Chapter 6
Principles of Binocular Stereoscopic Imaging

Geoff Ogram
Stereoscopic Society, UK

ABSTRACT

The relation between visual perception and the recorded image is discussed in this chapter, emphasising the historical growth of the understanding of depth perception and its visual cues. The stereoscopic principle is explained in detail, and figures are given for comfortable viewing of stereoscopic images.

SPACE: MAPPING OUR SURROUNDINGS

Our world is three-dimensional. We were born into this three-dimensional space, we grew up in it and we handle it as part of our daily lives without conscious effort. We rarely contemplate its characteristics in detail or philosophise on its significance; we take it for granted. Nevertheless, if we wish to understand the principles underlying 3-D vision (stereopsis) or 3-D imaging (stereoscopy), we need to examine three-dimensional space in detail.

The word ‘space’ is not easy to define, although we may use it freely in common parlance. A typical dictionary definition is ‘Space, noun: that in which bodies have extension; a portion of extension; room; intervening distance; an interval; an open or empty place.’ A more helpful interpretation of ‘space’ can be achieved by mathematics. One of the many functions of mathematics is to describe the physical world in a succinct and accurate way. The dictionary definition covers both 2-D and 3-D space, but this chapter concentrates only on the latter. Using mathematical techniques we can map out our space to locate key points and show the relationships between various features within that space, such as the distance between two points or the angle between two lines. All such mapping enables us to understand the structure of that space in terms of the objects within it, their sizes, orientations and locations. These details can be established accurately with reference to some fixed point or scale. This principle will be familiar to anyone conversant with map-reading. Britain, for example, can be mapped onto flat sheets of paper despite the curvature of the Earth and the topography of the landscape. The space is two-dimensional and the reference lines are the N-S and E-W axes on which the Ordnance Survey references are based.

Principles of Binocular Stereoscopic Imaging

Thus the map reader can locate places and estimate distances from the scale of the map. Such maps are based upon the principle of a Cartesian graph plotted on a flat sheet with orthogonal x- and y-axes. Flat objects can be mapped in terms of width and depth (x- and y- coordinates). In a three-dimensional world we can apply the same principles, but we need an extra axis, the z-axis, perpendicular to both the x- and y- axes (Figure 1).

Whereas in 2-D representations the y-axis represents depth and the x-axis width, in 3-D space height is now represented by the z-axis. In this convention, distances in the y-direction are measured away from the observer. (Note, however, that in many mathematics textbooks the convention is to have the y-axis pointing out of the page.)

VISION: SEEING OUR SURROUNDINGS

Sight is arguably the most valuable of the senses, because it enables us to explore and understand our three-dimensional world to a greater extent than do hearing, taste, smell and touch. From the effects of visible light entering the eyes we are able to interpret information about our surroundings. The result of looking at our environment is usually described as ‘creating an image’ of that environment. The term ‘image’ is also applied to photographs and to what we see on a cinema or TV screen. An image matches what we see; it is an echo of reality. But the image we ‘see’ is not like that projected onto a screen or printed on paper. Light passes through the cornea and lens of the eye, then falls on the retina, and light-sensitive molecules within the rods and cones of the retina absorb the energy, resulting in a change in electrical potential. The detected signal is sent to the brain via the numerous nerve fibres. These signals are processed by the brain to produce what we call the ‘image’. Like all real images produced by a convex lens, the retinal image is inverted, though the brain interprets it as being erect. The physiology of vision was discussed in Chapter 2.

Figure 1. (a) Three axes (x, y & z) mutually at 90°, representing width, depth, and height; (b) one of the conventional labelling methods for axes