

## Chapter 3

# Visual Psychophysics

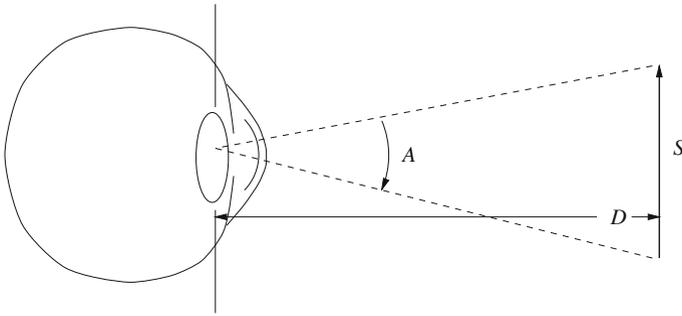
Given the underlying physiological substrate of the human visual system, measurable performance parameters often (but not always!) fall within ranges predicted by the limitations of the neurological substrate. Visual performance parameters, such as visual acuity, are often measured following established experimental paradigms, generally derived in the field of psychophysics (e.g., Receiver Operating Characteristics, or ROC paradigm, is one of the more popular experimental methods).

Unexpected observed visual performance is often a consequence of complex visual processes (e.g., visual illusions), or combinations of several factors. For example, the well-known Contrast Sensitivity Function, or CSF, describing the human visual system's response to stimuli of varying contrast and resolution, depends not only on the organization of the retinal mosaic, but also on the response characteristics of complex cellular combinations, e.g., receptive fields.

In this book, the primary concern is visual attention, and so the book primarily considers the distinction between foveo-peripheral vision. This subject, although complex, is discussed here in a fairly simplified manner, with the aim of elucidating only the most dramatic differences between what is perceived foveally and peripherally. In particular, visual (spatial) acuity is arguably the most studied distinction and is possibly the simplest parameter to alter in eye-based interaction systems (at least at this time). It is therefore the topic covered in greatest detail, in comparison to the other distinctions covered here briefly: temporal and chromatic foveo-peripheral differences.

### 3.1 Spatial Vision

Dimensions of retinal features are usually described in terms of projected scene dimensions in units of degrees visual angle, defined as



**Fig. 3.1** Visual angle. Adapted from Haber and Hershenson (1973) ©1973. Reprinted with permission of Brooks/Cole, an imprint of the Wadsworth Group, a division of Thomson Learning

**Table 3.1** Common visual angles

Object	Distance	Angle subtended
Thumbnail	Arm's length	1.5°–2°
Sun or moon	—	0.5° or 30' of arc
U.S. quarter coin	Arm's length	2°
U.S. quarter coin	85 m	1' (1 min of arc)
U.S. quarter coin	5 km	1" (1 s of arc)

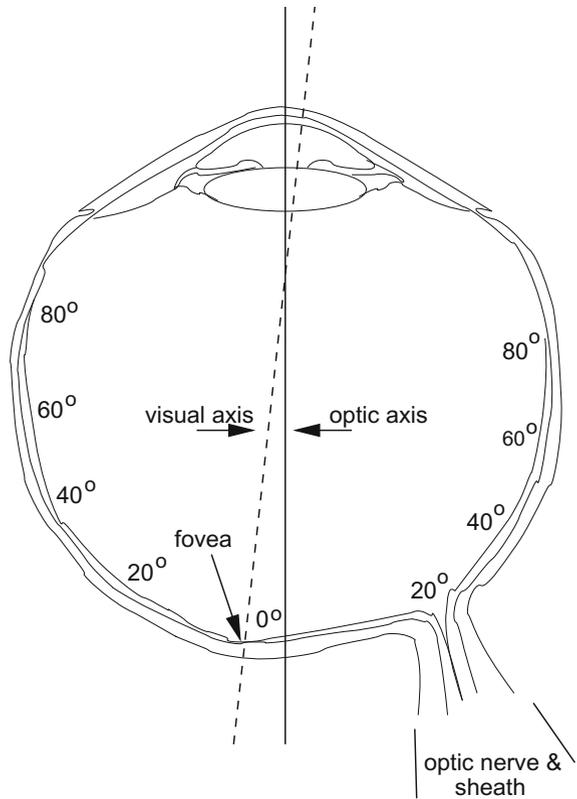
$$A = 2 \arctan \frac{S}{2D},$$

where  $S$  is the size of the scene object and  $D$  is the distance to the object (see Fig. 3.1). Common visual angles are given in Table 3.1.

The innermost region is the fovea centralis (or foveola) which measures 400  $\mu\text{m}$  in diameter and contains 25,000 cones. The fovea proper measures 1500  $\mu\text{m}$  in diameter and holds 100,000 cones. The macula (or central retina) is 5000  $\mu\text{m}$  in diameter, and contains 650,000 cones. One degree visual angle corresponds to approximately 300  $\mu\text{m}$  distance on the human retina (Valois and Valois 1988). The foveola, measuring 400  $\mu\text{m}$  subtends 1.3° visual angle, and the fovea and macula subtend 5° and 16.7°, respectively (see Fig. 3.2). Figure 3.3 shows the retinal distribution of rod and cone receptors. The fovea contains 147,000 cones/ $\text{mm}^2$  and a slightly smaller number of rods. At about 10° the number of cones drops sharply to less than 20,000 cones/ $\text{mm}^2$  and at 30° the number of rods in the periphery drops to about 100,000 rods/ $\text{mm}^2$  (Haber and Hershenson 1973).

The entire visual field roughly corresponds to a 23,400 square degree area defined by an ellipsoid with the horizontal major axis subtending 180° visual angle, and the minor vertical axis subtending 130°. The diameter of the highest acuity circular region subtends 2°, the parafovea (zone of high acuity) extends to about 4° or 5°, and acuity drops off sharply beyond. At 5°, acuity is only 50% (Irwin 1992). The so-called “useful” visual field extends to about 30°. The rest of the visual field has

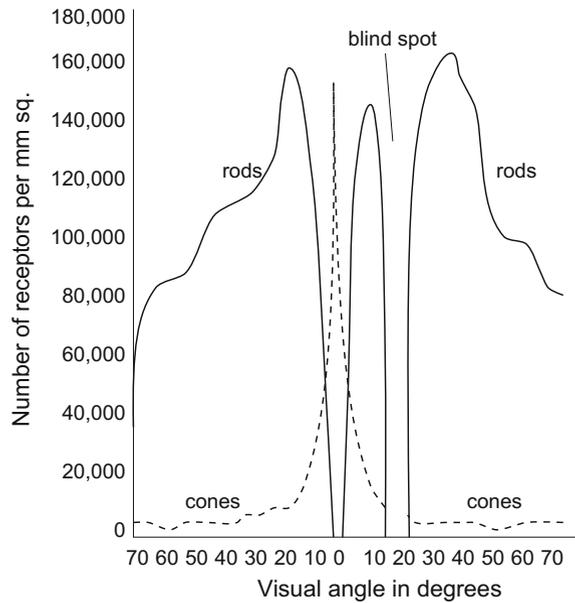
**Fig. 3.2** Density distributions of rod and cone receptors across the retinal surface: visual angle. Adapted from Pirenne (1967); as cited in Haber and Hershenson (1973)



very poor resolvable power and is mostly used for perception of ambient motion. With increasing eccentricity the cones increase in size, whereas the rods do not (Valois and Valois 1988). Cones, not rods, make the largest contribution to the information going to deeper brain centers, and provide most of the fine-grained spatial resolvability of the visual system.

The Modulation Transfer Function (MTF) theoretically describes the spatial resolvability of retinal photoreceptors by considering the cells as a finite array of sampling units. The  $400\ \mu\text{m}$ -diameter rod-free foveola contains 25,000 cones. Using the area of a circle,  $25000 = \pi r^2$ , approximately  $2\sqrt{25000/\pi} = 178.41$  cones occupy a  $400\ \mu\text{m}$  linear cross-section of the foveola with an estimated average linear inter-cone spacing of  $2.24\ \mu\text{m}$ . Cones in this region measure about  $1\ \mu\text{m}$  in diameter. Because one degree visual angle corresponds to approximately  $300\ \mu\text{m}$  distance on the human retina, roughly 133 cones are packed per degree visual angle in the foveola. By the sampling theorem, this suggests a resolvable spatial Nyquist frequency of  $66\ \text{c/deg}$ . Subjective resolution has in fact been measured at about  $60\ \text{c/deg}$  (Valois and Valois 1988). In the fovea, a similar estimate based on the foveal diameter of  $1500\ \mu\text{m}$  and a 100,000 cone population, gives an approximate linear cone

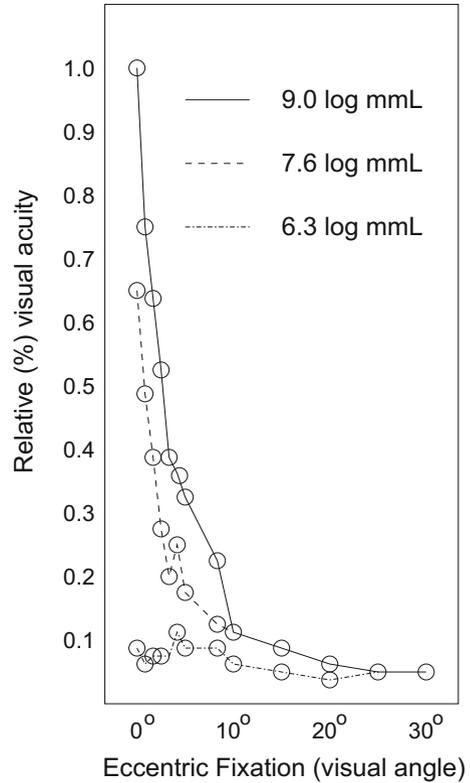
**Fig. 3.3** Density distributions of rod and cone receptors across the retinal surface: rod/cone density. Adapted from Pirenne (1967); as cited in Haber and Hershenson (1973)



distribution of  $2\sqrt{100000/\pi} = 356.82$  cones per  $1500 \mu\text{m}$ . The average linear inter-cone spacing is then 71 cones/deg suggesting a maximum resolvable frequency of 35 cycles/deg, roughly half the resolvability within the foveola. This is somewhat of an underestimate because cone diameters increase twofold by the edge of the fovea suggesting a slightly milder acuity degradation. These one-dimensional approximations are not fully generalizable to the two-dimensional photoreceptor array although they provide insight into the theoretic resolution limits of the eye. Effective relative visual acuity measures are usually obtained through psychophysical experimentation.

At photopic light levels (day, or cone vision), foveal acuity is fairly constant within the central  $2^\circ$ , and drops approximately linearly from there to the  $5^\circ$  foveal border. Beyond the  $5^\circ$ , acuity drops sharply (approximately exponentially). At scotopic light levels (night, or rod-vision), acuity is poor at all eccentricities. Figure 3.4 shows the variation of visual acuity at various eccentricities and light intensity levels. Intensity is shown varying from 9.0 to 4.6 log micromicrolamberts, denoted by log mmL [9.0 log micromicrolamberts =  $10^9$  micromicrolamberts = 1 mL, see Davson (1980, p. 311)]. The correspondence between foveal receptor spacing and optical limits generally holds in foveal regions of the retina, but not necessarily in the periphery. In contrast to the approximate 60 c/deg resolvability of foveal cones, the highest spatial frequencies resolvable by rods are on the order of 5 cycles/deg, suggesting poor resolvability in the relatively cone-free periphery. Although visual acuity correlates fairly well with cone distribution density, it is important to note that synaptic organization and later neural elements (e.g., ganglion cells concentrated in the central retina) are also contributing factors in determining visual acuity.

**Fig. 3.4** Visual acuity at various eccentricities and light levels. Adapted from Davson (1980) with permission ©1980 Academic Press



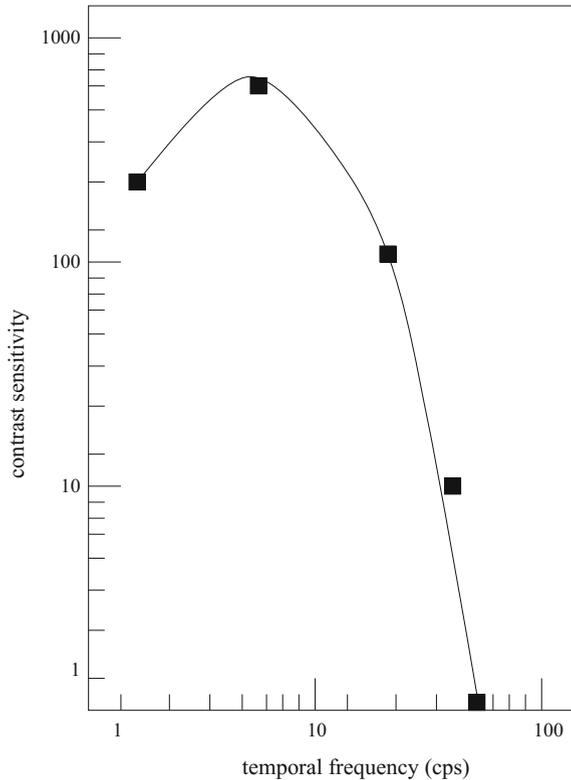
### 3.2 Temporal Vision

Human visual response to motion is characterized by two distinct facts: the *persistence of vision* and the *phi phenomenon* (Gregory 1990). The former essentially describes the temporal sampling rate of the HVS, and the latter describes a threshold above which the HVS detects apparent movement. Both facts are exploited in television, cinema, and graphics to elicit perception of motion from successively displayed still images.

Persistence of vision describes the inability of the retina to sample rapidly changing intensities. A stimulus flashing at about 50–60 Hz (cycles per second) will appear steady (depending on contrast and luminance conditions and observers). This is known as the critical fusion frequency (CFF).<sup>1</sup> A stylized representation of the CFF, based on measurements of response to temporal stimuli of varying contrast (i.e., a temporal contrast sensitivity function) is shown in Fig. 3.5. Incidentally, the curve of the CFF resembles the shape of the curve of the contrast sensitivity function (CSF) which describes retinal spatial frequency response. The CFF explains why flicker

<sup>1</sup>Also sometimes referred to as the Critical Flicker Frequency.

**Fig. 3.5** Critical fusion frequency. Adapted from Bass (1995) ©1995 McGraw-Hill. Reproduced with permission of The McGraw-Hill Companies



is not seen when viewing a sequence of (still) images at a high enough rate. The CFF illusion is maintained in cinema because frames are shown at 24 frames per second (fps, equivalent to Hz), but a three-bladed shutter raises the flicker rate to 72 Hz (three for each picture). Television also achieves the CFF by displaying the signal at 60 fields per second. Television's analog to cinema's three-bladed shutter is the interlacing scheme: the typical television frame rate is about 30 frames per second (depending on the standard use, e.g., NTSC in North America, PAL in other regions), but only the even or odd scanlines (fields) are shown per cycle. Although the CFF explains why flicker is effectively eliminated in motion picture (and computer) displays, it does not fully explain why motion is perceived.

The second fact that explains why movies, television, and graphics work is the phi phenomenon, or stroboscopic motion, or apparent motion. This fact explains the illusion of old-fashioned moving neon signs whose stationary lights are turned on in quick succession. This illusion can also be demonstrated with just two lights, provided the delay between successive light flashes is no less than about 62 Hz (Brinkmann 1999). Inverting this value gives a rate of about 16 fps which is considered a bare minimum to facilitate the illusion of apparent motion.

### 3.2.1 Perception of Motion in the Visual Periphery

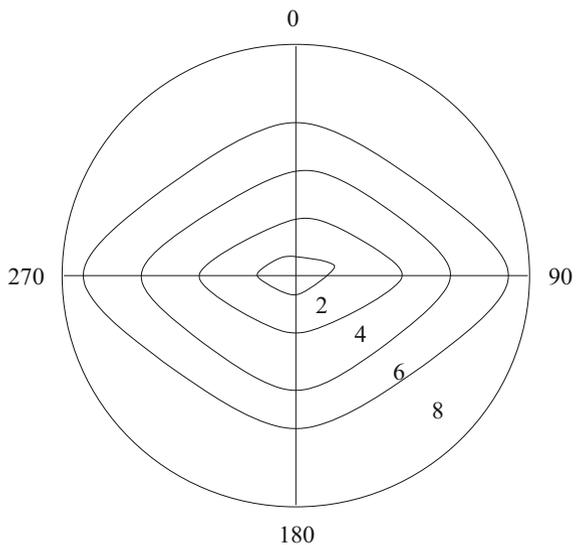
In the context of visual attention and foveo–peripheral vision, the temporal response of the HVS is not homogeneous across the visual field. In terms of motion responsiveness, Koenderink et al. (1985) provide support that the foveal region is more receptive to slower motion than the periphery, although motion is perceived uniformly across the visual field. Sensitivity to target motion decreases monotonically with retinal eccentricity for slow and very slow motion [cycle/deg; Boff and Lincoln (1988)]. That is, the velocity of a moving target appears slower in the periphery than in the fovea. Conversely, a higher rate of motion (e.g., frequency of rotation of grated disk) is needed in the periphery to match the apparent stimulus velocity in the fovea. At higher velocities, the effect is reversed.

Despite the decreased sensitivity in the periphery, movement is more salient there than in the central field of view (fovea). That is, the periphery is more sensitive to moving targets than to stationary ones. It is easier to peripherally detect a moving target than it is a stationary one. In essence, motion detection is the periphery’s major task; it is a kind of early warning system for moving targets entering the visual field.

### 3.2.2 Sensitivity to Direction of Motion in the Visual Periphery

The periphery is approximately twice as sensitive to horizontal-axis movement as to vertical-axis movement (Boff and Lincoln 1988). Directional motion sensitivity is shown in Fig. 3.6.

**Fig. 3.6** Absolute threshold isograms for detecting peripheral rotary movement. Numbers are rates of pointer movement in revolutions per minute. Adapted from McColgin (1960) with permission ©1960 Optical Society of America

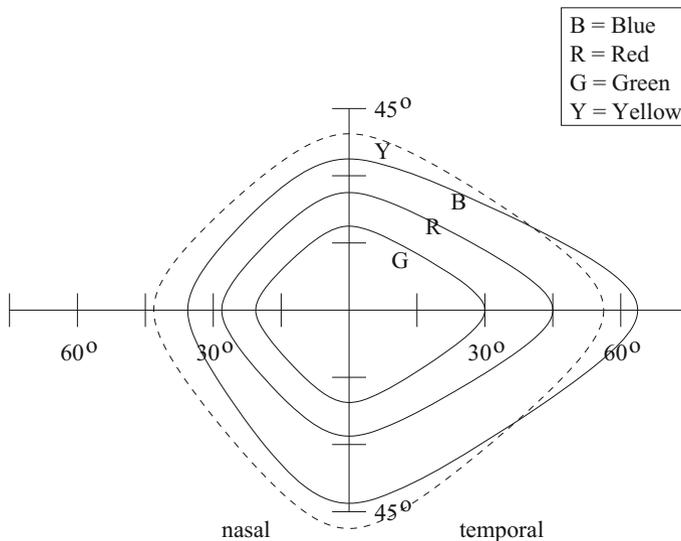


### 3.3 Color Vision

Foveal color vision is facilitated by the three types of retinal cone photoreceptors. The three main spectral sensitivity curves for retinal cone photoreceptors peak at approximately 450, 520, and 555 nm wavelengths, for each of the blue, green, and red photoreceptors, respectively. A great deal is known about color vision in the fovea, however, relatively little is known about peripheral color vision. Of the seven million cones, most are packed tightly into the central 30° region of the fovea with scarcely any cones found beyond. This cone distribution suggests that peripheral color vision is quite poor in comparison to the color sensitivity of the central retinal region. Visual fields for monocular color vision are shown in Fig. 3.7. Fields are shown for the right eye; fields for the left eye would be mirror images of those for the right eye. Blue and yellow fields are larger than the red and green fields; no chromatic visual fields have a definite border; instead, sensitivity drops off gradually and irregularly over a range of 15–30° visual angle (Boff and Lincoln 1988).

Quantification of perceptual performance is not easily found in the literature. Compared to investigation of foveal color vision, only a few experiments have been performed to measure peripheral color sensitivity. Two studies, of particular relevance to peripheral location of color CRTs in an aircraft cockpit environment, investigated the chromatic discrimination of peripheral targets.

In the first study, Doyal (1991) concludes that peripheral color discrimination can approximate foveal discrimination when relatively small field sizes are presented (e.g., 2° at 10° eccentricity, and less than 4° at 25°). Although this sounds



**Fig. 3.7** Visual fields for monocular color vision (right eye). Adapted from Boff and Lincoln (1988) with permission ©1988 Wright-Patterson AFB

encouraging, color discrimination was tested at limited peripheral eccentricities (within the central 30°).

In the second, Ancman (1991) tested color discrimination at much greater eccentricities, up to about 80° visual angle. She found that subjects wrongly identified the color of a peripherally located 1.3° circle displayed on a CRT 5% of the time if it was blue, 63% of the time if red, and 62% of the time if green. Furthermore, blue could not be seen farther than 83.1° off the fovea (along the  $x$ -axis); red had to be closer than 76.3° and green nearer than 74.3° before subjects could identify the color.

There is much yet to be learned about peripheral color vision. Being able to verify a subject's direction of gaze during peripheral testing would be of significant benefit to these experiments. This type of psychophysical testing is but one of several research areas where eye tracking studies could play an important supporting role.

### 3.4 Implications for Attentional Design of Visual Displays

Both the structure and functionality of human visual system components place constraints on the design parameters of a visual communication system. In particular, the design of a gaze-contingent system must distinguish the characteristics of foveal and peripheral vision (see Sect. 24.2). A *visuotopic* representation model for imagery based on these observations is proposed:

1. *Spatial resolution* should remain high within the foveal region and smoothly degrade within the periphery, matching human visual acuity. High spatial frequency features in the periphery must be made visible “just in time” to anticipate gaze-contingent fixation changes.
2. *Temporal resolution* must be available in the periphery. Sudden onset events are potential attentional attractors. At low speeds, motion of peripheral targets should be increased to match apparent motion in the central field of view.
3. *Luminance* should be coded for high visibility in the peripheral areas because the periphery is sensitive to dim objects.
4. *Chrominance* should be coded for high exposure almost exclusively in the foveal region, with chromaticity decreasing sharply into the periphery. This requirement is a direct consequence of the high density of cones and parvocellular ganglion cells in the fovea.
5. *Contrast* sensitivity should be high in the periphery, corresponding to the sensitivity of the magnocellular ganglion cells found mainly outside the fovea.

Special consideration should be given to sudden onset, luminous, high-frequency objects (i.e., suddenly appearing bright edges).

A gaze-contingent visual system faces an implementational difficulty not yet addressed: matching the dynamics of human eye movement. Any system designed to incorporate an eye-slaved high resolution of interest, for example, must deal with the inherent delay imposed by the processing required to track and process real-time

eye tracking data. To consider the temporal constraints that need to be met by such systems, the dynamics of human eye movements must be evaluated. This topic is considered in the following chapter.

### 3.5 Summary and Further Reading

Psychophysical information may be the most usable form of literature for the design of graphical displays, attentional in nature or otherwise. Introductory texts may include function plots of some aspect of vision (e.g., acuity) which may readily be used to guide the design of visual displays. However, one often needs to evaluate the experimental design used in psychophysical experiments to determine the generalizability of reported results. Furthermore, similar caution should be employed as in reading neurological literature: psychophysical results may often deal with a certain specific aspect of vision, which may or may not be readily applicable to display design. For example, visual acuity may suggest the use of relatively sized fonts on a Web page (larger font in the periphery), but acuity alone may not be sufficient to determine the required resolution in something like an attentional image or video display program. For the latter, one may need to piece together information concerning the visual contrast sensitivity function, temporal sensitivity, and so on. Furthermore, psychophysical studies may involve relatively simple stimuli (sine wave gratings), the results of which may or may not generalize to more complex stimuli such as imagery.

For a good introductory book on visual perception, see Hendee and Wells (1997). This text includes a good introductory chapter on the neurological basis of vision. Another good introductory book which also includes an interesting perspective on the perception of art is Solso (1999). For a somewhat terse but fairly complete psychophysical reference, see the USAF Engineering Data Compendium (Boff and Lincoln 1988). This is an excellent “quick” guide to visual performance.