ABSTRACT

Much of the guidance provided to designers of visual displays is highly simplified because of historical limitations of visual display hardware and software. Vast improvements have been made in processors, communication channel bandwidth, and display screen performance over the past 10 years; and the pace of these visual system improvements is accelerating. It is now time to undertake a critical review of the true performance capability of the human visual system (HVS). Designers can now expect to realize systems that optimize (increase) worker productivity rather than minimize the negative impact on human effectiveness of hardware artifacts like low resolution—spatial, gray level, and temporal—relative to the real world. Myths and realities of human vision are examined to show where some assumptions used by designers are not based on solid research. Some needed new human vision studies are identified. An ideal display system is described that would enable, rather than limit, full exploitation of HVS capability.

Keywords: Human visual system, acuity, color, brightness, luminance, illuminance, contrast ratio, refresh rate, display

1. INTRODUCTION

The future of displays depends upon how well it will be possible to design a display device that can display ‘real world’ information rather than merely fool the human eye. Much of the guidance provided to designers of visual displays is incomplete or drastically simplified with regard to the true optical performance of the human visual system. Often the information provided corresponds only to the performance levels at which image quality becomes objectionable, or represents just the minimum appropriate for text or normal (current), but not future, workstation tasks. More detailed information regarding operator visual characteristics would be beneficial in determining guidelines for the design of a visual display system that would meet or exceed the known capacity of the viewer. Performance parameters of human vision are examined to identify the some key characteristics of an ideal display.

As display design has progressed, rapid improvements have occurred in the image quality. Over the past 15 years, displays having a resolution of 0.3 megapixel (formatted as 640 x 480 pixels) with 8-bits per color pixel (256 colors) have been replaced by current day units that are typically 24-bit per color pixel (17 million colors) with mass production device resolutions as high as 2 megapixels (formatted as 1600 x 1200, or 1920 x 1080 pixels). Display technology development spirals, however, suffer the same fate as most other products as they move from their earliest stages toward a mature technology—a steepening slope on the cost vs. performance trade-off curve. In the early stages of a technology spiral, comparatively modest investments lead to robust performance gains. As a particular spiral matures, increases in performance require progressively more expenditure of time and money for comparatively smaller improvements in capability. Designers at the mature point within a particular product development spiral usually begin to change their focus from the question of “Can it be made better?” to “How much is good enough?” Then a disruptive change (breakthrough) in the underlying device or software technologies starts the next cycle of spiral product developments. Even after such a disruption has occurred, many research and development communities coast along assuming that current generation products are completely adequate for all possible uses and that improvements in display technology are neither necessary nor desirable. What is often missed is the fact that improved capabilities often lead to new avenues of employment.
In the case of visual display quality, an end goal would be to specify a system that was capable of matching the performance of human visual system in terms of quality of output and that would be visually indistinguishable from the real world (other than the fact that it would be two- rather than three-dimensional.) Obviously, such a system would be overkill for some applications (e.g., the performance of a text-only display should only be driven by performance limits where the assimilation of such material is maximized). Equally as obvious, the goal of matching the performance of the human visual system would provide a built-in benchmark to compare against; development of greater capability than this would represent a waste of both time and effort, since there would be no way for the observer to appreciate higher quality imagery were it to be made available. In order to determine developmental targets, it would seem to be simply a matter of determining the visual capabilities of the human operator and then constructing a display system to match. There is, however, a slight problem with this approach—the peak performance of the human visual system varies depending on a number of factors. Many performance related parameters vary with the conditions under which they are measured. Further, many reference works do not adequately define exactly what they are referring to when describing performance capabilities; actual maximum performance levels are usually not described in detail.

Often, the material presented in reference works represents an acceptable minimum performance standard rather than providing design goals for an ideal one. For example, desired display refresh rates are often expressed as “at least 60 Hz,” rather than citing an ideal level beyond which performance does not improve for a particular task. One reason for this situation is that such references are not normally based on applied vision research that defines the known limits of the performance envelope. Instead, the goal of such guidance materials is to define specifications that will provide the best performance possible when using existing or near term apparatus and the question asked is “How do we best use what we have?” rather than “What is the best we can use?” Such an approach is adequate for employment of current technology or for defining near-term improvements, but not for identifying opportunities for mid- and long-term developmental efforts.

2. OCULAR PHYSIOLOGY

At least some knowledge of ocular physiology is required to further this discussion. As illustrated in Figure 1, the eye can be thought of as a hollow sphere approximately one inch in diameter and filled with a viscous fluid according to Graham. The cornea makes up one-sixth of the outer layer of the eye and is a transparent dome that serves as a window into the interior of the eye. It is the most powerful refractive structure within the eye, providing about two-thirds of the bending of light that is required for proper focusing. The sclera is the white, opaque portion structure that makes up remaining five sixths of the outer portion of the eye, and provides an attachment point for the muscles that move and position the eye. Light entering the eye is converged first by the cornea and then the lens, and ideally is brought into focus on the retina. The macula lutea is the central portion of the retina, which provides the clearest and most distinct vision. At the very center of the macula is the fovea centralis, an area composed only of cone cells about 1.5 mm in diameter, which is the point of the eye’s greatest acuity. The foveola is the centermost area of the fovea.

![Figure 1. Anatomy of the human eye.](image-url)
The iris, a thin diaphragm lying between the cornea and the crystalline lens, controls the entry of light into the eye and functions similarly to the aperture stop and shutter in a camera. The pupil, a variable diameter opening within the iris, determines the amount of light that passes through to the retina, and is analogous to the camera's aperture. The iris varies in width to alter the size of the pupil as a function of the level of illumination falling onto the retina. A number of other factors such as the size and region of the retina being stimulated, spectral and temporal characteristics of the available light, and emotional reactions also influence the size of the pupil and iris. To a large degree, the quality of the retinal image is a function of the size of the pupil. When the pupil is very small, the retinal image quality is degraded by diffraction. When the pupil is very large, spherical and chromatic aberrations caused by the cornea and lens are most pronounced.

The lens is located immediately behind the iris and provides additional refractive power, although its primary role is in distance compensation to maintain focus on the retinal surface. The lens naturally assumes a rounded, almost globular shape, which is optimized for viewing at near distances. When the ciliary muscle relaxes, the ciliary processes pull on the suspensory ligaments which in turn pull on the equator of the lens, causing it to flatten or become less convex allowing it to focus light coming from a greater distance. When the ciliary muscle contracts, tension is released and the lens becomes more convex. The process of adjusting the shape of the lens to focus on objects at different distances is referred to as accommodation. The aqueous humor is a clear fluid lying between the cornea and the lens and is primarily responsible for maintenance of interocular pressure levels. The vitreous humor is a clear gel that fills the rear compartment of the eye which occupies about 80% of the volume of the eyeball. Light enters the eye through the cornea, passes through the pupil, aqueous humor, and lens, and then passes through the vitreous humor to the retina.

The retina is the innermost layer of the eye and is composed of nerve tissue that senses the light entering the eye. It is about 250 \( \mu \text{m} \) thick over most of its extent and contains approximately 200 million nerve cells directly associated with the processing of visual information. The retina itself is comprised of ten layers. Moving from the back of the eye forward, they are:

1. Pigmented epithelium
2. Photoreceptors (inner and outer segments of the rod and cone receptors)
3. External limiting membrane
4. Outer nuclear (cell bodies of rods and cones)
5. Outer plexiform (rod and cone axons, horizontal cell dendrites, bipolar dendrites)
6. Inner nuclear (nuclei of horizontal cells, bipolar cells, amacrine cells, and Mueller cells)
7. Inner plexiform (axons of bipolar cells and amacrine cells, dendrites of ganglion cells)
8. Ganglion cells (nuclei of ganglion cells and displaced amacrine cells)
9. Axons (nerve fibers from ganglion cells traversing the retina to leave the eye at the optic disk)
10. Internal limiting membrane (separates the retina from the vitreous humor)

The optic nerve is a continuation of the axons of the ganglion cells in the retina. The optic nerve passes back from the eye to the optic chiasm, the point where the nerve fibers from the nasal half of each retina cross over to the other side (the fibers from the temporal side do not cross). Beyond the optic chiasm, the nerve fibers become the optic tract, and pass through the thalamus to the visual cortex located in the occipital lobe at the rear of the brain.

The retina contains two differing types of photoreceptive cells—rods and cones. Each eye contains approximately 125,000,000 rod cells and 6,400,000 cone cells according to Osterberg.\(^2\) The rod cells are highly sensitive to light intensity (from \(3 \times 10^{-6}\) to about \(10^{-10}\) nt where rod saturation begins), but only offer limited acuity for fixed images and no color perception (see Boff and Lincoln\(^3\)). Cone cells provide both high acuity and color capability, but contribute progressively less to visual perception as light levels decrease (threshold for cone perception is approximately \(10^{-5}\) nt, while discomfort begins at about \(10^{-9}\) nt and damage may occur at levels above \(3.18 \times 10^{8}\) nt) according to Boff and Lincoln.\(^3\) The distribution of both rod and cone cells is non-uniform across the retinal mosaic. In the fovea, only cone cells are present, with a density of approximately 200,000/mm\(^2\) according to Westheimer.\(^4\) The critical area of the fovea has a diameter of about 300 \(\mu\text{m}\) that accommodates some 15,000 individual cone cells. The light sensitive area of each foveal cone is between 1.0 and 1.5 \(\mu\text{m}\) in diameter (they are both shorter and wider elsewhere on the retina), with a spacing of approximately of between 2.0 and 3.0 \(\mu\text{m}\) from center to center according to Pugh\(^5\) and Snyder and Miller\(^6\). Cone density is greatest within the fovea and drops off rapidly out to about \(10^{8}\) from the eye's fixation direction (visual axis, measured

\[1 \text{ nt} = 1 \text{ nit} = 1 \text{ cd m}^{-2}\]

60 Proc. of SPIE Vol. 5080
relative to origin in the lens), whereupon the cones become more or less uniformly distributed across the remainder of the retina. According to Boynton\textsuperscript{7} the density of cones in the foveal area does not imply that the remainder of the retina is a "cone free zone." Only about 4\% of the cone cells are concentrated in the fovea; the remainder distributed throughout the more peripheral regions of the retina.

Rod cells are densest in the near visual periphery, reaching a maximum at about 18° from the direction of eye fixation. At this point rod density begins to decrease, though they maintain a much higher density than cone cells out to the limits of the visual field. These rod and cone cell distribution functions are illustrated in Figure 2. It should be noted that the peripheral cone cells are much larger in cross section than peripheral rod cells. To envision the relative likelihood of a particular photon exciting a rod or a cone cell, Boynton states that the relative cone sensitivities in the peripheral retina would have to be multiplied by a factor of ten.

![Figure 2. Density distribution functions of rods and cones in human eye.](image)

Another difference between the two types of receptors (rods and cones) lies in their connections to the ganglion cells. Outputs from multiple rod cells converge onto a single ganglion cell, while information from cone cells typically undergoes much less convergence. Experimentation has shown that while it takes only a single quantum of light energy to activate a rod, several quanta striking within a relatively small region and short period of time are normally necessary to elicit a threshold visual response (10 quanta absorbed within 100 ms inside a 10 arc-minute disk on the retina according to Cornsweet\textsuperscript{9}).

A heuristic treatment of temporal response is provided in a later section. Also, the size, height, and spacing of receptors varies significantly, but the contributions these variations to visual perception are not currently well understood.

3. VISUAL ACUITY

The maximum spatial resolution for an ideal display should be based on ability of human vision to distinguish fine details. The metric for this ability is visual acuity. There are a variety of different ways of defining "acuity," each of which leads to a different value for the maximum performance of the human eye. Reference materials often describe just a single type of acuity or provide a value without specifying the type of "visual acuity." The result is that many researchers and technologists come to the erroneous conclusion that only one measure of visual acuity exists, or to an erroneous expectation regarding the level of usable visual acuity that a display should support to avoid operator performance degradation.
3.1. Theoretical Acuity

3.1.1. Geometric Optics. Using the principles of geometric optics, the minimum angular separation between two points that can be resolved can be calculated. The formula for characterizing angular resolving power ($\Delta \theta_{\text{min}}$) for an optical device imaging two objects at an infinite distance away is

$$\Delta \theta_{\text{min}} = 1.22\lambda / D,$$

where $\Delta \theta$ is in radian angular measure, $D$ is the diameter of the entrance pupil of the device and $\lambda$ is the wavelength of the light. This means that for a human eye with a pupil diameter of 2.5 mm and light with a wavelength of 555 nm (the wavelength at which photopic, or cone-mediated vision is most sensitive), the maximum resolution would be expressed by the relationship,

$$\Delta \theta_{\text{min}} = \frac{1.22 \times 555 \times 10^{-9} \text{ m}}{2.5 \times 10^{-3} \text{ m}} = 2.684 \times 10^{-4} \text{ radians},$$

or about 0.92 arc-minutes. When tested, the highest sinusoidal grating that can be resolved in normal viewing lies between 50 and 60 cy/deg at a pupil diameter of about 2.5 mm. Theoretically, the highest angular frequency that could be passed by a diffraction-limited system at a wavelength of 555 nm with a pupil of this size is approximately 78 cy/deg (0.4 arc-minutes). The actual cutoff in the human eye, however, is limited by both optical aberrations and scatter according to Atkinson et al. 10

3.1.2. The Nyquist Limit. The theoretical resolving power of the eye, or Nyquist limit, represents the finest separation of a series of evenly spaced parallel lines that an individual should be able to differentiate as stripes rather than as a uniform gray field. This value is based upon the spacing between the centers of the cone cells of the foveola. The formula for calculating this value is:

$$v_N = \frac{f}{d_{ec} \sqrt{3} (57.3)}$$

where $v_N$ is given in cy/deg, $f = 16.7$ mm is the posterior nodal distance, $d_{ec}$ is the cone spacing in millimeters, 57.3 is the factor for converting from radian measure to degrees, and $\sqrt{3}$ is a factor from the two-dimensional sampling theorem for a hexagonal array. The value obtained for $v_N$ from this computation assuming a 3 $\mu$m cone spacing is about 56 cy/deg or about 0.54 arc-minutes. Computations involving a 2 $\mu$m cone spacing would result in a Nyquist Limit of about 84 cy/deg or about 0.36 arc-minutes according to Atkinson et al. 8

3.1.3. Point-Spread Function of the Eye. A number of physical and geometric factors prevent a point source of light from being imaged on the retina as a point. When such a source is imaged, the result is actually a bell-shaped distribution that follows a curve with the equation:

$$Q(\rho) = 0.952 \exp(-2.59|\rho|1.36) + 0.048 \exp(-2.43|\rho|1.74)$$

where $\rho$ is the distance from the center of the geometric image in arc-minutes of visual angle and the $Q(\rho)$ is the intensity of the light in the image at distance $\rho$ in relation to its maximum intensity. A number of factors contribute to this spread, chief among which is diffraction. At pupil diameters below ~2 to 2.5 mm in diameter, the eye acts as a diffraction limited optical device, and the width of the point spread function changes inversely with pupil size as shown in the function below:

$$I\rho = \left[ \frac{J_1(\pi a \rho / \lambda)}{\pi a \rho / \lambda} \right]^2$$

where $I$ is the intensity of the light in the image, $\rho$ is the radial distance from the geometrical point image in radians of visual angle, $a$ is the entrance pupil diameter measured in linear units such as meters, $\lambda$ is the wavelength of light expressed in the same units as the entrance pupil diameter, and $J_1$ is a first-order Bessel function. 8 At pupil diameters in excess of about 5.0 mm, the point spread function widens due to aberrations of the eye according to Boff and Lincoln. 3

3.1.4. Line Spread Function of the Eye. A single, sharp, narrow line is not imaged on the retina as a sharp line; the actual distribution of the light on the retina is a bell-shaped in cross-section due to the effects of diffraction. The distribution on the retina can be approximated by envisioning the line as a series of closely spaced point sources; the image of the line therefore would be a superimposition of the series of images of the individual points of which it is comprised. For a radially symmetrical point-spread function, the corresponding line-spread function would be:

$$A(\alpha) = \frac{1}{2} \int_0^\infty s(\rho)(\rho^2 - \alpha^2)^{-1/2} \rho d\rho$$
where $\alpha$ is the distance from the geometrical image of the line in a direction normal to the line, and $\rho$ is the radial distance from the center of the geometrical point image, with both being expressed in arc-minutes of visual angle.

3.2. Measured Acuities

3.2.1. Snellen Acuity. The standard eye test found in most optometrist office’s (the chart with the large “E” at the top) was originally developed by a Dutch ophthalmologist in 1862, and consists of several lines of high contrast text (black numbers and letters against a white background), with each line of characters presented in order of descending size. Using this test, acuity is based upon the smallest text size that that an individual is able to correctly read at a specified distance, divided by the distance at which a person with “normal” visual acuity would be able to read them. In the U.S., the test is commonly performed at a distance of 20 ft under 10 to 20 fc of illumination, though charts for other distances have been developed by changing text sizes as appropriate. In practice, “20/20” indicates that the individual being tested can correctly read letters that are 5 arc-minutes in height at a distance of twenty feet. For the letter “E”, assuming that the stroke width of the horizontal components of the letter and the spaces between them were equal, this would lead to a value of 1 arc-minute for each horizontal component of the figure. Under optimum viewing conditions with younger subjects, an acuity of 0.8 arc-minute is common according to Boff and Lincoln. 3

Similar results are derived using Landolt-C tests, wherein the individual is asked to determine the orientation of the gap in the letter C. Correct determination of letter orientation when the gap subtends an arc of 1 arc-minute of visual angle is taken to be equivalent of “20/20” vision. In practice, acuity levels measured with Landolt-C tests tend to be better than for Snellen letters (about 1.3 x Snellen acuity).

3.2.2. Contrast Sensitivity. One of the problems with standard measures of assessing visual acuity is that they often reflect conditions that are found only in laboratory settings (e.g., high contrast targets that are well illuminated), rather than those found in the real world. Contrast sensitivity testing is a method for determining the visual system’s capability to detect objects across a variety of spatial frequency and contrast conditions. Contrast sensitivity testing is performed by showing the observer a series of sine wave gratings that vary in orientation. The maximum contrast sensitivity for a given spatial frequency is defined as the minimum contrast required to distinguish between a sine wave grating and a uniform background of the same average luminance. There are two elements to contrast sensitivity: spatial frequency and contrast sensitivity (contrast/illumination).

3.2.2.1. Spatial Frequency. A “cycle” is defined to be one dark and one light band of the sine wave form. The spatial frequency is the number of cycles subtending one arc-degree (cy/deg) at the center of the observer's pupil. A low spatial frequency comprises wide bands, while a high spatial frequency, narrow bands. The sensitivity of the human eye peaks at mid-spatial frequencies (2 to 5 cy/deg) and is less sensitive at higher and lower spatial frequencies. 3 Under optimum viewing conditions, maximum visual acuity using sine wave gratings (detection of orientation or separation) is on the order of 60 cy/deg (30 arc-seconds, or 20/10) according to Boff and Lincoln. 3

3.2.2.2. Contrast Sensitivity. Contrast sensitivity is a function of both the display contrast and the ambient illumination. For a person with normal vision relatively little contrast is required to detect the light and dark bands of a sine wave grating. The contrast threshold is the minimum contrast required for pattern detection. The reciprocal of the contrast threshold (the minimum amount of contrast needed to see the pattern) is called contrast sensitivity. Contrast thresholds for various spatial frequencies can be measured and plotted as contrast sensitivity vs. spatial frequency. Such a graph is called a contrast sensitivity function. Contrast sensitivity at low spatial frequencies relates how well large objects and shapes are perceived. Contrast sensitivity at higher spatial frequencies is related to the ability to see lines, edges, and fine detail.

3.2.3. Minimum separable acuity. Another “visual acuity” often cited in reference literature is what is technically referred to as “minimum separable acuity”, and is based on the smallest separation at which two parallel lines can be discriminated from one another. This value is highly influenced by both contrast and the luminance level of both the target and background. Under laboratory conditions involving close to 100% contrast and 100 fl ($\approx 350$ nt) luminance, acuity values as high as 0.5 arc-minutes are derived for normal subjects. A contrast level of 50% and luminance of 10 fl ($\approx 35$ nt) produces a value of approximately 1.0 arc-minute. This one arc-minute value is the most commonly cited value for the “visual acuity” of the human eye.

3.2.4. Hyperacuity. Minimum separable acuity does not represent the highest level of performance of which the visual system is capable. In practice, it is actually possible for the eye to determine the position of an object with a precision that corresponds to only a fraction of the diameter of a photoreceptor according to Westheimer. 3 This phenomenon is normally
referred to as hyperacuity. It was at first thought that hyperacuity would violate the constraints for signal reconstruction and information transmission given by the Nyquist limit. However, hyperacuity tasks do not involve the resolution of a pattern, but rather are concerned with the detection of the position of an object. In other words, the critical issue involves the phase of the spatial frequencies constituting an object which itself can easily be resolved. This phase reconstruction is accomplished when the light levels in neighboring receptors are not completely uncorrelated with one another. Any spatial frequency of a signal that is lower than half the sampling frequency will be sampled unambiguously, and it is only the signal-to-noise ratio in the sampling channels that limits the precision with which the phase can be determined. Hyperacuity thus involves a population code, with the phase information being spread over a number of channels.

3.2.4.1. Vernier acuity. "Vernier acuity" represents the minimum misalignment that can be detected in the colinearity of two lines. Figure 3 demonstrates the underlying principle behind this type of acuity. The array of rectangles at the base of the figure represents the photoreceptors in an area on the retina. Assume that a single light spot (i.e., the left circle at the top of the figure) gives rise to a bell-shaped intensity distribution on the retina (solid curve). If the position of a second spot is slightly different (right circle at the top of the figure), the intensity distribution will be shifted slightly (dashed curve) by an amount that corresponds to half a receptor diameter. The photoreceptor excitations caused by the light distributions are illustrated by the vertical bars. For the receptor cells at the edges of the two curves, the shift can make significant difference in perceived light intensity even though the displacement between the two spots is very small. If the signal-to-noise ratio in the receptors is high enough to discriminate these intensity level differences, it is possible for one to easily detect spot position differences that are just a fraction of a receptor diameter. Resolution for vernier acuity is actually far higher than that which can be accounted for by cone cell spacing in the fovea, and falls into the 1.0 arc-second range under ideal conditions according to Klein and Levi.11

3.2.4.2. Minimum perceptible acuity. Vernier acuity does not represent the maximum acuity of the human eye. "Minimum perceptible acuity" is the ability of the eye to detect the presence of a single dark line several degrees of arc in length against a bright background (e.g., a human hair on a sheet of white paper). The critical dimension for this measure seems to be the width of the line itself. Normal values for minimum perceptible acuity range from about 3 arc-seconds at 0.10 nt to about 0.5 arc-second at 100 nt.

3.3. Required Display Resolution
As could be expected from the many ways for defining "visual acuity" discussed above, the determination of a single figure-of-merit for the resolution required to match the maximum performance of the human eye is difficult. If a display were to design to match the 1.0 arc-minute value, 20/20 Snellen acuity, a monitor placed at a typical arm's length viewing distance of 20 in. from the eyes should have individual pixels no more than 148 μm (0.00582 in.) across. This corresponds to a linear pixel density of 172 pixels/in. A 17-in display would thus have a resolution of 4.1 megapixels (format: 2335 x 1750).

Figure 3. Hyperacuity illustration; area difference automatically determined in vision system.

---

***The monitor image at 20 in. originates in the near field while the Snellen chart image is commonly at 20 ft. (far field).
A desired resolution of 0.5 arc-minute/pixel (matching minimum separable acuity under optimum conditions) would result in a pixel size of 74 \( \mu \text{m} \) (0.0029 in) leading to a target display resolution format of 4670 x 3500 or 345 pixels/in. This would employ over 16 million pixels for a 17-in. display. For display quality to match the maximum vernier acuity of the eye (1.0 arc-second), the same display would require a pixel size of 2.5 \( \mu \text{m} \) (0.000097 in.) or 10,300 pixels/in. for a total of 14.7 billion pixels for the entire display. Matching minimum perceptible acuity (0.5 arc-second) would correspond to a pixel diameter of 1.2 \( \mu \text{m} \) (0.000048 in) or almost 20,600 pixels/in. at a viewing distance of 20-in. To accommodate minimum perceptible acuity, the number of pixels on a 17-in. desktop display would have to be increased by a factor of more than 14,000 over that for the commonly cited "one arc-minute/pixel" value. For a display that would cover the entire 210 x 135° binocular field of view, nearly 1.5 trillion pixels would be required.

Typical computer workstation monitors today have a maximum resolution of about 2 megapixels (format: 1600 x 1200 pixels) with individual pixel sizes on the order of 250 \( \mu \text{m} \); each pixel in such cases subtends 1.69 arc-minutes of visual angle at 20 inches. Some CAD/CAM workstations with higher resolution (format: 1920 x 1600 pixels) and smaller pixel sizes (210 \( \mu \text{m} \)) are available. These more closely approach this target specification at relatively reasonable costs. The best currently available display provides a total of 9.2 million individually addressable pixels in a 22.2-in display with 16:10 aspect ratio (pixel size 124 \( \mu \text{m} \)) according to Hopper et al.\textsuperscript{12}

Research into determining the potential benefits of increased resolution that matches the more extreme capabilities of the human eye is lacking. Obviously, the use of such extreme displays for text-based applications would be overkill. It is rarely necessary to read text at very minute sizes and speed and accuracy both suffer at such presentation sizes. However, other activities such as viewing medical or photoreconnaissance imagery may benefit from improved quality in terms of increased ease and accuracy of interpretation or decreased time to detect elements of interest within the displayed image.

4. TEMPORAL FACTORS

Display refresh rate is another factor that is often misleadingly specified. Most guidelines are based on the critical flicker frequency (CFF). A light that is temporally modulated (flickers) below the CFF rate will be seen as flickering while no flicker will be seen in one that is modulated above the CFF. The CFF value usually also represents the minimum frequency needed to provide the illusion of constant luminance for a direct-write display, such as a cathode ray tube (CRT). For 95% of the population, the CFF of a CRT falls between 60 and 72 Hz. The CFF for the fovea is modeled by the Ferry-Porter Law expressed as the equation

\[
\text{CFF} = a \log L + b
\]

where a and b are constants and L is the luminance of a stimulus flickering at a constant frequency. In practice, the determination of the CFF is not quite this straightforward. There are a number of variables that may influence the perception of flicker in addition to display refresh rate, such as display luminance, phosphor persistence, line jitter, screen size, and temporal contrast. New active matrix technologies such as active matrix liquid crystal displays (AMLCD) present full images to the eye continuously, making the very concept of flicker irrelevant.

In practice, the use of CRT display refresh rates that significantly exceed the CFF seems to benefit foveal information processing. A number of studies have been performed comparing the reading of text from CRT video display terminals (VDTs) versus hard copy. The studies reported either worse performance with VDTs or no significant differences between the two. Muter, Latremouille, and Beam\textsuperscript{14} reported a drop in reading speed of about 30% when using VDTs. Gould and Grischkowsky\textsuperscript{14} reported a drop in proofreading speed of between 20 and 30%. One theory for the loss of reading speed on a VDT is attributed to the interference of space constancy during saccadic eye movements. Space constancy refers to the perception that the world remains completely stationary during saccades. At lower CRT refresh rates, the visual system appears to misjudge the extent of target displacement during the saccades (see Macknik, Fisher, and Bridgeman\textsuperscript{15} and Wilkins\textsuperscript{16}). Targets then appear to be jumping in a direction opposite to that of the saccade. This results in a requirement for additional corrective sub-movements by the eye following each saccade, thus slowing the reading process.

An alternative, simpler hypothesis is based on the assumption that saccadic suppression prevents the eye from collecting information during the actual movement itself and that for a scanned image display there are short periods when no information may immediately be available at the point where the saccade ends (due to phosphor decay in CRT-type displays). A 60 Hz refresh rate results in each pixel experiencing approximately 16.67 ms of dead time (time while the
image is progressively degrading) between screen rewrites. On average then, since saccades are not correlated with screen rewrites, there is an average delay of 8.33 ms following the end of each saccade before information again is refreshed for the reader. Higher refresh rates result in commensurately shorter dead times. Montegut, Bridgeman, and Sykes\(^\text{17}\) conducted studies involving reading text from a CRT-based VDT at screen refresh rates of 60 and 500 Hz at constant luminance. Subjects said the text appeared identical under both conditions, but reading speeds for text on the screen exhibited a 3% increase at the higher refresh rate.

A further area in which imagery update rate may come into play involves extremely short duration stimuli. The eye is capable of detecting briefly presented bright stimuli against a dark background with durations as short as 1 ms. Some tasks may require the detection of stimuli presented with such short durations, even though most displayed imagery is more or less static in nature (or at least updates at a much slower rate that those used in the cited studies).

The central and peripheral retina vary in their temporal responses. Screen size may also be of considerable importance with respect to the perception of CRT flicker, principally in cases involving lack of a uniform luminance level across the display. At a given viewing distance, an increase in display size means that a greater percentage of the displayed image will fall onto the peripheral retina. The sensitivity to a change in amplitude of the light varies between foveal and peripheral vision. The fovea is less sensitive to variations in light level, so relatively large changes in luminance must occur before an image on the fovea appears to flicker. In the periphery, smaller changes in luminance may be detectable. An example of this phenomenon occurs when a variation in room illumination is detected in the "corner of the eye". When viewed directly with the less sensitive fovea, the same change may be undetectable. Thus, from a display design standpoint, luminance variations across the display may become critical as the size of the display increases.

### 5. LUMINANCE

The range of light intensity (luminance range) over which the human eye is functional covers about 11 orders of magnitude from an absolute threshold of \(3.18 \times 10^{-6}\) nt to the upper tolerance level somewhat higher than \(3 \times 10^{5}\) nt according to Boff and Lincoln\(^3\). As illustrated in Figure 4, the range \(10^{-6}\) to \(10^{-2}\) nt is termed "scotopic" vision and only the rod cells are active. The range from about \(10^{-2}\) to \(3\) cd/m\(^2\) is termed "mesopic" vision; both rods and cones are active in this range with the cones gradually increasing in dominance as luminance increases. The range between 3 and \(10^{6}\) nt is termed "photopic" vision; cones dominate the perceptual process throughout this range, with rods reaching their response limit at light levels in excess of \(10^{5}\) nt. Color vision begins to occur within the mesopic region as the cones become active. Peaks in both color perception and visual acuity occur in the photopic region. Acuity approaches an asymptote at approximately 1000 nt, though levels as high as 10,000 to 20,000 nt are routinely used for very specialized tasks involving objects of small size and low contrast according to Boff and Lincoln\(^3\) and Sanders and McCormick\(^\text{18}\). Visual performance deteriorates above \(10^{5}\) nt, discomfort occurs at levels above \(10^{6}\), and retinal damage can be experienced at levels above \(10^{8}\) nt according to Boff and Lincoln\(^3\).

A display that could cover all luminance levels over which the eye can function would span the luminance range between \(10^{-8}\) and \(10^{6}\) nt. Ideally, such a display could be designed to match the performance characteristics of the human eye, displaying monochromatic imagery in the luminance range where only the rod cells were active (from \(10^{-8}\) to \(10^{-3}\) nt) and only displaying color imagery when luminance levels exceeded this range.

![Figure 4. Luminance sensitivity ranges for elements of human vision system.\(^{12}\)](image-url)
6. COLOR AND GRAY SCALE

The total number of discernable colors is unknown. Hunt and Judd have estimated the number to be in the range of 10 million, while most computer graphics engines now provide at least 24-bit color capability. Such a computer display system is capable of generating visual images displaying over 16 million colors, or theoretically more than the human vision system is capable of perceiving—except for the pesky detail that these 16 million colors are crammed into just over half of the color space perceivable by human vision. Young examined primate color vision from the perspective of creating a low-cost opponent-color LCD for automobile dashboards; his objective is typical of many commercial applications that focus on retaining a wide range of perceived color rather than examining the limits of human vision, as is the focus of the present paper.

Psychophysical scaling experiments have shown that the human visual system can perceive 500 to 600 just-noticeable difference levels in brightness according to Hurvich and Jameson. A number of studies have been performed to investigate the effects of differing levels of gray that would be ideal for a visual display. All of these have recommended considerably fewer luminance levels than the 500 to 600 that can be discriminated. Anthony and Farrell conducted an experiment using a display apparatus capable of displaying pictorial images at 1200 dpi and exposed subjects to imagery with 2, 4, 8, 12, 16, 24, 32, and 256 levels of gray and required them to decide which of a pair of images they preferred. Their results indicated that subjects were unable to perceptually differentiate between the 32 and 256 level images. In a similar paradigm, Shen and Wang found that 55 gray levels per pixel were sufficient to represent an image under normal viewing conditions, and that 64 levels were sufficient under any viewing condition. It is important to note, however, that both of these experiments were based on a simple preference paradigm (i.e., subjects were simply asked to selected the higher quality image). It is possible that use of real-world tasks involving greater attention to minute detail within the imagery (such as photo-reconnaissance imagery analysis or the evaluation of medical X-ray images) may have produced considerably different results.

Today's display technology can produce more brightness and color scale levels than can be discriminated, this does not, however, necessarily mean that a display can fully represent the range of stimulus values that can be perceived. As stated earlier, the visual system functions over an 11 log unit range. A typical desktop visual display has a luminance range from about 10 ft-L to 100 ft-L (a 1 log unit range), while those used in a high performance bubble canopy fighter ranges from .01 to 250 ft-L (slightly over a 4 log unit range). With respect to color, the range that can be produced by most VDTs is severely limited compared to the sensitivity of the human visual system. The gamut of colors that can be displayed on most monitors includes only 50-60% of that which lies within the human visual envelope.

7. VIEWING AXIS

Performance difference between display types based on viewing axis may also be a consideration. Images presented on CRT and LCD displays vary greatly in this respect. Pixel definition with LCDs is much higher with than on CRTs, since CRTs tend to produce a blurred Gaussian distribution of light at each pixel while LCDs produce a much sharper edge. Such sharp edges do result in some advantages, but may result in degraded performance due to aliasing with curved lines.

In operational environments, viewing angle may be an important consideration. Angular luminance and chromaticity both suffer with off-axis viewing on LCD displays according to Selhuber and Parker. In off-axis viewing on LCDs, the light

---

An alternative approach to the analysis of color and grayscale argues that over a large range, the increment of light required to provoke a threshold increment of perception is proportional to the background light from which the change in intensity is measured. This observation, known as Weber's Law, is written mathematically as $k = \Delta I / I$ where $k$ is a constant, $I$ is the background light, and $\Delta I$ is the increment of light on the background according to Hood and Finkelstein. The value of $k$, known as the Weber Fraction, is about 1%.

The exact number is open to debate, and results largely depend on experimental conditions; tested values range from as low as 0.5% up to 2%. In practice, Weber's Law brakes down, both at high and at low extremes of intensity. At high levels of background illumination, the slope increases beyond unity and the visual system is said to be saturated. If one were to assume that the 1% Weber fraction for luminance discrimination were to hold true for the color visual range of 1 to 1,000,000 trolands (the retinal illuminance produced by a distal stimulus of 1 nt for a clear eye pupil area of 1 mm²), one could deduce that a minimum of 1,389 discrete luminance steps would be required to fully sample the entire photopic range (1.01 $= 1.005,514$). Use of a 2% Weber fraction would result in a value of 698 discrete steps to span the photopic range (1.02$^{698} = 1,006,746$ for the 2% fraction). Such an argument assumes that the Weber fraction would remain a constant fraction across the entire range under consideration.

---

Proc. of SPIE Vol. 5080 67
from each pixel takes a longer path through the liquid crystal medium and undergoes a greater twist than when viewed from
directly ahead. The result of this is a drop in luminance for “on” pixels and increases for “off” pixels. While off-axis problems with luminance and color distortion are also present in CRT type displays, they are commonly less pronounced.

In a recent study, Hollands et al.\textsuperscript{27} showed that at off-axis viewing angles of 60°, subjects showed a reduced sensitivity for red and blue symbols on an LCD compared to on-axis viewing of a LCD and to either on- or off-axis viewing of a CRT. Both detection latency and absolute detectability suffered in the off-axis LCD condition for the colored symbols, though white targets suffered no decrement. Although the 60° offset used in their study is a relatively extreme viewing angle, with current increases in the size of commonly available displays it is possible that some operators may find themselves in a position where this effect becomes critical. Since the Hollands et al. study, LCD display viewing angle characteristics have improved dramatically; critical applications now have devices designed with no viewing problems at 60° and higher angles.

### 8. SUMMARY

A summary of comparisons made throughout this paper for various acuities and their corresponding typical pixel size and implied display resolution, actual or theoretical, is provided in Table I. For displays that mimic real world visual scenes one needs data regarding maximum, not minimum, human visual performance characteristics. The present review and analysis indicates that, ideally, such a visual display system should be capable of the following performance:

1. **Acuity:** 0.5 arc-second (2 μrad) of visual angle. For a 17-in. display at a typical, 20-in. viewing distance, this would correspond to an individual pixel pitch (diameter) of 1.2 μm;
2. **Video update rate:** between 500 and 1000 Hz for CRTs and other direct-write display technologies that build images by scanning a single spot; values for active matrix technologies (flat panel projection), are much lower;
3. **Number of luminance levels:** at least 500 to 600 (under some arguments, this might be as high as 1400);
4. **Luminance range:** $10^{-6}$ to $10^{6}$ nt. This range in itself does not correspond to the total range of the human visual system, but is constrained at the high end to prevent discomfort or danger of retinal damage to the viewer.

#### Table I. Comparison of display pixel sizes and image resolution for several metrics for spatial visual acuity.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value (arc-seconds)</th>
<th>Pixel Pitch (μm)</th>
<th>Resolution (megapixels)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workstations</td>
<td>101.4</td>
<td>250</td>
<td>1.43</td>
<td>17-in, 4:3 aspect (345 x 259 mm)</td>
</tr>
<tr>
<td>(circa 2002)</td>
<td>87.6</td>
<td>216</td>
<td>1.92</td>
<td></td>
</tr>
<tr>
<td>Snellen 20/20</td>
<td>60.0 @ 85 nt</td>
<td>148</td>
<td>4.08</td>
<td>17-in, 4:3 aspect (345 x 259 mm)</td>
</tr>
<tr>
<td>Best Available</td>
<td>50.5 @ 100 nt</td>
<td>124</td>
<td>9.216</td>
<td>22.2-in., 16:10 aspect</td>
</tr>
<tr>
<td>Minimum</td>
<td>60.0 @ 1 nt</td>
<td>74</td>
<td>16.316</td>
<td>17-in, 4:3 aspect (345 x 259 mm)</td>
</tr>
<tr>
<td>Separable</td>
<td>30.0 @ 100 nt</td>
<td>74</td>
<td>16.316</td>
<td>17-in, 4:3 aspect (345 x 259 mm)</td>
</tr>
<tr>
<td>Vernier</td>
<td>1.0</td>
<td>2.5</td>
<td>15,000</td>
<td>17-in, 4:3 aspect (345 x 259 mm)</td>
</tr>
<tr>
<td>Minimum</td>
<td>3.0 @ 0.1 nt</td>
<td>1.2</td>
<td>58,000</td>
<td>17-in, 4:3 aspect (345 x 259 mm)</td>
</tr>
<tr>
<td>Perceptible</td>
<td>0.5 @ 100 nt</td>
<td>1.2</td>
<td>1,500,000</td>
<td>210 x 135° human binocular field-of-view</td>
</tr>
</tbody>
</table>

68 Proc. of SPIE Vol. 5080
9. ACKNOWLEDGEMENTS

We thank Drs. David Post and Paul Havig for reading the initial manuscript and providing comments.

10. REFERENCES