Review

Camouflage and visual perception

Tom Troscianko[1,*], Christopher P. Benton[1], P. George Lovell[1], David J. Tolhurst[2] and Zygmunt Pizlo[3]

[1]*Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK*
[2]*Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3EG, UK*
[3]*Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-2081, USA*

How does an animal conceal itself from visual detection by other animals? This review paper seeks to identify general principles that may apply in this broad area. It considers mechanisms of visual encoding, of grouping and object encoding, and of search. In most cases, the evidence base comes from studies of humans or species whose vision approximates to that of humans. The effort is hampered by a relatively sparse literature on visual function in natural environments and with complex foraging tasks. However, some general constraints emerge as being potentially powerful principles in understanding concealment—a 'constraint' here means a set of simplifying assumptions. Strategies that disrupt the unambiguous encoding of discontinuities of intensity (edges), and of other key visual attributes, such as motion, are key here. Similar strategies may also defeat grouping and object-encoding mechanisms. Finally, the paper considers how we may understand the processes of search for complex targets in complex scenes. The aim is to provide a number of pointers towards issues, which may be of assistance in understanding camouflage and concealment, particularly with reference to how visual systems can detect the shape of complex, concealed objects.

Keywords: camouflage; vision; edges; motion; search; object recognition

1. INTRODUCTION

(a) *Illumination and objects*

The visual sense is very useful to many animals. It allows the detection and identification of distant objects. The properties of visual systems vary considerably between different animals (e.g. Walls 1942; Autrum *et al.* 1973; Weckstrom & Laughlin 1995; Bowmaker & Hunt 2006), but the main issues concern the directional sensitivity (acuity) of the system; the light levels under which it operates; the field of view, including any areas of binocular overlap; the extent to which specific features such as spectral or motion information are extracted from the visual environment; and the spatial and temporal characteristics of sampling the environment.

The key property of visual objects is the extent to which they modify the incident light. The spectrum and geometry of the incident light are modified by the media through which it is transmitted—usually air or water. It is also modified by reflections from surfaces. Scattering by fluids, and inter-reflections, typically introduce a diffuse component to the propagation of light; objects are therefore illuminated in a variety of ways. The rules governing these effects are necessarily complex, and best understood by the computer graphics community (Ward & Shakespeare 2004). However, some simple consequences of this aspect of the behaviour of light are given below.

(i) Material properties

These determine both the spectral composition of the diffuse component of reflected light and its intensity. For Lambertian (matte) surfaces, this component does not vary markedly with viewing angle. For the specular component of (glossy) surfaces, the intensity and spatial properties change markedly with viewing angle. The diffuse component is therefore the more stable property of a surface.

(ii) Intensity borders

The intensity of a surface is determined by its material composition. If the surface has a well-defined border, then the intensity at the border will be different from that of the immediate background. The detection of a border can therefore be robustly encoded by detecting a sudden change in intensity in the scene. The strategy of finding edges based on intensity changes is ubiquitous in computer vision, e.g. the edge detectors of Marr & Hildreth (1980) and Canny (1986).

(iii) Other kinds of border

There are two types of intensity edges that are not coincident with the edge of an object. The first is an illumination edge, commonly known as a shadow.

* Author for correspondence (tom.troscianko@bristol.ac.uk).

Shadows are dark, and to a first approximation modify only the intensity of the region which they fill. However, the spectral composition of this region will also vary if the directional component of the illumination is different from the diffuse component. This is the case with sunlight that has undergone Rayleigh scattering in the atmosphere. Rayleigh scattering is a process in which short-wavelength light is more likely to be scattered by the atmosphere, resulting in blue sky. Since a shadow area receives illumination from the diffuse component of illumination which, in Rayleigh scattering terms, has an excess of short wavelengths, therefore shadows are rich in such short wavelengths. For humans, shadows are therefore both dark and blue; for animals with UV vision they are dark and UV coloured.

The second type of intensity edge that is non-coincident with the edge of an object is an internal marking. This may be coincident with an internal feature of the object, i.e. an object at a different scale—e.g. the abdomen of a moth. However, an intensity edge may also represent a change in reflectance without a change in the nature of the object. Such markings are commonly referred to as 'texture'. One of the characteristics of visual textures (Julesz 1971) is that the exact position of the elements is not important. The grain of a piece of wood is a characteristic property of the wood, but the exact positions of the fine grain are not important. Rather, it is the statistical distribution of properties of the texture that is a characteristic feature of the object in question—thus, oak bark has a different texture to beech bark, even if the intensities and spectral properties of both barks are to be similar.

It becomes clear that the existence of these two types of illumination edges that are not coincident with object boundaries in the classical sense poses a problem for systems that detect objects simply by locating intensity edges. The artificial vision system proposed by Marr (1982), and implemented by an interdisciplinary team of researchers, became known as TINA. Early implementations of TINA (Porrill *et al.* 1988), based on the Canny edge detector, would fail in situations where there are strong shadows or textures. Such failures therefore provide pointers to situations which an animal may exploit to make simple identification difficult—high-contrast edges that are non-coincident with an object boundary, and which are not a texture in the classical sense of the term, cause difficulties for object segmentation systems.

(iv) *Spectral information*
We have already alluded to the fact that specular and diffuse reflection components, and also direct versus scattered illumination, have different spectral properties. By 'spectral', we refer to the wavelength composition of light. Light emanating from the Sun has a broad spectrum ranging from 300 to approximately 1000 nm. As weather and time of day change, the actions of Rayleigh and Mie scattering affect the spectral composition of both direct and diffuse light. Mie scattering is the process by which the sky surrounding the Sun appears to take on the Sun's colour. Unlike Rayleigh scattering, this process does not favour short wavelengths. The variation in atmospheric colour due to these processes is primarily

along an axis that, in primate vision, is in the yellow–blue direction (Lovell *et al.* 2005). This means that, as weather and time of day change, the main effect is to alter the balance of long-wavelength light energy to short-wavelength one (Finlayson & Funt 1994; Barnard *et al.* 1997; Lovell *et al.* 2005). This situation changes near sunset, when Mie scattering becomes increasingly important, and (in primate terms) the red–green balance changes dramatically. In human vision, this results in significant failures of 'colour constancy'. Colour constancy is the principle by which a visual system may discount the spectral properties of illumination and encode the more important reflected colour. To give a simple example: a white wall will appear white to human observers under a wide variety of weather conditions. However, the same wall may appear pink from approximately 10 min before sunset. The period when spectral properties of light are changing rapidly therefore provides challenges for object-classification systems that rely on (say) the red–green balance remaining roughly constant as a function of weather conditions.

An important benefit of being able to sense spectral information lies in the ability to disambiguate illumination edges from object edges. If we assume that the spectral composition of a shadow is the same as that of a non-shadow area, then the identification of a shadow is facilitated by the observation that the spectral properties (colour) are the same on both sides of the shadow boundary. This assumption is violated if an object boundary coincides with a shadow boundary; however, this is difficult both to achieve and to maintain over time. If we assume that a shadow is both dark and rich in short wavelengths, then such a combination may lead to the robust identification of shadows (primarily in regions that have little cloud cover and in which shadows are therefore strongly blue/UV). There have been speculations that some insects detect shadows in this way (Steverding & Troscianko 2004).

In the same way as colour can be used to disambiguate shadows, it can also be used to augment the perceived uniformity of a region that is rich in texture-based intensity variation. Thus, tree bark varies considerably in intensity, but its spectral signature remains relatively constant. Also, owing to changes in lighting, a given object may not vary in intensity compared with its background—but it will often vary in colour. This is particularly true for fruits among foliage—a monochrome version of the scene fails to render the fruit visible, especially under 'dappled' lighting conditions. However, the spectra of edible fruit and leaves are readily distinguishable from their leafy background. This principle has been argued to have driven the development of primate trichromacy (Osorio & Vorobyev 1996; Regan *et al.* 2001; Párraga *et al.* 2002) and, therefore, provides an important constraint in object identification. We can summarize this, and the preceding point about texture, thus: if colour changes suddenly at a point in the scene, we can be relatively confident that this point coincides with an object boundary. If both colour and intensity change together, we can also be confident that an object boundary has been detected, unless the change is to a dark blue/UV colour associated with shadows.

(v) *Change over time*

The visual environment is often surprisingly static. A large object tends to have high mass, and high mass cannot be moved without a great expenditure of energy. Large objects therefore tend to remain stationary. However, lighter objects, such as foliage, can move as a result of wind and contact with moving animals. Such movement is stochastic in nature, and often does not result in a significant overall translational movement over time. The movement of leaves etc. is therefore a movement equivalent to a texture in which the statistical properties of a given portion of the scene are indicative of the likely cause. A movement-sensing system therefore requires low-level detectors of motion, which is a vector quantity, encoding both speed and direction. A scene segmentation based on motion will therefore often attempt to group component motions together with a 'common fate' principle in operation.

One consequence of the relative stability of the visual environment is that it can serve as an external memory for scene content. Most of the information in a scene remains stable over time courses of seconds, and often longer. This property of the world has been invoked to account for the ability to sample scenes in a stochastic manner, such as with eye movements in humans or with a partly random flight path in insects (Land & Nilsson 2001). If most of a scene does not change with time, the exact order of sampling information does not matter greatly; nor is there a need to generate a detailed internal model of the scene, since the same information remains available for a long time in the environment. However, if this assumption is violated, as would be particularly the case for animals moving in groups, and those operating in an environment of high change, such as moving water or airflow, one would expect the 'external memory' assumption of the world to be violated. This would be expected to result in either a larger stored memory component, or a more rapid, or more parallel, sampling of the environment.

(b) *Summary and implications for camouflage*

We have considered how evidence for a change in an object is made available by the behaviour of light. We have seen how spatial, temporal and spectral factors interplay in likely solutions to this problem. An object boundary may be detected by sensing an abrupt change in colour or intensity, but neither process is immune from errors. Such errors may arise from spurious boundaries caused by illumination changes or by internal structure in the object in question. Separate detectors may therefore be needed for such confounding cases—in particular, for detecting textures and shadows. It follows that any system that seeks to conceal its presence by making its body less clear as a detectable/recognizable object, may benefit from some, or all, of the following strategies:

— To make object boundaries hard to detect by making them similar in spectral content, and intensity, to the immediate likely background. Note that this similarity only needs to apply to sensing systems from which the animal wishes to remain concealed.
— To provide false evidence of boundaries being a texture or due to an illumination change.

— To introduce high-contrast internal detail that is more salient than the edge, and sufficiently random not to serve as an independent cue for an identification process.
— To mimic the movement (or its absence) of the immediate surroundings or for the movement to be sufficiently random (whole body, or parts of it) to disable common fate detectors.

We have outlined some constraints on object detection and recognition. We will now consider how research, primarily on human vision (or animals deemed to be similar to humans), has informed us about the likely operation of relevant mechanisms.

2. EDGE DETECTION PROCESSES: DISRUPTION AND CAMOUFLAGE

Identification of an object (or *figure-ground processing*) has two stages. First, there is a low-level process whereby individual neurons detect the locations, polarity and orientation of small edge segments; the neurons might be the *simple cells* in mammalian primary visual cortex, V1 (Hubel & Wiesel 1959, 1962). However, we shall show that V1 *edge detectors* are rather weak at the task compared with detectors proposed for computer vision (Marr & Hildreth 1980; Canny 1986). The second (higher-level) stage groups the local edge information (resolves *border ownership*), identifying those edges that belong to a single object and rejecting others that belong to the background (Lamme 1995; Grossberg *et al.* 1997).

In principle, the neurophysiological steps of edge detection and edge grouping might be exploited in two ways by a prey animal, making it less visible to predators. First, its coloration or markings might make the small edge segments difficult to discern, most obviously, if the animal's body is of very similar colour and brightness as the background. Of course, 'colour' and 'brightness' depend upon mechanisms in the predator's eyes that determine the range of light wavelengths that are visible to it. A prey animal must evolve to be invisible to its predator specifically. However, even if the animal is not of the same colour and brightness as the background, there are properties of V1 neurons that may be exploitable to make edge segments harder to discern.

A second way of exploiting the mechanisms of edge processing would be to disrupt the grouping of the small edge segments to form a coherent outline of a whole object. Even if most of the individual edge segments are visible, it might be possible to confuse the edge grouping processes by deleting some edge information, by distorting the location and polarity information about edges that are present, and by inserting misleading information about edges that are not actually present (Stevens & Cuthill 2006).

(a) *'Edge detectors' in V1*

Each *simple cell* in V1 has a *receptive field* that occupies a small part of visual space, and typically consists of 2–5 parallel, elongated regions in which small spots of light have differing effects (see figure 1*a*,*b*). In alternating regions of the field, light causes excitation, while in
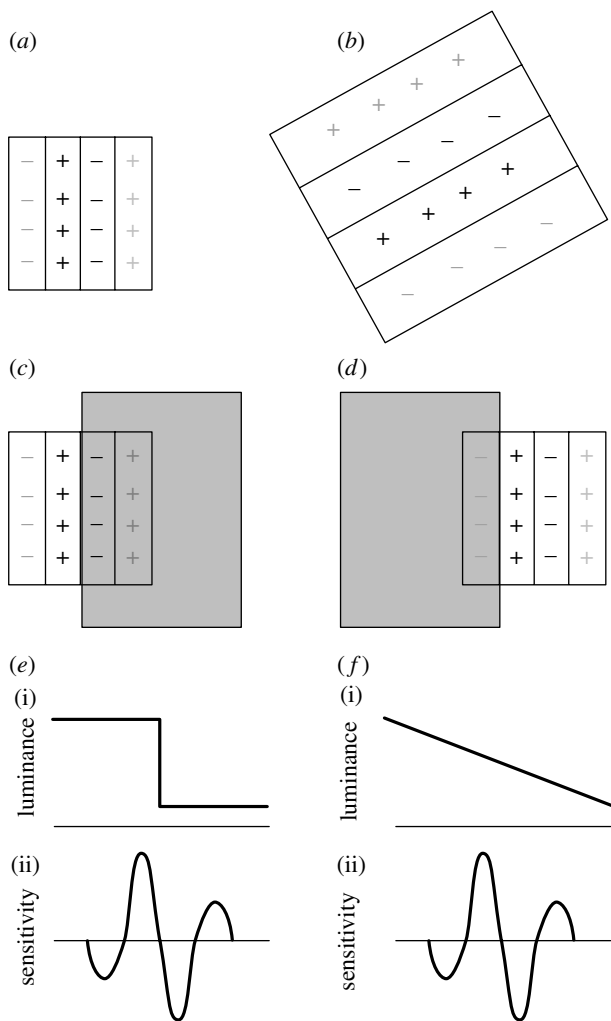
Figure 1. (*a*, *b*) Schematics of receptive fields of two V1 'edge detector' neurons. Light falling in excitatory regions (pluses) increases the neuron's firing rate, while light falling in the inhibitory region (minuses) will reduce firing. Conversely, darkness in the inhibitory region will increase firing. The grey symbols show that the outer flanking regions are weaker than the central pair. Simple cell receptive fields come in a great variety of spatial geometries. (*c*) The edge detector of (*a*) will be best stimulated by a light–dark edge as shown, which exactly falls along the border between the strongest excitatory and inhibitory regions. However, it will also respond weakly to an edge of reversed polarity that lies on the border between other receptive field regions (*d*). (*e*(i)) Shows the sharp luminance transition of the most effective edge, like that in (*c*) and (ii) shows how the responsiveness of edge detector (*a*) changes continuously across the receptive field. (*f*) The same edge detector will respond very weakly if the luminance changes gradually across the receptive field.

other regions it causes inhibition. Simple cells do not fall into neat classes of edge detector and 'bar detector' (Field & Tolhurst 1986; Ringach 2002), but they will respond well to borders between bright and dark objects, provided that the borders are of just the right orientation and location for the receptive field, falling exactly along the main excitatory–inhibitory border in the field (figure 1*c*). Confusingly, and significantly for camouflage, the neuron can respond to edges of the wrong polarity if they are located appropriately on the border between weaker excitatory and inhibitory regions (figure 1*d*). A single strong edge will stimulate

multiple neurons, apparently signalling several parallel edges of different polarity. Moreover, edge detectors do not only detect features such as edges or line segments; they will respond to any feature that has any similarity to an edge, provided that the feature is intense enough (Maffei & Fiorentini 1973; Movshon *et al.* 1978; Jones & Palmer 1987; Smyth *et al.* 2003). Edge detectors devised for computer vision (Marr & Hildreth 1980; Canny 1986) include nonlinear processes unlike real neurons (Tolhurst & Dean 1987), to restrict responses only to frank edges and to prevent such ambiguities.

(b) *Stopping edge detectors responding to edges*

Different simple cells respond to different orientations and over different spatial scales (compare figure 1*a*,*b*). In practice, this means that neurons prefer sharp edges, with pronounced step changes in brightness across the border (figure 1*e*); they respond less well if brightness changes gradually between dark and bright areas (figure 1*f*). A potential disruptive coloration strategy would be to make one's border edges 'blurry' by having graded pigmentation along the outline. Although to a first approximation, simple cells act as linear filters, there are a variety of nonlinear interactions among populations of V1 neurons (Carandini *et al.* 2005), which affect the way in which individual neurons respond to their best features, perhaps contributing towards resolution of border ownership (Lamme 1995, 2003; Zhou *et al.* 2000). One interaction is *non-specific suppression* (or *contrast normalization*); all the simple cells (with a whole range of different receptive field configurations) subserving a small part of the visual field drive an inhibitory pool, which feeds back to inhibit all the same simple cells (Bonds 1989; Heeger 1992; Tolhurst & Heeger 1997; and see Marr 1969). The functions of the inhibitory pool have been debated (Heeger 1992; Schwartz & Simoncelli 2001; Lauritzen & Tolhurst 2005) but, in the present context, non-specific suppression can act powerfully to suppress the response of one neuron to its best stimulus when other strong features are being detected by 'rival' simple cells whose receptive fields are in much the same part of the visual field; it has long been known that strong line stimuli, for instance, can make a weaker stimulus invisible (Tolhurst 1972; Weisstein & Bisaha 1972; Harmon & Julesz 1973).

Stimuli in the areas *surrounding* a simple cell's receptive field may be antagonistic (Blakemore & Tobin 1972; Cavanaugh *et al.* 2002*a*,*b*). *End inhibition* is caused particularly by stimuli outside the field of the same orientation as those that excite the neuron when presented within the field (figure 2*a*); thus, a short edge confined to the receptive field might be excitatory, while a long edge extending beyond the field might not (Hubel & Wiesel 1965; Gilbert 1977). In fact, the orientation tuning of this surround inhibition is rather complex (Cavanaugh *et al.* 2002*a*,*b*), so that perpendicular stimuli can also suppress if they are presented to the sides of the receptive field rather than along its axis of elongation (figure 2*b*).

Thus, pigmentation making strong edges near to or perpendicular to the animal's outline might suppress the information about the true outline, providing
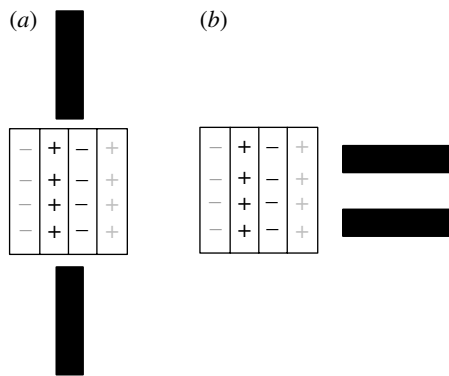
Figure 2. V1 neurons can be inhibited by stimuli that fall outside their 'classical' receptive field. (*a*) One example is 'end inhibition', when stimuli of the preferred orientation (the two dark bars) can be inhibitory if they are presented outside the receptive field, but in line with the axis of elongation. (*b*) Perpendicular stimuli (the dark bars again) can also be inhibitory if they are presented to the side of the receptive field. High-contrast 'tick marks' perpendicular to an animal's outline might disrupt the responses of the edge detectors that would otherwise signal the location of the outline.

disruptive information about non-coherent edges at erroneous locations and at erroneous orientations (Cuthill *et al.* 2005).

(c) *Making edge detectors respond to non-existent features*

Many visual illusions include *illusory contours* (figure 1*a*,*b* in supplementary materials); within a geometric figure, there may *appear* to be a shape or edges between bright and dark when, in truth, there are no such borders. Illusory contours can arise when sharply defined geometric shapes act typically as the ends or corners of non-existent lines or borders (figure 1*c* in the electronic supplementary material). Such geometric shapes might fool a predator's visual system into believing that there are other edges in other locations or even that there are coherent objects that do not resemble the outline of prey. Illusory contours probably arise during border owner-ship resolution, rather than the initial stages of edge segment detection, but neurons even in V1 respond to illusory geometric figures as if the contours are really there (von der Heydt *et al.* 1984; Grosof *et al.* 1993; Mendola *et al.* 1999).

3. MOTION
(a) *Encoding of motion*

In primates and cats, neurons sensitive to motion arise as early as primary visual cortex (Hubel & Wiesel 1959, 1962); in other species (for example rabbits and frogs) they may be found within retinal processing (Barlow *et al.* 1964; Finkelstein & Grüsser 1965). Such neurons are subject to the *aperture problem* (Adelson & Movshon 1982)—imagine drawing a straight line on a sheet of paper and then placing it under another piece of paper with a small hole cut in it. You can move the below piece of paper in many different directions to end up with what appears to be the translation of a line from one side of the hole to another. Any one velocity across the aperture can be the result of many different

combinations of the speed and direction of the underlying sheet.

How then do we extract unambiguous object motion? There are two basic schemes, almost certainly complementary. First, imagine placing a line end in an aperture. When you move this around, the direction of motion is unambiguous. Therefore, if you have cells in primary visual cortex that respond not to straight contours but to line ends or corners, these cells can correctly indicate object motion. Such cells are termed *end stop* or *hypercomplex* (Hubel & Wiesel 1965; Gilbert 1977) and there is good evidence that they play an important part in motion perception (Pack *et al.* 2003). If outputs from such cells are to be used to determine object motion, they must constrain the perceived motion in parts of the object not characterized by corners or line ends.

The other complementary method of extracting object motion also necessitates the spatial integration of motion signals. The motion of a single straight contour is ambiguous. It can be caused by a range of possible velocities. However, the motion of a straight contour lying at a different angle may be caused by a different range of velocities. The velocity common to these two sets (the *intersection of constraints*) gives the true motion of the object (Adelson & Movshon 1982). It therefore seems clear that the spatial integration of motion signals plays a critical role in motion processing. In macaque, the extrastriate area middle temporal (MT), also termed V5, appears to be an area largely dedicated to motion processing. MT neurons appear to integrate inputs of motion sensitive cells from primary visual cortex and have receptive field sizes that are approximately an order of magnitude larger (Born & Bradley 2005). The existence of a human homologue of this area is well established (Zeki *et al.* 1991; Tootell *et al.* 1995).

The process of integrating separately moving areas of an object moving in depth into a single object is termed *structure from motion*; the process appears to be dependent on neural structures within the motion processing hierarchy from area MT upwards (Orban *et al.* 1999; Vanduffel *et al.* 2002). In many ways the use of motion information for detecting the presence of animals is an exercise in recovering structure from motion although, in this case, the motion of the animal will most likely consist of a variety of differently moving parts. The recognition of the natural motion of animals falls into the field of *biological motion* (Blake & Shiffrar 2007). Typically, this is studied by degrading the stimulus to a series of dots attached to various important points such as ankles, knees, pelvis, etc. When such a *point-light walker* is animated the agent and the nature of its motion are readily recognized (Johansson 1973; Dittrich *et al.* 1996).

In initial accounts, the ability of humans to detect biological motion was taken as evidence for a special sensitivity to motion of this type (Hiris 2007). This view has recently been undermined by the finding that the addition of form to non-biological motion results in levels of performance similar to those found with biological motion (Hiris 2007). Human sensitivity to biological motion may well therefore reflect a general sensitivity to structured motion. On the other hand, the specific trajectories shown within biological motion

stimuli appear to follow a certain form; a two-thirds power law relating their tangential velocities and local curvature (Ivanenko *et al.* 2002). Functional imaging has demonstrated that humans show a more widespread and stronger response to motion of this type than they do to other comparable motion (Dayan *et al.* 2007).

(b) *Motion camouflage*

Given the widespread sensitivity to motion, how can motion be camouflaged? There seems three manners in which this may occur: motion signal minimization (MSM); optic flow mimicry (OFM); and motion disruption (MD). Camouflage through MSM is associated with the prevention of low-level detectors indicating motion activity. Camouflage through OFM is associated with an attempt to mimic the background or surrounding motion so that (although the motion is detected) it does not provide a cue for segmentation. MD involves a breaking or misrepresentation of motion cues to distort the perception of that motion.

MSM can be split into two further subtypes. First, actually minimizing motion itself (and therefore the motion signal) and second, minimizing the motion signal created by any given motion. The former is probably the most obvious technique for camouflaging motion. It is used, for example, by predators trying to approach stationary prey and simply involves moving slowly. All things being equal, the most obvious approach trajectory will be directly towards the prey. When this is done, the only motion cue is one of the predator looming, a strategy that again minimizes the motion signalled by the predator to the prey. The minimization of the motion signal for a given motion depends on reducing the signal available to the motion processing system. For example, when settling on stripe patterns, cuttlefish (*Sepia officinalis*) orient their bodies so that their major axis lies perpendicular to the stripes (Shohet *et al.* 2006). Shohet *et al.* suggest that this reduces motion signals created by the cuttlefish's occlusion of the underlying pattern.

The term *optic flow* refers to the motion of elements relative to an observer moving through an environment. The basic concept underlying OFM is simple; a shadower wishing to hide itself from a translating shadowee moves in such a way that its motion is indistinguishable from the optic flow perceived by the shadowee (Srinivasan & Davey 1995). Note that the term *shadowee* refers to an agent wishing to hide its motion, while *shadower* refers to the agent from which the motion is hidden. Take a prey animal moving through an environment. If the predator simply heads straight towards the moving prey then, from the point of view of the latter, the predator will appear to both loom and to have a sideways component in its relative motion that will distinguish it from the background optical flow perceived by the prey. On the other hand, the predator can choose a fixed point in the environment and then approach its prey in such a way that the predator's position always lies directly between its prey and that fixed point. In this case, the predator will (if we ignore looming) have the same optic flow component as the chosen fixed point from the point of view of the prey.

The strategy has been shown to be used by dragonflies (Mizutani *et al.* 2003) and hoverflies

(Srinivasan & Davey 1995) and has been demonstrated to be an effective method for the camouflaging of approaches to human observers (Anderson & McOwan 2003*a*). The movement of the shadower can be viewed in terms of epochs where, at the start of each epoch, the shadower makes a decision about the direction and speed that they should move in. To successfully implement OFM, the shadower needs to be aware of (i) their current position with respect to the chosen fixed point, (ii) the current position of the shadowee and (iii) the motion of the shadowee. Recent work has shown that a simple neural network architecture relying on visual information available to the shadower can successfully implement OFM (Anderson & McOwan 2003*b*).

There have been a number of recent mathematical approaches to OFM; these can be split into two camps, one where the chosen fixed point is the start of the shadower's motion (Glendinning 2004) and the other where the fixed point lies at infinity (Justh & Krishnaprasad 2006; Reddy *et al.* 2007). The difference between these can be made clear if one thinks of a line connecting shadower and shadowee. When the chosen point is at the start of the shadower's motion, the shadower–shadowee line will always run through that chosen point, rotating about it as the shadowee moves through the environment. On the other hand, when the fixed point lies at infinity, then the shadower–shadowee line does not change its compass bearing; it has no rotational component.

The first of the above is clearly the best in terms of OFM as, when the shadower begins to move, there is no optic flow discontinuity. From the point of view of the shadowee, the shadower begins to loom. The *infinity-point* strategy would be ineffective against an obvious close background but would work well against, for example, sky. Additionally, the computational demands of the infinity-point strategy are probably less than those of any non-infinity-point (or *real point*) strategy as the position of the shadower in relation to its start point does not need to be calculated.

The *infinity-point* strategy might well therefore be the preferred choice with aerial predators, particularly, if they approach their prey from above. Indeed, Mizutani *et al.* (2003) show that dragonflies employ both real-point and infinity-point strategies. Recent evidence has shown that echolocating bats use what appears to be point at infinity approach when attempting to capture flying insects (Ghose *et al.* 2006). Ghose *et al.* characterize their approach trajectory as a *constant absolute target direction strategy* and show that it minimizes the time needed for the bats to intercept their prey. In terms of the present discussion, this finding is important because the bats are not clearly camouflaged; their approach can be identified by the noise they make as an intrinsic part of their echolocation. What might appear on the surface to be OFM is actually driven by other criteria. However, it is worth emphasizing that both hoverflies and dragonflies do appear to use real-point OFM as part of their behavioural repertoire (Srinivasan & Davey 1995; Mizutani *et al.* 2003).

MD involves the manipulation of contours and form to create a misperception of motion in the perceiver.

When an object is defined by high-contrast contours, its perceived direction of motion can be biased by the orientation of those stripes (Wuerger *et al.* 1996). This is basically a reflection of the aperture problem and reflects the influence within the motion integration process of mechanisms that signal motion orthogonal to contours. Whether or not MD is a motive for the striping patterns seen in many animals is moot. However, during the First World War, dazzle paint (called razzle–dazzle in the US) was applied to allied shipping in an attempt to reduce the toll from attacks by submarines.

Dazzle paint involved painting high-contrast striped coloured patterns on to shipping. Its primary purpose was to confuse the perceived motion of the ship in terms of both its speed and heading (Behrens 1999). Note that part of this was undoubtedly figural deception rather than motion deception; many dazzle paint schemes create the impression of a false bow. Misconstrual of a ship's motion could prevent a submarine getting into a good attack position and misperception of target motion could reduce the effectiveness of any weapons targeted at the camouflaged vessel.

In conclusion, there are potentially a variety of ways in which motion can be camouflaged. This ranges from the obvious 'move as little (or slowly) as possible' to more complicated techniques where a shadower mimics the optic flow background from the shadowee's point of view. Additionally, there are good theoretical reasons to think that the manipulation of configural information can create a misperception of an object's or animal's motion. A deliberate attempt to do this has been through the dazzle painting of ships; however, the British Admiralty, in a report towards the end of the First World War, noted that there was no evidence for dazzle painting's effectiveness (Behrens 1999). The role of MD as a possible camouflage technique is therefore currently open to debate.

4. OBJECTS AND SHAPE

As indicated at the beginning of this paper, the main task of vision is to detect and identify objects in the environment. In the context of camouflage, animate objects are of primary interest. Animals are best identified by their shapes. Colour, texture or size does not uniquely identify an animal. Colour, texture and size are secondary in the sense that they allow fine discriminations of an object's details, after the shape of the object has been recovered. In this treatment, shape is defined conventionally as those global geometrical properties of the object that are not affected by rigid motion and overall size scaling. Shape carries a lot of information about an object because it is 'complex' (Pizlo 2008).

An animal's visual system is faced with the difficult problem of how to recognize a three-dimensional shape from incomplete two-dimensional retinal information. Our knowledge of three-dimensional shape perception is limited because it comes almost exclusively from the study of human subjects. A brief overview of the reconstruction, recognition and detection of three-dimensional shapes by humans will be presented next. It is followed by a discussion of the means available to animals that can be used

to prevent the correct perception of their three-dimensional shape (camouflage).

(a) *How three-dimensional shapes are perceived*

There are at least three tasks related to the perception of three-dimensional shapes that the visual system may need to accomplish: (i) detection of the presence of a shape, (ii) recognition of a familiar shape, and (iii) reconstruction of a shape. Conventionally, shape reconstruction has been considered to be the most difficult of the three (Marr 1982). We will begin with (iii), shape 'reconstruction', because this task is the most fundamental. Note that in our approach, it is more appropriate to talk about three-dimensional shape 'recovery', than reconstruction because the term reconstruction as used by Marr refers to rebuilding three-dimensional shapes from local surface patches. Our term recovery emphasizes the fact that the percept of three-dimensional shapes is not built from its elements. Instead, the three-dimensional shape percept is formed by the application of abstract shape properties, such as symmetry. In this approach, shape recovery proves to be simple, requiring only relatively few computations, making it potentially effective with primitive, as well as sophisticated, vision systems. An approach to shape perception as ours should provide clues to the nature and effectiveness of visual camouflage throughout the animal kingdom.

(i) *Shape recovery*

According to Marr (1982), reconstruction of a three-dimensional shape from a two-dimensional image is computationally difficult because the information about depth has been lost in the projection from the three-dimensional space to the two-dimensional image. In this view, the visual system must try to collect additional images of the same three-dimensional shape by moving relative to the object and/or by using binocular stereo vision (Julesz 1971; Ullman 1979; Longuet-Higgins 1981). There is also another, easier way to recover the shapes of objects. Note that most (probably all) animals are symmetric (Thompson 1992). It was shown by one of us that using three-dimensional symmetry and three-dimensional compactness as constraints leads to accurate recovery of a three-dimensional shape from one of its two-dimensional images (Pizlo 2008; Sawada & Pizlo 2008; Li *et al.* in press). Three-dimensional compactness is defined as the ratio between object's volume squared and its surface area cubed (V^2/S^3).

Now, consider an example of three-dimensional shape recovery using symmetry and maximum three-dimensional compactness constraints. Figure 2a (electronic supplementary material) is a two-dimensional image of a symmetric insect, a mantis. Figure 2b shows the main contours drawn by hand and superimposed on the image of the mantis. These contours were used for the three-dimensional shape recovery. The pairs of symmetric contours (lines) were marked by hand before the recovery. This two-dimensional shape was then used to produce a three-dimensional shape whose three-dimensional symmetry and three-dimensional compactness are maximal. Two views of the recovered three-dimensional shape are shown in

(c) and (d). Note that the body of the mantis does not have a lot of volume, and that the contours, drawn by hand, have zero volume. For these reasons, the volume and the surface area of a convex hull of the three-dimensional contours were used in the computations. Recall that a convex hull of a set of three-dimensional points is the smallest convex three-dimensional region that contains all the points in the set. This example shows that three-dimensional symmetry, if detected and described in the two-dimensional image, allows the three-dimensional shape to be recovered reliably. The entire symmetric three-dimensional shape may often be recovered even when part of the shape is occluded. Recovery and recognition of the shape of a predator or its prey is likely to fail if its symmetry is not detected, or if its critical contours are not extracted. Symmetry and contours provide the primary mechanisms underlying the use of camouflage.

Note that three-dimensional shape recovery does not require motion or binocular disparity. Three-dimensional shape recovery can be done reliably from a single two-dimensional image because all animals are symmetric. Note that the animal's visual system has to find an object in the two-dimensional image before it can recover its three-dimensional shape. Finding objects in a two-dimensional image is called 'figure-ground organization'. Specifically, figure-ground organization refers to (i) specifying two-dimensional contours that represent contours of the three-dimensional shape, (ii) determining which pairs of features are symmetric in the three-dimensional interpretation, and (iii) determining which contours are planar in the three-dimensional interpretation. If figure-ground organization fails, the three-dimensional object will not be perceived (the object is camouflaged).

(ii) *Recognition*

Recognition is, in principle, easier than three-dimensional recovery because recognition of a three-dimensional shape can be based on the characteristic parts of the shape. This is the main idea behind Biederman's (1987) 'recognition by components' theory. In order to recognize a three-dimensional shape, the animal has to be familiar with the specific shape or, at least familiar with the category of shapes to which the specific shape belongs (cats, birds, etc.). This raises the obvious question of whether animals learn the shapes of important objects (prey, predators) or are born with this information?

Recognition of a three-dimensional shape could be done by matching three-dimensional shapes or their parts, stored in the memory, to the two-dimensional retinal image, or to the three-dimensional recovered shape. The former seems more direct, in the sense that it does not require three-dimensional shape recovery, so it is not surprising that several algorithms have been proposed for matching three-dimensional shapes with two-dimensional retinal images (Lowe 1985; Biederman 1987; Basri & Ullman 1993; Pizlo & Loubier 2000). If the object is almost planar, or has planar parts (e.g. a moth sitting on the ground), recognition may involve affine or projective invariants (Mundy & Zisserman 1992; Weiss 1993).

(iii) *Detection*

Detection of objects in two-dimensional images involves (i) detecting a feature not part of the background, (ii) identifying a region in the image representing an object, (iii) describing its contours (two-dimensional shape) and (iv) verifying that the two-dimensional shape is produced by an object. The first step involves visual search (see §5), in which some discontinuity of the background is detected. The discontinuity may be defined along any of a number of perceptual dimensions: lightness; colour; motion; depth; and texture. Note that visual search does not have to result in object identification (i.e. provide an answer to the 'what' question), but only in the location of something unusual (i.e. provide an answer to the 'where' question). Currently, it is commonly accepted that these two aspects of an object (its presence and location versus identity) are processed separately in the brain. There is, however, an ongoing discussion about the functional role of the anatomical pathways involved (i.e. of the dorsal versus ventral stream), as well as about the order of processing of these two aspects (i.e. detection before identification versus recursive computations in which identification may feedback to detection). The second and third steps (analysis of texture and contours) are described above. In the fourth step, the visual system verifies whether the two-dimensional shape is produced by a three-dimensional shape. How can this be done? If a three-dimensional object is symmetric, then the line segments connecting images of symmetric features are all parallel to one another in a two-dimensional orthographic image and, furthermore, their midpoints are *not* collinear. If the midpoints are collinear, the symmetric shape 'out there' is planar. The parallelism of several line segments in the two-dimensional retinal image should not be difficult to verify. This kind of computation is probably done in the early stages of the visual processing.

(b) *How to make object recognition difficult*

A three-dimensional shape will not be seen if any of the four steps enumerated above fails. First, if there is no sign of background discontinuity, the observer (prey or predator) will not allocate its attention to this part of the visual field. Once the attention is allocated, a distinctive region representing the object must be found. Otherwise, the object will not be seen. This can happen when an animal's skin has a texture similar to that of the background. Even when a distinctive region is found, its two-dimensional shape may not be described adequately. This can happen when an animal's skin has distinctive contours whose geometry is unrelated to the animal's three-dimensional shape. The zebra's stripes are a good example. Next, the three-dimensional shape may not be perceived as an object if the symmetry in the image indicates two-dimensional, rather than three-dimensional symmetry out there. For example, high-contrast texture and contours on the back of some frogs form two-dimensional rather than three-dimensional symmetric patterns because the frog's back is approximately planar. The viewer may overlook the shape of the three-dimensional frog, if the viewer detects the two-dimensional symmetry of these patterns. Using symmetric coloration has major
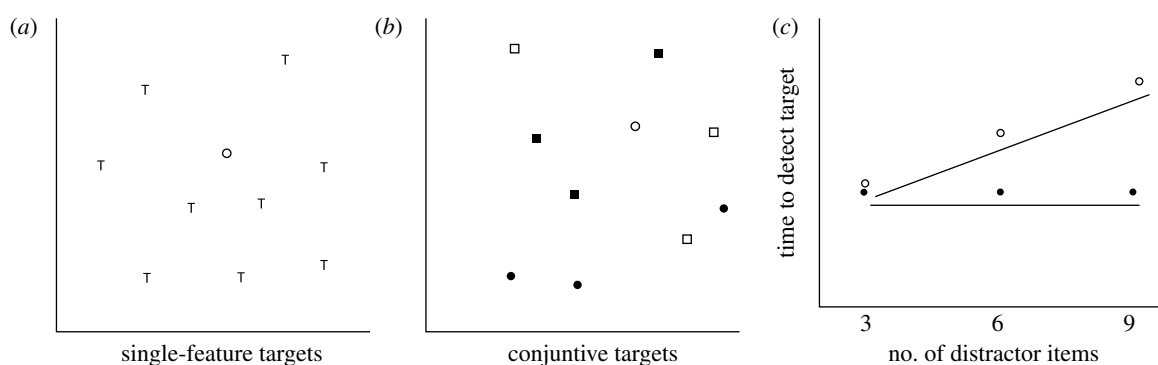
Figure 3. Examples of visual search stimuli with typical results. (*a*) A single feature search where the target is the open circle and (*b*) a conjunctive search with the same target. Note that it is not sufficient to find the hollow object, nor a circular object. The target is defined by a conjunction of colour and shape. (*c*) Typical experimental results, where the reaction times increase as a function of the number of distractors for conjunctive search but not for single-feature search.

benefits for camouflage. Symmetry is particularly effective for camouflage because its analysis takes time and resources. It takes time because symmetry is a spatially global feature. It does not 'pop-out'. So, if a predator begins with an analysis of the symmetry of a cryptic pattern on the prey's body, the prey may have time to escape.

All of the perceptual mechanisms described in this section operate in the human vision system, but it is not clear at the time of writing as to which, if any, other animals share these mechanisms. The fact that the camouflage widely used by animals can be explained in terms of a known human visual system's mechanisms suggests that visual perceptions of animals are very similar to the humans.

5. VISUAL SEARCH: FEATURES ACROSS THE SCENE

(a) Search image and search target

Thus far, we have considered the function of simple neural units which can respond to changes in the optical array such as may be caused by edges of important objects. However, most natural environments contain other objects and textures which may not be important to the perceiver. For example, the perceiver may wish to locate an edible item located somewhere among (inedible) foliage. This simple, but ubiquitous, problem has been of fundamental importance in vision science. The problem is called 'visual search'. A typical experiment investigates the perceiver's ability to detect the presence of a 'target' among other elements called 'distractors'. The participant has to signal whether the target is present or absent on a given trial. The dependent variables are usually the reaction time for a response, and the accuracy of the responses. Where the target is easy to find (e.g. a bright red item among green items), the response time is independent of the number of distractors and the search is said to be 'efficient'. In an efficient search, it is not necessary to inspect each part of the image to find out whether the target is present. Alternatively, inefficient search results in increases of response time with increasing numbers of distractors. The latter search typically necessitates detailed inspection of several parts of the scene before a decision is reached. Efficient and inefficient searches

are therefore distinguished by the slope of the function relating response time to number of items, called the 'search slope' (see below).

Before discussing visual search, it is necessary to distinguish the research domain of visual search from the concept of a 'search image' used by ecologists (see Tinbergen 1960 and Dawkins 1971 for a definition; but also see Lawrence & Allen 1983 for a useful clarification). Briefly, a search image is an internal representation of the prey species, or some characteristic of the prey, which is used to aid its detection. For cryptic species, this image may consist solely of the telltale cues that camouflage has failed to conceal. Other behavioural habits that might influence predation rates are specifically excluded; these include biases to specific locations and learning behaviours that might increase the likelihood of capturing particular prey (Dawkins 1971; Krebs 1973). The stated aim of visual search is to investigate attentional mechanisms underlying the detection of target items. It is an implicit assumption of this paradigm that the observer must have some internal representation of the visual characteristics of the target object, and some description of the physical properties that allow its selection from a background of different objects. One might conclude that the search image defined by ecologists must be equivalent to the internal representation of the *target* sought by participants in visual-search experiments.

Given the nature of the visual search task, such studies are likely to be relevant to our understanding of camouflage. The target may be defined by various features such as shape, colour, texture or movement, or in 'conjunction' search the target may be defined by combinations (figure 3*b*) of the aforementioned features (Treisman 1988). Distractors will vary in their similarity to the target. The number of distractors within a stimulus is generally manipulated in order that a search slope (ms per number of distractors) can be calculated (figure 3*c*). Interest has centred on search efficiency (e.g. Treisman & Gelade 1980), which in turn relies upon measurements of search slope. Initially, it was presumed that preattentive search must be based upon visual properties available in early visual processing areas, such as colour, luminance and orientation. However, later

studies have demonstrated that complex scene properties can also pop-out, for example targets with differences that can only be based upon object properties, rather than low-level features such as lines and shading, can be detected with apparently pre-attentive levels of efficiency (Ramachandran 1988; Enns & Rensink 1990*a*,*b*). A high search slope (inefficient search) typically results when the scene contains more items that resemble the target. This is exactly the situation that camouflaged items are trying to achieve.

It is easy to see how studies of visual search should inform our understanding of camouflage; however, the majority of search studies have used very simple, synthetic stimuli with backgrounds consisting of punctate elements rather than a continuous, complex, visual environment (Wolfe 1994*a*). Targets and distractors tend to be capital letters or simple geometric shapes (e.g. Treisman 1988). In real-world environments, where organisms seek to camouflage themselves, the visual world is a continuous array of overlapping objects and textures (see Rosenholtz *et al.* 2007 for a useful summary of the differences between traditional visual search stimuli and real-world scenes). Traditionally, interest has centred upon search efficiency. The degree of efficiency of search is usually expressed as a search slope—defined as the increase in response time when one further distractor is added to the scene. Search slopes around zero indicate efficient search, whereas search slopes approximately 60 ms per item (in humans) indicate inefficient search. There is a continuum of search efficiencies between these two extremes. Increasing inefficiency is thought to result in a greater need to deploy attentional resources to various parts of the scene, resulting in a (partly) serial inspection strategy.

Apart from a few exceptions (notably, the feature congestion model, Rosenholtz *et al.* 2007), models that attempt to predict visual search speed tend to take the number of distractors as a known quantity (e.g. the Guided Search Model, Wolfe 1994*b*)—something that would be difficult to define in a natural scene.

Duncan & Humphreys (1989, 1992) formalized the effect of distractor–distractor heterogeneity upon vision search times. As heterogeneity increases, search times become slower, but only where the target bears some similarity to the background. How does this relate to camouflage? Essentially, camouflaged objects, i.e. objects that aim to look similar to their background, will be harder to find if the background itself is more heterogeneous; so if you are a moth trying to hide among leaves, then you would be better choosing a plant that has more variable leaves. Figure 3 (electronic supplementary material) illustrates the search surface described by Duncan & Humphreys. It is evident that the search is the most difficult, i.e. the search-slope is the steepest, where (*a*) the target is similar to the distractors, and (*b*) the distractors are heterogeneous.

(b) Search in natural scenes

In a recent study, Lovell *et al.* (2008) used photographs of natural objects, pebbles, as targets and distractors. In another trial, observers were asked to locate one of four target pebbles hidden among distractors (figure 4 in the electronic supplementary material); both were drawn from a population of 180. Observers were asked to indicate whether the target pebble in any particular trial was to the left or to the right of the centre of the stimulus—there were no target-absent trials. Stimuli featured 4, 9 or 14 randomly selected distractors; consequently, there should be a range of target–distractor and distractor–distractor differences. While the stimulus still featured a uniform background and punctate objects, the stimulations predicting the observer reaction times were based upon examination of the whole stimulus image, so arguably the model should generalize to search with more natural scenes. Target–stimulus difference was calculated by estimating the visual difference of the target pebble from the scene as a whole. This was achieved using a visual-difference predictor (VDP) model of contrast encoding by cells in primate visual cortex (Párraga *et al.* 2005; Lovell *et al.* 2006). The output of the VDP model results in an 18-dimensional array of difference maps (the product of six spatial frequencies and the three chromatic opponent channels). By estimating the Euclidian distance between the differences at each pixel location, it is possible to achieve an approximate measure of heterogeneity. If two vectors point in different directions, then it is likely that these regions of the original image are different; in other words, these image regions are heterogeneous. Small target-stimulus differences and large target-stimulus ones are summed separately and along with the heterogeneity measure are fed into a neural network. Following training with cross validation, the neural net was able to successfully predict observer reaction times ($r = 0.68$). Finally, the results demonstrated that the Duncan & Humphreys prediction (figure 3 in the electronic supplementary material) of the influence of distractor–distractor heterogeneity upon the shape of the search surface was confirmed, even for search among natural objects.

6. CONCLUSIONS

This review has concentrated on those optical and visual processes that appear to be central to an understanding of visual concealment, but which have rarely been considered in the literature on concealment. Key properties of the light environment, and its sensing by neural systems, suggest that the encoding of certain discontinuities (in pattern and motion, i.e. in space and time) is central to the encoding of complex scenes. Principles of grouping and pattern allow the two-dimensional retinal sampling to be translated into three-dimensional structures. Finally, these structures need to be found in complex, cluttered scenes. This last area is probably one in which least progress has been made to date. We have tried to indicate possible ways in which this could be understood, but the work in this area is least advanced.

REFERENCES

Adelson, E. H. & Movshon, J. T. 1982 Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525. (doi:10.1038/300523a0)

Anderson, A. J. & McOwan, P. W. 2003*a* Humans deceived by predatory stealth strategy camouflaging motion. *Proc. R. Soc. B* **270**, S18–S20. (doi:10.1098/rsbl.2003.0030)

Anderson, A. J. & McOwan, P. W. 2003*b* Model of a predatory stealth behaviour camouflaging motion. *Proc. R. Soc. B* **270**, 489–495. (doi:10.1098/rspb.2002.2259)

Autrum, H., Jung, R., Loewenstein, W. R., Mackay, D. M. & Teuber, H. L. 1973. *Handbook of sensory physiology*, vol. VII/5. New York, NY: Springer.

Barlow, H. B., Hill, R. M. & Levick, W. R. 1964 Retinal ganglion cells responding selectively to direction + speed of image motion in rabbit. *J. Physiol.* (*Lond.*) **173**, 377–407.

Barnard, K., Finlayson, G. & Funt, B. 1997 Colour constancy for scenes with varying illumination. *Comput. Vis. Image Underst.* **65**, 311–321. (doi:10.1006/cviu.1996.0567)

Basri, R. & Ullman, S. 1993 The alignment of objects with smooth surfaces. *CVGIP: Image Underst.* **57**, 331–345. (doi:10.1006/ciun.1993.1022)

Behrens, R. R. 1999 The role of artists in ship camouflage during World War I. *Leonardo* **32**, 53–59. (doi:10.1162/002409499553000)

Biederman, I. 1987 Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* **94**, 115–147. (doi:10.1037/0033-295X.94.2.115)

Blake, R. & Shiffrar, M. 2007 Perception of human motion. *Annu. Rev. Psychol.* **58**, 47–73. (doi:10.1146/annurev.psych.57.102904.190152)

Blakemore, C. & Tobin, E. 1972 Lateral inhibition between orientation detectors in cat's visual cortex. *Exp. Brain Res.* **15**, 439–440. (doi:10.1007/BF00234129)

Bonds, A. B. 1989 Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis. Neurosci.* **2**, 41–55.

Born, R. T. & Bradley, D. C. 2005 Structure and function of visual area MT. *Annu. Rev. Neurosci.* **28**, 157–189. (doi:10.1146/annurev.neuro.26.041002.131052)

Bowmaker, J. K. & Hunt, D. M. 2006 Evolution of vertebrate visual pigments. *Curr. Biol.* **16**, R484–R489. (doi:10.1016/j.cub.2006.06.016)

Canny, J. 1986 A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J. I. & Rust, N. C. 2005 Do we know what the early visual system does? *J. Neurosci.* **25**, 10 577–10 597. (doi:10.1523/JNEUROSCI.3726-05.2005)

Cavanaugh, J. R., Bair, W. & Movshon, J. A. 2002*a* Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J. Neurophysiol.* **88**, 2530–2546. (doi:10.1152/jn.00692.2001)

Cavanaugh, J. R., Bair, W. & Movshon, J. A. 2002*b* Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J. Neurophysiol.* **88**, 2547–2556. (doi:10.1152/jn.00693.2001)

Cuthill, I. C., Stevens, M., Sheppard, J., Maddocks, T., Párraga, C. A. & Troscianko, T. 2005 Disruptive coloration and background pattern matching. *Nature* **434**, 72–74. (doi:10.1038/nature03312)

Dawkins, M. 1971 Perceptual changes in chicks: another look at the 'search image' concept. *Anim. Behav.* **19**, 566–574. (doi:10.1016/S0003-3472(71)80113-6)

Dayan, E., Casile, A., Levit-Binnun, N., Giese, M. A., Hendler, T. & Flash, T. 2007 Neural representations of kinematic laws of motion: evidence for action–perception coupling. *Proc. Natl Acad. Sci. USA* **104**, 20 582–20 587. (doi:10.1073/pnas.0710033104)

Dittrich, W. H., Troscianko, T., Lea, S. E. G. & Morgan, D. 1996 Perception of emotion from dynamic point-light displays represented in dance. *Perception* **25**, 727–738. (doi:10.1068/p250727)

Duncan, J. & Humphreys, G. W. 1989 Visual search and stimulus similarity. *Psychol. Rev.* **96**, 433–458. (doi:10.1037/0033-295X.96.3.433)

Duncan, J. & Humphreys, G. W. 1992 Beyond the search surface: visual search and attentional engagement. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 578–588. (doi:10.1037/0096-1523.18.2.578)

Enns, J. T. & Rensink, R. A. 1990*a* Influence of scene-based properties on visual-search. *Science* **247**, 721–723. (doi:10.1126/science.2300824)

Enns, J. T. & Rensink, R. A. 1990*b* Sensitivity to 3-dimensional orientation in visual-search. *Psychol. Sci.* **1**, 323–326. (doi:10.1111/j.1467-9280.1990.tb00227.x)

Field, D. J. & Tolhurst, D. J. 1986 The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc. R. Soc. B* **228**, 379–400. (doi:10.1098/rspb.1986.0060)

Finkelstein, D. & Grüsser, O. J. 1965 Frog retina—detection of movement. *Science* **150**, 1050–1051. (doi:10.1126/science.150.3699.1050)

Finlayson, G. D. & Funt, B. V. 1994 Color constancy using shadows. *Perception* **23**, 89–90.

Ghose, K., Horiuchi, T. K., Krishnaprasad, P. S. & Moss, C. F. 2006 Echolocating bats use a nearly time-optimal strategy to intercept prey. *PLOS Biol.* **4**, 865–873. (doi:10.1371/journal.pbio.0040108)

Gilbert, C. D. 1977 Laminar differences in receptive field properties of cells in cat primary visual cortex. *J. Physiol.* **268**, 391–421.

Glendinning, P. 2004 The mathematics of motion camouflage. *Proc. R. Soc. B* **271**, 477–481. (doi:10.1098/rspb.2003.2622)

Grosof, D. H., Shapley, R. M. & Hawken, M. J. 1993 Macaque VI neurons can signal 'illusory' contours. *Nature* **365**, 550–552. (doi:10.1038/365550a0)

Grossberg, S., Mingolla, E. & Ross, W. D. 1997 Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends Neurosci.* **20**, 106–111. (doi:10.1016/S0166-2236(96)01002-8)

Harmon, L. D. & Julesz, B. 1973 Masking in visual recognition: effects of two-dimensional filtered noise. *Science* **180**, 1194–1197. (doi:10.1126/science.180.4091.1194)

Heeger, D. J. 1992 Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* **9**, 181–197.

Hiris, E. 2007 Detection of biological and nonbiological motion. *J. Vis.* **7**, 1–16. (doi:10.1167/7.12.4)

Hubel, D. H. & Wiesel, T. N. 1959 Receptive fields of single neurones in the cats striate cortex. *J. Physiol.* (*Lond.*) **148**, 574–591.

Hubel, D. H. & Wiesel, T. N. 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154.

Hubel, D. H. & Wiesel, T. N. 1965 Receptive fields and functional architecture in two nonstriate areas (18 and 19) of the cat. *J. Neurophysiol.* **28**, 229–289. (doi:10.1159/000161635)

Ivanenko, Y. P., Grasso, R., Macellari, V. & Lacquaniti, F. 2002 Two-thirds power law in human locomotion: role of ground contact forces. *Neuroreport* **13**, 1171–1174. (doi:10.1097/00001756-200207020-00020)

Johansson, G. 1973 Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **14**, 201–211.

Jones, J. P. & Palmer, L. A. 1987 An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**, 1233–1258.

Julesz, B. 1971 *Foundations of cyclopean perception.* Chicago, IL: University of Chicago Press.

Justh, E. W. & Krishnaprasad, P. S. 2006 Steering laws for motion camouflage. *Proc. Math. Phys. Eng. Sci.* **462**, 3629–3643. (doi:10.1098/rspa.2006.1742)

Krebs, J. R. 1973 Behavioural aspects of predation. In *Perspectives in ethology* (eds P. P. G. Bateson & P. H. Klopfer), pp. 73–111. New York, NY: Plenum Press.

Lamme, V. A. 1995 The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15**, 1605–1615.

Lamme, V. A. F. 2003 Why visual attention and awareness are different. *Trends Cogn. Sci.* **7**, 12–18. (doi:10.1016/S1364-6613(02)00013-X)

Land, M. F. & Nilsson, D.-E. 2001 *Animal eyes.* Oxford, UK: Oxford University Press.

Lauritzen, J. S. & Tolhurst, D. J. 2005 Contrast constancy in natural scenes in shadow or direct light—a proposed role for contrast-normalisation (non-specific suppression) in visual cortex. *Netw. Comput. Neur. Syst.* **16**, 151–173. (doi:10.1080/03091920500304582)

Lawrence, E. S. & Allen, J. A. 1983 On the term 'search image'. *Oikos* **40**, 313–314. (doi:10.2307/3544597)

Li, Y., Pizlo, Z. & Steinman, R. M. In press. A computational model that recovers the 3D shape of an object from a single 2D retinal representation. *Vis. Res.* (doi:10.1016/j.visres.2008.05.013)

Longuet-Higgins, H. C. 1981 A computer algorithm for reconstructing a scene from two projections. *Nature* **293**, 133–135. (doi:10.1038/293133a0)

Lovell, P. G., Tolhurst, D. J., Párraga, C. A., Baddeley, R., Leonards, U., Troscianko, J. & Troscianko, T. 2005 On the stability of the color-opponent signals under changes of illuminant in natural scenes. *J. Opt. Soc. Am. A* **22**, 2060–2071. (doi:10.1364/JOSAA.22.002060)

Lovell, P. G., Párraga, C. A., Ripamonti, C., Troscianko, T. & Tolhurst, D. 2006 Evaluation of a multi-scale color model for visual difference prediction. *Trans. Appl. Percept.* **3**, 155–178. (doi:10.1145/1166087.1166089)

Lovell, P. G., Gilchrist, I. D., Tolhurst, D. J., To, M. & Troscianko, T. 2008 Predicting search efficiency with a low-level visual difference model. *J. Vis.* **8**, 1082–1082.

Lowe, D. G. 1985 *Perceptual organization and visual recognition.* Boston, MA: Kluwer.

Maffei, L. & Fiorentini, A. 1973 The visual cortex as a spatial frequency analyzer. *Vis. Res.* **13**, 1255–1267. (doi:10.1016/0042-6989(73)90201-0)

Marr, D. 1969 A theory of cerebral cortex. *Proc. R. Soc. B* **174**, 161–234.

Marr, D. 1982 *Vision.* San Francisco, CA: W. H. Freeman.

Marr, D. & Hildreth, E. 1980 Theory of edge detection. *Proc. R. Soc. B* **207**, 187–217. (doi:10.1098/rspb.1980.0020)

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K. & Tootell, R. B. H. 1999 The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J. Neurosci.* **19**, 8560–8572.

Mizutani, A., Chahl, J. S. & Srinivasan, M. V. 2003 Insect behaviour: motion camouflage in dragonflies. *Nature* **423**, 604. (doi:10.1038/423604a)

Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. 1978 Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.* **283**, 53–77.

Mundy, J. L. & Zisserman, A. 1992 *Geometric invariance in computer vision.* Cambridge, MA: MIT Press.

Orban, G. A., Sunaert, S., Todd, J. T., Van Hecke, P. & Marchal, G. 1999 Human cortical regions involved in extracting depth from motion. *Neuron* **24**, 929–940. (doi:10.1016/S0896-6273(00)81040-5)

Osorio, D. & Vorobyev, M. 1996 Colour vision as an adaptation to frugivory in primates. *Proc. R. Soc. B* **263**, 593–599. (doi:10.1098/rspb.1996.0089)

Pack, C. C., Livingstone, M. S., Duffy, K. R. & Born, R. T. 2003 End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. *Neuron* **39**, 671–680. (doi:10.1016/S0896-6273(03)00439-2)

Párraga, C. A., Troscianko, T. & Tolhurst, D. J. 2002 Spatio-chromatic properties of natural images and human vision. *Curr. Biol.* **12**, 483–487. (doi:10.1016/S0960-9822(02)00718-2)

Párraga, C. A., Troscianko, T. & Tolhurst, D. J. 2005 The effects of amplitude–spectrum statistics on foveal and peripheral discrimination of changes in natural images, and a multi-resolution model. *Vis. Res.* **45**, 3145–3168. (doi:10.1016/j.visres.2005.08.006)

Pizlo, Z. 2008 *3D shape: its unique place in visual perception.* Cambridge, MA: MIT Press.

Pizlo, Z. & Loubier, K. 2000 Recognition of a solid shape from its single perspective image obtained by a calibrated camera. *Pattern Recognit.* **33**, 1675–1681. (doi:10.1016/S0031-3203(99)00170-3)

Porrill, J., Pollard, S., Pridmore, T. P., Bowen, J. B., Mayhew, J. E. W. & Frisby, J. P. 1988 TINA: a 3D vision system for pick and place. *Image Vis. Comput.* **6**, 91–99. (doi:10.1016/0262-8856(88)90004-2)

Ramachandran, V. S. 1988 Perception of shape from shading. *Nature* **331**, 163–166. (doi:10.1038/331163a0)

Reddy, P. V., Justh, E. W. & Krishnaprasad, P. S. 2007 Motion camouflage with sensorimotor delay. In *Proc. 46th IEEE Conf. on Decision and Control*, vols 12–14, pp. 1660–1665. (doi:10.1109/CDC.2007.4434522)

Regan, B. C., Julliot, C., Simmen, B., Vienot, F., Charles-Dominique, P. & Mollon, J. D. 2001 Fruits foliage and the evolution of primate colour vision. *Phil. Trans. R. Soc. B* **356**, 229–283. (doi:10.1098/rstb.2000.0773)

Ringach, D. L. 2002 Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *J. Neurophysiol.* **88**, 455–463.

Rosenholtz, R., Li, Y. Z. & Nakano, L. 2007 Measuring visual clutter. *J. Vis.* **7**, 1–22.

Sawada, T. & Pizlo, Z. 2008 Detecting mirror-symmetry of a volumetric shape from its single 2D image. In *Proc. Workshop on Perceptual Organization in Computer Vision, IEEE Int. Conf. Computer Vision and Pattern Recognition, Anchorage, Alaska, June 23.*

Schwartz, O. & Simoncelli, E. P. 2001 Natural signal statistics and sensory gain control. *Nat. Neurosci.* **4**, 819–825. (doi:10.1038/90526)

Shohet, A. J., Baddeley, R. J., Anderson, J. C., Kelman, E. J. & Osorio, D. 2006 Cuttlefish responses to visual orientation of substrates, water flow and a model of motion camouflage. *J. Exp. Biol.* **209**, 4717–4723. (doi:10.1242/jeb.02580)

Smyth, D., Willmore, B., Thompson, I. D., Baker, G. E. & Tolhurst, D. J. 2003 The receptive-field organisation of simple cells in primary visual cortex (V1) of ferrets under natural scene stimulation. *J. Neurosci.* **23**, 4746–4759.

Srinivasan, M. V. & Davey, M. 1995 Strategies for active camouflage of motion. *Proc. R. Soc. B* **259**, 19–25. (doi:10.1098/rspb.1995.0004)

Stevens, M. & Cuthill, I. C. 2006 Disruptive coloration, crypsis and edge detection in early visual processing. *Proc. R. Soc. B* **273**, 2141–2147. (doi:10.1098/rspb.2006.3556)

Steverding, D. & Troscianko, T. 2004 On the role of blue shadows in the visual behaviour of tsetse flies. *Proc. R. Soc. B* **271**, S16–S17. (doi:10.1098/rsbl.2003.0121)

Thompson, D. W. 1992 *On growth and form*. New York, NY: Dover.

Tinbergen, L. 1960 The natural control of insects in pine woods. 1. Factors influencing the intensity of predation by songbirds. *Arch. Neerl. Zool.* **13**, 265–343.

Tolhurst, D. J. 1972 On the possible existence of edge detectors in the human visual system. *Vis. Res.* **12**, 797–804. (doi:10.1016/0042-6989(72)90006-5)

Tolhurst, D. J. & Dean, A. F. 1987 Spatial summation by simple cells in the striate cortex of the cat. *Exp. Brain Res.* **66**, 607–620. (doi:10.1007/BF00270694)

Tolhurst, D. J. & Heeger, D. J. 1997 Comparison of contrast–normalization and threshold models of the responses of simple cells in cat striate cortex. *Vis. Neurosci.* **14**, 293–309.

Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R. & Belliveau, J. W. 1995 Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J. Neurosci.* **15**, 3215–3230.

Treisman, A. 1988 Features and objects—the 14th Bartlett memorial lecture. *Q. J. Exp. Psychol. A* **40**, 201–237.

Treisman, A. M. & Gelade, G. 1980 Feature-integration theory of attention. *Cognit. Psychol.* **12**, 97–136. (doi:10.1016/0010-0285(80)90005-5)

Ullman, S. 1979 *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J. T. & Orban, G. A. 2002 Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science* **298**, 413–415. (doi:10.1126/science.1073574)

von der Heydt, R., Peterhans, E. & Baumgartner, G. 1984 Illusory contours and cortical neuron responses. *Science* **224**, 1260–1262. (doi:10.1126/science.6539501)

Walls, G. L. 1942 *The vertebrate eye and its adaptive radiation*. New York, NY: Hafner.

Ward, G. & Shakespeare, R. 2004 *Rendering with radiance: the art and science of lighting visualization*. Charleston, SC: Booksurge Press.

Weckstrom, M. & Laughlin, S. B. 1995 Visual ecology and voltage-gated ion channels in insect photoreceptors. *Trends Neurosci.* **18**, 17–21. (doi:10.1016/0166-2236(95)93945-T)

Weiss, I. 1993 Geometric invariants and object recognition. *Int. J. Comput. Vis.* **10**, 207–231. (doi:10.1007/BF01539536)

Weisstein, N. & Bisaha, J. 1972 Gratings mask bars and bars mask gratings: visual frequency response to aperiodic stimuli. *Science* **176**, 1047–1049. (doi:10.1126/science.176.4038.1047)

Wolfe, J. M. 1994*a* Guided search 2.0—a revised model of visual-search. *Psychon. Bull. Rev.* **1**, 202–238.

Wolfe, J. M. 1994*b* Visual-search in continuous, naturalistic stimuli. *Vis. Res.* **34**, 1187–1195. (doi:10.1016/0042-6989(94)90300-X)

Wuerger, S., Shapley, R. & Rubin, N. 1996 'On the visually perceived direction of motion' by Hans Wallach: 60 years later. *Perception* **25**, 1317–1367. (doi:10.1068/p251317)

Zeki, S., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C. & Frackowiak, R. S. J. 1991 A direct demonstration of the functional specialisation in the human visual cortex. *J. Neurosci.* **11**, 641–649.

Zhou, H., Friedman, H. S. & von der Heydt, R. 2000 Coding of border ownership in monkey visual cortex. *J. Neurosci.* **20**, 6594–6611.