Chapter 2

The human visual system

It may be surprising to find a chapter on the human visual system in a book on digital image processing, but there are two reasons to include it:

- We need to be aware that there is a large difference between the image we *display* and the image we actually *perceive*. There is, for instance, a difference between the *luminance* of a pixel on a computer screen and the perceived *brightness* of this pixel: if we double the screen luminance, this does not imply that the perceived brightness is doubled too. The brightness also depends on other factors, such as contrast around the pixel and various cognitive processes[1].
- The human visual system can perform a number of image processing tasks in a manner vastly superior to anything we are presently able to do with computers. If we want to mimic such processing, we need to carefully study the way our eyes and brain do this.

2.1 The human eye

The human visual system consists of two functional parts, the eye and (part of the) brain. The brain does all of the complex image processing, while the eye functions as the biological equivalent of a camera.

Figure 2.1 shows a cross section of the human eye and identifies its most important parts. What our eyes perceive of a scene is determined by the *light rays* emitted or reflected from that scene. When these light rays are strong enough (have enough energy),

---

[1] In this book we will use the term *luminance* for the actual physical brightness, and *brightness* for the perceived brightness of an object, pixel, *etc.*

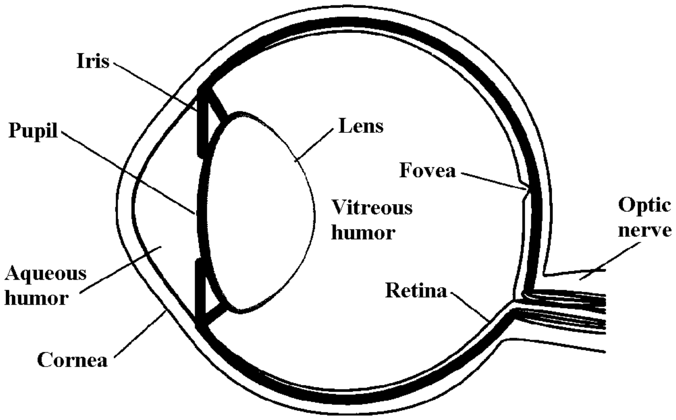

Figure 2.1 Cross section of the human eye.

and are within the right range of the electromagnetic spectrum (about 300 to 700 *nm*), the healthy eye will react to such a ray by sending an electric signal to the brain through the optic nerve.

When a light ray hits the eye, it will first pass through the *cornea*, then subsequently through the *aqueous humor*, the *iris*, the *lens*, and the *vitreous humor* before finally reaching the *retina*. The cornea is a transparent protective layer, which acts as a lens and refracts the light. The iris forms a round aperture that can vary in size and so determines the amount of light that can pass through. Under dark circumstances the iris is wide open, letting through as much light as possible. In normal daylight, the iris constricts to a small hole. The lens can vary its shape to focus the perceived image onto the retina.

In the retina, the light rays are detected and converted to electrical signals by *photoreceptors*. The eye has two types of photoreceptors: *rods* and *cones*, named after their approximate shape. The rods are abundant, about 100 million in a human eye, and spread evenly about the retina, except at the fovea, where there are almost none. The fovea is the area of the retina where our vision is sharpest. There are much fewer cones, about 6 to 7 million, which are mainly located around the fovea, but can be found in a low density in the entire retina. No photoreceptors are found at the point where the optic nerve attaches to the eye (the so-called *blind spot*), so we cannot perceive anything there. Since rods are more responsive to light than cones we can identify three types of vision, depending on the amount of light that reaches the eye. Under dark circumstances, practically only the rods are active. Since rods cannot discriminate colors, we perceive only shades of grey. We call this *scotopic* or night vision. Under daylight cicumstances, the cones are most active, and we experience *photopic* or day vision. In dimly lighted circumstances there is an intermediate stage where both rods and cones are active called *mesopic* vision.

We are able to distinguish colors because there are three distinct types of cones, each sensitive to a different band of the electromagnetic spectrum. The relative sensitivity as a function of wavelength of the cones is shown in figure 2.2. We can see that one type of cone is almost exclusively active in the region of 400 to 500 *nm*, and hence enables us to perceive colors in the violet-blue part of the spectrum[2]. The other two types have only a slightly different response curve and enable us to see colors in the spectrum from cyan to red. By combining the response of the three types of cones in a small area of the retina we are able to perceive any color. This process is called *trichromacy*[3].

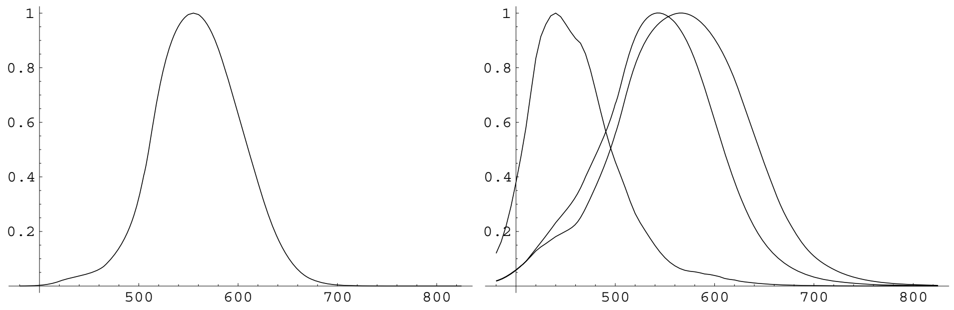

Figure 2.2 Photopic luminous efficiency function of the human visual system (left) and relative sensitivity of the three types of cones (right) as a function of wavelength in *nm*.

Example

If light with a wavelength of 550 *nm* (*i.e.*, green) is projected onto the retina, the response of the three types of cones (from left to right in figure 2.2) will respectively be 0.1, 0.975, and 0.975 of the maximum responses, which unique combination will be recognized as “green” by the brain.

Vice versa, a response of 0.0, 0.2, and 0.525 indicates the color red (around 640 *nm*).

Because of the trichromatic way the human eye operates, we also need three distinct colors to *reproduce* a certain color on a medium. For example, if you look up close at a TV screen, you will notice the screen is built up of tiny dots or stripes of either red, blue, or green.

---

[2] The visible spectrum of colors ranges from about 380 to 780 *nm* forming the familiar “rainbow” of colors going from violet to red. The primary colors of blue, green, and red are approximately found at the areas around 450, 550, and 650 *nm* respectively.

[3] From the Greek *chroma*, color.

2.2 Reflectivity and luminance

Suppose we have a light source that emits light rays with energy $E(\lambda)$, where λ is the wavelength of the emitted light, then the light I reflected from a certain object can be written as

$$I(\lambda) = \rho(\lambda)E(\lambda),$$

where $\rho(\lambda)$ is the reflectivity of that object. The reflectivity is a function that takes values between 0 and 1. Suppose $\rho(\lambda_1) = 1$, then this means that all of the light with wavelength λ_1 is reflected. Alternatively, if $\rho(\lambda_2) = 0$, this means none of the light at wavelength λ_2 is reflected, *i.e.*, all of the light at this wavelength is absorbed by the object. Effectively, the reflectivity function $\rho(\lambda)$ determines the color of an object. For example: if a certain object reflects only light with wavelengths λ of about 650 *nm* (*i.e.*, ρ is zero for other values of λ), we call it “red”.

The *luminance* L of an object is defined as

$$L = \int_0^\infty I(\lambda)V(\lambda)\, d\lambda,$$

where $V(\lambda)$ is the *luminous efficiency function* of a visual system. This function $V(\lambda)$ tells us how well a visual system is able to detect light of a certain wavelength. The typical luminous efficiency function of the human visual system is shown in figure 2.2. Obviously, we are best equipped to see light of a wavelength of 550 *nm*. At this wavelength, a light ray needs only a little energy to trigger a photoreceptor in our visual system. Alternatively, an electromagnetic ray with a wavelength of 1000 *nm* will not trigger anything, no matter how large its energy is.

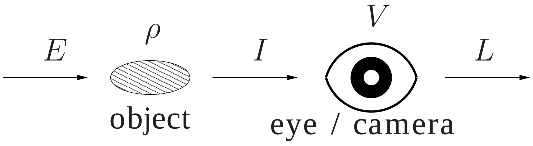

Figure 2.3 Schematic of object luminance.

Figure 2.3 shows a schematic of how object luminance relates to light energy, reflectivity, and luminous efficiency.

Intermezzo

Different visual systems will generally have different luminous efficiency functions. So what is “seen” by a rabbit, a snake, or a camera is generally different from human vision, amongst others because their visual system reacts differently (either weaker or stronger) to light of certain wavelengths. In biological visual systems, the main cause for this is the variety in cones that can be found. Some insects have been shown to have one or two types (mono- and dichromatic vision). Some turtles have two types, but each type of cone can also contain four different kinds of colored oil, so in effect there are eight types of cone. Pigeons have at least three types of cone, and five oil colors. In the color area we call red, a pigeon has six active forms of cones, and so it can distinguish far more shades of red than humans can.

The difference in cones is only one of many when comparing biological visual systems. If biological vision research has shown anything, it is that almost any species has it own unique visual system in terms of color and luminance perception. Even in species that have the approximate trichromatic vision of humans –like honeybees or bumblebees– there are large differences. For example: bumblebees cannot perceive ordinary white light.

The upshot of this is that we must be careful in taking our perception of an image as a "gold standard"; what an animal, a camera, or, –by extension– a computer perceives in the same image may be something completely different.

Because of various physical and psychophysical reasons, the luminance L of an object is not the same as the perceived brightness. We will show some concepts of psychophysics in the next section. The purely physical reasons include adaptation effects: when we "dark adapt" our eyes (*i.e.*, stay in complete darkness for over an hour), our scotopic vision has a luminous efficiency function that is shifted in comparison to the normal photopic function of figure 2.2. So under these circumstances our eyes will perceive an object as brighter than when using photopic vision. Another adaptation effect is that after seeing a light flash, the cells in the retina need a short period of time to recover. The effect of this recovery is noticeable as the occurence of so-called "after-images" if we look away from a brightly lit object.

2.3 Some psychophysics

The ability to detect a spot of light does not depend so much on the luminance of the spot itself as on the *difference* in luminance of spot and background, *i.e.*, the *contrast*. Of course, the luminance must be above some minimal value, but it is the contrast of spot and background that must be above a certain threshold before we can detect the spot. We call this threshold the *just noticeable difference*. *Weber's law* states that the just noticeable difference ΔL is proportional to the background luminance L. In other words, the higher the background luminance, the higher the contrast needs to be before we detect a difference.

In addition to Weber's law there are a number of other contrast effects that influence the perceived brightness, like the *Mach band effect* and the *simultaneous contrast* effect. The Mach band effect is shown in figure 2.4: even though each bar is uniformly grey, our visual system enhances luminance changes. As the figure shows, overshoots and undershoots appear in the brightness graph: the bright side of a contrast jump appears to be extra light and the dark side appears extra dark. The Mach band effect shows us that our visual system sharpens the edges of the objects we perceive by adding a little contrast.

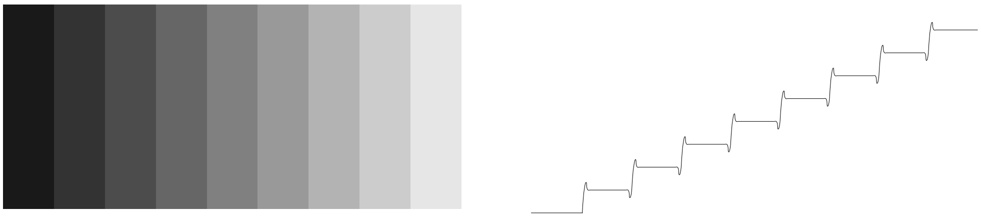

Figure 2.4 The Mach band effect: each bar is uniformly grey, but our visual system enhances each contrast jump. The graph at the right approximately shows the perceived brightness.

The *simultaneous contrast effect* is another effect that shows us that the perceived brightness depends on the contrast, in this case on the contrast with the local background. Figure 2.5 shows four squares of equal luminance, each with a background of different luminance. As the background gets darker, the perceived brightness of the squares increases. Simultaneous contrast is not just a 'low-level' physiological phenomenon, but requires a complex decision of what exactly is "background" in the image. See, *e.g.*, figure 2.6, the so-called *Benussi* ring, where a small change in the image suddenly changes our perception of background for the entire image.

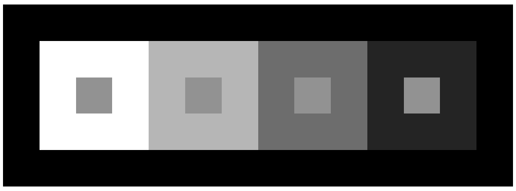

Figure 2.5 Simultaneous contrast effect. Even though the center squares are of equal luminance, they appear to be brighter if the background is darker. the one with the largest contrast to the background appears darkest.

Differences in brightness can even be induced by the mere *suggestion* of object edges, as shown in figure 2.7. Such effects are purely psychological in nature. Another psychological effect is the human tendency to view parts of images that appear to belong

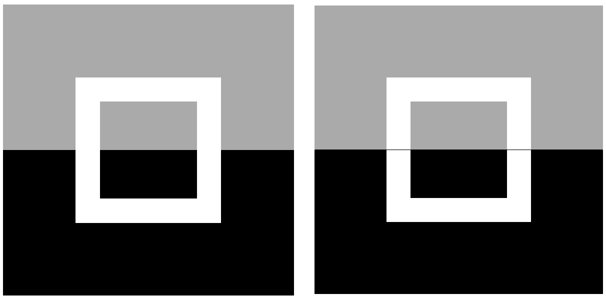

Figure 2.6 The Benussi ring. The brightness of the white ring in the left image appears to be uniform. In fact, the luminance *is* uniform. When we draw a small line as in the right image, the upper and lower parts of the ring are suddenly perceived as belonging to different backgrounds with different contrasts, and the upper and lower brightness is different.

together as a unit. This is a phenomenon known as *Gestalt*: an arrangement of separate elements whose configuration appears to be integrated is viewed as a single unit. The human visual system is uncanny in its ability in finding such units and their integrating patterns. This clustering ability works even in the absence of many object edges, as figure 2.8 shows. The integrating pattern determines what we perceive in such images. Figure 2.9 shows what happens if there are two obvious integrating patterns: we are not able to view both patterns simultaneously. Only with some difficulty are we able to make the *Gestalt switch* between the patterns.

Although explaining such psychological effects is beyond the scope of this book, it can be useful to acknowledge such effects when comparing human vision and computer vision.

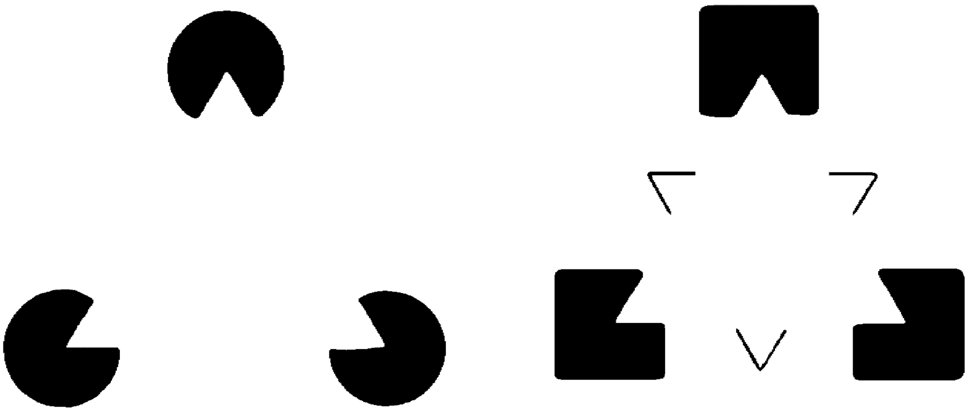

Figure 2.7 Examples of brightness differences induced by suggestion of edges.

Figure 2.8 Clustering ability of the human visual system. After a few seconds, this image of blobs contains a Dalmatian. . .

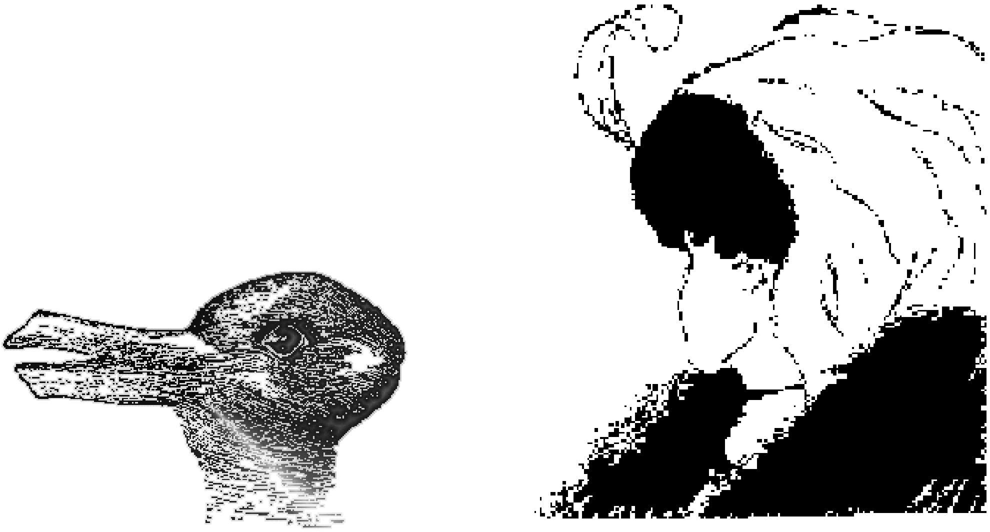

Figure 2.9 Example of images that contain two Gestalt patterns. Although the human visual system is able to detect a certain pattern very rapidly, we are unable to view both patterns simultaneously.