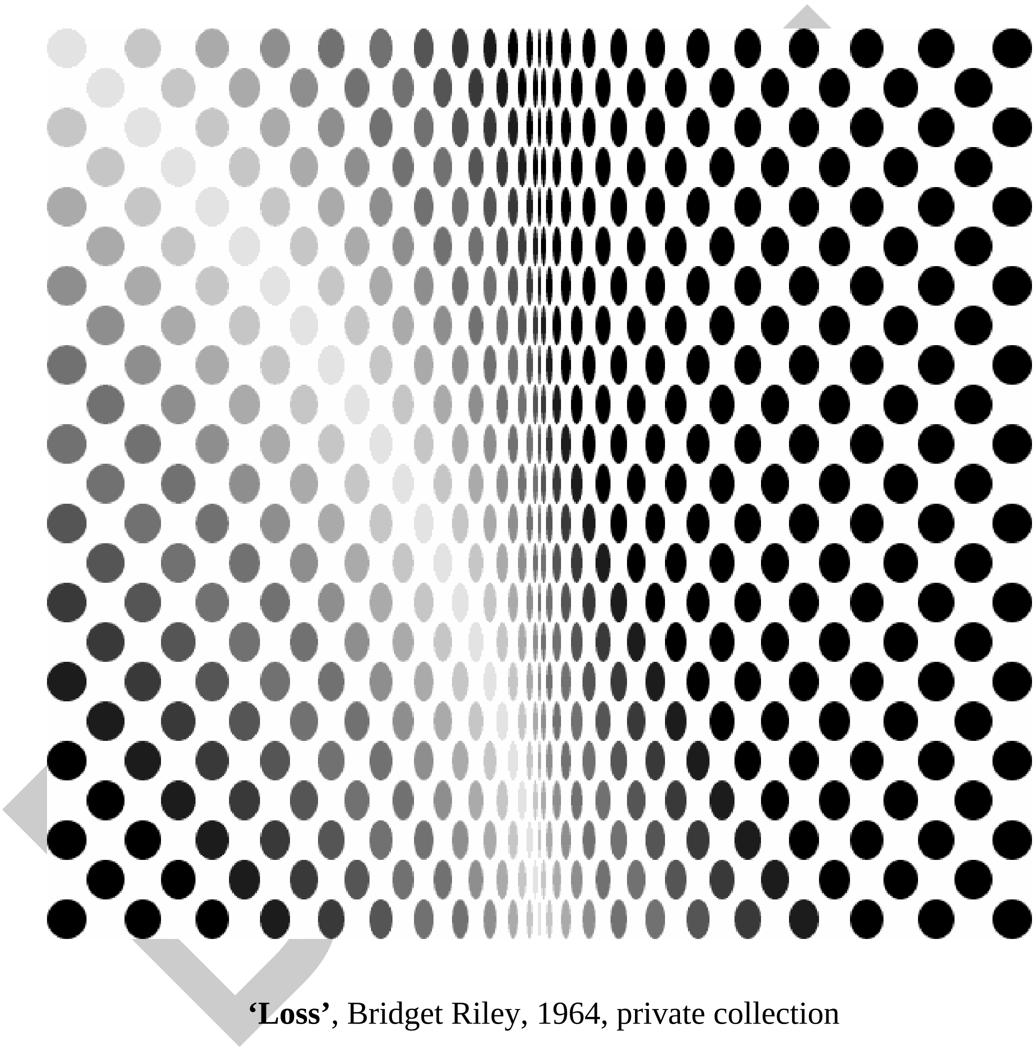
Sensation, Perception, Action – An Evolutionary Perspective

Chapter 3

Vision 1: Brightness



‘Loss’, Bridget Riley, 1964, private collection

Black, white, and shades of grey in static images are the most basic local stimulus properties that the visual system encodes from the environment – and yet they can generate a rich world of perceptions, as so effectively demonstrated by Op Artists.

Overview

The physical luminance of the light emitted or reflected by an object or surface is converted into perceived brightness in the visual stream. This mapping of the outside world on mental representations is the result of neural encoding strategies that reduce redundancy and enhance local contrast. Such processing mechanisms lead to illusions, which often are interpreted as ‘errors’ of representation, but equally can be used as a tool to understand the workings of the brain. Brightness contrast illusions are revisited as a starting point to look at the encoding of light intensity in static images as basic visual processing – seeing bright and dark is arguably the most ordinary task for the visual system. Perceived brightness serves as an input to many other steps of visual information processing, and as such has a fundamental importance for more complex tasks like three-dimensional shape or face perception. Neural encoding through opponency filters explains the resolution of spatial detail in spatial frequency channels, and also gives initial clues for the function of contrast enhancement, redundancy reduction and image compression. These filters therefore can be well understood as a product of the evolution of optimal filter mechanisms. More complex brightness phenomena, such as the conservation of uniform shades of grey perceived within regions enclosed by uninterrupted boundaries, give some initial indication about higher-level processes beyond the initial filtering stage, in particular a filling-in mechanism that is needed to reconstruct the brightness for uniform surfaces. Finally, it will be shown that in less trivial configurations, including the appearance of transparent surfaces and shadows, the visual system provides effective tools to extract the information from images which is most relevant for meaningful interaction with the surrounding world.

Simple visual stimuli

Many readers, I presume, would be most interested in questions about perception that are related to high-level interactions of humans, like the one raised at the end of chapter 2: how can we perceive, interpret, and understand the mysterious smile of Mona Lisa? We have seen in chapter 2 that a dense network of serial and parallel information processing is the basis of visual perception, converting images of the visual field to retinotopic representations of image features (‘neural images’): the three-dimensional environment is captured as flat two-dimensional (2D) images in two eyes and then the information is processed along the visual stream in a set of brain regions (see figure 3.1). The various analysis steps in the visual stream, in the eyes of an engineer, reduce the information in the stimulus and extract the relevant aspects for any particular task, such as face perception, and the recognition of facial expressions. In this framework, the retina works like a CCD camera with some initial data compression, the optic nerve serves as highly parallel data bus, and after a relay station in the LGN the information is distributed across at the early areas of the visual cortex for initial image analysis and storage (Hubel and Wiesel 1979;

van Essen et al 1992). In particular, area V3 is generally thought to be concerned with the representation of form, area V4 serves the representation of colour, and area V5 is specialized for motion processing (Zeki 1993).

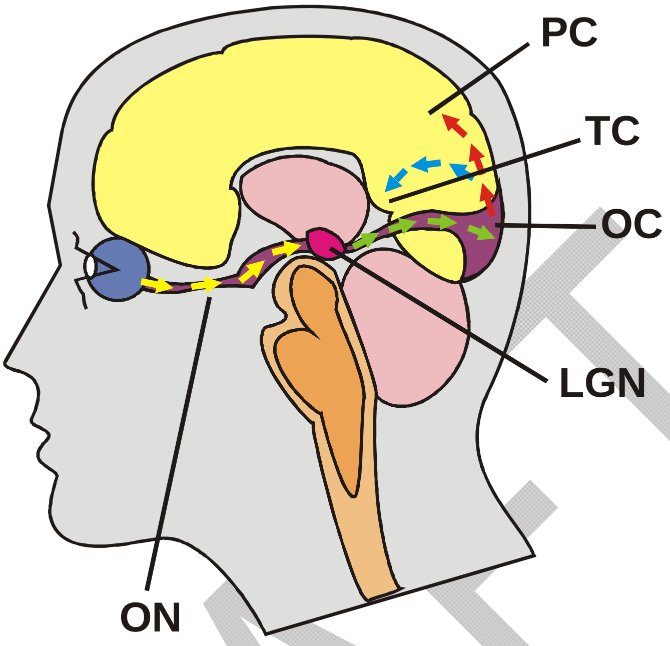


Figure 3.1: Processing in the visual stream. The visual scene is captured by the left and the right eye and transmitted through the optic nerve (ON) and the Lateral Geniculate Nucleus (LGN) into the occipital cortex (OC), from where it is distributed into various parts of the parietal (PC) and temporal(TC) regions of the cortex. In this network of serial and parallel processing, different components of the incoming information (features) are dealt with separately, to be later recombined (in a process call 'feature binding' by psychologists, or 'data fusion' by engineers) to form unitary perceptions for a range of high-level processes, such as recognition, decision making, or motor planning and control. (see Colour Plate II.1)

Given the complexity of the underlying information processing stream, and of the interactions between the brain regions involved in such processing, searching for an understanding of how we can read Mona Lisa's smile is like searching for a needle in a haystack. In chapter 1 it was argued that the key to understanding perception and brain function is to dissect complex problems into simple steps. A quick look at image reproduction technologies can give some clues to how the richness of a visual scene can be reduced to simple representations, thus providing less information whilst retaining the core of an image. Line drawings, or etchings (each image point is black or white, carrying 1 bit of information) were the traditional means of picture reproductions in books. Photography started with grey-level pictures (in digital images usually encoded with 8 bits per image point, or pixel), went on with colour images (usually 24 bits per pixel), and led to cinematography, sequences of moving images (whereas a still photograph can be stored using a couple of hundreds of Kilobytes in compressed format, a movie requires a many Megabytes, and a high-capacity DVD as storage medium). Therefore, in an attempt to reduce the complexity of image processing under investigation, in the first instance (i) we restrict our analysis to static images (motion will be covered in chapter 6) and (ii) focus on greylevel images, ignoring colour (covered in chapter 4).

The Op Art image of Bridget Riley on the frontpage of this chapter gives an indication of how rich perceptual experience can be in absence of colour or movement. In Akiyoshi Kitaoka's collection of simple static patterns (Kitaoka 2002) you can admire the many facets of such images, how they trick the eye, and see how they are augmented by the addition of colour. When studying the perception of brightness, we therefore will have another, and a closer look at a number of illusions. Illusions are often regarded as failures of the sensory system to provide a veridical representation of the outside world – but we will see that they do provide crucial pointers towards understanding the mechanisms of its information processing mechanisms. In this context, illusions are a key to reality!

Descriptors of greylevel images

Let us start with a brief look at the physics of the intensity of light entering the eye. Light can either be emitted by an object, such as a candle or a lightbulb, a it can be reflected from a surface, such as a sheet of paper in sunlight. The amount of visible light that comes to the eye from a surface is called '*luminance*', whereas the amount of light that is shining on a surface is called '*illuminance*' and the proportion of light reflected from a surface is called '*reflectance*' (see figure 3.2). These purely physical properties can be measured using physical devices, such as a luminance meter that measures light intensity, or tallies the number of light particles arising from a small region of a scene. Perceptually, there is an interesting difference between the direct appearance of a surface and its physical properties. Perceived luminance, or the perceptual correlate of the light that arises from a surface, is referred to as '*brightness*', whereas perceived reflectance, or the perceptual correlate of the light that bounces back from an illuminated surface is referred to as '*lightness*'. A piece of white paper looks white in the sunlight, and it also looks white in moonlight, despite the fact that there is much less amount of light coming from its surface to enter the eye. There would be no difficulty for a human observer to notice the difference of brightness under these two illuminations, and still the piece of paper would be recognised as white rather than grey – which is a surface property independent of illumination (Gilchrist 1979). We will return to this interesting aspect of the visual system, which is referred to as 'brightness constancy', when looking at shadows at the end of this chapter.

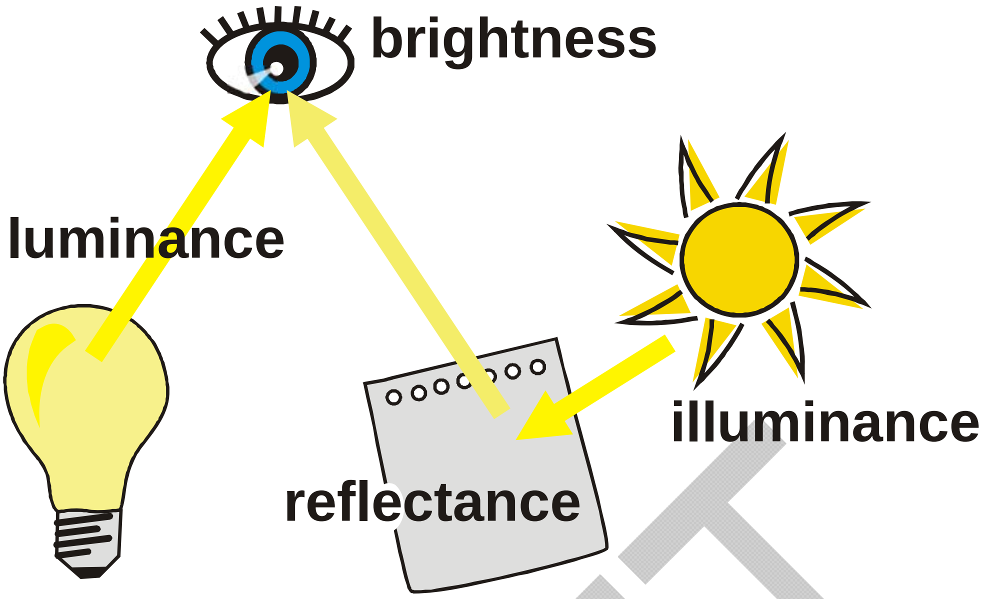


Figure 3.2: The physics of light intensity, as seen by the eye. Light emitted from an object (light bulb) or reflected by a surface (sheet of paper in sunlight) is entering the eye, giving rise to a luminance distribution which is perceived as brightness; the object property that determines the relationship between illuminance – the amount of light hat falls on a surface – and the amount of light reflected from this surface, its ‘reflectance’ is perceived as lightness (white or grey paper).

When it comes to perceiving brightness, perhaps the most basic question is how many different shades of grey a human observer can discriminate. The answer commonly offered to this question, based on discrimination thresholds for boundaries between similar shades of grey at different base levels, is that we are able to discriminate about 5000 shades (4096 different levels could be encoded with 12 bits of information). Notably, this is much more than the 256 grey levels provided by a typical computer screen, or an 8-bit digital image, the typical bitmap format.

Figure 3.3: Perception of brightness. In the grey wedge on the left side, ranging from white to black (surrounded by a frame of opposite polarity) a human observer can discriminate approximately 5000 shades of grey, about 20 times more than the levels provided by a typical 8-bit image: boundaries between areas of a single unit difference (1/256 of range) are visible, as can be seen by close inspection of the set of four squares on the right.

This observation raises a number of scientific questions. What is the best description, and neural representation, of a greylevel stimulus? In chapter 2, we have seen that neural encoding strategies emphasize luminance changes. Indeed, the most efficient description of brightness does not relate to absolute levels of grey, but to differences between neighbouring regions, or contrast. It will shown in the following how these mechanisms affect brightness perception. We will ask what is the relationship between objective (physical) luminance and subjective (perceived) brightness, can we interpret the visual system as measurement device similar to a ruler which is used to measure length. The simple statement about ca. 5000 discriminable levels of brightness is an oversimplification because brightness is not perceived in a linear fashion, which means that equal steps of luminance do not lead to equal steps in brightness. Instead, larger luminance differences are needed at high average luminance levels than at low average luminance levels, in order to perceive a difference in brightness. This fundamental relationship between the intensity of a stimulus and the smallest difference between two different intensities that can be discriminated, is a very common property for all kinds of sensory systems and can be generally described by 'Weber-Fechner Law', which claims that the proportion of change needed to detect an intensity step is approximately constant, later refined as 'Stevens' Power Law' (Stevens 1961; Stevens 1970). Based on this observation, we can get abetter idea about the metrics of brightness, i.e. how the 5000 visible shades of grey are mapped on a luminance scale? Because the smallest visible steps do not have the same size, such a scale is non-linear, unlike the linear scale of a ruler, where a millimetre covers exactly the same distance at 1 cm as at 1 m, for instance. Such scaling extends the overall range of luminance levels that can be discriminated, the so-called 'operating range. In many technical devices, such as video monitors, this nonlinear behaviour is incorporated in the display by means of 'Gamma-correction'.

The next basic question is related to the spatial resolution of the human visual system: how much spatial detail can we see? Usually, you would not be able to see individual pixels on your computer or television screen (typically 1028 by 768 pixels), but I you use a lower screen resolution (640 by 480 pixels) or a larger screen, you will see that it is made up of small image points, or pixels, if you get close enough. This 'spatial resolution' of the visual system can be investigated systematically by line patterns, or gratings (see figure 3.4). It is important to note that increasing or decreasing the viewing distance changes the number of lines that fill a degree of visual angle. The number or grating cycles in a degree of visual angle (cycles/deg) is called by physicists and engineers 'spatial frequency'. The reason why you take a closer look of an image or an object when you attempt to see its fine spatial detail is that the visible detail is limited by resolution of the visual system, in points per degree (or cycles/deg), and by reducing viewing distance you are growing the size of the image (or the cycles of a test pattern), as sketched on the right of figure 3.4. When our eyesight gets worse, for instance in old age, it means that the resolution is reduced, and we need magnifying glasses to recognise fine detail, or read small texts. An optometrist or ophthalmologists using standardised tests to measure visual acuity, and 100% visual acuity corresponds to the resolution of approximately 50 cycles per degree (Adams et al 1988; Westheimer 1984).

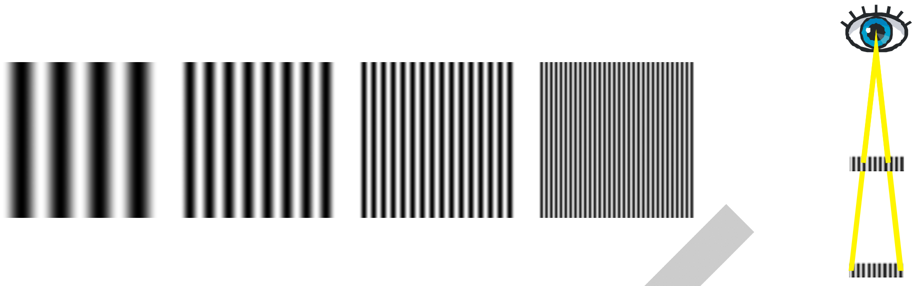


Figure 3.4: Perception of spatial detail can be tested with textures or line gratings of increasingly fine grain, usually measured as spatial frequency (number of lines, or grating cycles per degree of visual angle). Looking at the page from a distance of 80 cm, at which each of the four squares (14 x 14 mm) covered with gratings has the angular size of ca. 1 deg, the 4 gratings have a spatial frequency of 4, 8, 16 and 32 cycles/deg. If you increase your viewing distance further, at some point the fine gratings blur into a grey field and are no longer visible: this is your resolution limit. The sketch on the right shows how spatial frequency of a grating image increases with viewing distance: the 10 degrees of visual angle between the two yellow lines is filled with 5 grating cycles (i.e. 5 cycles/deg) at the smaller viewing distance, and 10 grating cycles (i.e. 10 cycles/deg) at the larger viewing distance.

Again, there are a number of scientific questions to be asked. How can spatial detail be described, and what is the metric of such detail? As mentioned above, for simple gratings a straight forward measure of spatial variation is the number of cycles per degree, or spatial frequency. It turns out that this is a very general and useful measure because physicists use a mathematical framework, Fourier Theory (Bracewell 1986), to compose arbitrary spatial patterns from a combination of sinewave gratings. Whereas a psychology student not necessarily needs to understand workings of Fourier Theory, it is important to not that it is very common to investigate the function of the sensory systems with sinewave gratings and to express its performance in terms of frequency. In the case of spatial vision this is spatial frequency, measured in cycles/deg (Campbell and Robson 1968; Watson et al 1983). What are the experimental approaches to investigate the spatial properties of the human visual system? For instance, we can measure detection thresholds for gratings at different contrasts, or how well different gratings can be discriminated. Starting from this, the performance of less simple stimuli can be determined, and by applying Fourier Theory mathematical models for the processing of arbitrary patterns can be developed (Wilson 1991). In the nervous system, the tuning of individual neurons to spatial frequency can be determined and compared for a range of neurons and brain regions (Tootell et al 1981). This brings us back to the simple receptive field structures of retinal ganglion cells introduced in chapter 2, which already show size, or spatial frequency, tuning. Each neuron in the early visual system has characteristic spatial frequency tuning, where the optimum frequency corresponds to the receptive field size (cf. figure 2.16). Groups of neurons with different optimum frequencies cover the full range of spatial frequencies visible to the human observer, the so-called Contrast-Sensitivity Function (CSF, see figure 3.5). The universality of such neuronal properties, and the power of this approach, led to scientific discussions whether this is a complete picture of early visual processing, and whether the visual brain can be understood as a ‘Fourier Analyser’ (Maffei and Fiorentini 1973; Ochs 1979).

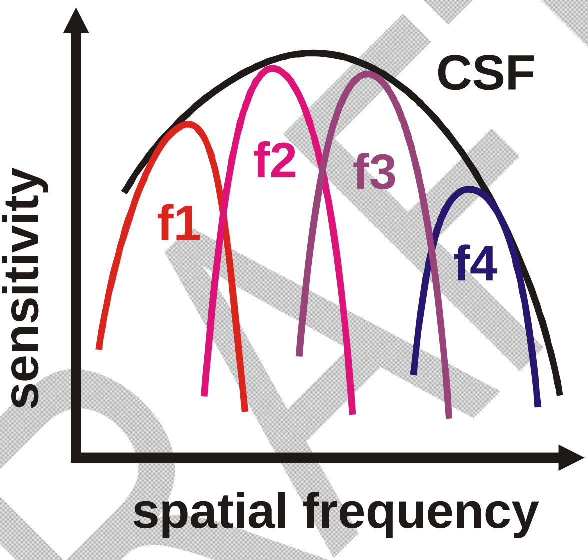


Figure 3.5: Schematic sketch of the contrast sensitivity function (CSF, black line), which represents the sensitivity of the human visual system for gratings with variable spatial frequency. It can be understood ass the combination of a set of neurons tuned to different optimum spatial frequencies, f1 – f4 (coloured lines), or spatial frequency channels.

The fundamental concept: filtering

Whether or whether not the brain can be regarded as Fourier analyser, there is a bulk of experimental evidence (Campbell and Maffei 1974) that visual function can be explained to a large degree by *spatial frequency channels* that transmit and represent fine- to coarse-grain versions of an image, which are processed, stored, used in separate neurons. Such filter mechanism are interpreted as a property emerging from the receptive field structure of sensory neurons, which are organised in parallel sets not only for spatial frequency (figure 3.5), but equally for other properties, such as colour, orientation, or velocity (see chapter 2). When the action of sensory neurons is described as such filtering - what does this mean in more conceptual terms? The parallel sets of neurons are interpreted as pattern analyzers, which are organised as filter banks operating in the visual stream. The size and structure of receptive fields determines the optimum size, or spatial frequency, of encoded stimulus, which is then described as the filter mechanism processing only a selection of the information available. To understand the function of such spatial frequency channels at an intuitive level, they can be compared to a simple mechanical system – a gravel pit. In the sorting engine of the gravel pit, mixed sediments which were dug up from a river bed, are separated into materials of different grain: sand, pebbles, and rocks. A simple way of doing this is involves sets of sieves which allow material of a particular grain pass (i.e., filters), whereas larger grain is retained in the flow (see figure 3.6 left). At the end of such processing the different components included in the original mix can be collected in separate containers, which then contain material of a particular grain.

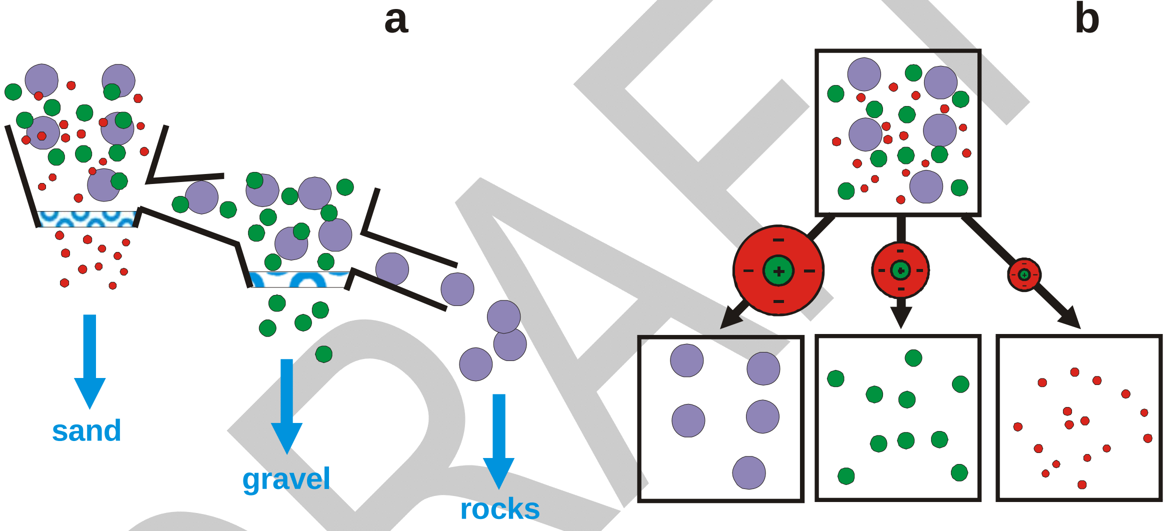


Figure 3.6: Filtering. (a) Think of the brain as a gravel pit: when the sorting engine is fed with a mix of rocks (blue disks), pebbles (green), sand (red), different grains are separated by a set of sieves which act as a filter – and different grains can be collected in different containers. (b) In the same way, think of the visual stream in the brain as a set of filters (spatial frequency channels, indicated by concentric receptive fields) that separate features of different spatial scale (or size) – and image representations with coarse (blue discs) to fine spatial grain (red dots) are generated in the brain. (see Colour Plate II.2)

In the same way, in the visual system parallel sets of receptive fields are operating to separate different spatial frequency components from that are contained in the retinal image. Through this filtering the visual scene is analysed and represented in a set of different spatial frequency channels (see Wilson 1991): Fourier analysis. At some later stage of the visual stream the information from these separate representations need to be combined, together with other image features such as colour or motion, to deliver a unitary and rich percept – therefore, as will be discussed in the final section of this chapter, such Fourier analysis is a good description of the early stages of visual systems, but does not capture all aspects of visual perception.

Illusions as key to reality

When was the last time that you were sitting on the train, waiting for it to leave the station, and when looking out of the window at the train right on the next track you got the impression that your train started moving – only to notice a few seconds later that you are still waiting at the platform, and that it was the other train which started to move? Quite often, your eyes are tricked when looking at the surrounding world, and you have to realise that your senses do not provide you with reliable or veridical information about the outside world, but generate perceptual illusions. Painters, when trying to provide 'realistic' representations of the world, are using a bag of tricks to create such illusions. For instance, in order to create an impression of depth on a flat canvas, the theory of geometric perspective and shading has been systematically developed and exploited since the Renaissance in Western Visual Arts (Gombrich 1977). Op Artists in the second half of the twentieth century started experimenting with 'optical illusions' created in simple line drawings or colour patterns (Riley 1995; Wade 2003). With the introduction of computer aided editing and animation techniques, in the 21st century the art form of cinematography reached new levels of presenting augmented, puzzling, or impossible – but visually absolutely convincing – worlds (just think of the special effects in movies like Harry Potter or Lord of the Rings). Even life television broadcasts should be watched with a certain level of scepticism (remember the fireworks in the opening ceremony of the Bejing Olympics?). Furthermore, illusions and other visual puzzles are frequently used in advertising, which may be less surprising in view of the fact that the reality of a product is rarely what customers want to buy.

The important thing to notice is that we are constantly exposed to illusions, and that our senses do not tell us in a reliable way what exactly is out there (Gregory and Gombrich 1973). Whereas we usually think that we should only believe what we see, we also could arrive at the conclusion that we never should believe what we see (Coren and Girgus 1978; Marshall 1987). Although this suggestion may be rather sobering, or perhaps even irritating, there is something very exciting about illusions: whilst puzzling our minds, these phenomena tell us a lot about how our brain works. In fact, visual illusions are rarely 'optical', and don't trick our eyes, but reflect how our brains are processing visual information – and for the very same reason illusions play a major role in the scientific study of sensory systems. So what can we learn from illusions about visual information processing? In science, minimal stimulus configurations are essential to study perception. Figure 3.7 shows a famous example of a very simple drawing that generates an illusion, the 'Kanizsa triangle'. We see a white triangle in front of three lines and three discs, and sometimes perceive the triangle brighter than the background – this is referred to as illusory or 'subjective' contours (Kanizsa 1976).

Figure 3.7: 'Illusory' Kanizsa Triangle. What looks like a bright triangle in front of three disks and three lines radiating from the image centre, in fact is a pattern of three short lines and three disk sections ('pacmen') oriented towards the centre.

Using illusions generated by such simple stimuli as a guidance to understand the visual system, two fundamental approaches to understand illusions have been. On one hand, we have supporters of 'Gestalt Psychology' who interpret illusions as reflection of perceptual organisation (Kanizsa 1979; Koffka 1935). Within this conceptual framework, perceptions are understood as unitary experiences emerging from and going beyond the components of a stimulus: "the whole is more than the sum of parts". In cases like the illusory triangle, the perceptual system chooses the best, simplest and most stable shape, according to the Gestalt law of 'Praegnanz'. More about this can be found in chapter 13. On the other hand, illusions are understood as by-products of the processing mechanisms underlying perception (see chapter 2), which under some specific conditions lead to discrepancies between the physical stimulus and its mental experience that could be regarded as representation 'errors'. In the last decades a wealth of evidence has been accumulated of properties of individual cortical neurons that can account for the perceived illusion of subjective contours (Paradiso et al 1989; Peterhans and von der Heydt 1991). This physiological approach is complemented by computational modelling, formalising the operations carried out by the neurons along the visual stream (Marr 1982). Illusions tell us a lot about the mechanisms of visual information processing.

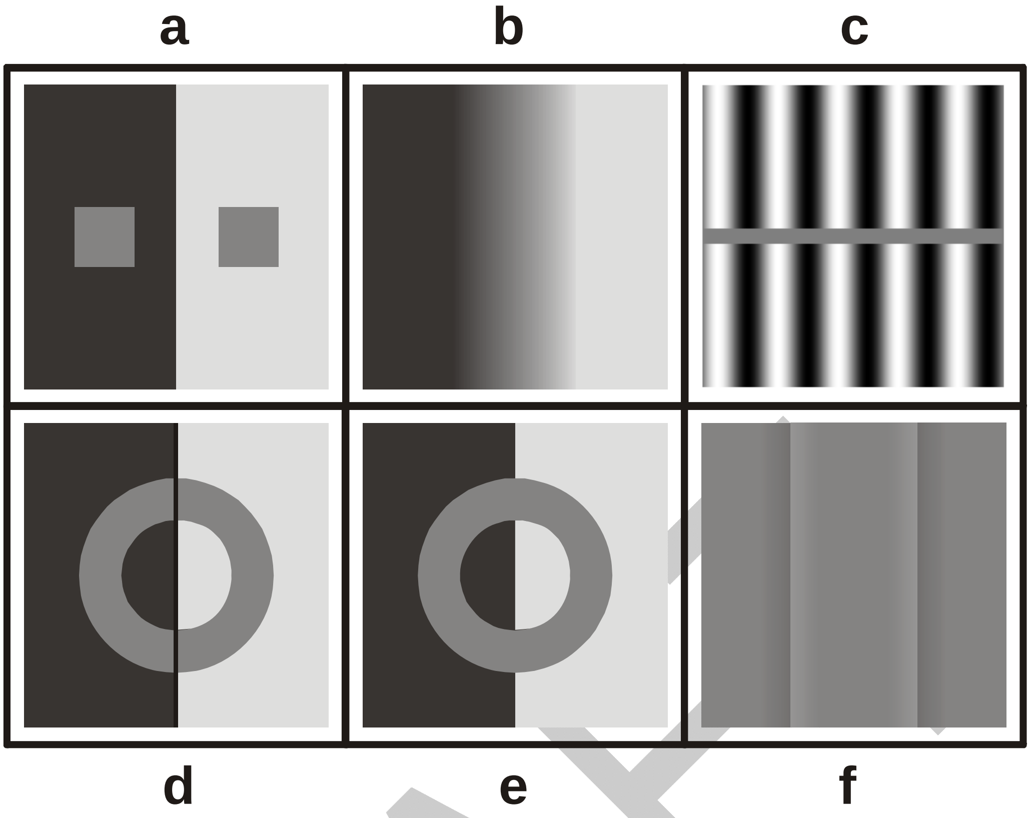


Figure 3.8: A zoo of brightness illusions. Top row: (a) simultaneous contrast – the brightness of the square is influenced by the shade of grey in the background, see chapter 2; (b) Mach bands – a dark vertical line and a bright vertical line are perceived when luminance starts and ends to increase gradually from one low level to high level; (c) grating induction – the grey horizontal stripe in the centre appears to be modulated in brightness in reverse to the vertical gratings above and below. Bottom row: (d) partitioned Koffka ring – simultaneous contrast; (e) open Koffka ring – in a single object crossing the background luminance border, simultaneous contrast breaks down and a unitary brightness is perceived for the whole object; (f) Craik-Cornsweet illusion – the vertical region in the centre looks brighter than the left and right regions, although they only differ in the boundary regions but not in the middle.

Some of the challenges for such an information processing approach is illustrated by some examples from the 'zoo of brightness illusions' shown in figure 3.8. In the top row of this figure some examples are shown of how the perceived brightness at a given location depends on its neighbourhood, like the simultaneous contrast illusion presented in chapter 2, and similar effects in Mach bands (Ross et al 1989) and grating induction (Blakeslee and McCourt 1997). The bottom row of this figure shows some examples of how the perceived brightness of a clearly outlined surface tends to be uniform, such as the Koffka ring (Koffka 1935) or the Craik-Cornsweet illusion (Cornsweet 1970). The functional significance, and the co-operative relationship between the processes of contrast enhancement and filling in, will be discussed in the two next sections.

Contrast enhancement in brightness encoding

In chapter 2 it was demonstrated that opponent filtering, as generated by the lateral inhibition mechanisms of centre-surround receptive fields, generates minimum and maximum responses at luminance boundaries and discards average luminance. This phenomenon of contrast enhancement is illustrated schematically in figure 3.9: the perceived brightness profile at a luminance boundary is generated through the opponent integration mechanism between the excitatory centre and inhibitory surround of a concentric ON-centre receptive field.

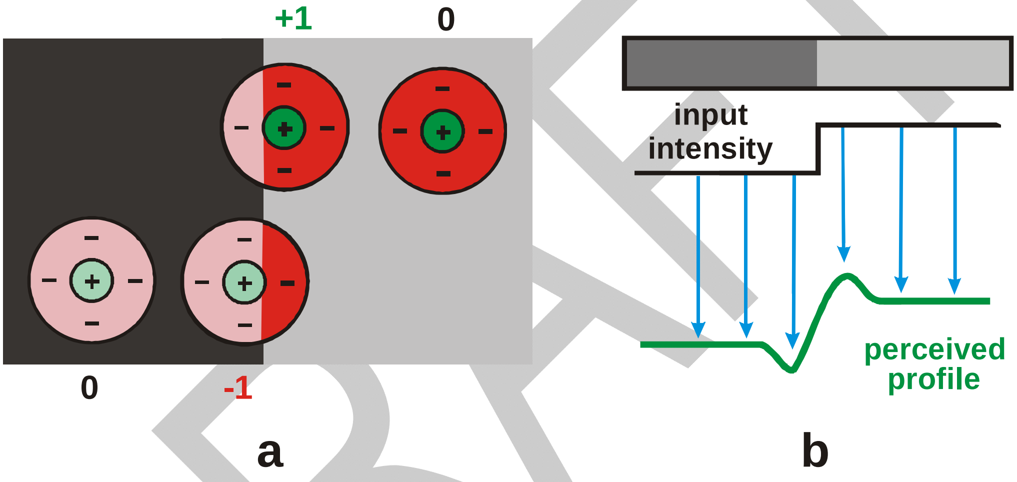


Figure 3.9: Contrast enhancement (similar to figure 2.17). (a) Opponent ON-centre receptive fields – shown as simplified opponency model (cf. figure 2.9) with a central excitatory (light and saturated green) and a surrounding inhibitory (red, pink) sub-region – at a different spatial positions of a pattern with a prominent luminance step; reduced response (0) in coherently dark and bright areas: excitation and inhibition cancel each other; negative response (-1 in arbitrary units) at dark side of boundary: more inhibition than excitation; positive response (+1) at bright side of boundary: more excitation than inhibition. (b) schematic sketch of the encoding of a luminance step function (intensity function, black line) is encoded through such opponency operators into a profile encoded in a activity profile which corresponds to the perceived contrast enhancement. (see Colour Plate II.3)

The same action of opponent receptive fields can account for illusions with less simple patterns, like the multiple luminance step function shown in figure 3.10a. Although each of the five vertical bars has a constant luminance, they are perceived as darker on the left than on the right. This can be easily understood as the influence of the contrast enhancement at each luminance step as illustrated in figure 3.10b.

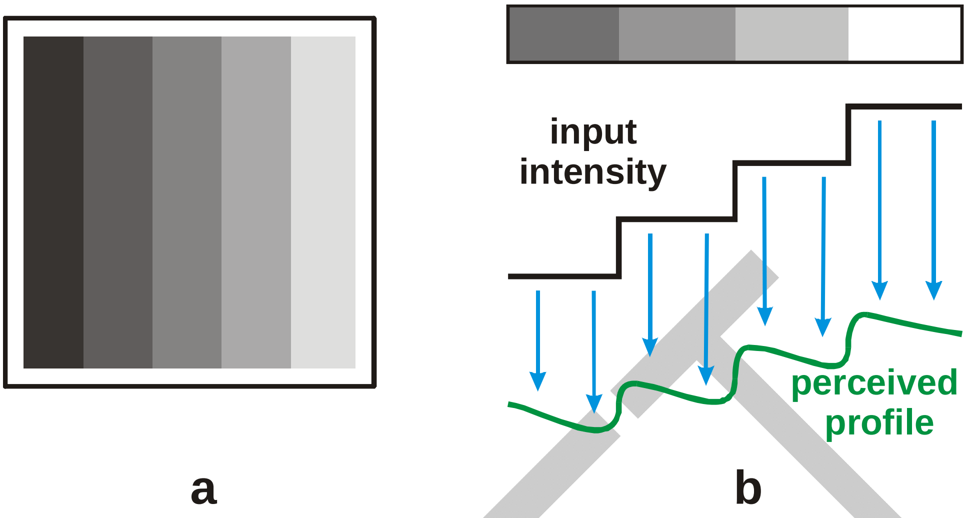


Figure 3.10: What you have is not what you see. The multiple intensity step function of the stimulus pattern (a) is transferred by the boundary contrast enhancing opponent receptive fields such that at each level of the luminance steps the grey seems to be brighter at the left boundary than at the right boundary (b).

The very same organisation of receptive fields which can be interpreted as coding strategy to amplify luminance differences across contours also operates in two spatial dimensions. The famous ‘Hermann Grid’ shown in figure 3.12a is made up of a regular pattern of black squares separated by small white regions (‘edges’). At the intersections of vertical and horizontal edges grey blobs are perceived, as long as they are not in the centre of the visual field. In the visual field centre the intersection is projected on the fovea, the retinal region of highest spatial resolution, where receptive fields are smaller and don’t match the size of the white regions as is necessary for the explanation of the perceived grey blobs suggested in the next paragraph.

In figure 3.11 a neural explanation is provided for the grey blobs appearing in the intersections of the white edges. In opponent filters (ON-centre-OFF-surround receptive fields), larger parts of the inhibitory surround are stimulated in the intersections of the edges, as compared to the white edge regions between the intersections, which explain the apparent reduction of brightness seen as grey blobs. This simple, schematic consideration demonstrates how the contrast enhancement delivered by opponent filters operates in two spatial dimensions, leading to a strong brightness illusion in the Hermann grid. An interesting amplification of this illusion is the ‘scintillating grid’, a quite recent discovery that raises interesting new questions about the underlying brain mechanisms (Schrauf et al 1997). Obviously the explanation put forward here is a very idealised picture, an despite the simple explanation of the illusory grey blobs, there are other aspects which are in want for an explanation: The opponent filter mechanism would predict darker regions close to boundaries of the squares than in the middle of the squares, but each square appears constantly black. Additional aspects of image encoding that can explain this effect will be discussed in the next section.

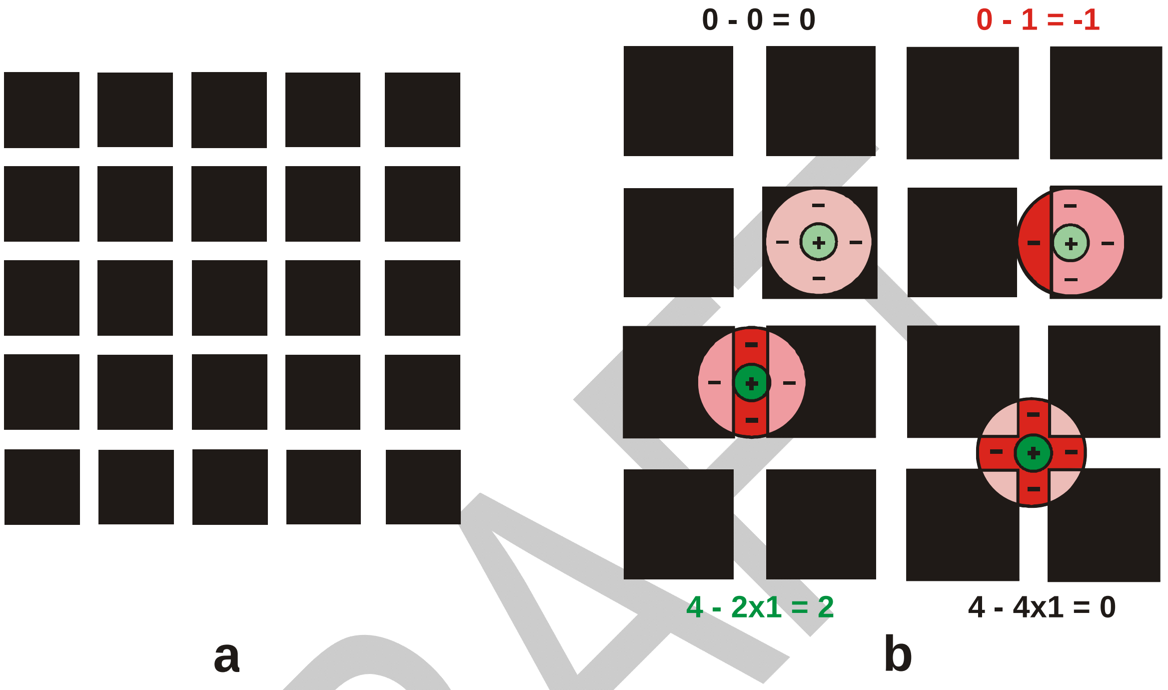


Figure 3.11: (a) can you see grey spots in the white intersections of lines of this ‘Hermann Grid’? (b) Opponent receptive fields – shown schematically as in figure 3.9 – at a different spatial positions of the Hermann grid would generate no response within the black squares (top left: no stimulation of excitatory and inhibitory regions), and a smaller response in the intersections of the edges (bottom right: full excitation and inhibition) than in the edges themselves (bottom left: full excitation, some inhibition). However, for the same reason the squares should be darker in regions close to the white edges (top right: some inhibition, no excitation). (see Colour Plate II.4)

Image compression in brightness encoding

Revisiting the variety of simultaneous contrast illusion between regions and at luminance boundaries (see figure 3.8) in the context of the observation made for the Hermann grid (figure 3.11) that each black square appears uniformly black, leaves us with a small paradox. The simple opponency mechanism discussed so far explains how the luminance of the background affects the brightness of the grey square in the foreground of figure 3.8a, and the underlying lateral inhibition would predict a particular enhancement of luminance borders, because the modulation of activation from nearby regions should be largest at a luminance step. However, whilst such a contrast enhancement is perceived for the multiple luminance step pattern shown in figure 3.10, it is not visible at the boundaries of the squares in figure 3.8a or figure 3.11 – the squares clearly have the same brightness in the complete area surrounded by the luminance border. A similar uniform level of grey is perceived in the open Koffka ring shown in figure 3.8e, where the simultaneous contrast illusion that is visible in the partitioned ring (figure 3.8d) breaks down for a coherent shape crossing a luminance boundary in the background. This discrepancy suggests that the operation of simple receptive fields does not provide a full explanation of how we perceive brightness, and perhaps – not surprisingly – some additional, higher level mechanisms need to be considered. Considering the limitations of the most simple models will bring us back to the fundamental questions of how perceived features reflect physical properties, and what the function ('purpose') of opponency encoding might.

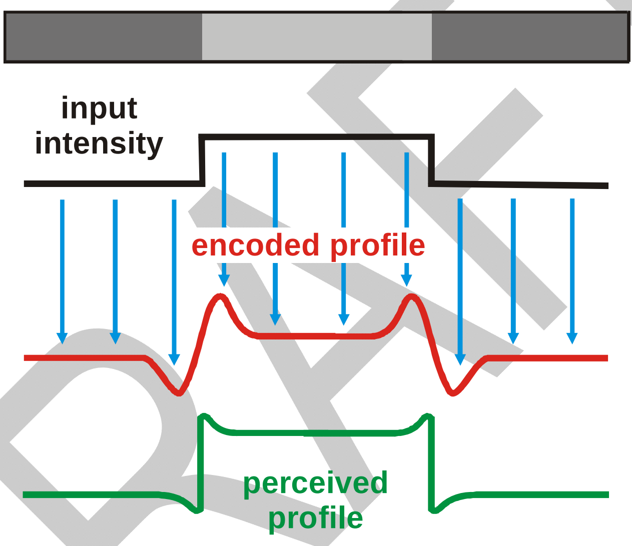


Figure 3.12: What you have (black input intensity) and what you see (green profile), in this example, is not what you what you should see if simple opponency filtering (symbolised by blue arrows) would be the only encoding mechanism in the visual stream (red encoded profile) – a region of uniform luminance embedded in dark background should have enhanced contrast in the border regions and reduced brightness in central region (compare with figures 3.9 and 3.10, for instance).

This discrepancy between filtering expectation and the observation of uniform grey levels inside borders can be best seen for the dark-bright-dark profile shown in figure 3.12: the expectation for the encoded profile, if the neural image generated by opponency filtering, would show strong contrast enhancement at the luminance boundaries, and very similar activation levels in the inner regions of the brighter and darker sections. Because the exclusive encoding of spatial changes through opponent processes removes redundant image components that do not contain information, it can be interpreted as an effective image compression strategy. When no signal is transmitted for regions without luminance change, and information only is encoded at luminance boundaries, the information about the uniform luminance level between such boundaries is no longer represented. The fact that for the stimulus illustrated in figure 3.12 perception is dominated by a clearly outlined region of enhanced brightness therefore raises the question of how the average luminance level of objects is reconstructed in the brain. There could be various possible mechanisms, such as combining information from different spatial scales, which could support such a reconstruction process. The observation that the enclosure of an image region is so important for performing uniform brightness (see figure 3.8d & e) led to the proposal of a cortical 'filling-in' mechanism by which regions that are surrounded by clear boundaries assume the same brightness (Pessoa et al 1995). The result of such a reconstruction mechanism, which also can account for the uniform black squares in the Hermann grid (figure 3.11) is illustrated in figure 3.12 and 3.13.

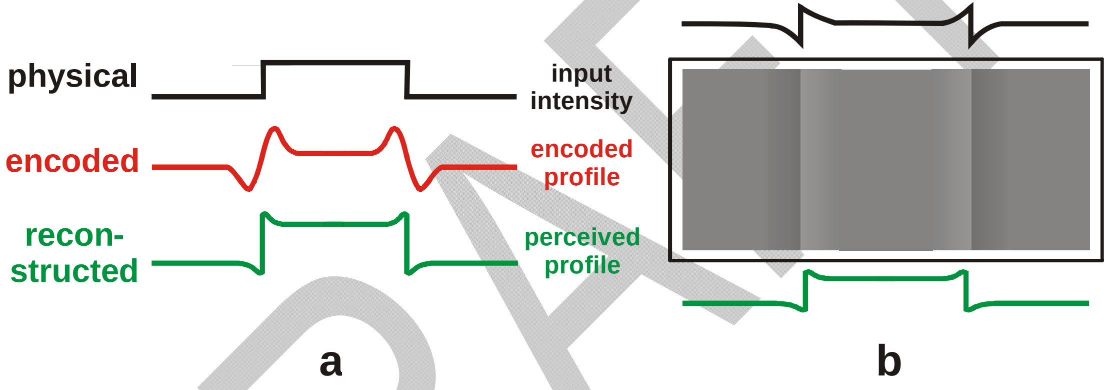


Figure 3.13: (a) The discrepancy between physical luminance profiles (black line) and activity profiles as encoded by opponent filters (red) is compensated by filling-in mechanisms that reconstruct the percept close to the original profile for the stimulus illustrated in figure 12 (green). (b) Craik-Cornsweet illusion. The physical luminance profile (black line above the stimulus pattern) of this stimulus is characterised by contrast edges of opposite sign without changes of mean luminance – rather different from the luminance steps shown in (a). Filling-in mechanisms, however, generate a reconstructed profile (green line below stimulus pattern) that resembles that in (a), which explains the illusory perception of a bright region embedded in a darker background.

In essence, a filling-in mechanism reconstructs the redundant image regions by assuming that regions enclosed by contiguous boundaries originally have the same intensity. In view of such processing one could expect a stimulus that has a physical profile, which looks like the encoded profile that only indicates luminance borders without any change of average luminance, would lead to a very similar encoded profile and after the reconstruction should be perceived like a change of mean brightness. This effect is nicely demonstrated by the Craik-Cornsweet illusion (Cornsweet 1970) illustrated in figure 3.13b. The central area appears brighter than the regions on the left and right, although the physical luminance is identical in the centre and the peripheral regions, and is changing only in the visible boundary regions. This illusion demonstrates the ‘filling in’ mechanism: surfaces between boundaries are apparently filled with uniform brightness. Note that this is an ‘inverted’ illusion, where the perceived brightness profile is reconstructed from an ambiguous luminance profile, assuming certain properties of the outside world like uniform coloration. The Craik-Cornsweet illusion demonstrates that if in the attempt to reduce redundancy only changes intensity are encoded, uniform properties of the physical properties need to be assumed between such changes in higher-level interpretation from sparse data. In consequence, subtle gradients are ignored.

Advanced brightness perception: Does the brain know the laws of physics?

The Craik-Cornsweet illusion is a first indication that in perception the relation between physical properties and internal representations can be non-veridical in such a way that the most likely surface property is recovered. When a not perfectly flat sheet of paper that is placed on a black background is illuminated from the side, for instance, the uneven paper surface will show slight variations of luminance - shading. Still it will be interpreted not as a piece of paper with faint grey stains but as perfectly white, the visual system reconstructs the uniform white surface property between the strong luminance boundaries. We encounter capabilities of the visual system like this, which go beyond the most basic effects of visual encoding covered so far, in a range of contexts that are related to recognition of objects and making them available to interaction. Three examples of such phenomena are described briefly to close the chapter with a wider perspective, which will further expanded in chapter 13.

Looking through a transparent surface, like tinted glass or wrapping paper, is a not uncommon situation that changes the luminance (or colour) of all objects behind the transparent material. When reading a newspaper whilst wearing dark sunglasses, however, you don’t get the impression that it is printed on grey paper (and just from wearing pink sunglasses you will not get the impression you are reading the Financial Times). Somehow, therefore, the brain is able to account for transparency. At the top of figure 3.14a you see a light grey rectangle that is partially covered by a slightly darker, transparent rectangle. When looking at this image you would not interpret it as a three adjacent shapes covered with three different paints, which is however the preferred interpretation as soon as the figure is disintegrated into three separate shapes (bottom of 3.14a). The perception of transparency is a very powerful tool for the visual system that helps us to recognise physical properties of the world surrounding us (Metelli 1974).

In the Checker Shadow Illusion (figure 3.14b) a diffuse shadow is cast over a regular pattern of dark and bright tiles, which changes the physical luminance of individual elements considerably. In our internal representation, however, the surface is still covered with dark and bright tiles. This leads to the effect that two tiles that are perceived as belonging to the set of bright and dark tiles, respectively, actually have the same physical luminance. Once again, it seems as if visual system uses good knowledge of the laws of physics, and by realising that a shadow is cast over the tiling pattern is able to retrieve the surface properties – in this case the colouring of each tile, independent of illumination condition – of objects for the internal representation of the outside world. It is important to note that it is the intrinsic properties of objects what is important to understand and remember, and that this goes far beyond just taking and storing images (Adelson 2001).

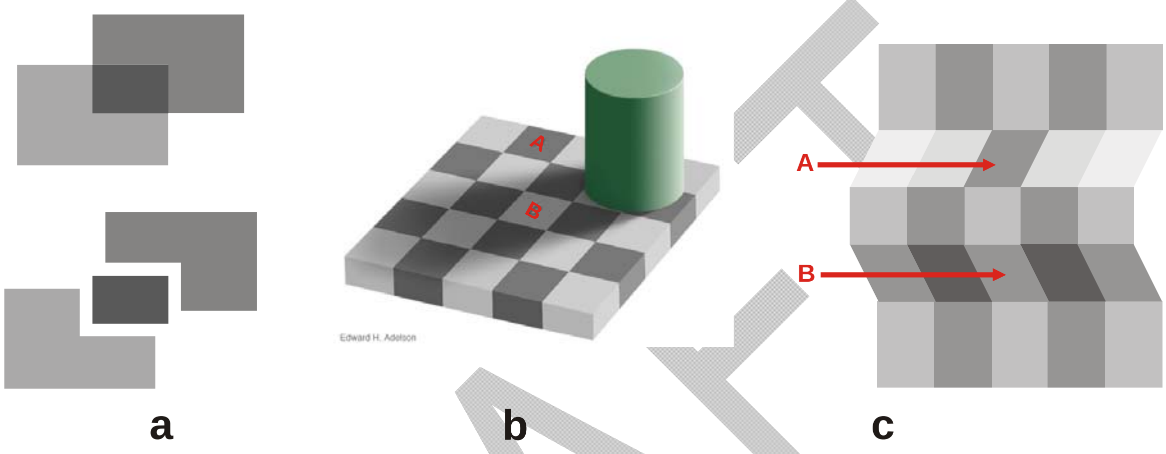


Figure 3.14: Higher-level aspects of brightness perception. (a) Transparency: when 3 shapes covered with a particular combination of grey (bottom) are arranged in a special way (top), transparency is perceived. (b) Checker Shadow Illusion: perceived brightness in regions of a regular tiling pattern that are overcast by a shadow differs from the physical luminance of these regions: tiles A and B (red labels) have the same luminance but differ in perceived brightness (©1995, Edward H. Adelson) (c) The Corrugated Plaid Illusion demonstrates that a similar difference of brightness for areas of identical luminance (regions A and B, red arrows) can be perceived in the absence of explicit shadows as well, if they seem to be exposed to different illumination.

When thinking about the mechanisms for this correction of perceived brightness, a possible candidate could be to detect the shadow and then subtract its contribution to the illumination of the tiles, and thus recover the original luminance of each tile. The Corrugated Plaid Illusion, shown in figure 3.14c, demonstrates that a similar difference of brightness for rectangles of identical luminance (indicated by red arrows) can be perceived in the absence of explicit shadows as well. Here it seems to be the three-dimensional interpretation of the image as a folded set of squares painted with different greys which is crucial for the illusion, which might be related to an implicit reduction of illumination on the seemingly downward facing surface region. When the set of squares is perceived as folded horizontally such that the two rectangles appear to sit on the same plane, luminance and brightness both are identical. Whereas one might think that this illusion could suggest that the brain somehow has to reconstruct some rather complex variations of surface orientation and lighting conditions, there are proposals how the same behaviour can be modelled with some rather simple and pragmatic computations (Adelson 2000).

Take home messages

- visual information is re-organized, compressed, and categorized by parallel and hierarchical processing mechanisms to encode information efficiently in the visual stream
- the encoding of levels of grey in static images is one of the most basic tasks performed in the visual stream, determining perceived brightness
- opponent filters encoding by centre-surround receptive fields is a crucial strategy to enhance local contrast and remove redundancy, and crucial for brightness perception
- opponency can account for a number of illusions such as simultaneous contrast, but additional mechanisms (like filling-in) are required to deal with other aspects of perception
- illusions often are explained as high level mechanisms that construct solutions to puzzles; many illusions are a consequence of basic information processing strategies, others require higher-level mechanisms that are helpful to reconstruct the most likely properties physical objects

Discussion Questions

- How many different shades of grey can humans discriminate? How does this compare to technical systems?
- Describe the action of concentric receptive field, and how they can act as spatial filters.
- What is meant by cortical 'filling-in'?
- Discuss the phenomenon of transparency in the context of brightness representation.

References

Adams A J, Wong L S, Wong L, Gould B, 1988 "Visual Acuity Changes with Age: Some New Perspectives" *American Journal of Optometry and Physiological Optics* **65** 403-406

Adelson E H, 2000 "Lightness Perception and Lightness Illusions", in *The New Cognitive Neurosciences* Ed M S Gazzaniga (Boston: Boston) pp 339-351

Adelson E H, 2001 "On seeing stuff: the perception of materials by humans and machines", in *Proceedings of the. SPIE Vol. 4299, Human Vision and Electronic Imaging VI* Eds B E Rogowitz and T N Pappas pp 1—12

Blakeslee B, McCourt M E, 1997 "Similar Mechanisms Underlie Simultaneous Brightness Contrast and Grating Induction" *Vision Research* **37** 2849-2869

Bracewell R N, 1986 *The Fourier Transform and its Applications* (Singapore: McGraw Hill)

Campbell F W, Maffei L, 1974 "Contrast and Spatial Frequency" *Sci.Am.* **231** 106-114

Campbell F W, Robson J R, 1968 "Application of fourier analysis to the visibility of gratings" *J.Physiol.* **197** 551-566

Coren S, Girgus J S, 1978 *Seeing is Deceiving: The Psychology of Visual Illusions* (Hillsdale: L. Erlbaum Ass.)

Cornsweet T N, 1970 *Visual Perception* (Orlando: Academic Press)

Gilchrist A L, 1979 "The Perception of Surface Blacks and Whites" *Scient.Am.* **240** 88-97

Gombrich E H, 1977 *Art and Illusion* (Oxford: Phaidon)

Gregory R L, Gombrich E H, 1973 *Illusion in Nature and Art* (London: Duckworth & co. Ltd.)

Hubel D H, Wiesel T N, 1979 "Brain Mechanisms of Vision" *Scient.Am.* **241** 130-144

Kanizsa G, 1976 "Subjective Contours" *Scient.Am.* **234** 48-52

Kanizsa G, 1979 *Organization in Vision. Essays on Gestalt Perception* (New York: Praeger Publishers)

Kitaoka A, 2002 *Trick eyes* (Tokyo: Kanzen)

Koffka K, 1935 *Principles of Gestalt Psychology* (London: Routledge)

Maffei L, Fiorentini A, 1973 "The visual cortex as spatial frequency analyser" *Vision Research* **13** 1255-1267

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

Marshall J C, 1987 "Is seeing believing?" *Nature* **325** 583-584

Metelli F, 1974 "The Perception of Transparency" *Scient.Am.* **4/74**

Ochs A L, 1979 "Is Fourier analysis performed by the visual system or by the investigator" *Journal of the Optical Society of America* **69** 95-98

Paradiso M A, Shimojo S, Nakayama K, 1989 "Subjective contours, tilt aftereffect, and visual cortical organization" *Vision Research* **29** 1205-1213

Pessoa L, Mingolla E, Neumann H, 1995 "A Contrast- and Luninace-driven Multiscale Network Model of Brightness Perception" *Vision Research* **35** 2201-2223

Peterhans E, von der Heydt R, 1991 "Subjective contours - bridging the gap between psychophysics and physiology" *TINS* **14** 112-119

Riley B, 1995 *Dialogues on Art* (London: Zwemmer)

Ross J, Morrone M C, Burr D C, 1989 "The conditions under which Mach bands are visible" *Vision Research* **29** 699-715

Schrauf M, Lingelbach B, Wist E R, 1997 "The Scintillating Grid Illusion" *Vision Research* **37** 1033-1038

Stevens S S, 1961 "To Honor Fechner and Repeal His Law: A power function, not a log function, describes the operating characteristic of a sensory system" *Science* **133** 80-86

Stevens S S, 1970 "Neural Events and the Psychophysical Law" *Science* **170** 1043-1050

Tootell R B, Silverman M S, DeValois R L, 1981 "Spatial Frequency Columns in Primary Visual Cortex" *Science* **214** 813-815

van Essen D C, Anderson C H, Felleman D J, 1992 "Information processing in the primate visual system: An integrated systems perspective" *Science* **255** 419-423

Wade N J, 2003 "Movements in art: from Rosso to Riley" *Perception* **32** 1029-1036

Watson A B, Barlow H B, Robson J G, 1983 "What does the eye see best?" *Nature* **302** 119

Westheimer G, 1984 "Spatial Vision" *Annual Reviews of Psychology* **35** 201-226

Wilson H R, 1991 "Pattern discrimination, Visual Filters, and Spatial Sampling Irregularity", in *Computational Models of Visual Processing* Eds M S Landy and J A Movshon (Cambridge MA: Cambridge MA) pp 153-168

Zeki S, 1993 *A Vision of the Brain* (Oxford: Blackwell Scientific Publications)