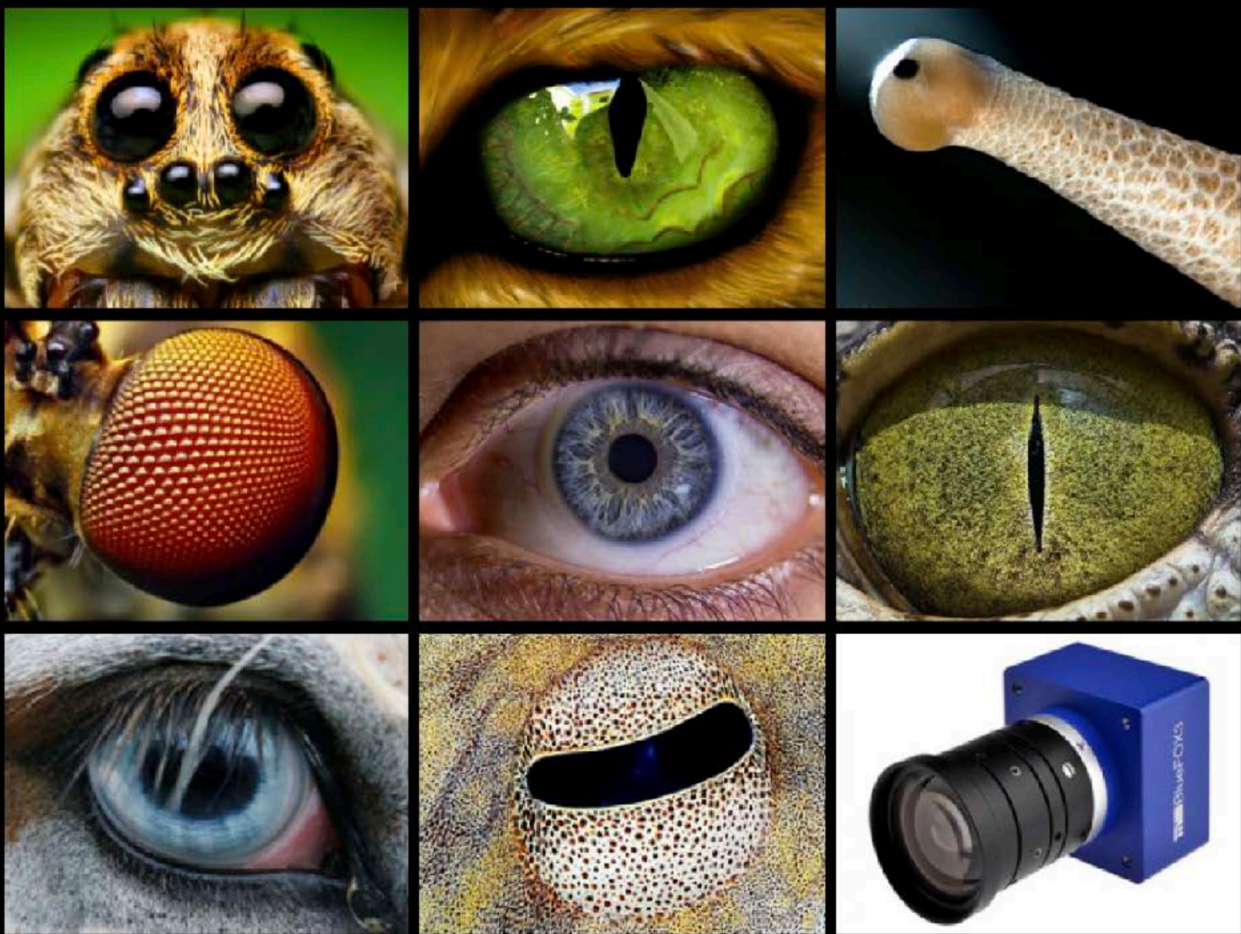
Chapter 1

Vision in Flies, Fish & Folk



Machine Vision

Blank page

More Than Forty Kinds of Vision

Vision is critical for the survival of most species of fish, amphibians, reptiles, birds and mammals. Many simpler animals, including insects, arachnids, crustacea and molluscs possess light-sensing organs that allow them to identify and locate food, shelter, a mate, or a potential predator. Many plants and unicellular organisms, such as amoeba and paramecium, are photo-sensitive, although they cannot base complex behaviour on images. Vision bestows great advantages on any motile organism. Foraging and hunting are both made more efficient by reducing energy requirements. Identifying a crevice that might provide a safe haven, or nesting site, greatly improves the chances of evading a predator. The latter may itself be using vision to hunt. Animals that can detect a predator approaching are able to take appropriate evasive action and reduce the likelihood of becoming merely a tasty snack for some other creature. Many animals rely on vision to identify a potential mate, by looking for specific

Vision is just as important to people as it is to animals; life in a forest and a city would be impossible without it. Human beings have developed highly refined techniques to communicate visually. (You are, I hope, enjoying doing so now!) People dress in a way that indicates mood, social standing and sexual status. Commerce was developed around hand-written and printed documents that are, of course, sensed visually. The process is called reading!

Authors of technical books rely on the reader's ability to understand complex abstract ideas expressed through the visual medium. Electronic visual communication media are

Blank page

ubiquitous. Education also depends on visual information displays: models, white-boards, books, maps, diagrams, photographs, charts, video and computer screens. The leisure, fashion and entertainment industries are built on the premise that vision is the dominant human sense. As further evidence of the great importance of vision, consider its loss: blindness is one of the most feared disabilities.

For many years, people have dreamed of building machines to perform routine vision-based tasks, such as reading. That ambition is now on the process of being realised. In these pages, we will explore the achievements and potential of machine vision systems in detail. We will have to put aside all of our prior concepts about how human beings see. Machine Vision cannot perform some ordinary everyday tasks that we might imagine are trivial. Yet, machines can already do some things that human beings cannot.

Surely it should not be difficult to build a machine that, can, for example, read a page of printed text, or recognise a given face in a crowd? All we have to do is ask a person to tell us how he/she sees and then build a machine, or write a computer program, to perform the same functions. *This is a widely held belief but it is totally wrong!* The experience of numerous researchers, working over many years, has shown, quite conclusively, that introspection about human vision never works nearly as well as most people anticipate. Many newcomers to vision research are totally convinced that they know how a human being recognises, say, an upper-case letter A or numeral 2. However, when tested experimentally, the rules they formulate are invariably unsatisfactory. We will encounter an illustration of this later.

Machine Vision

Blank page

Vision in Nature

In the animal kingdom, there are more than forty fundamentally different paradigms (models) for vision; human vision is just one of them. We will briefly discuss just a few of them here, in order to make two important points:

- Animals often have very different capabilities from human beings.
- Sometimes, there can be great value in a vision system that has a very limited range of functions.

Some of the curiosities mentioned below have parallels in modern vision systems, as will be seen later in this book.

Protozoa

Amoeba is a single-cell organism that exhibits simple sensitivity to light, although it does not have "eyes" and cannot distinguish images. Paramecium can even learn to discriminate between different brightness levels. Again, there is no ability to form and analyse images.

Cnidaria

The notoriously dangerous *Box Jellyfish (Chironex fleckeri)* has a total of twenty-four eyes. These are distributed in four groups of six around the bell-like body. Some of these eyes seem to be capable of forming images. Vision allows Box Jellyfish to avoid obstacles when swimming but it is uncertain how well they can recognise objects. Experiments have shown them to be repelled by red, aware of black but unresponsive to white. They do not possess a central nervous system and the processes involved in processing the visual information are not well understood.

Blank page

Insects

The beautiful compound eyes of the fly or bee are well known and are clearly very different from those in frogs, fish, horses or human beings. **[Figure 1.1]** The compound eye consists of numerous light-sensing organelles called *ommatidia*. Each ommatidium provides the brain with one "point" measurement of light intensity. The brain forms an image from these independent picture elements ("*pixels*"). The number of ommatidia determines the resolution of the eye and can vary from just a handful in some primitive insects to around 30,000 in moths and dragonflies. (A simple calculation shows that these insects are only capable of resolving crude pictures containing about 170x170 pixels. ($\sqrt{30000} = 173.2$) Many insects are *dichromatic*. That is, they are able to distinguish light in two different wave bands. However, honey bees are *trichromatic* and are sensitive to blue, green *ultra-violet (UV)* light.

Figure 1.1 The compound eye of an insect (*Calliphora vomitoria*) consists of numerous light-sensing organelles, called *ommatidia*. Each of these provides the insect brain with one picture element (“pixel”). The number of ommatidia determines the resolution of the eye and can vary from just a handful in some primitive insects to around 30,000 in dragonflies and moths. (A crude calculation shows that this limits the insect's spatial resolution to about

People cannot normally see UV. Bees can distinguish yellow, orange, blue-green, violet, purple, as these are combinations of their three primary colors. Some flowers exploit insects' UV sensitivity and have very distinct markings that are only visible in the ultra-violet wave-band. **[Figure 1.2]** UV imaging enables honeybees to distinguish flowers that to the human eye may look very similar. High-contrast ultra-violet patterns exist on some butterfly wings that look drab to human eyes. Ultraviolet light helps monarch butterflies to navigate in their extraordinary 2000 mile migrations across North America.

Figure 1.2 Insects are sensitive to ultraviolet, which people cannot normally detect. (a) Flower viewed in visible light. (b) Same flower viewed in ultraviolet. A bee would be able to sense the shading on the petals but would not be able to resolve as much detail as we see.

The *stalk-eyed fly* is a spectacular creature found in the jungles of South East Asia and Africa. These flies possess long projections from the sides of the head, with eyes and antennae at the ends. Male flies have the ability to enlarge their eyestalks to impress the females. **[Figure 1.3]**

Figure 1.3 Stalk-eyed fly (*Teleopsis dalmanni*) [Source: Stalk-eyed fly, Wikipedia, *http://en.wikipedia.org/wiki/Stalk-eyed_fly*, accessed 13th October 2011

Spiders

Spiders may have two, four, six or eight eyes. **[Figure 1.4]** Members of one particular family, the *Ogre-faced spiders* (Family *Deinopidae*) have one pair of large eyes, for night vision, and three pairs of much smaller eyes, for seeing in day-light. Their superb night vision is possible because their big eyes can collect a lot of light. Some spiders have more sensitive night vision than cats, or even owls.

Crustaceans

In the larval stage, members of the *Nauplius* genus of *crustaceans* have a single eye at the top of the head. This so-called *parietal eye* has a lens and senses direction of light but cannot form images. A parietal eye is anatomically different from normal eyes and uses a different biochemical method for detecting light.

Lobster, *shrimp* and ***prawn*** eyes form images by reflection, rather than refraction. The surface of this type of eye has numerous tiny square facets, making it look like graph paper. The grid-like appearance is formed by the ends of many tiny square tubes aligned radially within the spherical eye. The side walls within these tubes are flat and shiny, which act like mirrors, guiding light onto the photo-detector array. Their precise geometrical arrangement means that a light beam consisting of numerous parallel light rays is focussed to a smal area. This type of eye contains an effective imaging reflector that allows the animal to see even when it is quite dark. In bright light, the lobster's eye moves opaque pigment so that reflection on the tube side-walls is inhibited. When present, this pigment reduces the intensity of all light rays to the retina, except those parallel to the

Figure 1.4 Jumping Spider with eight eyes. [Source: *Spider, Wikipedia,*
http://en.wikipedia.org/wiki/Spider,
accessed 13th October 2011

tubes. Without it, the sensitivity is greater but the acuity is reduced. This remarkable structure is not found elsewhere in nature!

The *mantis shrimp* is actually a crustacean, from the order *Stomatopoda*. These animals have compound eyes, like those found in insects, although they have a far smaller number of individual light-sensing organelles (ommatidia). In the mantis shrimp, each row of ommatidia has a particular function. For example, some of them detect the presence of light, others colour. Mantis shrimp eyes have 12 types of colour receptors. (Human beings have only three, while dogs have two.) Moreover, mantis shrimp can detect ultraviolet, infrared and the polarisation of light. They can justifiably claim to have one of the most complex eyes of any known animal. The eyes are located at the end of stalks and can be moved independently and rotated by as much as 70˚. Visual information is processed by the eyes, not the brain, as it is in mammals. Each of the mantis shrimp's eyes is divided in three sections, which allows it to see an object simultaneously in three different ways and enables it to judge distance with a single eye. The beautiful *Peacock Mantis Shrimp* is shown in **Figure 1.5**.

Figure 1.5 Peacock Mantis Shrimp (*Odontodactylus scyllarus*) (Source: Mantis Shrimp, Wikipedia, URL *http://en.wikipedia.org/wiki/Mantis_shrimp* accessed 13th October 2011)

One of the most curious eyes found in nature is that of a small crustacean, Copilia Quadarata, which is found in the Mediterranean Sea. **[Figure 1.6]** Only the female possesses the remarkable eyes which occupy more than half of its body. Each eye consists of a single light-sensing organelle. Light is collected by the anterior lens and transmitted by a crystalline conical lens onto a cluster of pigmented light-sensitive cells. This structure oscillates back and forth across the optical axis of the anterior lens. The oscillatory movement is 'sawtooth' in form; the photo-receptors move rapidly towards one another other and then move apart comparatively slowly. This is reminiscent of the early experiments in television which used a mechanical spinning disc (Nipkow Disc) and a single photo-detector.

Molluscs

Land snails have two pairs of tentacles with eyes located at the tips of the longer pair. The animal can move its tentacles independently, up, down, left and right, to adjust its view. When the snail is threatened, the tentacles retract, protecting the eyes from injury. A snail's eyes are only able to provide a crude low-resolution image (approximately 8*8 pixels) of its surroundings.

The *Colossal Squid* (*Architeuthis Dux*) is the largest known invertebrate and has the biggest eyes in nature; each one can be up to 300 mm across. These football-size eyes allow the squid to see in the dim light existing at ocean depths of 2000 metres. These eyes have another amazing feature: each one has a built-in “headlight”, known as a *photophore*, which can produce enough light for the squid to hunt for prey in the dark.

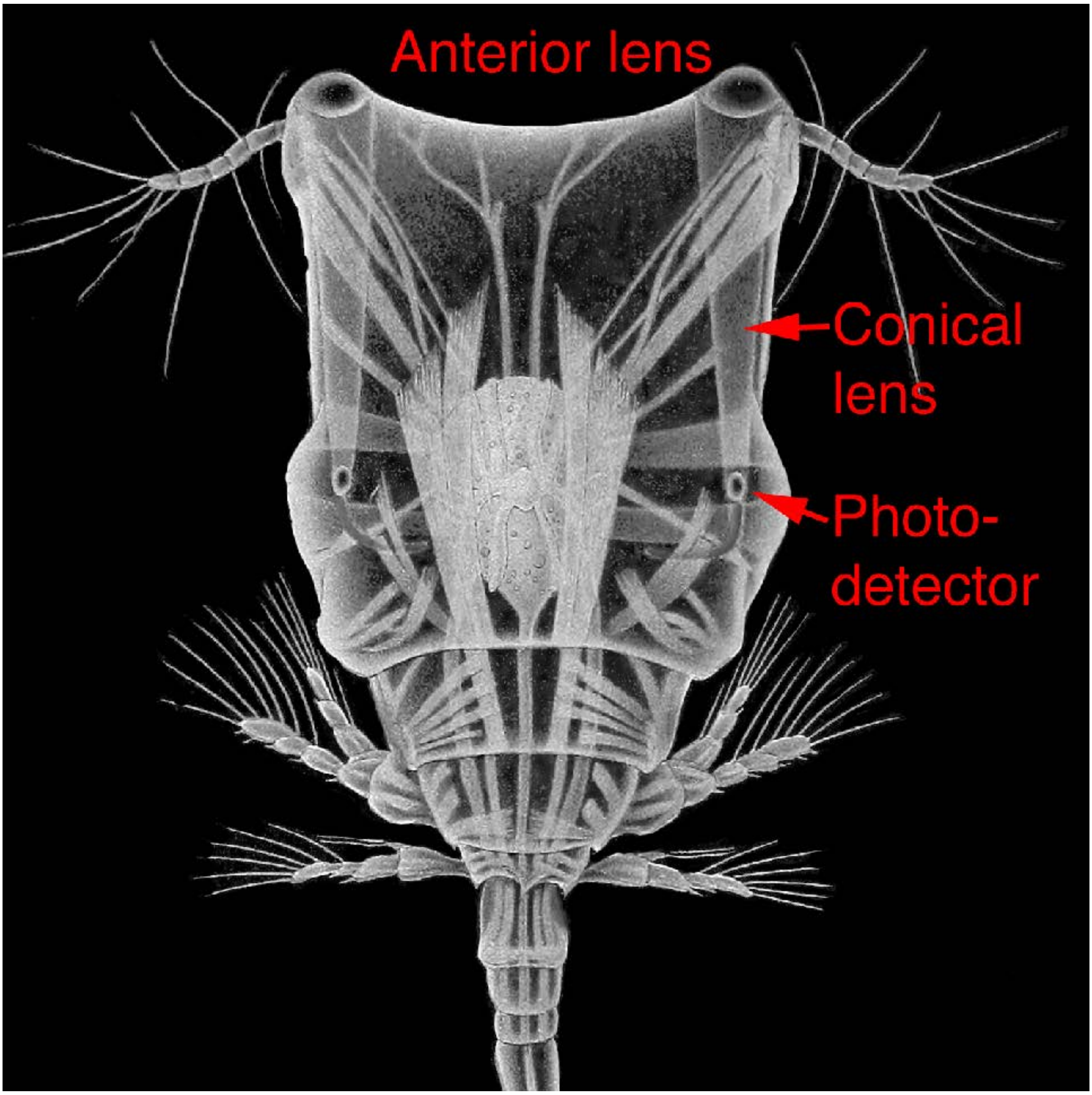


Figure 1.6 Copilia Quadarata (Based on an image found at *http://gallery.obs-vlfr.fr*, *accessed 7th November 2016)*

Fish

The attenuation of light, even in clear water, is far greater than that in air. For this reason, it is not normally necessary, nor possible, for fish to see more than about 50 metres. In muddy, turbulent water, or where there is a dense concentration of micro-organisms, visibility may be only a few centimetres.

The light levels available for fish vision decreases exponentially with depth. At depths of 150 to 750 metres, twilight exists, even in clean, clear water. Below 1,000 metres, there is virtually no light from the sun. Nevertheless, many fish spend some, or all of, their lives at far greater depths. Water absorbs long wavelengths (red) more readily than short wavelengths (blue). Hence, the ability to see the range of colours that human beings experience is impossible below about 100 metres.

Sharks and *rays* do not have the ability to sense colour at all. Most *teleost fish* (possessing bony skeletons) living in shallow water have color vision. *Rainbow Trout* and *Goldfish* can detect *ultra-violet light*, while certain species of *Ciclid fish* can detect prey using infrared light [IR, wavelengths above 780 nm]. It appears that this may also be important for mating, as males and females can be distinguished by their IR markings. Infrared vision is unusual in nature.

The lens of the typical fish eye is spherical and has a high refractive index, compared to other vertebrates. This compensates in part for the fact that the lens operates in water. Since the lens is fixed in its shape, focus control is achieved by muscles which adjust the distance between it and the retina. Although the lens is spherical, the eye itself

Blank page

is elliptical. This means that it can achieve close focus at the front (for hunting/foraging for food), but retain remote focus at the side/behind. This enables it to gain early warning of the threat from a predator. Many fish have a greater monocular visual field than a human being but a smaller binocular field. **[Figure 1.7]**

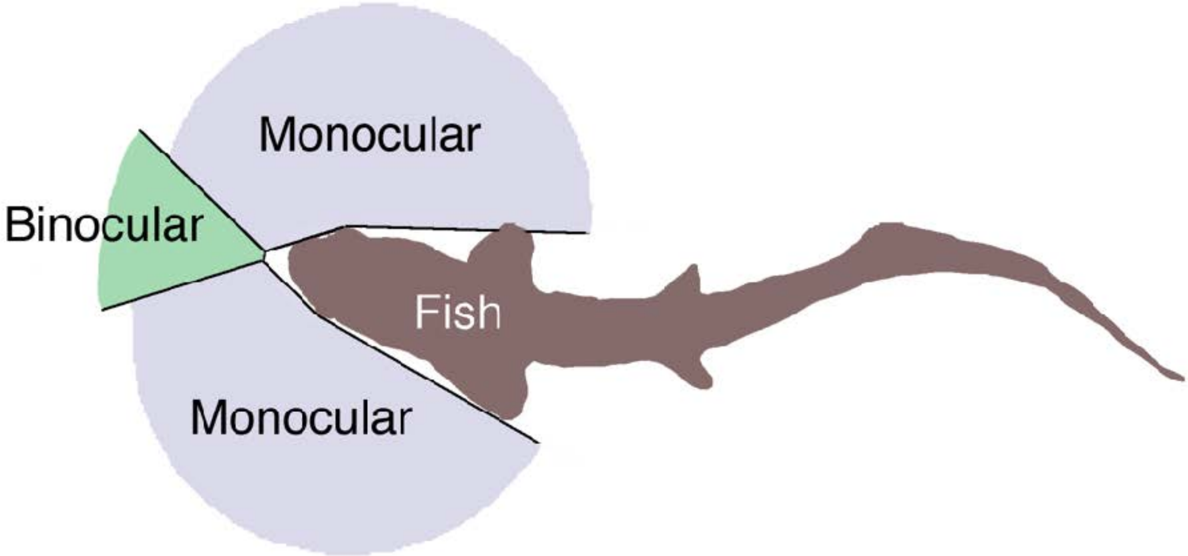


Figure 1.7 A free-swimming fish needs to see forward for foraging or hunting and have a wide field of view to detect predators. Binocular (stereo) vision enables an animal to sense distance.

In the young of *Idiacanthus Fasciola* the eyes are found on the end of stalks projecting from the upper sides of the head. Many of the deepest living fish, and those that live deep in caves have very small and often non-functional eyes. **[Figure 1.8]**

The *Four-eyed Fish (Anableps Tetrophthalmus)* has eyes that rise above its body. Each eye is divided into two parts with the upper part being adapted for seeing in the air and the lower part being adapted for seeing under water (like beetles of the genus *Gyrinus*). Four-eyed Fish feed mostly on insects, so they spend most of their time swimming at the surface. Despite their name, they only have two eyes, each divided by a horizontal band of tissue. The upper part of the eye is adapted for vision in air, while the lower half is adapted for vision under water. Each half of the eye has a separate pupil of its own. Although both parts of the eye use the same lens, its thickness and curvature are different in the upper and lower eye halves. This compensates for the different behavior of light in air and water, allowing the fish to see clearly above and below the water-line simultaneously. Hence, it can look for both prey and predators at the same time.

Some fish have a light-sensitive *pineal gland*. Some *lampreys* have light-sensitive cells in their body skin, even their tails. This enables them to detect local intensity gradients but they cannot form images.

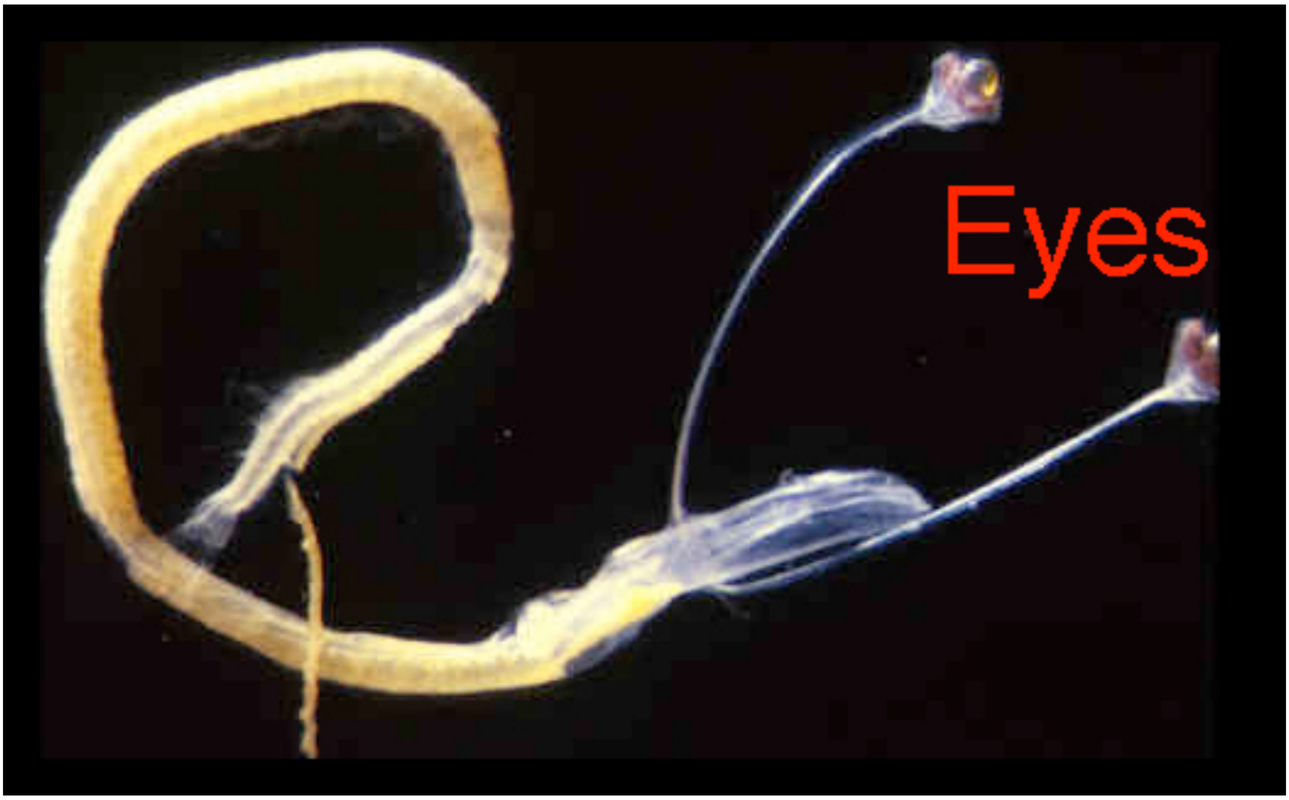


Figure 1.8 Idiacanthus fasciola.
(Source: Black Dragonfish, Featured Creature, URL:
https://featuredcreature.com
accessed 7th November 2016)

Frogs

Most *frogs* have large protruding eyes, giving them almost 360° vision, and are well suited for low-light conditions. While they can differentiate some colours, frogs cannot easily detect red light. They are most responsive to yellow. Frogs are alert to movement in the visual field but have difficulty detecting still prey. Some frogs have *parietal eyes*. **[Figure 1.9]** Most nocturnal frogs have good night vision, enhanced by the presence of a reflective layer behind the retina. Frogs focus by moving the lens relative to the retina. One of the most interesting features of frog vision, is the very limited class of visual events that the animal can detect:

- Local sharp edges and contrast
- The curvature of edge of a dark object
- The movement of edges
- The local dimming produced by movement, or rapid general darkening.

A frog responds to a stimulus consisting of a short line in a way that depends on its orientation with respect to the direction of motion. When a line-stimulus is horizontal and moves horizontally, the frog behaves as if it were a worm or fly. The frog approaches and tries to bite it. When the line-stimulus is vertical and moves horizontally, the frog senses it as a threat. The frog then crouches and puffs up its body. Curvature of the edge of a dark object, combined with appropriate (i.e. horizontal) movement, indicates the presence of another frog, possibly a potential mate.Frogs interprets dark areas as crevices, possible places of sanctuary.A frog's needs are modest: food, refuge, a mate

Figure 1.9 Parietal eye of a frog (*Rana catesbeiana*)
[Source: Parietal eye, Wikipedia, URL
! *http://en.wikipedia.org/wiki/Parietal_eye*
accessed 13th October 2011]

and the ability to avoid predators. Although it is very simple compared to that of a human being, the frog's vision system provides the means to detect all of these. The frog's response to this limited class of stimuli is completely hardwired and does not depend on its experience. This type of stereotypical response to a small set of visual stimuli is exceptional among the higher animals. Most reptiles, birds and mammals have much more sophisticated visual perception controlling behaviour that is modifiable by experience.

Frogs have survived for about 250 million years relying on a vision system with a distinctly limited set of recognisable stimuli and responses. The fact that relatively unsophisticated visual perception can be so useful in practice is an encouragement and inspiration for designers of Machine Vision systems.

Machine
Vision

Blank page

Birds

Good vision is essential for safe flight. **Birds** rely on sight more than any other sense and possess a number of adaptations that enhance their vision beyond the abilities of other vertebrate groups. The shape of the lens can be changed quickly and to a greater extent than is possible in mammals. Birds have the largest eyes relative to their size within the animal kingdom. They have four types of colour receptors in the eye, whereas most mammals have two. (Primates, including human beings have three.) Birds have the ability to perceive visible and ultraviolet light, as well as sense polarisation. The last of these is important during migration in cloudy conditions. Raptors have a very high density of photo-receptors and other adaptations that maximise their visual acuity. The placement of their eyes provides them with good **binocular vision**, which is needed to judge distances. This is, of course, essential for a bird when it is hunting. Nocturnal birds have low numbers of colour detectors, but a high density of more sensitive rod cells. **Terns**, **gulls** and **albatrosses** have eyes that are specially adapted to seeing in hazy conditions. **Owls** can see far better than humans at night and have a broad angular **field of binocular vision** **[Figure 1.10]** On the other hand, **pigeons** have only a small field of binocular vision but almost 330° of **monocular vision**.

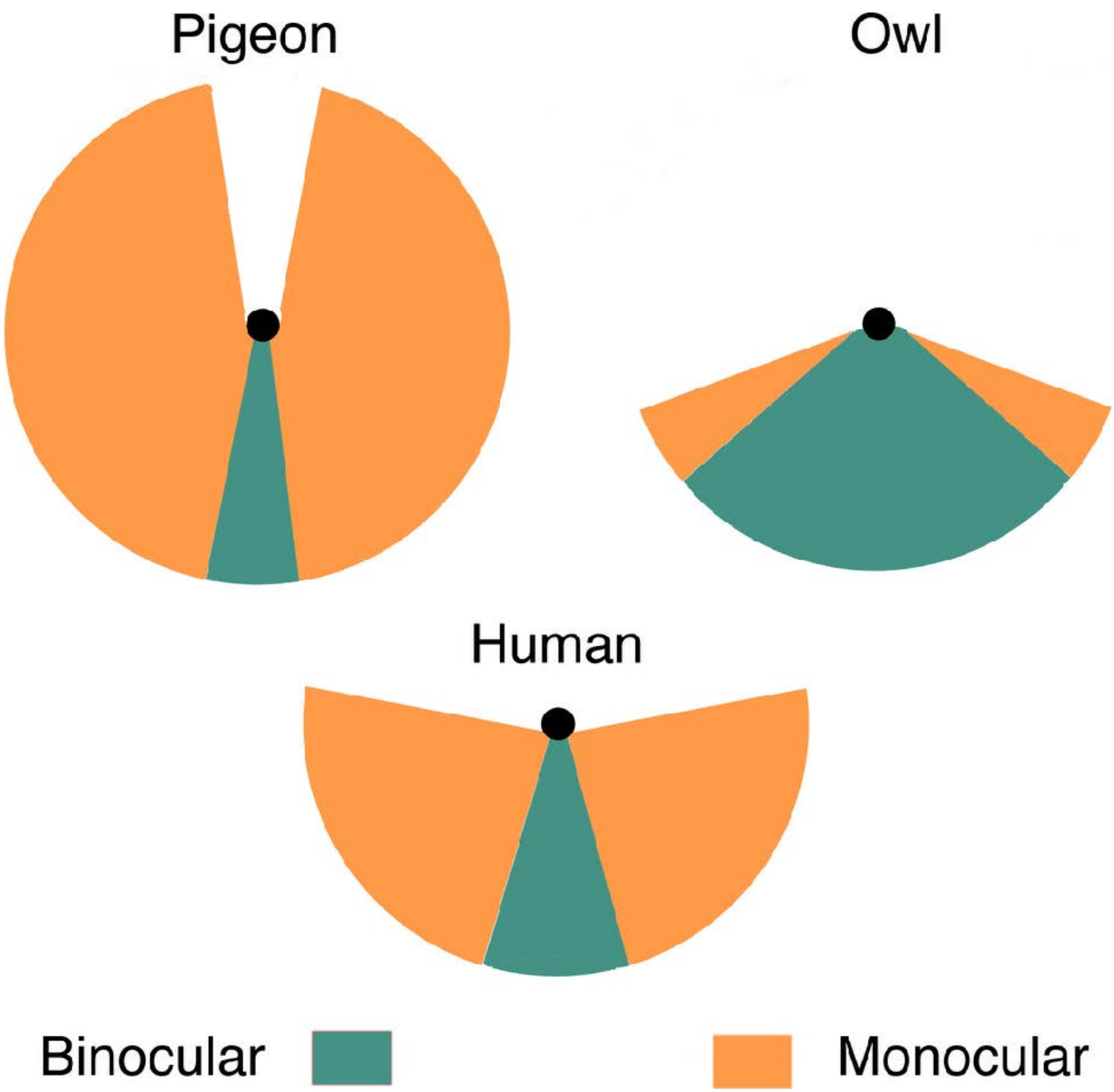


Figure 1.10 Comparing the fields of vision for two birds and human beings. A ground feeder, such as a pigeon, requires a good all-round view, to detect predators. On the other hand, a hunter (owl) has forward binocular stereo vision to sense range.

Reptiles

Since most reptiles are diurnal, their vision is typically well developed and adapted to daylight conditions. In addition to their eyes, some **pit vipers**, **boas** and **pythons** have eyes and additional imaging pit-like sensors that detect **thermal-band infrared** light. **[Figure 1.11]** These pits are sensitive to local temperature differences of a little as 0.2° and allow the **snakes** to hunt for warm-blooded prey at night. **Lizards** and **turtles** can perceive ultraviolet as well as colour.

Nocturnal reptiles usually have smaller eyes than diurnal ones but have relatively large pupils. This improves their ability to operate in **low-light conditions** but reduces **visual acuity**. Some **geckos** can even see colors at night. The **chameleon** is notable for the fact that it can move its eyes independently

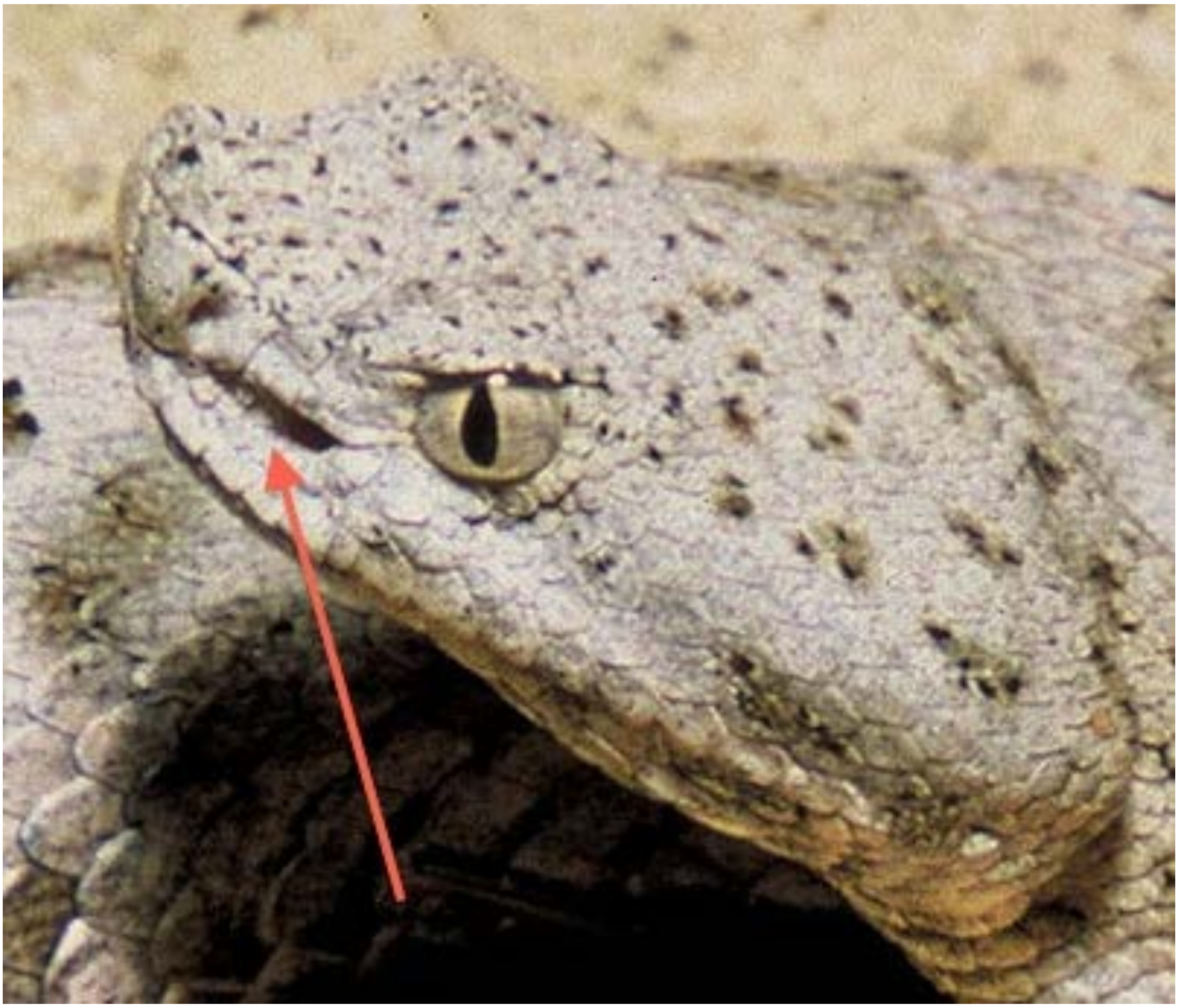


Figure 1.11 Heat sensing organ of the Mexican Ridged Nosed Rattlesnake (*Crotalus willardi*. Source: Wikipedia, URL
http://en.wikipedia.org/wiki/Pit_viper
accessed 15th November 2011)

Mammals

There has, of course, been a great deal of research into the structure and functioning of the human visual system and that of of our closest animal relatives. Many experiments, especially those involving invasive investigation, have been conducted on proxies, most notably **primates** and other mammals. Human vision differs from that of many mammals in being able to sense three different colour components. (**Trichromatic**) **Dogs**, for example, have only two different colour sensors and are therefore said to be **dichromatic**. Some idea of what this implies is evident in **Figure 1.12**.

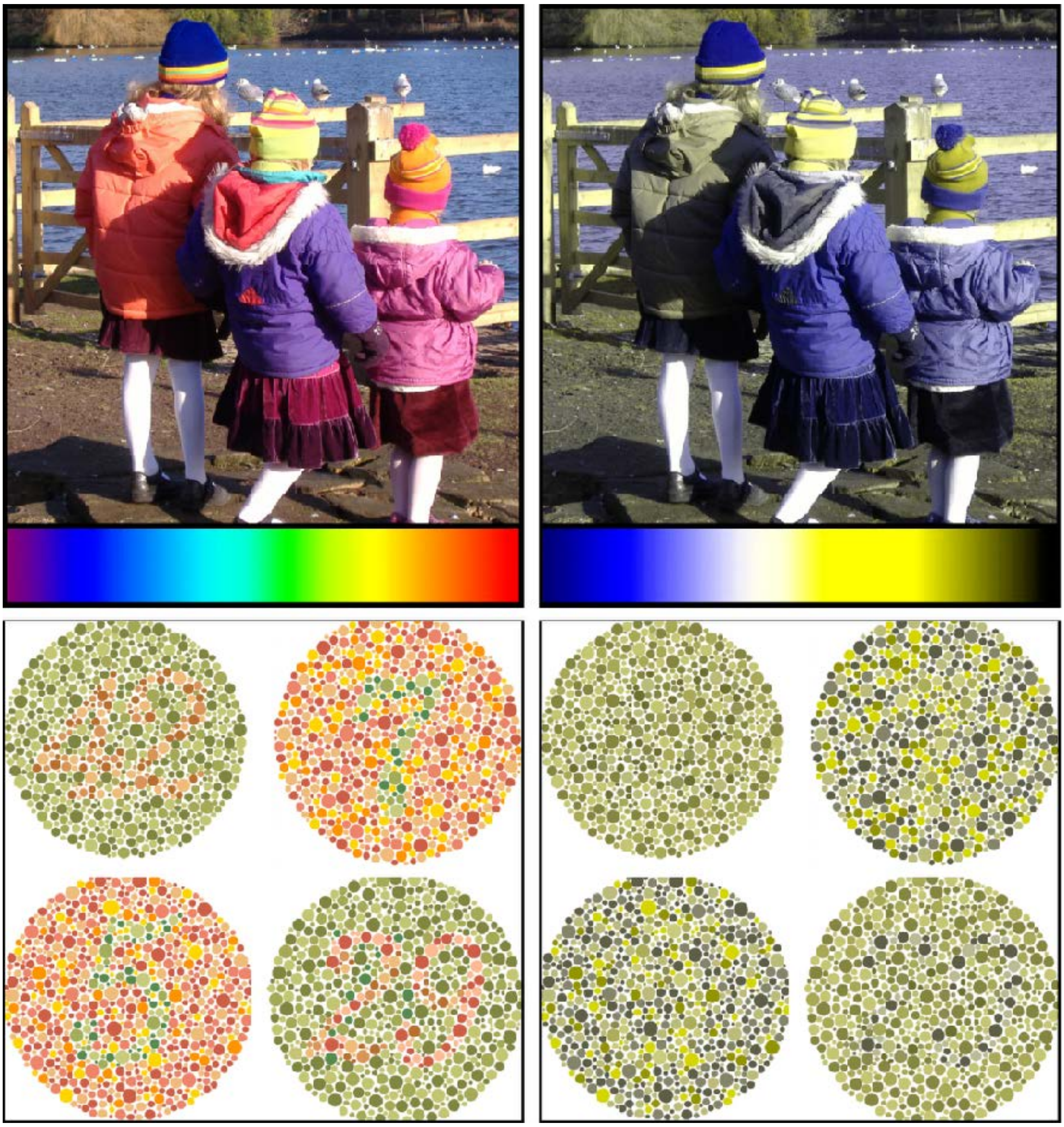


Figure 1.12 Dichromaticity (simulation). *This figure can only be seen properly by people with normal colour vision.* (Left) Normal trichromatic images. The four circular discs form part of the Ishihara colour vision test. (Right) What dogs, and people with red-green colour blindness (Deuteranopia), can see. [This is a gross approximation, which was calculated by setting the R-image to be equal to the G-image. The G- and B-image are unchanged.

Cats are able to see well in low-light conditions that a person finds difficult. Many herbivores, such as the horse, have a wide field of vision, so that they can detect the approach of a would-be predator. Since a horse responds to a threat by running fast, it also needs to be able to see forwards as well. Its eyes are set so that both wide-angle monocular and narrow-angle binocular vision are possible simultaneously. **[Figure 1.13[L]]** The eyes of the horse have a special shape that is well adapted to that of a ground-grazing herbivore. A horse spends much of its life with its head down while it is eating, so it needs to detect predators on the horizon at the same time as looking at the vegetation near its feet. So the horse's eye is asymmetrical **[Figure 1.13[R]]** Arboreal animals, including the primates, require the ability to see in three dimensions, which requires that the visual fields of the two eyes overlap.

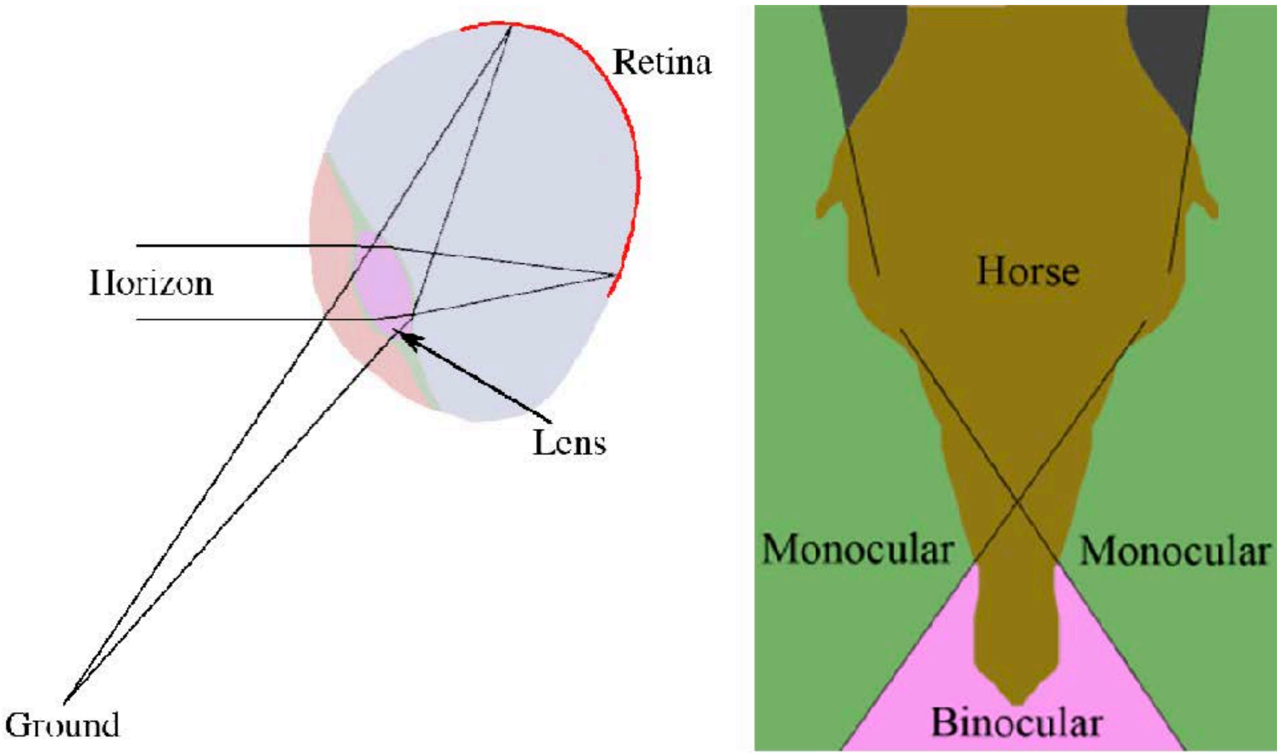


Figure 1.13 Horse. (Left) Aspherical eye is able to focus on horizon and the grass beneath the animal's feet at the same time. (Right) The horse has a wide monocular field view (to sense potential threats in the the distance and binocular vision (to allow it to run safely at speed).

Primates

Human beings are not the only creatures that enjoy visual stimulation. At the age of two, Congo, a male chimpanzee, was given the opportunity to paint and draw. By the age of four, Congo had completed 400 drawings and paintings. When a picture was taken away that he did not consider complete, Congo would scream and remonstrate. Furthermore, once he considered one of his works to be finished, he would refuse to modify or extend it. What is even more remarkable is that Pablo Picasso was a fan of Congo's paintings. Three of them once sold for $26000! Congo was not unique among the Great Apes in his enjoyment of creating and viewing visual art.

Machine

Vision

Blank page

Physiological Studies

However, it is the cat that will complete our brief review of animal vision. In 1959, D. H Hubel and T. Wiesel inserted micro-electrodes into the primary visual cortex of an anesthetized cat. They then projected patterns of light and dark onto a screen in front of the cat. They discovered that some neurons fired rapidly when the cat was presented with lines at a certain angle. Other cells responded to lines at different angles. Some neurons responded differently to light and dark patterns. Other cells detected light-dark transitions (edges) regardless of where they were located in the receptive field. Hubel and Wiesel also found neurons that detect motion in certain specific directions. Their studies showed, for the first time, how the mammalian visual system constructs complex representations of visual information from simple stimulus features. Some of these discoveries have parallels in the processing of images within a computer. At one time, there was hope that this line of research would lead to ideas that could be implemented directly within a computer to give it the same power as human vision. Biomimetic research like this does continue but it has not yielded the same level of insight that was envisaged in the 1960s. Nowadays, non-invasive methods, based on PET scanners, are commonly used for research, following Hubel and Wiesel's lead.

Eyes

Spider	Cat	Snail
Insect	Human	Crocodile
Horse	Octopus	Camera

Human Vision: Features We Need Not Emulate

Research in human vision has advanced rapidly in recent decades but there is still much to be learned. Study of the effects of localised brain injury has provided insight into some of its secrets. Numerous physiological and psychological studies have been devoted to understanding more about how people see. *Visual illusions* (often mistakenly called *optical illusions*) have provided a wealth of evidence too. Of all the fascinating facts that have emerged, one stands out in my mind: if a simple high-contrast image, such as a circle enclosing a cross, is fixed on the retina, the pattern soon begins to disintegrate; quadrants of the circle and arms of the cross may disappear and reappear. This demonstrates that motion of the image relative to the retina is essential for perception. I did not expect that! There are, of course, many other fascinating facts that we could list but we must keep to our central thesis: vision in humans and animals is extremely subtle, very diverse and always likely to surprise. What will become clearer as we progress through the following pages is that biomimetic research has done very little to inspire engineers to build visual machines whose structure, function and performance resemble that of human beings. However, the varied models of vision that exist in animals allow us to accept that machine vision can be different, without being inferior to that in human beings. It is not necessary, or helpful, to judge a visual machine by human standards; we must take a much broader approach and accept that a machine may "see" in a very different way from a person, or indeed any known animal. While many engineered vision systems have parallels in nature, their development usually took place without serious reference to that as a source of inspiration.

Blank page

Human vision has, of course, been studied very extensively, over many years but it would not be helpful to try to review all of the accumulated knowledge about it here. The result of all this work may be summarised in a few words:

- The human visual system is extremely complicated
- Much remains to be learned about human vision; a lot more research is needed before we can claim to understand it properly
- People do not see the world as it really is
- There are many very subtle interacting effects that together to affect our perception of the world
- A person cannot reliably describe how he/she recognises an object.
- Intelligence is an integral and essential part of vision.

Vision research involves a wide range of technical and artistic specialists; physiologists, psychologists, philosophers, physicists, engineers, mathematicians, programmers, artists, photographers, printers, chemists, doctors and linguists have all contributed to our knowledge about vision. Even sculptors and architects have contributed! I shall try to justify these points by drawing attention to a few key experiments and observations. The purpose of this is simply to demolish preconceived ideas about what human vision is and thereby make the novel concepts of artificial vision more acceptable.

Machine

Vision

Blank page

Colour Is a Purely Human Concept

Grass is not really green; it only looks green to an English speaker! There is nothing in grass that says it must be called "green". In some languages, it customary to use colour terms that have no exact equivalent in English. Grass is only perceived as being green by a person when it is illuminated by sunshine, or some other light source with a similar spectrum. In the dark, or when no person is present, grass has no colour at all. (At this point, do not complicate matters by considering deferred viewing via photography or video.) Colour is a purely human concept and does not exist outside the human mind. As a society, we have chosen to associate a certain class of spectra with the name for a colour (e.g. "green"). However, we ultimately rely on the fact that a person (or a group of people) has made this arbitrary choice first. I originally learned the "meanings" of the colour terms "green", "yellow", etc from my parents. Since then, my understanding has been refined/reinforced numerous times, whenever somebody associates the name of a colour with a surface that they and I can both see. Beneath a sodium street lamp at night, grass does not look green. In moonlight, grass appears to be grey. Whatever the viewing conditions, chlorophyl molecules are the principal agents responsible for defining the spectrum of the light reflected from grass. If we had eyes that were sensitive to infrared we would see the grass as appearing very bright in daylight. The colour shading on fruit would even disappear. **[Figure 1.14]** Our unaided eyes cannot assess the infrared or ultraviolet reflectivity of grass. (Also see **Figure 1.1**.) Whatever, we say about the colour of grass, we cannot truthfully claim that we see grass as it really is; our gaze on the world is limited by a narrow sample of the spectrum.

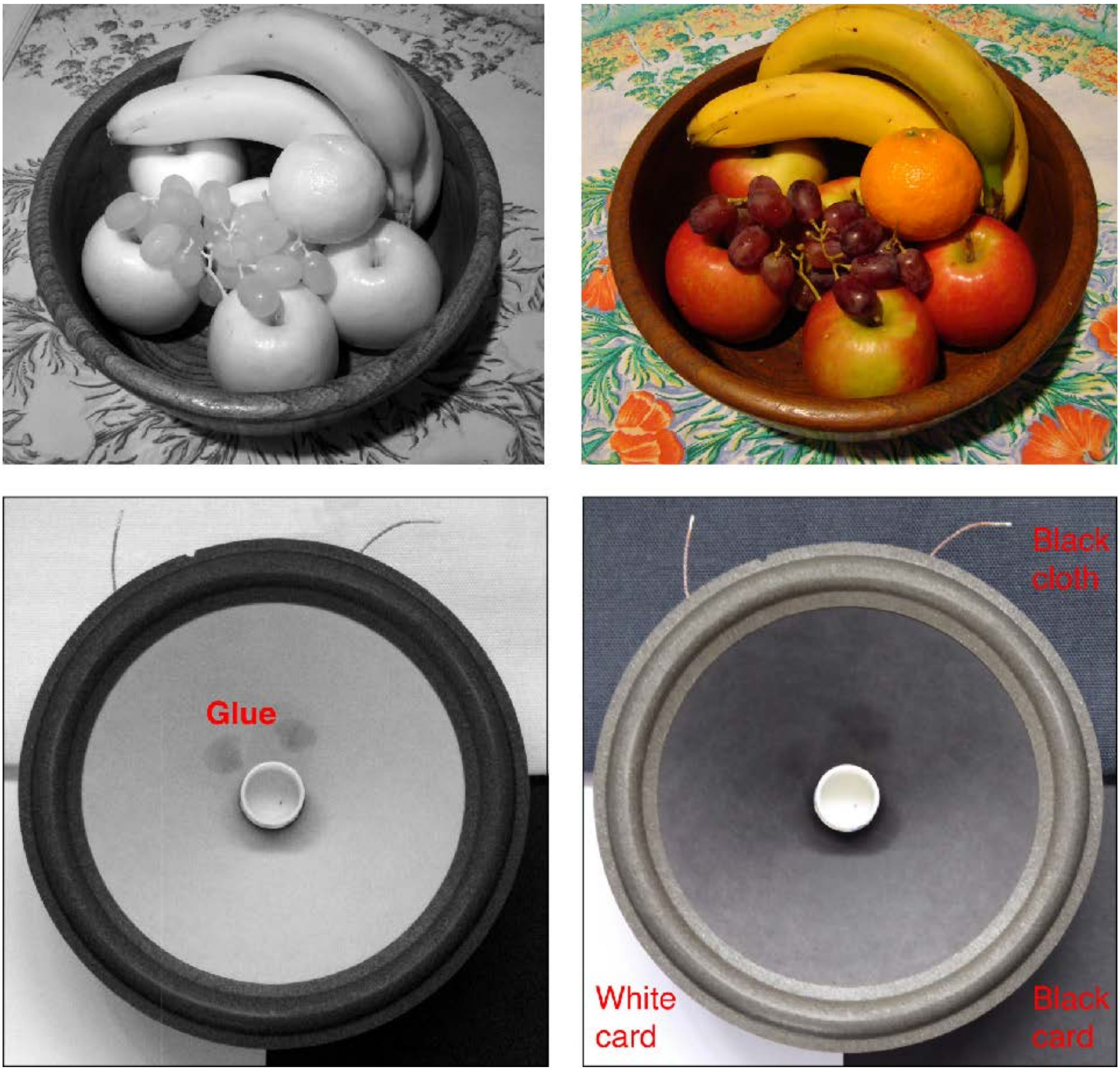


Figure 1.14 People and animals do not sense infrared, even though it has properties that are sometimes invaluable. (Left) Infrared images [The camera is sensitive to light in the 750-1100 nm waveband.] Notice that the colouration of the fruit is obscured by the IR camera, whereas some materials (e.g. black cloth and black card) that are visually similar behave quite differently. (Right) Images obtained from ordinary visible light

Colour Constancy

Another example of seeing the world in a way that does not coincide with reality is found in the phenomenon of colour constancy. Grass is still called "green" if it is viewed at mid-day, or dusk, under a fluorescent or incandescent filament lamp. Although these sources have appreciably different spectra, our brains compensate, giving nearly the same perceptual sensation.

After-Images

After-images give a temporary and local change of perceived colour. This effect is often exploited in visual illusions. **Figure 1.15** is a still taken from a sequence of twelve images in which the pattern of violet spots seems to rotate in steps. When an observer stares at the central cross, the violet dots fade into the grey background and a single fuzzy green spot appears to rotate instead. This is an illustration of the eye creating complementary colour as a negative after-effect. There is no green spot in any of the displayed images. Like many other subtleties of human vision this can be revealed in visual illusions.

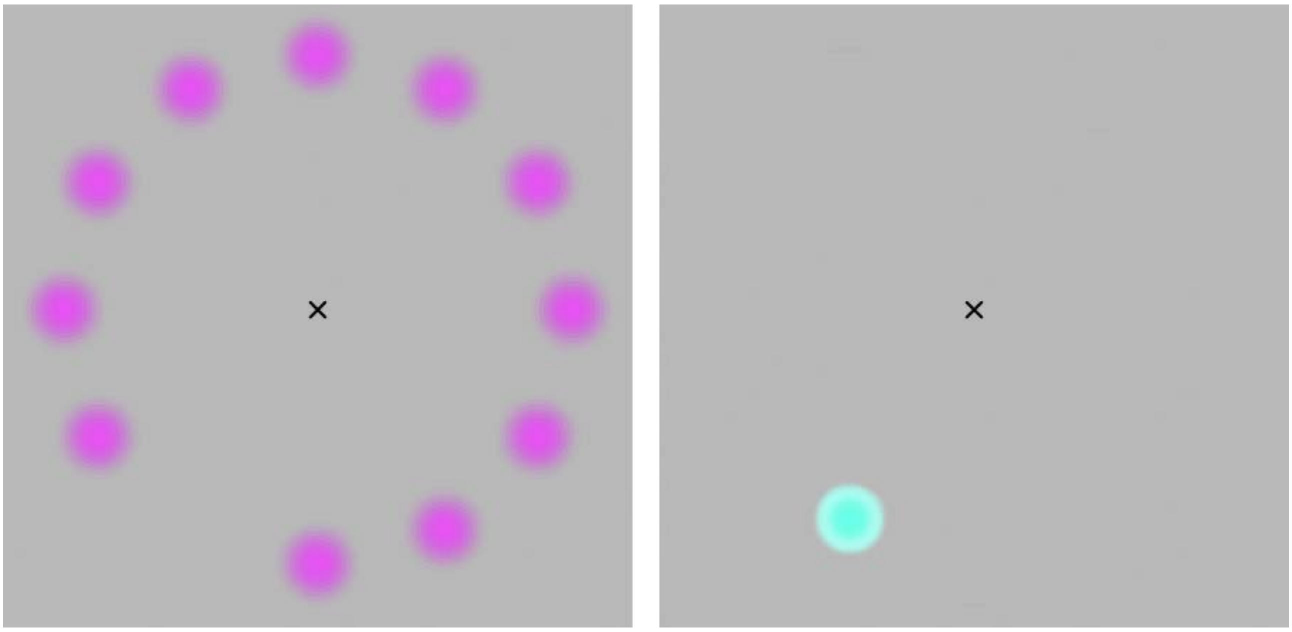


Figure 1.15 Subtleties of human vision are often revealed in visual illusions. (Left) In a web-based animated sequence, eleven fuzzy violet spots are shown, with the “empty space” rotating around the centre of the image. (Right) When an observer stares at the central cross, the violet dots fade into the grey background and a single fuzzy green spot appears to rotate instead. No green spot is displayed. This is an example of **Troxler fading** and is due to adaptation of the eye and the creation of complementary colour as an after-effect.

[To view the illusion visit, visit *Visual Phenomena & Optical Illusions*, Michael Bach, URL
http://www.michaelbach.de/ot/
accessed 26th July 2015]

Naming Colours

Changes in colour perception take place in certain illnesses, such as dementia and stroke. Alcohol and drug use can also cause changes. There is no universally agreed names for colours. For example, my wife and I often disagree about the names of colours, near the blue-green boundary. The difficulty of placing precise limits on what is " green" or "yellow" can be demonstrated in a simple experiment. **[Figure 1.16]**

Past Experience

Perception depends upon past experience. People who work in a world where fine differences of colour really matter, such as dermatologists, gardeners, artists, printers, car-body repairers, etc. often develop a very fine sensitivity to colour that evades the rest of us. The Ford Motor Company alone has used over ten different paint finishes called "silver". When I visited a body repair workshop with my "silver" Ford car recently, the repair technician was able to identify the colour precisely. I cannot even come close to matching that level of colour discrimination

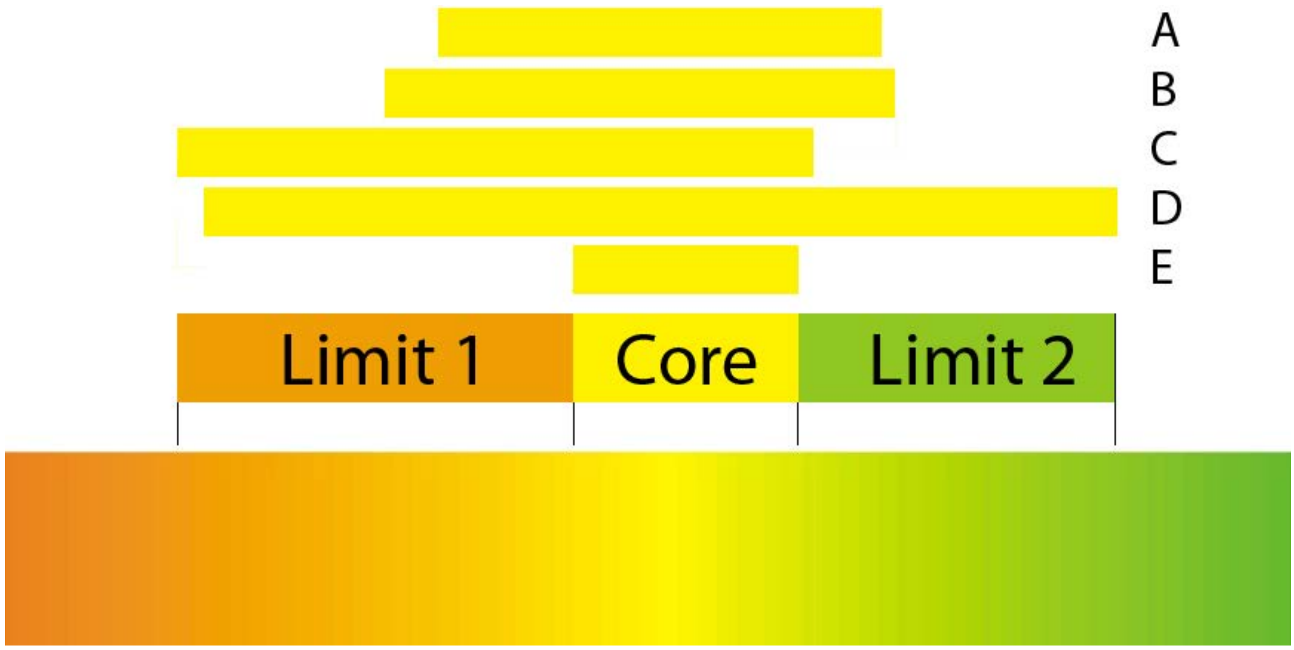


Figure 1.16 Named colour categories are ill-defined. Five people were asked to place limits defining the band of colours that they call *"yellow"* on the orange-green continuum. A - D are adults. E is a 3-year old child. All agree that *“Core”* is yellow but opinions differ in the bands labelled *“Limit1”* and *“Limit 2”*.

Colour Names in Different Languages

Studies of colour names in different languages show that there are subtle differences in nomenclature. I was deceived when I was told at school that *"yellow"* can be translated into French as *"jaune".* [*http://french.about.com/library/begin/bl_colors.htm*] This is an approximation, since there is no exact word in French that corresponds precisely with the English word *"yellow"*. When I was lecturing, I used to tease my Welsh-speaking students by refuting their assertion that *"green"* can be translated is *"gwyrdd"* and *"blue"* as "glas". (I do not speak Welsh.) After a few minutes of verbal jousting, I pointed out that *"glas"* is often used in Welsh to refer to the colour of both (*"green"*) grass and (*"blue"*) sky. **[Figure 1.17]** This bears no reference to the "blue grass" of Kentucky.

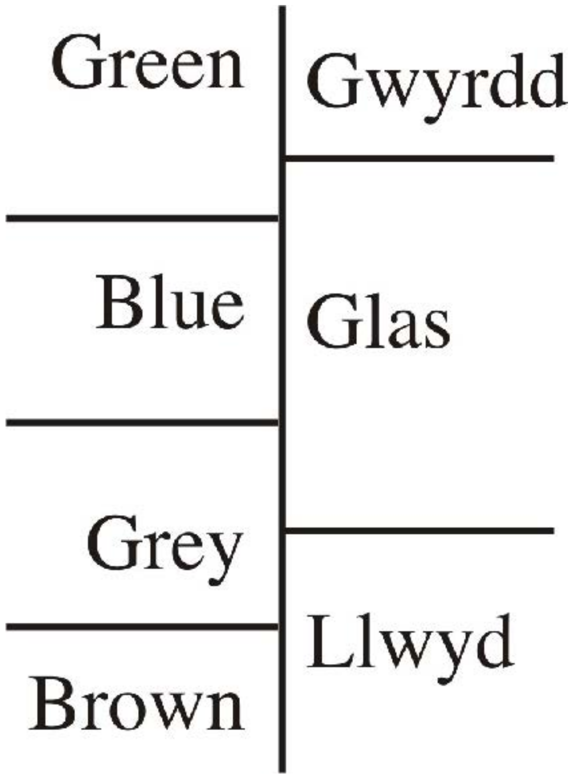


Figure 1.17 Colour terms in English and Welsh do not correspond exactly, despite co-existing next to one another for over a thousand years.

See D. Crystal, *"The Cambridge Encyclopedia of Language",* 2nd ed., Cambridge University Press, Cambridge, UK, 1997, ISBN 0-521-55967-7.

Competing Demands on the Visual System

Figure1.18 demonstrates the interaction of two distinct visual processes: reading and recognising colours. The reader will soon discover how difficult it is to read the text embedded in this image quickly and accurately. The alternative task, identifying the colours and ignoring the meaning of the text, is also difficult.

Figure 1.18 Demonstrating the interference of one visual process with another. First, read the text out loud quickly. Then, perform the reverse operation: identify the colours. How many mistakes did you make? Colour recognition is probably easier if you squint.

Visual Illusions Involving Colour

Numerous visual illusions (often incorrectly called *optical illusions*) reveal a great deal about the complexity of human colour perception. One family of illusions shows that colour perception at one point in a scene can be influenced by the colours present in another part. **[Figure 1.19]**

Colour Created by Motion

During the 1950s, BBC Television experimented with attempts to create a sensation of colour by a rapidly moving black-white pattern. It succeeded in part but the results were unreliable, since one person's perception of colour was not identical to everybody else’s.

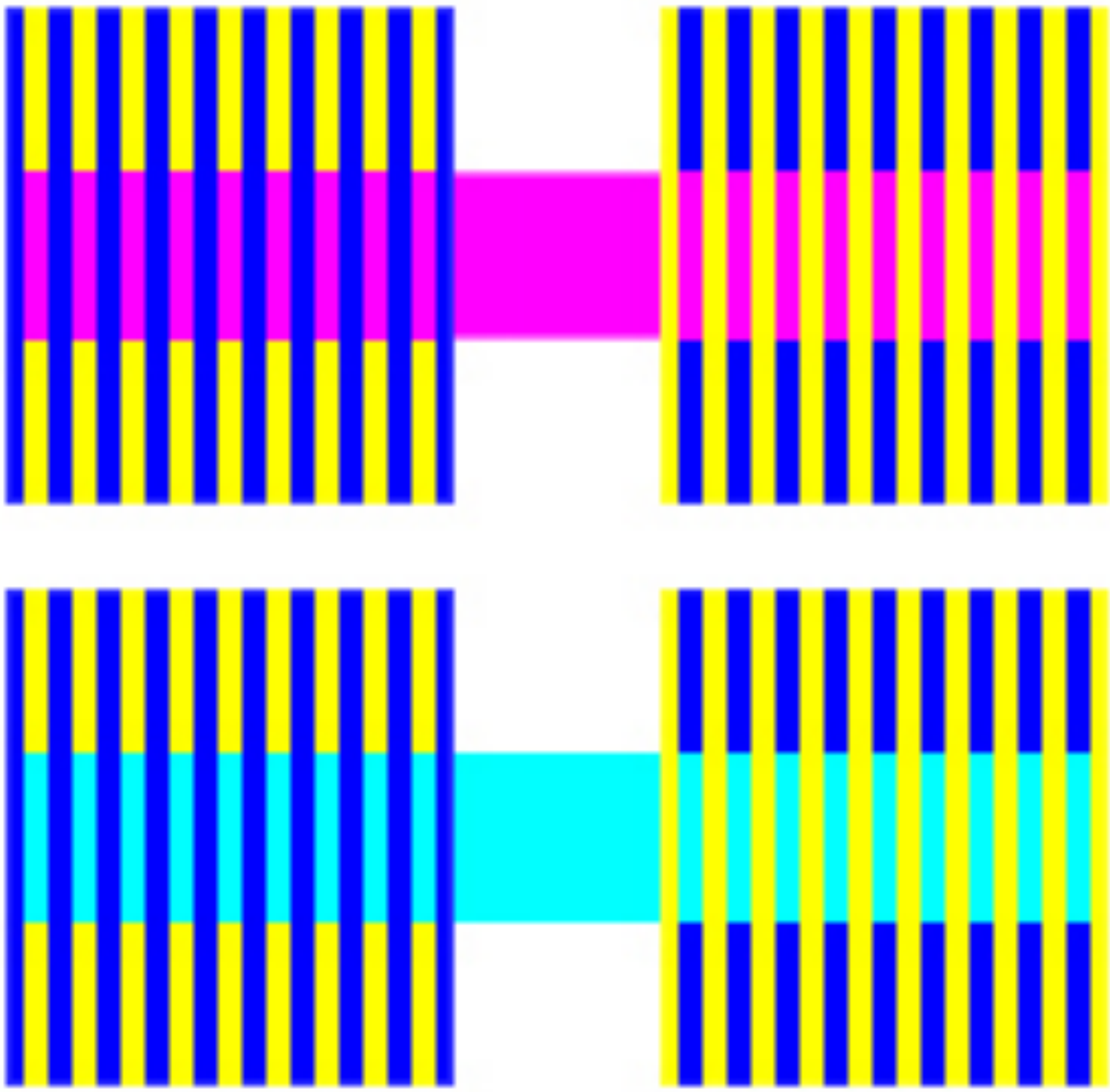


Figure 1.19 Colour perception can be disturbed by the presence of other colours nearby. (Bezold Effect) The two horizontal bands are of constant colour.

Perception of Size

The perception of size is not always accurate, as the famous Müller-Lyer illusion shows **[Figure 1.20[T]]** The equally famous Ames Room **[Figure 1.20[C]]** causes confusion because the visual cues for distance are skewed to make one person appear larger than he/she actually is in relation to the other. Michelangelo's statue of David shows a distorted image of the human body; the head and hands are unusually large. This feature may be deliberate, as the statue was originally intended to be placed on the roofline of a cathedral and viewed from below. Another example of the brain automatically adjusting the perceived size of objects is found in the fact that the Moon appears larger when it is close to the horizon than it does when it is at high elevation.

Warping of Space

Visual illusions that distort space are also well known. An example is shown in **[Figure 1.20[B]]**. This is the so-called and was actually discovered on the tiled wall outside a café in Bristol, England. Many illusions exist in which straight lines appear bent. When I first wore spectacles with variable-focus lenses, I was very much aware of some geometric distortion of my visual field. Now, after several years of wearing these spectacles, the warping has mysteriously disappeared. This is due, of course, to my brain compensating for the optical distortion.

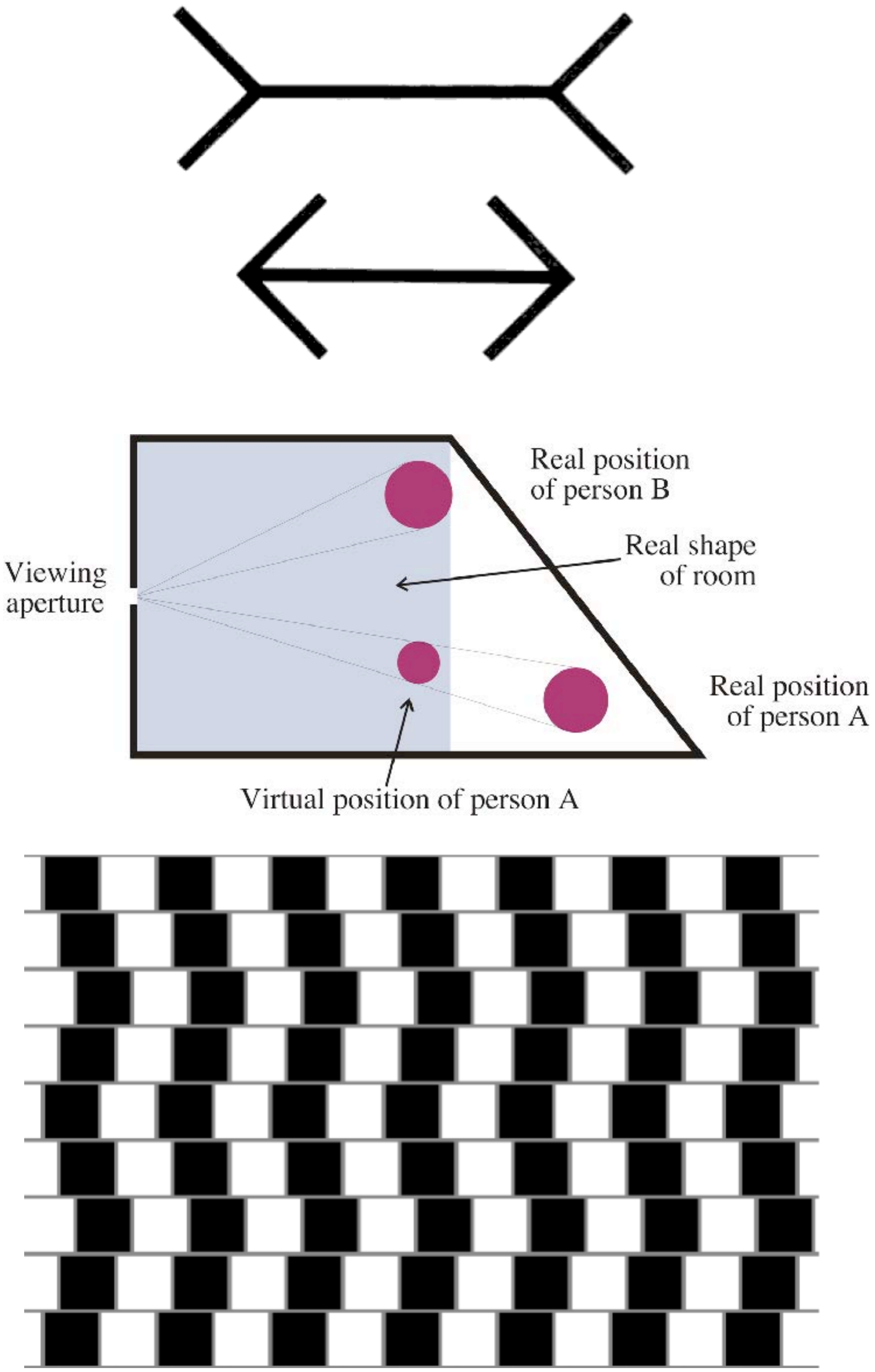


Figure 1.20 Visual illusions. (Top) Müller-Lyer illusion. The horizontal lines are of equal length. (Centre) Ames Room (Bottom) The Café Wall Effect.

Intelligence Is Part of Vision

Intelligence is an essential part of vision. **[Figure 1.21]** No two views of the same face are exactly identical, even if they were taken just a few seconds apart and from the same view point. From this single fact, it is clear that the brain interprets the signals it receives from the eyes. The brain is constantly analysing those signals, adjusting the inferences it draws about the scene being viewed in the light of past experience. This is done sub-consciously automatically and rapidly. Signals take about 100ms delay to travel through the nervous pathways of the eye-brain complex, so the top-levels must predict what the world looks like, based on observations taken a tenth of a second earlier. Playing a fast interactive game, such as tennis, would be impossible without this predictive process.

Identifying a person's sex from an image is, for most people, a very simple, straightforward task but it involves the consolidation of numerous subtle visual cues. Estimating their age, on a coarse scale (child to adult to elderly) can also be performed rapidly and easily. Similarly, judging a person’s mood from facial expressions is another visual processing task that most people are able to perform sub-consciously perform while holding a conversation. These tasks are all done so rapidly that we are unaware of any effort, or delay while the brain processes data from the eyes. These judgements are fundamental to our functioning as social creatures but are far from easy to replicate in a machine. For the type of task we will discuss in relation to Machine Vision we do not need to do so.

Figure 1.21 Intelligence is an essential part of vision.

It is a simple matter to demonstrate the presence of intelligent interpretation of visual signals by presenting situations that lead to conflict between expectation and reality. Visual illusions are intended to do just that. So-called "impossible objects" illustrate the point by showing deficiencies in the inference mechanism relating to our expectation of the nature of our world. **Figure 1.22[T]** shows images that upset this prediction process in the way the brain anticipates the geometry of everyday objects. Many so-called "impossible" objects are not impossible at all; we are simply not used to seeing objects like that. **[Figure 1.22[B]]**

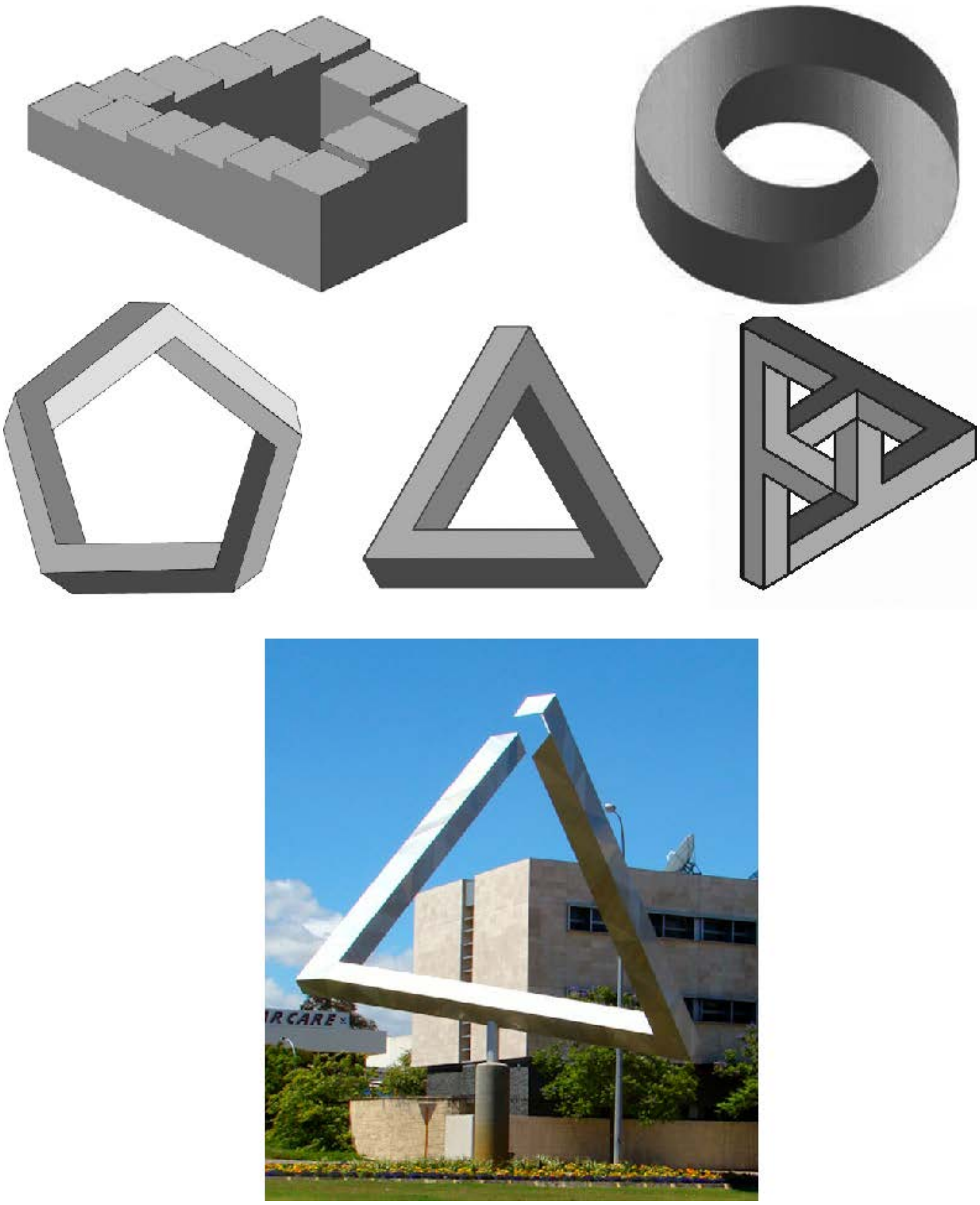


Figure 1.22 Impossible" objects. The brain misinterprets the 2-dimensional projection of a body with an unfamiliar 3-d geometry. The tribar [CC] has been a popular subject for street art. This example was built by Brian MacKay & Ahmad Abas and is in Perth, Australia.

People from cultures where there are no straight lines perceive things differently from those of us who grew up in cities where the buildings have straight vertical walls. If straight lines, edges and corners are unfamiliar, certain visual illusions do not work as they do for people from industrialised nations. For example, some people in Melanesia and members of a small group in India show significantly smaller errors than do their British counterparts in judging the relative lengths of the lines in the Müller-Lyer Illusion. Children of the hunter-gatherers from the Kalahari Desert are completely immune to this illusion. Similarly, “impossible” objects are only perceived as such by people who have certain expectations based on experience. The logic that makes those predictions is part of human intelligence. Very often, our understanding of images depends upon high-level knowledge; in **Figure 1.23**, three examples are shown where image interpretation depends upon geometric shapes, properties of materials, how things are made and English spelling.

Recognition of individual letters is just part of the process of reading and understanding printed text. Very large variations in the form of printed characters can be tolerated as can poor print quality that would vex OCR software. The difficulty of proof-reading accurately is evidence of the fact that our brains can predict what was intended, rather than being disturbed by simple typographical errors. This is ability is so strong that the letters within words can be jumbled without serious loss of understanding. **[Figure 1.23[B]]**

It deosnt mttaer in waht oredr the ltteers in a wrod are. The olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a total mses and you can sitll raed it wouthit a porbelm

Figure 1.23 Images whose interpretation depends upon high-level knowledge. (Top) Exploiting the viewer's familiarity with geometric shapes. At one specific view-point, the disjointed yellow markings on the lock gates and quay-side suddenly take on the appearance of "floating" rings. [Cardiff Bay, Cardiff, Wales, UK] (Centre) This "impossible object" only appears so if the viewer has a high level of understanding of materials and how things are made. (Bottom) This is meaningless to somebody who is not fluent in reading English.

Things People Cannot Do

Human beings cope with previously unseen events by applying their intelligence. In consequence, a person is very often able to answer open questions such as *"What is this?"*

Pictorial quizzes are often based on this by presenting photographs of familiar objects taken from unusual angles. **[Figure 1.24]** On the other hand, Machine Vision systems do not normally attempt to answer this type of question; they are much better at answering precisely defined questions: for example *"Is this a well made widget?"* or *"How long is this widget?"*

This is a crucial difference that affects our whole approach to the subject of Machine Vision and its effects permeate this whole of this essay.

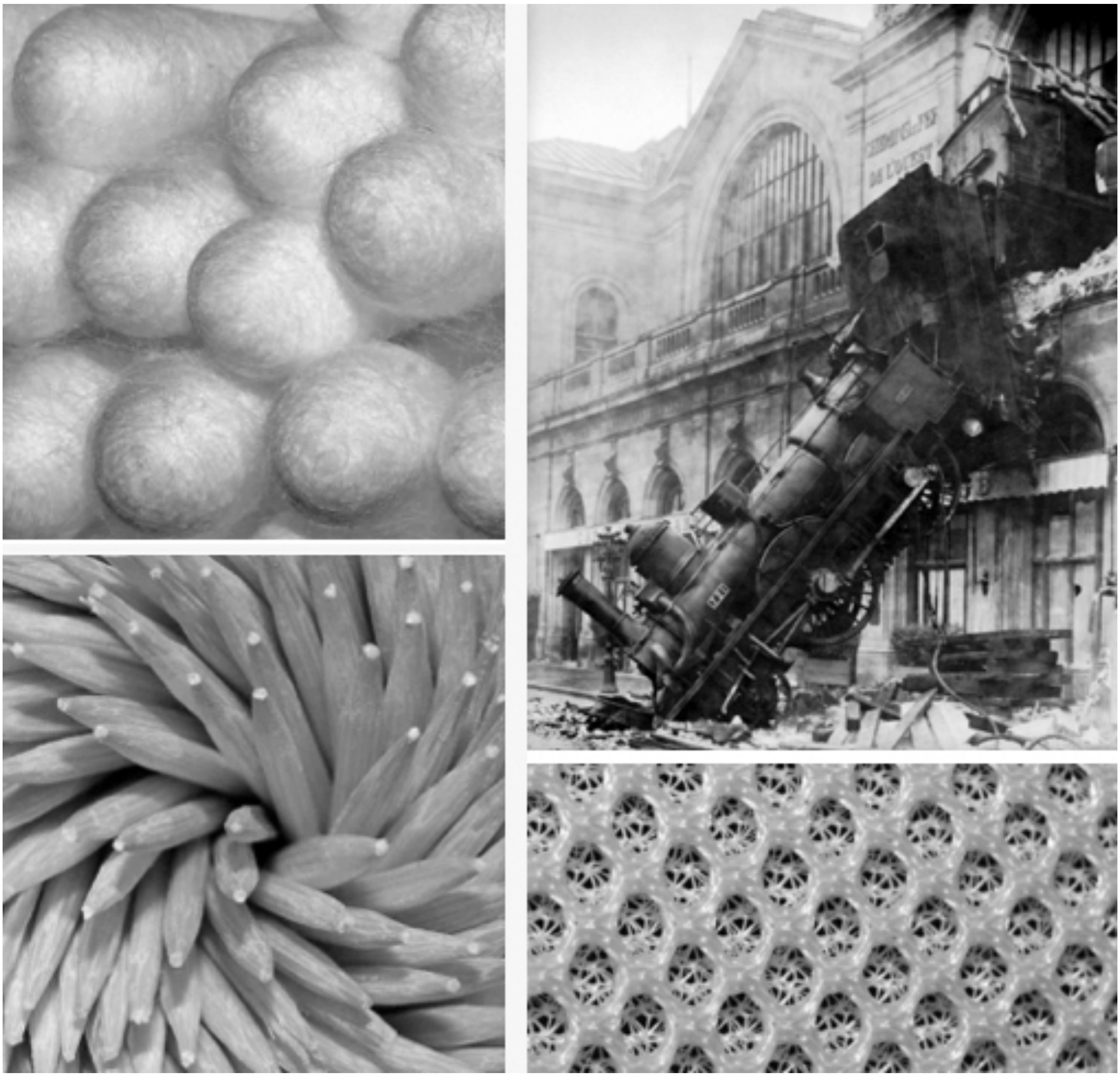


Figure 1.24 The open-ended question “*What is this?”* is popular for pictorial quizzes but has little relevance for Machine Vision. (Top-left) Cotton buds. (Top-right) Rail accident, Montparnasse, France, 1895. (Bottom-left) Wooden tooth-picks. (Bottom-right) Synthetic fabric for sports shoes.

People quickly become bored and tire quickly when they perform routine repetitive tasks, such as inspecting a stream of bottles travelling rapidly and relentlessly along a conveyor belt. Moreover, a human inspector is easily distracted, perhaps by nothing more than a fleeting glimpse of another person walking past. He/she can miss dangerous defects as a result of even a moment's inattention. Bodily discomfort due to poor seating, heat, cold, illness, pain, uncorrected optical deficiencies, alcohol and drugs also impair visual accuracy, as does poor lighting. Flashing lights can cause serious problems: migraine, epilepsy and making moving machinery appear stationary. When a person is asked to perform a task, such as detecting faulty products on a high-speed production line, as many as 30% of defects may escape detection. The performance is even worse if subtle visual judgements are required, or defective product items are rare. Clearly, this cannot be tolerated in situations where safety-critical decisions are to be made. Imagine the consequences of missing 30% of cracks in an air-frame structure, or malignant cancer cells in a pathology laboratory! Another problem is that people, even experts, are inconsistent, both with other people and compared to themselves performing the same task on a different day. The correlation between them may be as low as 50%. Such a difficulty, is not restricted to vision-based tasks.

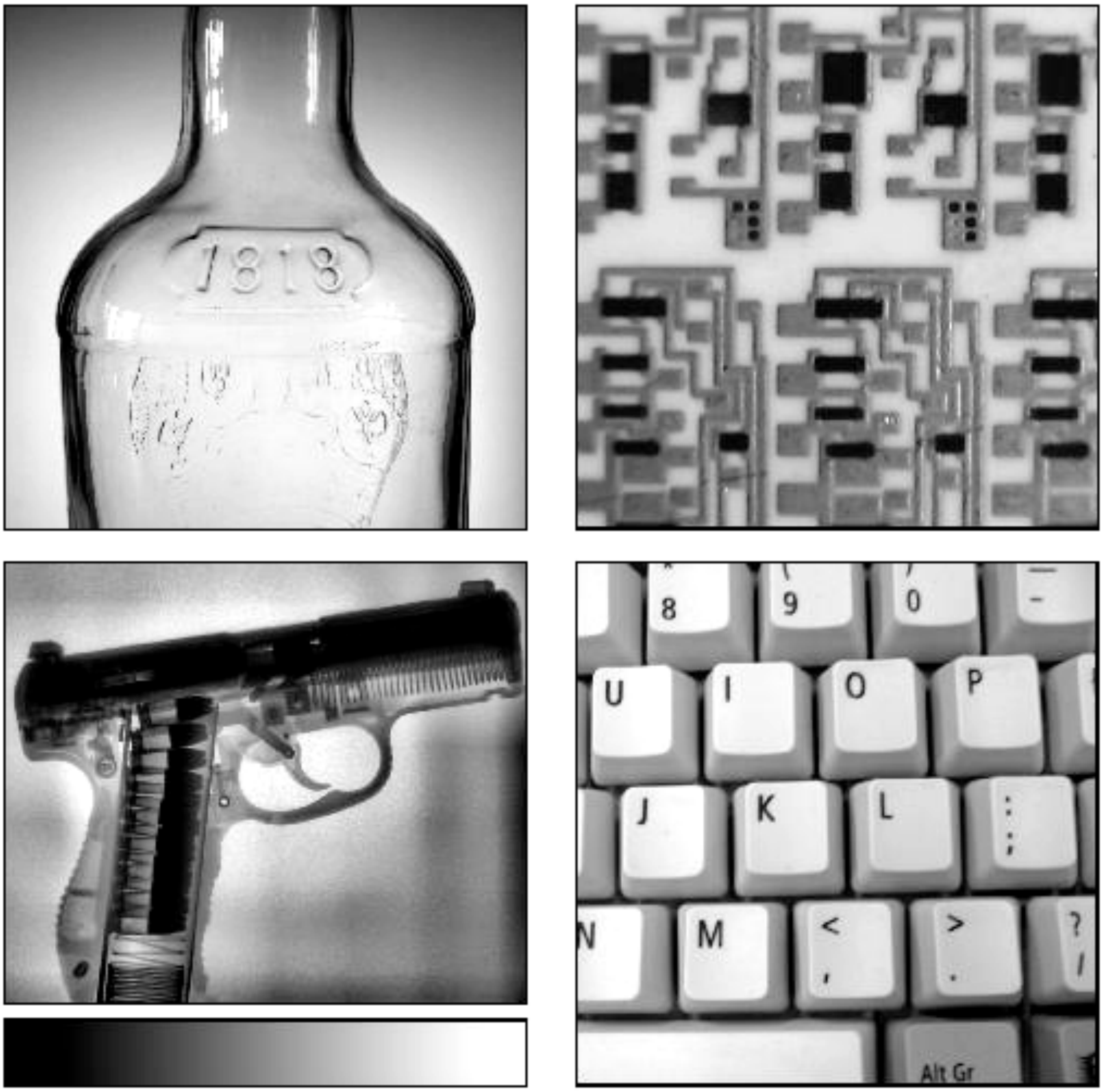
Many decision-making tasks cannot be defined objectively; a judgement can often be made by a person but he/she cannot explain properly how that decision was reached. Examples of this include grading marble tiles and vegetables, inspecting car-body paintwork, etc. There is no mathematical formula, or precisely defined logical rule, that allows a *"good"* cake, or loaf, to be distinguished from an

Blank page

"unacceptable" one. A person knows a good-looking cake when he/she sees one but often cannot reliably explain the criteria that lead him to that decision.

Obvious Ideas Do Not Always Work

There are two visual processes that require special mention here. The first relates to a process, called **thresholding**, which converts a monochrome (grey) image to another, which contains only black and white points. Every point in the former is compared to a fixed brightness value. If that point is brighter, the corresponding point in the resulting image is made white. Otherwise, it is made black. This simple operation seems an obvious way to separate an object from its background. I have had many discussions with students and other newcomers to Machine Vision who are initially utterly convinced that this would provide a sound basis for building an inspection system. So it would if the background were of uniform and constant brightness and there is good local contrast all round the edge of the silhouette. Unfortunately, this is rarely the case, so a machine based solely on fixed-level thresholding will almost certainly be unreliable. [In fact, thresholding is a very useful operation but as part of more complex procedure.] **Movie** shows that thresholding alone fails to provide a satisfactory way to separate the parts of even a good quality images as judged by eye. Initially, most people are totally confident that they know how to perform tasks like thresholding. What seems to be an obvious way to analyse images is rarely so.



Movie Thresholding does not provide a reliable way to segment images.

Another common mistake that newcomers to our subject make is believing that recognising, say, a printed letter ***'a'*** is performed in the human mind by placing a template (a mental model) of an idealised ***'a'*** over the character to be identified and then performing a point-by-point comparison. (This process is called **Template Matching**.) This is rarely successful when the template is based on the whole of a letter. It is most useful when it is combined with other procedures, using a template that resembles part of a letter. Sometimes, debate over the effectiveness of template matching has become quite lively! It is so "obvious" to many people that this is the way to recognise letters that they have some difficulty in accepting the fact that it is not. Following the maxim that a seeing is worth ten thousand tellings, study **Figure 1.25**, where it is clear that no template could properly match all examples of even perfectly formed letters. When they are imperfect, the situation is even worse.

There is a fundamental theoretical objection to template matching, based on the fact that there is a truly gigantic number of possible patterns that can be drawn on an image like a chess-board. Each of these patterns would all have to be drawn, classified, stored and analysed for template matching to be effective. Human life is far too short for learning such templates! Even worse, the Universe is too small and too young for such a training exercise to be possible! Whatever method a person uses to recognise printed letters, it is *not* template matching based on learned the shapes of complete characters.

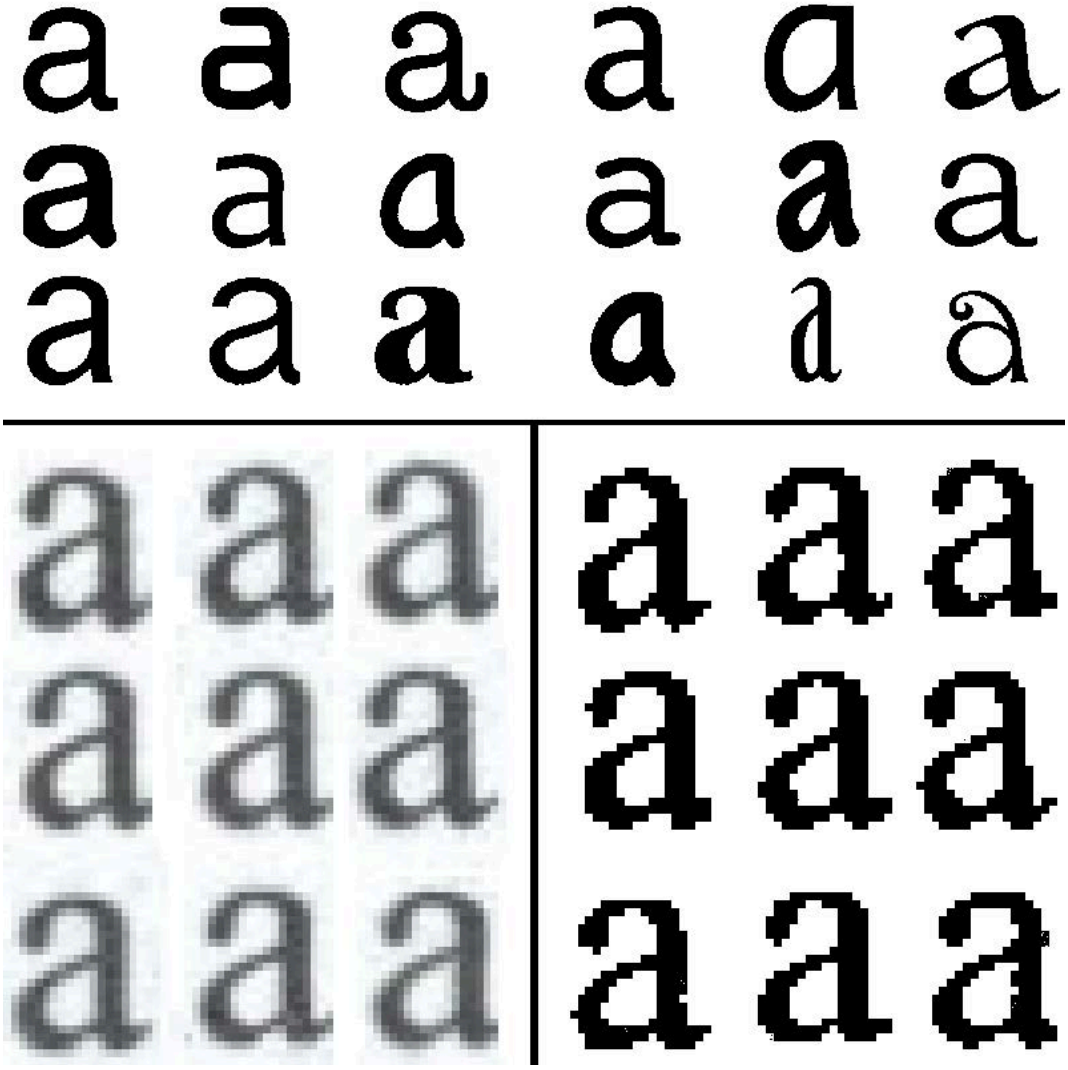


Figure 1.25 (Top) Letter *'a'* in twelve different fonts. Clearly, no template could properly fit all of these. (Bottom-left) Scanned type-set text. This image was obtained using a standard flat-bed scanner. There has been no processing. A person finds this print quite acceptable and comfortable to read, although it appears to be of indifferent quality when magnified. (Bottom-right) Image obtained by thresholding [BL]. The threshold level was optimised manually. Notice that no two samples are the same. Even for a single font, with letters of fixed size and orientation, template matching is inaccurate.

Meret Oppenheim's artistic creation depicting a furry cup, saucer and spoon provokes a potent reaction, even revulsion and nausea. **[Figure 1.26]** The brain analyses the visual signal and, once again, finds a marked difference between its expectations and observations of the visual world. Much of surrealist art and visual illusions are intended to highlight what is essentially a failure of the human visual system to see the world as it really is. Television, photography and colour printing also rely on deceiving the eye-brain complex in various ways. In addition, cartoonists exploit the fact that even a very simple sketch can evoke a powerful response. Virtual reality systems, such as flight and maritime simulators can be made simpler by displaying only those visual features that are needed to mislead the human visual system sufficiently. Even with relatively crude graphics, a simulator can make a person feel nauseous. Why should we want to build a machine that feels queazy at the sight of a furry cup and saucer, even if we could?

Surrealist Art: Furry cup, saucer and spoon, created by Meret Oppenheim

Conclusions

Our subject is Machine Vision, so why did we begin by discussing Natural Vision? The reason is that we needed to establish the scope and variety of abilities of successful vision systems that are different from our own. Machine Vision can be much simpler than human vision, yet still be very effective. Frogs have done rather well, over many millions of years with, what seems to be a very limited form of vision. Insects, spiders, molluscs, fish, reptiles, birds and non-primate mammals all have different types of vision from each other and from us. Again, they are successful, even though they do not achieve exactly what we can do. Nor can human beings accomplish all that animals can do. We tend to think that the world is exactly as we see it. *It is not!* Our brains are always interpreting the signals received by the eye. Machine Vision systems will not deceived by the same visual illusions as we are because they have only a very limited ability to relate what they "see" to past experience. Machine Vision systems cannot yet relate a photograph of a child to the same person in late adult life. They cannot view a pen and realise that it serves a similar purpose to a pencil, or computer printer. Machines are not susceptible to visual illusions in the same way as a person. They are unable to discern "impossible objects". Nor do they feel nauseous when seeing a seriously injured person, or a furry cup and saucer. Machines do not readily compensate for changes in lighting. There is a wide range of applications for Machine Vision where none of these differences from human vision matter. We should not necessarily expect machines to "see" as human beings do. This is not a deficiency, because visual machines can perform some important tasks that people cannot. Human beings cannot (normally) sense x-rays,

Machine

Vision

Blank page

ultraviolet, or infrared light. Machines can! A person cannot accurately measure linear dimensions, surface area, or weigh pigs by looking at them. Machines can! Visual machines do not lose concentration, become tired, or distracted. With good engineering, machines are much more consistent in their judgements than people. They can work at a far greater speed than people and can work in hot, corrosive, toxic and radioactive environments, and where there is a serious biohazard. Machine Vision can complement human vision and should not be expected always to replicate it.

Blank page

References

1. Eye, Wikipedia, *http://en.wikipedia.org/wiki/Eye*, accessed 13th October 2011

2. A. Richards, *"Alien Vision"*, SPIE Press, Bellingham WA, USA, ISBN 0-8194-4142-2, 2001

3. Stalk-eyed Fly, Wikipedia, *http://en.wikipedia.org/ wiki/Stalk-eyed_fly*, accessed 9th August 2011

4. Jumping Spider Eyes, Wikipedia, *http://en.wikipedia.org/wiki/File:Jumping_Spider_Eyes.jpg*, accessed 3rd August 2011

5. Mantis Shrimp, Wikipedia, *http://en.wikipedia.org/* wiki/Mantis_shrimp, accessed 4th August 2011

6. Parietal eye , Wikipedia, *http://en.wikipedia.org/wiki/Parietal_eye*, accessed 3rd August 2011

7. Pit Viper, *Wikipedia, http://en.wikipedia.org/wiki/Pit_viper, accessed 15th November 2011*

8. 92 Visual Phenomena & Optical Illusions, Michael Bach, *http://www.michaelbach.de, acccessed 14th November 2011.*

9. D. Crystal, *"The Cambridge Encyclopedia of Language"*, 2nd ed., Cambridge University Press, Cambridge, UK, 1997, ISBN 0-521-55967-7

10. J. Kingston, *"Visual Illusions"*, TAJ Books Int'l, Cobham, UK, 2009, ISBN 978-1-84406-125-9

11. Café Wall Effect, Wikipedia, *http://en.wikipedia.org/wiki/Café_wall_illusion*, accessed 15th November 2011

12. Perth Impossible Triangle, Wikipedia, *http://en.wikipedia.org/wiki/File:Perth_Impossible_Triangle.jpg*, accessed 15th November 2011

Further Reading

- Apple Snail Website, http://www.applesnail.net/content/anatomy/senses.php, accessed 9th January, 2012.
- Novel technique for making highly efficient biomimetic devices, Raúl J. Martin-Palma, Carlo G. Pantano, and Akhlesh Lakhtakia, 13 October 2008, SPIE Newsroom. DOI: 10.1117/2.1200809.1297
- Bird vision, Wikepedia, *http://en.wikipedia.org/wiki/* Bird_vision, accessed 3rd August 2011
- Chironex fleckeri, Wikipedia, *http://en.wikipedia.org/wiki/Chironex_fleckeri*, accessed, 9th January 2012
- Congo (Chimpanzee), Wikipedia, *http://en.wikipedia.org/wiki/Congo_(chimpanzee)*, accessed 8th August 2011
- D.H. Hubel & T. N. Wiesel, *"Receptive Fields Of Single Neurones In The Cat's Striate Cortex"*, Journal of Physiology, (I959) I48, 574-59I.
- Deinopidae, E. Nieuwenhuys, *http://ednieuw.home.xs4all.nl/australian/Deinopidae/Deinopidae.html*, accessed 9th August 2011
- Engravings by Maurits Escher, M C Escher Official Website, *http://www.mcescher.com/*, accessed 5th October 2011
- Impossible Illusions Made Possible, *http://www.amusingplanet.com/*, accessed 5th October 2011
- Impossible Object, Wikipedia, *http://en.wikipedia.org/wiki/Impossible_object*, accessed 5th October 2011
- Insect colour vision, Don Glass, *http://indianapublicmedia.org/amomentofscience/insect-color-vision/*, accessed 4th August 2011
- Insect eye inspires future vision, Jonathan Fildes, BBC News, 27 April 2006, *http://news.bbc.co.uk/1/hi/sci/tech/4946452.stm*, accessed 3rd August 2011

- Jumping Spider, Wikipedia, *http://en.wikipedia.org/wiki/Jumping_spider*, accessed 13th October 2011
- Landing strips for bees, Don Glass, *http://indianapublicmedia.org/amomentofscience/landing-strips-for-bees/*, accessed 4th August 2011
- Lobster eyes—brilliant geometric design, J Sarfati, *http://creation.com/lobster-eyes-brilliant-geometric-design*, accessed 10th August 2011
- M. C. Escher Official Website, *http://www.mcescher.com/*, accessed 15th November 2011
- M. Oppenheim, *"Fur-covered cup, saucer, and spoon"*, *http://www.moma.org/visit/calendar/exhibitions/960*, accessed 15th November 2011
- Mantis Shrimp, Wikipedia, *http://en.wikipedia.org/wiki/Mantis_shrimp*, accessed 13th October 2011 Stalk-eyed fly, Wikipedia, *http://en.wikipedia.org/wiki/Stalk-eyed_fly*, accessed 13th October 2011
- Molecular genetics and the evolution of ultraviolet vision in vertebrates, Proc. National Academy of USA, *http://www.pnas.org/content/98/20/11731/F5.expansion.html*, accessed 16th November 2011
- O. Reutersvärd and V Alexeev, *http://im-possible.info/english/a*rt/reutersvard/, accessed 5th October 2011
- Paramecium, Wikipedia, *http://en.wikipedia.org/wiki/Paramecium*, accessed 4th August 2011
- Parietal eye, Wikipedia, *http://en.wikipedia.org/wiki/Parietal_eye*, accessed 13th October 2011

- Penrose Triangle, Wikipedia, *http://en.wikipedia.org/wiki/Penrose_triangle*, accessed 5th October 2011
- Perception Lecture Notes: Recognition, D Heeger, *http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/recognition/recognition.html*, accessed 5th October 2011
- R. L. Gregory, H. E. Ross and N. Moray, *"The curious eye of Copilia"*, Nature, 201, 4925, 1166 (1964), An animated view of the oscillating eye of Copilia may be viewed at *http://www.richardgregory.org/papers/copilia/curious-eye-copila.htm*, accessed 12th October 2011
- Sight in Fish, Earthlife, *http://www.earthlife.net/fish/sight.html*, accessed 3rd August 2011
- Spider, Wikipedia, *http://en.wikipedia.org/wiki/Spider*, accessed 13th October 2011
- Spiders, Stefan Sollfors, *http://www.eurospiders.com/*, accessed 3rd August 2011
- Ten animals with incredible eyes, Science & Nature, *http://listverse.com/2010/12/12/10-animals-with-incredible-eyes/*, accessed 4th August 2011
- Ultra Violet Vision, Kimball's Biology Pages, *http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/C/CompoundEye.html*, accessed 3rd August 2011