Neural basis of pattern vision

From the retina to the cerebral cortex, the early stages of the visual system decompose the visual scene into a number of relevant features, while discarding redundant information. Successive stages of the visual system perform an increasingly complex analysis of the visual scene.
Introduction

The main task of the mammalian visual system is to analyse the images projected on the retinas. These images are two-dimensional and contain both informative features and irrelevant or redundant features. "Pattern vision" is devoted to extracting the important information, so that the observer can detect and recognise objects, and ultimately construct an internal, three-dimensional representation of the world.

Pattern vision is achieved by a succession of neural operations performed in a variety of brain structures. These neural operations have been studied extensively in cats and in primates such as macaque monkeys. Studies in these animals have revealed many principles of neuronal function and cortical organisation that most likely apply to humans as well.

Most studies focus on the responses of individual neurones, whose activity can be recorded with microelectrodes. Microelectrodes detect electrical signals called action potentials or "spikes", which are generated by the cell body and transmitted along the cell's axon. The rate of spike occurrence in any given cell is strongly influenced by the images projected on a small portion of the animal's retinas, which is called the cell's "receptive field".

This article reviews how the early stages of the visual system select, encode, and transmit the important features of an image. We focus on the treatment of image features that are particularly relevant for pattern vision, and we do not discuss other aspects of vision, such as the perception of colour or motion.
Basic anatomy of the visual system

The optical apparatus of the eye focuses the visual images onto the retina, where photoreceptors transform light information into electrical signals. Different kinds of photoreceptor absorb light of different wavelengths and intensities, enabling the perception of colour and extending the range of vision to lighting conditions as different as night and day. Through a network of retinal cells, the electrical signals generated by the photoreceptors are transmitted to the retinal ganglion cells. The ganglion cells axons are bundled to form the optic nerve, which carries the spikes of all ganglion cells to the brain. Ganglion cells receiving signals originating in the left visual field send their axons to the right side of the brain, and vice-versa, as shown in Figure 1. After this partial crossover, most of the axons in the optic nerve terminate in the Lateral Geniculate Nucleus (LGN).

The optic nerve carries about one million axons, much fewer than the pixels in a high-quality digital camera. Precise vision is achieved in the face of this limitation by devoting most of the axons to the transmission of information about the center of gaze. Indeed, the concentration of photoreceptors in the retina decreases with the distance from the center, and is by far the highest in a small central region called the fovea. Most animals, including humans, move their eyes continuously to place the center of gaze so that the images of interest fall on the fovea. This arrangement may have resulted from a sort of evolutionary chicken-and-egg problem: to deal with a small optic nerve, one needs a small fovea and a very mobile eye, but to keep the eye mobile one needs a small optic nerve, one that is not too bulky. To have a feeling for how small your fovea is, and how important it is to move your eyes around, try the
following experiment: look at *this* word, and notice how few words you can read around it without moving your eyes.

Neurones in the LGN are arranged in layers, with each layer receiving input from only one eye. Within each layer, the spatial organisation of the optic nerve is preserved (we say that each layer is organised in a "retinotopic" fashion). The resulting map of the retinal image is homogeneous but distorted by the fact that more neurones are devoted to the central regions than to the peripheral regions of the retina.

LGN neurones send their axons to the cerebral cortex, where they terminate in a large area called the primary visual cortex, or V1 (Figure 1). The primary visual cortex, like the LGN, is organised retinotopically, and contains a disproportional representation of the fovea. Like the rest of the cerebral cortex, the primary visual cortex is made of 6 cell layers. The LGN fibres terminate in layer 4, which forms a thick band parallel to the cortical surface. Because this thick band is unique to it, V1 is also known as "striate cortex" (in Latin "stria" means "band").

V1 is the largest visual cortical area and projects to the rest of the cortex. It has numerous connections with other visual cortical areas, grouped under the name of "extrastriate" visual areas. In primates, over 30 extrastriate visual areas have been discovered. The visual cortex also has numerous connections with other cortical regions that are not specifically devoted to the analysis of visual information. It is through these connections that the visual system communicates with other sensory modalities and participates in the elaboration of motor commands and, well, all the other things that brains do.
Functional aspects of pattern vision

The receptive field of a visual cell is defined as that region of the visual field where the turning on, or off of a light can elicit a change in the cell's response. This change manifests itself either by an increase, or a decrease in the frequency of the action potentials generated by the cell. The study of the receptive field properties of visual neurones has proven crucial to our understanding of visual function, and provides the basis for the current knowledge on pattern vision. In particular, receptive fields can be seen as filters through which the cells view the visual scenes, and filtering is a well-known concept in engineering and computer science.

Image filtering in the retina and the thalamus

The successive operations performed on an image by the initial stages of the visual pathways are illustrated in Figure 2. The visual stimulus (A) is a bit dark because it was obtained in the evening, when there are few photons available. This stimulus is analysed by successive stages in the visual pathway, starting with the photoreceptors in the retina (B).

Photoreceptor cells absorb photons and generate electrical signals. Their receptive field (B, bottom) is small and positive, that is, increasing light on it increases the response of the cell. Because light intensity can vary over many orders of magnitude, however, the photoreceptors need to adjust their gain according to the prevalent light conditions. This gain adjustment is called "light adaptation", and results from a series of chemical events that maintain the photoreceptor in a roughly constant, ready state. In bright light, the gain is turned all the way down, and a given electrical response may require millions of photons. In dim light, the gain is turned up, and the same
electrical response may be obtained with just a few photons. Thus, despite enormous variations in illumination levels during the day, visual signals can be transmitted to the rest of the retinal network using a limited range of available responses. We are mostly unaware of this mechanism, but can become aware of it when we step from a dark room into bright daylight, or vice versa. After a few moments of blindness, vision is progressively restored.

Thanks to light adaptation, rather than signalling absolute light intensity, photoreceptors encode stimulus contrast, i.e. whether the light on their receptive field is stronger or weaker than it has been in the recent past in that local region of the retina. This action of photoreceptors can be observed in the top portion of Figure 2B, which illustrates the "neural image", a map of the responses that would be obtained if an array of identical cells were looking at the physical image in all possible positions. In a neural image, grey indicates no response, whereas white and black indicate higher and lower responses. Even though the image Figure 2A is obtained in a dark room, the neural image obtained by the photoreceptors shown in Figure 2B uses the full available range of response intensities.

The output of the retina is given by ganglion cells. Ganglion cells have a concentric receptive field made of a roughly circular centre and an antagonistic surround (Figure 2C, bottom). The cell in our example is an "ON-centre" cell, meaning that it has a positive centre and a negative surround. Such a cell will increase its activity when a light spot is turned on at in the centre of its receptive field, but decrease its activity if the spot is in the surround region just outside the centre. There are also "OFF-centre" cells, which are organised in the opposite way. In both kinds of cells, if the spot of light is enlarged to simultaneously cover both the centre and surround, the resulting enhancement and suppression will partially cancel out. Thus, ganglion cells do not
respond well to changes in uniform illumination, but respond vigorously if there are differences between the light intensity falling in the centre and that in the surround.

The neural image obtained with ON-centre retinal ganglion cells is illustrated in Figure 2C, top. In the locations corresponding to the cheek or to the dark region behind the face, the intensity of the stimulus is rather uniform, and the cells do not give any response (the neural image is grey). By contrast, in the locations corresponding to features such as the edges of the face, the mouth and the eyes, the cells give strong negative or positive responses. For example, the cells that are stimulated with the mole just to the right of the nose (a "beauty mark" according to the French) give strong negative responses: the mole is dark, and the surrounding skin is light, the opposite of what gives a positive response in an ON-centre cell.

From the retina, the axons of ganglion cells contact the neurones in the LGN. This brain structure seems mostly concerned with gating the visual signals so as to control arousal and other physiological states. The image processing operations performed in the LGN, however, do not distinguish themselves significantly from those of the retina. Indeed, most LGN neurones have receptive fields that resemble those of their ganglion cell afferents. They have a concentric, centre-surround organisation, and are either ON- or OFF-centre. The signals that the LGN sends to the cortex, as a result, are not radically different from those that it receives from the retina.

**Image filtering in the primary visual cortex**

Most V1 cells show receptive field properties that are absent at earlier stages. As revealed by the Nobel Prize winning work of Torsten Wiesel and David Hubel in the 1950s, cells in the primary visual cortex are extremely selective for stimulus
orientation. Some of these cells are called "simple", and have a receptive field that is made of distinct elongated ON and OFF regions (Figure 2D, bottom). Others are called "complex", and have a receptive field where the elongated ON and OFF regions are superimposed (Figure 2E, bottom). Stimuli that elicit strong responses in V1 cells are bars and edges. Thanks to the elongation of the regions and the alternation of ON and OFF sub-regions, V1 cells can be extremely selective for the orientation of such stimuli. By contrast, the symmetry in the receptive fields of retinal ganglion cells causes their responses to be independent of stimulus orientation.

The effects of orientation selectivity can be observed in the neural image obtained with our example of simple cell (Figure 2D). The cell's receptive field is composed of four horizontal regions, two stronger central ones flanked by two weaker ones on the top and on the bottom. The optimal stimulus to obtain a strong positive response is one that stimulates the ON regions with positive intensity while stimulating the OFF regions with negative intensity. In our example, this is precisely what happens near the mouth, the eye, and the eyebrow. By contrast, the edge of the face is mostly vertical, so it stimulates the ON regions and the OFF regions at the same time. The inputs from these regions cancel out, and the resulting response of the cell is zero, which is indicated by grey in the neural image.

Simple cells are sensitive to the precise location of a stimulus within their receptive field. For example, the negative response evoked by the eyebrow is flanked by two strong responses, evoked by the regions above it and below it. Moreover, simple cells are sensitive to the "polarity" of a stimulus: a simple cell's response differs depending on whether a given stimulus is brighter or darker than the background.
This dependence on stimulus location and polarity is overcome by complex cells. A complex cell gives roughly the same response to a bar presented anywhere within its receptive field, whether it is a light bar or a dark bar. As a result, the neural image produced by complex cells (Figure 2E, top) is characterised by only positive responses, here corresponding to the location of the mouth, eye, and eyebrow.

Complex cells can be thought of as summing the positive outputs of many simple cells whose receptive fields are elongated in the same direction, and whose spatial location is displaced. For example, the complex cell receptive field at the bottom of Figure 2E could be obtained by the summing the outputs of a few simple cells with receptive fields as in Figure 2D. The receptive fields would be spatially displaced so that the sum of their positive outputs is insensitive to the precise spatial location and polarity of the stimulus.

While in our example we have chosen cells that are selective for horizontal orientations, the primary visual cortex contains cells tuned for the full range of orientations. In addition, V1 cells are also tuned for spatial scale. An example of this variety is illustrated in Figure 3, which shows the receptive fields of four simple cells and the resulting neural images. The two cells in the left quadrants are selective for horizontal orientations, whereas the two cells in the right quadrants are selective for vertical orientations. The two cells in the top quadrants are selective for lower spatial scales than the two cells at the bottom. From the neural images (Figure 3C), it is evident that these cells respond to very different features in the visual stimulus. The cell in the top right gives strong negative responses to the edge of the face, where the ON region is strongly stimulated by the dark background, and the OFF region is stimulated by the lighter skin. None of the other cells respond well to these features.
By contrast, the cell in the lower right quadrant responds strongly to features such as the mouth and eyebrows, which are missed by the other cells.

Taken together, the various cell properties described above allow us to imagine the role of V1 in pattern vision. When an observer sees the profile shown in Figure 3A, orientation sensitive simple and complex cells whose receptive fields overlap the edges of the face will give a strong response. Those centred on the background or on uniform parts of the face will remain mostly silent. Thus, the face's edges and the contours of the face's various features will be extracted from the visual image. Simple cells will provide an accurate description of the location of the edges, as well as the indication of their polarity. For example, they will indicate that the face is brighter than the background. Finally, note that the invariable presence of small eye movements, which displace the edges within the cells' receptive fields, will induce different populations of simple cells to become successively active. The population of complex cells, on the other hand, will continue to signal the edges without interruption, allowing the observer to recognise that the same, unique face is continuously present. In summary, cortical cells could be described as feature detectors: they extract relevant elements of the visual image (lines, edges, corners, etc.). The combined activity of simple and complex cells could yield a stable, faithful representation of the objects present in the environment.

**Functional architecture of the primary visual cortex**

The orientation preference of individual V1 cells can be determined by recording their electrical activity with microelectrodes, while the eyes are stimulated with bars of varying orientation. If the electrode is advanced into the cortex perpendicular to the cortical surface, the cells encountered share the same preferred orientation. If the
microelectrode penetration is oblique to the surface, the cells' preferred orientation show a smooth progression, eventually covering the whole range of possible orientations.

**Maps of orientation preference and ocular dominance**
The smooth progression of preferred orientations indicates that stimulus orientation is represented orderly across the surface of V1. Indeed, this orderly distribution can be admired in maps such as that in Figure 4A, which was obtained by recording the activity of several square millimetres of cortex. Rather than with microelectrodes, which record from few cells at a time, such a map was obtained with optical means, taking advantage of changes in the reflectance of the brain regions where the neurons produce spikes. The colour code in the map indicates the preferred orientations of the cells in a region: For example, red indicates regions where cells are selective for horizontal stimuli, and green indicates regions where cells are selective for vertical stimuli. Regions responding to similar orientations are arranged in curved bands, which are arranged radially around special points where regions of different orientation preference appear to meet. These special points are called "pinwheels", and are a particularity of orientation preference maps. Around them, all orientation preferences are represented.

In addition to orientation bands and pinwheels, the functional architecture of V1 includes a map of ocular preference, an example of which is illustrated in Figure 4B. This map refers to the same region of cortex as that in Figure 4A, but codes for a different stimulus attribute. Here, white regions are those where cells respond the most to stimulation of the left eye, whereas dark regions are those where cells respond mostly to stimulation of the right eye. Whereas cells in the retina and in the LGN are
responsive only to stimulation in one eye, in V1 the visual signals originating from both eyes converge onto individual cells. This convergence usually involves unequal proportions. Thus, most V1 cells are preferentially excited by visual stimuli arriving in one eye, but can also respond to signals coming from the other one. Cells that prefer the same eye are found in clusters, forming vertical columns extending from the cortical surface down to the white matter. As illustrated in Figure 4B, regions sharing the same ocular dominance form stripes on the surface of the cortex.

The fact that V1 contains both ocular dominance and orientation columns has led to the suggestion that it is made of functional modules, each devoted to the analysis of a small region of the visual field. These modules would occupy approximately one square millimetre of the cortical surface and would analyse the full range of orientations and ocular dominance for a given spatial location. The modules would tile the surface but would not be precisely separated from each other. Indications as to the structure of these modules can be derived from the interactions between maps. For example, the borders of ocular dominance bands tend to be orthogonal to the borders of orientation preference bands. In other words, if one moves in a direction along the cortex where the orientation preference is changing rapidly, then the ocular dominance will be almost constant. Conversely, in those directions in which ocular dominance changes the most, orientation preference is mostly constant.

**Retinotopic organisation**

In addition to being organised by orientation preference and by ocular dominance, the surface of V1 is also retinotopically organised. This organisation is derived from that of the retina, of the optic nerve and of the LGN, and ensures that nearby regions in a visual stimulus will correspond to nearby regions in the cortex. An example of
retinotopy can be observed in Figure 5A. When the stimulus is a bullseye (left), the map of the responses in V1 (right) is a distorted but continuous representation of the stimulus. Note, however, that while this representation is continuous, it is also grossly distorted. Just as the retina and the LGN, the visual cortex devotes most of its cells to the centre of fixation, the fovea. Indeed, while the portion of the bullseye that is inside the second ring takes up about one fifth of the surface of the stimulus, the region that represents it occupies more than half of the cortical area covered by the whole stimulus. Similarly, the portion of the bullseye that is inside the third ring occupies only about 1/25 of the whole stimulus, but the area that represents it is about 1/3 of the total.

Another striking difference between the stimulus and the corresponding cortical map of activity shown in Figure 5A is the geometrical arrangement of the features. The concentric circles and radial lines present in the stimulus become stripes of activity in the cortical map that are vertical and horizontal. This geometrical distortion can be modeled mathematically (Figure 5B), and could be important to help the rest of the brain recognise objects regardless of their distance and of their orientation. Indeed, when we get closer to an object, its size on the retina is scaled up. If we tilt our head (or the object), the object's image on the retina rotates around the fovea. Thanks to the transformation operated by the cortex, such that scalings and rotations become simple translations, which may be easier to analyse for the subsequent stages of visual processing.

A striking example of the consequences of cortical magnification is illustrated in Figure 5C. In this simulation, a mathematical model was fed a photograph of a room (left) and generated a prediction of the distribution of responses across the primary visual cortex in one hemisphere (right). The model simulates an observer who points
her eyes to the letter O on an eyechart (see detail) located on the back of the room.

The representation in V1 is amazingly distorted. Most of the cortical surface devoted to the stimulus is taken up by the eyechart, and particularly by the letter O and by the nearby letters. Barely any cortical space is given to the rest of the room, which yet occupies by far the largest portion of the visual stimulus.

**Long-range interactions**

In addition to depending on the properties of the visual stimulus for which they are selective, the responses of V1 neurons also depend on those of nearby cortical regions. Indeed, while different regions of the cortex respond to different spatial locations in an image and to different visual attributes such as stimulus orientation, the responses in each region are not independent of each other.

The effects of the interactions among neurons in different regions appears to be mostly that of response suppression. An example of response suppression is known as "cross-orientation suppression". A stimulus such as a cross, which contains two orthogonal orientations, will stimulate two largely separate groups of V1 cells, which will suppress each other's response. The response to a cross is then less than the sum of the responses to the individual bars that compose it. Another example of response suppression is known as "surround suppression". Here, stimuli that occupy different portion of the visual scene can have a suppressive effect on each other, especially if they have the same orientation. For example, if one records from a V1 neuron and increases the length of an optimally oriented bar presented within a cell's receptive field, the cell's response increases only up to a point. In many cells, extending the bar further suppresses the responses, even though this suppression is originating from a region in the visual field that is outside the receptive field of the cell. This suppression
from beyond the receptive field is thought to originate in other V1 cells that have displaced receptive fields and are being stimulated by the extended stimulus.

Cells showing this property are often called "end-stopped". When looking at a visual stimulus such as that in Figure 3, end-stopped cells located at the positions such as the tip of the nose, or the chin, will be strongly activated. The detection of features such as end-points, or corners is considered important to our ability to identify various forms and faces.

In addition to suppression, interactions between different regions of the receptive field may also involve enhancement. This enhancement may result from a sort of double-suppression: if a population of neurons normally suppresses the activity of another population, and the activity of the first population is suppressed, the activity of the second population may be enhanced. Anatomically, the interactions between different regions are thought to be mediated by local circuits as well as by long fibres, running parallel to the cortical surface, that connect V1 cells over distances of several millimetres.

While their role in pattern vision is not fully understood, long-range connections were proposed to perform at least two functions. On the one hand, it was shown that they tend to connect cells that share the same preferred orientation. Moreover, they preferentially connect cells along an axis that corresponds to the cells' preferred orientation. As a result, these connections could serve to affect the responses of cells to stimuli that are "collinear". An enhanced representation of collinearity is very useful for the encoding of contours. A second putative role for the long-range connections is to synchronise the activity of cells that are coding for the same object.
They could thus be useful to segment an image into distinct, identifiable objects. This particular hypothesis is discussed in more detail below.

Regardless of its precise role in pattern vision the existence of long-range interactions underlines the importance of context in perception. It has been known for centuries that a given visual stimulus is perceived differently in different contexts. The properties of the long-range, intrinsic connections of V1 may provide a physiological explanation for these phenomena.

**Binocularity, stereopsis, and rivalry**

In normal viewing conditions, both eyes are focused on the same point of the world, and the images falling on the two retinas are very similar, as seen in Figure 6. These two images are effortlessly fused in the visual cortex, through the activity of binocular neurones. If the two eyes are looking into different directions, as in people with strabismus, or if one wears a prism in front of one eye, the two retinal images are different and cannot be fused. In that case, the observer is subject to binocular rivalry. Rivalry can take different forms. In some cases, one eye is completely suppressed, and the observer perceives only the image coming from the other eye. Sometimes, the suppression is only partial, and the observer perceives a meaningless, patchy juxtaposition of image components coming from the two eyes. In yet other situations, the observer can experience diplopia (double vision). Although it has these potential pitfalls, a binocular visual system also has advantages. Among them is the capability to extract depth information from the two retinal images, which is called stereopsis.

When the two eyes are focused on the fixation point, the image of that point falls on each eye's fovea. Similarly, all the visible points in the world that are at the same
distance from the observer as the fixation point will fall on corresponding parts of the
two retinas. As seen in Figure 6, a point that is further from the observer than the
fixation point will have its image formed on the left of the fovea in the right eye, and
on the right of the fovea in the other eye. Using this simple asymmetry, the brain
correctly deduces that this point is further away than fixation. An analogous situation
applies to objects located closer than the fixation point. This simple geometrical fact
is called "binocular disparity". Larger distance differences between an object and the
fixation point result in larger amounts of binocular disparity. By comparing the
images formed on the two retinas, it is thus possible to determine the depth of the
image elements relative to the fixation point.

It is now firmly established that there exist cells in V1, and other cortical areas as
well, that can compute binocular disparity. Measured independently in each eye, the
two receptive fields of those cells can be slightly displaced relative to one another, or
their internal organisation can be slightly different, giving them unequal positional
sensitivity. Such a cell is most optimally activated when a stimulus falls on regions of
the retinas that are slightly offset relative to each other, and will thus signal if a
stimulus is not at the same distance as the fixation point. Using this simple
computation of binocular disparity, the visual system is thus able to extract depth
from the two-dimensional images falling on the retinas.

### Beyond the primary visual cortex

The primary visual cortex has numerous connections with other visual cortical areas,
grouped under the name of "extrastriate" visual areas. In primates, over 30 extrastriate
visual areas have been discovered. Cortical area V2 surrounds V1 and projects to
other areas such as V3 or V4, and to areas whose names reflect their anatomical location (for example MT, for Middle Temporal area).

**Cortical area V2 and the perception of illusory contours**

The next stage of visual processing after V1 is performed in the secondary visual area, V2. At equivalent positions in the visual field, the receptive fields of V2 cells are slightly larger than in V1. Most other receptive field properties are rather similar in V2 and V1. A functional property that appears to be created “de novo” in V2 cells is their ability to respond to illusory contours.

Illusory contours are contours that an observer perceives when none is physically present. A well-known example is that of the "Kanizsa triangle", illustrated in Figure 7. Most observers report seeing the triangle’s edges between the partially occluded discs. Although the light distribution on a cell’s receptive field located in region is perfectly uniform, it has been shown that some V2 cells give a response to this illusory edge. Only cells whose preferred orientation corresponds to that of the illusory edge give responses, and the response vanishes if the three indented discs are slightly rotated, thereby destroying the illusion of the triangle, and of its illusory edges.

While the mechanisms involved to create such responses in V2 are still mysterious, the stage could be set by the “collinearity” enhancement performed by V1 cells. The possible role of end-stopped cells has also been stressed by researchers trying to explain this property. In any case, area V2 appears to perform operations that could underlie important aspects of pattern vision. Indeed, the capacity to interpolate the
contours of partially occluded objects is necessary to achieve a meaningful segmentation of the visual scene into various objects.

**Cortical area V4: towards the perception of form**

V2 has numerous connections with V4, a visual area of the temporal lobe. The properties of V4 cells are less well understood than those at earlier stages, but several results suggest that V4 cells have an important role in pattern vision. The receptive fields of V4 cells are considerably larger than in V1 or V2. This makes individual V4 cells suitable for the analysis of large areas of the visual field. Thus, it is generally thought that area V4 could be setting the stage for the detection and identification of whole objects. Indeed, V4 transmits information to the Inferior Temporal cortex, a cortical region that is known to be crucial for object recognition.

Several properties of V4 cells are consistent with the notion that this area is performing important operations for pattern vision. Most V4 neurones, like those of V1 and V2, are selective for the orientation of a visual stimulus. Many V4 cells are also selective for binocular disparity. Moreover, the foveal representation is even more enhanced in V4 than it is in either V1 or V2. Taken together, these facts strongly suggest that V4 is important for pattern vision. Recent physiological results, although still controversial, add support to this view. Researchers have shown that a number of V4 neurones show a preference for stimuli that differ from the bars and edges that seemed optimal to stimulate cells at earlier levels. For example, some V4 neurones respond better to stimuli made of concentric rings, or of radially oriented line segments than to simpler bar and edge stimuli. Thus, individual V4 cells seem capable to encode visual patterns more complex than those encoded by individual cells at earlier levels.
A number of studies have shown that the activity of V4 cells is strongly modulated by attention. The response of a V4 cell to a given stimulus is high if the stimulus as a behavioural relevance (i.e. if the observer is paying attention to it), and low otherwise. The strength of this attentional modulation has also been taken as evidence that V4 is performing a detailed analysis of the visual stimuli, and thus plays an important role in pattern vision.

**The binding problem**

Most of the theoretical and experimental work described above implies that the early stages of the visual system decompose an image into features, or components. Our visual perception, however, is one of unity and coherence, not of a juxtaposition of independent elements. This suggests that at some stage of the visual pathways, the image components must be bound together to form a coherent perception. To date, the solution to this problem is largely unknown.

In an attempt to provide an answer to this question, it was proposed that the neuronal substrate that “glues” different features together is found in the temporal pattern of neuronal activity. In that view, two cells that code for features belonging to the same object would synchronise their activity. The synchronicity of cell firing would thus serve as a “tag” for a given object, which is propagated through the various stages of visual processing. The so-called “binding by synchrony” hypothesis received some experimental support with the discovery of cells in the visual cortex of cats and monkeys that, for example, synchronise their activity when they are simultaneously activated by a single bar, and not when activated by two distinctly separate bars.
The “binding by synchrony” is the subject of intense debate and experimental scrutiny and there is, to date, no consensus on its validity.

**Conclusions**

It is now firmly established that the early stages of the visual system extract the relevant components of an image, while discarding the redundant information. Thanks to their retinotopic organisation, the early processing stages preserve the spatial relationships between the image components. Ample evidence has demonstrated that the orientations as well as the relative depth of contours are encoded in the visual cortex, and that this information is preserved across several cortical visual areas. Moreover, it appears that successive processing stages are sensitive to increasingly complex features, thus setting the stage for an efficient and effortless recognition of objects. Although the nature of these image-processing operations is well understood, their underlying biological mechanisms are still largely unknown. Moreover, the link between the activity of cortical neurones and the subjective quality of visual perception remains an almost complete mystery.

**Suggested readings:**


Glossary:

Neurone, action potential, spike, receptive field, axon, cerebral cortex, cortical area.

Figure captions
**Figure 1.** The basic organisation of the visual pathways. The images originating in the left and right visual hemifields are projected onto opposite parts of the two retinas. The retina's ganglion cells send their axons (which form the optic nerve) to the Lateral Geniculate Nucleus. LGN neurons then project to the primary visual cortex. The visual information is then further analysed in an array of extrastriate visual areas. Note that the left visual hemifield is analysed in the right cerebral hemisphere, and vice-versa.

**Figure 2.** Receptive fields and neural images in the retina and V1. **A:** A physical image obtained in conditions of dim illumination. **B-E, top row:** Receptive fields of a photoreceptor, an ON-center ganglion cell, a horizontally-tuned V1 simple cell and a horizontally-tuned complex cell. **B-E, bottom row:** Neural images corresponding to the cell types above. A neural image is a map of the responses of a given cell as it is moved over a whole image. White indicates positive responses, black indicates negative responses, and grey means no response.

**Figure 3.** Neural images in V1. **A:** Physical image. **B:** The two-dimensional receptive fields of 8 simple cells, each sensitive to vertical, horizontal, or oblique orientations. The top set encodes orientations at a lower spatial scale than the bottom set. **C:** The neural images resulting from the analysis of the image shown in A by the two sets of cells shown in B. Each cell extracts one particular orientation at a given spatial scale. Adding together the 8 images in C would result in a coarse, but faithful representation of the physical image A.

**Figure 4.** Maps in V1. Orientation preference map and ocular dominance map. (From fig. 1 in Obermayer, K. and G. G. Blasdel (1993). “Geometry of orientation and ocular dominance columns in monkey striate cortex.” J. Neurosci., 13: 4114-4129).

**Figure 5.** Retinotopy in the primary visual cortex. **A:** Stimulus (left) and map of the responses in V1 (right). The visual stimulus was shown to a macaque monkey, whose eyes were fixating on the bullseye. The map of the responses was obtained by flattening the V1 cortex, and reveal which regions were active during the stimulus presentation (from Tootell, R. B. H., et al. (1988). “Functional-anatomy of macaque striate cortex. II. Retinotopic organization” J. Neurosci., 8(5): 1531-1568. Figs 1 and 2B). **B:** Mathematical models of flattened retina (left) and flattened cortex (right). The transformation from one to the other was taken to be the logarithm of a complex number. Red and green curves indicate the transformation of concentric and radial features into vertical and horizontal lines of activity (from Frederick, C. and E. L. Schwartz (1990). “Conformal image warping.” IEEE Computer Graphics and Applications (March): 54-61. Figs 6a and b.). **C:** Application of the model to simulate the representation in primary visual cortex of the image of a room. In the back of the room is an eyechart (see detail), and the retina is centered on the letter "O". The representation in the cortex greatly magnifies the letters in the eye chart to the expense of the rest (Figures 6a, b and d from Schwartz, E. L., B. Merker, et al. (1988). “Applications of computer graphics and image processing to 2D and 3D modeling of the functional architecture of visual cortex.” IEEE Computer Graphics and Applications (July): 13-23).

**Figure 6.** Stereopsis. While the images of the fixation point fall on the fovea of each retina, the images of points located further or closer than the fixation locus fall on non-corresponding parts of the retinas. This geometrical fact is used by the brain to
compute the relative depth of objects (see text). (Figure from Kandel, E.R. (1991). "Perception of motion, depth, and form", in "Principles of Neural Science", 3rd Ed. Kandel, Schwartz, and Jessel, Eds. Figure 30-16.)

**Figure 7.** The "Kanisza" triangle, a well-known example of illusory contours. Most observers report perceiving the triangle's edges between the partially occluded discs.
Fixation point

Right eye

Left eye

Optic nerve

Lateral Geniculate Nucleus (LGN)

Primary visual cortex (V1)

Extrastriate cortex

Extrastriate cortex
Physical image

- Photoreceptor
- Ganglion cell
- V1 simple cell
- V1 complex cell
A

Physical image

B

Receptive fields

C

Neural images
A Map of orientation preference

B Map of ocular dominance