James Larimer, Ph.D., Technical Monitor

May 6, 1996

MAY 13 1996

CASI

SPATIAL RESOLUTION, GRAYSCALE, and ERROR DIFFUSION TRADE-OFFS: IMPACT ON DISPLAY SYSTEM DESIGN

Jennifer Gille[1], *Russel Martin*[2], *James Larimer*[3]

[1] Western Aerospace Laboratories, NASA/Ames MS 269-6, Moffett Field, CA 94035-1000
[2] Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304
[3] NASA/Ames Research Center, NASA/Ames MS 269-6, Moffett Field, CA 94035-1000

Abstract: We examine technology trade-offs related to grayscale resolution, spatial resolution, and error diffusion for tessellated display systems. We present new empirical results from our psychophysical study of these trade-offs and compare them to the predictions of a model of human vision.

Introduction

The technologies used in display systems have implicit trade-offs in the cost and ease of manufacture. For a matrixed display such as an AMLCD, grayscale and spatial resolution directly affect cost, and increasing either increases manufacturing costs. Grayscale and spatial resolution can be traded off [1,2] and error diffusion can be used to reduce the visibility of stair-stepping artifacts that result from quantizing grayscale. The trade-offs among these technologies can be difficult to assess. Which trade-offs are effective is constrained by the visual system of the user. The ultimate criterion for judging display quality resides in the human visual system, so technology trade-offs must be evaluated in terms of their impact on system performance with the human viewer as an essential system component. A modeling environment [3] which uses a model of human vision [4] has been developed as a design tool for making these trade-offs. But models must be validated to assure the user that they correctly evaluate the visual phenomena that result from various combinations of display technologies.

This paper will present an empirical investigation of the grayscale/spatial-resolution trade-off and how that trade-off is affected by error diffusion. These data confirm the predictions of a model of human spatial vision that can be used to evaluate design trade-offs produced by various combinations of these technologies.

Spatial resolution

Device manufacturers typically specify resolution in terms of dots per inch (dpi). This measure maps directly into manufacturing cost considerations. The resolution of a typical laptop computer display is given in terms of pixel counts, for example, 640X480 pixels (VGA). The appropriate measure of resolution for the human visual system, however, is pixels per degree visual angle, or cycles per degree. The limiting resolution of the human visual system which is generally taken to be around 60 cycles per degree, or 120 pixels per degree, at 100% contrast for the average person. For a display viewed at 0.5 m, this would be about 350 dpi. Laptop computer displays are limited in diagonal measure. At 0.5m viewing distance, a 0.25m diagonal VGA display (about 80 dpi) would be approximately 30 pixels per degree.

Some of the manufacturing factors that are controlled by these dimensions in an AMLCD are: (1) pixel aperture, (2) lithographic feature size or "design rule", (3) liquid crystal domain size, and (4) manufacturing yield. These factors are not independent. As the pixel size decreases the aperture decreases. This affects the relative size of the design rule, since TFT and storage capacitor features are usually fixed in size. As the design rule is reduced in size to accommodate higher pixel counts with reasonable aperture ratios, manufacturing yields are reduced.

Grayscale

In printing grayscale is achieved by halftoning. Here very small dots of ink are used in a spatial pattern to achieve larger areas of gray out of composites of many dots of ink. In a TN LC device, grayscale can be achieved by varying the electric field on the pixel. But this has several problems. First, carefully controlling the voltage at each pixel location in an AMLCD requires a high degree of uniformity in the TFT operating characteristics. It is typically difficult to achieve uniform drive voltages in a-Si TFT AMLCDs. These problems imply greater manufacturing consistency which means lower yields and higher costs. Drivers are an additional problem. Drivers that can produce more voltage levels are more costly and require more inputs to the driver circuitry.

Error diffusion

Manufacturers have attempted to achieve grayscale in AMLCDs by techniques similar to halftoning in printing. These techniques use groups of pixels whose aggregate values are controlled spatially and temporally by a special controller chip which produces a dither pattern. Error diffusion, the dithering scheme used in LCDs, is a type of dither where the error in grayscale value introduced by quantization is spread over small regions in space and time making the average value in this region as close as possible to the average value for the corresponding region in the continuous tone image. Implementing an error diffusion scheme trades off the costs associated with driver and TFT complexity for the costs of the controller chip that operates between the frame store and display.

Human Visual System

The human visual system is limited in its ability to perceive grayscale steps and spatial resolution. There is a trade-off within the visual system where one of these can be traded for the other. This is why halftoning works in printing and anti-aliasing works for lower-resolution displays. Specifying the regions of trade-off, therefore, will be useful to the display system designer who must trade cost of manufacture against system performance to achieve an efficient design.

Modeling and empirical methods

Our previous studies of the grayscale/spatial resolution trade-off, reported at SPIE '94 and SID '94 [1,2], indicated that tessellated displays are prone to several artifacts of spatial and grayscale quantization, including jaggies, apparent pixel width, banding, false luminance contours, and simultaneous contrast effects. For this reason, low-resolution versions of images are always distinguishable from one another based on changing patterns of artifacts. For lower-spatial-resolution images, we could not establish a grayscale/spatial resolution trade-off. For the present study, we made three changes in this regard. First, we greatly increased the equivalent spatial resolution of our simulation. Second, we compared each image to the continuous-tone high-spatial-resolution version in order to find discriminabilities that are equivalent, even though based on different artifacts. Third, we incorporated error diffusion as an additional manipulation of stimuli.

Our previous results also indicated that the Sarnoff Human Vision Model was an excellent predictor of discriminability for our stimuli. The present study was designed to further test the model, but we also used its computational results as pilot data to guide the creation of our sets of psychophysical stimuli.

The empirical study included two phases. In the first phase, the Sarnoff Human Vision Model was used to predict the discriminability of images differing in grayscale and resolution from a continuous-tone high-resolution version. In the second phase, discrimination thresholds for these images were measured in three observers and compared to the predictions of the model.

Figure 1. Examples of zone-plate stimuli

Stimuli

The single test image against which all others were compared is shown in Figure 1a. This type of image, which we call a zone plate, is a radially symmetric spatial frequency chirp. Our previous work has shown that the zone plate test image embodies many of the qualities of natural images that are prone to produce artifacts in low-resolution (both spatial and grayscale) versions of those images. Displayed on the simulation-device CRT screen, the zone plate test image subtended 1.48 x 0.99 degrees of visual angle, with an equivalent resolution (for viewing at 0.5 meter) of 1200dpi and 256 levels of gray. Comparison images that differed in equivalent spatial resolution were created by using aggregated CRT pixels to simulate square-pixel tesselation. Quantized grayscale resolution was varied by linear binning of luminance levels. Multi-bit error diffusion used the Floyd-Steinberg technique. Figure 1b shows a zone plate with one-eighth the spatial resolution of the test image and four levels of gray with simple quantization. Figure 1c shows a zone plate with one-fourth the spatial resolution of the test image and two levels of gray with error diffusion.

Figure 2. Predictions of the Vision Model

Sarnoff Human Vision Model

The Sarnoff Human Vision Model is a model of visual discrimination performance. It takes as input two digitized images in luminance units, with specifications of viewing parameters such as stimulus distance, image width, observer fixation state, stimulus eccentricity, screen reflectivity and illuminance [5]. The output of the model is a map of the differences between the two images, for a human observer, expressed as the number of Just Noticeable Differences (JNDs) at each location. This array of JNDs for the comparison was summarized by computing the root-mean-square (RMS) JND of the entire array. The RMS JND for the comparison of two images will predict their discriminability. Threshold for discrimination should be at RMS JND = k, where we expect k=1. The exact value of k depends on the task. A different JND summary statistic would have a different k value.

In the present study, the discriminabilities of lower-resolution versions of the zone plate compared to the high-resolution test image were computed using the Sarnoff model. Model predictions, in RMS JND, for the comparison of the high-resolution test image to lower-resolution zone plates are shown in Figures 2a and 2b, for quantized grayscale and error-diffused versions, respectively. Again, threshold for discrimination should be at about RMS JND = 1. The significance of these results is discussed below.

Discrimination thresholds

Discrimination thresholds for three observers, JG, JL and RM, in several conditions were measured using the following procedure. Stimuli were displayed on two calibrated Barco CRTs. Observers viewed the monitors using two first-surface mirrors such that the folded light path was 32' 4" long. At this distance, the resolution of the CRTs was the equivalent of viewing 1200 dpi screens at about 20", a standard viewing distance for desk-top work. Viewing distance was held constant by a head rest, but otherwise there was free viewing. The psychophysical procedure was a four-alternative, forced-choice, one-up-two-down double random staircase. On each trial of a given staircase, two zone plates were displayed on each of the monitors. Of the four zone plates, three were the standard continuous-tone high-resolution image, and one was the lower-resolution comparison image. The observer's task was to indicate which of the four zone plates was different from the other three.

Observer thresholds were measured for several conditions of resolution; in each condition either the grayscale was held constant and the spatial resolution threshold measured, or *vice versa*. Some conditions were limiting conditions, e.g. for a low spatial resolution condition, a threshold grayscale resolution could not be found, because the comparison stimuli were always discriminable from the standard on the basis of visible tessellation.

Figures 3a and 3b show the measured thresholds for quantized grayscale and error diffusion, respectively. The data from all three observers is shown together. Also plotted on these graphs are the contour lines from the surface generated by the predictions of the vision model. Empirical thresholds cluster around the contour lines for RMS JND = 1.0 as predicted, validating the model as measuring discriminability. The significance of these results is discussed below.

Figure 3. Discrimination thresholds for three observers compared to isodiscrimination contours computed by the vision model.

Results

Modeling

The results of the modeling phase of our study are shown in Figure 2. The images formed by each combination of eight spatial resolutions and five grayscale resolutions, with and without error diffusion, were compared to the continuous-tone high-resolution zone plate using the Sarnoff model. RMS JNDs were calculated and displayed here against spatial resolution on the abscissa, number of gray levels as separate lines, and simple quantization or error diffusion in two separate plots. Spatial resolution is measured in bits as $\log_2$(equivalent dpi^2); for convenience the points are labelled by their equivalent dpi. Predicted threshold for discrimination is at RMS JND = 1.

Quantized gray levels are shown in Figure 2a. There are five main features of this graph. First, below about 50 dpi, spatial resolution dominates; that is, increasing the number of gray levels does little to improve the appearance of the image. Second, above about 150 dpi, grayscale dominates, and increasing spatial resolution does not improve the image for a given number of gray levels. Third, the region where a grayscale/resolution trade-off could be said to exist is between 50 and 150dpi. For example, a 75 dpi image with eight levels of gray and a 130 dpi image with four levels of gray will be about equally different from the high-resolution image, although all three will be mutually discriminable. Fourth, increasing the number of gray levels above 16 has little value. Fifth, in order to produce an image that is indistinguishable from the high-resolution image, it must have both at least 300 dpi spatial resolution and 16 levels of gray. If the intent is to design to the point where the image is indistinguishable from the high-resolution standard, lower resolution in either variable cannot be compensated for by higher resolution in the other.

Multi-bit error diffusion is shown in Figure 2b. Again, below about 50 dpi, spatial resolution dominates, increasing the number of gray levels does not result in an improvement, and the use of error diffusion has no significant effect over simple quantization. However, in the previously grayscale-dominated region, there is significant improvement with the use of error diffusion. In this region, for instance, 150 dpi with 8 gray levels, 300 dpi with 4 gray levels, and 600 dpi with 2 gray levels should all be conservatively below threshold for discrimination from the 1200 dpi with 256 gray level image. Overall, there is little to be gained by increasing the number of gray levels above eight.

Human psychophysical thresholds/model validation

The results of the threshold measurement phase of our study are shown in Figure 3. The variable space for the study is plotted as grayscale *vs.* spatial resolution in bits ($\log_2$[number of gray levels] and $\log_2$[equivalent dpi^2], respectively). The labelled lines are the projections of contour lines from the surface generated by the predictions of the Sarnoff model. Since it is predicted that empirical thresholds should fall where RMS JND = 1, data points should fall near the contours labelled "1.0", and regions above and to the right of these contours should represent below-threshold regions where the lower resolution images are visually equal to the high-resolution standard. The data from the observers JG, JL and RM are plotted as upright triangles, inverted triangles, and discs, respectively.

There are five main features to these plots. First, the data points for all three observers cluster around the contour for RMS JND = 1 as predicted, strongly validating the model for threshold measurements. Second, the below-threshold region for images formed using error diffusion is considerably larger than that for using simple quantization, indicating the utility of this technique for tessellated displays. Third, for simple quantization, at least 16 levels of gray are needed to adequately reproduce the high-resolution image, but for error diffusion there is a region of trade-off between grayscale and spatial resolution. Fourth, neither of the two RMS JND = 1 contours go below about 150 dpi in spatial resolution, indicating a practical spatial resolution limit for good reproduction of the high-resolution image. Fifth, the RMS JND =1 contours for the two plots meet at about 150 dpi with 32 gray levels, indicating that at this spatial/grayscale resolution, there is no advantage to the use of error diffusion for good reproduction of the high-resolution image.

Application to system design

The simulation of human visual response provides us with a measure of the difference between the "perfect" image and an image on a matrix display with square pixels. The psychophysical measurements have given us a value for the threshold of perceptible differences between the display image and the "perfect" image. From this we may deduce the requirements for a display matched to the limits of human vision. The region where spatial frequency dominates in Figure 2 is below 50 dpi, while the region where grayscale dominates is above 150 dpi. The threshold of perception is approximately 1 JND.

Most displays today are in the range between the two regimes; they typically run from 72 dpi to 120 dpi. An exception would be the 6.3 Million Pixel display reported by Xerox [6]. This display has two gray levels (binary drive) and 284 dpi resolution. The number of gray levels available on displays are determined by the column driver and by the display design. Most displays are currently 8 or 16 levels while up to 256 level drivers are becoming available [7]. In the intermediate range, increasing the gray levels to at least 16 will always improve the image for quantized images. Above that level there was no improvement for these test images. This implies diminishing returns for the addition of extra gray levels beyond 16. The decrease in difference from the "perfect" image continues with increasing spatial resolution for the entire regime. Therefore, for improved image quality it is desirable to reduce the pixels size to at least a density of 150 dpi.

The requirements for gray scale can be dramatically reduced by the use of error diffusion. This, however, comes at a system cost. By its use one can reduce the needed grayscale for the 150 dpi displays from 16 to 8 levels. It will improve the quality of images that are above the discrimination threshold as well. The advantages are balanced by the cost of dithering, which may be done in either hardware or software. In software the cost is processing time, although advanced systems can perform dithering relatively quickly. Hardware can also implement dithering. Blue Noise masking [8] is particularly well adapted to hardware because it is a point operation and can be performed at video rates [9].

These results have been borne out by the performance of the Xerox 6.3 Million Pixel Display. This approximately 300 dpi binary display shows obvious artifacts with quantized images having moderate spatial frequencies, as with the stimuli used in this paper. However, when such images are shown with error diffusion few artifacts are apparent.

There are limitation of this work which should be noted. First, although the zone plates used in this test have a wide range of spatial frequencies and gray levels, they do not cover all stimuli. For example, large regions of very low (but not zero) gray level would have occasional non-zero pixels if error diffused. Because of the sensitivity of human vision to fractional contrast, these pixels stand out strongly against the background. Another point to be aware of is that although these tests all assume a 0.5 m viewing distance, there is nothing to keep the final consumer of a display system from working at a closer distance. When viewed from shorter distances, dithering artifacts can become obvious. Finally, one must recognize that this discussion addresses only image quality, not pixel count. As pixel size is reduced, if pixel count is not increased, the amount of information available on a display may be reduced, decreasing its usefulness.

Appendix: Sarnoff Human Vision Model

The Sarnoff Human Vision Model calculates the discriminability of two images. It includes representations of the eye's optics, early adaptation, and selectivity of orientation and spatial scale in the visual system. The elements of the model are: (1) a transformation of the image file to a standard format that reflects viewing conditions; (2) a stage modeling photoreceptor sampling; (3) the generation of a seven-level contrast pyramid, based on a standard observer, that reflects local change within each of seven spatial frequency bands; (4) the decomposition of the contrast pyramid into four oriented pyramids that reflect orientation asymmetries in visual processing; (5) non-linear transformation; (6) pooling stage; (7) the construction of a discriminability map; and (8) the calculation of the RMS JND. The free parameters of the model have been previously fixed by fitting it to two standard data sets. Model parameters are therefore independent of the data reported in these experiments. The model has been validated using a variety of human psychophysical data.

References

1. J. Gille, R. Samadani, R. Martin, and J. Larimer, "Grayscale/resolution tradeoff," *Proceedings of the SPIE* , *2179,* 47-59 (1994).
2. J. Gille, R. Martin, R. Samadani and J. Larimer, "A human vision discrimination model analysis of the grayscale/resolution tradeoff: displays with 150 dpi or less resolution," *SID Digest,* Vol. XXV, 194-497 (1994).
3. J. Larimer, et al, "A Video Display Engineering and Optimizaton System: ViDEOS," *SID Digest,* Vol. XXV, 197-200 (1994).
4. J. Lubin, J. Bergen "Cockpit display visibility modeling," *NASA Contractor Report* 177623 (1993).
5. J. Larimer, M. Prevost, A. Arditi, S. Azueta, J. Bergen and J. Lubin, "Human visual performance model for crewstation design," *Proceedings of the SPIE,* 1456, 196-210 (1991).
6. R. Martin, et al, "A 6.3 Million Pixel AMLCD," *Technical Digest of the Society for Information Display Symposium,* 704-707 (1993).
7. A. Erhart, "A High-Voltage High-Gray-Scale LCD Column Driver Utilizing a Standard 1.3-um CMOS Process," *Technical Digest of the Society for Information Display Symposium,* 471-4 (1994).
8. T. Mitsa and K. Parker, "Digital Halftoning using a Blue Noise Mask," *Proceedings of ICASSP, Toronto,* 2809-12 (1991).
9. Da Costa, Victor, Xerox PARC, personal communication.

Acknowledgment

This work was funded by a NASA research grant awarded to Western Aerospace Laboratories under NCC 2-788 (Dr. James Larimer, Technical Monitor), and the ARPA High-Definition Systems Program.

Figure 1. Examples of zone plate stimuli that vary in spatial resolution, grayscale and use of error diffusion. (a) standard: high-resolution, continuous-tone (b) one-eighth the spatial resolution, 4 levels of gray (c) one-fourth the spatial resolution, 2 levels of gray, error diffusion

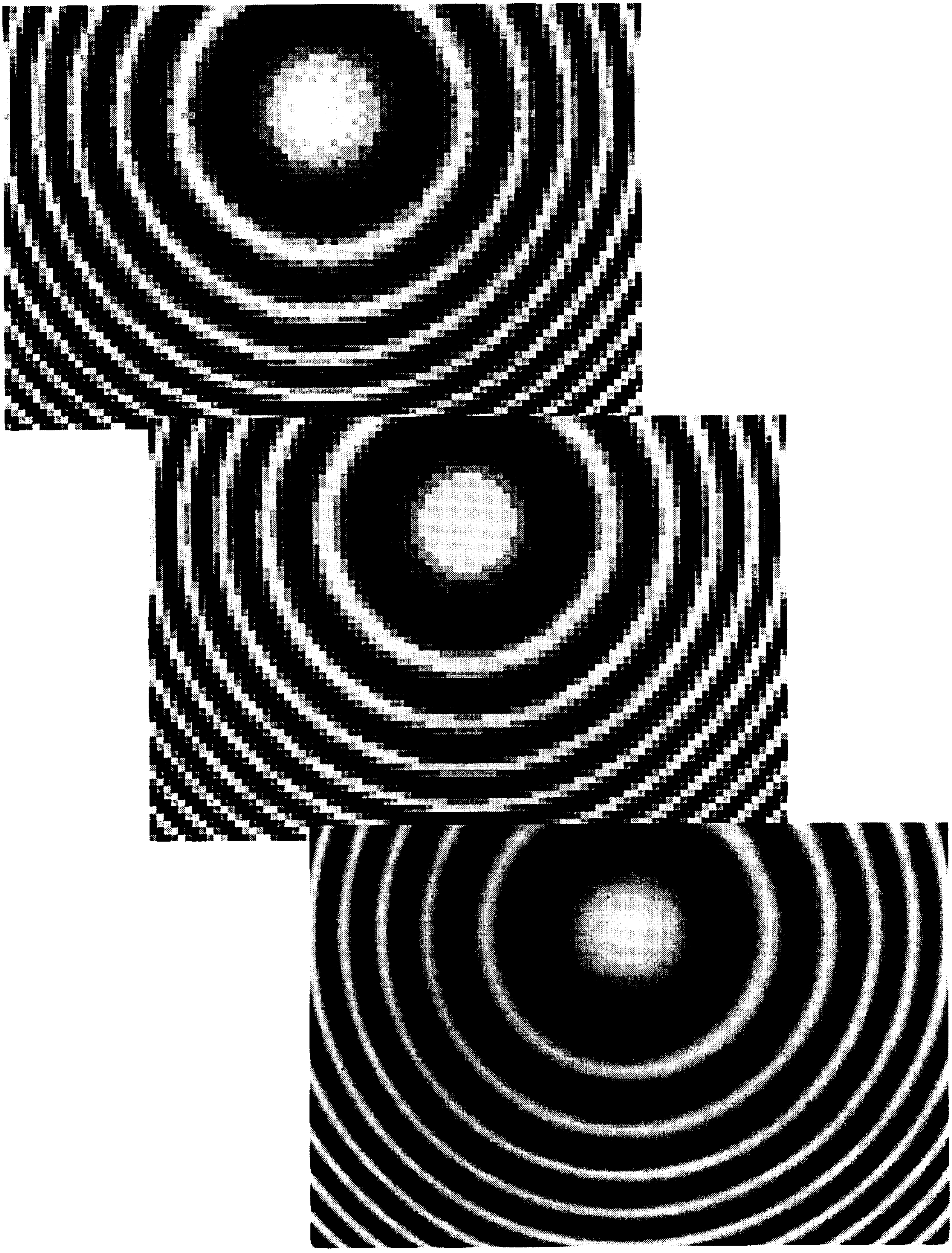
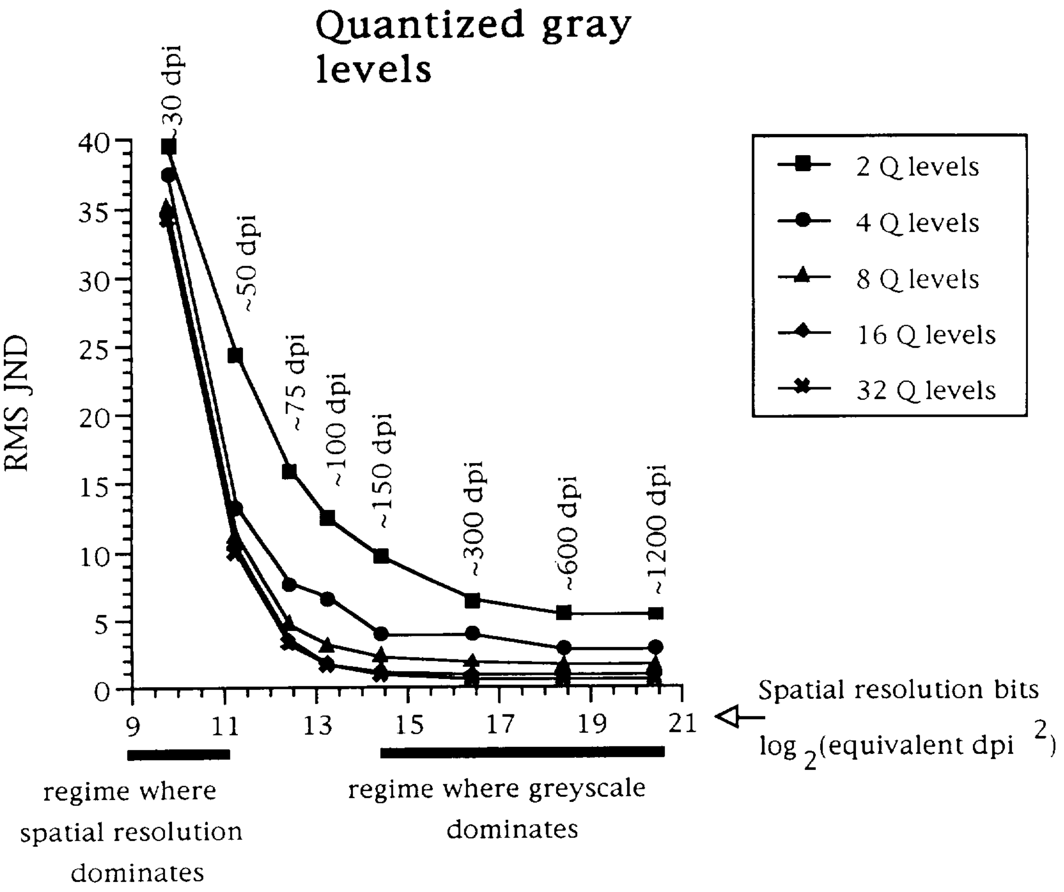
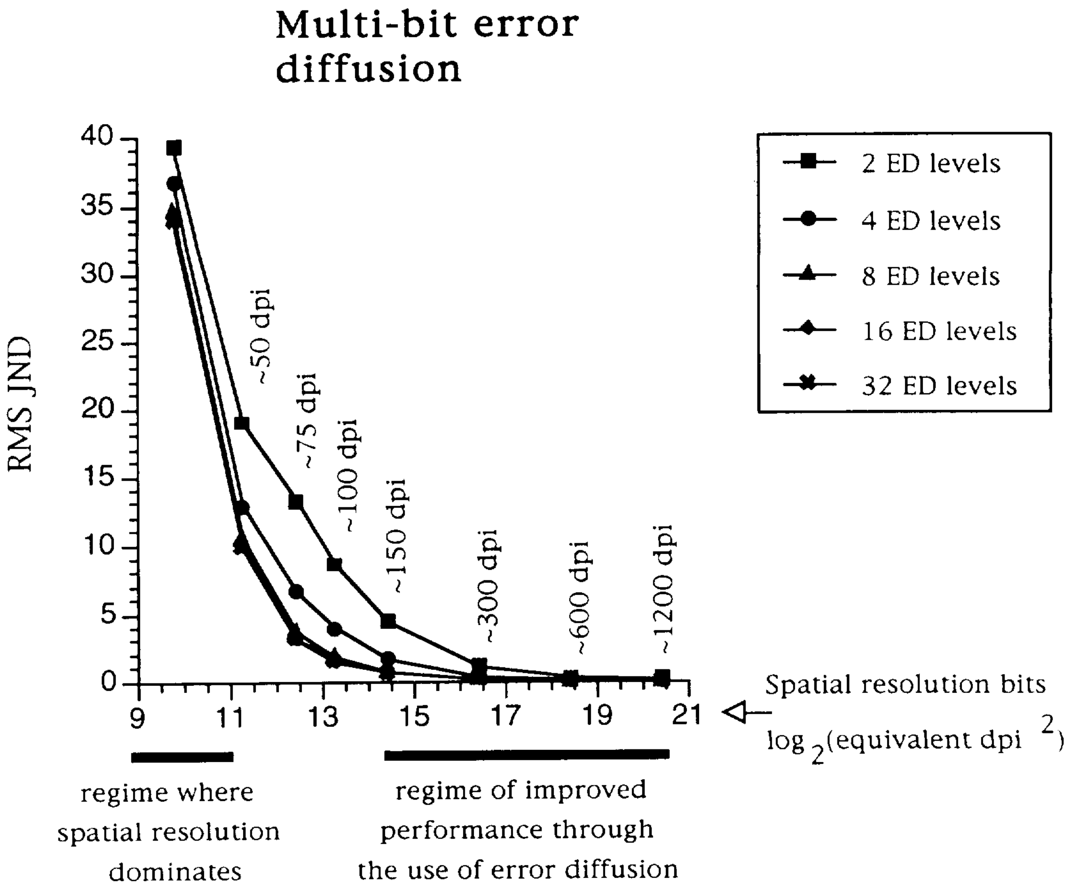


Figure 2. Predictions of the Sarnoff Human Vision Model for discriminabilities from a high-resolution standard of images varying in spatial and grayscale resolution. (a) simple quantization (b) with error diffusion

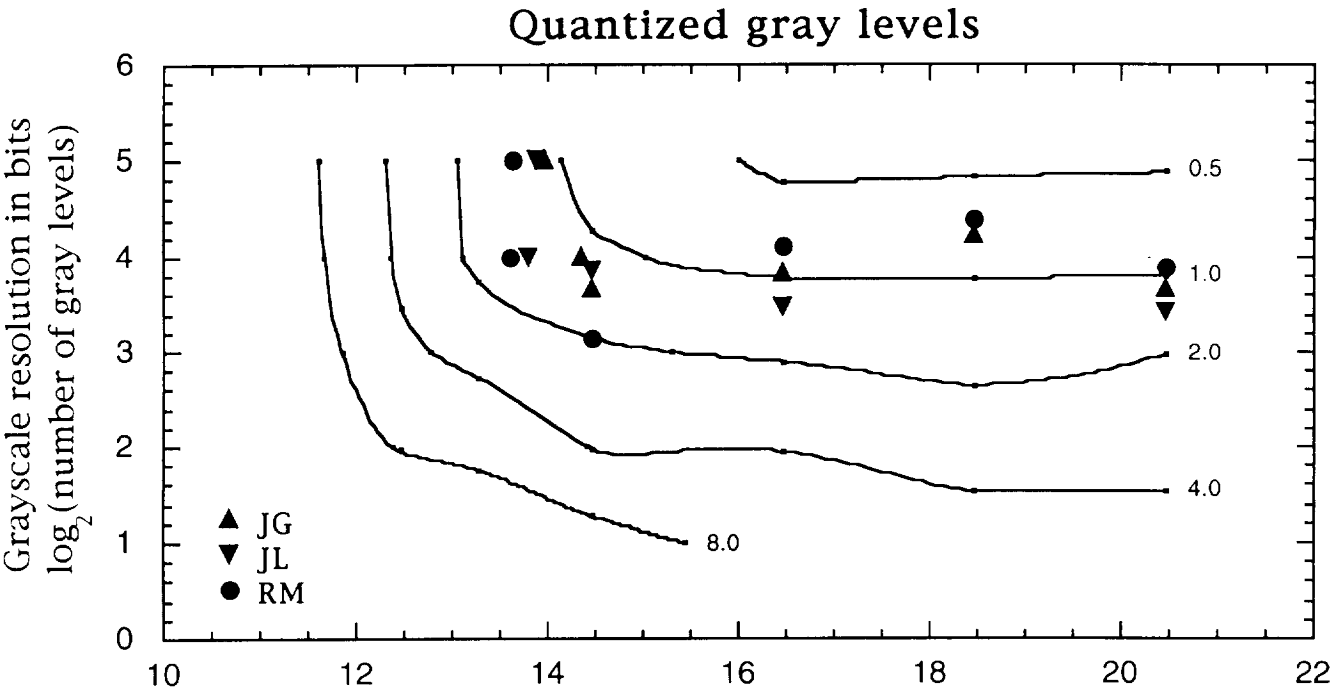


Quantized gray levels

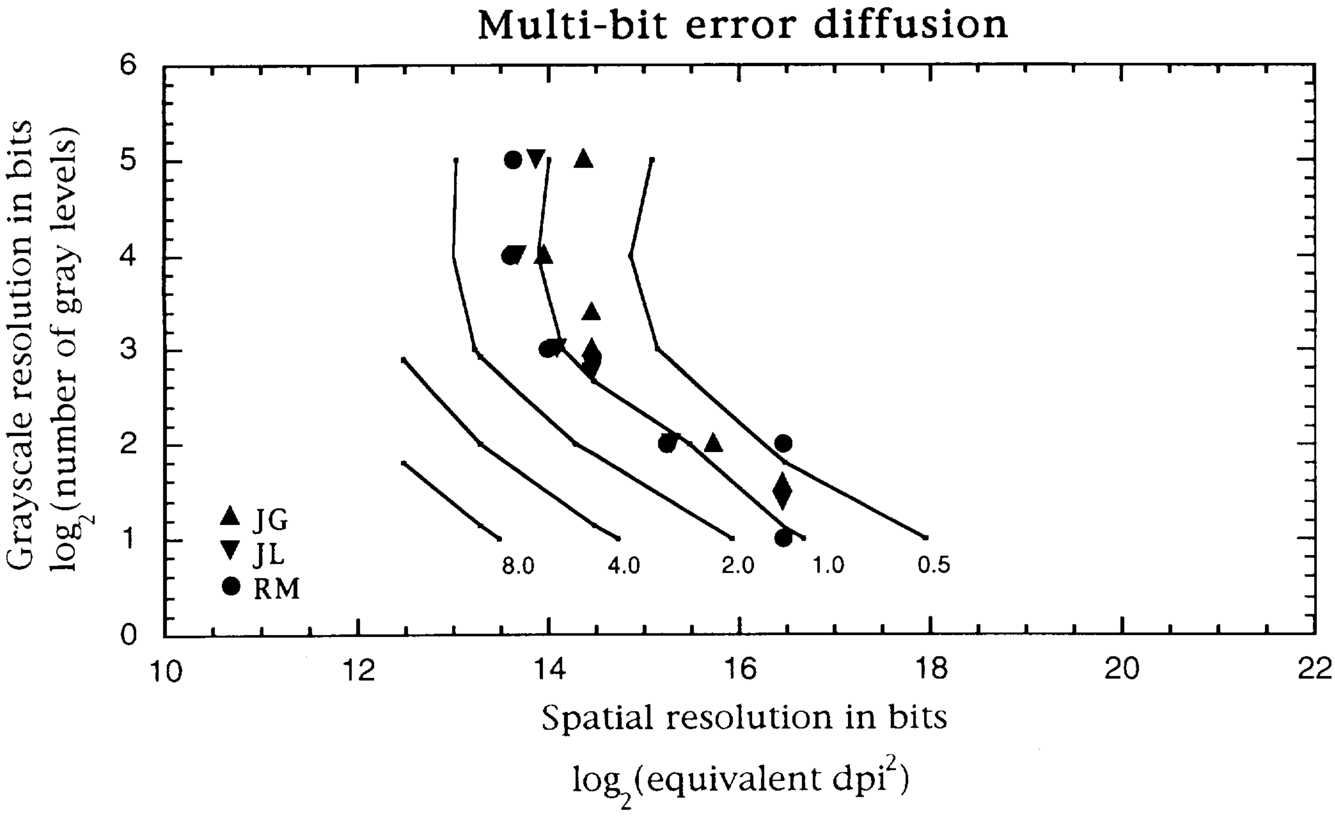


Multi-bit error diffusion

Figure 3. Discrimination thresholds for three observers compared to isodiscrimination contours computed by the vision model. (a) simple quantization (b) with error diffusion



Quantized gray levels



Multi-bit error diffusion