Visual Perception

John Lansdown
Centre for Electronic Arts
Middlesex University

May 1996

(This PDF version June 1998)

Contents

0. Introduction

1. The physical aspects of perception

2. Depth Perception

4. Colour perception

5. The principles of perceptual organisation

6. Concluding remarks and some theories

7. References

8. Appendix 1: Using the RGB model

9. Glossary index

0. Introduction

This report, made possible by support to the Centre for Electronic Arts (CEA) from British Telecom, looks at **visual perception.** At the heart of the study is a belief that a knowledge of visual perception is essential for the proper exploitation of those subjects central to the interests of the CEA: interactive multimedia, human-computer interaction, computer art, computer graphics and virtual reality. The study, however, does not cover the equally important and relevant areas of acoustic or haptic perception (touch and so on). These might be covered in a later document.

The structure of the report

The report is in six parts of varying lengths.

The first deals with the **physical aspects of perception** and attempts to bring together current information on the anatomy and physiology of the eye and the visual pathway. Strictly this section is not essential to the central theme of the report but is there for completeness and to enable understanding of other literature on vision.

The second part looks at **depth perception** and identifies a feature of the subject of perception as a whole — the fact that, despite intensive study going on for decades, perhaps centuries — there are still substantial areas where our understanding of the processes of perception is far from complete.

The third part covers **movement perception** — something that science has not given the attention it deserves although the mechanisms of motion perception in the cortex of the monkey have been intensively studied. The shortness of this section reflects the relative paucity of literature on the perception of movement in human beings.

The fourth section deals with **colour perception** where, despite an enormous literature — both scientific and artistic — much is still to be learned.

Drawing on the work of Gestalt psychologists, the fifth section looks at **perceptual organisation —** how we favour one interpretation of what we see over another that might be equally plausible. This section sets out, too, to show how we can exploit for design purposes the way we organise what we see.

The sixth, and final, section looks briefly at **theories of visual perception** and tries to draw together lessons of the whole study for those whose job it is to create works in the area of interactive multimedia, human-computer interaction, computer graphics and virtual reality.

The report concludes with a full list of **references** together with a **glossary index** and an **appendix.**

Coloured images are on the accompanying disc and can be viewed on Macintosh computers or PCs using the appropriate Acrobat viewer.

General remarks

Because of the special interests of the CEA, the report is written from the standpoint of art and design although bases its conclusions mainly on the scientific literature. The whole, though, is set in historical and cultural context.

It is worth stressing that much of what is believed about the physical aspects of our visual system is derived from studies on animals other than humans. There may be dangers in this. Although in many ways monkeys, especially macaque monkeys (the animals most favoured for study), have visual mechanisms similar to ours, they are not the same. For example, humans seem to have twice as many cones sensitive to long wavelengths as to medium wavelengths. This is not the case in other primates (Kaplan et al 1990). However, Kaas (1992) reports favourably on the similarities between the monkey and human visual systems, and Engel (1994), reviewing a publication by Gulyás et al (1993), says that the various studies of the human visual cortex:

> . . . have yielded convergent results that suggest the basic functional structure of our visual cortex can be compared to that of other primate species that have been studied even though the precise homology of visual areas in man and monkey is still a matter of debate. This conclusion is important since animal research in visual neurobiology is largely motivated and justified by the aspiration to understand the neural basis of our own perceptual capacities (p43).

Speaking of their own studies, Shipp et al (1995) are also positive:

> The results we obtained . . . confirmed that the initial prestriate visual areas of the human are mapped in a similar fashion to those of the monkey (p125).

Recent developments in non-invasive medical investigation techniques such as scanning by positron emission tomography (PET) are allowing study of the human brain in action and promise new insights into the way in which our visual systems work (Kushner et al 1988, Lueck et al 1989, Raichle 1990, Shipp et al 1995). The information gained from such studies will undoubtedly make us revise our views on the mechanisms of human visual perception and, perhaps sooner rather than later, lead us to resolve some of the confusions and problems that beset the subject.

Meanwhile we have to make do with the partial understanding we have.

1. The physical aspects of perception

Light and the eye

Followers of Pythagoras (BC 582-500) believed that sight was somewhat like the sense of touch and that light travelled outwards from the eye to ‘touch’ objects in order for us to see them. They thought that there was a ‘fire’ within the eye and Theophrastus (BC 372-286), justified this view by observing that ‘when one is struck, [the inner fire] flashes out’. Plato (BC 428-348) believed that the inner fire and daylight came together in a special way to enable us to see. Aristotle (BC 384 - 322), on the other hand, rejected the idea of light emissions from the eye and postulated that air was the necessary medium to complete the touching.

There was confusion in this area until the time of the great Arab physician, Ibn Al-Haithen, also known as Alhazen (935 - 1039). He surmised that rays of light came from objects to the eye. Because of a misunderstanding of his terminology, it was once believed that Alhazen thought that the image we see was formed on the interior back surface of the lens. More recently, it has been realised that he was probably aware of the fact that that the image formed on the interior back surface of the eye itself (Polyak 1941).

It was not until the beginning of the seventeenth century that Kepler (1571-1630), brought together the then current knowledge of optics, light and the anatomy of the eye sufficiently to explain vision more or less in the terms that we know it today. He believed the eye was an object similar to a camera — or rather, to a *camera obscura* — a notion that had been put forward by Leonardo da Vinci (1452-1519) about 80 years earlier.

The significance of Leonardo The work and words of Leonardo play a significant part in sections of this report and readers could be forgiven for questioning the need for such emphasis in a document that aims to bring together relevant modern material. There are a number of reasons that motivate the emphasis. Leonardo wrote a great deal about perception and seems to have had unique insights into the subject. Kenneth Clark (1939 p132), goes so far as to suggest that Leonardo had superhuman eyesight — a position that is supported by Ritchie Calder (1970 p214) who felt that Leonardo’s drawings of birds in flight could not have been made so accurately by someone with ordinary perceptual abilities. It has to be said, however, that many artists other than Leonardo were and are able to capture in drawings aspects of moving form that non-artists do not see. Rodin, for instance, developed a method of drawing for dealing with shapes of human movement. See also Leake (1993, 1995) for a different approach.

Of course, Leonardo had problems with the details of perception. For example, apart from his Paradox (which we deal with in the next Section), he could also not accept that an image projected onto the retina would be upside down. Hence he postulated the existence of a correcting lens in the centre of the eye in addition to the one at the front. He did this despite the evidence of his own anatomical investigations — or, perhaps, because of them. Calder (1970 p63) suggests that Leonardo was misled either by using an ox’s eye or, if a human eye was used, because the lens was pushed out of place during cutting. It was left to Kepler in 1604 to suggest that, even though the image on the retina is upside down, we see things the right way up because of mental processes or, as he thought of it, because of ‘activity of the Soul’.

Above all, though, Leonardo had views of vision that bridge the gap between science and art and between the ancient / mediaeval and renaissance viewpoints in a unique way. It is thus natural that this report, which also aims to set the subject in an art / science and historical context, should use Leonardo as its ‘patron’.

The eye as a camera

From the physical point of view, Kepler's camera analogy is a reasonably accurate one, as Wald (1950 / 1972) confirms. The eye is a more or less light-tight, roughly spherical chamber about 26 mm across with a lens system at the front and a light-sensitive surface, the retina, at the back (Figure 1. 1). The lens system serves to focus an image onto the retina in the way that a camera lens focusses an image onto photographic film. Unlike a camera, though, where focussing is achieved by moving the position of the lens relative to the film, the eye focusses by changing the shape of its lens.

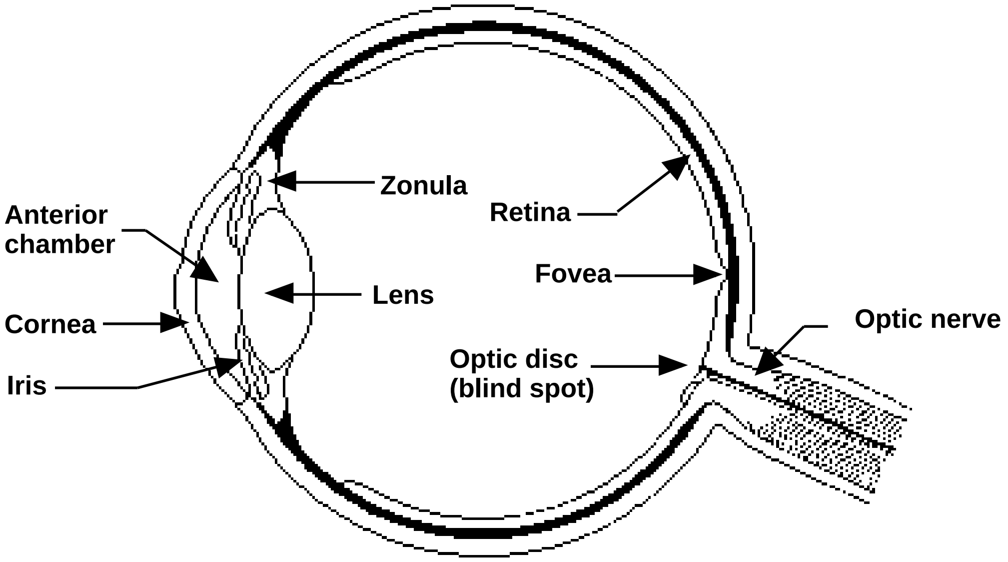


Fig 1.1 Diagrammatic section of the human eye

Something like a camera, too, the human eye is lined with a dark, almost black, matt surface to prevent too much back-scattering of light. This surface is called the ‘choroid membrane’. It is interesting to note that not all animals’ eyes have this dark lining. On the contrary, the eyes of nocturnal animals — cats for instance — have a shiny lining which reflects incident light back into the eye. This improves its light-gathering properties but inevitably gives considerable reduction in sharpness. The presence of this reflecting surface, the ‘tapetum’, is the reason that cats’ eyes seem to glow in the dark and might have contributed to the the early notions of the eyes sending out beams.

Problems with the camera analogy

But we can be misled by taking the camera analogy too far. In at least three respects the eye does not resemble a camera at all. First, and perhaps trivially, unlike in a camera, the space between the lens and the light sensitive surface is filled with a liquid. Sometimes this liquid contains the remains of dead cells, ‘floaters’, that interfere with our vision. Second, the light sensitive surface seems to be facing the wrong way — an inexplicable feature that we will come to again when we discuss the retina. Third, and most importantly, the eye is not a passive receiver of images. In general, we must actively look in order to see (Neisser and Beklen 1974). The concept, ‘I am a camera’ is not a correct one. Perception requires attention. In this respect the Pythagoreans were right: the analogy with reaching out to touch is much more accurate than it superficially appears.

The retina

Rods and cones

Our retinas incorporate two types of photoreceptors or light-sensitive cells: rods — which deal with low light conditions; and cones — which deal with normal light and colour. The rods and cones — so named because of their shapes — are connected in bundles to the optic nerve which channels the signals they produce to the brain for interpretation. In each eye there are about 130 million rods and cones but these are not distributed equally over the whole retina and rods outnumber cones by about 20 to 1. The majority of the 6 million or so cones are concentrated in and around one tiny, slightly depressed spot, no larger than the head of a pin, where there are no rods at all. This feature is known as the fovea, (Latin for ‘pit’), and in order to see anything really clearly, we must direct its image onto the fovea using saccades and fixations (of which more later).

A design anomaly?

A curious feature of the retina (and one which seems to defy good engineering practice) is that, except at the fovea, the receptors are not the first thing that incoming light encounters as it reaches the retina. The rods and cones are facing towards the back of the eye rather than the front and, except at the fovea, light is interrupted by a network of blood vessels, nerve fibres and cells before it reaches the photo-sensitive cells (Figure 1.2).

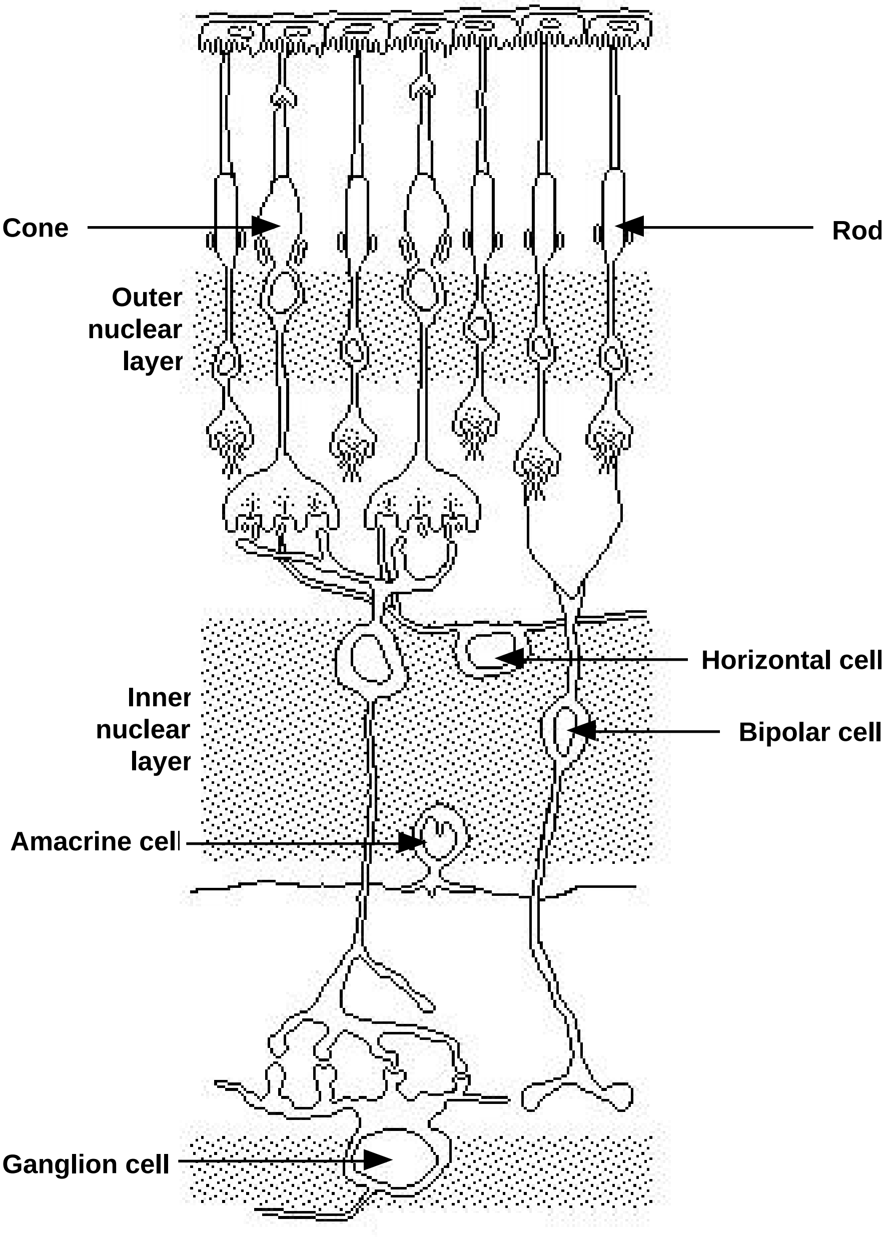


Fig 1.2. Diagrammatic section through the retina showing some of its different types of cell. Note that light from the lens travels upwards through the diagram. Thus it has to pass through the mass of ganglion and bipolar cells, as well as blood vessels before reaching the rods and cones. The pigmented cells are at the back of the eye (top of the diagram) and present a dark surface to prevent back-scattering of light. There are many more types of cell than are shown.

You can confirm this strange fact by sitting in a darkened room and shining a torchlight sideways into one eye for a few moments. When the torch is switched off you will see a spider-like after-image which is formed by the shadows of the blood vessels impinging on an unexpected area of your retina.

Something like modern radar systems, we seem to have a mechanism for cancelling out images that are usually stationary with respect to the retina. At all events, the presence of the blood vessels and nerves does not affect normal vision. Saccadic eye movements seem to be part of this fixed image cancelling mechanism. If we artificially prevent these movements from taking place and look at a picture, our view of it gradually disappears. This disappearance occurs not in the manner of a film-fade but in a piecemeal fashion until the whole field becomes grey. The length of time the image persists differs from person to person. I have personal experience of an experiment on fixing an image carried out on a group of people by the late Dr Chris Evans. In that case, partial persistence lasted from a few minutes (for most) to some days (for one person). She was naturally upset by the experience. Pritchard (1964) illustrates and comments on the phenomenon.

The sensitivity of rods and cones

An individual rod is so light-sensitive that as little as a single photon will excite it. However we need to have about 7 rods activated by photons before we become aware of the sensation of light (Hecht, Schlaer and Pirenne 1942). Individual cones seem to need about 5 photons before they become excited. If, on the other hand, we compare the total performance of rod vision with cone vision as in Figure 1. 3, at some wavelengths it appears that rod vision is about 100 times more sensitive than cone vision.

The reason for this is that the rods are grouped together with cells which produce a summation effect. This effect seems to be much less prominent in cone groupings.

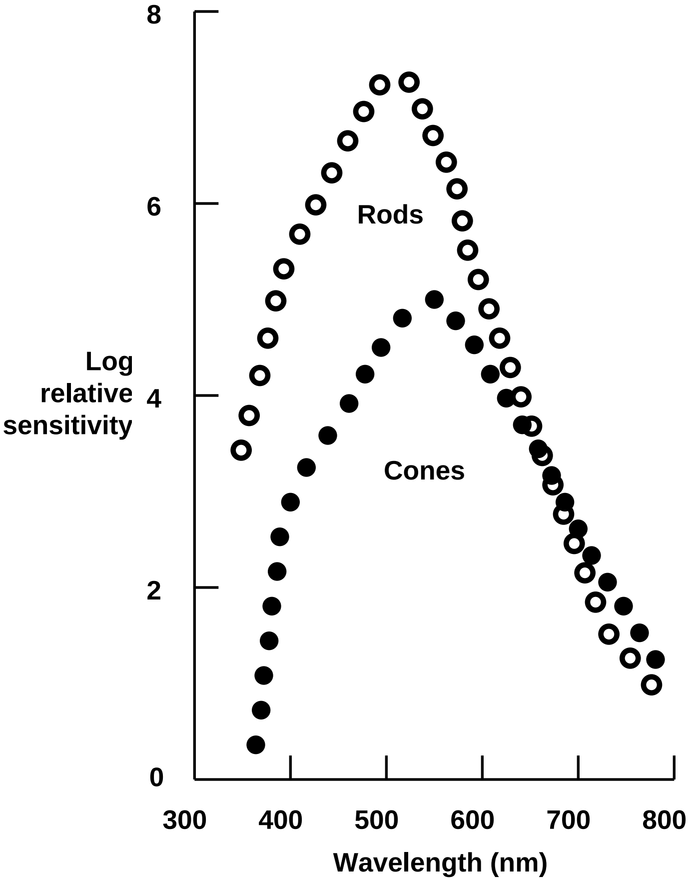


Fig 1.3. Relative sensitivities of rods and cones

The cells of the retina

As the diagrammatic section (Figure 1.2) shows, in addition to the rods and cones there are a variety of cells, that, in layered fashion, make up the retina. Each type of cell has its subtypes — indeed, there are at least 50 - 60 distinguishably different type of neural cell in the retina — a number that is increasing almost yearly as new discoveries are made (Sharp and Philips 1993 p7; Piccolino 1995 p14).

Kaplan et al (1990) tell us that:

> The primate retina, like all vertebrate retinae, contains 6 major cell types, organised in 5 layers: 3 layers of cell bodies (the outer nuclear layer, the inner nuclear layer, and the ganglion cell layers), and 2 layers of synaptic connections (the inner and outer plexiform layers) (p276).

The six major cell types are: the two photoreceptors: rods and cones which are in the outer nuclear layer; bipolar cells and horizontal cells which are in the inner nuclear layer; amacrine and ganglion cells which are in the ganglion cell layer.

Right at the back of the eye, as part of the choroid membrane, are the cells containing the black pigment, melanin, absorbing stray light and preventing it scattering back into the eye and affecting vision. These cells also serve to restore the light-sensitive chemicals (visual pigments) in the rods and cones when they have been bleached out by light.

Bipolar and horizontal cells In the inner nuclear layer of the retina are two basic forms of cell: bipolar, and horizontal. Bipolar cells take their input from the rods and cones and pass this on to the ganglion cells which are in the layer closest to the light. Some bipolar cells take input from a large number of photoreceptors. In the fovea, however, where there are no rods and the cones are packed together at a density of something like 150,000 per sq mm, it is likely that there is a 1-to-1 relationship between cones and bipolar cells. For these foveal cells, too, it is possible that the 1-to-1 relationship is retained throughout the visual pathway (Barlow 1982). This possibility carries with it the implication that a particular receptor in the fovea is matched to a particular cell in the visual cortex. Note too that, in some cases, there may be more than one bipolar cell associated with a given cone.

Horizontal cells connect over a larger area than the normal bipolar cells and seem to have the function of modulating input from the cells nearest to them by getting the more distant cells to act in an opposing fashion. Thus, if the nearest cells display a high level of activity, the more distant cells are made to show less activity and vice versa. Horizontal cells also play a part in lateral inhibition, something we return to in the discussion of Mach banding. These cells connect to between 7 and 12 receptors.

Amacrine and ganglion cells The inner nuclear layer of the retina is composed of amacrine and ganglion cells. The exact role of the amacrine cells is not yet known. They sometime stretch out for relatively long distances (that is, about 4mm), and, as the noted zoologist, JZ Young (1978) suggests, may make synaptic contact with neighbouring cells to influence them without propagating action potentials. Thus they may fulfil a feedback role similar to that of the horizontal cells in the bipolar layer (Weale 1982). This is also the view of Arden (1972) who says:

> Amacrines receive inputs from bipolar cells and other amacrines, . . . they output not only to ganglion cells and other amacrines, but also to bipolar cells via reciprocal synapses. Thus the bipolar cell which excites an amacrine may be influenced by feedback from the same cell which it is exciting. In addition, bipolar cells may be influenced by remote amacrines, which send dendrites laterally across the retina (p296).

Ganglion cells collect together the information from the other cells and pass it down the visual pathway. It has been known for a hundred years that there are many different anatomical types of ganglion cell. Only now — at least in animals other than humans — has it been able to match the anatomical with the physiological diversity. A comparatively recent discovery is that the ganglions can be divided into two broadly similar types: the magnocellular type, all of whose output goes via magnocellular nerves, and the parvocellular type, all of whose output goes via parvocellular nerves. This separation into types seems to remain throughout the visual pathway. The magnocellular route seems to be concerned with movement analysis and the parvocellular route with form, colour and stereo analysis.

The uniqueness of the human retinal structure

Unlike in the construction of the rods and cones — where there seems to be little structural difference between species — it is in the two layers, the inner nuclear layer and the ganglion layer, that the most differences between species occur. Weale (1982 p35) comments that, ‘from an evolutionary point of view this cannot occasion surprise’. Working from data established by Dowling

(Dowling and Boycott 1966; Dowling 1987), who has done a great deal to elucidate the anatomy of the retina, Willmer (1982 p42) suggests that the main forms of grouping of the cells are:
(i) numerous rods -> rod bipolar -> diffuse ganglion (Fig 1. 4a)
(ii) single cone -> midget bipolar -> midget ganglion (Fig 1. 4b)
(iii) several cones -> flat-top bipolar -> diffuse ganglion (Fig 1. 4c).

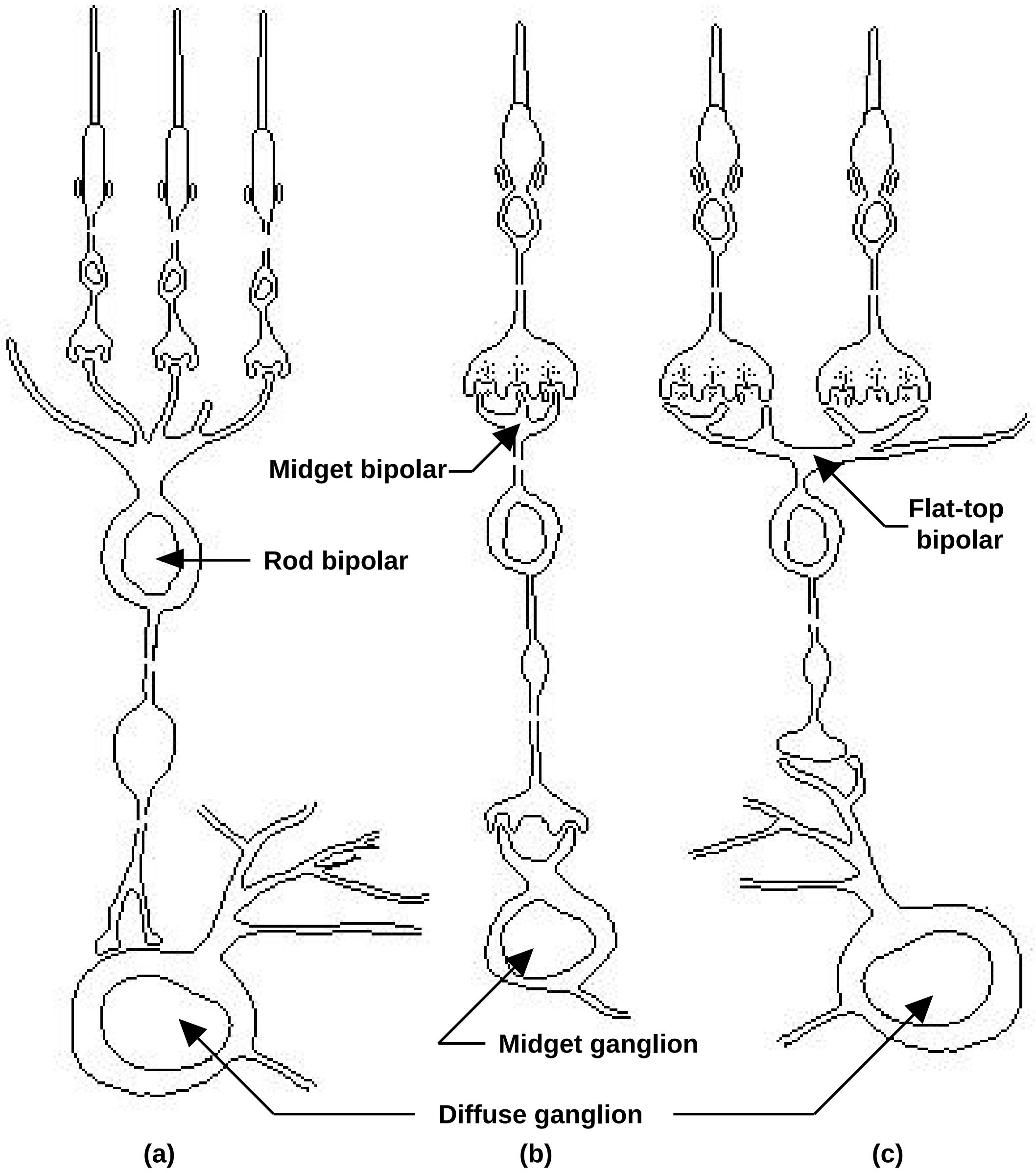


Fig 1.4. Basic photoreceptor types

Interconnections

The extent of the interconnections can be deduced from the fact that we have roughly 130 million photoreceptors, but only about 800,000 to 1,000,000 ganglion cells in each eye. Similarly, there are only about 1,000,000 fibres in each optic nerve. This means that the ratio of receptors to ganglions and nerve fibres is around 130 - 160 to 1. Thus, considerable filtering and, perhaps, other processing of information must take place at the retina before signals are passed to the optic nerves. To assist in this processing, a single ganglion gets input from anywhere from 1 to 1000 bipolar cells. Piccolino (1995) puts it thus:

> In most vertebrates, photoreceptors greatly outnumber ganglion cells. This suggests that the functional design of the retina should be optimised in order to preserve important information to be transmitted by optical nerve fibres, in spite of the limitations in the information-carrying capacity of the available neural channels (p14).

The roles of different retinal cells

Following Morgan (1991), we can list three possible functions for the various types of retinal cells. They can:

(1) play a direct role in the transmission of information through the retina,

(2) shape or regulate the transmission of information,

(3) enable the transmission of information.

In many cases we are not yet sure which or how many of these roles individual cell types play.

Synapses

As a result of the pioneering work of the Spanish neuroanatomist, Santiago Ramon y Cajal (1852 - 1934) in the last century, we know that nerves are not joined together in the same way that we join electrical wires, that is, with direct physical contact. At junction points there is a minute gap, known as the synaptic gap or synapse, between one nerve ending and the next (Figure 1. 5). When a signal is to be sent along a pathway of nerve fibres, a chemical is squirted across the synapse at the appropriate moment and this propagates the signal onwards.

It is important to note that some synapses are not on or off digital switches as was supposed until comparatively recently. Some synapses, particularly those from retinal bipolar and horizontal cells are capable of graded output. Ganglion and amacrine cells, however, give a spiked, digital output and it is this that is passed on down the optic fibres (Barlow 1972 p110).

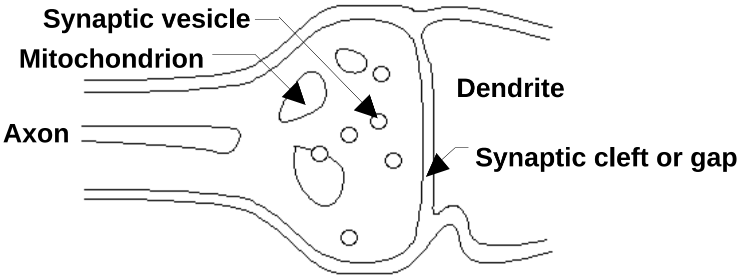


Fig 1.5. Diagrammatic section of an axon-dendrite synapse showing the synaptic cleft across which a chemical signal passes.

Ramon y Cajal and Golgi

In his investigation of the structure of the brain and the eye, Cajal used and improved the method invented by the Italian pathologist, Camillo Golgi (1842 - 1926), to show up neurons in high contrast to enable microscopic examination. In the course of his investigations Cajal discovered the neuron and Cajal and Golgi were jointly awarded the Nobel prize for this work in 1926. The broad details of the anatomy of the retina were established by Cajal using the microscope techniques of 100 years ago. Even with the development of the electron microscope, only relatively minor modifications to Cajal’s analysis have been made.

Retinal synapses

There are two broad sets of synapses in the retina, the outer synapses which join the receptors to the layer of bipolar and interconnecting cells, and the inner synapses, which connect this layer to the ganglion cells at the start of the optic nerve fibres. Signals from the photoreceptors have to pass through two to four synapses before reaching the ganglions. As has been pointed out, there are many different types of ganglion cells in the retina and their sensitivity and functions seem to be very different. Some monitor the illumination differences in small areas of the visual field; some are sensitive to green or to red; but the exact functions of each type in humans have yet to be determined (Barlow 1982).

Ganglion cell activity

It has been known for some time however, that the ganglion cells are continuously active and discharge electrical impulses at a rate of between 5 and 60 per second depending on type. Most deal with a small patch of visual field comprising a number of retinal receptors arranged in a roughly circular pattern. Sometimes the arrangement is such that the inner receptors of the circle are excitory (that is, they increase the discharge rate) and the outer ones are inhibitory (that is, they decrease the discharge rate). Sometimes it is exactly the reverse. From this, it is deduced that ganglion cells are not concerned with assessing exact levels of illumination but with comparing the level in a tiny area of the field with a slightly larger area surrounding it — a role assisted by the presence of horizontal cells.

A perceptual anomaly: Mach banding

This arrangement helps to gives rise to a perceptual anomaly that has become troublesome in computer graphics — so-called Mach banding. The effect was known to Andrea Mantegna (1431 - 1506) and to Leonardo but was first scientifically studied by the Austrian physicist, Ernst Mach (1838 - 1916), in 1865 who found that, when viewing two adjoining areas of contrasting lightness, he could see a thin, extra-bright band on the lighter side of the join between the two areas. There was also a thin, extra-dark band on the darker side (Ratliff 1965). Although, at that time, there was no physiological evidence to back up the idea, Mach postulated that the bands were due to some mechanism of lateral inhibition. The circular arrangement of receptors outlined above now gives us some evidence to support his contention. The visual brain increases the simultaneous lighting contrast at edges by the way some receptors inhibit the activity of others. At the darker side of an edge, the inhibition makes the receptors ‘undershoot’ the objective lightness; on the lighter side, the inhibition makes them slightly ‘overshoot’ it (Figure 1. 6). In computer graphics, Phong shading (Bui-Tong

Phong 1971) was introduced partly to help overcome the effect of Mach bands often present in Gouraud-shaded pictures (Gouraud 1971).

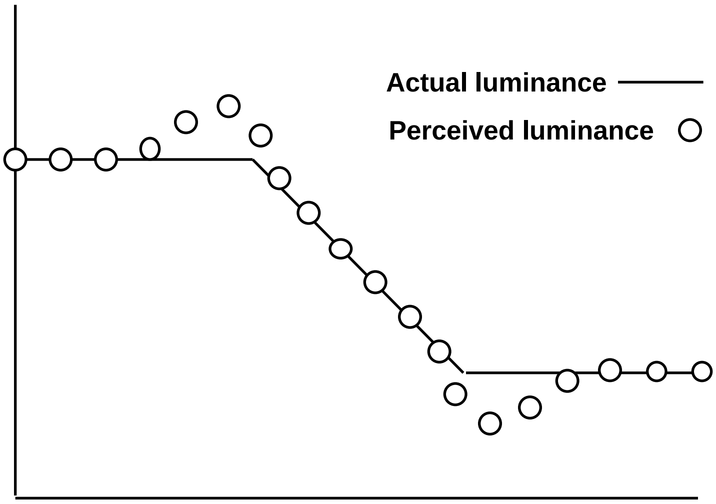


Fig 1.6. The Mach banding phenomenon arises because we perceive greater luminance or lesser luminance at sharp changes in actual luminance.

Exploiting Mach bands

Ratliff (1972 / 1976) gives some fascinating examples of the way in which the perceptual mechanisms that make us see Mach bands have been used by artists to improve the luminosity of parts of paintings and other art works. One way of doing this is to edge adjoining surfaces of similar brightnesses with a darker colour. This makes the edged area seem brighter than its surround. In discussing the way in which this and similar ideas have been used in Chinese pottery, Ratliff points out:

> Such techniques date back as far as the Sung dynasty (960 - 1279) and they are still employed in Oriental art.

He might have added too that the techniques are still taught to Western students of painting even today.

The visual pathway

As Figure 1. 7 shows, the optic nerves pass from each eye into parts of the brain called the ‘lateral geniculate nuclei’ (LGN) or the ‘lateral geniculate body’ (LGB). Some of the nerve fibres from the left eye go to the right LGN and some of those from the right eye go to the left LGN. The nerves that make the crossover are those dealing with the nasal sides of the visual field. The crossover point occurs in an area known as the ‘optic chiasma’. In the chiasma, the nerves also divide into new groupings to enable some of the left eye fibres (those dealing with the temple side of the visual field) to go to the left LGN and some of the right eye temple-side fibres to go to the right LGN.

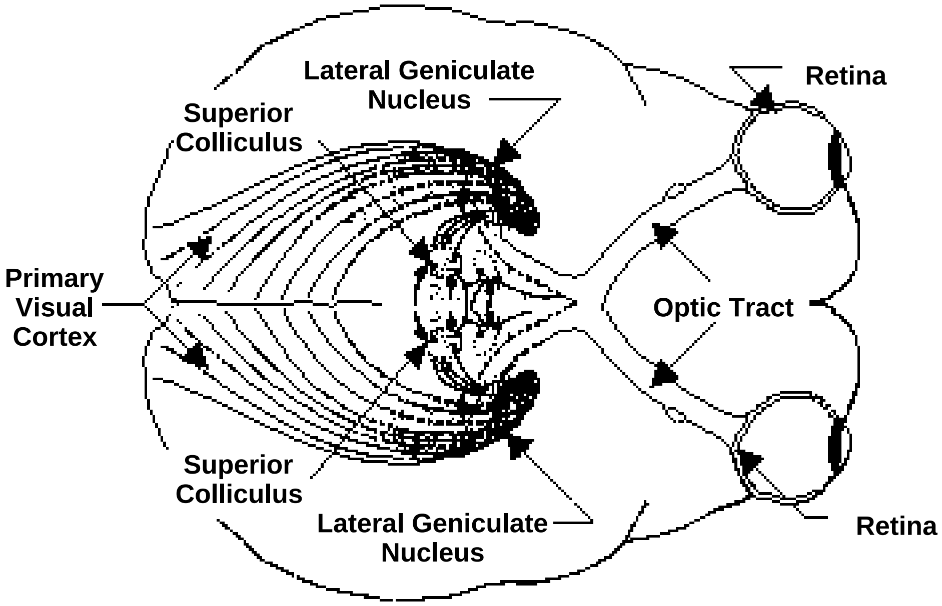


Fig 1.7. Diagrammatic plan view of the brain showing the visual pathway.

In addition, a comparatively few fibres go to two areas called the ‘superior colliculi’ (SC). The function of the superior colliculus is to perform the necessary computations for determining the amount of movement of the eyes, head and body, in order to view a scene properly. When the computations are done, the SC send messages to other areas of the brain which control the necessary movements.

After some processing in the LGN, the visual signals are synapsed to nerve fibres running to the visual cortex (VC) which is right at the back of the brain. From the vision point of view, this area is made up of (Figure 1. 8): the striate cortex or Brodmann’s area 17, the parastriate cortex or area 18, and the peristriate cortex or area 19 (and there may be other areas too). Here again we have a lateral split in specialisation with the right VC dealing mainly with the left nasal side vision and the left VC mainly with the right. Each VC, however, receives some input from both eyes. At this point, there are several further synapses before the visual signals are sent from area 17 to other parts of the VC and to other areas of the brain.

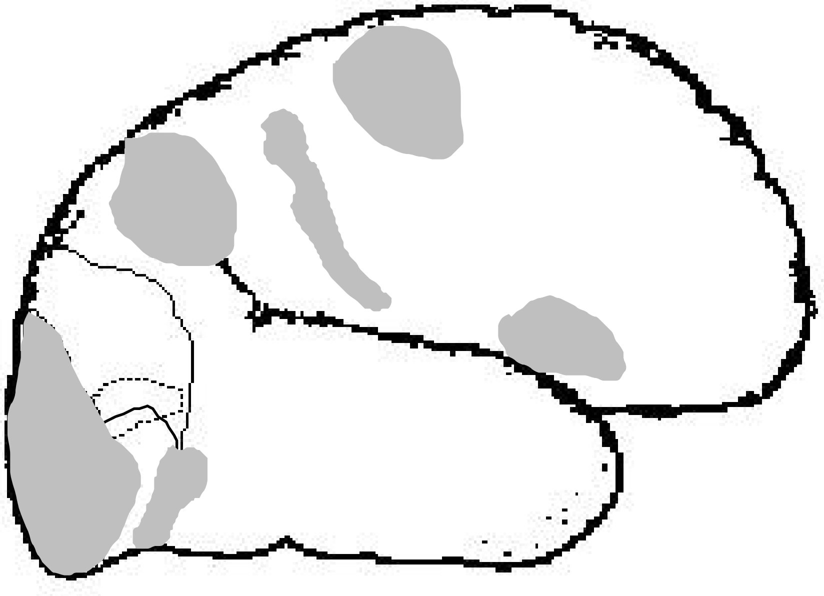


Fig 1.8. Diagrammatic side view showing of the human brain marked, at the rear, with Brodmann's cortical areas 17 (V1), the striate cortex; area 18 (V2), the parastriate cortex; and area 19, the peristriate cortex. Also shown are the possible locations of areas V4 and V5 together with the extent of area 17 at the middle of the brain.

There is even a set of nerve fibres that return some of the signals back to the LGN, but the function of these is not understood. Large portions of the brain seem to be involved in the visual process and it has been suggested that excitation of a few cells in the fovea will sometimes cause 10,000 times that number to come into play in the VC. It is as if the VC is a hugely expanded map of the retina (Mason and Kandel 1991). In fact, each degree of field at the fovea is represented by about 6 mm of the striate cortex. The two hemispheres of the cortex are connected by the corpus callosum which serves to coordinate the visual functions of both parts of the VC. The major eye to brain connections can, then, be represented diagrammatically as in Figure 1.9. Note that the nerve fibres forming the VC are of two sorts: P-type, carrying signals from the parvocellular ganglion cells, and M-type, carrying signals from the magnocellular ganglion cells.

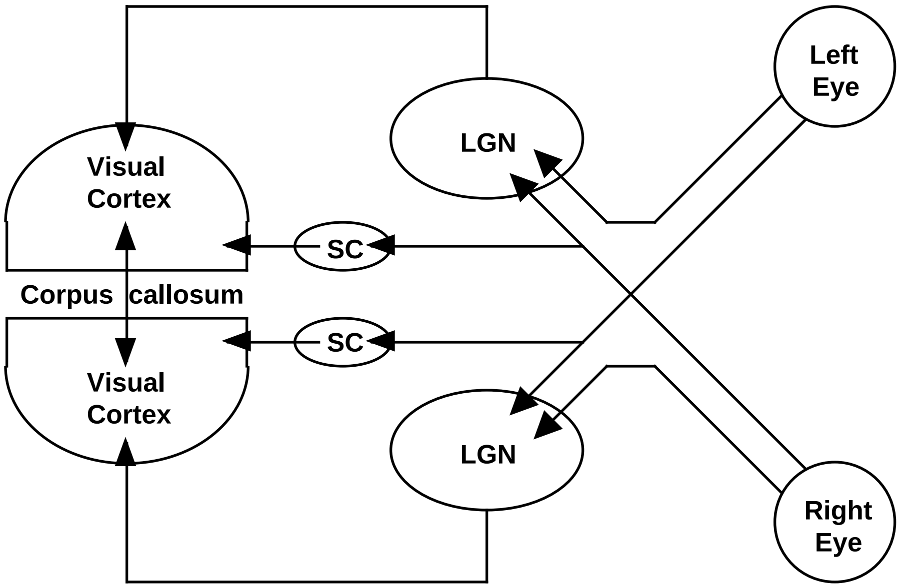


Fig 1.9. The visual pathway showing diagrammatically the interconnections to the lateral geniculate nucleus (LGN), the superior colliculus (SC), and the visual cortex

The structure of the LGN and VC

Rather in the manner layer cakes or club sandwiches both the LGN and the VC are made up of layers of cells of different forms and densities (Figures 1. 10 and 1. 11). According to Hubel (1995):

> For anatomical richness, in its complexity of layering, area 17 exceeds every other part of the cortex (p97)

The LGN comprises six layers four of which contain cells of relatively small size (the parvocellular layers) and two of which contain cells of relatively large size (the magnocellular layers). The corresponding ganglion cells in the retina project to one or other of these layers in the LGN.

Area 17 of the VC has nine layers which neurophysiologists code as 1, 2, 3, 4A, 4B, 4Cα, 4Cβ, 5 and 6. Layer 4C has the highest density of cells, while layers 1, 4B and 5 are the most loosely packed.

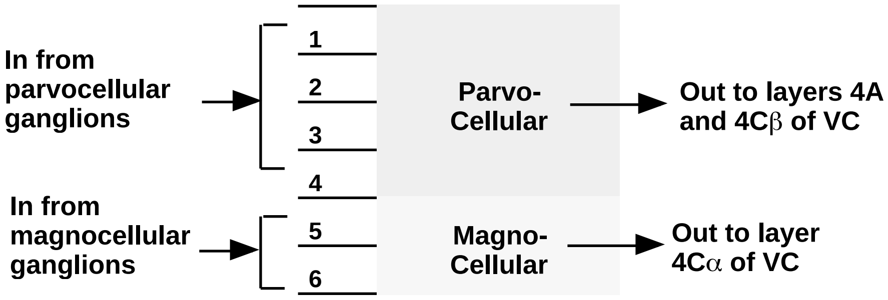


Fig 1. 10. Diagrammatically showing the layers of the LGN.

Layer 4Cα receives input from the magnocellular layers of the LGN and layers 4A and 4Cβ receive input from the parvocellular layers of the LGN. Layer 4B seems to specialise in movement analysis and layers 1, 2, 3 and 4A in the analysis of form, colour and depth (Hubel 1995).

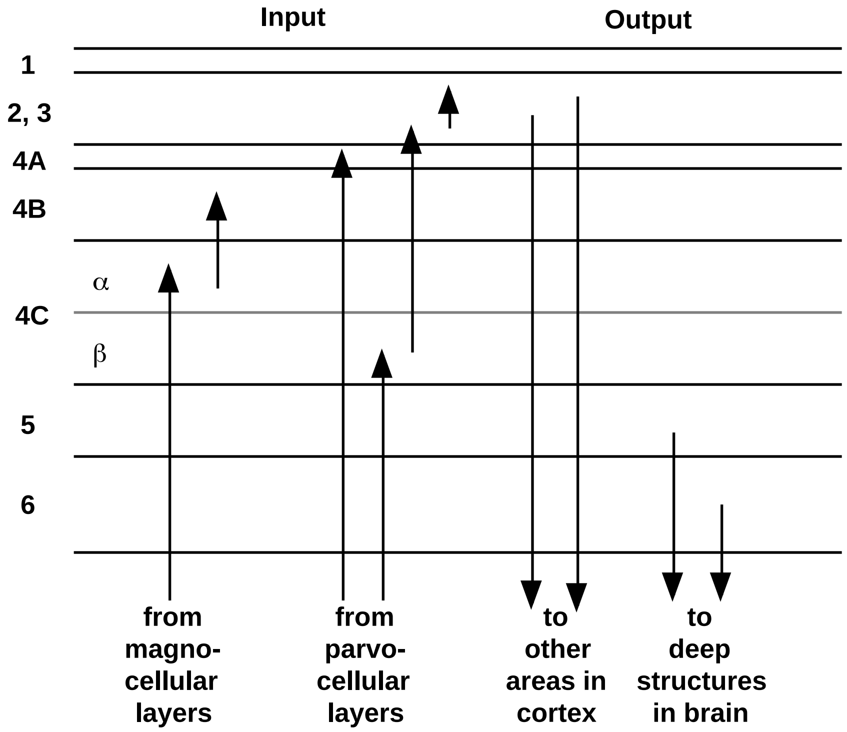


Fig 1. 11. The diagrammatic layered structure of the VC

Despite the fact that the information channels in the visual pathway seem to be equally distributed to both the right and left sides of the visual cortex, PET scan studies by Kushner et al (1988) show much more activity in the right side of the brain and this suggests that the right side of the cortex is specialised for visual processing.

Information flow

In their passage from receptors to the visual cortex into our consciousness and on into memory, the signals from the eye are synapsed such a comparatively large number of times that this must have an influence on the information flow. Steinbach (1962) has estimated the flow levels in bits per second (bus) at various stages in the processing and his conclusions are summarised in Figure 1. 12.

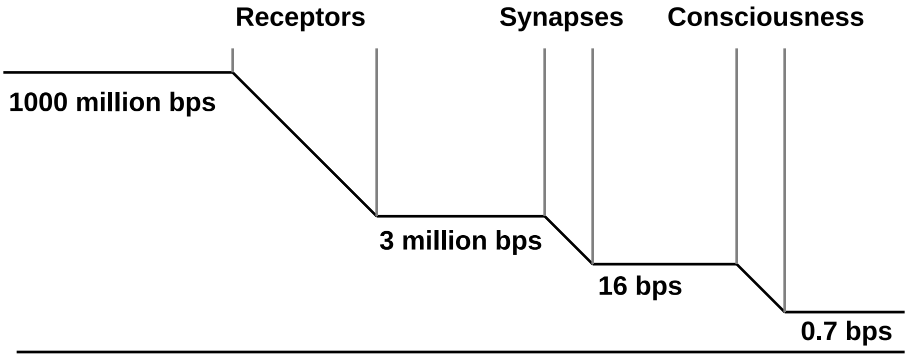


Fig 1.12. Information flow rates in bits per second showing enormous reduction from the light entering the eye to conscious perception (after Steinbach 1962).

Specialised cells

In both the LGN and the VC, the activity of individual nerve cells (neurons) is correlated with the visual pattern falling on the retina although, as mentioned above, the area of activity might be much larger. Experiments with monkeys suggest that some neurons in both the LGN and the VC are specialised to respond to lines, bars and edges and are only activated when these figures are seen in particular orientations (Hubel and Wiesel 1979). In some areas there seem to be cells which respond only to particular colours. Zeki (1974) suggests that there are (at least) five areas containing specialised neurons.

These areas deal with orientation, binocularity, depth, colour and movement, and are arranged so that processing passes from one area to another through a network of interconnecting pathways. Working on monkeys, Gross and his colleagues (Gross, Rocha-

Miranda and Bender 1972), have found cells so specialised that they only respond to the appearance of the monkey's own hand. Neurons of this hyper-specialisation have come to be known as 'gnostic' cells (because they know about the world). Humourists have suggested that, if monkeys have 'monkey hand recognition cells' then humans might have 'grandmother's face or yellow Volkswagen recognition cells' — a question that remains open! Although it is fairly clear that we do have specialist neurons, it is somewhat unlikely that gnostic cells of such hyper-specialism actually exist. It would surely require too many different sorts to make sense of even the elementary items that we need to recognise for day-to-day living let alone all the new patterns that we encounter.

Saccades and fixations

Our eyes are continuously in motion and we are constantly by making small jerky movements of the eye, known as saccades — a word which comes from the old French term for 'flick of a sail'. During a saccade, the image is 'blurred' across the retina, but we are unaware of any blurring of our vision. This implies that we must have perceptual mechanisms to cancel out such effects.

Yarbus (1967), who after Buswell (1935), made the first comprehensive study of eye movements in picture viewing activities showed that we use both saccades and fixations in the process. A fixation lasts for about 0.2 to 0.5 seconds and, during that period, our eyes are stationary as we concentrate on a particular area of picture (not more that about 5 degrees of visual angle). We then move the eye but rarely more than by 15 degrees of visual angle. The saccades are very speedy and occupy only about 5 of the time we spend scanning a picture. Points that arise from Yarbus's work — and that of others (such as Cooper 1974, Mackworth and Morandi 1967, Antes 1974, Loftus and Mackworth 1978) are:

- A large proportion of the fixations are on the most informative parts of the picture.

- The fixations are not necessarily on the lightest or the darkest areas nor on those with the most detail (unless these are the most informative).
- We do not seem to follow the contours of objects except when scanning the profiles of faces.
- The scanning and fixation patterns we use are dependent on our purpose in viewing the picture and are sensitive to the prior information we have about it.
- We do not need to concentrate on moving our eyes.
- Saccades and fixations seem to have some relationship with memory and recall.

Lansdown (1994) proposed a yet untested idea to use saccades and fixations in retrieving data from visual databases (Figure 1. 13).

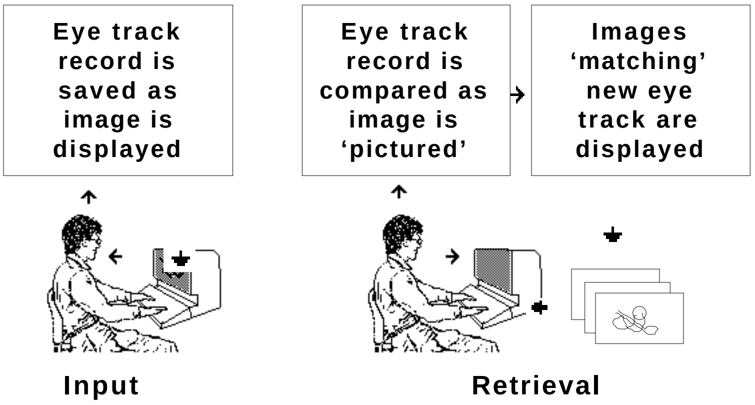


Fig 1. 13. A proposal for the use of saccades and fixations in retrieving images from visual databases. The pattern of saccades and fixations is recorded on input. When wishing to retrieve the image, the user 'pictures' it on the screen and the eye track of this picturing is compared with stored tracks to allow the computer to display 'matching' images

The role that saccades and fixations play in design of images is looked at in part 5 of this report.

Summary

This part of the report examined some of the physiological attributes of our visual perception system. Much of what is known about the system in operation is derived from studies of monkeys (and some other animals — notably cats and rabbits). Although caution needs to be expressed concerning extrapolating from these studies to humans, recent PET investigations of the living human brain suggest that a great deal of what is known about the monkey visual system seems to apply to us too.

The important lesson to come out of all these studies is the active nature of visual perception — the need to look in order to see. One of the purposes of designing is to guide and channel that act of looking to enable information to be more readily comprehended.

2. Depth Perception

The mechanisms of depth perception

A number of mechanisms come into play to enable us to perceive depth. They include:

- stereopsis
- motion parallax and motion perspective
- texture gradients
- linear and aerial perspective
- overlapping and partial occlusion
- retinal image size
- shadows and shading.

These mechanisms and their interactions are still not wholly understood. As Loomis et al (1992) point out:

> It might be thought that visual space perception, having been the focus of study for a number of decades, would be properly understood in functional terms if not in terms of the underlying mechanisms as well. Yet a proper understanding, even of function, remains elusive, in part, because of the diversity of theoretical approaches and empirical findings that exist without any serious attempt at integration.

The process of perception of depth and distance was certainly a puzzle to the philosopher, George Berkeley (1685-1753) who wrote:

> . . . distance of itself . . . cannot be seen. For distance . . . projects to only one point in the fund of the eye. Which point remains invariably the same, whether the distance be longer or shorter.

Of course, as many have pointed out through the ages, we do not see ‘depth’ at all and hence to talk about ‘depth perception’ is possibly somewhat misleading. What we see is not ‘depth’ but surfaces and textures of objects. However, for convenience, we will refer to this phenomenon as ‘depth perception’.

Binocular and monocular vision

We know now that some of the processes of depth perception depend on the fact that we have two eyes. Others are entirely monocular. In fact — and perhaps surprisingly — *most* of the mechanisms that we use to determine three-dimensional arrangement or layout do not depend on binocular vision at all (Figure 2. 1). Indeed, often the information presented to our eyes stereoscopically can be overridden by other depth information. For example, a pair of stereo images of a scene in linear perspective will still appear to have proper depth even when the images are presented to the wrong eyes (left image to right eye / right image to left eye). The well-known illusion of viewing a hollow mould of a face is another example of stereoscopic information being overridden by other factors. In this case, despite all the stereoscopic cues about depth, it is very hard to see the mould as anything but the sculpture of a face. Hill and Bruce (1993) show that this effect is not wholly dependent on lighting conditions and they put the illusion down to an apparent cognitive preference we have for convexity rather than concavity.

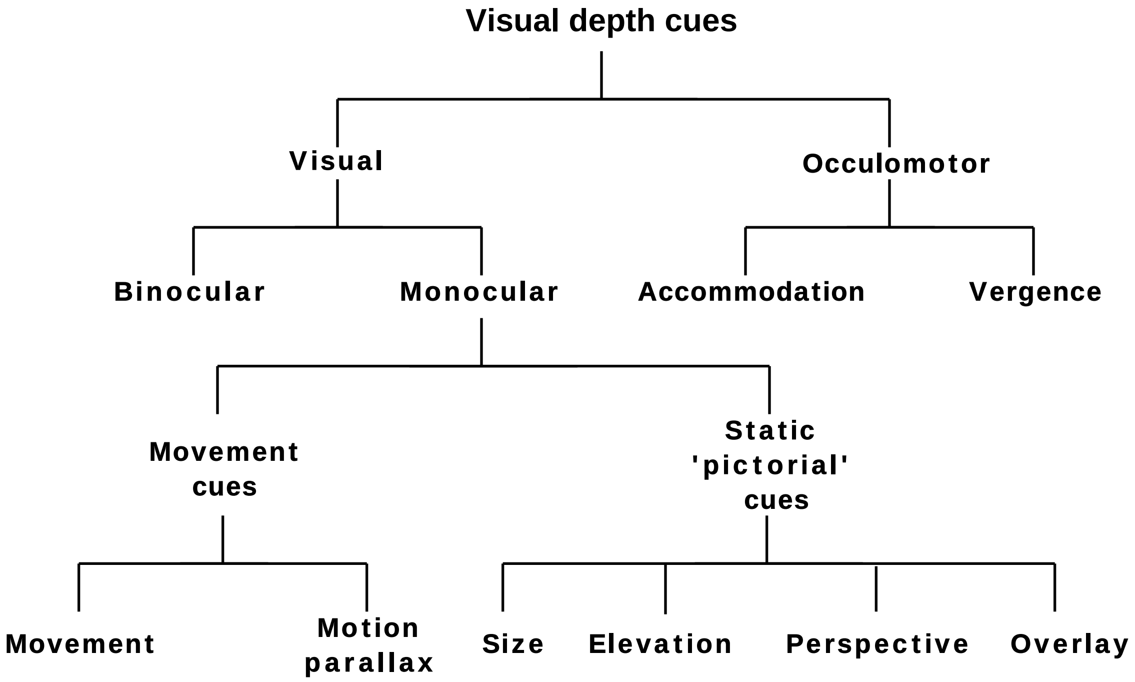


Fig 2.1. A hierarchy of visual depth cues (based on Fig 10.1 of Day 1994). Note the use of the word 'vergence' which refers to the capacity for moving the eyes and includes convergence.

In my view, this is an inadequate explanation, although a study by Massironi and Savardi (1991) of the Andrea Pozzo frescoes on the cylindrically vaulted ceiling of St Ignatius Church in Rome, does suggest that odd perceptual problems arise when we try to view images on concave surfaces. Pirenne (1970 pp79-94) deals in detail with the ceiling and Pozzo’s method of projection. He includes many pictures of the artwork viewed from the ‘correct’ position (where it appears to continue the real architecture in 3-dimensions). Also shown are pictures viewed from points away from the correct position and there the images appear to be distorted. This example may, of course, have nothing to do with our perception of curved surfaces but be just an exaggerated case of the normal situation in linear perspective which is strictly correct only when viewed from the station point.

Stereopsis

Stereoscopic vision is a binocular effect that was known to and commented on by Euclid (BC450-374) and the Greek physician, Galen (c130-c200) although both seemed to take the view that, more than to help in depth perception, binocularity was primarily necessary for us to see more of a scene than we could with one eye alone. References to the subject occurred throughout the Renaissance notably by Leonardo da Vinci (1452-1519); the Belgian monk, François d’Aguilon (1546-1617), also known as Aguilonius; the Italian philosopher, Giambattista della Porta (1535-1615); and the French philosophers, Pierre Gassendi (1592-1655) — particularly René Descartes (1596-1650), but the role of binocular vision seems not to have been fully studied until the early 1800s.

Leonardo — whose curiosity knew no bounds — was especially puzzled by binocularity and wondered why, as we see a slightly different view with each eye, some solid objects didn’t appear to be transparent (Figure 2. 2).

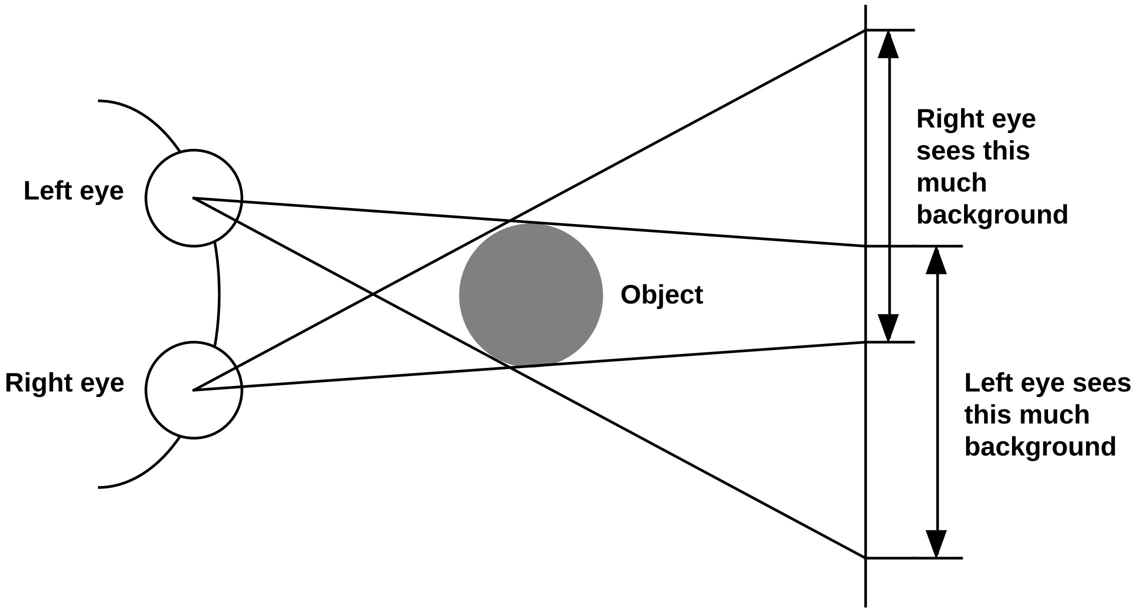


Fig 2.2. Leonardo's Paradox: Leonardo suggested that, because all the background can be seen by the overlapping views from each eye, the object should appear to be transparent.

One eye, or two?

Prior to 1833 — when the British scientist, Charles Wheatstone (1802-1875) made the first stereoscope — there seems to have been little understanding of the fact that the eyes act together and it was thought either that the vision of one eye suppressed the vision of the other or, as Giambattista della Porta believed, that vision alternated between the two eyes. Remarkably, this alternation of vision idea lasted as a possible explanation of stereopsis well into this century. It was shown however that, because stereoscopic perception can occur with very short presentations as in scenes lit by lightning flashes for instance, the alternation would have to be very rapid if we were not to be aware of it. No known biological processes could alternate fast enough to satisfy this requirement.

The Stereoscope The invention of the Wheatstone stereoscope, which was a laboratory instrument, moved the study of binocular perception along greatly but it wasn't until after the 1850s that stereoscopy caught the popular imagination.

The device we currently think of as a stereoscope — handheld and with a viewing hood — was invented in about 1860 by the American author, physician, and founder of the magazine, *Atlantic Monthly*, Oliver Wendell Holmes (1809-1894). Holmes was a great protagonist for stereoscopy and, at some financial loss, did not patent his invention because he wanted humankind in general to benefit from it. He seems to have admired the stereoscopic pictures of the day (of which he had thousands) more than the work of painters:

> The first effect of looking at a good photograph through a stereoscope is a surprise such as no painting ever produced . . . A painter shows us masses; the stereoscopic figure spares us nothing, — all must be there, every stick, straw, scratch, as faithfully as the dome of St Peters . . . (Holmes 1859 p743)

In the period 1859-1863, he wrote about the stereoscope in the *Atlantic Monthly* on three occasions (Holmes 1859, 1861, 1863) and had the same exaggerated hopes for the potential of the device to re-create realities as do some of today's enthusiasts for computer-based virtual reality:

> Form is henceforth divorced from matter. In fact matter as a visible object is of no great use any longer, except as the mould on which form is shaped. Give us a few negatives of a thing worth seeing, taken from different points of view, and that is all we want of it. Pull it down or burn it up, if you please . . . (Holmes 1859 p747)

In two of his articles (1859 and 1861) he also extols the virtue of stereography as a form of surrogate travel:

> I stroll through Rhenish vineyards, I sit under Roman arches, I walk the streets of once buried cities, I look into the chasms of Alpine glaciers, and on the rush of wasteful cataracts . . . The time will come when a man who wishes to see any object, natural or artificial, will go to the Imperial, National or City Stereographic Library and call for its skin or form, as he would for a book in any common library. (Holmes 1859 p746 and p748)

The general Victorian enthusiasm for stereoscopic images declined sharply — perhaps due to the rise of the even more exciting, and 'veridical' moving imagery of the cinema. Except for a short resurgence of interest in the 1950s and 1960s and some experimental work for television and computer graphics more recently, stereoscopy for entertainment now seems to have died out as a central concern of image makers.

Binocular disparity and convergence

Stereoscopic vision relies on two factors: binocular disparity and convergence (or to use the term referring to the more general ability to move our eyes, ‘vergence’. **Binocular disparity** is the name given to the fact that, because our eyes are set about 6 - 7cm apart, our retinas receive slightly different, but considerably overlapping, images. The overlapping of the views is a characteristic of the higher animals. In other species — where the eyes are not frontally set — the amount of overlap is minimised and two images seem to be used not for stereopsis but more to give a wide, panoramic view. Joseph Harris (1702-1764), in his posthumously published *Treatise of Opticks* (1775) was perhaps the first to suggest that the two disparate images we receive in binocular vision might be used to produce depth although, from the evidence of a number of his writings, Descartes (1596-1650) was clearly aware that the separate images from the two eyes were fused to create a single image. He believed that this fusion took place in the pineal gland — the ‘seat of the mind’ (Blakemore 1990, Pastore 1971).

Convergence The eyes can converge on points in the distance and we have to change focus to do this (a process known as **accommodation**). Theoretically, these two mechanisms could give enough information for us to calculate distances in the manner of a rangefinder. Descartes believed that ‘natural geometry’ was used to do just this but Berkeley, who conceded that convergence was a useful means for depth perception, rejected the idea of ‘natural geometry’. Just being aware of the differences in angular disparity between two objects would often be sufficient for us to determine which was in front of which. But Berkeley concluded that the necessary lines and angles for calculation purposes ‘. . . are themselves not at all perceived’, so he thought that it was the *sensation* of convergence and accommodation that gave us the cue.

Stereoscopic vision, though, does not appear to depend on the rangefinder effect. This can be confirmed by the fact that it is possible to see stereoscopic images that have been stabilised on the retinas (by briefly flashing a bright image into each eye). In this case, such things as convergence, focus and angle calculations can play no part. However, convergence and accommodation can be a factor in distance perception up to about 2 m.

The accuracy of binocular vision

Some suggest that the information we receive from stereoscopic vision is extremely accurate although, surprisingly, there is dispute on this:

> Stereoscopic depth discrimination is exquisitely sensitive: under optimal conditions, disparities as small as 2 sec of arc can be detected. This would correspond, for instance, to a depth difference of less than 0.05 mm at 50 cm, or 4 mm at 5 m. This very high stereoacuity is an example of the high positional accuracy that the visual system also demonstrates . . . implying that retinal spatial information is transmitted for stereoscopic comparisons with all the precision available (Braddick 1982 p194)
>
> Binocular distance perception is not veridical. A constant binocular disparity corresponds neither to a constant perceived depth nor to a constant perceived distance ratio (Foley 1980 p411)
>
> We know from our own experience and every-day observations that there exists a remarkably exact visual information both in man and animals about absolute distances, especially in the near space of the organisms . . . (Johansson 1973 p135)
>
> Though stereoscopic vision is often regarded as a means of gauging distance, human stereo depth judgments depart strikingly from the predictions of geometry. Not only are judgments of absolute distance highly inaccurate . . ., but also the perceived distance of an object is strongly affected by other objects nearby . . . , with the result that the relative distance is often incorrectly estimated (Mitchison & Westheimer 1984 p301)

The contradictory views on this, presumably easily testable, aspect of human vision leads one to suspect that the authors are, in fact, addressing different phenomena — that, despite the use of similar terminology, they are not talking about the same things.

The development of stereoscopic vision

On the other hand there is reasonable agreement on the view that stereoscopic vision is good for detecting local protuberances and depressions. Stereoscopic vision could there fore help us to separate objects from very close backgrounds and could have evolved for the purposes of ‘camouflage breaking’:

> Stereopsis may have evolved to aid the primary task of image analysis by defining those edges invisible to monocular inspection but derivable from retinal disparity cues.The emergence of stereopsis seems more likely to be related to three other aspects in which binocular parallax does better than monocular parallax . . .
> (1) absence of the need for the predator to move . . .
> (2) absence of the need for the prey to move . . .
> (3) lower susceptibility to photonic noise . . . (Pettigrew 1990 p288)

Whatever the accuracy of stereoscopic vision, Pettigrew, who has made an extensive study of the phenomenon in birds, tells us that:

> Improved precision *per se* is unlikely to provide a good account of the evolutionary origins of stereopsis . . . many primates and owls supplement binocular depth judgments with monocular parallax by the use of 'head bobbing', so it is hard to argue for the superiority of the binocular system on the grounds of precision alone (Pettigrew 1990 p287).

The binocular effect reduces as distances become larger although, once again, there is little agreement in texts as to the point at which binocular vision ceases to be effective — estimates vary between a few metres to as much as 130 metres and beyond.

> Stereopsis is not just a near vision sense . . . It is useful at considerable distances (Gillam 1990 p45).

An ecological view of depth perception

Warren (1995), also comments on our relative inability to make metrical judgments on depth information and gives us a Gibsonian ecological interpretation of this:

> There is long-standing evidence that perceptual judgments of distance, size, and shape are markedly inaccurate and non-

> Euclidean . . . However, metric information may be unnecessary for many tasks ranging from obstacle avoidance to object recognition, and metric tasks such as reaching could be governed by specific visuomotor mappings to which such perceptual ‘distortions’ are transparent. On this view, the goal of perception is not to recover a general-purpose scene description but to extract *task-specific information* for the activity at hand (Warren 1995 p264. His italics)

This view is essentially an *ecological* one: that visual perception is to enable us to get along in the world and that laboratory-type tests of its component parts will inevitably throw up anomalies:

> . . . the human visual system deals much better with natural images and multiple depth cues than with single depth cues in synthetic images (Bülthoff and Mallot 1990 p119).

The ecological approach to perception was spelled out in detail by JJ Gibson (1904-1979), the American experimental psychologist, mainly in three still influential and, in their time, somewhat controversial books: *The Perception of the Visual World* (1950), *The Senses considered as Perceptual Systems* (1966), and *The Ecological Approach to Visual Perception* (1979). The ecological approach of Gibson is often presented as uniquely original. It was certainly ground-breaking but, as Looren de Jong (1995) points out, some its essential features can be found in the work of two Austrian-American psychologists, Egon Brunswik (1903-1955) and Fritz Heider (1896-1988). Heider believed that:

> The performance of the perceptual system is to a great extent determined by the structure of the environment.

Stereoscopic vision generally

Under normal viewing conditions almost identical images fall on corresponding retinal areas of our two eyes. For binocular vision to work, however, our perceptual mechanisms must somehow be able to determine that two disparate image points on the left and right retinas refer to the same point in space and it is not easy to see how this happens. And of course, images which are too disparate do not produce depth. Reimann and Haken (1994) suggest a process that might model the capability we have for relating the disparate images.

From the evidence of tests originally carried out by Julesz (1960, 1971), however, it is clear that it is not necessary for the disparate points to be part of a meaningful scene — where context might help the decision process. Julesz used stereo pairs made up of a dense field of identical random dot images except that, in one of the pairs, a central square of dots was moved very slightly sideways to produce disparity. He showed that those viewing the stereo pair see the central square to be in front of the background (or behind it, if the images were reversed to the eyes). Stereopsis is the only mechanism that can be called into play for this to happen. Further, it confirms that the perceptual system can deal with minute individual differences in an image and yet detect an overall pattern.

A strangely anomalous effect however is that when two gratings set at an angle to one another are presented one to each eye, moiré fringes are not seen — as they would be if the gratings were either seen together simultaneously by both eyes or just by one eye alone (Spillman 1993). This seems to imply that we have the ability to fuse two disparate images for stereoscopic purposes but not for more general functions. Although he does not mention the moiré pattern anomaly, de Weert (1984) gives several examples of phenomena that are either purely monocular or purely binocular.

Specialised cells

More than twenty-five years ago Barlow, Blakemore and Pettigrew (1967) and Hubel and Wiesel (1970) identified certain neurons in the brains of cats and monkeys that are activated only if there is input from *both* eyes so there is some evidence in these cases for there being specialised depth perception cells but, in general, the processes of binocular depth perception are still a mystery. It is also not clear that humans have these specialised cells although it is likely that we do:

> It is generally thought that binocularly driven neurons in the primary visual cortex form the first stage for the processing of

> binocular disparity, because cells at lower levels in the visual pathway (ie the retina and the lateral geniculate nucleus) are not driven by stimulation through both eyes. However, the particular neural mechanisms that underlie depth discrimination are not well understood (DeAngelis et al 1995).

Vision defects

It should be noted that about 4-5% of the population cannot see stereoscopically even though they have otherwise normal vision. If we add to this the number who are, for example, monocular, we find that a surprisingly large number of the population cannot see stereoscopically at all.

Motion parallax and motion perspective

As we move our heads, objects in the distance appear to move in different directions and at different speeds— an effect known as motion parallax. The amount of movement of an object is directly proportional to the velocity of movement and inversely proportional to the square of the distance of the object from the observer. This provides excellent cues for depth perception. Indeed motion, either of ourselves or of the objects we are viewing, *could* be the major source of information on depth and distance. Helmholtz (1821-1894) says that a man in a forest finds that:

> it is impossible to distinguish, except vaguely and roughly, . . . what belongs to one tree and what to another . . . But the moment he begins to move forward, everything disentangles itself, and immediately he gets an experience of the material contents of the woods and their relations to each other in space, just as if he were looking at a good stereoscopic view of it.

The ‘forest effect’ is very clearly seen in cases such as shown in Figure 2. 3 where the overlapping edges are not detected until there is relative movement between the textures (Kaplan 1969). Experiments have been performed, however, that suggest that, whilst observers might be aware of information from motion parallax, they do not (or perhaps cannot) always use this information to determine depth and distance (Gibson et al 1959; Gogel and Tietz 1980).

Fig 2.3. This figure consists of two overlapping rectangles but one cannot see the edge of the overlaying rectangle until it moves.

Gibson took the view that it is not just motion parallax but the more general concept of *motion perspective* that is the primary factor in our perception of depth:

> We perceive in order to move, but we must also move in order to perceive (Gibson 1979 p223)

Motion perspective is the continuous change in the way objects — especially collections of objects — look as an observer moves about. One change in particular can be noted, that of the texture of the surfaces of the objects and the ground on which they stand — the **texture gradient.**

> The change is from motion in one direction through zero to motion in the opposite direction, and it also has a vanishing line, at right angles to the gradient, which passes through the centre of clear vision (Gibson 1950 p140).

As the ground stretches before us we are acutely aware of the way in which the surface texture becomes more and more dense as it recedes into the distance and that this happens regardless of the scale or orientation of the elements that compose the texture.

We relate objects in the distance to this texture and, hence, are able to judge both their distance and size from what we see. Furthermore, it is from the gradient of the texture that we determine the angle of the surface: a horizontal surface has one texture gradient, a vertical surface another, and a surface tilting away from us yet another (Figure 2. 4). As we move over undulating surfaces their gradient changes and we are able to understand them from this fact alone.

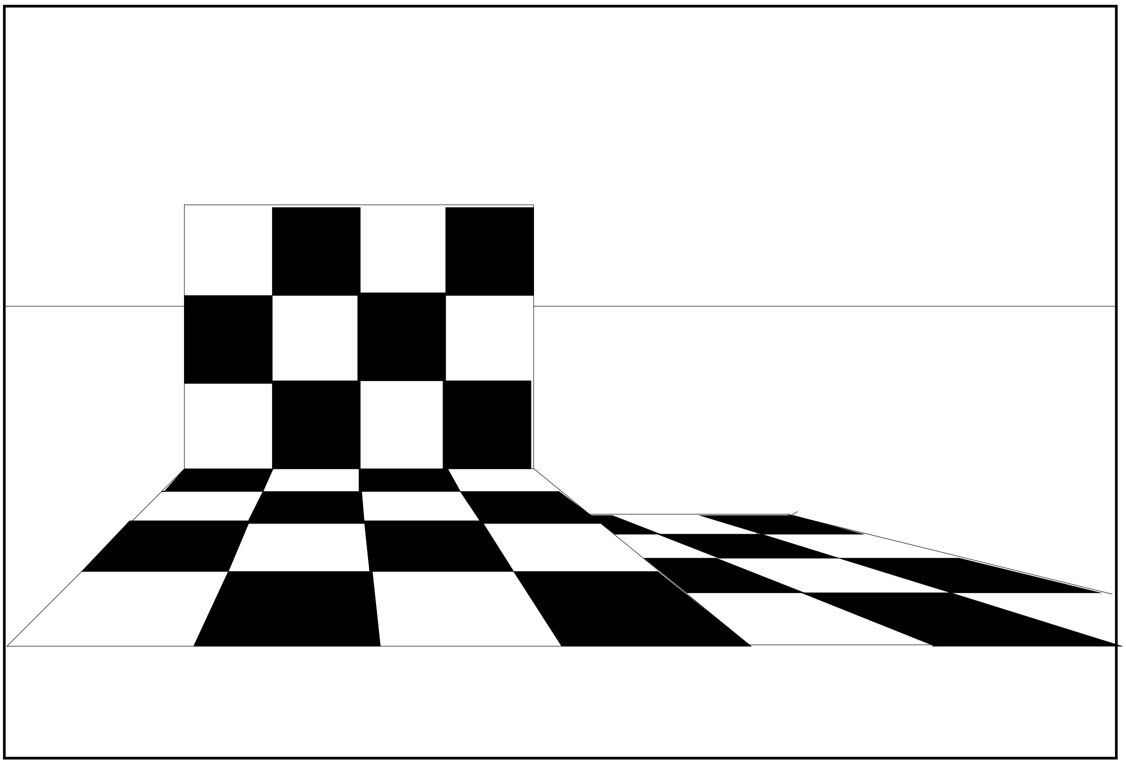


Figure 2.4. The variation in the appearance of texture according to the slope of a surface.

The later works of Gibson (1966, 1970) particularly Gibson (1979) stress the importance of motion perspective above all else as a cue to depth perception. Others take the view that the role of movement in perception, while important, could be exaggerated:

> . . . we think the role of motion in our perception of layout [the authors' preferred term for depth / distance perception] may have been overplayed . . . we would expect motion information to be relatively important for the perception of layout, but not all encompassing (Cutting and Vishton 1995 p89).

Superposition, overlay and partial occlusion

One of the elements we use in depth perception is our knowledge of the fact that objects that are nearer partially obscure objects that are farther away. As Boring (1942) has pointed out, this long-understood principle has been used to indicate depth in the earliest known drawings. It also, of course, plays its part in linear perspective. Undoubtedly, the information we receive from overlaying is used to convey an understanding of depth but it is not without its problems. Gibson (1950 p142) suggests that:

> the phenomenon of the superposition of objects is actually not a clue to the depth of objects but a perception which requires explanation. A man knows that a near object can partially obscure a far object but his retina does not, and a retinal explanation should be sought first.

He proposes that the contours of partially occluded objects are different from those that are not occluded — something that Helmholtz also thought significant when he said:

> . . . the mere fact that the contour line of the covering object does not change its direction where it joins the contour of the one behind it . . . [will generally enable us to decide which object is in front of which].

Helmholtz's remarks on contours and the 'forest effect' of motion perspective rely essentially on Gestalt principles. The first on the concept of 'good continuation' and the second on the fact that 'elements which move together are grouped together'. These are things that the later Gestalt psychologists made much of as an important principles of perceptual organisation (see Part 3 for a description of these principles).

Overlay is a very significant cue to relative depth and works over virtually any distance. It also seems to be an effective cue in the youngest of children (Craton and Yonas 1990).

Linear and aerial perspective

Leonardo writes of:

. . . three branches of perspective; the first deals with the reasons of the (apparent) diminution of objects as they recede from the eye, . . ; the second contains the way in which colours vary as they recede from the eye; the third and last explains how objects should appear less distinct in proportion as they are more remote. (Richter 1952 pp118-119).

Although Leonardo was writing about painting, perceptually these 'branches' (linear and aerial perspective) afford us further means for judging depth and distance. Later, in discussing aerial perspective, Leonardo goes on to say that ‘If one [object] is to be five times as distant, make it five times bluer’ (Boring 1942 p267). Exactly what being ‘five times bluer’ means, however, is difficult to deduce. He was right when he suggests elsewhere in his notebooks that height diminishes linearly with distance (linear perspective) but it is not clear that the variation of colour with distance (aerial perspective) is a linear property. Admittedly, Foley et al (1990 p727), dealing with approximate modelling of atmospheric attenuation in computer graphics by means of depth cueing, do suggest a linear relationship between a near colour and the same colour far off but Schacter (1983), discussing rendering for flight simulators, gives an exponential relationship. Nishita et al (1993) tell us that light is scattered (and hence attenuated) by the atmosphere inversely proportional to the 4th power of its wavelength, Thus short wavelengths are very heavily attenuated and long wavelengths are relatively unaffected (but this relationship is not linear). Giving a slightly different slant, Cutting and Vishton (1995 p88) tell us that:

> Unlike all other sources of [perceptual] information, the effectiveness of aerial perspective increases with the logarithm of distance.

It has been suggested that when (perhaps mythical) prewar hikers from towns went for walks in mountain areas, they frequently grossly underestimated distance to their target peaks. They did this because they were used to judging how far away things were in the smoky atmospheres of towns. When they came to the clear air of the countryside things seemed much nearer than they actually were. Higashayami and Shimono (1994), on the other hand, show that we tend consistently to overestimate the distance of far away (1-3 km or so) terrestrial objects. They found that this overestimation is proportional to the actual distance — so that the further away an object is, the more we overestimate its distance. To extent to which the observers in the Higashayami and Shimono study, who viewed the objects across water, were subject to the effects of aerial perspective is not clear.

Aerial perspective, was not of course invented by Leonardo. He was describing a well-established and universal technique that has featured in painting since Roman times. In some Japanese and Chinese paintings — and, later, in the works of JMW Turner (1775-1851) in particular — aerial perspective is often the only depth clue used (although, of course, in many Oriental works, overlay also features prominently). Arnheim (1956), though, suggests that aerial perspective is effective in paintings not because it imitates reality but ‘because it produces a perceptual gradient’ (p224). Strangely, Margaret Hagen (1986) in her thorough investigation of realism in painting does not touch on aerial perspective. Perhaps to be expected in the light of her book’s subtitle, she discusses Japanese painting only in terms of its projective methods although she does illustrate the ‘cloud convention’. This convention allowed artists to include isolated elements in their paintings as if seen through a clearing in the mist. The technique was used as a form of narrative.

On the face of it Leonardo outlines a rather strange situation as his example of where aerial perspective is important:

> when we see several buildings beyond a wall, all of which are seen above the top of the wall look to be the same size; . . you should make the building which is nearest above the wall of its natural colour, but make the more distant ones less defined and bluer . . .’ (Richter 1952 pp139-140).

The fact that he chose this example — where the ground texture was invisible — suggests that he was also aware of the defining role of two other things in depth perception and rendering depth in paintings. One was the texture gradient of the ground plane and the other, **the relationships of objects to the horizon**.

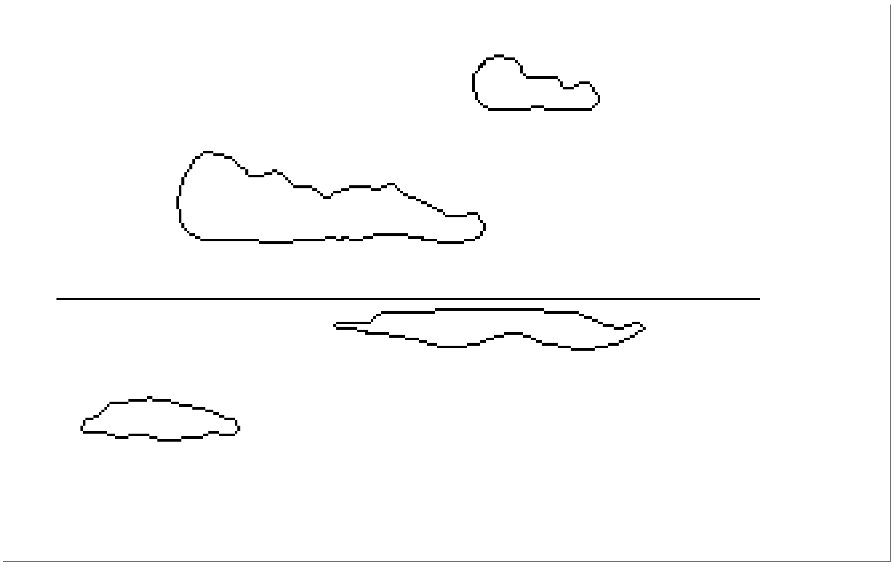


Fig 2. 5. In this rough sketch of a landscape the top cloud, though smaller than the lower one, is perceived to be nearer to us because of its greater height above the horizon line. Similarly the bottom pond, though smaller than the upper one, is perceived to be nearer than it because of its greater depth below the horizon line. We can use this cue when other linear or aerial perspective cues are missing (although this is a rare occurrence).

The height of objects above or below the horizon is an important cue to relative distance, something that, in part, was known to Euclid. Significantly, this cue can often be used to tell whether one object is nearer than another regardless of their relative sizes (Figure 2. 6). Artists have exploited this cue for centuries and it is particularly evident in some classical landscape paintings (see Plate 1. 1 [Koninck] and Plate 2. 2 [Corot] although, in these examples, other cues also exist).

In their excellent and comprehensive review of factors affecting depth perception, Cutting and Vishton (1995) follow Euclid and make the assumption that it is only to objects below the horizon that this cue applies. They also suggest that objects must be touching the ground plane for it to be useful. As examples in Figure 2. 5, and Plates 2. 1 and 2. 2 show, this is not necessarily the case.

The horizon-ratio relation

Sedgewick (1973, 1980) identified the relationship between objects and horizon line as a significant cue to depth and distance perception. One of the relationships he dealt with was the horizon-ratio relation. This states that objects of the same height are intersected in the same proportion by the horizon (Figure 2. 6).

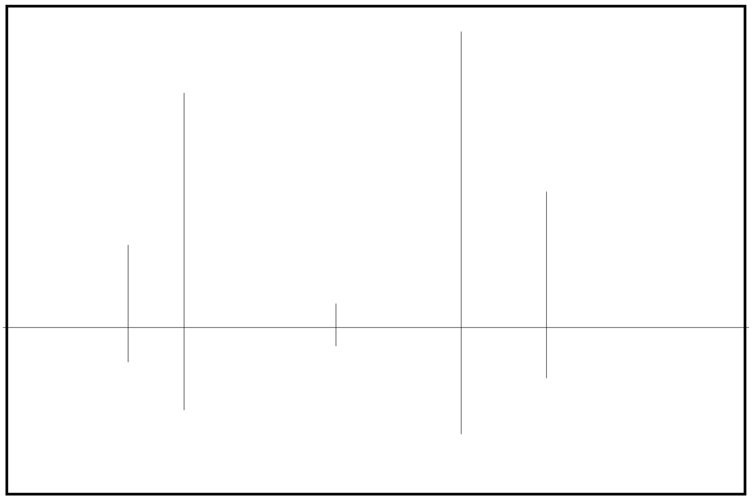


Fig 2. 6. The horizon cuts the randomly placed equal-height poles in the same proportion regardless of distance.

Tests carried out by Rogers (1985) suggest that, in drawings, we can very accurately judge relative distance of objects using the horizon-ratio relationship and that this result is largely independent of a subject's training in drawing. My own *ad hoc* tests with non-artists, however, have not convinced me that we make use of the horizon-ratio in drawings and I am not at all clear that we use it in real scenes even, say, at the seaside where the horizon is usually unambiguously revealed.

Retinal image size

The fact that the projection of the image of an object onto the retina is directly proportional to the object's size and inversely proportional to its distance is a clue to depth. Thus, given two objects of the same size, the retinal image of the one at distance X will be twice the size of the image of the object at distance 2X. Alternatively, given two objects at the same distance, the retinal image of the object of size 2X will be twice the size of the image of the object of size X. However, in order to process the retinal information on, say, two unfamiliar objects whose image sizes differ, we either have to guess something about their sizes or make some estimate of the distance each object is away from us! It is therefore hard to agree with Ogle (1962), who, in a primary work of reference, says:

> Despite the lack of certainty, one may suggest that among the . . [various perceptual] clues the size of the retinal image is the most important . . .

If this is true at all, it surely can only be the case where we are familiar with the sizes of objects in the scene being viewed. This may, of course, be the most common situation but we discuss later the problems that this gives rise to in traffic accidents to children.

Shadows and shading

Under some conditions shadows, or rather the relationship of objects to their shadows, give a cue to depth (Figure 2. 7). In contrast — despite playing a significant role in artists' renderings of objects — there seems to be some doubt on the value of shading in helping us to recover depth or even shape information in real scenes.

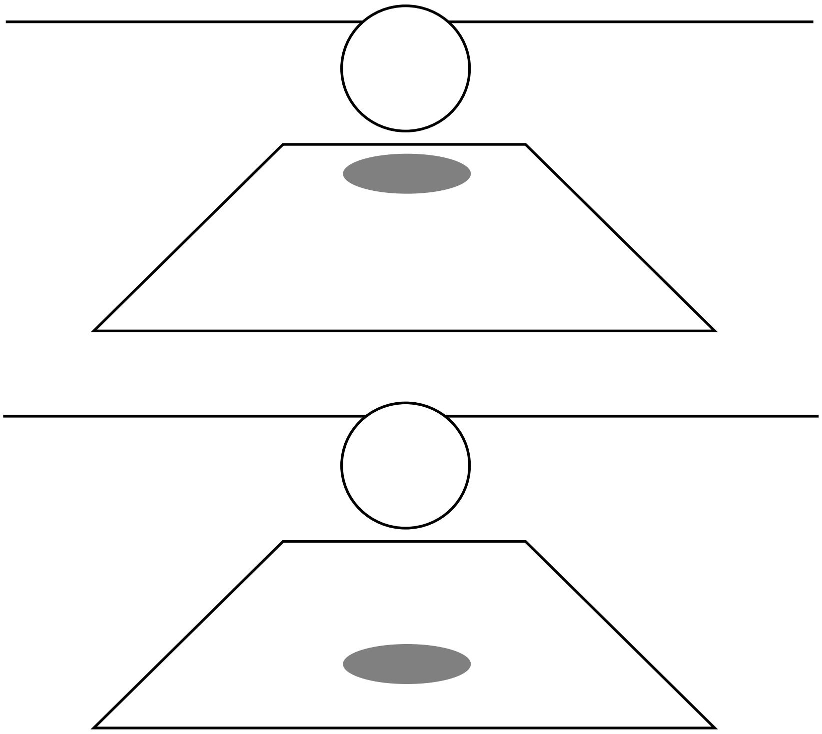


Fig 2. 7. Shadows help to determine height and distance as can be seen when a cast shadow is added in different positions to a fixed drawing

Todd and Mingola (1983) show that substantial errors in shading of images can be introduced before we are fooled about an object's true shape. Kleffner and Ramachandran (1992), however, report that the extraction of 3-D shape from shading is a compelling phenomenon but one that is poorly understood. They suggest that when we recover shape from from shading information we make an important simplifying assumption. This is that the scene is lit by a single light source coming from above. Strangely, 'above' in this context applies to a coordinate system centred on the eye rather than the scene itself so that, if we tilt our heads, we have difficulty in extracting shape unless the source of the light moves in sympathy. They also report a slight preference for convexity rather than concavity in our interpretation of shading (something we discussed when dealing with Andrea Pozzo's ceiling painting).

Leonardo defined two types of shadow: the primary shadow which is attached to the object and helps us determine or delineate its shape; and the cast shadow which is disengaged from the object and helps us (as in Figure 2. 7) to determine the object's height and distance. In his practical hints to artists on light and shade, Leonardo also had something important to say about the direction of lighting:

> Above all, see that the figures you paint are broadly lighted and from above . . . for you will see that all the people you meet out in the street are lighted from above, and you must know that if you saw your most intimate friend with a light [on his face] from below you would find it difficult to recognise him (Richter 1883/1970 p331)

Because of the strange quality it imparts, the technique of lighting upwards is often exploited in horror films to add to the impact of frightening scenes. Moreover, buildings floodlit from below often look unreal.

Illusory contours and brightness

Related to shading is the effect that, in our perception of the 3-dimensional world, brightly lit surfaces appear to be nearer than darker ones. This could be due to the fact that more light allows us to see detail on these surfaces more clearly. If that were the case, Leonardo’s branch of perspective where

> the third and last explains how objects should appear less distinct in proportion as they are more remote

would thus come into play giving the formula:

clearer detail = closer.

Does this effect works in reverse too? Can something that is *thought* to be nearer be seen as being brighter than its surround? These questions are prompted by an effect that is very clearly illustrated by some false contour illusions, such at that in Figure 2. 8.

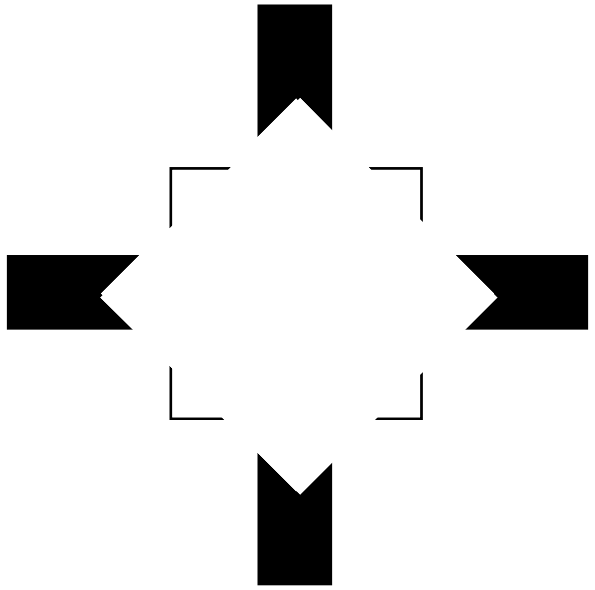


Fig 2. 8. Illusory contour where the central diagonal square seems brighter than its surround

The diagonally arranged square that we see in the centre seems to be brighter than the background although it cannot possibly be. One explanation for this phenomenon is that, because we think the square is occluding the background we assume it is nearer and, hence, must be brighter. Parks and Rock (1990) show the well-known Kanisza triangle illusion in various 3-dimensional forms and use these examples to justify this explanation for apparent brightness. More recently Purghé (1993) has given ingenious counterexamples of the effect where extra brightness is apparent but where occlusion could not be the explanation. Exactly what is happening, however, is not at all clear and Purghé suggests that:

> An occlusion clue may be, . . , a factor affecting but not determining the emergence of illusory figures (Purghé 1993 p816)

Spillman and Dresp (1995 p1336) concur:

> Although brightness enhancement and illusory contours usually occur together, there are example when they do not. Under certain conditions, illusory contours may emerge without enhanced brightness and, conversely, brightness enhancement may exist in the absence of illusory borders.

As a further complication apparent brightness is displayed in the Ehrenstein illusion (Figure 2. 9a) when there are no contours present but is absent when we actually include the contours (Figure 2. 9b). This could be because we see Figure 2. 9a as a 3-dimensional representation but Figure 2. 9b as a 2-dimensional one but why this should be is a mystery.

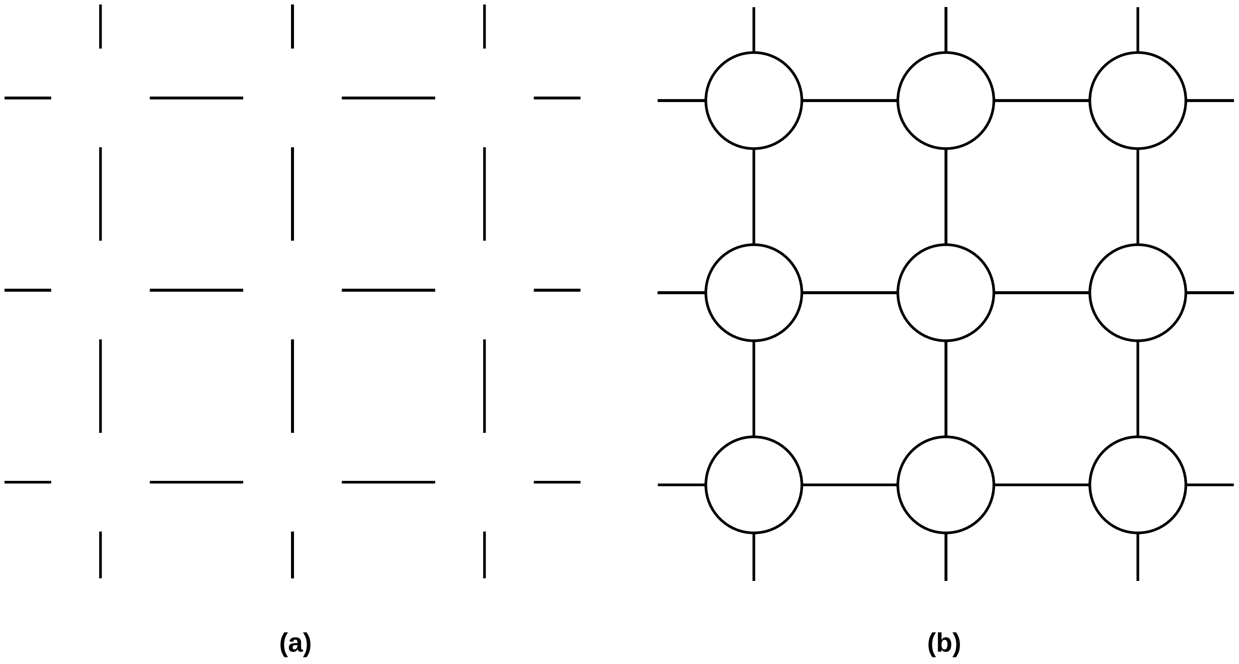


Fig 2. 9. The circles in (a) look brighter than the background whereas the circles in (b) do not.

The importance of multiple views

Although stereopsis and motion parallax (and motion perspective) are, on the face of it, different phenonema, they have one thing in common: they present us with more than one view of a scene. Binocular disparity gives us two slightly different views simultaneously: motion parallax gives us two or more different views consecutively. Thus it is probably safe to say that our most significant information for depth perception comes from the fact that we are receiving, over time, multiple views of a scene and that in these multiple views many perceptual cues are present. It is this richness of information that enables Helmholtz's man in the forest to gain his 'apperception of the material contents of the woods' . It is also the reason that Holmes, above, says:

> Give us *a few* negatives of a thing worth seeing, *taken from different points of view,* and that is all we want of it. (My emphasis)

The need for different points of view is paramount to our understanding. For example, most of us have only seen from photographs just one view of the Taj Mahal (1632-1654) and that is the one from the front garden across the formal pool (only part of which we are normally shown). We cannot know from this view the relationship of the building to its surroundings, particularly to the river Jumna at the rear. Nor can we tell that the mausoleum is completely symmetrical in plan (although we might guess that this is the case) and that it is flanked by two smaller buildings.

Integrating the information

In laboratory conditions it is possible to isolate and study each of the perceptual depth cues separately. In the everyday situation, however, we encounter many of the cues simultaneously and need to integrate the information from each to provide us with an overall view of the scene. Sometimes — indeed perhaps more often than we are conscious of — contradictory information is provided by the cues, and sometimes this confuses us. In normal circumstances, this confusion might not matter because in the next few moments further information arrives to allow us to disambiguate the percepts. At other times, events may be happening too fast for us to benefit from later information and problems may arise — as in the case of air and traffic accidents.

Perceptual problems and accidents Human factors problems generally are now thought to be responsible for about 90% of air accidents. Such problems are due to perceptual, cognitive or biomechanical difficulties experienced by aircrew (Haber 1987). Stewart et al (1993) suggest that the relatively high number of pedestrian accidents to children is not due, as was previously thought, mainly to children's' inexperience or recklessness. They identify as a more likely cause the fact that drivers misperceive the time to impact and hence do not stop quickly enough. This misperception happens because they overestimate the height of the child in front of them and thus their distance from him or her. (A short child close to subtends the same visual angle as an adult or taller child further away, see Figure 2. 10). Stewart et al claim that this distance/speed misperception is the cause of 50% of traffic accidents to children.

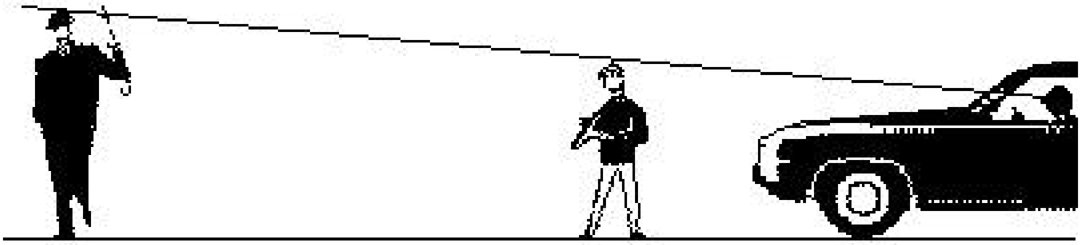


Fig 2. 10. A child close to subtends the same visual angle as an adult further away.

Is the world like we see it?

It is not surprising that Wagner (1985 p493) tells us that:

> In general, the visual world approaches the Euclidean ideal of veridical perception as the quantity and quality of perceptual information increases.

However, in examining the way in which stereoscopy, shading and texture interact to give us depth perception, Bülthoff and Mallot (1990) found that:

> Depth information can be collected from different cues and performance should improve as more information becomes available. Our data show that it is not the reliability which

> improves, but the perceived depth which increases (p141).

They suggest that, where ambiguity exists, some cues veto or inhibit the information from others. They further show that, rather than being separate sources of information, stereopsis, shading and texture are closely linked. More than this, they suggest that:

> Integration of cues is one of the key features of vision that underlie its performance and robustness (Bülthoff and Mallot 1990 p119)

Summary

Although it is true that much has still to be learned about depth perception it is clear that a great many visual mechanisms are called into play in order for us to understand the three-dimensional world. What is surprising is that so few of them depend on our having two eyes.

3. Movement perception

The visual perception of movement is important to virtually every living species. Indeed even organisms without vision usually have sensors to detect movement. Obviously detection of movement plays a vital role in the survival of animals: they must be good at perceiving movement of predators and of likely prey. Inability to do this would result in disaster and it is often more important to detect immediately that *something* has moved rather than to know straight away what that something is (or even in which precise direction it has moved). Sekuler (1975) proposes that:

> During evolution, motion perception was probably shaped by selective pressures that were stronger and more direct than those shaping other aspects of vision. . . As a result of such selective pressures, our visual systems contain neural mechanisms specialised for the analysis of motion (p385).

The importance of movement perception

We have at least two needs for motion perception: one, to make sense of the world as we move through it — self-motion or ‘egomotion’— and, two, to understand objects that move about us. Much of the time, of course, both these needs come into play simultaneously. However, it is puzzling how easily we can normally distinguish between the movement of objects and our own movements although, as far as the retinal image is concerned, there does not seem enough information to do this. Note, too, that the retinal image is also altered by movements of the eyes or head. Harris, Freeman and Williams (1992) postulate perceptual mechanisms that would, if they exist, help to explain how we are able to separate out these different retinal effects. Cutting et al (1977) suggest that movement detected in parts of the retina at the periphery of vision allows us to perceive self-motion whereas the same stimulus at the centre of vision allows us perceive object motion.

In their comprehensive review of aspects of visual motion analysis from a computational perspective, Hildreth and Koch (1987) point out that:

> The pattern of movement in a changing image is not given to the visual system directly, but must be inferred from the changing intensities that reach the eye. The 3-D shape of object surfaces, the locations of object boundaries, and the movement of the observer relative to the scene can in turn be inferred from the pattern of image motion. Typically, the overall analysis of motion is divided into two stages: first, the measurement of movement in the changing 2-D image, and second, the use of motion measurements, for example to recover the 3-D layout of the environment. It is not clear whether motion analysis in biological systems is necessarily performed in two distinct stages, but this division has served to facilitate theoretical studies of motion analysis and to focus empirical questions for perceptual and physical studies (p480).

The need for good movement perception

Nakayama (1985) suggests seven possible reasons why movement perception is important. These are:

to enable us to

- derive the third dimension
- calculate time to collision
- distinguish figure from ground
- ascertain information about our own movement
- stimulate eye movements
- understand pattern
- perceive moving objects

Warren (1995) goes further and stresses the importance of movement and action to perception as a whole:

> Traditionally, the problems of perception and action have been treated as logically independent. It has been assumed that the goal of perception is to recover objective quantities such as size, distance, shape, colour and motion, yielding a general-purpose description of the scene that can provide the basis for any subsequent behaviour. Considering vision in the context of action has important implications for this view of perception (p264).

He goes on to show that some of the anomalies of perception that are apparent when we consider it in a still ‘snapshot’ way unrelated to action may not matter so much when we see perception as something that unfolds over time. Viewed thus:

> . . . judgments at any instant may be qualitative or even nonveridical, and yet over the course of the act adaptive behaviour emerges from the animal-environment interaction (p264).

All this points to the fact that the detection of motion is highly significant to us. It is all the more surprising therefore that research into this aspect of our visual perception mechanisms seems to lag behind work on other aspects. While colour vision and stereopsis have received the greatest attention by researchers, as Nakayama (1985) says,

> . . . it is clear that colour processing is not present in all species and that binocular vision is restricted in animals with laterally placed eyes. As such numerous animals either lack colour vision or significant binocular vision or both. No animals have been found that lack mechanisms for motion processing (p627).

Development of motion perception

Referring to the growth of space and movement perception in children, Kellman (1995) tells us,

> From the earliest ages, motion attracts attention, and infants orient toward moving stimuli by using head and eye movements . . . The causal direction of the connection between motion and information is not known. Infants might be hard-wired to attend to moving things — a useful adaptation for learning about objects and events. Alternatively, it may be information, not motion, that guides attention. Infants may preferentially attend to events more than static scenes because more or better information about spaces and objects is available to them from kinematic sources than from static ones (p346).

It appears from studies by Kaufmann (1995) that a child’s ability to detect very rapid motion is fully developed soon after birth whereas the ability to detect very slow motion seems to improve gradually with age. As adults we can detect 3-6 mm of movement per second of an object when it is 1 m away. However, it appears that we substantially underestimate the velocity of moving objects (Caelli 1981 pp145-171).

Ullman (1987) suggests that the importance of motion detection has led to particular physical results:

> In view of the central role of motion perception, it is not surprising that the analysis of visual motion is wired into the system from the earliest processing stages. In some species, including the pigeon and the rabbit, rudimentary motion analysis is performed as early as the retinal level (p1280).

We know too that, after recovery from damage to the human visual processing mechanisms, it is our perception of movement that returns first. Furthermore, unlike for some other aspects of vision, most of the retina is available for motion detection and the thresholds for discovery are more or less the same over the whole area (Krumhansl, 1984).

The mechanisms of movement detection

In view of the significance of motion perception and the different forms it has to take, it has been postulated that several mechanisms are at work. For example, Juola and Breitmeyer (1989) tell us that:

> The sheer variety of motion phenomenon that can be perceived argues for the necessity of having more than one detection system, perhaps located at different levels in the nervous system. It is not the case that an image moving across the retina is sufficient stimulus for the perception of movement. For example, the retinal image moves whenever a saccadic eye movement is executed, yet we are unaware of any motion. Conversely, we can track a moving object against a featureless without necessitating changes in the retinal image, yet a strong sense of movement results (p251).

As with other aspects of visual perception it is possible to locate specialised areas. Anderson (1989) points out that:

> A substantial body of evidence suggests that visual motion analysis is treated by specific brain regions. There are several accounts of brain lesions in humans that produced deficits in motion perception without deficits in other forms of vision (p383).

Anderson locates the ‘pinnacle of a prescribed hierarchy in motion processing’ in the posterior parietal cortex — Brodman areas 5 and 7. Zeki (1990) confirms that, in macaque monkeys:

> The ‘funnel’ for the motion pathways of the visual cortex is area V5 . . . All its cells are responsive to motion in the field of view

> and the overwhelming majority are directionally selective . . . The fact that none is wavelength selective and the majority are not orientation selective either . . . led me to propose that it is a cortical visual area specialised for the detection of motion in the field of view (pp321-322).

From their examination of healthy and brain-damaged patients, Schenk and Zihl (1995) conclude that there are two distinct mechanisms involved in motion detection. One of these is concerned with the analysis of global motion. The other deals with extracting form from motion.

In their discussion of human motion perception and the physiological nature of motion detectors, Burr and Ross (1986) point out that accurate perception of form in motion involves visual integration of a type that eliminates smear. Image motion does not produce the same problems for the eye as it does for the camera. People can clearly see objects in motion, and they can see motion on cinema or TV when what they are shown is actually a sequence of stills. They look at mechanisms that allow for understanding when there is real motion, and interpolation when motion is sampled as in cinema. They suggest that such visual abilities may be explained by specialisations of visual neurons.

In examining moving images of a Mondrian-like nature, Zeki and Lamb (1994) conclude three things:

- that an image of the visual world is not impressed upon the retina, but assembled together in the visual cortex;
- that separate attributes of the visual scene are processed in geographically separate parts of the visual cortex, before being combined to give a unified and coherent picture of the visual world;
- that the attributes that are separated, and separately processed, in the cerebral cortex are those that have primacy in vision. These are colour, form, motion, and possibly depth.

Using PET scanning DuPont et al (1994) confirm that several brain areas are involved in the perception of movement. In addition to sites on either side of the brain at the border between Brodmann areas 19 and 37, a V1/V2 focus and a focus in the cuneus, they observed activity in other visual areas in the cerebellum and in two presumed vestibular areas — at the posterior part of lateral sulcus and at the border of Brodmann Areas 2 and 40 (Figure 3. 1).

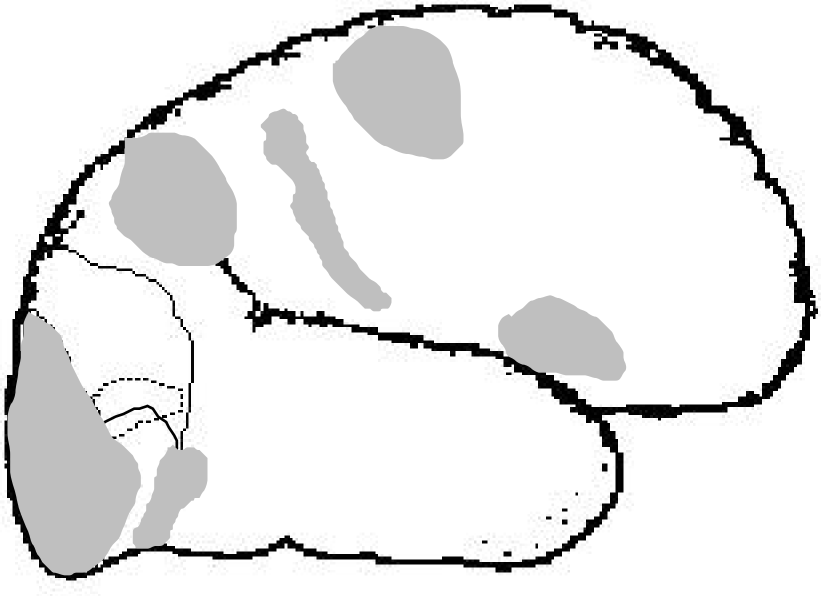


Fig 3. 1 Diagrammatic view of the brain areas identified by DuPont et al (1994) as being involved in the perception of movement

Early work

In their overview of motion detection research, Smith and Snowden (1994) suggest that the work of the Austrian scientist and colleague of Helmholtz, Siegmund Exner (1846-1926), provided the starting point in the latter part of the nineteenth century.

> He demonstrated that motion could be perceived from two stationary images (sparks of electricity in his case) presented in quick succession — a fact exploited by television and 'movies'. This had been known for some time previously . . . However, Exner's great insight was that this perception of movement could be elicited from two sparks that were so close together in space that they could not be distinguished. Under these conditions it seems impossible that the observer could infer (consciously or unconsciously) motion from a knowledge of position and time. It therefore follows that motion perception must be a sensation in its own right, not one derived from a sense of position and time (p5).

Previous to Exner's studies, others had commented on aspects of motion such as optical illusions but little detailed study had been carried out. However, the stroboscope — for examining objects in motion — had been in existence since the 1830s (see Boring, 1942 pp588-596). About that time, too, Robert Addams (1834) reported on a perceptual anomaly about motion that he had encountered on a trip to Scotland:

> During a recent tour through the Highlands of Scotland, I visited the celebrated Falls of Foyers on the border of Loch Ness, and there noticed the following phenomenon.
> Having steadfastly looked for a few seconds at a particular part of the cascade, . . . , and then suddenly directed my eyes to the left, to observe the vertical face of the sombre age-worn rocks immediately contiguous to the waterfall, I saw the rocky surface as if in motion upwards, and with an apparent velocity equal to that of the descending water, which the moment before had prepared my eyes to behold this singular deception.

In parenthesis it is worth noting the style of language used in this description of what has now become known as the 'waterfall illusion'. It was written for a scientific journal and yet the style does not differ markedly from the sort that was used a few short years later by, for example, John Ruskin (1819-1900), in his descriptions of natural phenomena in his books on art. The differences we now see between scientific reporting and writing on art seem to have come about in this century: the two cultures of CP Snow are clearly a modern concept.

Despite some earlier work and remarks on motion phenomena, it was not until the late nineteenth century that systematic studies of motion occurred. We should also remember that it was in the period 1870-1890 that the English photographer, Eadweard Muybridge (1830-1904) in the US and Etienne-Jules Maray (1830-1904) in France, carried out and published studies on the movement of humans and animals. As well as leading in the mid-1890s to the development of the motion picture, these studies must have also spurred an interest in the phenomenon of movement perception.

The optic array and optic flow

Many believe, with Gibson, that the major element in motion perception (and, perhaps, in visual perception generally) is the so-called ‘optic flow’ (Figure 3. 2).

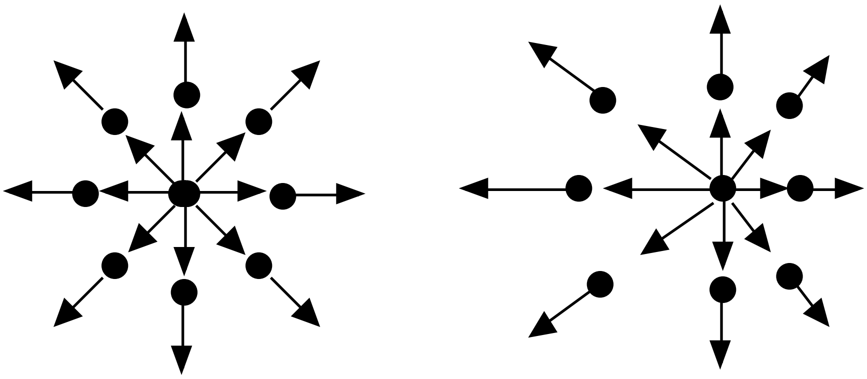


Fig 3.2 On the left is the diagrammatic situation when we approach a surface at right angles to us. On the right, when we approach a surface at an angle, with the left edge nearest to us

As Harris (1994) perceptively comments:

> As we move about and scan the visual world, the images of things move about, changing their relationships in a complex dance. JJ Gibson . . . was the first to understand this dance . . . Probably Gibson's greatest contribution was to redefine the dance-floor, emphasising the amount of information potentially available to an observer in the transforming optic array rather than the instantaneous fragments provided by a pair of retinal images.
>
> The optic array is the three-dimensional bundle of light rays that impinge from all directions upon each point in an illuminated world. Objects in the world can be thought of as labelling specific rays, so producing a global pattern of light intensities. A retinal image provides access to only part of the optic array at any one time, but a stationary observer can sample different parts by eye movements and head rotations. By changing position, the observer can sample the different optic arrays impinging on neighbouring points in space. However, sampling in this case should not be thought of as a discrete process. Rather, as the observer gradually moves, so each ray gradually moves, thus producing the smooth transformation in the optic array that Gibson called the optic flow (p307-308).

Gibson believed that we ‘picked up’ on the optic array and optic flow in order to understand the world. Unfortunately, he proposed no explanation as to how the ‘picking up’ might be achieved. Johansson (1994) agrees with Gibson and suggests that the principles for decoding the optic flow are ‘built-in to the specific visual systems of the species’ (p311). He outlines his approach to decoding the optic flow and what he calls the ‘optic sphere’ in Johansson and Börjesson (1989).

Warren and Hannon (1988) conclude from studies of computer screen images that optic flow is sufficient at least for perceiving the direction of self-motion.

Many workers have made proposals for how we might cognitively deal with optic flow. Harris, Freeman and Williams (1992), already referred to, develop a concept by Koenderink and van Doorn (1976) who treat optic flow as a vector field in which the operations of translation, div (divergence), curl (rotation) and def (deformation, shear) apply (Figure 3. 3).

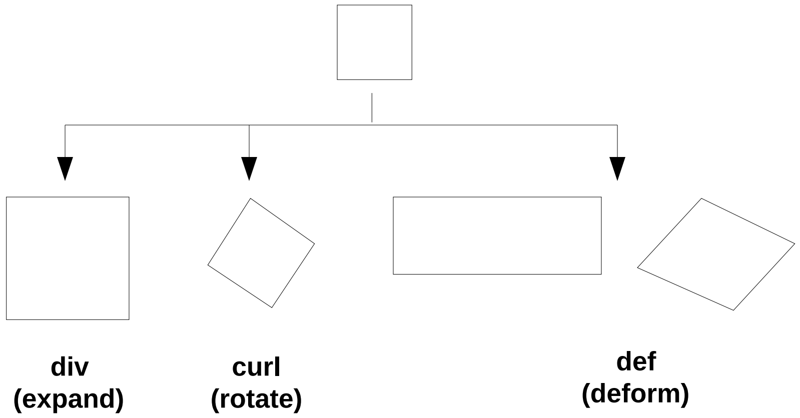


Fig 3. 3 showing the vector field operations of div, curl and def

De Bruyn and Orban (1990) give evidence to suggest that we are able to detect div, curl and def over much of the field of view.

Div, curl and translate into operations on the optic flow lines as in Figure 3. 4.

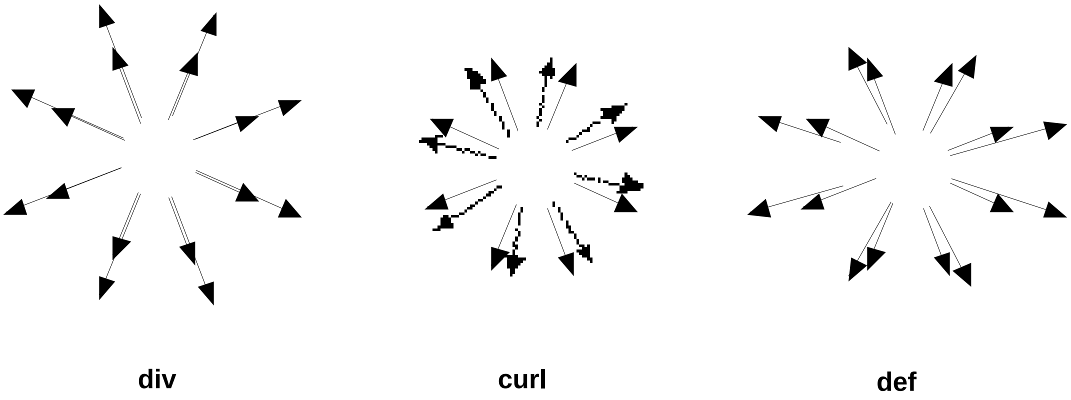


Fig 3. 4 showing how div, curl and def affect the optic vector field

Dodwell (1983), basing his work on that of Hoffman (1996), uses the vector field ideas of Lie algebra to show how, with appropriate perceptual mechanisms, we might simply process optic flow (see also Hoffman 1970; Hoffman and Dodwell 1985). However, Caelli (1981 p143) takes the view that Hoffman's formulation is a meta-language that allows us conveniently to talk about some aspects of perception rather than giving us an explanation of it. He suggests that lower level models are needed before we can understand the situation. Harris, Freeman and Williams (1992) try to provide such models.

Biological motion detection

A curious ability we possess was first demonstrated by Johansson (1973a) who showed that, if small lights are placed on the ankles, knees, hips, wrist and elbows of a black-clad actor, and a movie is made of the person walking in a dark room where only the lights are visible, we can easily understand what we are seeing. Conversely, we cannot understand what we are seeing if no movement takes place.

This phenomenon has been studied by a number of workers. Kozlowski and Cutting (1977) show that the gender of an actor can be deduced from the moving pattern of lights. Runeson and Frykholm (1983) show that people are remarkably accurate in guessing the distance that the black-clad actors are able to throw small sandbags even though unable to see the bags or anything more than the lightspots on the actors’ joints. Pavlova (1992) showed that quite young children are able to recognise the patterns — Berthenthal et al (1985) estimate that the ability to judge biological motion develops at around 6 to 9 months. Mather and West (1993) show that animals can also be recognised from this apparently limited information source. Sumi (1984) shows that, even if the film is run backwards and upside-down, subjects still recognise a human walker. What is more they perceive it as a walker with a peculiar gait rather than as an inverted image of a person moving backwards. Sumi concludes that this conception seemed to arise from the fact that the actor's arms were perceived as legs and vice versa. Dittrich (1993) shows that recognition of what is happening is more readily achieved when the figures are waking or running than carrying out social or instrumental actions. Further, that recognition is not greatly impaired if the lights are placed not on the joints but between them.

It is probably our ability to detect biological motion that explains the results of studies by Campbell (1979) who found that, in getting young children to understand still drawings of running and walking figures in implied motion, it was more important to show the figures in active postures than to employ motion cues like speed lines such as cartoonists sometimes use. Not surprisingly, Friedman and Stevenson (1975) found that older children were able to understand both postural cues and the cartoonists’ conventions. They too, however, seemed to favour active postures.

We return briefly to the phenomenon of biological motion detection in part 5 when we discuss perceptual organisation (p5- 13).

Movement in graphics and virtual reality

There is little doubt that movement helps in our understanding of objects. Frequently, drawings that make little sense when shown as stills are understandable when they are animated. This is particularly the case in viewing ‘wire-frame’ drawings. Wire frame drawings of three-dimensional scenes often contain too many lines for us to disambiguate when seen in still form. When animated to allow for panning across the scene and rotating objects, however, they spring to life and the third dimension appears effortlessly. Often, indeed, a wire-frame animation seems to convey more information about the form of an object than does a movie of the same thing (although, of course, it cannot tell us anything about the surface properties of the object).

The attractions of three-dimensional drawings, particularly animated three-dimensional drawings, are such that, sometimes, they are used inappropriately — for example, to illustrate two-dimensional data. When this happens not only do they *not* enhance the data set, they sometimes falsify it. We should also note that, although we seem to have separate mechanisms for dealing with motion, colour, and form (at least), it is clear that we can better understand subtle changes in a scene when we attend to one of these aspects only (Corbetta et al 1990). From this, and other studies, I conclude that, if we wish to have someone focus their attention on a particular element in an image or animated scene, then we should not create conflict by changing motion, colour and form together. This would be especially the case if abstract data is being shown.

Virtual Reality

The current resurgence of interest in virtual reality both for representing concrete and abstract forms of data founds much of its rationale on the exploratory appeal of the user being able to move about within virtual space. Total immersion virtual reality when headsets (and, sometimes, special suits) are worn goes further and aims to recreate fully the appearance of reality. Whilst this approach has its good points (although in its current manifestation does not remotely approach the appearance of actuality), it also has some drawbacks. Evidence is arising that disorientation and nausea can result from total immersion virtual reality. Regan (1995) reviews some of these problems.

Interestingly, in an admittedly slightly limited and uncontrolled study by Felix (1995), we find that there is no evidence that being, as it were, inside a virtual space and walking about it, as opposed to viewing it from outside from different viewpoints, necessarily improves understanding of it. In both cases, motion assists. But immersion in the space does not necessarily provide good understanding of it.

Summary

The detection of motion is a strong element in visual perception and we gain much information from it. Indeed, some believe motion to be the prime source of our understanding of the world. Sometimes, it is only through motion that we can disambiguate conflicting signals and it is likely that our ability to process optic flow is primal.

4. Colour perception

‘All men’, said Ruskin (1819 - 1900), ‘completely organised and justly tempered, enjoy colour; it is meant for the perpetual comfort and delight of the human heart . . .’ Although Ruskin himself appeared to be rather far from being ‘completely organised and justly tempered’ (see, for example, Abse 1980), one cannot but agree with his remark. His view that ‘colour is the most sacred element in all visible things’, however, is more difficult to accept. Colour has certainly exercised fascination throughout the ages and is often spoken of as if it had magic powers. Many, too, including Newton (1642 - 1727) himself, have tried to relate colours to musical notes (Caviano 1994; Pridmore 1992; Sebba 1991). Indeed Newton’s penchant for this idea seems to have made him abandon his earlier view that colours were infinitely graded and led him to name just seven colours of the spectrum. He could, of course, have chosen to name just three or four, or as many as ten or twenty colours.

Hue, lightness and saturation

We seem to be able to distinguish three features which contribute to our perception of colour: hue, lightness and saturation.

- **Hue** is the feature determined by the dominant wavelength of the light seen directly from a source, or indirectly from reflections off surfaces. It is the feature we use to give a colour its name — red (dominant wavelength about 680 nm), green (520 nm), yellow (580 nm) and so on.

- **Lightness** of a colour is its degree of perceived luminance relative to the luminance of another colour or the surroundings. We usually speak of the lightness of a surface colour but of the intensity or brightness of lights and CRT phosphors. Although, technically, these are different concepts (Agoston 1979), in daily life and in computing we tend to use the words synonymously.

- **Saturation** is the apparent purity of the hue. In the case of a surface colour, this is the degree to which the colour is undiluted by white. In the case of lights or phosphors, saturation depends on the relative amounts of luminous intensity held by the various wavelengths that make up the colour. In both cases, the more one wavelength dominates, the greater is the saturation. Black, grey and white, where no wavelength dominates, have the same saturation — they differ only in lightness.

The interdependence of hue, lightness and saturation

It should not be assumed, however, that hue, lightness and saturation are independent attributes. On the contrary, as has been known since the end of the last century, our perception of both hue and saturation changes as lightness changes — a phenomenon known as the Bezold-Brücke effect (Purdy 1931). In fact, it is only at three particular wavelengths that varying the lightness will leave our perception of hue unchanged. Roughly, these are at 470 nm (vivid blue), at 505 nm (green), and 572 nm (greenish-yellow). At all other wavelengths, we tend to see different hues as we change their lightness or saturation. Valberg et al (1991) suggest that the Bezold-Brücke effect arises because of the nonmonotonic response of retinal and LGN cells.

White light

In nature, white light is made up of equal radiance levels of all wavelengths — as Newton showed by prismatically splitting up sunlight into its spectral parts. However, white can also be produced from only three properly chosen colours (again of equal radiance levels). For example, a white approximately matching daylight at 6500 degrees Kelvin can be produced by mixing red, green and blue lights in the proportions 0.26, 0.66 and 0.8. By using complementary hues, only two colours need be used to make up white. For example, blue at 480 nm and yellow at 580 nm would do. The odd thing is that we seem unable to detect the differences in these whites although they are created in distinct ways and embody different wavelengths.

Distinguishable colours

Within the visible range of about 300 nm we can distinguish about 150 - 200 different hues providing we are allowed to compare one with another. (If not, less than a dozen are distinguishable using just our memory for colours). But note that we are talking here about hue. Well over 150,000 and perhaps as many as 2 million, different hues with their shades and tints can be detected if we are allowed to compare examples with one another.

Cowlishaw (1984) showed that 21 bits of colour are sufficient to encode images of natural scenes (which implies about 2.1 million colours). In some cases, as little as 17 bits of colour are sufficient (which implies about 131,000 colours). In both cases, however, it is necessary to encode green and red with more bits than blue. This is because our sensitivity to greens is greater than that for reds by about 1 bit and that for blues by about 2 bits. At some points in the spectrum these differences in sensitivity allow us to distinguish colours whose wavelengths differ by as little as 1 nm.

We are particularly good at *relative* judgment of brightness, where about 500 steps can be distinguished when comparisons are made. On the other hand — and perhaps surprisingly in view of our success are relative judgments — we are poor at making *absolute* judgments about brightness. Less than half a dozen steps of brightness can be remembered.

Colour receptors

Our ability to perceive colours comes from a combination of neural mechanisms and the characteristics of the cones, of which it is likely that we have three types (MacNichol 1964) — although recent investigations by Neitz, Neitz and Jacobs (1993) suggest that some aspects of colour vision are better explained if more than three types of cone exist. The three types are S-type (short wavelength), M-type (medium wavelength), and L-type (long wavelength). These have peak sensitivities at around 420 nm (blue), 535 nm (green) and 560 nm (yellow-red) respectively (Figure 4. 1).

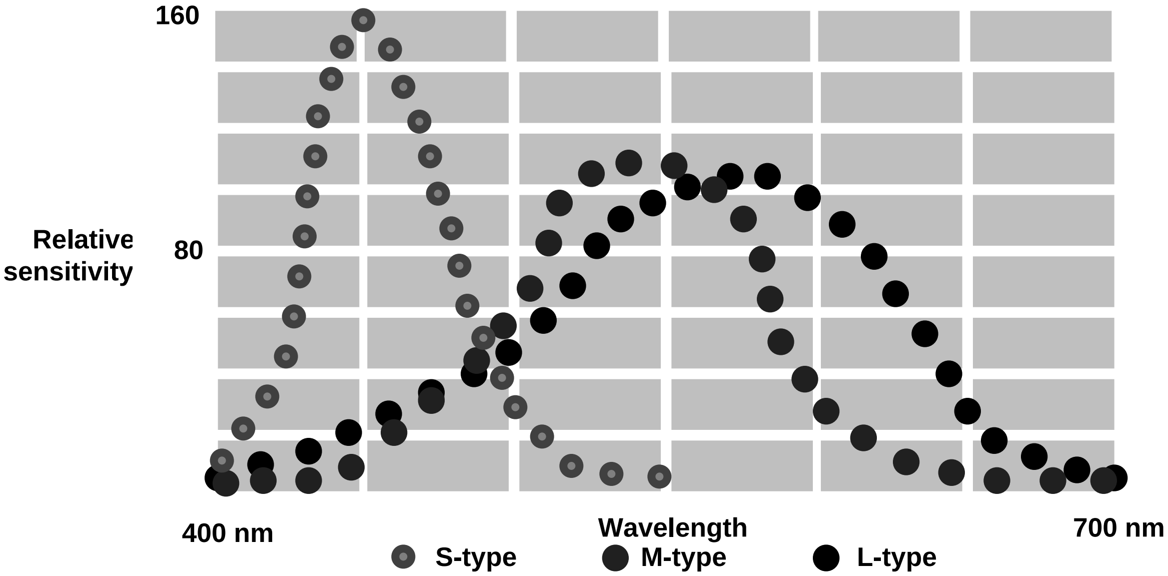


Fig 4. 1 Relative colour sensitivities of cones

Rods, on the other hand, peak at around 490 - 500 nm. The sensitivities of cones are broad-banded and there is considerable overlap between them. The absolute sensitivities of each type of cone vary, with M-type cones being the most sensitive and L-type cones slightly less so. S-type cones are more sensitive than the other two but in a much narrower bandwidth.

Strangely, too, the difference between the cones exists not only in their responsiveness. We have considerably fewer S- type than M-type or L-type cones, even in the fovea where they are most needed. S-types seem also to be more vulnerable to diseases of the retina. A consequence of the differences in responsiveness and numbers of the cones is that, as stated previously, we are more sensitive to differences in greens and reds than blues.

Because of its broad-band sensitivity, a photoreceptor cannot recognise a particular wavelength. When photons are absorbed, a photoreceptor can only respond more or less vigorously and all information about actual wavelength is lost. Indeed, a change in either the intensity or the wavelength of the stimulus is likely to produce the same response in an individual receptor. This has been called the ‘Principle of Univariance’ (Naka and Rushton 1966) and it implies that colour discrimination depends on special neural mechanisms to compare the outputs of different types of cone.

Primary hues

Our perception of hues is slightly strange in that we are able to see and identify mixtures of some colours — such as reddish yellow (around 600 nm wavelength) or bluish green (495 nm) — but not others such as bluish yellow. In addition, some hues seem more fundamental and basic than others — red, green and blue (which cannot be mixed from one another) fall into this category but so, too, does yellow (which can be mixed from red and green lights). The four colours, red, yellow, green, and blue are often called the ‘psychological primaries’ because of their ‘basic’ quality. Artists as far back as Leonardo have accepted these four colours as primary.

Opponent theory

One theory designed to explain, among other things, the fact that, although we apparently have cones of only three colour response types, we seem to see four primaries, was originally proposed more than 100 years ago by the German physiologist, Ewald Hering (1834 - 1918). After further research and development, this theory is now receiving much support and versions of it are now fairly generally accepted (Hering 1920/1964, De Valois and De Valois 1975).

In the 1870s, Hering gave a series of lectures to the German Imperial Academy of Sciences under the title, *Zur Lehre vom Lichtsinne* (On the Theory of Light Sense), in which he pronounced his view that:

> there are six elementary colours, each of which shows no resemblance to the others . . . all perceived colours can be described by their resemblance to the elementary colours only.

Hering's six elementary colours were black, white, yellow, blue, red, and green. Aware as the rest of us are that there are no yellowish-blues or reddish-greens, he considered the colour pairs yellow and blue, and red and green as, in a sense, *opposed* to one another.

According to Hering's theory, our perceptual mechanisms have the ability to determine not the hue, saturation and lightness of light but its blackness-whiteness, its yellowness-blueness, and its redness-greenness. If a mixture of equal yellow and blue light meets the eye, not only does the yellowness-greenness mechanism come into play but also the blackness-whiteness one too. However, the equal amounts of yellow and blue cancel one another out and have no effect but the blackness-whiteness mechanism would convey the impression of a neutrally-coloured light.

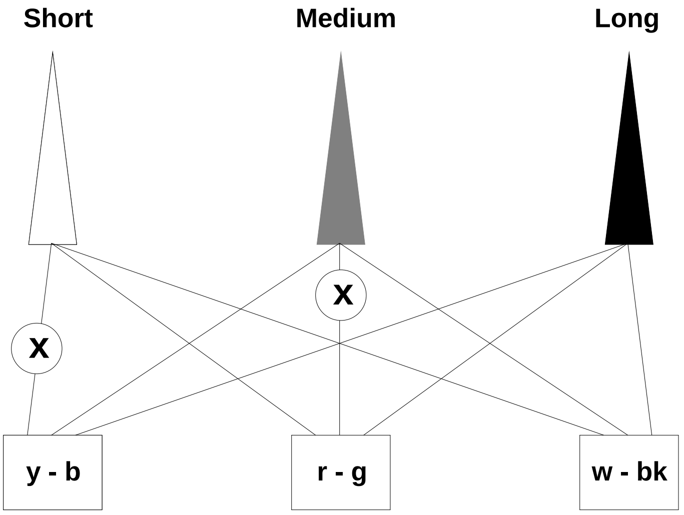


Fig 4. 2 A possible opponent mechanism for colour vision.The connections marked with x are inhibitory

Thus the theory suggests that, in neural processing, there are three components that deal with colour information (Figure 4. 2). One component, **w - bk**, signals either black or white (but not both simultaneously) — another, **r - g**, similarly signals either red or green; and the third, **y - b**, signals either blue or yellow. An orange hue at about 590 nm would result from the **r - g** component signalling red and the **y - b** component signalling yellow, both in equal strengths. The theory, then, gives us an explanation of why we can see bluish green but not bluish yellow: bluish green can be seen by the simultaneous action of the **r - g** and **y - b** components;

bluish yellow cannot be seen because the **y - b** component can signal either blue or yellow but not both together.

The usefulness of opponent theory

Opponent theory also helps explain saturation. Saturation depends on the relative amounts of achromatic (colourless) and hued light that are present in a colour. The theory claims that the **w - bk** component provides the achromatic signal which the LGN and VC compares with the **r - g** or **y - b** signal to estimate the level of saturation (or desaturation).

Some aspects of anomalous colour vision are also explained by this theory. For example, if someone has difficulty in distinguishing reds or greens (the most common form of ‘colour blindness’), this would be explained by a defect in the **r - g** opponent channel.

The general acceptance of Hering's opponent process theory by scientists was long delayed and it is only in the last thirty years or so that it has taken its proper place in our understanding of colour perception (Hurvich 1969, Abramov and Gordon 1994). The major stumbling block to its acceptance was the knowledge that any colour could be mixed by percentages of three red, green, and blue (RGB) primaries and the fact that only three types of cones seem to be present in the eye. But Buchsbaum and Gottschalk (1983) tell us that, if it is necessary to transmit signals with substantial spectral overlap, then it is very efficient to code them in terms of their common characteristics and their differences. Lennie (1984) suggests that:

> By transmitting chromatic signals in opponent pathways the nervous system makes more efficient use of limited information-carrying capacity (p243).

Because of its information-carrying efficiency as well as psychophysical evidence we now accept that the concept of opponent processing is not just confined to colour vision. There is a general use of opponent mechanisms throughout the visual system.

The Natural Colour System

Almost from the start, Hering's colour ideas were wholeheartedly taken up in Sweden, particularly by the physicist, Tryggve Johansson (1905 - 1950), and they form the basis of the Natural Colour System which is the standard for colour description by Swedish architects, designers and others concerned with colour specification (Hård 1966, Hård and Sivik 1981). In this system, neither hue nor saturation are regarded as fundamental colour attributes: hue is the relationship between redness-greenness on the one hand and yellowness-blueness on the other; saturation is the relationship between whiteness and the colour attributes (chromaticity).

A four-colour model

The opponent concept give rise to the three-axis colour model shown in Figure 4. 3 Incidentally, this is rather like that devised by the Renaissance architect, Alberti (1404-1472), who also did so much to establish the form of linear perspective we use today.

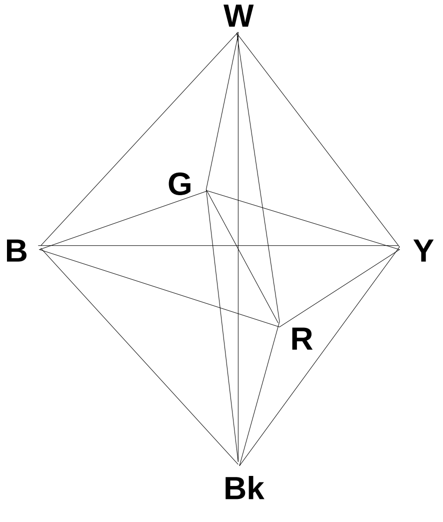


Fig 4.3. A three-axis colour model based on the opponent process. A similar model was used by Alberti (c1470) and by Hoffler (1883)

The OSA colour model devised over the period 1947 -1977 by the Optical Society of America uses a Hering-like approach. In this colours are described in terms of their lightness, yellowness and greenness. A 30% grey is used as lightness 0 and the range runs between -17 to +5. Yellowness ranges from -6 to +11 with negative values indicating blues and positive values indicating yellows. Greenness ranges from -9 to +9 with negative values indicating reds and positive indicating greens.

Processing colour

What sort of mechanism would explain the opponent process? A clue is given in the widely overlapping responses of the cones as illustrated in Figure 4. 1 and that, in fact, cones do not measure wavelength directly. Abramov (1981) puts it thus:

> A cone can signal only the rate at which its pigment traps photons and cannot provide unequivocal information about the wavelength of light — its response is univariant — and thus a cone is not a 'colour' receptor. Colour vision depends on interactions, at subsequent stages of the visual system, between responses derived from different types of cone.

The importance of those subsequent stages in our perception of colour cannot be overstated and, as Boynton (1983) points out:

> Most importantly, most of the mechanisms of colour discrimination are not retinally based. Theories that assume that there are mechanisms capable of registering differences between receptor outputs in the retina . . . seem highly suspect, since they do not account for, or even take account of, the effects of chromatic induction; nor can they explain the good quality of dichoptic discrimination or an ability to make consistent judgments with respect to certain remembered colours. Our results indicate that the monoptic-dichoptic differences in chromatic discrimination may result from an extra enhancement of colour difference from chromatic induction, the mechanisms of which lie mainly in the monocular visual pathways, but not to any great extent within the retina.

De Valois and de Valois (1993) set out a four-stage model of colour vision, which explores the extent to which essentially random connectivity (as suggested by recent studies of anatomy) might result in a sensible colour organisation. The first stage of the model has three sorts of cones, L-, M-, and S-type in ratios of 10: 5: 1. In the second stage, retinal connections lead to three pairs of cone-opponent, and one pair of cone-nonopponent systems. At a third stage of colour processing, in the cortex the S-opponent cells are added to or subtracted from the L- and M-opponent units to split the parvo-geniculate response axis into separate **r - g** and **y - b** colour axes, and separate luminance from colour. This model is currently gaining favour as a possible mechanism for explaining colour vision.

D'Zmura (1991), on the other hand, shows that some coloured targets ‘popped out’ of displays under conditions in which the standard **r - g**, **y - b**, and **w - bk** mechanisms could not deal with directly. He suggests that observers possess chromatic detection mechanisms tuned to intermediate hues such as orange as well as to the standard colour-opponent hues.

Colour vision defects

About 8% of the male population has anomalous colour vision and is unable to distinguish colours in some part of the visual range. The defect is gender-related: only about 0.5% of the female population are similarly afflicted. Very few people are totally colour blind in the sense that they see only in monochrome. Those that have this defect usually also suffer from very poor vision generally although there are rare cases where the ability to match or identify colours is absent without any other vision loss. Such cases suggest that people can process signals from the three types of cone but must be unable to compare their differences.

The majority of those who suffer from what we normally call ‘colour blindness’ fail to distinguish colours either in the red range (their vision is called ‘protanamalous’) or in the green range (their vision is called ‘deuteranomalous’). It is thought that these problems arise because of the absence of L- or M-type cones. Alternatively, it is possible that the cones exist but peak at abnormal wavelengths. Those who fail to see in the blue range are rare when compared with those who have problems in the red-green range. Those who cannot distinguish red from green at all are called ‘dichromats’.

Personal and cultural differences

Colour seems to be a subject in perception where both personal and cultural differences might show themselves in a marked way and it would not be prudent to assume that colours mean the same things to everyone. There is no evidence, however, that different people *see* colours differently although this might possibly be the case as, for instance, the extent of colour blindness is known to vary markedly around the world. Pickford (1972), for example, informs us that there are very many less examples of red / green defects in the vision of black people and Australian aborigines than there are in white people. But, clearly, it is likely that the *meaning* of perceptions and preferences for certain colours might depend on contextual and cultural rather than perceptual factors. Over the years the literature abounds with instances of these. Pickford (1972) summarises some of the work up to the 1960s.

However, Davidoff (1991) is rightly sceptical about some of the published effects of colour preference differences and complains that many of the experiments that report on these are badly designed and executed. There is little doubt that this is true and that some reports on the meaning and effects of colours are little more than anecdotal. Furthermore the reports often put down phenomena just to changes in hue and fail to separate the effects of hue, lightness and saturation. As Davidoff (1991) puts it:

> There is said to be, in general, a preference for blue and a dislike of yellow (McManus, Jones and Cottrell 1981). The same has even been reported to be true for monkeys (Humphrey 1972). However, Pickford (1972) in his review of the research on cross-cultural and personality preferences for colours concludes that the lack of stimulus control makes it impossible to be adamant about the preference for blue (p118).

There is, though, *some* good evidence of cultural differences in colour perception. Fussell and Haaland (1978), for example, outline their studies of colour understanding in Nepal. In these they asked more than 400 village people to pick out their colour preferences for certain purposes from a set of cards. The colours most preferred for happy occasions and for women were purple, pink and red. As the

authors point out, this may be because Hindu brides wear red. Orange and yellow were also liked and these were colours chosen ‘for gods’. Buddhist monks, of course, wear these colours. Thus cultural influences may have been at work in relating colours to purposes. As the least preferred colours were black, dark brown and grey, one also wonders whether people were making preferences based on lightness or saturation rather than hue.
Saito (1996) shows that the Japanese in particular have a very strong preference for white (more than 25% choosing it as their preferred colour which they see as being clean and pure). Saito puts this down to the influence of Shintoism. He concludes that colour preferences depend on age, gender and geographical distribution (finding, for example, some significant differences in preference in different Asian cities). He also suggests that, particularly among the young, fashion plays a significant role in preference.

In studies of students in Beirut, Choungourian (1968) found red and blue as highest in preference for Americans but lowest for Kuwaitis. Blue-green on the other hand was ranked lowest by Americans but highest for Iranians and Kuwaitis.

In examining US college students, Silver (1988), found that black subjects preferred colours in the red-purple-black range but that white students preferred blues and greens. Grieve (1991) found that subjects from Senegal and the Transkei preferred red and black but that their traditional cultural beliefs were not as strong an influence on colour preference as might be expected. The general order of Western preference for colours in adults is given by the list blue, red, green, violet, orange and yellow. Smets (1982) found that, almost always, fully saturated colours are preferred to desaturated ones. Goethe (1749 - 1832), who made extensive studies of colour and developed an elaborate theory about it, suggested that this preference might be a primitive appreciation:

> Men in a state of nature, uncivilised nations, children, have a great fondness for colours in their utmost brightness, and especially for yellow-red . . .

On p4- 26 we look at some contradictory evidence to suggest that saturation might not play as significant effect as Goethe and others thought.

In a rather surprising US study of colour preference between white cheese and yellow cheese, Scanlon (1985) found that 30% of black subjects chose white cheese and 70% chose yellow. On the other hand, 53% of white subjects chose white cheese and 47% chose yellow. When people of Hispanic origin were tested, 32% chose white and 68% yellow. To ensure that they were choosing solely on the basis of colour, the subjects were given blindfold tests before making the colour choices and were clearly unable to distinguish different tastes in the cheeses.

Many of tests on colour preferences are made with colour patches under laboratory conditions and, hence, even when they are properly conducted, need to have their results treated with caution. In her tests Lind (1993) found some evidence that many people used the same criteria in making preference choices in coloured test patches and clothing, but up to a quarter of those tested did not. This is not surprising. It is often hard and slightly unmotivating to make preferences in the abstract, and the colours that one prefers for one purpose would not necessarily be chosen for another.

Colour and language

Trandis et al (1973) tell us:

> Although different languages encode in their vocabularies different numbers of basic colour categories, there are exactly eleven basic categories from which from which the colour terms of every language always draw.

Based on the work of Berlin and Kay (1969), they list these as shown in Figure 4. 4.

White
Black
Red
Yellow or Green
Yellow and Green
Blue
Brown
Purple
Pink
Orange
Grey

Fig 4.4 Showing the ordered list of colour words found by Berlin and Kay (1969) to be shared by most languages

Thus, the most primitive of languages would have words for black and white, less primitive languages would include black, white, and red, and so on. Only well-developed languages would have words for orange and grey. The colour ‘grue’, a combination of green and blue, has often been suggested as a ‘universal’ primitive because many languages once did not have different words for blue and green (see also Zollinger 1984 and Zimmer 1984).

Of course, it is not the case that people who do not have words for different colours cannot see differences. It is only that they the differences are not regarded as of sufficient importance to name. Davies et al (1991) examined whether there were any differences in perception among those whose languages differ in the number of categorisations of colours. Russian speakers, who have two basic terms for blue, were compared with English speakers, who have only one basic term for the same colour-range. Some of their experiments tried to find whether the Russians could find more perceptual differences in the blue region than non-Russians. All their experiments failed to find any differences between the two groups.

There is some evidence that the order in which children learn to name and distinguish colours also falls in to the hierarchy of Figure 4. 4. Incidentally, Davidoff (1991) comments on the fact that, at the turn of the century, Alfred Binet (1857 - 1911), the French psychologist who made a study of child development and attainment, suggested that at the age of 7 or 8, children should be able to recognise and name the four primaries! Nowadays, of course, a child of 3 or 4 who was unable to do this would be regarded as somewhat backward and teachers expect children entering British primary schools to already have a well-developed vocabulary of colour names (see also Johnson and Tomiie 1985).

Wierzbicka (1990) argues that colour concepts are anchored in certain 'universals of human experience' such as night, day, fire, the sun, vegetation, the sky and ground and proposes a new, as it were, ecological interpretation of the evolutionary sequence of basic colours suggested by Berlin and Kay (1969). Courtney (1986), on the other hand, shows that red — associated by many with fire, blood and danger — is not seen universally in this way.

Von Waltenwyl and Zollinger (1979) make the point that:

> If the opponent scheme of Hering . . . is reflected in colour terms, words for black, white, red, green , yellow and blue should be primary terms, whereas all other colour words should have lower salience (p281).

They found that, when asked to list the most necessary words for colour in their respective languages, 95% of German, French and English speakers listed red, blue and green as the most necessary (with yellow also in this category for French speakers). 95% of Japanese speakers on the other hand, listed black, white and green. Von Wattenwyl and Zollinger put this difference down to the influence on Japanese culture of Zen Buddhism. At the end of their study, they conclude that:

> . . . the linguistics of colour terms *corroborate* the neurophysical basis of the opponent colour theory; they do not conflict with it.

Also of interest here is the point that, if the most preferred colour is blue and the least preferred, yellow, as is reported in many studies on Western subjects:

> . . . it is not a coincidence that the dimension blue-yellow represents a dimension both of aesthetic preference and also of

> the opponent colour system (McManus, Jones and Cottrell 1981 p665).

Other studies seem to confirm that there is a strong correlation between Berlin and Kay's list of 'universal' colour terms and the opponent primaries. Martindale and Moore (1988) also show that, if one tests people with 'good' examples of these colours (that is those examples that one would regard as prototypical versions) against 'bad', non-prototypical versions, the prototypical versions are universally preferred.

Strangely though, it has been known for more than a century that recognising a colour takes longer than recognising a word naming the colour (Seifert and Johnson 1994). This leads to the conclusion that words and colours are differently processed and that words seem to hold a privileged position in our processing capacity.

In Chapter 21 of her book, Rossotti (1983) has some useful insights on words and colour.

Colour and context

Perhaps more than any other feature of perception, our appreciation of colour is highly influenced by context. The appearance of a patch of colour depends very much on its surroundings and this applies to all three aspects of a colour — hue, lightness and saturation. 'Colours', said Leonardo, 'appear what they are not, according to the ground that surrounds them'.

For example, a small patch of colour against a black background will appear more saturated and darker than a larger patch of the same colour seen against a black background (Plate 4. 1). A white square seen against a black background will seem larger than a black square on a white ground (Fig 4. 5).

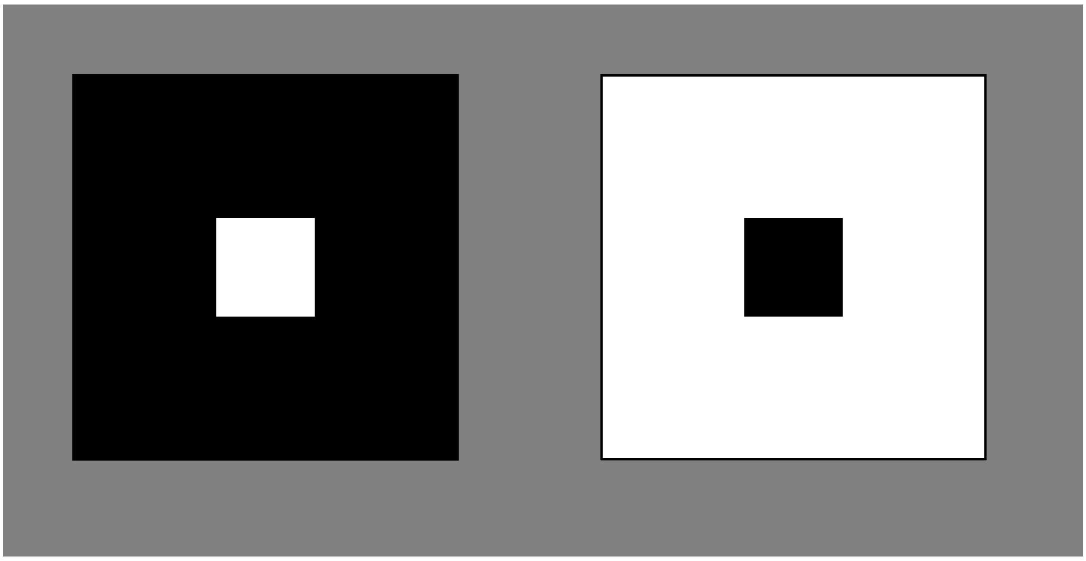


Fig 4.5 The small white square on the left seems larger than the small black one on the right

Because it depends on lightness rather than hue, this phenomenon applies to coloured squares too. Thus three equally-sized squares of yellow, green and blue will appear to be of different sizes (Plate 4. 1).

Not only does the apparent colour of a patch appear to change according to the colour of its background but the sharpness of its edges alters too (Plate 4. 2). This is because we cannot focus on the edges between two colours of equal lightness even when they are of different hues — a phenomenon perhaps having something to do with Mach banding where, because there are no differences in lightness, overshoot or undershoot does not come into play. Thus edges are not enhanced.

The importance of contrast to legibility

For legibility of lettering, then, it is therefore essential that there are contrasting lightnesses between lettering and its background. It is surprising how often this factor is forgotten. We frequently see examples of blue lettering on black or yellow on white. Sometimes this is done deliberately for effect and its consequences are not serious. Sometimes, though — as in instructions for medicines — it can give rise to severe problems. Note too that these problems increase with age. The older we get, the more contrast is needed for legibility. Chen and Yu (1996) show that contrast in brightness is more important than contrast in hue but that either can be effective if contrast is high enough. They also found evidence of an additive effect between brightness and colour contrast.

Specifying colours to computers

Computer monitors produce their colours by excitation of red, green and blue (RGB) phosphor dots too small for the unaided eye to separate. These are arranged in groups of three RGB dots and a number of these groups make up each pixel. Different colours are obtained by the excitation (by electron guns) of each of the phosphor dots in combination. All the red, green or blue dots in a pixel are illuminated together — with amount of illumination of each set varying in steps from 0 to 100%. If the illumination of each RGB set can only take the value 0 or 1, then just 8 colours are possible: black and white (either all on or all off); red, green, blue (two off and one on), magenta, yellow, cyan (two on and one off). If the illumination of each set can take 256 levels (8 bits on each of R, G and B) then 16.7 million colours are possible.

The RGB model

One way of specifying colours to the computer is to use the RGB model. The model works because, give or take some gamma correction, we directly specify a desired colour by telling the system the intensity of each of the three electron guns to be used in making it up. However, this is not always a straightforward thing to do when many steps of colour are possible and, sometimes, considerable trial and error is needed before success in subtle colour specification is achieved. A method I have used for some years to teach students how to manipulate colours in RGB works quite well and is given in Appendix 1, but the RGB model is, essentially, a low-level one which does not properly encapsulate the perceptual characteristics of colour.

It has been suggested elsewhere, though, that the three-dimensional spatial nature of the RGB model (Fig 4. 6) can be exploited to give designers a new way of working with colour relationships (Mitchell 1986, Lansdown 1987).

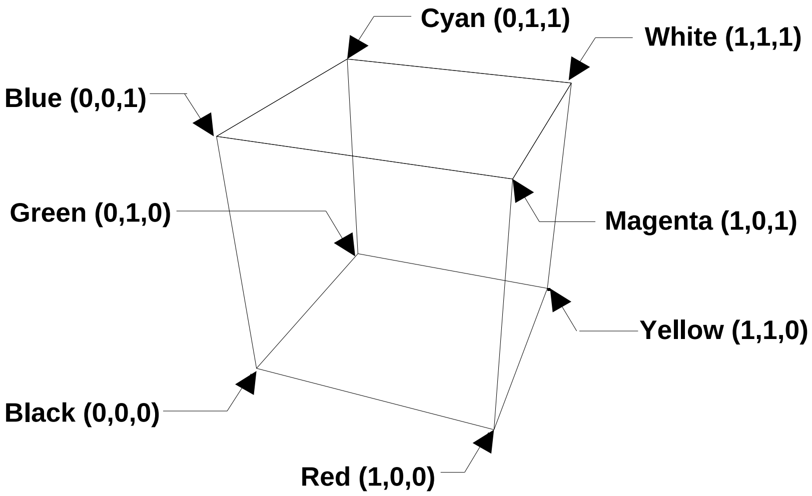


Fig 4. 6. The RGB cube model in which all displayable colours lie within or on the surfaces of the cube

Assessing lightness in the RGB model

The fact that a three-primary model seems to match both the physical structure of the eye and the way in which colour monitors operate seems to be a powerful argument in favour of the RGB model. Because of its direct association with the physical process of creating colours, too, the RGB model ought to assist us in assessing how light a particular colour should appear. To produce white we must have all the guns full on (RGB specification 1,1,1). Thus white will appear the lightest (and brightest) of all because more light is reaching us than with any other arrangement. To produce cyan (0,1,1), yellow (1,1,0) and magenta (1,0,1), two of the guns must be full on so these colours should be less bright than white but brighter than the primaries (1,0,0), (0,1,0) and (0,0,1) which only have one-third of the phosphors glowing. This, though, is not our perception. White (1,1,1) certainly is the brightest; next comes yellow (1,1,0) and cyan (0,1,1); but then comes green (0,1,0) and red (1,0,0); finally we have magenta (1,0,1) followed by blue (0,0,1). This, of course, is because the hues of are not of equal intrinsic brightness.

Gerritsen (1979) tells us that, from the time of Aristotle to about the middle of the seventeenth century, people always listed colours in apparent lightness order rather than the spectral order we have tended to use since Newton. The historic list was white, yellow, green and red, blue, black — roughly the order just discussed.

HLS and other models

Aware of the difficulties in manipulating the RGB model, the US computer company,Tektronix, developed a colour model in 1977 loosely based on the Munsell colour system well-known to many artists and designers (Munsell 1916). In this model (Figure 4. 7), a particular colour is specified in terms of its hue (given as an angle in the range 0°-360° starting at blue = 0°, passing through red = 120° and green = 240°); its lightness (varying from 0% to 100%); and its saturation (also varying from 0% to 100%).

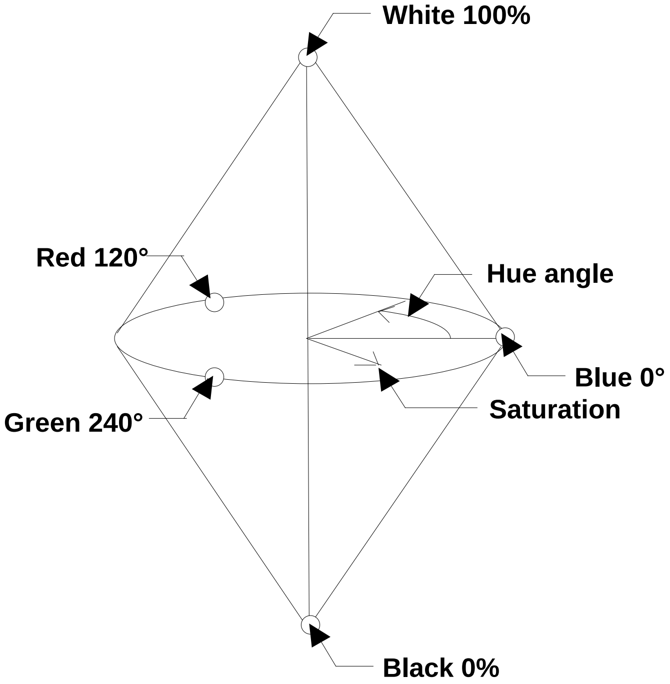


Fig 4. 7. The HLS double-cone model with hues situated at various angles about a circle, lightness on the vertical axis and saturaion given by the distance out from the axis. Fully saturated colours lie on the circumference of the centre circle.

Although useful — and despite the plaudits of its champions — the HLS model does not really have any firmer theoretical status than the RGB model. This is because, as we have already pointed out, it is not possible to specify lightness and saturation independently of one another: we notice a change in saturation when we change lightness and *vice versa.* Furthermore, as can be deduced from the Ancient Greek and folk-ordering of colours mentioned above, our perception of lightness also changes with hue and *vice versa.*

All these problems of colour recognition and specification — plus the fact that our perception of a given patch of colour depends on its size and the colours that surround it, as well as the colour of the ambient light in which we view the screen — might suggest that it is hardly worthwhile worrying too much about the sophistication and accuracy of computer colour models. It might be better to rely on crude approximations which are then fine-tuned by trial and error to suit particular cases. There are all sorts of reasons why such an experimental approach would not be justified and considerable work goes on to find devise models which try both to match the physics and the psychology of the situation. Apart from the RGB and HLS models just touched upon, the Hue, Saturation, Value (HSV) model (Smith 1978), the Colour Notation System (Berk, Brownston and Kaufman 1982) and the Uniform Colour Space model (Taylor, Murch and McManus 1987) are all valuable attempts at trying to achieve this matching. The YIQ colour model used in television, where most of the signal is devoted to luminance rather than hue, is also of interest.

Opponent colour models in computing

There have been two well-known attempts to use the opponent process in computer graphics. These are by Schwarz et al (1984) and Naiman (1985). In both these cases the luminance information is separated from the chrominance information in a way similar to that used in the YIQ television model although neither system does this in a very intuitive way. Schwarz et al (1987) however tested their model against the others in controlled experiments and found that it had some advantages. Naiman includes a table comparing his variation of Schwarz's model with the original and six of the others. It too shows a claimed superiority for an opponent type approach. Lansdown (1988) proposed an easily implementable opponent colour model sitting on top of the RGB model that allows a more rapid specification of colours.

CIE colour models

There are, then, many different models available but implementations of the computer graphics standards, PHIGS and PHIGS PLUS, are required only to support the colour model known as CIELUV because it has greater perceptual uniformity than any other. CIELUV is a version of the CIE 1931 model that has been an international standard for more than sixty years.

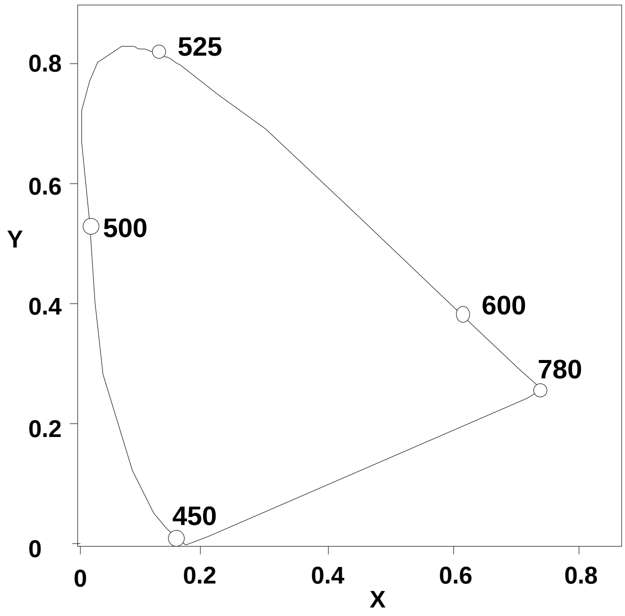


Fig 4 . 8. The CIE 1931 colour space. All perceivable colours lie within and on the perimeter of the horseshoe.

The CIE 1931 model (Figure 4. 8) was derived from experiments in which subjects were required to match given light sources by mixing red, green and blue lights in measured proportions. The amount of each light, x, y and z, needed to make the matches was then normalised so that the sum of colours totalled 1. The value of this normalisation where $X + Y + Z = 1$ is that a three-dimensional colour space (X, Y, Z) can be plotted correctly in two-dimensions (X, Y) allowing the third dimension, Z, to be calculated from the other two. The CIE 1931 model appears as a horseshoe shape in which all perceivable colours lie. The most saturated spectral colours are on the perimeter of the horseshoe. Note that the non-spectral colours (magentas, browns) lie near and on the straight line at the bottom of the horseshoe.

Although this model is good for many purposes, it is highly non-linear as far as our perception of colours is concerned and many attempts have been made over the years to reconfigure it in a more linear way. Figure 4. 9 shows the most recent agreed form of this: the CIE 1976 uniform chromaticity space (UCS).

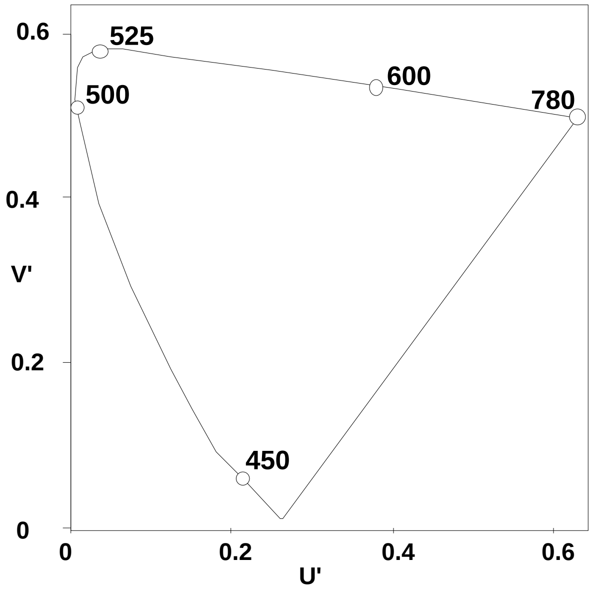


Fig 4. 9. The CIE 1986 uniform chromaticity space designed to relate colours in a more linear fashion than the CIE 1931 model.

Colour monitors

Given that colour monitors have just three types of phosphor that emit red, green or blue light, it is clear that no computer can show all the colours that can be perceived. This is the case even if the phosphors could emit colours that are saturated to the spectral limit. Why this is so becomes clear if we plot the wavelengths of the light emitted from the phosphors on the CIE diagram. These form a triangle and no triangle can fill the whole horseshoe space. In fact monitors normally emit light at very much lower colour saturations than those on the light spectrum and a typical colour monitor gamut is shown in Figure 4. 10.

It is important to note that the ‘best’ and most prototypical colours in monitors are not always those that are produced by the selection of hues on the boundaries of the gamut triangle (Kauffman and O'Neill 1993). Thus the 'best' red, green and blue, for example, are not necessarily given by using (1, 0, 0), (0, 1, 0) and (0, 0, 1). Very often, experimentation is needed to choose the most representative colours and this is a factor that should also be taken into account in the choice of monitors.

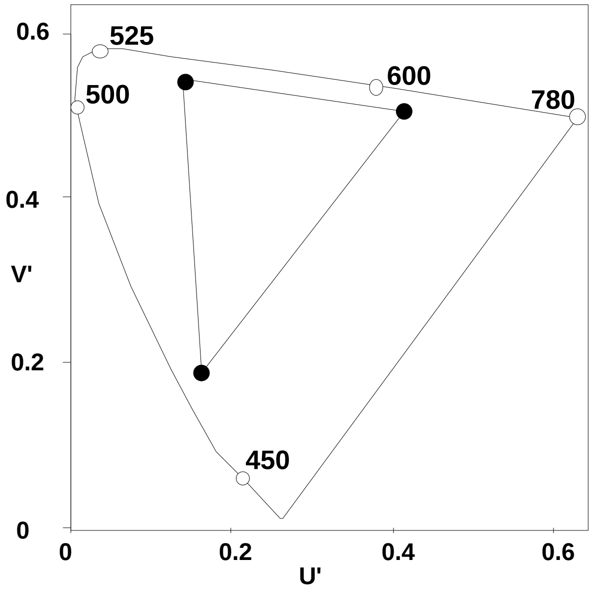


Fig 4. 10 showing the gamut of a typical computer monitor plotted on the CIE 1976 UCS. From this we can see that, whilst the monitor is able to display well-saturated yellow and orange, its display of blues, greens and reds as well as non-spectral magentas is very limited.

Using colour in graphics

Wright and Lickorish (1988) investigate the use of colour in helping people remember where they are in lengthy texts. Although some of their experiments confirmed the value of coloured paper for printed texts, three experiments failed to confirm similar advantages for screen information. They suggest that, for screen information, colour might work better as borders or strips or if the colours were assigned by the readers rather than the writers.

Haber and Wilkinson (1982) stress the now fairly standard view that, to indicate values on an increasing or decreasing scale, colours should be arranged in their spectral order — red through yellow, to green to blue. Many experiments confirm that, if values are to be related to hue, spectral ordering is certainly essential. Livingstone (1988), on the other hand, in a fascinating study of the way in which the human brain processes information, shows that more useful information arises from arranging the colour representations in order of brightness rather than hue (Plate 4. 3). This, of course, has been the way cartographers have tended to display coloured contour information for many years — although Bruce and Foster (1982) are right to warn us that studies of colour in printed matter are unlikely to apply directly to screen images.

The addition of colour can be used to emphasise similarity or difference. All too often though in computer graphics, colour is used arbitrarily without regard for the way in which it may influence or disturb perceptual grouping and there is now sufficient evidence to show that the arbitrary or thoughtless use of colour can prevent rather than facilitate visual communication of information. See, for example, Schontz, Trumm and Williams (1971), Katzman and Nyenhuis (1972), Knoll (1977), Elio and Reutener (1978), Haber and Wilkinson (1982), Murch (1985).

Many studies carried out over the years suggest that there are some colour principles to be followed in creating graphics. For example, and referring to a colour circle:

- any hue displays in strong contrast to its complementary colour (hue + 180°)
- colours 120° apart are easy to distinguish (often even by those with defective colour vision)
- colours near in hue have a harmonious effect.

Surprisingly, experiments by Polzella and Montgomery (1993) suggest that, in determining the ‘pleasantness’ and harmony of colours, we seem only to to take into account their hue and lightness. Saturation does not seem to be a significant factor. This result is in direct contradiction to Smets (1982) who showed that saturation was by far the most significant determinant of ‘pleasantness’ in his experimental subjects. This is an example of the difficulty in finding the sort of agreement in perceptual studies that is necessary if we are to establish guiding principles for colour use. Despite the plethora of reported experiments on the psychophysical effects of colour, it is clear that much still remains to be studied.

Summary

Colour plays a significant part in our daily lives but is still a phenomenon that is not completely understood. It is clear that there are prototypical colours that are more representative of Berlin and Kay’s universal colours than others and that there are differences in personal colour preferences that, at least partially, can be put down to cultural differences. However, studies in this area needed to be treated with caution.

The use of colour can enhance information by helping ‘pop out’ but, equally, can sometimes confuse and obfuscate. Colour, then is problematical and, despite hundreds of years of study, still is a mysterious phenomenon. It should therefore be used judiciously and not with the belief that it always enhances meaning.

5. The principles of perceptual organisation

A number of organisational principles seem to be at the heart of our visual perception of scenes. They seem to make us prefer some explanations of what we see to others that, on the face of it, would be just as plausible. The principles were set down in the 1920s and 1930s by a group of psychologists, in particular, Max Wertheimer (1908-1943), Wolfgang Köhler (1887-1967) and Kurt Koffka (1886-1941). These psychologist, known as the 'Gestalt psychologists', challenged the then prevailing views of human psychology which they believed to be too simplistic. Taking the standpoint that 'the whole is greater than the sum of its parts', they believed that the piecemeal stimulus-response approach adopted by the classical psychologists was inadequate to explain the subtlety and complexity of human behaviour. They performed a series of striking visual experiments that convinced them (and now most of us) that the brain actively and dynamically organises what we see and imposes a structure on it.

Some Gestalt psychologists, though, went on to suggest that the reason these principles hold is that the brain performs its organisational tasks in a what might be thought of as a physically analogous way — something they termed 'isomorphism': so that, for example, if we see a circular pattern, the brain actually organises a 'circular trace of field forces' within its cells — a viewpoint that had no neurological evidence to support it. Dodwell (1986) writes:

> Although the Gestalt psychologists believed passionately in the importance of organising forces, and postulated that the origin of these forces was to be found in physiologically determined properties of the visual cortex, they in fact lacked virtually all of the detailed knowledge of the way the visual brain operates. Small wonder then that their theories, at least in terms of their physiological embodiment, came to nothing (p320).

As Henle (1984) is at pains to point out, however, many of the problems about the Gestalt ideas on 'physiological embodiment' and especially isomorphism, arise because of gross misunderstandings about what the Gestaltists believed and, in many cases, actually said.

Gestalt isomorphism applies more to topological relationships than to geometric traces.

Whether through misunderstandings or not, nowadays the Gestalt approach to psychology as a whole has been largely abandoned and we think of their work as descriptions of visual organisation rather than explanations of it — although Epstein (1988) asks whether it is time to look again in a more favourable light at the whole of Gestalt theory. Gestalt principles that describe the way in which some perceptions are more likely to occur than others are now widely accepted and need to be known by those whose job it is to make images.

Organisational principles

There are many different lists of Gestalt organisational principles in the literature. In some cases, over 100 are listed but we will confine ourselves to some of the ones that can be exploited in the design of graphics and interfaces. These are:

- smooth (or good) continuation
- proximity
- similarity
- orientation
- closure
- relative size (figure and ground)
- symmetry
- familiarity and context
- common fate

Smooth continuation

We group together in a single structure those parts which seem to align or continue smoothly. Thus, in Figure 5. 1, we see two curved lines crossing at right angles as in Figure 5. 2 rather than two V-shaped forms meeting at a point as in Figure 5. 3. This principle is also known in some texts as ‘good continuation’.

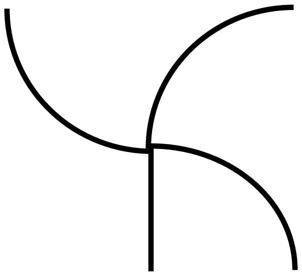


Fig 5.1 Smooth continuation of lines

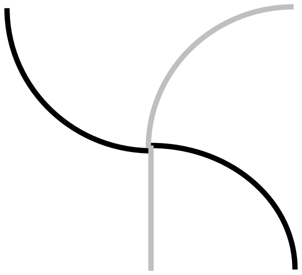


Fig 5. 2 Preferred visual grouping of lines

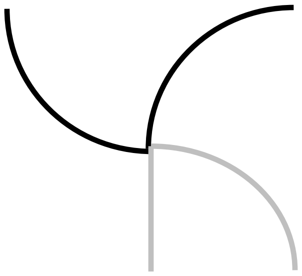


Fig 5. 3 This visual grouping of lines is normally rejected (even though it was the grouping that was used to create Fig 5. 1)

Perhaps this feature derives from our experience of the 'objectness' of things. Thus, we see objects as a whole even when they are partially occluded by other things because we know that they do not break up when they disappear from view, as Figure 5. 4 shows. Objects overlaying others have continuous outlines. The overlaid objects have interrupted outlines but, by the principle of smooth continuation, we can deduce continuity behind the overlaying object. Kellman and Spelke (1983), however, show that 3-4 month old infants only appreciate this if the objects are moving relative to one another.

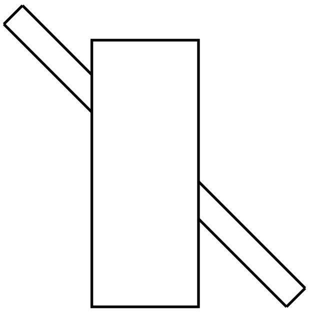


Fig 5. 4 Smooth continuation contributes to our feeling of 'objectness' and allows us to judge that the thin rectangle might be a complete object behind a rectangle in front

Presumably we learn to judge the still-image situation from experience although, as Figure 5.5 illustrates, our tendency to accept objectness through the principle of smooth continuation can sometimes mislead us.

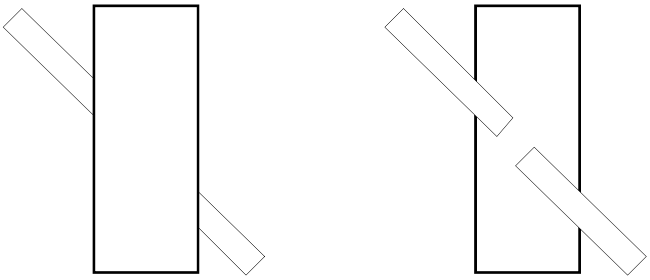


Fig 5. 5 Sometimes objectness overrides smooth continuation and gives us the mistaken impression that lines continue when they do not. Not only is the diagonal rectangle in the left hand side of the figure not continuous, the two diagonal rectangles (as can be seen in the figure to the right) do not even align

Proximity

We group together those parts that are closest together. As can be seen in Figure 5. 6, we perceive the group (a) as three vertical lines of dots and the group (b) as three horizontal lines of dots. The dots in (c) are equally spaced and do not suggest an orientation.

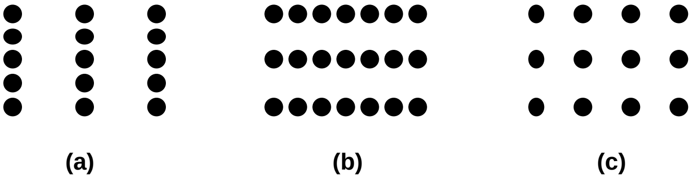


Fig 5. 6 The principle of proximity determines our interpretation of the groups

Obviously proximity and size of the elements that make up the pattern are related factors here (Zucker and Davis 1988).

Similarity

We group together those parts that appear ‘similar’. Hence in Figure 5. 7, we see separate white diagonal lines and black diagonal lines rather than vertical or horizontal lines of black and white dots.

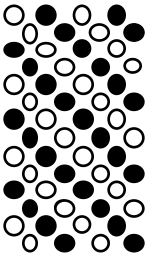


Fig 5.7 The principle of similarity determines our interpretation of the groups in this case and thus we tend to see diagonal lines

Sometimes, similarity can override proximity as the organising principle (Figure 5. 8).

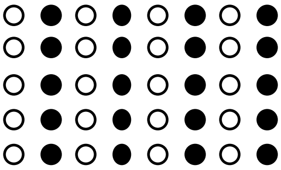


Fig 5. 8 The principle of similarity ensures that we group the dots in vertical columns

Orientation

We group together items which are arranged in a vertical or horizontal orientation in preference to those orientated on different axes. Thus orientation seems often to be a stronger grouping principle than similarity (Figure 5. 9).

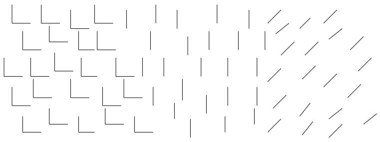


Fig 5. 9 We group together items having similar orientations rather than similar shapes

Orientation, as we shall see later when we look at relationship to frame, also affects our interpretation of a shape.

Closure

We group together parts that give the appearance of closed shapes (Figure 5. 10).

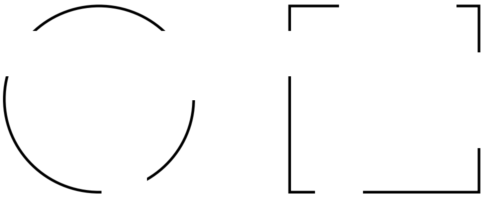


Fig 5. 10 The principle of closure determines that we see these interrupted lines as forming closed figures

Thus it is often possible to suggest a virtual frame around a figure by only drawing its corners. The organisational principle of closure seems to come to the fore when we interpret sketch drawings — which are often incomplete but which we normally have little difficulty in understanding (Figure 5. 11).

Fig 5. 11 Closure comes into play in our recognition of sketches

Relative size: figure and ground

Given two superimposed areas, we will tend to see the smaller as a figure against the larger background rather than vice versa (Figure 5. 12).

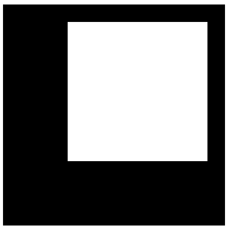


Fig 5. 12 We tend to read this image as a white square on a black one rather than a black square with a hole in it

When there is little difference in the size of the parts, ambiguity can result and we are unable to fix exactly which is the figure and which is the ground. Sometimes this ambiguity can be exploited for art purposes. Figures 5. 13 illustrate this.

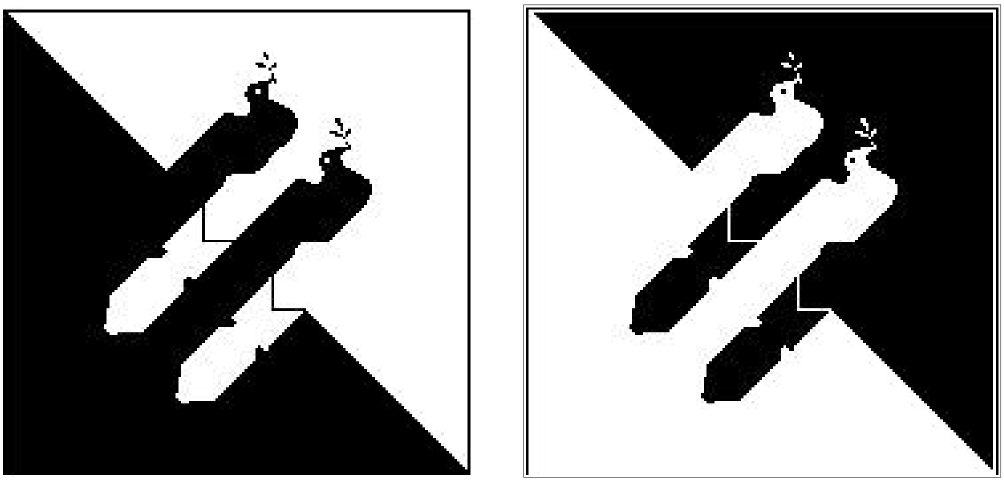


Fig 5. 13 Two versions of an image in which the relative sizes of the parts are similar so that it is difficult to distinguish figure from ground (based on a CND poster)

In cases like these we can chose arbitrarily which is the figure and which is the ground. One choice allows us to perceive one meaning, the alternative choice allows us to perceive a different meaning. This effect is the basis of a number of well-known optical illusions such as the faces / vase illusion of Figure 5. 14.

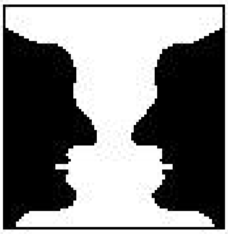


Fig 5. 14 Different information arises depending on whether we take the black or the white to be the ground

Part of the ambiguity of Figure 5. 14 seems to derive from the mixture of convexity and concavity of the black and white parts. Kanizsa and Gerbino (1976) show that shapes that are symmetrical about a vertical axis are usually seen as figures against ground but that this is not always the case if the forms are concave. Carrying out experiments with diagrams similar to those shown in Figure 5. 15, they discovered that over 92% of people tested see the convex shapes as the figures and the concave ones as ground. This is independent of whether the convex shapes are black as in (a) and the concave ones white or vice versa as in (b).

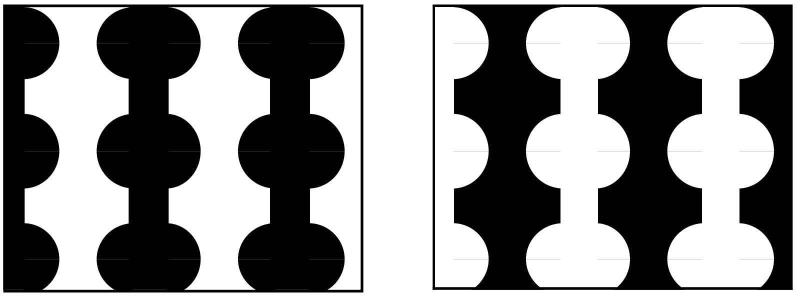
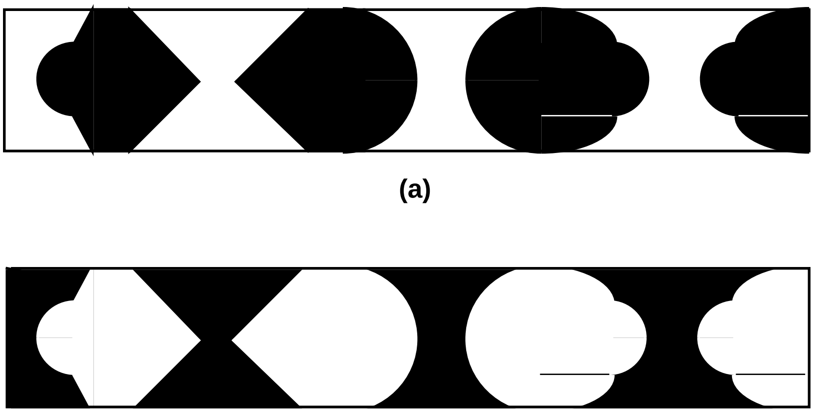

(a) (b)

Fig 5. 15 In general it is the convex shapes that are seen as figures against ground

To check whether the matter is influenced by degree of symmetry rather than convexity / concavity, Kanizsa and Gerbino (1976) went on to test subjects using diagrams similar to those in Figure 5. 16 in which the concave shapes are symmetrical about the vertical and horizontal axes but the convex figures are only symmetrical about the horizontal axis. Once again the overwhelming majority of subjects saw the convex shapes as the figures whether they were drawn in black, as in Figure 5. 16 (a) or in white, as in Figure 5. 16 (b).


(a)

(b)

Fig 5. 16 Here again, the convex shapes are seen as figures on a ground of the opposite colour

This seems to confirm an innate preference for convexity in two-dimensional shapes. It is not clear, though, that this preference transfers to a preference for convexity in the third dimension (although this can probably be inferred).

As with the other grouping principles, the figure / ground, relative size principle can sometimes be overridden by a different preference: in this case for certain orientations (as in Figure 5. 17 where, in both images it is the upright one that dominates).

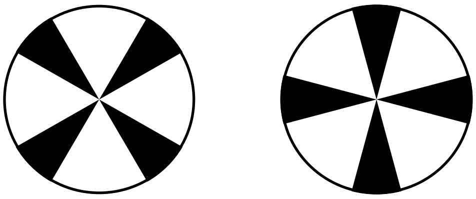


Fig 5. 17 Our perception of what is figure and what is ground changes with the orientation of the figure suggesting that orthogonal relationships are preferred

Symmetry

We group together symmetrically arranged items and find it easier to make sense of symmetrical groupings than asymmetrical ones (Figure 5. 18).

{ } { } { } { }

} } { } } { { }

Fig 5. 18 Symmetry is a strong grouping principle. It is very much easier to make sense of the top line of the figure than the bottom one

It is known that we make less eye movements when dealing with symmetrical figures and it is probable that they take less cognitive resources to process. Like all the principles, symmetry can also be upset by context. Thus the dot in the centre of the square in Figure 5. 19 can appear not to be in the centre when an additional off-centre square is added.

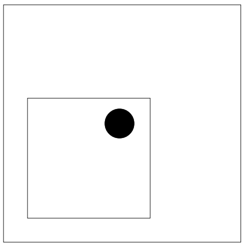


Fig 5. 19 The addition of an off-centre square upsets our perception of the central dot

Familiarity and context

Familiarity with a scene and its context affects our grouping — sometimes bringing about substantial changes in our understanding of what we see (Figure 5. 20).

Fig 5. 20 This familiar image is hard to recognise in this orientation. Turn the image through 90 degrees clockwise to view in ‘correct’ orientation

Our sense of ‘objectness’ often comes into play: knowing that we are looking at a partially occluded object immediately affects our ability to group. This is well illustrated in Figures 5. 21 and 5. 22.

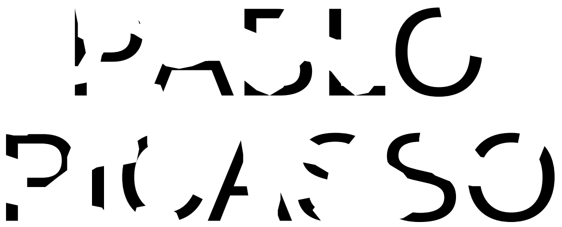


Fig 5. 21 It is virtually impossible to make sense of this image . . .

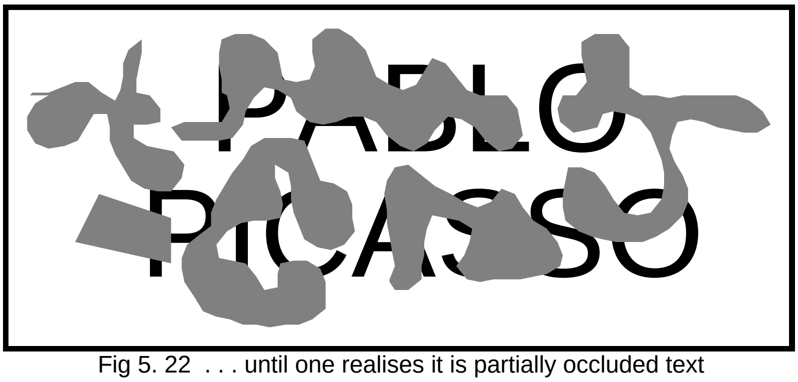


Fig 5. 22 . . . until one realises it is partially occluded text

The context with which we frame a drawing is also significant. The squares and diamonds in Figure 5. 23 are perceived differently according not, as might be supposed, to their relation to some absolute frame of reference, but to their placement within the drawing frame.

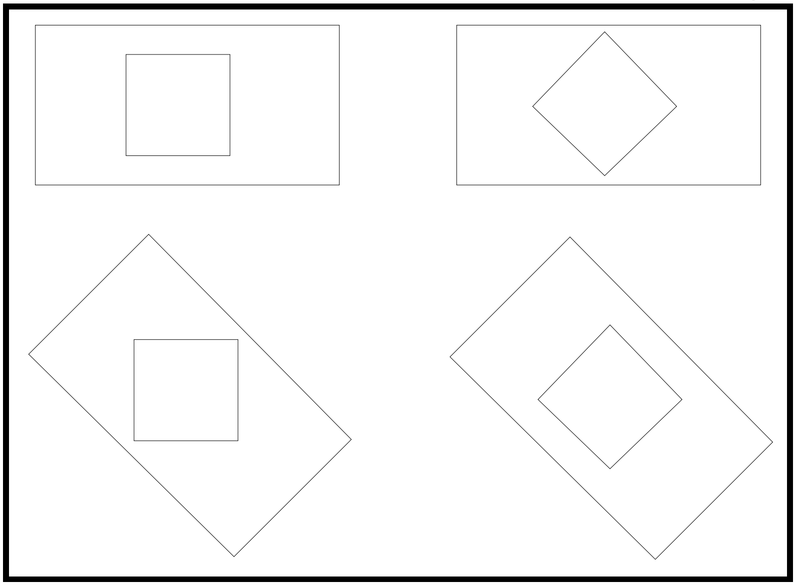


Fig 5. 23 The relationship of the internal figure to the orientation of the frame determines whether we will perceive it as a square or a diamond

Palmer (1992) has shown that context is also significant in allowing us to orientate ambiguously pointing shapes such as equilateral triangles. When seen alone, an equilateral triangle can appear to point in any of three directions (In Fig 5. 24 a the triangle can be seen as pointing to 3 o’clock, 7 o’clock or 11 o’clock). When seen grouped in company with similarly orientated triangles, all seem simultaneously to point in one or other of these directions (Fig 5. 24b). The preferred direction of pointing changes when the triangles are aligned along an axis (Fig 5. 24c). When they are aligned along one of their sides, they seem to point in a direction at right angles to the alignment (Fig 5. 24d).

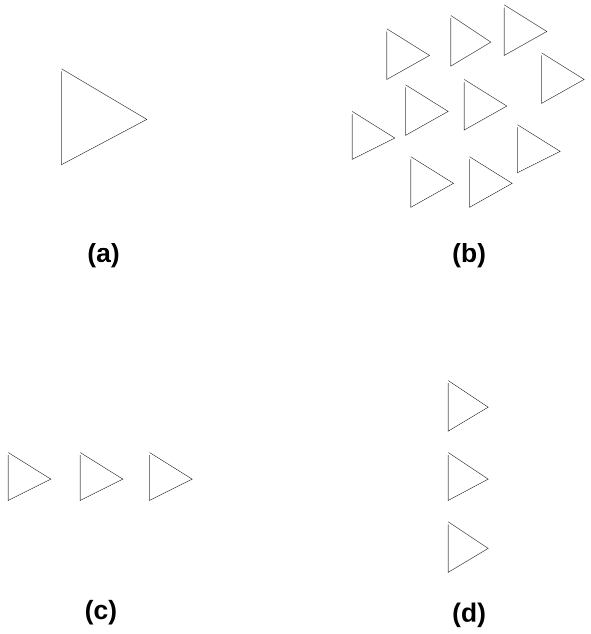


Fig 5. 24 Equilateral triangles have ambiguous orientations that can be directed by appropriate grouping

Note the even more ambiguous effect of arranging similarly orientated triangles around a circle (Fig 5. 25).

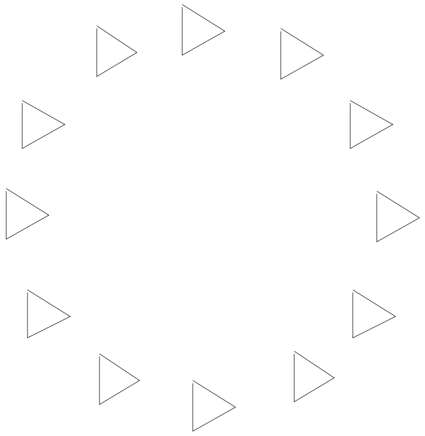


Fig 5. 25 It is hard to perceive these equilateral triangles as having the same orientation

Common Fate

Items that move together are grouped together. This organising principle cannot be adequately illustrated in a still picture but requires animation or movies to be fully appreciated. One of the most impressive manifestations of the principle of common fate was performed by Johansson (1973a, 1975) who attached lights to joints of a black-clad actor and filmed him as he moved across a darkened room. When the actor was stationary, no pattern could be discerned in the lights. But, as soon as he walked, it was easily possible to identify the very sparse pattern as that of a moving figure. Pavlova (1992) has shown that, when children aged 3-5 years were tested with animated cartoons consisting of moving dots attached to the main joints of an invisible man and an invisible animal moving as if on a treadmill, the 3-year olds were able to recognise the moving displays and the 5-year old's performance was as good as adults. Static versions of the display however were not recognised. Familiarity and context must also play a part here.

The minimal principle

At the heart of all these principles is one which the Gestaltists termed the 'minimal principle'. This can be summarised in the phrase:
'Of several geometrically possible organisations, the one that will be perceived is that which possesses the best, simplest and most stable shape'.

Thus four dots arranged at the four corners of an invisible square will be seen to represent a square (Figures 5. 26).

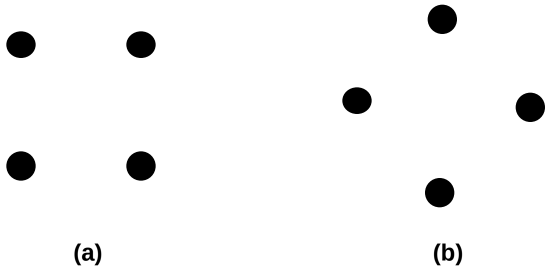


Fig 5.26 We perceive (a) not as isolated dots but as a square — even when rotated (b)

This law (as well as others) applies not only to the perception of external objects but to mental images of three-dimensional objects too. Parslow (1988) in an oft-repeated and continuing series of *ad hoc* experiments has shown that, when asked to mentally rotate a cube onto one of its corners, most people then perceive it to be an octahedron (Figure 5. 27).

Fig 5.27 People have no difficulty in mentally picturing a cube in its canonical position as shown on the left, but , if asked to mentally rotate it so that it stands on one of its corners, many people imagine a shape like that on the right

Hinton (1979), who first described this experiment, gives an explanation for the apparent change in form which is essentially based on the minimal principle.

Hochberg (1957 p83) gives a more general statement of the minimal principle:

> . . . other things equal, that perceptual response to a stimulus will be obtained which requires the least amount of information to specify

The parsimonious approach of the minimal principle seems to pervade the whole of perception and Hatfield and Epstein (1985) look at why the visual system might favour economical representations.

Precedence

Pomerantz (1985, 1986) has suggested that it is the extent to which we divide our attention between parts of a figure that determines the way we organise it and that it is the composition of the parts that affect this. Whether we see the elements rather than the whole depends on whether local or global precedence dominates. In Pomerantz (1983), he distinguishes between those elements in a figure whose *position* contributes to the perception of the whole (type-P elements) and those whose *nature* as well as position contributes to the perception (type-N elements).

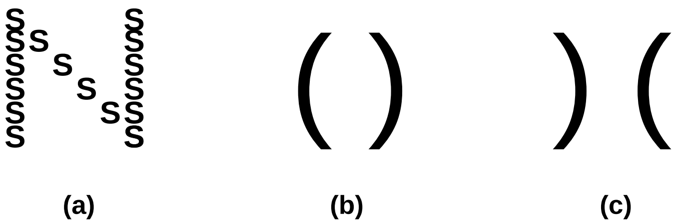

(a) (b) (c)

Fig 5. 28 Type-P (a) and type-N (b) and (c) elements produce different perceptual effects

Thus, in the case of Figure 5. 28 (a), it is only the position of the letters that determine the overall perceived form and any other set of letters could be substituted. In Figure 5. 28 (b) on the other hand, it is the shape of the elements that hint at the closed form, which they do not do when reversed as in Figure 5. 28 (c).

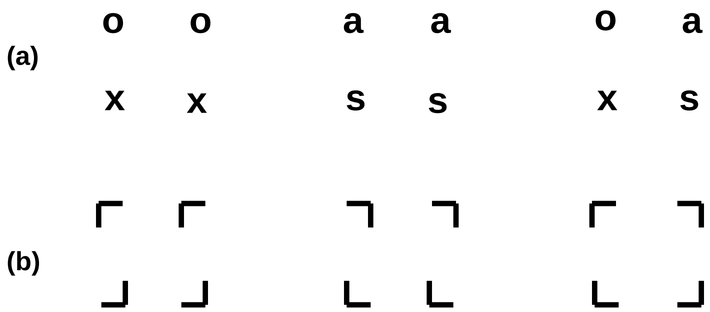


Fig 5. 29 Pomerantz' illustrates how type-P elements (a) differ from type-N

In Fig 6 of Pomerantz (1983), redrawn here as Figure 5. 29, he shows how his division into type-N and type-P affects our perception and that whilst we perceive all of the type-P elements as squares only the rightmost arrangement of the type-N elements gives this outcome (Figure 5. 29 (b)).

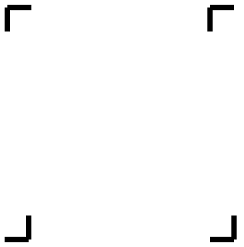


Fig 5. 30 Showing that, under some circumstances, even type-N elements can act as type-P

The scale of the elements in relation to the whole, though, affects the issue, as Figure 5. 30 shows. Here, Pomerantz's type-N elements of

Figure 5. 29 have been moved apart far enough for them to act simply as markers to the corner of a square — where they become type-P elements. It appears to me, then, that the type-N elements of Figure 5. 29 (b) work more by the nature of closure and good continuation than anything else. When they are placed far enough apart for these factors not to work, they become type-P.

The practicalities of perceptual grouping

Of course, in viewing ordinary pictures or in real life, the individual organising principles we use cannot be as readily identified as they can in all the foregoing examples. In addition, more than one principle is usually brought into play. However, it is now difficult to doubt that ordering principles are involved in our perception and cognition and that we do have preferred ways of perceiving arrangements. Knowledge of the way in which we do this can help guide the design or presentation of information.

Exploiting perceptual grouping

Thus, when we design any information presentation we should arrange the elements to take into account the way in which they are likely to be perceptually grouped by those who view them. For example, by closely aligning images and their associated captions, we can ensure that they are read together. In addition, as illustrated in Figure 5. 31, by keeping related elements together we can ensure that they are perceived as being related.

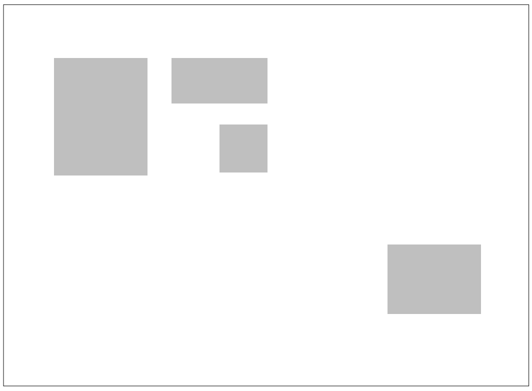


Fig 5. 31 The proximity principle ensures that we see the three parts to the left as a single related group

Tabular information is also enhanced by taking into account Gestalt principles. Thus, by use of the proximity principle, we can ensure that the emphasis is on particular directions, vertical or horizontal, in the table. Thus, in Figure 5. 32, by grouping the elements closer horizontally than vertically we can place perceptual emphasis on the rows of the table.

	90	**92**	**94**
France	24	26	25
Germany	27	25	29
Holland	12	14	17

Fig 5. 32 Showing that, even with only three columns, it is possible to place perceptual emphasis on the rows (the country series)

This is in contrast to Figure 5. 33 where the emphasis is on the columns of figures.

	1990	**1992**	**1994**
France	24	26	25
Germany	27	25	29
Holland	12	14	17

Fig 5. 33 Here, the perceptual emphasis is on the columns (the date series)

Because of the difficulties that sometimes occur in computer setting when placing emphasis on rows, it is often more convenient to rely on other cues rather than proximity for these cases (Figure 5. 34).

	1990	**1992**	**1994**
France	24	26	25
Germany	27	25	29
Holland	12	14	17

Fig 5. 34 Here, rules rather than proximity are used to emphasise the importance of the row data

Because, in the West, we are used to reading from left to right and top to bottom, images arranged on a page or on screen will tend to be read in this order too. If we wish to contradict this ‘natural’ tendency we must use make use of grouping principles such as proximity or closure. Thus, the images arranged as shown in Figure 5. 35 (a) will tend to be viewed in the order 1, 2, 3, 4 illustrated in Figure 5. 35 (b).

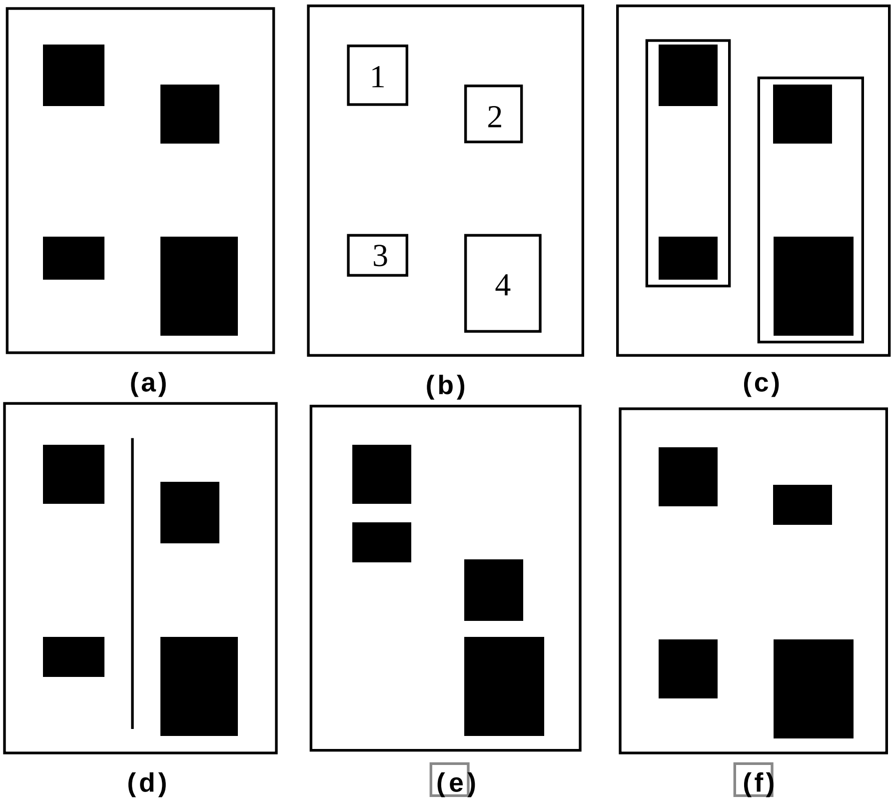


Fig 5. 35 Arranging illustrations on a page

If we wish them to be read in the order 1, 3 , 2, 4, we can either put them in boxes in groups as in Figure 5 .35 (c), or if that is likely to be problematical, divide the page into two with a rule as in Figure 5. 35 (d). This device will interrupt the flow and have the effect of making the viewer see a two column layout where the natural reading is top to bottom, left to right. Perhaps the better solution is to rearrange the illustrations either making use of proximity, as in Figure 5. 35 (e) or the ‘natural’ ordering as in Figure 5. 35 (f).

It should be emphasised that the ‘natural’ ordering is probably culture-dependent whereas the basic gestalt principles seem to be universal. However, in his now classic studies, Buswell (1935) examined the eye fixations of subjects when viewing many different types of images, from art works to advertisements. He did this with both Western and Oriental subjects and could find no consistent differences in the way they viewed the images. He accepted, though, that his sample of Oriental subjects might have been too small to draw valid conclusions on cultural differences.

Because of the principle of similarity, the odd object in a field of similar objects will often tend to 'pop-out' of the field and be easier to locate, as in Figure 5. 36.

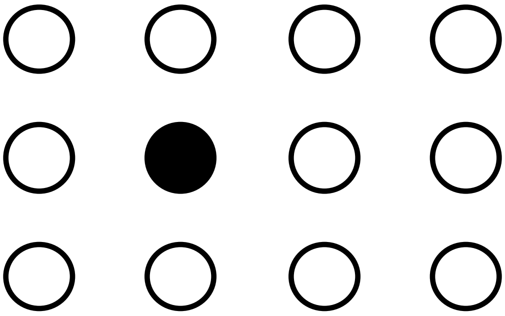


Fig 5. 36 Because of its dissimilarity to the rest, the black circle tends to pop-out — making it particularly easy to see

Thus an icon will tend to pop-out of a field of other icons providing it is sufficiently dissimilar to them (Figure 5. 37).

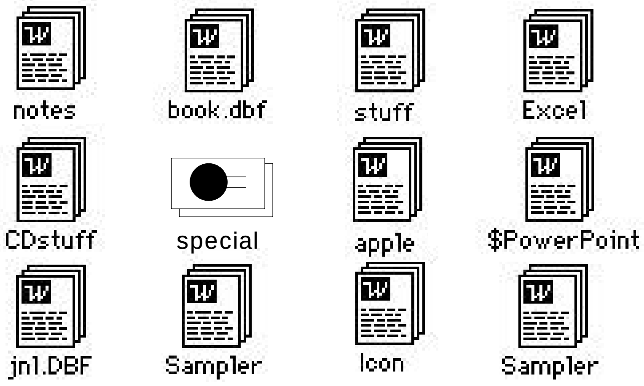


Fig 5. 37 An icon dissimilar to a field will tend to pop-out

The extent and type of the dissimilarity is important. It might be thought, for example, that a colour change is sufficient. Theeuwes (1995) shows that the significant element, though, is brightness (luminance) rather than colour. This can be confirmed by Apple Macintosh users. If the 'Label' facility of the Finder is used to change the colour of one icon in a field of icons, it will be seen that this does not generally improve noticeability unless the colour change also radically alters the relative brightnesses of the field and the target icon. Figure 5. 37 also illustrates the fact that a set of similar icons will be identified with a common purpose or subject. Thus, not only does the dissimilar icon pop-out, but we also assume that it represents something different in kind from the rest.

The importance of movement and attention

We have already commented on the fact that, in order for us to see anything at all, our eyes must be kept in a more or less constant state of movement. We make these movements not only voluntarily by turning our eyes, head and body, but also involuntarily by virtue of the saccades. It is clear, too, that perception requires attention (Neisser and Beklen 1974). We do not really see things that we do not attend to. We can take a Gibsonian standpoint and say that these movements are necessary in order for us actively to sample the optic array. Alternatively, we can take a more conventional standpoint and say that they are necessary in order for us to create retinal images for processing edges, contours and so on. At all events, we must acknowledge the importance of eye movements.

When we look at any image we do not perceive it immediately as a whole but in sections although, while of doing so, we do not seem to be aware of the fact. This is to say we scan an image rather than take it in completely all at once. This scanning process is not like the ordered rastering of a television camera. It is rather more like a sampling process where we jump from one part of the image to another, lingering momentarily on one point before rapidly jumping to another. The jumpy movements are known as ‘saccades’, and the periods of lingering, ‘fixations’. Saccades only take about 5 per cent of the total time we look at an image so they are very rapid indeed and normally they don’t move us more than about 15° at a time. During the saccades the image on the retina must be blurred but, as we are unaware of this, there must be perceptual mechanisms to cancel out the effect. We fixate for between 200 and 500 milliseconds.

The whole process of saccades and fixations is remarkable in that we seem not to need to examine large areas of an image. We can fixate very rapidly on areas that give us the most information even though these may not be the ones that have the most detail. In particular, we fixate longest on those parts of an image that are the least expected. Thus, given two images like those in Figure 5. 38, we make more eye fixations on the asymmetrical pattern (a) than on the symmetrical pattern (b).

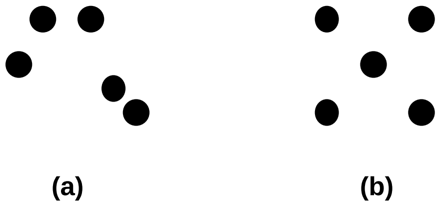


Fig 5. 38 We make more eye fixations to understand pattern (a) than pattern (b)

Eye movements and perceptual grouping

These factors may account for some of the perceptual grouping phenomena, particularly if we assume that some sort of ‘principle of least effort’ (another version of the minimal principle) applies. For example, proximity would be favoured because saccades would sweep narrower angles. Symmetry would be preferred to asymmetry and order to chaos because less fixations would be needed for us to come to an understanding of what we are seeing.

In a series of interesting experiments, Biederman (1972), presented subjects with images of scenes for periods of time so short (typically in the region of 200-300 milliseconds) that only one fixation was possible. The subjects were then asked to say whether or not one of a group of four items was in the scene. Because only one fixation was possible, the position where the object to be detected would appear was cued by the presence of an arrow shown just before the scene appeared. Two forms of image were shown: conventional photographs and the same photographs cut into six rectangular sections and randomly reassembled.

Biederman found that the detection rates for objects were significantly lower for the random pictures than the originals, presumably because, even in the short periods that the image was seen, context and familiarity played an important part in what was expected by the subjects.

Yarbus (1967) confirms that the composition (in the art sense) of an image makes a difference on how it is viewed. Showing his subjects a picture called ‘An Unexpected Visitor’ by IE Repin he found that, when freely viewing the image (which has the basic configuration shown in Figure 5. 39):

> the scale of the figures and objects and their position in the picture — in other words, everything which we call the composition of the picture — also has definite importance. The figure of the woman, being the largest and being centrally situated, attracts more attention than any of the other figures, Hence composition is the means whereby the artist to some extent may compel the viewer to perceived what is portrayed in the picture (p193).

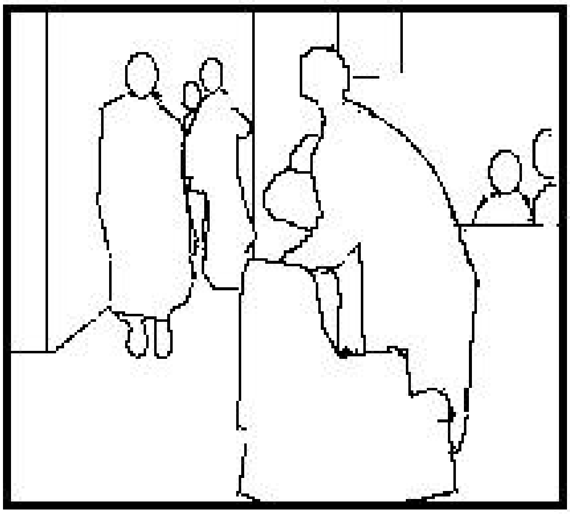


Fig 5. 39 The basic composition of Repin’s picture ‘An Unexpected Visitor’. The visitor is the figure furthest left. There are two women to the right of the visitor and two young girls seated at the table. The woman in the foreground is rising to greet the visitor and immediately to her left is a figure seated at a grand piano on which there is a sheet of music

As might be expected, viewers fixated frequently on the faces of the major figures in the picture. Scanning patterns were different from the free-viewing case if a search was being made for particular information in the picture. When, for example, viewers were asked to estimate the ages of the people shown they spent even greater percentage of fixations on the faces. When asked to estimate the social circumstances of the family, less fixations were made to the faces and more to the objects in the environment. Yarbus (p192-193) says:

> In response to the instruction ‘surmise what the family was doing before the arrival of the “unexpected visitor”’, the observer directed his attention particularly to the objects arranged on the table, the girl’s and the woman’s hands, and to the music. . . . Finally, the instruction ‘estimate how long the “unexpected visitor” had been away from the family,’ caused the observer to make particularly intensive movements of the eyes between the faces of the children and the face of the person entering the room. In this case he was undoubtedly trying to find the answer by studying the expressions on the faces and trying to determine whether the children recognised the visitor or not.

Yarbus also reports that when viewers are given a long time to view a picture, they do not spend time viewing secondary features in it but repeatedly return to the same parts of the image. Gale (1993 p126) says that:

> Fixation patterns are under voluntary control . . . In 95% of cases, objects in a picture which are subsequently remembered are fixated by a 3rd fixation and are then refixated several times.

This returning to the same parts of the image might result from the formation of cognitive models.

Scan paths and cognitive models

Although, in looking at traces of the paths of saccades, it might be initially assumed that they are random and uninfluenced by the viewer, Stark and Ellis (1981) used Markov chain probability matrices to show that as viewers develop cognitive models about the image, they change the way it is scanned. Indeed, their studies made them conclude that, in examining difficult to understand pictures, three successive stages come into play. At first, before subjects recognise the image, they ‘look without seeing’. At this stage eye movements are diffuse as they try out different explanatory cognitive models. Then, suddenly, they are able to ‘see without looking’. Stark and Ellis believe that what has been ‘seen’ at this point is the viewer’s own cognitive model. Finally, viewers are able to ‘look and see’. In this phase, the viewer makes a more ordered sequence of eye movements which, Stark and Ellis suggest, are used to confirm the cognitive model.

In studies on the way eye-movements are used in reading music, Goolsby (1994), reports that patterns of eye movement in experienced and less-experienced musicians were similar across melodies and encounters, but differed with the complexity of the music. Eye movement was reduced when performing melodies with more-concentrated visual information than when performing melodies with less-concentrated visual information. The study showed that music readers used fewer but longer fixations after practising the melodies and that skilled music readers look farther ahead in the score, and then back to the point of performance, when sightreading.

Picture composition and scan paths

Brandt (1945) carried out a wide survey of the way people scan pictures and, in respect of a symmetrical layout, concluded that:

> The results of the study reveal that the median of the first fixation for all subjects falls at a point above and to the left of centre of the observed field (pp30-31).

Other later studies have confirmed a slight tendency for the start of a scan path to move upwards to the left. Brandt also showed that layout of items on a page influenced the amount of time they were viewed and, once again identified a preference for attending to elements on the left hand rather than the right hand side. On the other hand, given a four-column layout of photographs such as shown in Figure 5. 40, more attention was paid to photographs placed in the middle two columns than to the the same photographs when they were placed in an alternative layout in the outer two columns.

> Whenever any one of the four columns is changed from an outside to an inside position, the percentage of time devoted to that area is increased. In cases where a columns is changed from a left outside position to an inside right position the increase is small (p43).

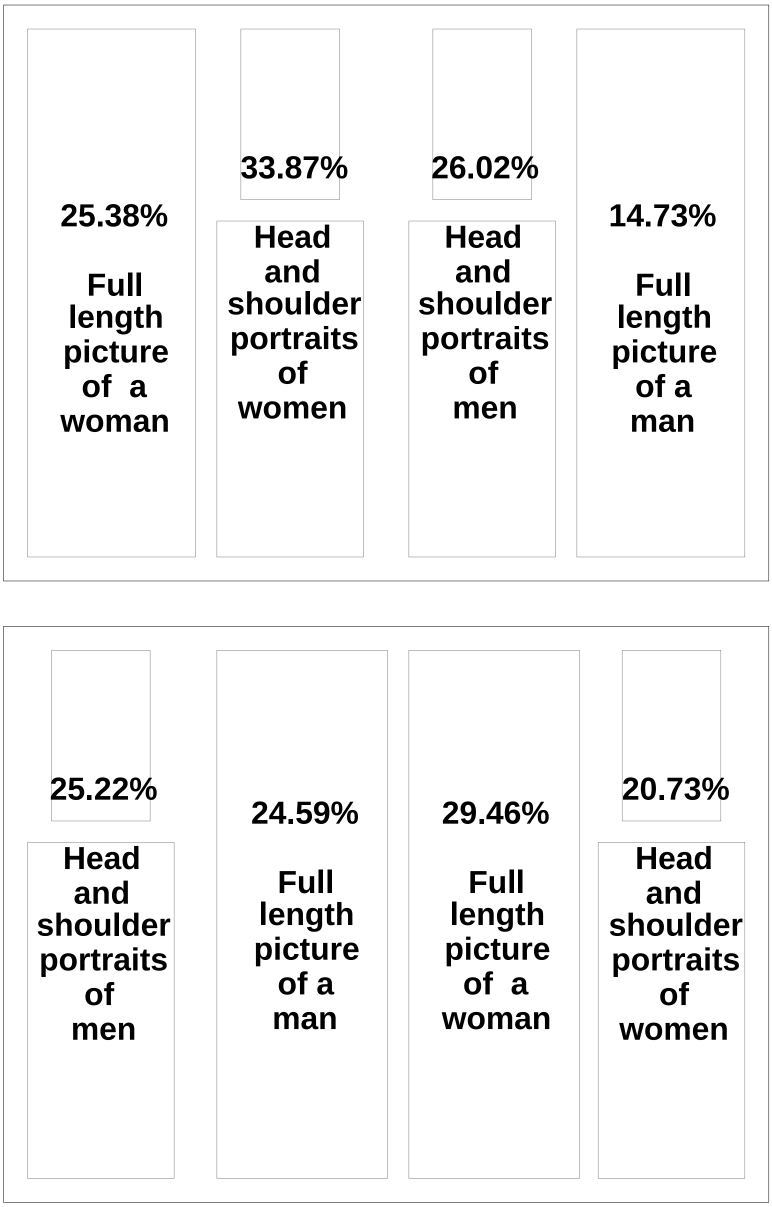


Fig 5. 40 Typical layout of photographs as investigated by Brandt. Greater attention was given to the pictures when they were placed in the centre than at the edges.

Brandt concluded that:

> Wherever an inside left competes against an outside right the difference in attention time is far greater than where an inside right competes with an outside left (p45).

All the studies on saccades and fixations confirm that a knowledge of their mechanisms will assist in the creation of image and interface design. It is surprising though that Brandt's work, which is now more than 50 years old, does not seem to have been followed up in detail by those concerned with graphic layout rather than the psychology of perception. Thus we still largely rely on Brandt’s work and more general art history principles (as well as the studies of Buswell — of more than 60 years ago) to describe the way composition and layout affect attention and, hence, understanding. This is strange because, undoubtedly, it is a rich field for further study.

Summary

The contribution of the Gestalt psychologists in describing the way we perceptually organise what we see cannot be overstated. The principles that they and later workers set down are of great significance to those who design. By working with the principles (and, perhaps, occasionally working against them), we can influence attention and guide the viewers’ understanding of images, layouts and interfaces. Without a knowledge of these principles we rely entirely on designer’s intuition — a notoriously fickle and unreliable source.

6. Concluding remarks and some theories

The seamlessness of visual perception

As we have seen in this Report, our visual perceptual system seems to be built up of separate mechanisms for dealing at least with layout, form, movement and colour. Different areas of the brain are involved in each of these mechanisms. Yet, in the day-to-day process of seeing, we are unaware of any separation between the elements that are being dealt with. We look and we see and, mostly, we easily understand what we are seeing. The elements of layout, form, colour and movement are seamlessly joined together and we appear to be instantaneously conscious of our surroundings. Certainly, we are unaware of anything like 'information processing' taking place. As Bruner and Potter (1964/1968) say:

> Under ordinary conditions, visual recognition operates effortlessly and with no discernible inference (p730).

But, as they go on to point out:

> If the clarity of the display is diminished in some manner, however, recognition understandably takes longer. Moreover, studies indicate that if a subject is *initially* exposed to a blurred image that he cannot recognise, subsequent recognition of the image in clearer form is substantially delayed (p730).

Certainly, in normal circumstances, there does not seem time for any 'processing' to happen. Our apparently instantaneous perception of the world, however, is illusory. We *do* take time to perceive and this time, when compared, say, to the time it takes a computer to perform some of its tasks, is appreciable.

The mystery of perception

Despite humankind's thousands of years of study of visual perception, though, we cannot yet explain *how* things are processed. Nor do we yet know how the signals that come to us via our eyes are integrated with what happens in our heads to produce the seamless awareness of the world that we see. Theories exist — some of them are touched on in what follows — but, in general, the whole process is still a mystery. We are, in many ways, like the legendary blind men of Toledo when confronted by a previously unknown object, namely an elephant. By touching parts of it they each concluded that it was something different. They knew much about the parts but were unable to integrate them into a picture of a single, whole animal. We now know much about the parts that make up our perceptual mechanisms but are still unable to explain how these fit together.

Types of theory

In briefly examining theories of visual perception we can see three rough categories:

- **Cognitive theories**
  those that take the view that perception is just another form of inferential process, albeit a largely unconscious one — a view originally set out by Helmholtz.
- **Ecological theories**
  those that assume that we have a special relationship with the environment and that we are, as it were, directly in touch with it — this is the view propounded by Gibson.
- **Computational theories**
  those that assume that perception is essentially an information processing task not fundamentally different from that which computers deal with — the standpoint adopted by Marr and others.

In addition, there are some theories that do not fall into any of the above categories. An example is the theory of Ramachandran (1990) who takes what he calls a ‘utilitarian’ viewpoint.

> According to this view perception does not involve intelligent reasoning as implied by some psychologists; does not involve resonance with the world as suggested by Gibsonians; and does not require creating elaborate internal representations or solving equations as implied by AI researchers. On could argue, instead, that perception is essentially a ‘bag of tricks’; . . . (p347)

Cognitive theories

Cognitive theories spring from the notion that part of what we perceive comes from what we see ‘out there’ and part from what is already in our minds. There is obvious force in this idea. It is clear that our perception *is* influenced by what we know — some of the examples in the previous section on the Gestalt approach confirm this. Those influential perception psychologists who take the Helmhotzian stance that we interpret what we see in the light of what we know include the Gregory (1970) in the UK and Rock (1983) in the US.

Writing about Helmholtz, Gregory (1987) says that he was,

> . . . philosophically a thoroughgoing empiricist, believing that sensory signals only have significance as the result of associations built up by learning. We are essentially separate from the world of objects, and isolated from external events, except for neural signals which, somewhat like language, must be learned and read and according to various assumptions, which may or may not be appropriate (p309).

The Russian worker, Demidov (1979 / 1986), tells us that cognition:

> . . . is a chain of hypotheses that are tested to be rejected as invalid or accepted and used as a guidance. Vision performs the same job. We do not notice this only because it goes on only at the at the subconscious level. The ‘resonable eye’ constructs hypotheses about space and arrangements of objects in it (p198).

Clearly, however, there are some aspects of visual perception that are more direct than this (at least they seem so). There are aspects of vision where we do not seem to have to form an hypothesis and test it with further information. We seem inately able to do certain things — scan and fixate for instance. Aspects such as this led to the development of ecological theories.

Ecological theories

Gibson (1979) took the view that:

. . . perceiving is an act, not a response, an act of attention, not a triggered impression, an achievement not a reflex.

Gibson believed that we do not, as it were , interpret and hypothesise about the environment but ‘resonate’ to it. We do this by being able to detect elements in our surroundings that are invariant to rotation and movement. In addition, we use the whole abient array of light, the optic array to allow us to understand our surroundings.

The Gibsonian concept of vision is a global one. It doesn't accept that perception comes from an accretion of small items of information assembled via 'atomic' feature detectors in the visual system (which is the more conventional standpoint that seems in line with the psychophysical evidence). To Gibson, texture, perspective and movement gradients in the ambient optic array are sufficient. It must be said though, that many who generally acknowledge the importance of Gibson's insights over the last 35 years of his life do not accept all the implications of his theory.

Twenty years ago, Dodwell (1975) made the point that:

> Gibson's conception of the global nature of sensory stimulation is quite out of tune with the atomistic findings of sensory neurophysiology. It seems that the visual system does . . . operate as a detector of small elements of pattern information . . physiological evidence of the more elaborate processing that would be required in Gibson's system is so far lacking . . . Perceptual judgments based on texture, perspective and movement gradients are not as uniform as Gibson's theory predicted.

Later neurological studies seem to support Dodwell’s view.

However, despite these and other cogent criticisms, Gibson's view of how we perceive should not be dismissed. His ideas on the optic array and the optic flow field (the fluctuating pattern of light reaching us caused by any relative movement between ourselves and the environment) are of importance to us in computer graphics (Tsotstos 1984). They also help to explain some phenomena which are otherwise extremely puzzling. In addition, his approach seems to give us a clue as to how we can make computer drawn pictures of sufficient variety to be convincing without involving us in excessive computational overhead — we need textures and gradients in addition to perspective.

Johansson (1994), who seems to take a standpoint between the ecological and the computational, takes the view that the decoding principles of vision are buit-in to specific systems in the species. They are, as it were, hard-wired.

> Given a digital computer, the desired output can be generated by a number of differently constructed algorithms. This does not apply to visual perception as I understand it. In the visual system a flexible software does not exist. it lacks the capacity for basic restructuring or plasticity. This distinguishes my view from the pure computational approach (p311).

Computational theories

David Marr (1982) expressed the view that the fatal shortcoming of Gibson's theory is that it fails to realise:

> First, the detection of physical invariants, like image surfaces, is exactly and precisely an information processing problem . . . And, second, he vastly underrated the sheer difficulty of such detection.

Marr (1945-1980), a British psychologist who worked at MIT, proposed in his posthumously published 1982 book, a detailed, computational explanation of vision which he suggested took place in three stages. The initial stage when we see an object is to form what he called the 'primal sketch' of it. This entails using the intensity values of the shading on it to detect edges and rough shapes. The second stage derives what he called the '21/2 D sketch'. At this point we are aware of the orientation and depth of surfaces. Finally, from this information, we derive the 3 D model. Thus, although we may not know anything else about the object that we are looking at, we can still understand a great deal about it from just the image that it presents to us.

Biederman (1988) proposes that we recognise objects by what he calls RBC or Recognition-by-Components. RBC assumes that:

> an image of an object is segmented at regions of deep concavity into an arrangement of simple convex generalised cone primitives, such as cylinders, bricks, wedges and cones . . (p370).

This theory seems to owe something both to Marr and to Winston (1973). Other computational and information processing theories abound. Spoer and Lehmkuhle (1982) outline some of these.

Summary

All of these theories have parts that go some way to explaining how we understand what we see. Yet none of them seems fully to capture the remarkable ability we have to make sense of what is 'out there'. Perhaps the tidiness that theoreticians seek in the detail of their theories is not realisable. Ramachandran (1990) is perhaps right when he says that perception is a bag of miscellaneous tricks that have evolved over time and that;

> . . . the brain appears to bring to bear a whole repertoire of strategies which exploit many different regularities in the natural world. And for each perceptual task, the simultaneous use of a wide range of such short-cuts allows the sytstem to achieve the same biological goal as it could with a single sophisticated algorithm (p359).

I would go further and suggest that such simultaneous use of short-cuts is *far more likely* to achieve a particular biological goal than any single-minded approach. We are, I believe, unlikely ever to be able to set down for visual perception laws and theories of the same sort as Newton's Laws. Biological events are of a different nature to these.

But, nonetheless, designers need to understand something of the mechanisms of human visual perception. They need to be able to exploit its regularities and its inconsistencies in order to more effectively convey and enhance information. It is hoped that this Report will assist in this task.

7. References

Abse J (1980) *John Ruskin: The Passionate Moralist*, Quartet Books, London

Abramov I (1981) ‘Physiological mechanisms of colour vision’, in Richter M (ed) *AIC Colour 1981*, AIC, Berlin

Abramov I and Gordon J (1994) ‘Colour appearance: On seeing red — or yellow, or green, or blue’, *Annual Review of Psychology* (45) pp451-485

Addams R (1834) ‘An Account of a Peculiar Optical Phenomenon Seen after Having Looked at a Moving Body’, *The London and Edinburgh Philosophical Magazine and Journal of Science* (5) pp373-374

Agoston GA (1979) *Colour Theory and its Applications in Art and Design*, Springer, Berlin

Andersen GJ and Braunstein ML (1983) ‘Dynamical occlusion in the perception of rotation in depth’, *Perception & Psychophysics* (34) 4 pp356-362

Anderson RA (1989) ‘Visual and eye movement functions of the posterior parietal cortex’ in Cowan WM, Shooter EM, Stevens CF and Thompson RF (eds), *Annual Review of Neurosciences* (12) pp377-403

Antes JR (1974) ‘The time course of picture viewing’, *Journal of Experimental Psychology* (103) pp 62 - 70

Antes JR and Kristjanson AF (1991) ‘Discriminating artists from non-artists by their eye-fixation patterns’, *Perceptual and Motor Skills* (73) pp893-894

Arden GB (1972) ‘The retina — Neurophysiology’, in Davson H (ed), *The Eye (2nd edition) Vol 3A: Visual Function in Man*, Academic Press, New York pp229-256

Arnheim R (1956) *Art and Visual Perception : A Psychology of the Creative Eye*, Faber and Faber, London

Attneave F (1954) ‘Some informational aspects of visual perception’, *Psychological Review* (61) pp 183 - 193

Barlow HB (1982) ‘Physiology of the retina’, in Barlow HB and Mollon JD (eds), *The Senses*, CUP, Cambridge

Barlow HB, Blakemore C and Pettigrew JD (1967) ‘The neural mechanism of binocular depth discrimination’, *Journal of Physiology* (193) pp327-342

Berk T, Brownston L and Kaufman A (1982) ‘A new colour-naming system for graphics languages’, *IEEE Computer Graphics and Applications* (2) 3 pp37-44

Berlin B and Kay P (1969) *Basic Colour Terms : Their Universality and Evolution*, University of California Press, Berkeley

Berthenthal BI, Proffitt DR, Spetner NB and Thomas MA (1985) ‘The development of infant sensitivity to biomechanical motions’, *Child Development* (56) p531-543

Biederman I (1972) ‘Perceiving real-world scenes’, *Science* (177) pp77-80

Biederman I (1988) ‘Aspects and extensions of a theory of human image understanding’, in Pylyshyn ZW (ed), *Computational Processes in Human Vision: An Interdisciplinary Perspective,* Ablex, Norwood pp370-428

Blakemore C (1990) ‘Understanding images in the brain’, in Barlow H, Blakemore C and Weston-Smith M (eds) *Images and Understanding,* CUP, Cambridge pp257-283

Bonnet C (1975) ‘A tentative model for visual motion detection’, *Psychologia* (18) 1 pp35-50

Boring EG (1942) *Sensation and Perception in the History of Experimental Psychology,* Appleton-Century-Crofts, New York

Boynton RM (1983) ‘Mechanics of chromatic discrimination’, in Mollon D and Sharpe LT (eds), *Colour Vision: Physiology and Psychophysics,* Academic Press, London pp409-423

Braddick OJ (1982) ‘Binocular vision’, in Barlow HB and Mollon JD (eds), *The Senses,* CUP, Cambridge pp192-200

Brandt H (1945) *The Psychology of Seeing,* Philosophical Library, New York

Bruce M and Foster JJ (1982) ‘The visibility of coloured characters on coloured backgrounds in viewdata displays’, *Visible Language,* (XVI) 4 pp382-390

Bruner JS and Potter MC (1964/1968) ‘Inference in visual recognition’, Science (144) pp424-425, reprinted in Haber RH (ed), *Contemporary Theory and Research in Visual Perception,* Holt, Rinehart and Winston, New York pp730-733

Buchsbaum G and Gottschalk A (1983) ‘Trichromacy, opponent colour coding and optimum colour information transmission in the retina’, *Proceedings of the Royal Society London* (B220) pp89-113

Bui-Tuong Phong (1971) ‘Illumination for computer-generated pictures’, *Communications of the ACM* (118) 6 pp 311 - 317

Burr and Ross (1986) ‘Visual processing of motion’, *Trends in Neurosciences* (9) 7 pp304-307

Buswell GT (1935) *How People look at Pictures: A Study in the Psychology of Perception in Art,* University of Chicago Press, Chicago

Bülthoff HH and Mallot HA (1990) ‘Integration of stereo, shading and texture’, in Blake A and Troscianko T (eds), *AI and the Eye,* Wiley, New York pp119-146

Calder R (1970) *Leonardo and the Age of the Eye,* Heinemann, London

Caelli T (1981) *Visual Perception: Theory and Practice,* Pergamon, London

Campbell PF (1979) ‘Artistic motion cues, number of pictures, and first-grade children’s interpretation of mathematics textbook pictures’, *Journal for Research in Mathematics Education* (10) 2 pp148-151

Caviano JL (1994) 'Colour and sound: Physical and psychophysical relations', *Colour Research and Applications* (19) 2 pp126-133

Chen T-L and Yu C-Y (1996) 'The relationship between visual acuity and colour contrast in the OSA uniform colour space', *Colour Research and Application* (21) 1 pp18-25

Choungourian A (1968) 'Colour preference and cultural variation', *Perceptual and Motor Skills* (26) pp1203-1206

Clark K (1939) *Leonardo da Vinci: An Account of his Development as an Artist*, CUP, Cambridge

Cooper RM (1974) 'The control of eye f ixation by the meaning of spoken language', *Cognitive Psychology,* (6) pp 84 - 107

Corbetta M, Miezin FM, Dobmeyer S, Shulman GL and et al (1990) 'Attentional modulation of neural processing of shape, color, and velocity in humans', *Science* (248) 4962 pp1556-1559

Courtney AJ (1986) 'Chinese population stereotypes: Colour associations, *Human Factors* (28) pp97-99

Cowlishaw M (1984) *Fundamental requirements for picture presentation*, IBM UK Scientific Centre, Winchester

Cutting JE, Proffitt DR, Kozlowski LT and Johansson G (1977) 'Studies on visual perception of locomotion', *Perception* (6) 4 pp365-376

Cutting JE and Vishton PM (1995) 'Perceiving layout and knowing distances', in Epstein W and Rogers S (eds), *Perception of Space and Motion* (2nd Edition), Academic, San Diego pp69-117

Davidoff J (1991) *Cognition through Colour,* MIT Press, Cambridge

Davies IR, Corbett GG, Laws G, McGurk H and et al (1991) 'Linguistic basicness and colour information processing', *International Journal of Psychology* (26) 3 pp311-327

Day RH (1994) 'The foundations of veridical and illusory perception', in Ballesteros S (ed), *Cognitive Approaches to Visual Perception*, Lawrence Erlbaum, Hillsdale pp243-268

DeAngelis GC, Ohazawa I and Freeman RD (1995) 'Neuronal mechanisms underlying stereopsis: How do simple cells in the visual cortex encode binocular disparity?', *Perception* (24) 1 pp3-31

de Bruyn B and Orban GA (1990) 'The role of direction information in the perception of geometric optic flow components', *Perception and Psychophysics* (47) 5 pp433-438

Demidov V (1979/1986) *How We See What We See* (trans) Repyev A, Mir Publishers, Moscow

De Valois RL and De Valois KK (1975) 'Neural coding of colour', in Cartarette EC and Friedman MP (eds), *Handbook of Perception: Vol V: Seeing,* Academic Press, New York pp 117 - 166

De Valois RL and De Valois KK (1993) 'A multi-stage colour model', *Vision Research* (33) 8 pp1053-1065

de Weert CMM (1984) 'Limits in vision with two eyes', in Van

Doorn AJ, van de Grind W and Koenderink JJ (eds), *Limits in Perception*, VNU Science, Utrecht pp235-282

de Weerd P, Vandenbussche E andOrban GA (1993) ‘Occlusion cues contribute to orientation judgements of occlusion-defined contours’, *Perception and Psychophysics*, (54) 6 pp706-715

Dittrich WH (1993) ‘Action categories and the perception of biological motion’, *Perception* (22) 1 pp15-22

Dodwell PC (1975) ‘Contemporary theoretical problems in seeing’, in Cartarette EC and Friedman MP (eds), *Handbook of perception: Vol V: Seeing*, Academic Pres, New York pp 57 - 77

Dodwell PC (1983) ‘The Lie transformation group approach to visual perception’, *Perception and Psychophysics* (34) pp1-16

Dodwell PC (1986) ‘The Mickey Mouse theory of perception’, *Canadian Psychology* (27) 4 pp319-331

Dowling JE (1987) *The Retina*, Harvard University Press, Harvard

Dowling JE and Boycott BB (1966) ‘Organisation of the primate retina: Electron microscopy’, *Proceedings of the Royal Society Series B* (166) pp80-111

Dupont P, Orban GA, De Bruyn B and Verbruggen A et al (1994) ‘Many areas in the human brain respond to visual motion’, *Journal of Neurophysiology* (72) 3 pp1420-1424

D’Zmura (1991) ‘Colour in visual search’, *Vision Research* (31) 6 pp951-966

Elio RE and Reutener DB (1978) ‘Colour context as a factor in encoding and as an organisation device for retrieval of word lists’, *Journal of General Psychology*, (99) 2 pp223-232

Engel AK (1994) ‘Review of Functional Organisation of the Human Visual Cortex’, *Trends in Neuroscience* (17) 1 pp43-44

Epstein W (1988) ‘Has the time come to rehabilitate Gestalt theory?’, *Psychological Research* (50) 1 pp2-6

Felix N (1995) *Analysis of Spatial Representation: From Traditional Techniques to Computer Graphics*, Unpublished PhD thesis, Strathclyde University, Glasgow

FoleyJD, van Dam A, Feiner SF and Hughes JF (1990) *Computer Graphics: Principles and Practice* (2nd Edition), Addison-Wesley, Reading

Foley JM (1980) ‘Binocular distance perception’, *Psychological Review* (87) 5 pp411-434

Friedman SL and Stevenson MB (1975) ‘Developmental changes in the understanding of implied motion in two-dimensional pictures’, *Child Development* (46) pp773-778

Fussell D and Haaland A (1978) ‘Communicating with pictures in Nepal: results of practical study used in visual education’, *Educational Broadcasting International* (11) 1 pp25-31

Gale AG (1993) ‘Human response to visual stimuli’, in Hendee WR and Wells PNT (eds) *The Perception of Visual Information*,

Springer, New York, pp115-133
Gerritsen F (1979) ‘Evolution of the colour diagram’, *Colour Research & Application* (4) 1 pp33-38
Gibson EJ, Gibson JJ, Smith OW, and Flock H (1959) ‘Motion parallax as a determinant of perceived depth’, *Journal of Experimental Psychology* (58) pp40-51
Gibson JJ (1950) *The Perception of the Visual World,* Houghton Mifflin, Boston
Gibson JJ (1966) *The Senses considered as Perceptual Systems,* Houghton Mifflin, Boston
Gibson JJ (1979) *The Ecological Approach to Visual Perception,* Houghton Mifflin, Boston
Gillam B (1995) ‘The perception of spatial layout’, in Epstein W and Rogers S (eds), *Perception of Space and Motion* (2nd Edition), Academic, San Diego pp23-67
Gogel WC and Tietz JD (1980) ‘Absolute motion parallax and the specific distance tendency’, *Perception and Psychophysics* (28) 4 pp321-328
Goolsby TW (1994) ‘Eye-movement in music reading: effects of reading-ability, notational complexity, and encounters’, *Music Perception* (12) 1 pp77-96
Gordon IE (1989) *Theories of Visual Perception,* Wiley, Chichester
Gouraud H (1971) ‘Continuous shading of curved surfaces’, *IEEE Transactions on Computers* (C-20) 6 pp 623 - 628
Gregory RL (1970) *The Intelligent Eye,* Weidenfield and Nicholson, London
Gregory RL (ed) (1987) *The Oxford Companion to the Mind,* OUP, Oxford
Grieve KW (1991) ‘Traditional beliefs and colour perception’, *Perceptual and Motor Skills* (72) 3 pt 2 pp1319-1323
Gross CG, Rocha-Miranda CE and Bender DB (1972) ‘Visual properties of neurons in inferotemporal cortex of the macaque’, *Journal of Neurophysiology* (35) pp 96 - 111
Gulyás B, Ottoson D and Roland PE (1993) *Functional Organisation of the Human Visual Cortex,* Pergamon, Oxford

Haber RN (1987) ‘Why low-flying fighter planes crash: Perceptual and attentional factors in collisions with the ground’, *Human Factors* (29) 5 pp519-532
Haber RN and Wilkinson L (1982) ‘Perceptual components of computer displays’, *IEEE Computer Graphics and Applications,* (2) pp23-35
Hagen MA (1986) *Varieties of Realism: Geometries of Representational Art,* CUP, Cambridge
Harris M, Freeman T and Williams G (1992) ‘Surface layout and retinal flow’, in Humphreys GW (ed), *Understanding Vision,* Blackwell, Oxford pp199-231
Harris MG (1994) ‘Optic and retinal flow’, in Smith AT and

Snowden RJ (eds), *Visual Detection of Motion,* Academic, London pp307-332

Hatfield G and Epstein W (1985) ‘The status of the minimum principle in the theoretical analysis of visual perception’, *Psychological Bulletin* (97) 2 pp155-186

Hård A (1966) ‘Philosophy of the Hering-Johansson natural colour system’, *International Colour Meeting 1,* Musterschmidt, Göttingen pp 357-365

Hård A and Sivik L (1981) ‘NCS - Natural Colour System: A Swedish standard for colour notation’, *Colour Research & Application* (6) 3 pp129-138

Hecht S, Schlaer S and Pirenne MH (1942) ‘Energy, quanta and vision’, *Journal of General Physiology* (25) pp 819 - 840

Henle M (1984) ‘Isomorphism: Setting the record straight’, *Psychological Research* (46) pp317-327

Hering E (1920/1964) *Outlines of a theory of the light sense,* Hurvich LM and Jameson D (trans), Harvard University Press, Cambridge

Hesselgren S (1984) ‘Why colour order systems?’, *Colour Research & Application* (9) 4 pp220-228

Higashayama A and Shimono K (1994) ‘How accurate is size and distance perception for very far terrestrial objects?’, *Perception and Psychology* (55) 4 pp429-442

Hildreth EC and Koch C (1987) ‘The analysis of visual motion: From computational theory to neuronal mechanisms’, *Annual Review of Neuroscience* (10) pp477-533

Hill H and Bruce V (1993) ‘Independent effects of lighting, orientation, and stereopsis on the hollow-face illusion’, *Perception* (22) 8 pp887-897

Hinton GE (1979) ‘Imagery with arrays’, *The Behavioural and Brain Sciences,* (2) pp555-556

Hochberg JE (1957) ‘The effects of the Gestalt revolution’, in Beardslee DC and Wertheimer M (eds), *Readings in Perception,* Van Nostrand, Princeton

Hoffman WC (1970) ‘Higher visual perception as prolongation of the basic Lie transformation group’, *Mathematical Biosciences* (6) pp437-471

Hoffman WC and Dodwell PC (1985) ‘Geometric psychology generates the visual Gestalt’, *Canadian Journal of Psychology* (39) 4 pp491-528

Holmes OW (1859) ‘The stereoscope and the stereograph’ *Atlantic Monthly* (3) pp738-750

Holmes OW (1861) ‘Sun-painting and sun-sculpture’ *Atlantic Monthly* (8) pp13-29

Holmes OW (1863) ‘Doings of the sunbeam’ *Atlantic Monthly* (12) pp1-15

Hubel DH (1995) *Eye, Brain and Vision,* Scientific American, New York

Hubel DH and Wiesel TH (1970) 'Stereopsis vision in the macaque monkey', *Nature* (225) pp41-42

Hubel DH and Wiesel TN (1979) 'Brain mechanisms of vision', in Morrison P (ed), *The Brain,* WH Freeman & Co, San Francisco, pp 84 - 96

Humphrey NK (1972) 'Interest and pleasure: Two determinants of a monkey's visual preferences', *Perception* (1) 385-416

Hurvich LM (1969) 'Hering and the scientific establishment', *American Psychologist* (24) pp497-514

Johansson G (1973) 'Monocular movement parallax and near-space perception', *Perception* (2) 2 pp135-146

Johansson G (1973a) 'Visual perception of biological motion and a model for its analysis', *Perception and Psychophysics* (14) pp201-211

Johansson G (1975) 'Visual motion perception', *Scientific American* (232) pp76-88

Johansson G (1994) 'Personal comments', in Jansson G, Bergstrom SS and Epstein W (eds) *Perceiving Events and Objects: Resources for Ecological Psychology,* Lawrence Erlbaum Associates, Hillsdale, pp495-503

Johansson G and Börjesson E (1989) 'Toward a new theory of vision: Studies in wide-angle space perception', *Ecological Psychology* (1) pp301-331 reprinted in Jansson G, Bergstrom SS and Epstein W (eds) *Perceiving Events and Objects: Resources for Ecological Psychology,* Lawrence Erlbaum Associates, Hillsdale, pp239-269

Johnson EG and Tomiie T (1985) 'The development of colour-naming in four- to seven-year old children: A cross-cultural study', *Psychologia: An International Journal of Psychology in the Orient* (28) 4 pp216-227

Julesz B (1960) 'Binocular depth perception of computer generated patterns' Bell Systems Journal (39) pp1125-1162

Julesz B (1971) *Foundations of Cyclopean Perception,* University of Chicago Press, Chicago

Juola JF and Breitmeyer BG (1989) 'A discussion of models of motion perception', in Elsendoorn BAG and Bouma H (eds), *Working Models of Human Perception,* Academic Press, London pp251-259

Kaas JH (1992) 'Do humans see what monkeys see?', *Trends in Neuroscience* (15) 1 pp1-3

Kanizsa G and Gerbino W (1976) 'Convexity and symmetry in figure-ground organisation', in Henle M, *Vision and Artifact,* Springer, New York

Kaplan GA (1969) 'Kinetic disruption of optical texture: The perception of depth at an edge', *Perception and Psychophysics* (6) pp193-198

Kaplan E, Lee BB, and Shapley RM (1990) ‘New views of the primate retinal function’, in Osborne NN and Chader GJ (eds), *Progress in Retinal Research 9*, Pergamon, Oxford pp273-336

Katzman N and Nyenhuis J (1972) ‘Colour vs black-and-white effects on learning, opinion and attention’, *AV Communication Review,* (20) 1 pp16-28

Kaufmann R and O’Neill MC (1993) ‘Colour names and focal colours on electronic displays’, *Ergonomics* (36) 8 pp881-890

Kaufmann F (1995) ‘Development of motion perception in early infancy’, *European Journal of Pediatrics* (154) 9 pp548-553

Kellman PJ (1995) ‘Ontogenesis of space and motion perception’, in Epstein W and Rogers S (eds), *Perception of Space and Motion* (2nd Edition), Academic, San Diego pp327-364

Kellman PJ and Spelke ES (1983) ‘Perception of partly occluded objects in infancy’, *Cognitive Psychology* (15) 4 pp483-524

Kennedy JM and Domander R (1985) ‘Shape and contour: the points of maximum change are least useful for recognition’, *Perception* (14) pp367-370

Knoll NE (1977) ‘Effects of irrelevant colour changes on speed of name decisions’, *Quarterly Journal of Experimental Psychology,* (29) 2 pp277-281

Koenderink JJ and van Doorn AJ (1976) ‘Local structure of movement parallax of the plane’, *Journal of the Optical Society of America* (66) pp717-723

Kozlowski LT and Cutting JE (1977) ‘Recognising the sex of a walker from a dynamic point light display’, *Perception and Psychophysics* (21) 6 pp575-580

Krumhansl CL (1984) ‘Independent processing of visual form and motion’, *Perception* (12) 4 pp417-423

Kushner MJ, Rosenquist A, Alavi A, Rosen M, et al (1988) ‘Cerebral metabolism and patterned visual stimulation: A positron emission tomographic study of the human visual cortex’, *Neurology* (38) 1 pp89-95

Lansdown J (1987) ‘Computer graphics in design: Parametric variation as a design method’, in Rogers DF and Earnshaw RA (eds) *Techniques for Computer Graphics,* Springer-Verlag, New York pp7-14

Lansdown (1988) *The Opponent Process Theory of Colour: Its Possibilities as a Computer Graphics Colour Model,* CEA Working Paper, Middlesex University

Lansdown J (1994) ‘Visualising design ideas’, in MacDonald L and Vince J (eds) *Interacting with Virtual Environments,* Wiley, New York, pp61-77

Leake I (1993) *Apprehending Movement of the Human Figure,* Unpublished PhD Thesis, University of Brighton

Leake I (1995) ‘A morphology of visual representation of human movement’, in Robertson A, *Proceedings of the 4D Dynamics*

Conference, De Montfort University, Leicester pp101-106

Lennie P (1984) ‘Recent trends in the physiology of colour vision’, *Trends in Neoroscience* (6) pp243-248

Lind C (1993) ‘Psychology of colour: Similarities between abstract and clothing colour preferences’, *Clothing and Textiles Research Journal* (12) 1 pp57-65

Livingstone MS (1988) ‘Art, illusion and the visual system’, *Scientific American,* (258) 1 pp68-75

Loftus GR and Mackworth NH (1978) ‘Cognitive determinants of fixed location during picture viewing’, *Journal of Experimental Psychology: Human Perception and Performance,* (4) pp 565 - 572

Loomis JM, Da Silva JA, Fujita N and Fukusima SS (1992) ‘Visual space perception and visually directed action’, *Journal of Experimental Psychology: Human Perception and Performance* (18) 4 pp906-921

Looren de Jong H (1995) ‘Ecological psychology and naturalism: Heider, Gibson and Marr’, *Theory and Psychology* (5) 2 pp251-269

Lueck CJ, Zeki S, Friston KJ, Deiber M-P et al (1989) ‘The colour centre in the cerebral cortex’, *Nature* (340) 6232 pp386-389

Mackworth NH and Morandi AJ (1967) ‘The gaze selects informative details within pictures’, *Perception and Psychophysics,* (2) pp 547 - 551

McManus IC, Jones AL and Cottrell J (1981) ‘The aesthetics of colour’, *Perception* (10) 6 pp651-666

MacNichol EF (1964) ‘Three-pigment colour vision’, *Scientific American Reprint 97*

Marr D (1982) *Vision: A Computational Investigation into Human Representation and Processing of Visual Information,* WH Freeman, San Francisco

Martindale and Moore (1988) ‘Priming, protypicality and preference’, *Journal of Experimental Psychology: Human Perception and Performance* (14) pp661-670

Masin SC (ed) (1993) *Foundations of Perceptual Theory,* North-Holland, Amsterdam

Mason L and Kandel ER (1991) ‘Central Visual Pathways’ in Kandel ER, Schwartz JH and Jessell TM *Principles of Neural Science (3rd ed)* Elsevier, New York pp420-437

Massironi M and Savardi U (1991) ‘The ceiling of the Church of St Ignatius and the perception of concave surfaces’, *Perception* (20) 6 pp771-787

Mather G and West S (1993) ‘Evidence for second order motion detectors’, *Vision Research* (33) 8 pp1109-1112

Mitchell WJ (1986) ‘Colour transformations in computer-aided design’, *Design Computing* (1) 1 pp20-28

Mitchison GJ and Westheimer G (1990) ‘Viewing geometry and gradients of horizontal disparity’, in Blakemore C (ed) (1990) *Vision: Coding and Efficiency,* CUP, Cambridge pp302-309

Morgan IG (1991) 'What do amacrine cells do?', in Osborne NN and Chader GJ (eds), *Progress in Retinal Research 11,* Pergamon, Oxford pp193-214

Munsell AH (1916) *A Colour Notation*, Ellis, Boston (1st ed)

Murch GM (1985) 'Colour graphics - blessing or ballyhoo?' *Computer Graphics Forum,* (4) pp127-135

Naiman A (1985) 'Colour spaces and colour contrast', *Visual Computer* (1) 3 pp194-201

Naka KI and Rushton WAH (1966) 'S-potentials from colour units in the retina of fish (cyprinidae)', *Journal of Physiology,* (185)

Nakayama K (1985) 'Biological image motion processing: A review', *Vision Research* (252) 5 pp625-660

Neisser U and Beklen R (1974) 'Selective looking: Attending to visually specified events', *Cognitive Psychology* (6) pp84-107

Neitz J, Neitz M and Jacobs GH (1993) 'More than three different cone pigments among people with normal colour vision', *Vision Research* (33) 1 pp117-122

Nishita T, Sirai T, Tadamura K and Nakamae E (1993) 'Display of the Earth taking into account atmospheric scattering', *Computer Graphics* (28) 3 pp175-182

Ogle KN (1962) 'Perception of distance and of size', in Davson H (ed) *The Eye (1st edition)Vol 4,* Academic Press, New York pp247-269

Palmer SE (1992) 'Modern theories of Gestalt perception', in Humphreys GW (ed), *Understanding Vision,* Blackwell, Oxford pp39-70

Parks TE and Rock I (1990) 'Illusory contours from pictorially 3-dimensional inducing elements', *Perception* (19) pp119-121

Parslow R (1988) 'Just an illusion', *Computer Images,* Mar-April, pp12-17

Pastore N (1971) *Selective History of Theories of Visual Perception: 1650-1950,* OUP, New York

Pavlova MA (1992) 'Preschoolers can recognise biological motion patterns', *Perception: Supplement 2 on European Conference on Visual Perception* (21) p20 (Abstract only)

Pettigrew (1990) 'Is there a single, most-efficient algorithm for stereopsis?', in Blakemore C (ed) (1990) *Vision: Coding and Efficiency,* CUP, Cambridge pp283-290

Piccolino M (1995) 'The feedback synapse from horizontal cells to cone receptors in the vertebrate retina', *Progress in Retinal and Eye Research 14,* Pergamon, Oxford pp141-196

Pickford RW (1972) *Psychology and Visual Aesthetics,* Hutchinson, London

Pirenne MH (1970) *Optics, Painting and Photography,* CUP, Cambridge

Polyak SL (1941) *The Retina,* University of Chicago Press, Chicago

Polzella DJ and Montgomery DA (1993) ‘Dimensions of colour harmony’, *Bulletin of the Psychonomic Society* (31) 5 pp423-435

Pomerantz JR (1983) ‘Global or local precedence: Selective attention in form and motion perception’, *Journal of Experimental Psychology: General,* (112) 4 pp516-540

Pomerantz JR (1985) ‘Perceptual organisation in information processing’, in Aitkenhead AM and Slack JM (eds), Issues in Cognitive Modelling, Lawrence Erlbaum, Hillsdale pp127-158

Pomerantz JR (1986) ‘Visual form perception: An overview’, in Schwab EC and Nusbaum HC (eds) *Pattern Recognition by Humans and Machines, Volume 2. Visual Perception,* Academic, New York

Pridmore RW (1992) ‘Music and colour: Relations in the psychophysical perspective’, *Colour Research and Applications* (17) 1 pp57-81

Pritchard RM (1964/1972) ‘Stabilised images on the retina’, in Held R and Richards W, *Perception: Mechanisms and Models, Readings from Scientific American,* WH Freeman and Co, San Francisco, 1972 pp 176 - 182

Purdy DM (1931) ‘Spectral hue as a function of intensity’, *American Journal of Psychology* (43) pp541-559

Purghé F (1993) ‘Illusory contours from pictorially 3-dimensional inducing elements’, *Perception* (22) 7 pp809-919

Raichle M (1990) ‘Images of the functioning human brain’, in Barlow H, Blakenore C and Weston-Smith M (eds), *Images and Understanding,* CUP, Cambridge pp284-296

Ramachandran VS (1990) ‘Interactions between motion, depth, colour and form: the utilitarian theory of perception’, in Blakemore C (ed) (1990) *Vision: Coding and Efficiency,* CUP, Cambridge pp346-360

Ratliff F (1965) *Mach Bands: Quantitatlve Studies on Neural Networks in the Retina,* Holden-Day, San Francisco

Ratliff F (1972/1976) ‘Contour and contrast’ in Held R and Richards W (eds), *Recent Progress in Perception,* WH Freeman, San Francisco

Reed ES (1987) ‘James Gibson’s ecological approach to cognition’, in Costall A and Still A (eds) *Cognitive Psychology in Question,* Harvester, Brighton

Regan EC (1995) ‘Some human factors issues in immersive virtual reality: fact and speculation’, in Earnshaw RA, Vince JA and Jones H, *Virtual Reality Applications,* Academic Press, London pp163-182

Reimann D and Haken H (1994) ‘Stereo vision by self-organization’, *Biological Cybernetics* (71) 1 pp17-26

Richter IA (1952) *Selections from the Notebooks of Leonardo da Vinci,* OUP, Oxford

Richter JP (1883/1970) *The Literary Works of Leonardo da Vinci (Vol 1),* Phaidon, London

Rock I (1983) *The Logic of Perception*, MIT Press, Cambridge

Rogers S (1985) *Representations and Reality*, Unpublished PhD Thesis, Royal College of Art, London

Rossotti H (1983) *Colour: Why the World isn't Grey*, Penguin Books Ltd, Harmonsworth

Runeson S and Frykholm G (1983) 'Kinematic specifications of dynamics as an informational basis for person-and-action perception: Expectation, gender-recognation, and deceptive intention', *Journal of Experimental Psychology: General* (112) p585-615

Saito M (1996) "Comparative studies on colour preference in Japan and other Asian regions, with special emphasis on the preference for white', *Colour Research and Application* (21) 1 pp35-49

Scanlon BA (1985) 'Race differences in selection of cheese colour', *Perceptual and Motor Skills* (61) p314

Schacter BJ (ed) (1983) *Computer Image Generation*, Wiley, New York

Schenk T and Zihl J (1995) 'Motion perception after brain damage', *Perception* (24) Supplement p54 (Abstract only)

Schontz WD, Trumm GA and Williams LG (1971) 'Colour coding for information location', *Human Factors,* (13) 3 pp237-246

Schwarz MW, Beatty JC, Cowan WB and Gentleman JF (1984) 'Towards an effective user interface for interactive colour manipulation', *Graphics Interface 84,* National Research Council of Canada, Toronto pp 187-196

Schwarz MW, Cowan WB and Beatty JC (1987) 'An experimental comparison of RGB, YIQ, LAB, HSV, and opponent colour models', *ACM Transactions on Graphics* (6) 2 pp123-158

Sebba R (1991) 'Structural correspondence between music and colour', *Colour Research and Applications* (16) 2 pp81-88

Sedgewick HA (1973) *The Visible Horizon,* Unpublished PhD Thesis, Cornell University

Sedgewick HA (1980) 'The geometry of spatial layout', in Hagen MA, *The Perception of Pictures* Volume 1, Academic, New York

Sekuler R (1975) 'Visual motion perception', in Carterette C and Freeman MP (eds), *Handbook of Perception (Vol 5),* Academic, New York pp385-430

Sharp PS and Philips R (1993) 'Physiological optics', in Hendee WR and Wells PNT (eds) *The Perception of Visual Information,* Springer, New York pp1-29

Shipp S, Watson JDG, Frackowiak RSJ and Zeki S (1995) 'Retinotopic maps in human prestriate visual cortex: The demarcation of areas V2 and V3, *Neuroimage* (2) pp125-132

Seifert LS and Johnson NF (1994) 'On the naming of colour words and colour patches', *Memory and Cognition* (22) 2 pp169-180

Sekular R (1975) ‘Visual motion perception’, in Carterette C and Freeman MP (eds), *Handbook of Perception (Vol 5),* Academic Press, New York pp385-430

Silver N (1988) ‘Sex and racial differences in colour and number preferences’, *Perceptual and Motor Skills* (66) pp295-299

Smets G (1982) ‘A tool for measuring relative effects of hue, brightness and saturation of colour pleasantness’, *Perceptual and Motor Skills* (55) pp1159-1164

Smith AR (1978) ‘Colour gamut transform pairs’, *Computer Graphics* (12) 3 pp12-19

Smith AT and Snowden RJ (1994) ‘Motion detection: An overview’, in Smith AT and Snowden RJ (eds), *Visual Detection of Motion,* Academic, London pp3-13

Spillman L (1993) ‘The perception of movement and depth in moiré patterns’, *Perception* (22) 3 pp287-308

Spillman L and Dresp B (1985) ‘Phenomenon of illusory form’, *Perception* (24) 11 pp1333-1364

Spoehr KT and Lehmkuhle SW (1982) *Visual Information Processing,* WH Freeman, San Francisco

Stark L and Ellis SR (1981) ‘Scanpaths revisited: Cognitive models direct active looking’, in Fisher DF, Monty RA and Senders JW (eds)*Eye Movements: Cognition and Visual Perception,* Lawrence Erlbaum, Hillside pp193-226

Spoehr KT and Lehmkuhle SW (1982)*Visual Information Processing,* WH Freeman & Co, San Francisco,

Steinbuch K (1962) ‘Information processing in man’, *Proceedings IRE International Conference on Human Factors in Electronics,* Long Beach

Stewart D , Cudworth CJ and Lishman JR (1993) ‘Misperception of time-to-collision by drivers in pedestrian accidents’, *Perception* (22) 10 p1227-1244

Sumi S (1984) ‘Upside down presentation of the Johansson moving light spot pattern’, *Perception* (13) 3 pp283-286

Sutherland NS (1989) ‘Simplicity is not enough’, in Elsendoorn BAG and Bouma H (eds), *Working Models of Human Perception,* Academic Press, London pp382-390

Taylor JM, Murch GM and McManus PA (1987) *Tektronix HVC, A uniform perceptual colour system for display users,* Tektronix Laboratories, Beaverton

Theeuwes J (1995) ‘Abrupt luminance change pops out; abrupt colour change does not’, *Perception and Psychophysics* (57) 5 pp637-644

Todd JT and Mingola E (1983) ‘Perception of surface curvature and direction of illumination from patterns of shading’, *Journal of Experimental Psychology: Human Perception and Performance,* (9) pp583-595

Tsotsos JK (1982) Knowledge of the visual process: content, form and use, *Proceedings International Conference on Pattern Recognition,* IEEE (2), pp 654 - 699

Tsotstos JK (1984) ‘Motion: representation and perception’, *Computer Graphics* (18) 1 pp7-27

Trandis HC, Malpass RS and Davidson AR (1973) ‘Psychology and culture’, *Annual Review of Psychology,* (24) pp 355 -378

Ullman S (1987) ‘Visual motion perception’, in Adelman G (ed), *Encyclopedia of Neuroscience,* Birkhäuser, Boston pp1280-1281

Valberg A, Lange-Malecki B and Seim T (1991) ‘Colour changes as a function of luminance contrast’, *Perception* (20) 5 pp655-668

Von Waltenwyl A and Zollinger H (1979) ‘Colour-term salience and neurophysiology of colour vision’, *American Anthropologist* (81) pp276-288

Wagner M (1985) ‘The metric of visual space’, *Perception & Psychophysics* (38) 6 pp483-495

Wald G (1950/1972) ‘Eye and camera’, in Held R and Richards W (eds), *Perception: Mechanisms and Models, Readings from Scientific American,* WH Freeman and Co, San Francisco, 1972 pp 94 - 103

Warren WH (1995) ‘Self-motion: Visual perception and visual control’, in Epstein W and Rogers S (eds), *Perception of Space and Motion* (2nd Edition), Academic, San Diego pp263-325

Warren WH and Hannon DJ (1988) ‘Direction of self-motion is perceived from optical flow’, *Nature* (336) 6195 pp162-163

Weale RA (1982) *Focus on Vision,* Hodder and Stoughton, London

Wierzbicka A (1990) ‘The meaning of colour terms: semantics, culture, and cognition’, *Cognitive Linguistics* (1) 1 pp99-150

Willmer (1982) ‘Anatomy of the visual system’, in Barlow HB and Mollon JD (eds), *The Senses,* CUP, Cambridge

Winston PH (1973) ‘Learning to identify toy block structures’, in Solso RL (ed), *Contemporary Issues in Cognitive Psychology,* McGraw-Hill, New York

Wright P and Lickorish A (1988) ‘Colour cues as location aids in lengthy texts on screen and paper’, *Behaviour and Information Technology* (7) 1 pp11-30

Yarbus AL (1967) *Eye Movements and Vision,* Plenum Press, New York

Young JZ (1978) *Programs of the Brain,* OUP, Oxford

Zeki S (1974) ‘The mosaic organisation of the visual cortex in the monkey’, in Bellairs R and Gray EG (eds), *Essays on the Nervous System: a Festschrift for Professor JZ Young,* OUP, Oxford

Zeki S (1990) ‘The motion pathways of the visual cortex’, in Blakemore C (ed) (1990) *Vision: Coding and Efficiency,* CUP, Cambridge pp321-345

Zeki S and Lamb M (1994) ‘The neurology of kinetic art’, *Brain* (117) 3 pp607-636

Zimmer AC (1984) ‘There is more than one level in color naming: A reply to Zollinger (1984)’, *Psychological Research* (46) 4 pp411-416

Zollinger H (1984) ‘Why just turquoise? Remarks on the evolution of color terms.’, *Psychological Research* (46) 4 pp403-409

Zucker SW and Davis S (1988) ‘Points and endpoints: A size - spacing constraint for dot grouping’, *Perception* (17) 2 pp229-247

Zusne L (1970) *Visual Perception of Form,* Academic Press, New York

8. Glossary index

Note that this glossary index attempts to give only informal definitions of some of the terms used in the Report.

Accommodation Adjustment of the optics of the eye to allow it to keep an object in focus on the retina as its distance from the eye varies. In the human eye this is achieved by changing the shape of the lens (p2- 6).

Adaptation A change in the sensitivity to light of a photoreceptor or the visual system as a whole allowing it to take account of the average light intensity. Adaptation to bright light is quicker than adaptation to low light.

Amacrine cells Cells within the retina that connect to bipolar cells, ganglion cells and to other amacrines. Some amacrine cells respond to sustained light and others only to transient light. About 25 different types of amacrines have been identified (p1- 9)

Axon Long nerve fibre that usually carries impulses away from the nerve cell (p1- 12).

Bezold-Brücke effect Other than at some fixed points in the spectrum, as intensity of light increases, colours shift slightly towards yellow if they are at the red end of the spectrum or towards blue if they in the green area. The effect is named after Wilhelm von Bezold (1837-1907), a German meteorologist, and Ernst Wilhelm von Brücke (1819-1892), an Austrian physiologist. They did not work together but each published on the phenomenon, Bezold in 1873 and Brücke in 1878 (p4- 1).

Bipolar cells Cells in the retina that transmit impulses from the photoreceptors to the ganglion cells. They are of two sorts: on and off (p1- 8).

Brodmann's areas Areas of the brain mapped by the German neuroanatomist, Korbinian Brodmann (1868-1918) at the beginning of the twentieth century and published in 1909. The areas are distinguished by their differences in thickness, density and cell size of the cortical layers as well as by functional specialisations. Although detail changes have been made to Brodmann's map and its about 50 areas, the broad arrangement is accepted even today and is used to index brain locations. The areas are numbered simply in the order that Brodmann studied them (p1- 16).

Camera obscura A device used by renaissance and later artists to assist in drawing perspective by projecting an image of a scene onto a surface where it could be traced. The invention of the device is sometimes credited to the architect, Leone Battista Alberti (1404 - 1472) and the name comes from the Latin for 'dark chamber'. Portable versions of camera obscura resembled photographic plate-cameras with a ground glass screen: others were boxes large enough for the artist to sit in. The name is also given to 18th and 19th century rooms into which an exterior panoramic scene is projected by means of a periscope-like apparatus for entertainment purposes. Versions of these fascinating devices are still in use in Edinburgh and Clifton, near Bristol. Research by Steadman suggests that Jan Veermer (1632 - 1675) used a camera obscura extensively in his paintings of interior scenes. In the 19th century, the camera obscura was largely replaced by the camera lucida (Latin: 'light chamber') which allowed controlled enlargement of the projection (p1- 1).

Cerebral cortex Thin layer of nerve and other cells over much of the outer surface of the brain.

Choroid membrane A five-layered middle coat of the eye lying between the *retina* and the *sclera* (p1- 3).

Cones Photoreceptors in the retina that mediate colour vision. They are so named because of their roughly conical shape. Cones are made up of four parts: an outer segment which is furthest from the light, an inner segment, a cell body, and a synaptic terminal which connects it to other cells. The outer segment is the light-sensitive part and contains photosensitive visual pigment of one of three sorts whose maximum absorbances are 558, 531 and 419 nm. As little as one photon will cause a response in the cone of about 10 μV but this is about 100 times less than given off by rods in the same circumstances. Cones interact with neighbouring cones and with neurons using *opponent-colour response* to produce signals for passing on to the *visual cortex* (pp1- 4 to 1- 10, 4- 2 to 4- 9).

Corpus callosum A bundle of fibres that connects the right and left hemispheres of the brain (p1- 17).

Cornea The transparent outer surface of the *sclera* that lies in front of the *iris* and lens.

Dendrites Short branching processes emerging from nerve cells (pp1- 9, 1- 12).

Dichromatic vision Vision with only two types of cone. This can arise in human individuals via ‘colour blindness’ or in whole species that only have two types of cone. In such species there will be one wavelength that appears without colour and hue will not vary continuously across the spectrum.

Ecological approach to perception The view of perception held by JJ Gibson and his followers that our perceptual mechanisms are tuned to allow us to get along in the world and that we ‘pick up’ on elements of the environment to allow us to understand it in a direct way (p2- 9)

Fixation The periods of time between saccades when our eye stops moving in order to allow us to concentrate on a detail of a scene (pp 1- 20 to 1- 22, 5- 19 to 5- 24).

Floaters Debris (such as dead cells) within the fluid of the eye that appear to us as dark spots (p1- 3).

Fovea A small (0.3 mm across) depression in the centre of the retina where we can see most clearly and where many of the cones of the eye are sited. There are no rods in this area but the density of cones is about 150,000 per sq mm. We use *saccades* to focus details of a scene onto the fovea (p1- 4).

Ganglion cells An aggregation of nerve cells situated outside the central nervous system. Ganglion cells make up the innermost layer of the retina and are the only cells with axons that leave the eye (pp1- 5 to 1- 17).

Gnostic cells Cells that specialise in the recognition of features of the environment at a level above that of a line or dot (p1- 20).

Hue The dominant wavelength of the light seen directly from a source, or indirectly from reflections off surfaces. It is the feature we use to give a colour its name — red, green, yellow and so on. We are able to distinguish differences in about 150-200 hues (pp4- 1 to 4- 2).

Iris Circular structure that forms the coloured part of the eye and acts as a diaphragm to control the size of the pupil opening.

Lateral inhibition The feature in which adjacent and neighbouring cells mutually inhibit one another's actions (p1- 8).

Lateral geniculate nucleus (LGN) The part of the brain where the axons of retinal *ganglion cells* terminate and from which *axons* run to the *visual cortex*. There are two LGN, one on each side of the brain and each comprises six known layers, four containing *parvocellular* material and two containing *magnocellular* material (pp1- 15 to 1- 19).

Lightness The lightness of a colour is its degree of perceived *luminance* relative to the luminance of another colour or the surroundings. We usually speak of the lightness of a surface colour but of the intensity or brightness of lights and CRT phosphors. Although, technically, these are different concepts, in daily life and in computing we tend to use the words synonymously. We are able to distinguish differences in about 500 steps of lightness (p4- 1).

Luminance The amount of light intensity per unit area reaching the eye from a surface (p).

Mach banding An anomaly of perception in which, when viewing areas in which there are abrupt changes of luminance, we see additional lines of different brightnesses. The misperception probably arises because of *lateral inhibition* and was first scientifically studied by Ernst Mach (1828 - 1916) although the phenomenon was known to renaissance artists (pp1- 13 to 1- 15).

Melatonin Discovered in 1958, melatonin is a hormone produced by the pineal gland which has an effect of circadian rhythms. There is a ten-fold increase in the production of melatonin at night as opposed to the day.

Monocular vision Vision involving just one eye (pp2- 1 to 2- 11).

Motion parallax Movement of the image of an object over the retina caused by the movement of the body or head of an observer or by the motion of the object relative to the observer. The rate of movement depends on the velocity of the object relative to the eye and its distance from the eye (pp2- 11 to 2- 12, 2- 22 to 2- 23).

Neurons Specialised cells (also known as neurones) that constitute the building blocks of the nervous system. They comprise a cell body (soma) with multiple *dendrites* that receive information from other neurons and a single *axon* that transmits information via a *synapse* to other neurons. An average neuron has from 1000 to 10000 synaptic contacts and receives input from about 1000 other neurons. There are at least 10^{11} neurons in the human brain. About 60 different types of retinal neurons have been identified (pp1- 12, 1-19 to 1- 20).

Occlusion The partial or complete hiding of one object behind another (pp2- 13, 2- 22).

Optic array A term due to JJ Gibson. It is the instantaneous pattern of light reaching a point from all directions. In different regions of the optic array the spatial pattern of light will differ according to the nature of the surface from which it has been reflected and this property can be used to assist in perception (p3- 8).

Optic disc The point at which the optic nerve joins the retina and where no photoreceptors occur giving rise to a blind cpot.

Optic flow The fluctuating pattern of light intensity reaching an observer caused by any relative movement between the observer and the environment (p3- 8).

Opponent colour response A cell is said to have opponent colour response when light of one wavelength falling on the receptive field causes it to fire more frequently than its resting state, and light of a different wavelength causes it to fire less frequently than its resting state (p4- 4 to 4- 8, 4- 12 to 4- 13, 4- 17).

PET scanner A large electromechanical device comprising a tube surrounded by equipment in which a patient lies to enable the measurement of the blood flow of the brain. This is done by means of detection of a radioactive tracer. The measurement takes place in notional slices through the brain and a computer converts the results into coloured areas in the form of maps with areas of brain activity shown up by increased blood flow. PET stands for positron emission tomography (p1- 21).

Pineal gland A tiny element at the centre of the brain and shaped like a pine-cone, hence its name. The French philosopher, René Descartes (1596 - 1650) believed it to be the 'seat of the mind' but its actual function seems to be associated with the biological clock. The gland secretes *melatonin.* Apparently the gland is larger in children than adults. (p2- 6)

Plexiform layers Layers in the retina in which virtually all the synapses occur. There are two plexiform layers, the inner and the outer (p1- 7).

Psychological primaries The four colours of light, red, yellow, green and blue, that are perceived to be 'primary' and not mixtures of other colours. Leonardo was one of the first to talk about these as primary colours (p4- 1).

Rods Photoreceptors in the retina that specialise in grey scale and low light activity. They are so named because of their shape. They are distributed all over the retina except at the fovea and there are about 160 million of them in each eye (pp1- 4 to 1- 10, 4- 2 to 4- 9).

Retina The photosensitive layer on the inside of the eye, approximately 270 sq mm in area. Contains the photoreceptors (rods and cones) as well as five other major classes of retinal neurons: horizontal, bipolar, amacrine, interplexiform and, ganglion (pp1- 2 to 1- 13, 1- 16 to 1- 20).

Saccades Abrupt and rapid movements of the eyes as they move from one fixation to another (pp 1- 20 to 1- 22, 5- 19 to 5- 24).

Saturation The apparent purity of a *hue*. In the case of a surface colour, this is the degree to which the colour is undiluted by white. In the case of lights or phosphors, saturation depends on the relative amounts of luminous intensity held by the various wavelengths that make up the colour. In both cases, the more one wavelength dominates, the greater is the saturation: black, grey and white, where no wavelength dominates, have the same saturation — they differ only in lightness. Because they emit light of only one wavelength, lasers are normally the only devices that provide fully saturated light (pp4- 1, 4- 5 to 4- 10, 4- 14 to 4-17).

Sclera The white, opaque, fibrous outer coating of the eye covering it everywhere except at the *cornea* (p).

Stereopsis Perception of depth dependent on the disparity in the images projected onto the retinas of the two eyes (p2- 3 to 2- 10, 2- 22, 2- 25).

Stereoscope An optical device for viewing pairs of photographs taken from viewpoints a few centimetres apart in order to see their images in three-dimensions. The original desk-top stereoscope, a scientific instrument, was invented in 1838 by Charles Wheatstone (1892-1875). This was a piece of laboratory equipment. In about 1860, the American author and physician, Oliver Wendell Holmes (1809 -1894) invented the hand-held version of the device which proved so extremely popular with Victorians (p4- 4 to 4- 5).

Superior colliculi Two areas in the central region of the brain. These are layered structures and seem to be exclusively concerned with vision. Each superior colliculus receives input from the retina and from the visual cortex. Damage to the superior colliculus in the right or left hemispheres of humans affects the action of the saccades (p1- 15).

Synapse Specialised zone of contact between nerve cells or between a nerve cell and a muscle cell. Synapses were first described in detail by Ramon y Cajal. The word comes from the Greek, synapsis = to clasp, to indicate a junction and was first coined by the British neurophysiologist, Charles Scott Sherrington (1857 - 1952), in 1879 (pp1- 11 to 1- 12).

Tapetum A dark, shiny layer in the retina of many nocturnal animals that reflects back some of the light into the eye, presumably to help its light-gathering properties. The effect causes the eyes of these animals to shine in low-light conditions and may have contributed to the Pythagoreans belief that light proceeded from the eye to illuminate objects (p1- 3).

Vergence Eye movements in opposite horizontal directions that allow the eyes to focus on a target (p2- 6).

Visual cortex That part of the rear brain that is specialised in vision. It has a layered structure and is made up of sheets of neurons (pp1- 8, 1- 16 to 1- 19).

A1. Appendix 1

Using the RGB model

The following technique is useful in specifying colours via the RGB model.

Assume that we have three vertical bars representing R, G and B slider that can vary in length from 0.0 to 1.0 (Figure A1. 1)

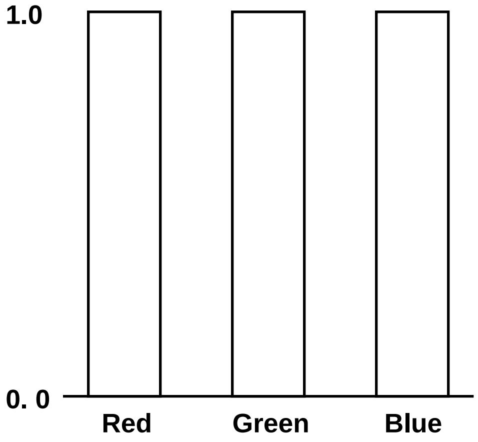


Fig A1. 1 The notional RGB sliders

Case 1: Only one bar of length > 0

1. When a single bar is of length 1, the colour red, green or blue that it represents is full on.
2. When a single bar is less than 1, the length represents the percentage of lightness of that primary — although this will not normally be the same as our perception of lightness (that is, for example, 0.8 will not appear to be twice as bright as 0.4 unless the necessary gamma correction is used).

Case 2: Two bars are of length > 0

1. If both bars are of length 1, the resulting colour will be full brightness cyan, magenta or yellow depending on which sliders are set.
2. If both bars are of equal length but < 1 the cyan, magenta or yellow will be of reduced brightness (again non-linearly).
3. If one bar is longer than another, draw a line across both bars at

the height of the minimum bar (Figure A1. 2). Below the line we then have two bars of equal height giving us an example of Case 2.2 (that is reduced brightness cyan, magenta or yellow) together with the dominant colour above the line. Thus in Figure A1. 2 representing (0.9, 0.5, 0), we have and equal mixture of red and green at 0.5 giving darkish yellow plus extra red. This is a quite bright reddish-yellow or orange.

Note that many of the colours that we wish to mix will be covered by this two-primary case. One of the major errors made by those learning to use RGB is that they assume every non-primary colour is made up of some percentage of each of R, G and B. This is not so. Mixing all three primaries does no more than add grey to the colour — hence reducing its saturation.

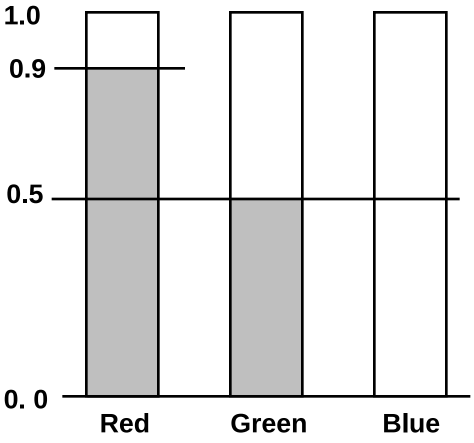


Fig A1. 2 Notional RGB sliders at (0.9, 0.5, 0)

Case 3: Three bars are of length > 0

1. If all bars are at full height, we have white.
2. If all bars are equal but < 1, we have a grey whose brightness depends on the height (but not always linearly).
3. If at least one bar is longer than the others, draw a line across all three bars at the level of the minimum bar (Figure A1. 3). Below this line we have an example of Case 3. 2 — a grey. Next draw a line across the remaining two bars at the height of the next highest bar. Here we have an example of Case 2. 2 — an equal two-primary mixture which gets added to the grey.

Finally, we have the top-most piece of the remaining bar which gives the dominant colour. The colour represented by (0.9, 0.5, 0.4) as in Figure A1. 3 then, will be grey at 0.4, yellow up to 0.5, and red above — that is, a quite grey orange or yellowish pink.

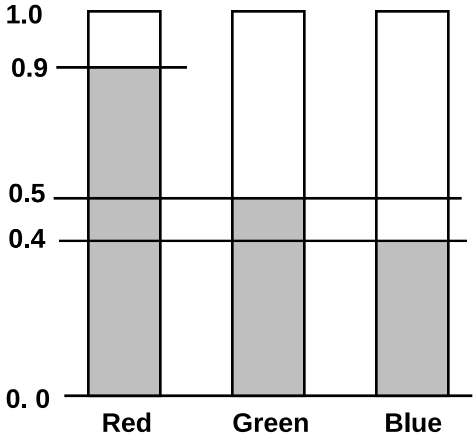


Fig A1. 3 Notional RGB sliders at (0.9, 0.5, 0.4)

Summary

Setting colours by means of the RGB model is simple if it is remembered that whenever all three notional RGB sliders are set then a grey exists at the level of the lowest slider (the lower, the darker). More saturated colours can be set using only one or two sliders.

Plate 2.1 Koninck: Landscape with a Road by a River 1655 (National Gallery, London)

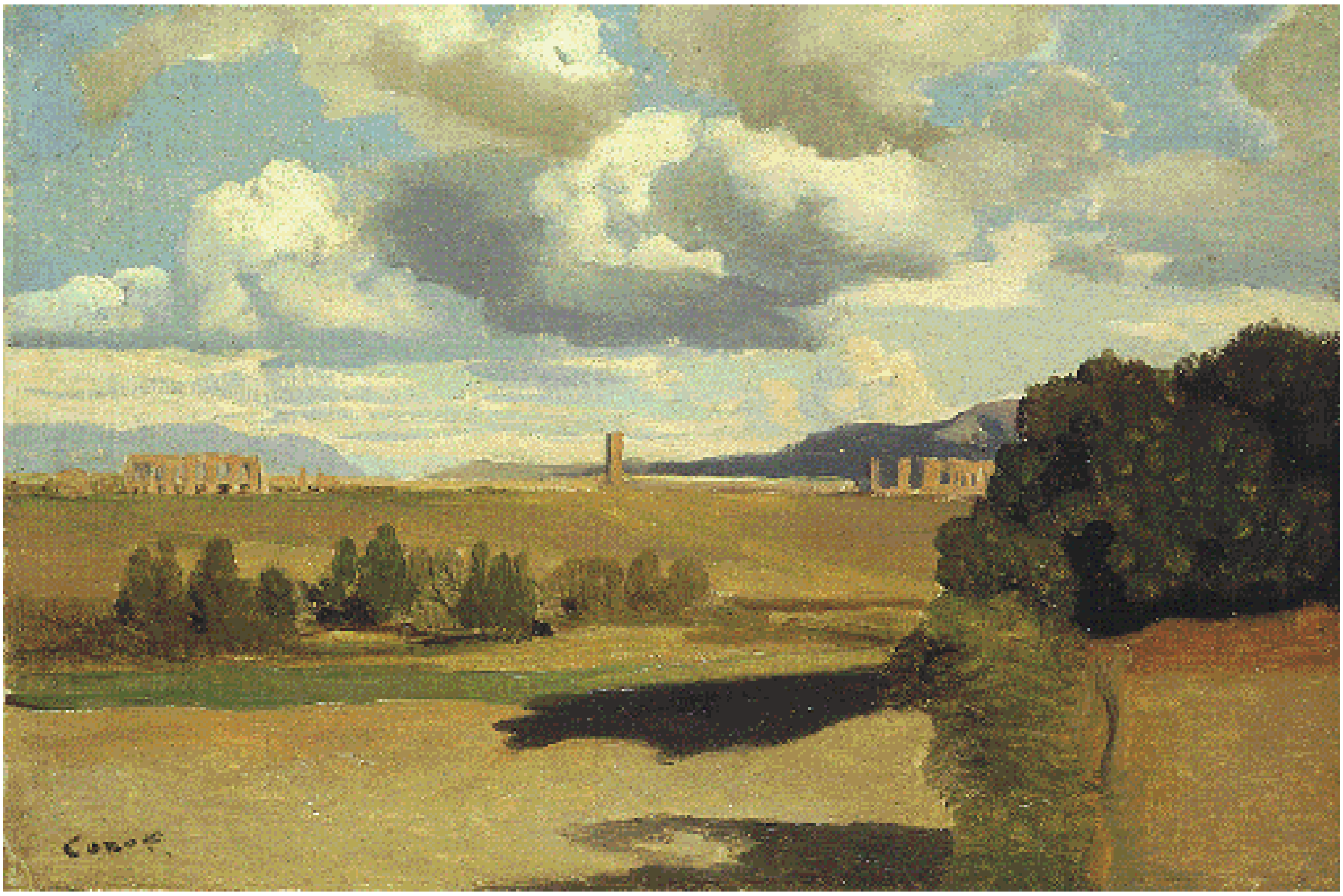


Plate 2.2 Corot: The Roman Campagna 1826-28 (National Gallery, London)

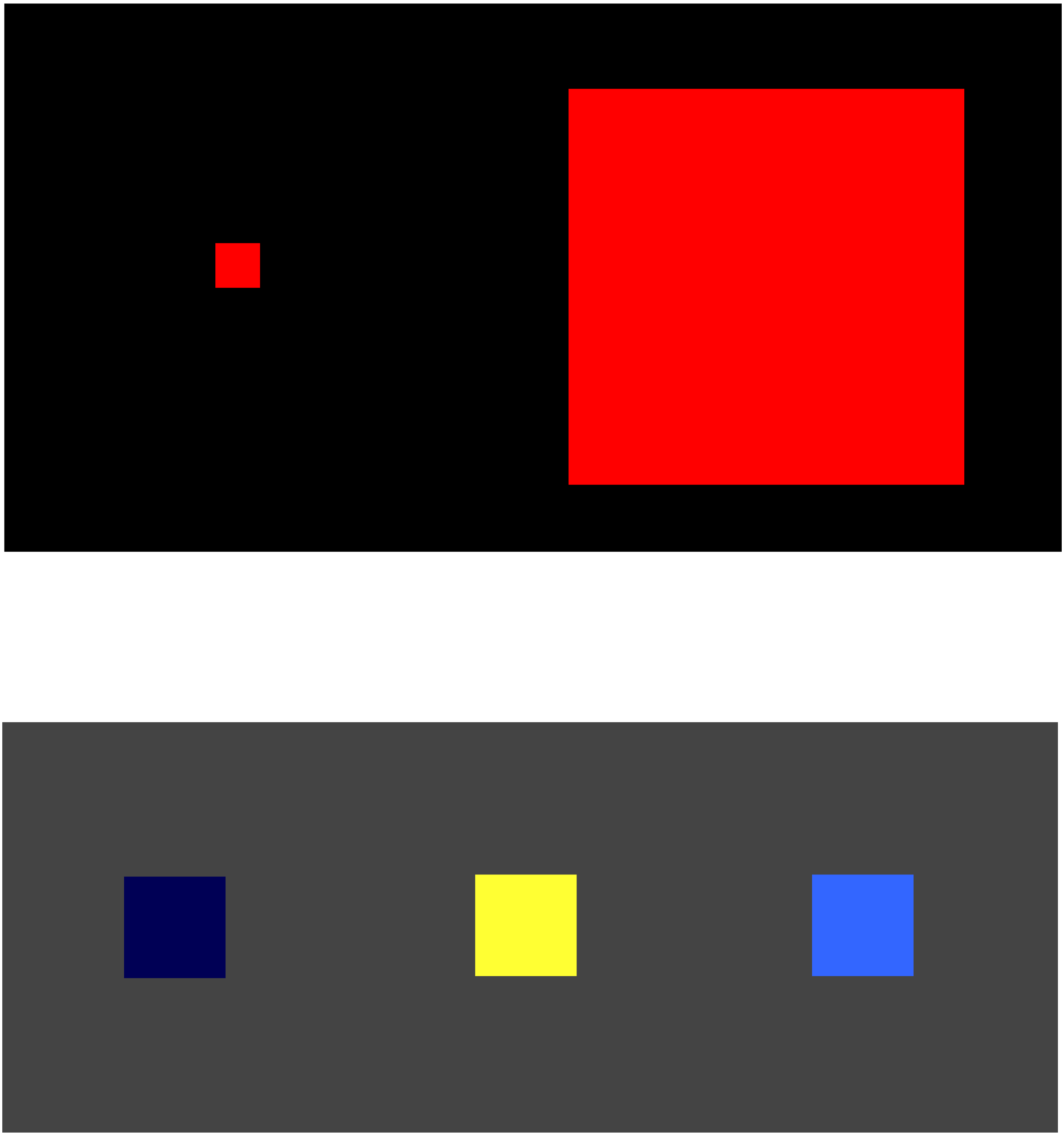


Plate 4.1 (Upper) A small patch of colour seems more saturated and darker than a larger patch of the same colour when seen against a black background. (Lower) Equal-size squares of different colours seem to be of different sizes

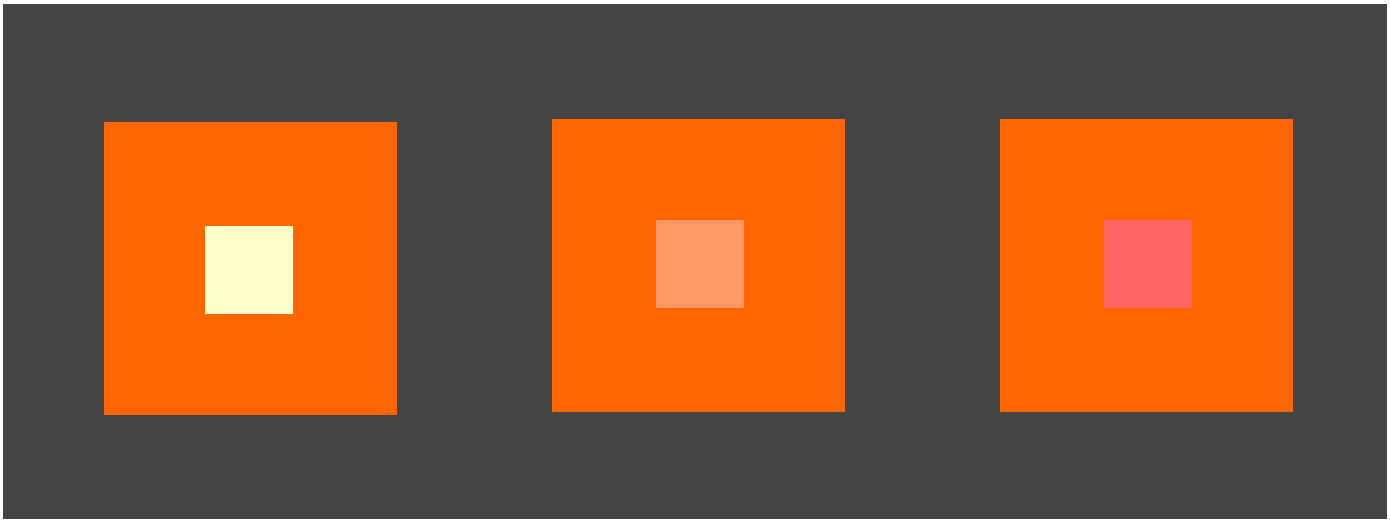


Plate 4.2 The edges of the smaller squares are equally sharp but do not appear to be so