5.1: Visual Performance Depends Upon Signal Resolution: Frame Rate, Dot Pitch & Bit Depth Guidelines

James Larimer[1] & Jennifer Gille[2]
[1]NASA Ames Research Center
[2]University of California Santa Cruz at NASA Ames

ABSTRACT
Digital imagery is sampled in three domains: spatial, temporal, and amplitude. Image quality depends upon signal resolution and display fidelity. Different visual tasks require different display resolutions. We review psychophysical data on text legibility and search, object search, motion and grayscale artifacts that result from resolution tradeoffs: frame rate, bit depth, and dot pitch.

INTRODUCTION
Display performance depends on the display device's ability to reconstruct the image signal and the conditions and environment in which the display will be used. A display that will be used in controlled artificial lighting or will only be viewed at a short distance will have different requirements than a display used out of doors and at a great distance. The signal content also drives display requirements. A segmented display may be fully adequate as a gas pump customer interface limited to conveying fuel quantity and pricing information whereas a pixilated display may be required for advertising and natural images.

The trade-offs implicit in finding the appropriate display for an application depend upon the visual task environment, the video signal content, and the human visual system. To the eye what matters is the image of the display formed on the retina. The retinal image scales with viewing distance, so a spatially coarse pixilated display will become equivalent to a high spatial resolution display as it is viewed from increasing distance. The retinal image gets smaller with increasing viewing distance so the tradeoff here would be image size for resolution. For the eye, visual angle is the appropriate measure of resolution; the retinal photoreceptor mosaic in the region of best vision samples at a constant rate, samples per degree, and this determines the spatial details that can be detected in the image. The appropriate measure of display resolution is pixels per inch (ppi). Display resolution drives manufacturing costs and complexity. An important task parameter that must always be made explicit is viewing distance as this links display resolution to visual system resolution.

There are three domains to the video signal: space, time and intensity. This paper will review each of them and report previous findings on the resolution required for each domain when applied to visual performance.

SPATIAL
Van Nes and Bouman (1) reported the classic study of human vision's spatial resolution in 1967. They measured the ability of an observer to detect a sinusoidal grating patch as a function of the mean luminance of the grating and the amplitude modulation required to detect the grating as opposed to a spatially flat field of the same mean luminance. The spatial tuning of the eye is bandpass with its peak sensitivity at approximately 2 cycles per degree visual angle (c/deg). Alternating black and white bars one twelfth of an inch wide viewed from a distance of approximately 19.7 inches produce this fundamental frequency. On a display of 100 ppi a bar pair would be approximately 17 pixels wide. Each doubling of viewing distance would shift this fundamental frequency by one octave. Viewed from 78 inches the bar pattern would be at 8 c/deg, at 8.7 yds the frequency would be 32 c/deg, and at 17.5 yds the spatial frequency of this bar pattern would be approximately 64 c/deg. If an entire 12.5 inch diagonal XGA, 100 ppi screen were viewed from this distance, it would subtend 0.9° in visual angle or roughly equal to 0.3 inches of the same screen viewed from the normal viewing distance of 19.7 inches for a workstation screen.

The eye's spatial cut-off frequency is approximately 60 c/deg, so the bars viewed at 17.5 yds on a 100 ppi screen might just be visible if the screen can produce sufficient contrast. Most screens cannot because the first surface reflection is more than sufficient to reduce the contrast to levels well below the almost 100% contrast required to see a grating at this frequency. Sixty c/deg is close to the diffraction limit for the eye (2) and corresponds well with the 120 cone photoreceptors per degree finding for cone packing densities in the foveal region of the eye (3). The Snellen acuity 20/20 refers to the standardized ability to correctly read letters with features as small as 1 arc-minute. A 200 ppi display viewed at 19.7 inches is capable of producing letters with details or features at this scale. At this resolution the pixels themselves are at the threshold of visibility so jaggies are not visible in letters or slanted edges on the screen. At 100 ppi both jaggies and the pixel structure of the screen is visible at the standard viewing distance.

Historically, for example in Colonial America, text was printed with 6 to 8 point fonts to conserve paper and ink. Even today novels, newspapers, and many printed media are produced with fonts in the 6 to 10 point range as these are quite legible in hard copy media. Only recently with the introduction of 200 ppi screens has it been possible to render legible text at this scale. We have found in search and editing tasks that 6 point fonts on a 100 ppi screen are illegible, but fonts as small as 4 points are readily legible on a 200 ppi screen or on paper (4). Text is the image content that demands the highest spatial resolution screens, although we have also demonstrated in a search task that features in natural objects at the resolution limit of 200 ppi screen are perfectly detectable when subjects are asked to identify object features at this spatial scale (5).

An important difference between electronic media and printing is spatial phase resolution. A1200 ppi printer can begin a letter at any of 1200 addressable locations per inch and it can separate letters by any multiple of 1/1200th inch. This space flexibility is called kerning in the print industry and letter spacing differences can be readily visible on displays of 100 to 150 ppi without appropriate signal processing to hide it. A 200 ppi display on the other hand can only select from 200 locations to begin a letter or edge. Software products like Microsoft's ClearType™ (6) or anti-aliased text fonts by Adobe use grayscale and signal processing both to hide pixilation and to produce more apparent phase control on softcopy displays. ClearType and other schemes like it use subpixel addressing to increase the phase resolution in one direction at a cost of sometimes generating a chroma error that is just at or below the threshold of visibility.

Subpixel addressing and anti-aliasing are related to the hyperacuity abilities of the eye. Westheimer (7) documented a visual skill that he called hyperacuity, wherein the eye can detect a vernier offset of approximately 3 arcsec. From a naïve perspective this is a remarkable ability and almost two orders of magnitude better than the spatial Nyquist limit for the eye based upon cone sampling densities. Hyperacuity sensitivity is not based on spatial sampling alone, it depends upon the visual nervous system's ability to sense and code the intensity distribution near edges imaged onto the retina. Seeing kerning differences, or hiding pixelation by anti-aliasing are examples applying this skill or suppressing its signal on the retina through signal processing that controls the intensity profile of the image, i.e., softening edges or moving the intensity distribution center of mass slightly. We reported a study where Vernier acuities were measured on displays of 100 and 200 ppi and found that equivalent sensitivities could be measured by simply changing the intensity distribution of the cross section of a line (8). Near record hyperacuity can be measured on ordinary pixelated displays. The 200 ppi screen produced slightly better Vernier offsets than the 100 ppi display. The ability of a display to match the retinal images by using grayscale helps to explain why signal processing schemes like sub-pixel addressing and text anti-aliasing work. The primary advantage of a 200 ppi display is that less signal processing is required and smaller line widths can be used to emulate soft copy lines.

INTENSITY

A just visible increment or decrement of light against a background is called the increment threshold. Stiles (9) in a series of classic experiments measured increment thresholds for the isolated cone systems. He found that the cone photoreceptors most responsible for spatial vision have an increment threshold of approximately 1.8% when carefully measured in an cone isolation experiment. This fraction, called the Weber Fraction, is invariant over the photopic range of vision once the threshold for seeing is surpassed. This dynamic intensity range is used in all high information content display applications.

The data reported earlier by van Nes and Bouman can also be considered as increment threshold data. They found that for grating patches, once the threshold for seeing was surpassed the Weber Fraction was a constant regardless of the background level. Additionally, they found that at 2 c/deg the fraction was 1%, changing to higher fractions as the grating patch moved off of the best spatial tuning frequency. At 30 c/deg the Weber Fraction is close to 6%, or equivalently more contrast is required to see the grating at every background level. The difference between the Stiles and van Nes and Bouman finding is due to the nature of the visual tasks. A grating is detectable when any region within the grating patch reaches the threshold of visibility. The large area of the grating patch provides the spatial mechanisms of vision more opportunities or locations in which to detect the patch so this difference is due to probability summation (10).

For graycale steps to be invisible they must not exceed the Weber Fraction at the corresponding spatial frequency content in the image. The rule is that Weber's Fraction is a constant; this means that the grayscale spacing must follow a logarithmic rule. On displays with limited dynamic range, say below 80 cd/m^2, the traditional power gamma rule with a power between 1.8 and 2.2 works quite well. As screen brightness improves, however, a power law transfer function introduces readily visible intensity steps that produce banding artifacts (11). Banding artifacts generated by large grayscale steps in inappropriate spatial frequency bands are mitigated by dithering (12). We found that above approximately 120 ppi dithering applied to soft copy displays allows dot pitch to be traded off for grayscale. Both ppi and grayscale can be expensive to provide requiring more drivers or more gates per display surface area or both, so trading off one for the other can potentially suppress image artifacts that would otherwise be visible in the image. The trade off is between manufacturing and component costs versus signal processing, signal processing hardware, power budgets, and addressing bandwidth limits.

TIME

Electronic displays, especially those based on CRT technology, have traditionally used frame rates above 30 frames per second to avoid flicker. On a CRT a full field flicker signal is caused by the rapid decay of light emission from the phosphor once the e-beam raster slews past a screen location. This stimulus is optimal for producing flicker because of its low spatial frequency component (13). An equivalent phenomenon in cinema is generated by the dark period between frames required to move the film to the next frame without visible image smearing. Cinema film projectors use a dual-bladed shutter to expose each frame of the film twice before moving to the next frame. This puts the flicker frequency at 48 Hz although the signal frame rate remains at 24 Hz. Flicker becomes increasingly visible as brightness levels increase (14) so cinema projectors would require a 3 or 4 blade shutter if peak whites exceeded typical theater presentation levels in the 8 to 12 fL range.

Displays with latching pixels produce brighter peak whites and do not exhibit flicker induced by dark states between image frames. Because the pixel is latched light is continuously present making the display brighter but also making any temporal variation more apparent. Generally any increase in mean brightness also increases the eye's frequency response to temporal and spatial edges in the signal, so display artifacts are also more visible as brightness levels increase. An LCD-based display can produce a temporal variation at half the frame rate due to field polarity swapping required to prevent image sticking, but this signal is low in amplitude and therefore usually below the threshold of visibility (15).

There are two kinds of motion artifacts that occur in electronic displays and both can be reduced in visibility by increasing the update rate of the signal. The first is produced by observer motion and is called motion blur. Motion blur has two distinct phenomena associated with it: motion blur in static images is produced by sweeping the image across a sensor array, either the retina in the case of human vision or film or a focal plane array in video capture; and a second phenomena, uniquely associated with discrete time domain reconstruction, is called judder (16). Judder occurs when your eye follows a moving object in a discretely sampled and reconstructed scene.

Motion blur in static images cannot be reduced by frame rate increases, indeed it is natural and due to the fact that the eye integrates signals in time just like any sensor does to improve signal to noise performance. Human vision differentially integrates the signal employing longer time periods for higher spatial frequency content in the image and shorter integration periods for lower spatial frequency content in the spatial-temporal signal (13).

Judder can be reduced in visibility by increasing frame rate at both capture and reconstruction. When an object moves in a static background or when the camera pans to track a moving object, the reconstruction signal is a series of discrete still images. The eye tracks smoothly, so during the brief time at which the image is stationary on the screen, the eye motion results in a shearing of the screen image on the retina. This shearing motion can produce a time domain variation at edges in the scene that is correctly coded by the nervous system as flicker when the contrast of this time domain variation exceeds the threshold of visibility (16). There is currently a debate about whether or not pixel latching increases the visibility of judder-induced edge flicker, but our data have not identified any boost in visibility due to image holding when the reconstruction rate exceeds 25 to 30 Hz. This remains a question for further study.

To render fast motion accurately the application of the sampling theorem requires higher sampling and reconstruction rates. However, the eye, due to its temporal integration strategy, is a temporal low pass filter, so there are diminishing returns from increasing frame rates with respect to motion fidelity. Two methods have been applied to up-sample the video signal during reconstruction. The first, motion vector interpolation, is more expensive in terms of computational resources (17). The second, fading, requires less computational power. Both are effective for up sampling content captured at one frequency, say 24 Hz, and reconstructing it at a higher frequency, say 48 Hz.

Pulse width modulation is a time domain method used to produce grayscale. Two kinds of artifacts can be generated by this method; we will mention but not discuss them in detail. The first is an aliasing-like phenomena where a time variation between levels 127 and 128 can appear instead as a variation between 0 and 255 to an observer. Bit scheduling and bit splitting, i.e., representing a higher order bit with two addressing periods, can mitigate this. Pursuit eye movements that can “re-sample” the bit stream on the retina generating false edges cause the second, more intractable, artifact. This too can be addressed by bit scheduling, up-sampling, i.e., adding addressing periods, and dithering of the bit ordering.

REFERENCES

1. van Nes, Floris L. & Bouman, M. A., (1967). Spatial modulation transfer in the human eye. *JOSA, 57(3)*, 401-406.
2. Campbell, F. W. & Gubisch, R. W., (1966). Optical quality of the human eye. *J. Physiol., 186*, 558-578.
3. Curcio, C.A., Sloan, K.R., Kalina, R.E. & Hendrickson, A.E. (1990). Human photoreceptor topography. *J. Comparative Neurology, 292*, 497-523.
4. Gille, Jennifer, Larimer, James, Powers, Maureen, & Liu, Hsien-Chang, (2004). Very-high-resolution displays: Productivity gains in word-processing and spreadsheet tasks. *2004 SID Intl. Sym. Digest, 35(2)*, 1366-1369. Powers, Maureen, Larimer, James, Gille, Jennifer, & Liu, Hsien-Chang, (2004). Reading performance with large fonts on high-resolution displays. *Human Vision and Electronic Imaging IX*, Proc. SPIE-IS&T Electronic Imaging, SPIE Vol. 5292, 254-259.
5. Larimer, J., Gille, J., Powers, M., & Liu, H-C., (2004). Search on high-resolution screens. *USDC High Information Content Display Systems Conference*. Presented at Arlington, VA, Sept. 2004.
6. Betrisey, C., Blinn, J. F., Dresevic, B., Hill, B., Hitchcock, G., Keely, B., Mitchell, D. P., Platt, J. C., & Whitted, T., (2000). Displaced filtering for patterned displays. *SID Intl. Sym. Digest, 31*, 296-299.
7. Westheimer, G. & McKee, S. P., (1977). Spatial configurations for visual hyperacuity. *Vision Research, 17*, 941-948.
8. Larimer, J., Gille, J., Powers, M., & Liu, H-C., (2004). Hyperacuity on high-resolution & very-high-resolution displays. *Human Vision and Electronic Imaging 10*, Proc. SPIE-IS&T Electronic Imaging, SPIE Vol. 5292, 211-217.
9. Stiles, W. S. (1978). *Mechanisms of Colour Vision*. London:Academic Press.
10. Savoy, R. L. & McCann, J. J. (1975). Visibility of low-spatial-frequency sine-wave targets. *JOSA 65(3)*, 343-350.
11. Larimer, J., Gille, J., & Luszcz, J. (1997) Display resolution trade-offs: grayscale spacing. Presented at the *International Display Research Conference* (IDRC '97) , p. 262. Toronto, Canada; September, 1997. Larimer, J., Gille, J., Liu, H-C., Powers, M. (2003) Trading spatial resolution for greyscale bit depth: How many bits are enough? Presented at the *3rd Annual High Information Content Display Systems Conference of the U.S. Display Consortium*, October 2003.
12. Gille, J., Martin, R. & Larimer, J. (1994). Spatial resolution, grayscale, and error diffusion trade-offs: Impact on display system design. *Conference Record of the 1994 International Display Research Conference*, 381-385. Gille, J., Samadani, R., Martin, R., and Larimer, J. (1994). A human visual discrimination model analysis of the grayscale/resolution trade-off: Displays with 150 dpi or less resolution. *Proceedings of the Society for Information Display, 25*, 494-497.
13. Kelly, D. H., (1972). Adaptation effects on spatio-temporal sine-wave thresholds. *Vision Research, 12*, 89-101.
14. Farrell, J. E., Casson, E. J., Haynie, C. R., & Benson, B. L. (1988). Designing flicker-free video display terminals. *Displays, 9(7)*, 115-122.
15. Wright, S. L., Millman, S., Kodate, M. (1999) Measurement and Digital Compensation of Crosstalk and Photoleakage in High-Resolution TFTLCDs SPIE 3636 pp 200-211.
16. Larimer, J., Gille J., & Wong, J. (2001). Judder-induced edge flicker in moving objects. *SID Intl. Sym. Digest*, *32*, 1094-1097. Feng, C., Larimer, J., Gille, J. & Cheung, V. (2003). Judder-Induced Edge Flicker at Zero Spatial Contrast. *SID Intl. Sym. Digest*, *34*.
17. Klompenhouwer, M. A. & de Haan, G. (2004). Video, display and processing, *2004 SID Intl. Sym. Digest, 35*, 1466-1469.