The role of monocularly visible regions in depth and surface perception

Julie M. Harris[1] **and Laurie M. Wilcox**[2]

1 School of Psychology, University of St. Andrews, South St., St. Andrews, KY16 9JP, Scotland

2 Department of Psychology Centre for Vision Research 4700 Keele St., North York Ontario, M3J 1P3 Canada

Abstract

The mainstream of binocular vision research has long been focused on understanding how binocular disparity is used for depth perception. In recent years, researchers have begun to explore how monocular regions in binocularly viewed scenes contribute to our perception of the three-dimensional world. Here we review the field as it currently stands, with a focus on understanding the extent to which the role of monocular regions in depth perception can be understood using extant theories of binocular vision.

Key words: stereopsis, binocular vision, da Vinci Stereopsis, monocular regions, half-occlusion, monocular

For publication in Vision Research.

(1) Introduction

(1.1) The problem: how do we get a 3-D view when the eyes see different things?

Because we have two forward facing eyes that are separated in the head, our visual system continuously receives two slightly different views of the world. For many decades scientists have been trying to understand how the visual system deals with the slight differences in the images that result from the lateral separation of the two eyes (since Wheatstone, 1838). Some have argued that the differences are so small that they are irrelevant. Instead, they suggest that the primary reason for having two eyes is for binocular concordance: that is, increasing visual efficiency (reducing noise) by having effectively 'a second go' at viewing each scene (e.g. Jones & Lee, 1982). This approach makes sense when one views a distant scene, as the two views are very similar.

However, much of the time we function in environments in which there are multiple objects at close range, and for such scenes, there are substantial differences in the retinal images. Rather than emphasising concordance, a successful approach has been to consider the extent to which the visual system can exploit discordance by measuring the tiny differences between the two eyes' views (binocular disparity), and by using disparity to represent the three-dimensional (3D) structure of the world. A rich literature has shown that binocular disparity is used by the visual system, and that we are exquisitely sensitive to it (e.g. Howard & Rogers, 2002). Measuring the differences between locations in the two eye's views first requires a solution to the binocular correspondence problem. That is, which point on one retina matches with a given point on the other. This is a potentially complicated operation given the complex 3-D structure of the natural environment, but it is made even more difficult by the fact that not all image points in one eye have a partner in the other. For example, Figure 1a shows a natural scene viewed through an occluding foreground 'fence'. We refer to this scene as the background because it is the furthest thing visible in the display. Figure 1b illustrates how there are features in one eye that are not present in the other, and vice versa. Notice that different portions of the background scene are occluded in each eye (in this extreme example none of the background is visible to both eyes).

For many years, researchers studying stereopsis treated these monocular regions as 'noise': a potential source of false matches and ambiguity, and therefore an obstacle to binocular disparity processing. Since the work of Gillam and Borsting (1988), and now many others, the prevailing view has changed. We have come to realize that these monocular regions or half-occlusions are useful, and play a potentially important role in binocular depth perception.

This review focuses on the utility of monocular regions for perception, discussing the relatively scant literature that has explored their utility, and how they are processed by the visual system. After briefly reviewing the history of our understanding of the phenomenon and the importance of monocular regions, several key topics will be explored. Our main aim in this review is to consider how information from monocular regions may be used for depth and surface perception. Many examples show that there appears to be more going on in depth perception than traditional theories of binocular stereopsis can account for. We will discuss many of those examples here and show that only a few provide major challenges to our understanding of binocular stereopsis. Second, we consider how monocular and binocular information are brought together to form a stable representation of the world. We go on to review work showing that monocular regions have a role in the perception of surfaces. Finally, we consider some recent biologically inspired models of the utility of monocular regions in depth perception. Our aim throughout is to explore the extent to which the monocular regions deliver useful information for depth and surface perception, and whether our perceptions can be accounted for using some common binocular visual mechanisms.

(1.2) History

The study of vision has a long history, stretching back to the ancient Greeks and Arabs. Howard (2002) gives an excellent introduction to the history of vision in general, and discusses binocular vision and occlusion in particular. He describes how Euclid, around 300 BC, first outlined the geometry of binocular vision and the fact that the two eyes see different parts of a sphere. Some 500 years later, Galen noted that when a foreground object is viewed, parts of objects lying behind that object are only seen by one eye. This point was later developed visually in a series of drawings from Leonardo da Vinci (Wade, Ono & Lillakas, 2001; Richter, 1977) who was the first to note that depth perception can arise when each eye sees different parts of an object. Some of Leonardo's drawings also illustrate that when looking through a hole, there are regions of the background scene that are only visible to one eye (see Strong, 1979).

Many of the beautiful and challenging examples of monocular regions that deliver depth perception in binocular vision, that we will review below, were anticipated and demonstrated by von Szily (1921, translated by Ehrenstein & Gillam, 1998). Figure 2 shows two examples of his demonstrations. Lawson and Gulick (1967) first demonstrated experimentally that monocular regions of a scene can deliver a perception of depth akin to that from stereopsis. Many important issues in this field were discussed in a PhD thesis (otherwise unpublished) by Barrand (1979). Kaye (1978) was the first to show that a sensation of depth can be obtained from viewing an isolated point monocularly, and that its perceived depth depends on its location in on the retina.

The first experimental evidence that monocular regions could specifically aid depth perception from binocular disparity came from Gillam's lab in the 1980's. She hypothesised that monocular regions could be used to identify the location of depth edges. This hypothesis was supported by evidence that perception of a depth edge occurs faster when explicitly textured monocular regions are present, than when they are left blank and the same colour as the surround (Gillam and Borsting, 1988). Perhaps the most well known study in this area is that which coined the term 'da Vinci stereopsis' to refer to the use of monocular regions in depth perception (Nakayama & Shimojo, 1990). In this extensive study of depth from monocular regions the authors used very simple stimuli, in which one eye's view contained a bar that was not visible to the other eye (figure 3). When a stimulus is designed so that both eyes view the rectangle, and the right eye views an additional vertical bar to the right of the rectangle, the observer perceives the bar to lie behind the rectangle. This is consistent with viewing geometry as shown in figure 3a, where the bar is occluded by the closer rectangle, in the left eye's view.

Though known of for many years, monocular regions have been by-passed by much of the binocular vision community. Instead, there has been a focus on the significant computational problem of extracting the correct binocular disparity information from a pair of disparate images. This problem has perhaps most famously been set out by Marr, who used it as an exemplar of how a computational approach could yield a richer understanding of the mechanisms and algorithms underlying perception (Marr & Poggio, 1976, 1979). His approach, and that of many modellers of both human and machine vision since, was to initially break the problem into two parts: first corresponding image points must be identified between the two eyes views, and then the disparity can be extracted. Over the years, a variety of rules, or constraints, have been suggested to guide this process, including the assumption that each point has a single match (uniqueness constraint) and that the world is composed of piece-wise smooth objects (smoothness constraint), and so on. The foundation of all such work is that *retinal points in the two eyes must be matched*. Points for which no match is found are rejected: they do not form part of the modelled scene. Clearly, monocular regions will pose problems for such approaches to disparity processing because there simply is no correct binocular correspondence. Not surprisingly, monocular regions are known to challenge many models of depth extraction, providing a source of noise, rather than a useful signal. Some recent developments in modelling do take monocular regions into account, as outlined by Egnal and Wildes (2002) who compare recent computational models that extract monocular regions. We discuss more recent, physiologically plausible models in section 5 below.

The exclusion of monocular regions from models of binocular disparity processing might be understandable if such phenomena were trivially rare. But in fact, monocular regions are abundant in scenes containing a cluttered foreground. Figure 1a shows a foreground 'picket-fence' occluding a background scene, with figure 1b portraying how much of that scene is visible to each eye's view. In this example, the background scene is only visible to one eye, or the other, never both. Whilst such an extreme example will occur rarely, monocular regions exist in real scenes at all vertical object boundaries where there is a significant depth discontinuity. In fact, such locations are arguably the regions of most interest in a scene, for they indicate where one object ends and the next begins (Gillam & Borsting, 1988; Anderson & Nakayama, 1994). Some recent modelling work has made this proposal explicit. For instance, Langer (2008) developed an artificial world in which square or spherical objects were randomly distributed through a volume (emulating natural cluttered scenes, like foliage or tree branches). He noted that occlusion in one eye's view of a given point (resulting in a monocular region visible to the other eye) becomes increasingly more likely as its distance from the observer increases. In a world densely populated with objects, the further away a point is from the observer, the more likely it is that there will be an occluding object along any particular line of sight. Another very recent paper has taken this idea a step further. Changizi and Shimojo (2008) suggest that the main reason why forward-facing eyes have evolved is not for stereopsis, but rather to take advantage of the increased proportion of background objects that can be viewed in cluttered scenes with two eyes. As we describe below, this hypothesis cannot account for the apparent utility of monocular regions in depth perception, so it surely cannot represent the whole story. However, the theoretical position adopted does illustrate the potential importance of monocularly visible regions for vision.

(1.3) Geometry and classification of monocular regions

In the real world objects are located at different depths. When viewed from certain locations, foreground objects result in partial occlusion of objects that are further away. These are different for each eye, because the eyes are laterally separated, which results in monocular regions of the scene; that is, regions that are visible to one eye only. There are several different ways in which this can occur. Here, we classify monocular regions into 3 types, that depend on the arrangement and features of the objects being viewed, as well as the relative location of the observer and her eyes. All the studies to be described in this review use stimuli that conform to one of these 3 types.

Type 1: monocular background

If foreground occluders are of a specific size and at a specific distance from a background scene, there will be regions of the images where the eyes are delivered completely different patterns and where there are no binocular regions that contain those patterns. Figure 1 is an example of this kind of stimulus, where the foreground occluders are just the right width, ensuring that no part of the background is simultaneously visible to both eyes (see also figures 4 and 10). This situation is rare in natural viewing, and although parts of these stimuli can appear rivalrous, scenes are perceived as stable and with a depth difference between foreground and background (Forte et al, 2002; Howard, 1995, Tsai & Victor, 2000). Phenomena linked to this configuration will be discussed in section 3.2.

Type 2: binocularly visible foreground

Figure 3a shows a top-down view of a foreground rectangle and featureless background (featureless regions are indicated by dashed lines in our figures). Any object positioned in the 'right eye only' region will be visible only to the right eye. Figure 3c shows stimuli, as used by Nakayama & Shimojo (1990), in their classic study on 'da Vinci stereopsis'. They found that, for this simple configuration, the bar is perceived as lying behind a foreground rectangle. In many real-world situations the foreground and background will both be visible, and differently textured or coloured. Under such conditions monocular regions are seen at the depth of the background.

Type 3: invisible foreground
Figure 3b shows a situation in which an observer views a background (featureless except for the binocularly visible rectangle) through a hole in a featureless foreground object. Such a scene is consistent with switching the two eye's views in figure 3c, so that the left eye now views the stimulus containing the monocular bar. This has been called the camouflage configuration (Howard & Rogers, 2002), and can only occur when an object has the same texture and luminance as the background (i.e. is camouflaged) in one eye, but not in the other (first described by Kaye, 1978). Figure 4a shows another example, where a small grey foreground line is camouflaged to the left eye (it occludes a grey section of the background and is therefore invisible to that eye) but not to the right eye (from the right eye's view the background region visible is striped). Stereo-pairs that simulate this configuration are shown in figure 4b. Although cases like this will be rare in the world, because they require the coincidence of identically patterned foreground and background, the literature provides several recent examples that do appear to support depth perception from such images (see section 4.1, figure 10; section 4.3, figure 11).

A key issue to notice when considering these 3 types of occlusion is that Type 2 is much more common in the world than either Types 1 or 3, but that all are consistent with real 3-D scenes, as shown in the figures (this point runs contrary to the original ideas put forward in Nakayama & Shimojo, 1990, and will be taken up in more detail in section 2.3 below). The extent to which these three types of monocular stimulation may be processed via different mechanisms and whether depth perception mechanisms need to 'know' the geometrical constraints that underlie occlusion, is described and discussed in the sections below.

2 Monocular regions and depth perception

A number of different lines of research have explored depth from monocular regions, most of which use distinct stimuli. The challenge is to integrate the diverse effects that have been discovered to form a coherent understanding. At the heart of our review is the question of whether the visual processing of monocular regions is distinctly different from classical stereopsis. Below we will examine some of the studies that have used stimuli containing monocular regions, where the monocular regions appear to be involved in depth perception. Some of these are directly aimed at understanding how depth is perceived from monocular regions, while others are related to this field by nature of the stimuli, but not by intent. We will consider these in separate sections reviewing what is widely accepted, what is controversial and what remains a puzzle.

2.1 Speed of depth perception

A monocular region is present whenever there is a significant depth discontinuity. Perhaps the simplest way that a monocular zone could provide information about depth is to signal the location of a depth discontinuity. Some work has demonstrated that monocular regions speed up depth processing. Saye and Frisby (1975) found that, for large disparities, monocular features did speed up depth detection in some configurations. Gillam & Borsting (1988) found that depth discontinuities were more rapidly detected when the appropriate monocular zone was filled with the same pattern as the immediate surrounding background, than if it was left blank. Grove & Ono (1999) explored whether longer latencies occurred because a monocular region was missing, or because it was different from the background pattern (a possible, but unlikely, real configuration). They found that response latencies were longer when the monocular region was differently patterned from the background, but not when it was missing, apparently contradicting the Gillam and Borsting study. Anderson and Nakayama (1994) provided evidence, and conceptual models, that went one step further: suggesting that monocular regions not only signal the location of depth discontinuities, but also help constrain the stereoscopic matching process.

A recent study used photographs of real objects in which half-occluded regions of boxes could be present, or absent (Wilcox & Lakra, 2007). Observers were asked to decide whether they were viewing scenes with correct depth configurations (where disparity information was congruent with depth from perspective, texture etc) or whether the disparity information had been reversed. For richly textured scenes reaction times were faster when monocular regions were present than when removed, but only for scenes where the disparity was congruent with other depth cues. This suggests that occlusion geometry must be consistent with other cues to depth for the rapid perception of depth ordering. One point to note about the stimuli used for this study is that the monocular regions were ‘self-occluded’ regions that were part of the object, rather than part of the flat background wall. Self-occlusions have been very rarely studied in detail but see work on contour stereopsis, e.g. Nefs, 2008). It is not known if monocular regions due to self-occlusion are processed differently from other forms of monocular region.

Overall, it appears that even though latency effects are subject to large individual differences and may be specific to particular stimulus characteristics, they suggest a facilitatory role for monocular regions in identifying depth discontinuities.

(2.2) Minimum requirements for depth from monocular regions

The simplest possible variant of a stimulus that delivers depth perception from stimulation of one eye is that of monoptic depth, where one eye views a point, or line, and the other a blank screen. Whether this phenomenon can be directly linked to depth from monocular regions is not yet fully understood.

In a systematic study of depth from monocular elements, Kaye (1978) showed that the perceived depth of a monocular element depends on its distance from the fovea. This issue has been studied in more detail recently, and experiments have shown that the phenomenon does not rest on the notion of a simple ‘local sign’. This concept, outlined by Hering (and discussed in detail in Howard, 2002) asserts that each location on the retina, in each eye, encodes a particular direction and relative distance. Wilcox and colleagues (Wilcox, Harris & McKee, 2007) ruled out the local sign account by showing that that no depth is perceived if the non-stimulated eye is patched, rather than viewing a blank screen. They also showed that the depth percept is lost at small disparities, and with eccentric fixation. Taken together their work suggests that the phenomenon is most likely due to a crude binocular mechanism that matches a point in one eye with the line of sight, or fovea, in the other eye.

Traditional stereoscopic mechanisms cannot account for monoptic depth phenomena. However, there is growing evidence for stereoscopic mechanisms that do not conform to our traditional understanding of binocular disparity processing. This topic is reviewed in more detail in a companion article (Wilcox & Allison, 2009). Whilst the conventional stereoscopic mechanism processes fine disparities present in luminance-defined stimuli such as bars and edges, there is at least one other type of disparity mechanism that is able to abstract over fine detail and provide a depth signal for the whole of an object regardless of the similarity of the interocular detail. This is commonly known as 2nd-order stereopsis (but is also referred to as coarse, or envelope, stereopsis). Stereoacuity using the 2nd-order mechanism is much poorer than for 1st order stereopsis, but delivers depth perception for diplopic targets (Wilcox and Hess, 1995) and for patches of uncorrelated noise (Wilcox & Hess, 1996).

Recent experiments by Fukuda, Wilcox, Allison and Howard (2009) have provided evidence for a linkage between monoptic depth and 2nd-order stereopsis. It has long been known that there is a large tolerance to vertical misalignment in stimuli containing binocular disparity (Mitchell, 1969). Fukuda et al (2009) showed that the same patterns of perceived depth are obtained for a single monoptic target as for targets with large vertical offsets. This work opens up a new possibility, perhaps a very simple mechanism based on the responses of binocular mechanisms to monocular regions, or even monoptic elements, could account for some types of depth processing from monocular regions.

(2.3) Does occlusion geometry constrain depth perception for monocular regions?

Nakayama and Shimojo (1990) suggested that the brain's knowledge, or experience, of 3-D occlusion geometry constrains our perceptions of depth from monocular regions. When a monocular region is adjacent to an unambiguous background and foreground (unambiguous due to the presence of shading, texture or colour differences: a type 2 region), its depth interpretation is straightforward: the monocular region is assigned the same depth as the unambiguous background (Julesz, 1971; Collett, 1985; Shimojo & Nakayama, 1990; Anderson & Nakayama, 1994). Further, monocular probe dots are located more reliably at a specific depth when they are located in a monocular region, than when they are placed in a binocularly visible part of a stimulus (Shimojo and Nakayama, 1994). When monocular regions are textured so that they are clearly not part of the background, less perceived depth results than if they have the same texture as the background (Grove et al, 2002).

A more ambiguous example, is one in which the scene is very sparse (figure 3a), and the binocular surface is not specified by a pattern (also Type 2). In principle, the geometry (as shown in figure 3a) dictates that monocular regions should be perceived as lying somewhere behind a foreground occluder but precisely where it should be in depth is not specified. Nakayama and Shimojo (1990) defined the depth constraint zone, as shown in figure 5 (striped region). Any real points or objects lying within that zone will only be seen by one eye. An example monocular point in the right eye (indicated by the solid line in figure 5) could correspond to a real point with a depth anywhere along that eye's line of sight, within the depth constraint zone. The zone extends back to an effectively infinite depth for a large foreground object, but could itself be constrained if the object were smaller, or if a textured background were present.

These geometric constraints raise the issue of whether the depth from monocular points is qualitative or quantitative in nature. That is, whether the perceived depth is somewhere within the constraint zone, but the location cannot be precisely identified, or at a specific location in depth, where the depth is matchable to depth from binocular disparity. This issue was dealt with in some detail by Nakayama & Shimojo (1990). They showed that, for a Type 2 monocular region, the perceived depth of a monocular point can be matched using a stereoscopic probe and that the depth matches are quantitative in nature: the perceived depth of the monocular point increased with increasing separation from the occluding edge. For points located close to the occluder (up to around 30 min arc in their hands, or around 10-15 min in a related study (Hakkinen & Nyman, 1996) the matched depth followed the forward edge of the depth constraint zone, suggesting that the depth assigned was the smallest possible that would be consistent with occlusion geometry. For larger separations the matched depth gradually fell to zero disparity. When the two eye's views were inter-changed (resulting in a Type 3 monocular region, figure 3b), Nakayama and Shimojo found that the matched depth was zero disparity, whatever the separation between binocular rectangle and monocular bar. This suggested that there are fundamental differences between the ways in which Type 2 and Type 3 monocular regions are processed in visual perception. Nakayama and Shimojo discussed this difference in terms of viewing geometry, with configurations such as those in figure 3b dubbed 'ecologically invalid', suggesting that they did not correspond to a monocular region that would be present in a real scene. This assertion, and indeed the generality of the result itself, was subsequently challenged.

First, as described in the geometry section above, figure 3b shows a real scene that could correspond to the so-called 'ecologically invalid' case. Although such real scenes may be rare, they are clearly possible. Potential real configurations that correspond to the cases dubbed 'invalid' have been discussed at some length by Assee & Qian (2007). They suggest that the differences in depth perception found in the Nakayama & Shimojo (1990) studies may have more to do with whether the monocular region in the stimulus appears rivalrous (as can occur for these Type 3 configurations and sometimes for Type 1) than with the ecological validity of the stimulus itself.

Second, Hakkinen and Nyman (1996) found no differences in depth perception between configurations very similar to Nakayama & Shimojo's 'valid' and 'invalid' cases. The main difference between the experiments was that stimuli in the Hakkinen & Nyman study also contained an additional binocularly visible plane. These authors showed that the relative depths between that plane and the occluding plane (the one near which the monocular element was placed) affected the perceived location in depth of the monocular element. As suggested by Assee & Qian (2007), it may be that, when the stimulus is particularly ambiguous due to the sparseness of a scene (as in figure 3a and 3b), different observers place different interpretations on why a monocular bar is monocular. If the configuration in 3b was perceived as if behind a featureless occluder, then quantitative depth perception may result. If not, then no depth would be perceived. Adding additional binocularly visible objects to the scene, as Hakkinen & Nyman (1996) did, could result in rather different binocular interpretations. Further evidence that different observers may use different interpretations comes from a recent study in which the perceived depth location (near or far) of a monocular bar could be manipulated by adjusting the pictorial cues of bar size and contrast (Makino & Yano, 2006). Again, scenes were very sparse and observers were idiosyncratic in their responses.

2.4 Can monocular regions be processed by standard disparity mechanisms?

Our understanding of Nakayama and Shimojo's (1990) elegant result has recently been called into question in other ways. The issue is whether what they dubbed 'da Vinci stereopsis' (using a sparse Type 2 geometric arrangement) requires a specific novel brain mechanism, or whether known binocular processes (like those that may be involved in depth from Panum's limiting case, or simply a coarse stereoscopic mechanism) can account for the depth perceived. A range of evidence supports the possibility that the depth perceived using their stimulus provides an example of Panum's limiting case, and thus could be detected using standard disparity processing mechanisms.

Panum's case arises when one eye views a single line and the other a pair of lines. If the left eye views the single line, then the left line appears closer than the right. This configuration is consistent with a pair of real lines, at different distances, that fall along the line of sight of the left eye (thus appearing as a single line in that eye, but two lines in the right eye). Ono et al (1992) pointed out that in a purely geometrical sense, this is an extreme example of occlusion in the da Vinci configuration, so the same mechanisms could govern the two phenomena.

Gillam et al (1995) showed that for small disparities, depth settings in Panum's case are made as precisely as for normal stereopsis, and that disparity curvature effects can also be revealed when the monocular line is matched to a curved binocular line. One of the proposed explanations for Panum's limiting case is that the lone target in one eye is matched to the already matched line in the other eye, a case of 'double-duty matching' (e.g. McKee et al, 1995). Gillam et al (2003) noted that the da Vinci configuration used by Nakayama and Shimojo (1990) was reminiscent of a Panum's limiting case stimulus, because the monocular bar was similar in width and length to the edge of the rectangular occluder. Gillam et al (2003) went on to demonstrate that when the stimuli are carefully controlled so that Panum's-like mechanisms cannot work, depth from monocular regions is qualitative: a depth direction (in front or behind) is perceived, but not an accurate or precise depth location. To rule out any depth perception due to double-duty matching in this stimulus, they replaced the vertical line with a monocular disc. The occluded disk was always perceived as lying behind the occluder, but there was no change in perceived depth as the disc was separated from the binocular bar. The da Vinci stereopsis configuration used by Nakayama & Shimojo (1990) therefore appears to be an example of Panum's limiting case, and standard stereoscopic mechanisms could be involved in its processing. This is backed up by recent computational modelling efforts (Assee and Qian, 2007), which demonstrated that both Panum's case, and the da Vinci configuration can be modelled using a variant of the disparity energy model (Chen and Qian, 2004).

So far, then it appears that standard stereoscopic mechanisms may be responsible for the perception of depth in the da Vinci configuration. As will become evident below, however, other effects evade such straightforward explanation.

2.5 Do monocular regions provide evidence for a separate, sophisticated depth mechanism?

When one eye views a black bar and the other a black bar with a central gap (figure 6b), the fused percept is of a pair of rectangles displaced in depth (another example of a Type 2 occlusion). Figure 6 (top) shows a possible occlusion situation, in which a pair of objects are located side by side at different depths. One eye sees through the gap between them to the featureless background. In the other eye this gap is occluded by the near object. This effect has been dubbed monocular gap stereopsis (Gillam, Blackburn & Nakayama, 1999)[1]. This is a potentially important stimulus configuration because the depth settings are precise, and cannot easily be accounted for by mechanisms responsive to traditional binocular disparity. Depth in these stimuli appears to be mediated by mechanisms specifically sensitive to the width of the monocular gap. Gillam et al (1999) showed that the amount of depth perceived increases with the size of the gap, and can be matched to a stimulus containing depth from binocular disparity. This effect is consistent with the visual system interpreting the central monocular gap as a gap between the objects that is occluded in the one eye's view due to a depth difference between the objects.

Depth thresholds for monocular gap stereopsis have been found to be very similar to those for standard stereopsis, and, importantly, adaptation to stereopsis results in shifts in perceived depth from monocular gaps, and vice versa (Pianta & Gillam, 2003a). Adaptation techniques are frequently used to explore whether different stimuli are processed by common sensory mechanisms. The logic used is that if a stimulus adapts a particular mechanism, for example a depth mechanism, then perceived depth should be affected in other stimuli that are processed by the same mechanism. Pianta and Gillam's powerful result, the first using adaptation to explore the perception of depth from monocular regions, suggests that the two forms of depth information may be processed by a common mechanism. However, an alternative interpretation cannot be ruled out: the depth could be processed by different mechanisms that converge on a later, common mechanism, which can be adapted. It has also been found that the perceived depth is closest to that provided by a disparity depth probe when the gap contains visual information consistent with it actually being a gap through to the background: if the 'gap' is a different colour or texture than the background region which surrounds it, less depth is perceived (Grove et al, 2002). Further, the stimulus configuration must be such that the gap and background can form a continuous surface (i.e. both are textured, or both featureless), or else perceived depth is attenuated (Grove, Sachtler & Gillam, 2006).

If monocular-gap stimuli and binocular disparity are processed by a common mechanism, do we need to rethink how disparity itself is processed, or could traditional disparity-processing mechanisms account for both the depth threshold and adaptation results? Pianta & Gillam (2003b) suggested (and then tested) two forms of depth processing that might occur. In the original monocular gap stimuli, the gap in one eye's view is obtained by 'pulling apart' the rectangle that forms the other eye's view: the resulting rectangle-with-gap is wider than the single rectangle in the other eye, wider by an amount equal to the gap width (see figure 6b). One way that depth could be perceived is if the visual system were to detect the disparity difference between the outer edges of the black rectangles and then to use the monocular region to simply label the location of the depth edge. Depth attributed to the rectangles would somehow need to propagate from the outer edges, a form of depth interpolation.

---

[1] There is also a dynamic version of this effect(Brooks & Gillam, 2006a).

Monocular-gap stimuli can also be built without outer-edge disparity. This can be achieved by showing each eye a rectangle of the same width, then superimposing a ‘gap’ onto the centre of one of them. Now, the outer-edge disparity would be zero. Pianta and Gillam argued that depth could only be perceived at the gap if the visual system were able to infer an implicit depth signal, inferring that the lack of a gap in one eye can only be due to a particular geometric arrangement, involving occlusion.

By measuring depth thresholds for no-gap stimuli (with no gap the observer sees a narrower rectangle in one eye than the other, resulting in the perception of slant around a vertical axis), for same-width gap stimuli, and gap-with-outer-edge disparity stimuli, Pianta & Gillam (2003b) demonstrated that the depth percept at the gap is robust when outer-edge disparity is present, and of the expected sign and magnitude. Importantly, depth is also perceived at the gap for same-width stimuli, when the outer-edge disparity is zero, and it still varies monotonically with gap size, but its magnitude is smaller than that found when outer-edge disparity is present. These results imply that some mechanism other than outer-edge disparity is at work when interpreting the depth in monocular gap stereopsis.

Another hint that there may be a separate mechanism for monocular gap stereopsis comes from work on how depth from monocular regions is scaled by changes in accommodation and/or viewing distance. Kuroki & Nakamizo (2006) showed that depth does not scale with distance for monocular gap stereopsis, as it does for other examples of monocular occlusion depth, and as it does for standard binocular disparity. Recently, models have been developed that use the output of disparity-detectors in ways that could make use of monocular gaps (see modelling section below, in particular Grossberg and Howe, 2003; Cao and Grossberg, 2005). These models rely on the use of outer-edge disparities and cannot explain the depth perceived when the outer-edge disparity is zero.

A further twist to this story is added by a stimulus containing even fewer clues to the presence of depth, the ‘stereoscopic sliver’ stimulus (Sachtler & Gillam, 2007). Here there is a monocular gap in one eye’s view, but no outer-edge disparity. The gap does not cover the full vertical extent of the stimulus but tapers in width from the centre until it disappears near the top and bottom. This is consistent with a torn piece of fabric, with a featureless background visible through the tear in one eye’s view, but the tear occluded in the other eye. Depth differences between the edges of the tear can be reliably discriminated.

One interesting point to note is that observers in these experiments did not see the gap itself in depth (Gillam et al, 1999). We have noticed that some observers perceive the monocular gap to be in depth with respect to the black surfaces, particularly with careful fixation on the gap itself. This percept is reminiscent of that found by Kumar (1995) who showed both eyes a light rectangle and one eye a superimposed dark bar at the centre, the other a lighter bar. For some configurations the two halves of the rectangle appeared at different depths, for others the central bar appeared in depth. The percept may also be similar to that for the perception of monoptic depth (see section 2.2), where a single monocular element can appear in depth. It is also not clear whether the effect of instructions, of eye movements or some other stimulus property could account for this alternative interpretation, and we currently do not know whether only a small proportion of the population achieve this alternative percept. Further studies with larger numbers of naïve observers, could help us to understand just how robust these percepts are.

3 Integration of binocular and half-occluded regions in 3D scenes

‘Leonardo’s constraint’ (Ono, Wade & Lillakas, 2002) is the constraint that two opaque objects cannot be seen in the same visual direction. An example of this is the situation where a background surface is visible to one eye, or the other, but not both, when an occluding object is smaller than the interocular separation. Figure 1 shows an example of such a scene where many foreground objects occlude the background surface such that any one part of the background is only seen by a single eye. We defined this above as a Type 1 occlusion. In principal, the visual information is available to 'see behind' the foreground object. Can the visual system do this?

This question can be addressed in at least 3 ways. First, we could ask whether our visual impressions are stable under conditions where large regions of the scene are monocular. We explore this in section 3.1 below. Second, it is well known that when the eyes are shown different images, binocular rivalry results. Rivalry consists of the dominance of one eye's view, which is periodically replaced with the other eye's view (e.g. see Blake and Logothetis, 2002). There is a considerable literature on the spatial and temporal properties of binocular rivalry, though it rarely occurs under natural viewing conditions. We discuss literature in section 3.2 which demonstrates that depth can be perceived in occlusion situations, despite rivalry. As we will see, some researchers believe that rivalry itself has an interpretation in terms of occlusion geometry. Third, when some parts of a scene are viewed by a single eye, and some by both eyes, it is not clear how our phenomenal perception of a single fully 'stitched together' world, can be obtained. Figure 7 illustrates the problem. Each eye views the world from a different direction, because the eyes are laterally separated. Yet we feel that we view the world as if from a single point, mid-way between the eyes. If some transformation occurred to deliver a representation from that point, the brain would have to squeeze the region defined by the separation between points a and d (the right eye view in figure 7), into a smaller region separating a' and d' (the view as if from a single central location). How, or even if, this is done is still a hotly debated topic and beyond the scope of this review, but the interested reader is directed to Erkelens & van Ee (2002) and Ono, Mapp & Howard (2002) and for reviews see Howard & Rogers (2002) and Ono et al (2009).

3.1 Scene stability despite large monocularly visible regions

Forte et al (2002) studied the stability of monocular regions when there was no binocularly visible background, but regions of the background surface were visible to one eye or the other (e.g. figure 1). In their stimuli both eyes viewed foreground 'occluders', which were arranged so that no part of the background surface was viewed by both eyes. In such displays qualitative measurements suggested that observers saw what was described as 'stable diplopia', perceiving the background surface as coherent and continuous. We do not know to what extent the information in such scenes is represented in the same way as for binocular regions, because the work of Forte et al (2002) has not been extended to quantitative predictive measures. In particular, they did not establish whether there was a processing advantage of stable diplopic viewing over a monocularly viewed scene. In other words, can the visual system use information from the stable diplopic scene as well as it can for a normal binocularly or monocularly viewed scene? Could this information be used to help recognise objects, or to more accurately measure the number of items present behind a 'screen' of fence, or long grass? It does appear that all the items in monocularly viewed regions are visible (Erkelens, Muijs & van Ee, 1996), though the detectability of a monocular point is poorer in a monocular region than when it is presented in a binocular location (Emoto & Mitsuhashi, 1998). It is even possible to read text behind a small foreground object when viewed binocularly (Ono, Lillakas, Grove & Suzuki, 2003), as can be demonstrated by holding a pencil between yourself and this text. However, it is not known whether the text is perceptually distorted or whether it is more difficult to read than normally viewed text.

3.2 Linking monocular regions to binocular rivalry

Shimojo and Nakayama (1990) were the first to notice that some monocular regions appear to undergo suppression due to binocular rivalry. They accounted for this by suggesting that 'ecologically valid' monocular regions are perceived as stable (as well as having a consistent depth), and invalid arrangements are not. As described above, Assee & Qian (2007) pointed out that those 'ecologically invalid' regions do have a 'valid' interpretation but that they are often seen as rivalrous. However, as demonstrated by Forte et al (2002) some depth can be perceived even when the regions of an image seen at a different depth are totally different in the two eyes views.

The ‘sieve effect’ (a Type 1 occlusion arrangement) is another example of a viewing situation in which there is no consistent binocular disparity information, but where perceived depth is attributed to mechanisms that may rely on knowledge of occlusion geometry. Figure 8 shows the viewing situation devised by Howard (1995). An observer fixates a near surface in which there are ‘holes’ through which a far surface can be seen. The hole size and background pattern are designed such that one eye sees a light patch in the background and the other eye a dark patch. This results in binocular rivalry within the patch, with a fluctuating percept between white and black, but the rivalrous region contained within the holes is perceived as lying behind a foreground occluder, into which the holes are punched. The effect requires some sort of surface percept, for instance it does not occur when there is only one hole. The real scene consistent with this situation is that of viewing a striped background surface through a number of holes (the ‘sieve’). Howard provides a number of examples of this stimulus, and finds that the depth effect is most robust when the holes are each smaller than 1 deg of visual angle, when there are many holes, when each is surrounded by a binocular rim that is clearly visible in both eyes. Rivalry occurs in addition to the perception of depth for many observers. It has been noted that not all observers perceive depth from sieve effect stimuli (Howard, 1995; Tsai and Victor, 2000), and it is not clear how commonly the effect delivers the full sieve effect without prompting.

A recent computational model, designed to account for how monocular regions could contribute to depth perception alongside disparity (Hayashi et al, 2004), delivers binocular rivalry as an apparent ‘side-effect’ (see the modelling section 5 below). This is an interesting point given that others have suggested that almost all rivalry stimuli could be interpreted as examples of occlusion (Ooi & He, 2006). These authors noticed that when one scene is presented to the left eye and a different one to the right eye, the viewing arrangement is consistent with the geometrical interpretation that the observer is viewing two different scenes, side by side, through a hole in a binocularly visible foreground (akin to the adjacent dark and light regions viewed through one of the holes in figure 8a). Ooi & He suggest that rivalry can be understood and explained via a 3-D interpretation, thus forming an intimate link between rivalry and binocular surface perception. For example, if one eye see a small patch of one texture, and the other eye a patch of another texture, as in a typical rivalry stimulus, this is consistent with viewing a background through a small foreground hole, with the proviso that the background consists of regions containing both kinds of texture and the hole-eye arrangement is such that one eye sees one texture and the other eye the other texture. This is akin to viewing the sieve-effect stimulus through a single foreground hole (figure 8). The same group goes on to demonstrate that when rivalry configurations are consistent with this, or other, 3-D interpretations, scene stability is more likely to occur than rivalry. For example, observers experience a stable percept when a monocular target is presented on a binocular background (Ooi & He, 2006) and when monocular regions are consistent with an invisible occluding foreground (a Type 3 occlusion, figure 9, Bogaert, Ooi & He, 2008). Rivalry is much more likely to occur when the two eyes images are switched. The group suggest that the visual system prefers selecting images that contain a monocular boundary contour, consistent with a 3-D occlusion, precisely because is it consistent with a real 3-D interpretation. This idea resonates with Nakayama and Shimojo’s (1990) idea of ‘valid’ and ‘invalid’ scenes, though Assee & Qian’s (2007) interpretation suggests to us that rivalry occurs when the real 3-D scene formed by stimuli containing monocular regions is very unlikely to occur.

The idea that rivalry in general, and the sieve effect in particular, occur because the visual system interprets different inputs to the left and right eye as due to particular geometrical configurations, is compelling. However, work exploring detailed predictions about the quantity of depth that should be perceived in the sieve effect, argues against the idea. Tsai and Victor (2000) discussed several predictions of the sieve effect, noting that to be consistent with occlusion geometry, perceived depth in the sieve effect should vary with the horizontal size of the viewing apertures. They found that precision of depth perception from sieve-effect stimuli was rather poor, around 10 times worse than for standard stereopsis, though the depth within the sieve elements was consistently seen as behind the occluder. This is similar precision to that found for judging depth from disparity for anti-correlated bars (Cogan et al, 1995), but notice that in the sieve-effect stimulus there is no disparity applied to the elements, so there must be a different mechanism at work.

In the same study, Tsai and Victor (2000) found other attributes of the sieve effect that were not consistent with an occlusion-based explanation. According to such an account, perceived depth should vary systematically with element width, but not height. Although the horizontal size of the holes did affect perceived depth a little, so did the vertical size. In a later study (Tsai and Victor, 2005), the relative locations and luminances of elements were varied to alter the minimum depth between the occluder and the background surface that would be consistent with occlusion geometry. Binocular viewing geometry dictates that if two elements are closer than the element width, their relative luminance polarity (same or different) will determine the perceived separation between occluder and background. The study found no evidence that the visual system could take account of these geometric constraints. These results cannot be explained by a standard stereoscopic mechanism, or by mechanisms that rely on appropriate occlusion geometry.

Some recent work corroborates this conclusion. Matsumiya et al (2007) measured depth from the sieve effect under a number of conditions and found it to be maximal when exclusive rivalry within the elements was also greatest (exclusive rivalry is defined as occurring when perception correlates with the view from one eye, or the other, rather than some intermediate or partial effect, e.g. see Blake et al, 1992). These authors suggested that the same mechanisms might be at work in the processing of rivalry and depth from the sieve effect, although they did not speculate further. In sum, whilst there clearly are links between rivalry and depth from monocular regions, the available evidence is not wholeheartedly behind the idea that they arise as part of the same processing mechanism.

4 Monocular regions and surfaces

Monocular regions in a scene can generate the percept of an illusory surface, and associated illusory contours, consistent with an invisible occluder, that appears in the foreground in depth. We have defined monocular regions due to this type of occlusion as Type 3 and noted that they will occur very rarely in natural scenes. That an invisible occluder could be perceived from monocular regions was first shown using sparse dot patterns via depth magnitude estimation (Lawson & Mount, 1967; Lawson & Gulick, 1967). An extreme example of the phenomenon was demonstrated by Nakayama & Shimojo (1990), who devised a stimulus where only 4 points in a sparse random dot stereogram are viewed monocularly, yet a clear illusory surface can be seen in depth (figure 10). Vertically oriented monocular regions have also been shown to generate a clear percept of an illusory surface (Anderson, 1994, figure 11d), and monocular regions presented at the edge of slanting binocular surfaces in a random dot stereogram can increase perceived slant (Gillam & Blackburn, 1998). In the sub-sections below we discuss depth perception from several surface-related instances of monocular regions like these.

4.1 Phantom stereopsis

Figure 11a shows the stimulus used by Liu, Stevenson & Schor (1994) to demonstrate ‘phantom stereopsis’. In this stimulus, depth can be seen in the central white patch despite an apparent lack of traditional stereoscopic matchable features. Liu et al showed that quantitative depth was obtained from this stimulus, and that it can drive vergence in a similar way to stimuli containing binocular disparity (Liu, Stevenson & Schor, 1998). This study provides a classic example of how clever stimuli are invented to deliver depth perception without the involvement of traditional stereopsis, but where, later, it is shown that this assumption is not correct. In this case, Gillam (1995) argued that if one considers only the horizontal contours (figure 11b), the perception of depth remains, and it is evident that the end points of the lines can be used for standard stereoscopic matching. Liu et al (1997) tested a simple model of disparity processing, finding that, although disparity mechanisms would respond differently to phantom-stereopsis stimuli than to stimuli containing standard binocular disparity, there was a disparity signal that could be used to obtain depth sign information that is consistent with the psychophysical results. These depth signals are required for surface interpolation: the observer sees a plane in depth, rather than individual points. The authors noted that mechanisms that use some knowledge of occlusion geometry to guide disparity selection could achieve the depth sign consistent with perception, but also that much simpler mechanisms, perhaps using disparity averaging or other simple heuristics, might also work. Clearly, additional careful experimentation is needed to test these ideas more thoroughly.

A start has been made in this direction. Gillam and Nakayama (1999) designed an elegant stimulus composed only of vertical lines with gaps in them (Figure 11c). A central rectangle is perceived, standing out in depth, with strong illusory horizontal contours, dubbed the phantom occluder (type 3 monocular regions). Observers were able to match the perceived depth of the rectangle with a stereoscopic probe target, but performance was variable between observers and not as accurate as for standard stereoscopic stimuli. An interesting feature of phantom stereopsis is that the apparent depth between the phantom and the background (featureless except for the vertical lines) is greater than the depth constraint zone would predict (Gillam & Nakayma, 1999; Grove Gillam and Ono, 2002). By using visual search in noise defined by either disparity or half-occlusion elements, Mitsudo, Nakamizo and Ono (2005) were able to show that depth from phantom stereopsis appears to be processed at an early stage of visual perception.

In an earlier study Anderson (1994) showed that *vertical* offsets between the left and right eye views, consistent with an invisible occluder (see figure 11d), give rise to perception of a phantom occluding surface. A more general theory was later developed to account for the depth perceived from occlusion junctions (the places where objects overlap). Due to occlusion of one object by another, the occlusion junctions can have both a horizontal and vertical separation between the two eyes' views. The separation between junctions was defined, by Malik, Anderson & Charowhas (1999), as 'pseudodisparity' and it was demonstrated that for such scenes there is a clear quantitative relationship between perceived depth and image pseudodisparity. Remarkably, the orientation of the clear illusory contours can also be precisely judged: a neat demonstration of the clarity and crispness of such contours. While it seems clear that pseuododisparity can be exploited in some instances to support depth perception, this is not always the case. More recent studies by van Ee, Anderson & Farid (2001) have shown that depth detection near disparity threshold is not improved by the presence of pseudodisparities.

Depth perception has been demonstrated in dynamic versions of the phantom stimulus described above. This was first done by Shimojo, Silverman & Nakayama (1998). They showed that sequentially stimulating each eye with a moving line that 'disappears' behind an occluder (so that the depth cue is the differing time and location of occlusion and reappearance) results in a clear percept of depth. This depth percept increases as the temporal gap becomes larger, consistent with the presence of an occluded object further away. Brooks & Gillam (2006b) used a similar stimulus and ruled out the possibility of depth being perceived via inter-ocular delay because the effect remains robust when the delay is eliminated. It is difficult to imagine a way to perceive depth in such spatial or temporal line stimuli that simply relies on conventional stereoscopic mechanisms.

Hakkinen & Nyman (2001) have argued that phantom stereopsis must be closely linked to conventional disparity processing. They showed that a monocularly induced phantom surface can influence the depth perceived from stereopsis when that information is ambiguous. This effect only occurs for specific physical arrangements, notably when the monocular regions appear to be part of a continuous surface also defined by an adjacent binocular region. Grove, Gillam & Ono (2002) explored how monocular regions affect perceived depth for a variety of different stimuli. They found that less depth was perceived when the monocular texture was dissimilar to the background texture. Studies agree (Hakkinen & Nyman, 2001, Grove et al, 2002) that these effects occur most strongly when the monocular regions are consistent with there being a foreground invisible occluder. This is an important point because it emphasises that, whatever mechanisms are at work, both local processing (to account for monocular regions requiring locally consistent binocular regions) and larger-scale, or long-distance processing (to account for the global occlusion geometry) must be involved. This point has recently been demonstrated by Mitsudo, Nakamizo and Ono (2006). They measured contrast sensitivity for detecting stereopairs in noise and found greater sensitivity for a phantom stereopsis stimulus (like that in figure 11c) than for an equivalent stimulus with the two eye's views switched round. In the latter case it is possible to contrive a 'real' scene arrangement that could deliver the left and right eye views, but they would be rarely encountered in the real world. Sensitivity to the foreground occluder configuration was also greater than for a stimulus composed of only the left-most bar in each eye. This work suggests that, to obtain and use the phantom surface, information is combined by large-scale processing mechanisms which process information across the full extent of the stimulus, rather than relying on individual elements.

4.2 Monocular transparency

Howard & Duke (2003) presented a novel effect that they named monocular transparency, in which perceived depth is attributed to geometrical rules related to transparency, rather than to occlusion geometry. One of their stimuli is depicted in figure 12. One eye views a white rectangle, occluded by a slightly offset transparent square, and the other the same rectangle, with the square aligned. In the configuration shown in figure 12, the square is seen to float in front of the rectangle. When the two eyes views are switched, it is seen behind. The depth of the square could be matched to a disparity-defined depth probe and delivered quantitative depth percepts. The key point to the design of this display is that occlusion cannot be required to explain the perception of depth because no part of the scene is occluded, and standard disparity processing cannot account for the depth because there are no vertical edges of the foreground object in one eye's view.

Can other explanations account for these findings? Howard and Duke considered the possibility that the vertical contours from the eye containing the target could be matched to vertical contours above and below the gap in the other eye. They ruled this out because the contours are of opposite polarity. Grove, Brooks, Anderson & Gillam (2006) noted that the extent to which depth can be seen via opposite contrast edges is controversial. They performed experiments showing that such matches can result in perceived depth, and suggested that the depth in some configurations could be obtained via disparity-processing mechanisms which are robust to local luminance contrast differences. This suggests that performance is instead mediated by a disparity mechanism that responds to the overall extent of the stimulus, perhaps akin to the 2nd-order mechanism proposed by Hess and Wilcox (1994) and referred to by Cogan et al (1995). In other configurations, where horizontal contours had the same polarity, Grove, Brooks et al (2006) found that disparity matches were more robust and consistent with standard stereoscopic matching of horizontal contours.

Grove, Brooks et al went on to study other versions of the transparent stimuli used by Howard and Duke and demonstrated that many effects do not require transparency. Instead, they appear to be examples akin to those described as monocular gap stereopsis (Gillam et al, 1999, section 2.4). In sum, it may not be necessary to invoke novel depth processing mechanisms to account for depth perceived in this class of visual stimuli.

4.3 Surface intrusion

Cook & Gillam (2004) devised an 'intrusion stereogram' (see figure 13) in which one eye views a black figure-of-eight and the other eye views that same figure with a white patch removed on the right side. When the intrusion is presented on the nasal side of the stimulus (corresponding to the temporal retina)), some observers (4/7) see a white patch floating behind and partially occluded by the black figure-of-eight, as if viewed through a hole in the white foreground. The patch 'intrudes' into the black surrounding region, both of which are seen as behind the hole. When the left and right eye views are switched, so that the intrusion is now temporal, all observers see the intrusion as an object floating in front of a background that consists of the black figure-of-eight object and the white surround. This is a camouflage situation similar to that shown in figure 4 (type 3 region). Lateral motion of the intrusion has also been shown to result in the perception of motion in depth (Brooks & Gillam, 2007), just as standard disparity change allows the perception of motion in depth (e.g. see Harris, Nefs & Grafton, 2008).

For some observers, the depth perceived in both configurations varies as a function of the position of the vertical intrusion edge (for others this only works for the temporal stimulus configuration), suggesting a quantitative mechanism is at work. The key question, as ever, is whether traditional stereoscopic mechanisms can account for performance. One way to explore this issue would be to exploit the individual differences in precision, noted by Cook & Gillam. This has never been done, though is certainly tractable via forced choice psychophysical methods, where depth thresholds for monocular intrusion stimuli could be compared with those for binocular occlusion stimuli (where both eyes see an intrusion, but the intrusion is larger in one eye's view, providing a traditional binocular disparity).

Cook and Gillam argued that an explanation based on traditional stereopsis is unlikely. First, because the three observers who could not see consistent depth in the nasal configuration could do so for the binocular occlusion equivalent. This is intriguing and requires further study. Second, they conducted a control experiment in which observers were asked to set a depth probe to the depth seen in a narrow bar, presented at the same location as the intruding edge in the intrusion displays. Depth did not vary consistently with bar position in these displays, leading the authors to conclude that intrusion is a necessary condition for quantitative depth perception. However, bars presented on the temporal side of the stimulus (nasal retina) were seen behind, and those on the nasal side, in front, reminiscent of the monoptic depth perceived from a monocular bar viewed alone (Kaye, 1978, Wilcox et al, 2007). The visual system may simply be matching the monocular bar to the luminance centroid of the figure-of-eight in the other eye. Thus, while this argument suggests that something different is occurring for the intrusion displays, it does not help clarify what this difference is.

Third, Cook and Gillam note that point-for-point stereo matching of each vertical location on the left eye's edge with that on the right eye's edge should result in perception of a complex 3-D shape at the edge, because the cusp-shape in one eye's view must be matched to a vertical line in the other eye's view. They demonstrated that this depth profile was perceived by three of their four observers in a control stimulus that delivered appropriate disparity information, but the complex profile was not perceived when the explicit disparity information was removed. The authors concluded that the depths reported in the intrusion stereogram can therefore not be explained using traditional stereoscopic mechanisms.

Yet depth perception involves more than the responses of single-stage detectors that engage in a point-for-point match, as discussed in some of the physiologically inspired models described in section 5 below. Other interpretations could be consistent with this study. For example, one could posit a simple alternative matching rule. Consider the stimuli as delivering a pair of edges (figure 13c). We will assume the edges to be located at the widest points visible. If those edges were matched and disparity extracted, the perceived depth would be consistent with that perceived by Cook & Gillam's (2004) observers. This scheme works if we assume that mechanisms exist that can extract the edges. Coarse-scale disparity detectors could do just this. One could quibble about the location of the edges (should they lie at the stimulus extreme, or should they be the vertical centroid of the curved figure-of-eight edge), but with the stimuli used thus far, this will make little difference to the perceived depth. This is just one example of how, with an additional small twist, standard stereoscopic mechanisms could be employed to account for perceived depth in these stimuli. However, what our suggestion fails to account for are the crisp surface edges and contours that form part of the percept. Clearly, more research is needed to fully understand this intriguing stimulus.

4.4 Occlusion and slant

Monocular occlusion effects have an interesting effect on slant perception, suggesting there is a complex interaction between binocular stereopsis and occlusion. If the left and right eyes view images of objects with different horizontal extents, the observer perceives an object slanted around a vertical axis in 3-D. This interpretation is consistent with the real-world conditions that might generate such images in the two eyes (Rogers & Graham, 1983). But differences in horizontal extent can also occur due to monocular occlusion. When a foreground object occludes an object further away, one eye's view will be delivered an image of the background object that has a smaller horizontal extent than the image of the background object in the other eye. Yet in this case, observers do not observe slant, rather they perceive one object as lying front of the other. Figure 2a shows a nice example of this from von Szily's work. This phenomenon has been quantified by Hakkinen & Nyman (1997), who showed that perceived slant of a rectangular region is much diminished when there is a consistent 3-D occlusion interpretation provided by a binocularly visible plane. More slant is perceived when the relationship between the binocular and small test plane are consistent with the presence of occlusion, than when they are inconsistent. Thus slant processing appears to be linked to the processing of binocular disparity with consideration of the presence of monocular occlusion.

Gillam & Grove (2004) have addressed the related issue of what visual information might be used to distinguish between global occlusion and local slant interpretations, when scenes consist of sets of horizontal lines. Occlusion of a set of horizontal lines by an invisible vertical foreground occluder results in each line being shorter by a fixed amount in one eyes view (figure 14a). Yet this stimulus would also be consistent with the lines having different local slants in depth. Gillam and Grove showed that observers perceive the former, global, occlusion interpretation, rather than the local slant interpretation. However, a local slant interpretation is perceived when the two eyes views are switched, when the stimulus is no longer consistent with occlusion by an invisible foreground object. Grove, Byrne & Gillam (2005) have extended this work further, showing that for stimuli comprised of a set of oblique lines (each with differing orientations, figure 14b), perception of the occluding contour is stronger than for a similar but horizontal line stimulus. The authors suggest that the vertical differences between left and right eye views of the lines (which we know deliver a strong perception of occlusion see Anderson, 1994, and figure 11d) disambiguate the ambiguous information provided by horizontal binocular disparity.

5 Models of depth from monocular regions: revealing underlying mechanisms

When modelling the behavioural results obtained using the variety of half-occlusion stimuli described in this review, the experimental approach used must be taken into account. Most of the experimental research has concentrated on designing stimuli that do not contain conventional binocular disparity information, yet which contain monocular regions consistent with 3-D occlusion geometry. Success in achieving this has been mixed, with some stimuli harbouring disparity information that is available to non-standard stereoscopic mechanisms. Those stimuli that have, so far, defied an explanation based on binocular disparity processing, have been discussed in terms of the geometrical constraints that they fulfil. Such accounts do not, of course, further our understanding of the neural mechanisms underlying the phenomena.

Tsai and Victor (2000) provide a helpful distinction, between 'purpose', where the objective is to understand how the world (and its geometry) might constrain perception, and 'process', where the aim is to understand what visual brain mechanisms might actually be at work. They suggest that more work is required to try to understand how the brain's basic binocular disparity processing mechanisms (whose understanding is fairly well advanced at the earliest levels, e.g. Ohzawa et al, 1990; Qian, 1994) might be used to incorporate monocular information in a relatively simple way. Some efforts in this direction are outlined below.

A recent suggestion has been directly inspired by the neurophysiology of stereopsis. Disparity-sensitive simple cells are known to code for disparity by both phase and position shifts (e.g. DeAngelis et al, 1991; Anzai et al, 1999, Prince et al, 2002, Tsao et al 2003). It has been suggested that some occlusion arrangements generate unusual combinations of phase- and position-shifts in simple cells and that an appropriate combination of phase- and position-coding could account for some half-occlusion phenomena (Tsao, Conway & Livingstone, 2003). This suggestion has yet to be formalised, but modelling along these lines could generate valuable, testable, predictions.

There have been several computationally implemented models of stereopsis that have attempted to deal with the presence of unpaired points, but not necessarily make use of them as a depth cue. We do not have space to review them all here (but see Egnal and Wildes for a comparison of several models, and Jones and Malik, 1992, for a model that takes account of monocular regions). Watanabe and Fukushima (1999) have developed a stereo algorithm that combines traditional binocular matching (based on Marr and Poggio's 1979 cooperative stereoalgorithm, that implements both a smoothness and uniqueness constraint) with monocular detectors that signal the presence of monocular regions. They use a cooperative stage to combine information from binocular disparity and neighbouring monocular regions, which relies on a constraint consistent with binocular occlusion geometry. In this sense the model is inspired by 'purpose': the known properties of the world are used to direct what depths are signalled. They demonstrate that their model can identify monocular regions that occur due to partial occlusion of one eye's view, to aid in finding depth discontinuities (see also a similar model by Zitnick & Kanade, 2000).

Hyashi et al (2004) have extended Watanabe & Fukushima's model. One of their aims was to use the known properties of early visual neurons to generate a physiologically plausible initial matching stage, thus linking process with purpose. To do this they started with a disparity energy model (Ohzawa et al, 1990) and constructed a monocular-region detector by monitoring the output of a population of binocular disparity energy neurons, each looking at the same region of a scene. The signature of a monocular region is broad activation across the population of disparity detectors, rather than specific activity over a narrow band of disparities, as would be delivered by strong disparity signals, signalling a particular depth. So these authors have ingeniously used the population response from a purely binocular mechanism to identify locations where there is no consistent binocular signal (and hence there must be a monocular region). Another key feature of the model describes what they call an additional occlusion constraint, that only one monocular representation can occur at one instant, in other words, that a monocular region is present in one eye, or the other, but not both. To achieve this, right and left eye representations for the monocular regions are designed to inhibit one another. Not only does this model account nicely for some of the psychophysical results in particular, it also provides a 'for-free' model of binocular rivalry in general. When the two eye's views are completely different, the interocular inhibition between the left- and right-eye monocular representations occurs across the whole scene, resulting in temporally alternating perception of one or the other views. This is the first model of stereopsis that can simultaneously accommodate conventional disparity processing, monocular regions as depth signals, and binocular rivalry.

In a recent model, Assee and Qian (2007) critique preceding approaches, noting that in most cases while the front-end of the model is inspired by neurophysiology, later stages are not. Assee and Qian develop an algorithm to explain how monocular regions are perceived at the depth of the background surface, using model V2 cells (von der Heydt, Zhou and Friedman, 2000). In their model, the population response from V2 model neurons is used to identify the location and eye-oforigin of monocular regions, then the geometric rules of occlusion are used to assign the same depth as the background.

An alternative model also uses simulated binocular neurons and includes input from monocular regions (Grossberg and McLoughlin, 1997, McLoughlin and Grossberg, 1998). Monocular regions are initially assigned all possible depths, then depth is determined based on higher-level assumptions including filling in. The model has been developed to account for some results from the original da Vinci stereopsis study, monocular gap stereopsis (Grossberg & Howe, 2003) and can be elaborated to account for additional half-occlusion effects (Cao & Grossberg, 2005). The model does not explicitly require knowledge of occlusion geometry to deliver an output consistent with human perception, but does use a variety of other high-level rules, such as the propagation of depth from edges into a figure.

What we do not know is how these various models compare with one another. Egnal and Wildes (2002) compared a variety of computational models designed to detect monocular regions, but the models were not designed to be physiologically plausible, or to use monocular regions as a source of depth information. It would be fascinating to compare the current models directly on sets of the more challenging occlusion stimuli that have been described in this review.

6 Conclusions

In this article we have attempted to provide a comprehensive review of the literature related to depth from monocular regions. We have laid out in detail the experimental evidence showing how these regions can be used as part of a binocular visual representation. It is clear that monocular regions are important for forming surface representations and for depth perception. Research has shown that information from monocular regions is not simply thrown away by mechanisms dedicated to forming a seamless representation of the world.

At the outset we classified monocular regions into 3 types, based on the form of the 3-D scene that would be delivered by the monocular region and its accompanying binocular information. Most of the examples described in this review fit into one of these categories. Our classification was not intended to delineate separate phenomena requiring distinct processing mechanisms, and we would not pretend to have done so, but some interesting points about each type have emerged.

Type 1 occurs when the geometrical arrangements of foreground occluders is such that each portion of a textured or patterned background is seen by only one eye. This is interesting, because depth perception from these the arrangements is accompanied by binocular rivalry, but only sometimes. Stimuli exploring this type of monocular region have provided tantalising clues for how binocular vision, rivalry and depth from monocular regions, can be understood together. We are still a long way from this aim.

Type 2 represents the most commonly occurring monocular region, which is hidden by a binocularly visible foreground. As we described, some examples of this type can be explained using standard stereoscopic mechanisms, and modelling efforts are beginning to explain how binocular mechanisms can be adapted to obtain depth from the most common case, where the monocular region is accompanied by binocularly visible foreground and background. However, there remain a few stubborn effects that have not been amenable to straightforward explanation, in particular monocular gap stereopsis.

Although Type 3 monocular regions are rare, because they require the binocular foreground to be invisible (camouflaged by being the same colour/texture as the surrounding space), they have spawned a whole series of intriguing stimuli, in particular the depth and phantom surface effects emerging from studying intrusion stimuli. These will provide challenges to future researchers engaged in forming theories and models of binocular vision.

We have shown that it is now possible to explain some phenomena that involve depth from monocular regions, using extensions of standard stereoscopic mechanisms. A parsimonious view would be that someday, all of these phenomena could be explained via elaborations of the binocular mechanisms that we know underly standard disparity processing. But this has certainly not yet been demonstrated. The key question for the future is the extent to which each of these monocular region phenomena require explanation via theories and mechanisms that are distinct from normal stereopsis and that require inferences about the 3-D geometry of the external world. If this is necessary, then surely the mechanisms responsible for stereopsis are far more sophisticated than we current know.

Acknowledgements

We thank the following individuals for providing comments and suggestions on early versions of this manuscript: Kevin Brooks, Barbara Gillam, Paul Hibbard, Ken Nakayama, Harold Nefs, Hiroshi Ono.

References

Anderson, B. L. (1994) The role of partial occlusion in stereopsis. Nature, 367, 365-368.

Anderson, B. L., & Nakayama, K. (1994). Towards a general theory of stereopsis: Binocular matching, occluding contours, and fusion. *Psychology Review*, 101, 414–445.

Anzai, A., Ohzawa, I., and Freeman, R.D. (1999a). Neural mechanisms for encoding binocular disparity: receptive field position versus phase. J. Neurophysiol. *82*, 874–890.

Assee A., & Qian, N. (2007) Solving da Vinci stereopsis with depth-edge-selective V2 cells. Vis. Res., 47, 2585-2602.

Barrand, A. G. (1979) An ecological approach to binocular perception: the neglected facts of occlusion. Cornell University PhD, 1979. Published by University Microfilms International, Michigan.

Blake, R., & Logothetis, N. K. (2002). Visual competition. Nature Reviews Neuroscience, 3, 13–21.

van Bogaerts, E.A., Ooi, T.L. & He, Z. J. The monocular-boundary-contour mechanism in binocular surface representation and suppression (2008), Perception, 37, 1197-1215.

Brooks, K. R., & Gillam, B. J. (2006a). The swinging doors of perception: Stereomotion without binocular matching. *Journal of Vision*, *6*(7):2, 685-695

Brooks, KR; Gillam, BJ (2006b) Quantitative Perceived Depth From Sequential Monocular Decamouflage, Vision Research, 46, 605-613.

Brooks, K. R., & Gillam, B. J. (2007). Stereomotion perception for a monocularly camouflaged stimulus. *Journal of Vision*, *7*(13):1, 1-14

Cao, Y. & Grossberg, S. (2005) A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. Spatial Vision, 18, 515-578.

Changizi, MA & Shimojo, S (2008) "X-ray vision" and the evolution of forward-facing eyes. Journal Of Theoretical Biology, 254, 756-767 .

Chen, Y., & Qian, N. (2004). A coarse-to-fine disparity energy model with both phase-shift and position-shift receptive field mechanisms. Neural Computation, 16, 1545–1577.

Cogan, AI; Kontsevich, LL; Lomakin, AJ, Halpern DL & Blake, R. (1995) Binocular Disparity Processing With Opposite-Contrast Stimuli. Perception, 24, 33-47.

Collett, TS (1985) Extrapolating And Interpolating Surfaces In Depth. Proceedings Of The Royal Society Of London Series B-Biological Sciences, 224, 43-56.

Cook, M., & Gillam, B. (2004). Depth of monocular elements in a binocular scene: The conditions for da Vinci stereopsis. Journal of Experimental Psychology. Human Perception and Performance, 30, 92–103.

Cumming, B.G., Shapiro, S.E. & Parker, A.J. (1998) Disparity detection in anticorrelated stereograms. Perception, 27, 1367-1377.

DeAngelis, G.C., Ohzawa, I., and Freeman, R.D. (1991). Depth is encoded in the visual cortex by a specialized receptive field structure. Nature *352*, 156–159.

van Ee, R., Anderson, B.L. & Farid, H. (2001) Occlusion junctions do not improve stereoacuity. Spatial Vision, 15, 45-59.

Ehrenstein, W. H., & Gillam, B. J. (1998). Early demonstrations of subjective contours, amodal completion, and depth from monocular regions: ‘‘Stereoscopic experiments with silhouettes’’ by Adolf von Szily (1921). Perception, 27, 1407–1416.

Egnal, G. and Wildes, R. P. (2002) Detecting binocular monocular regions: empirical comparisons of five approaches. IEEE Trans. Pattern Anal. Mach. Int., 24, 1127-1133.

Erkelens, C. J., & van Ee, R. (2002). The role of the cyclopean eye in vision: sometimes inappropriate, always irrelevant. Vision Research 42, 1157-1163

Erkelens, CJ; Muijs, AJM; van Ee, R (1996) Binocular alignment in different depth planes Vision Research, 36, 2141-2147.

Emoto, M. & Mitsuhashi, T. (1998) Interocular suppression of a half-occluded region of stereoscopic images. J. Opt. Soc. Am. A., 15, 2257-2262.

Forte, J; Peirce, JW; Lennie, P (2002) Binocular integration of partially occluded surfaces. Vision Research, 42, 1225-1235.

Fukuda, K; Wilcox, LM; Allison, RS, Howard, I (2009) A reevaluation of the tolerance to vertical misalignment in stereopsis. *Journal of Vision*, *9*(2):1, 1-8.

Gillam B (1971) Depth Processing Theory Of Poggendorff Illusion Perception & Psychophysics 10 : 211 1971.

Gillam, B (1995) Matching Needed For Stereopsis, Nature, 373, 202-203.

Gillam, B. & Grove, P.M. (2004) Slant or occlusion: global factors resolve stereoscopic ambiguity in sets of horizontal lines. Vision Research, 44, 2359-2366.

Gillam, B. & Borsting, E. (1988) The role of monocular regions in stereoscopic displays. Perception, 17, 603-608.

Gillam, B. J. & Blackburn, S. G. (1998) Surface separation descreases stereoscopic slant but a monocular aperture increases it. Perception, 27, 1267-1286.

Gillam, B., & Nakayama, K. (1999). Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. Vision Research, 39, 109–112.

Gillam B., Blackburn, S. & Cook, M. (1995) Panum's limiting case: double funsion, convergence error, or da Vinci stereopsis? Perception, 24, 333-346.

Gillam, B; Blackburn, S; Nakayama, K (1999) Stereopsis based on monocular gaps: Metrical encoding of depth and slant without matching contours. Vision Research, 39, 493-502.

Gillam, B., Cook, M., & Blackburn, S. (2003). Monocular discs in the occlusion zones of binocular surfaces do not have quantitative depth— A comparison with Panum's limiting case. Perception, 32, 1009–1019.

Grossberg, S., & McLoughlin, N. P. (1997). Cortical dynamics of three-dimensional surface perception: binocular and half-occluded scenic images. Neural Networks, 10, 1583–1605.

Grossberg, S; Howe, PDL (2003) A laminar cortical model of stereopsis and three-dimensional surface perception. Vision Research, 43, 801-829.

Grove, P. M., & Ono, H. (1999). Ecologically invalid monocular texture leads to longer perceptual latencies in random-dot stereograms. Perception, 28, 627–639.

Grove, PM; Gillam, B; Ono, H (2002) Content and context of monocular regions determine perceived depth in random dot, unpaired background and phantom stereograms. Vision Research, 42, 1859-1870.

Grove, PM; Ben Sachtler, WL; Gillam, BJ (2006) Amodal completion with background determines depth from monocular gap stereopsis.Vision Research, 46, 3771-3774.

Grove, PM, Byrne, J.M. & Gillam, B.J. (2005) How configurations of binocular disparity determine whether stereoscopic slant or stereoscopic occlusion is seen. Perception, 34, 1063-1094.

Grove, PM; Brooks, KR; Anderson, BL, Gillam, BJ (2006) Monocular transparency and unpaired stereopsis Vision Research, 46, 3041-3053.

Hakkinen, J., & Nyman, G. (1996). Depth asymmetry in da Vinci stereopsis. Vision Research, 36, 3815–3819.

Hakkinen, J., & Nyman, G. (1997) Occlusion constraints and stereoscopic slant. Perception, 26, 29-38.

Hakkinen, J., & Nyman, G. (2001). Phantom surface captures stereopsis Vision Research 41 (2001) 187–199

Harris, J.M., Nefs, H.T. & Grafton, C.J. (2008) Binocular vision and motion-in-depth. Spatial Vision, 21, 531-547.

Hayashi, R., Maeda, T., Shimojo, S., & Tachi, S. (2004). An integrative model of binocular vision: A stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry. Vision Research, 44, 2367–2380.

Hess, RF; Wilcox, LM (1994) Linear And Nonlinear Filtering In Stereopsis, Vision Research, 34, 2431-2438.

von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. Vision Research, 40, 1955–1967.

Howard, I. P. (1995). Depth from binocular rivalry without spatial disparity. Perception, 24, 67–74.

Howard, IP (2002) *Seeing in Depth*, *I* Porteous Press.

Howard, IP; Wade, NJ (1996) Ptolemy's contributions to the geometry of binocular vision Perception, 1189-1201.

Howard, IP, & Rogers, BJ (2002) *Seeing in Depth*, *II* Porteous Press.

Howard, IP; Duke, PA (2003) Monocular transparency generates quantitative depth. Vision Research, 43, 2615-2621.

Jones, R. K. & Lee, D. N.(1982) Why two eyes are bette than one: the two views of binocular vision. JEP:HPP, 7, 30-40.

Jones, DG; Malik, J (1992)Title: Computational Framework For Determining Stereo Correspondence From A Set Of Linear Spatial Filters. Image And Vision Computing, 10, 699-708.

Julesz, B. (1971). Foundations of cyclopean perception. University of Chicago Press.

Kaye, M. (1978) Stereopsis without binocular correlation. Vis. Res., 18, 1013-1022.

Kumar, T. (1995). Stereopsis due to luminance differences in the two eyes. *Vision Research*, 35 (2), 255–262.

Kuroki, D. & Nakamizo, S. (2006) Depth scaling in phantom and monocular gap stereopsis using absolute distance information. Vision Research, 46, 4206-4216.

Langer, M. S. (2008) Surface visibility probabilities in 3D cluttered scenes. 10th European Conference on Computer Vision.

Lawson, R.B. & Gulick, W. L. (1967) Stereopsis and anomalous contour. Vis. Res., 7, 271-297.

Lawson, RB; Mount, DC (1967) Minimum Condition For Stereopsis And Anomalous Contour Science, 158, 804-&

Liu, L; Stevenson, SB; Schor, CM (1994) Quantitative Stereoscopic Depth Without Binocular Correspondence, Nature, 367, 66-69.

Liu, L., Stevenson, S. B., & Schor, C. M. (1997). Binocular matching of dissimilar features in phantom stereopsis. Vision Research, 37, 633–644.

Liu, L., Stevenson, S. B., & Schor, C. M. (1998). Vergence eye movements illicited by stimuli without corresponding features. Perception, 27, 7-20.

Makino, Y. & Yano, M. (2006) Pictorial cues constrain depth in da Vinci stereopsis. Vision Research, 46, 91-105.

Malik, J. Anderson, B.L. & Charowhas, C.E. (1999) Stereoscopic occlusion junctions. Nature Neuroscience, 2, 840-843.

Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. Science, 194, 283–287.

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. Proceedings of the Royal Society of London. Series B. Biological sciences, 204, 301–328.

Matsumiya, K; Howard, IP; Kaneko, H (2007) Perceived depth in the 'sieve effect' and exclusive binocular rivalry. Perception, 36, 990-1002.

McKee, SP; Bravo, MJ; Smallman, HS, & Legge, GE (1995) The Uniqueness Constraint And Binocular Masking, Perception, 24, 49-65.

McLoughlin, N. P., & Grossberg, S. (1998). Cortical computation of stereo disparity. Vision Research, 38, 91–99.

Mitchell, DE (1969) Qualitative Depth Localization With Diplopic Images Of Dissimilar Shape Vision Research, 9, 991-&

Mitsudo, H; Nakamizo, S; Ono, H (2005) Greater depth seen with phantom stereopsis is coded at the early stages of visual processing. Vision Research, 45,1365-1374.

Mitsudo, H; Nakamizo, S; Ono, H (2006) A long-distance stereoscopic detector for partially occluding surfaces Vision Research, 46, 1180-1186.

Nakayama, K., & Shimojo, S. (1990). da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. Vision Research, 30, 1811–1825.

Nefs, H. T. (2008) Three-dimensional object shape from shading and contour disparities. *Journal of Vision*, *8*(11):11, 1-16

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. Science, 249, 1037–1041.

Ono, H., Wade, N.J. & Lillakasa, L. (2009) Binocular vision: defining the historical directions. Perception, in press. doi:10.1068/p6130

Ono, H., Shimono, K. & Shibuta, K. (1992) Occlusion as a depth cue in the Wheatstone-Panum limiting case. Paerception & Psychophysics, 51, 3-13.

Ono, H; Mapp, AP; Howard, IP (2002) The cyclopean eye in vision: the new and old data continue to hit you right between the eyes. Vision Research, 42, 1307-1324.

Ono H, Wade NJ, Lillakas L (2002) The pursuit of Leonardo's constraint. Perception, 31, 83-102.

Ono, H; Lillakas, L; Grove, PM & Suzuki, M. (2003) Leonardo's constraint: Two opaque objects cannot be seen in the same direction. Journal Of Experimental Psychology-General 132, 253-265.

Ooi, T.L. & He. Z.J. (2006) Binocualr rivalry and surface boundary processing. Perception, 35, 581-603.

Pianta, M. J. & Gillam, B. J. (2003a) Paired and unpaired features can be equally effective in human depth perception. Vision Research, 43, 1-6.

Pianta, M. J. & Gillam, B. J. (2003b) Monocular gap stereopsis: manipulation of the outer edge disparity and the shape of the gap. Vis. Res., 43, 1937-1950.

Prince, S.J.D., Pointon, A.D., Cumming, B.G., and Parker, A.J. (2002b). Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. J. Neurophysiol. *87*, 191–208.

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. Neural Computation, 6, 390–404.

Richter (1977) Selections from the notebooks of Leonrdo da Vinci, Oxford Univ. press.

Sachtler, WL & Gillam, B (2007) The stereoscopic sliver a comparison of duration thresholds for fully stereoscopic and unmatched versions. Perception, 36, 135-144.

Saye, A; Frisby, JP (1975) Role Of Monocularly Conspicuous Features In Facilitating Stereopsis From Random-Dot Stereograms, Perception, 4, 159-171.

Shimojo, S., & Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry. Vision Research, 30, 69–80.

Shimojo, S., & Nakayama, K. (1994). Interocularly unpaired zones escape local binocular matching. Vision Research, 34, 1875–1881.

Shimojo, S., G.H. Silverman and K. Nakayama. (1988) An occlusion-related mechanism of depth perception based on motion and interocular sequence. Nature 333: 265-268, 1988.

Strong, D.S. (1979) Leonardo on the eye. An English translation and critical commentary of MS. D in the Bibliotheque Nationale, Paris, with Studies on Leonardo’s methodology and theories on optics. Garland Publishing, New York.

Tsai, J. J., & Victor, J. D. (2000). Neither occlusion constraint nor binocular disparity accounts for the perceived depth in the ’sieve effect’. Vision Research, 40, 2265–2276.

Tsai, J. J., & Victor, J. D. (2005). Binocular depth perception from unpaired image points need not depend on scene organization. Vision Research, 45, 527–532.

Tsao, D.Y., Conway, B.R. & Livingstone, M. S. (2003) Receptive Fields of Disparity-Tuned Simple Cells in Macaque V1 Neuron, 38, 103–114,

Wade NJ, Ono H, Lillakas L (2001) Leonardo da Vinci's struggles with representations of reality. Leonardo, 34, 231-235.

Watanabe, O., & Fukushima, K. (1999). Stereo algorithm that extracts a depth cue from interocularly unpaired points. Neural Networks, 12, 569–578.

Wheatstone, C. (1838). Contributions to the physiology of vision. Part the first: On some remarkable and hitherto unobserved phenomena of binocular vision. Phil. Trans. R. Soc. II, 371-394.

Wilcox, LM & Allison R. (2009) 'Review of 2^{nd} order stereopsis' For Vision Research.

Wilcox, LM; Hess, RF (1995) D-Max For Stereopsis Depends On Size, Not Spatial-Frequency Content, Vision Research, 35, 1061-1069.

Wilcox, LM; Hess, RF (1996) Is the site of non-linear filtering in stereopsis before or after binocular combination? Vision Research, 36, 391-399.

Wilcox, LM & Lakra, DC (2007) Depth from binocular monocular regions in stereoscopic images of natural scenes. Perception, 36,830-839.

Wilcox, L. M., Harris, J. M. & McKee, S. P. (2007) The role of binocular stereopsis in monoptic depth perception. Vis. Res., 47, 2367-2377.

Zitnick CL, & Kanade T (2000) A cooperative algorithm for stereo matching and occlusion detection. IEEE Transactions on Pattern Analysis and Machine Intelligence., 22, 675-684.

Figure 1

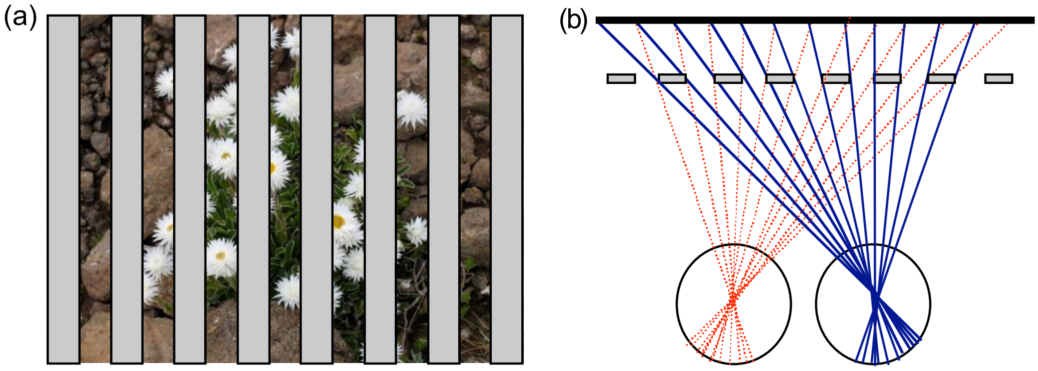
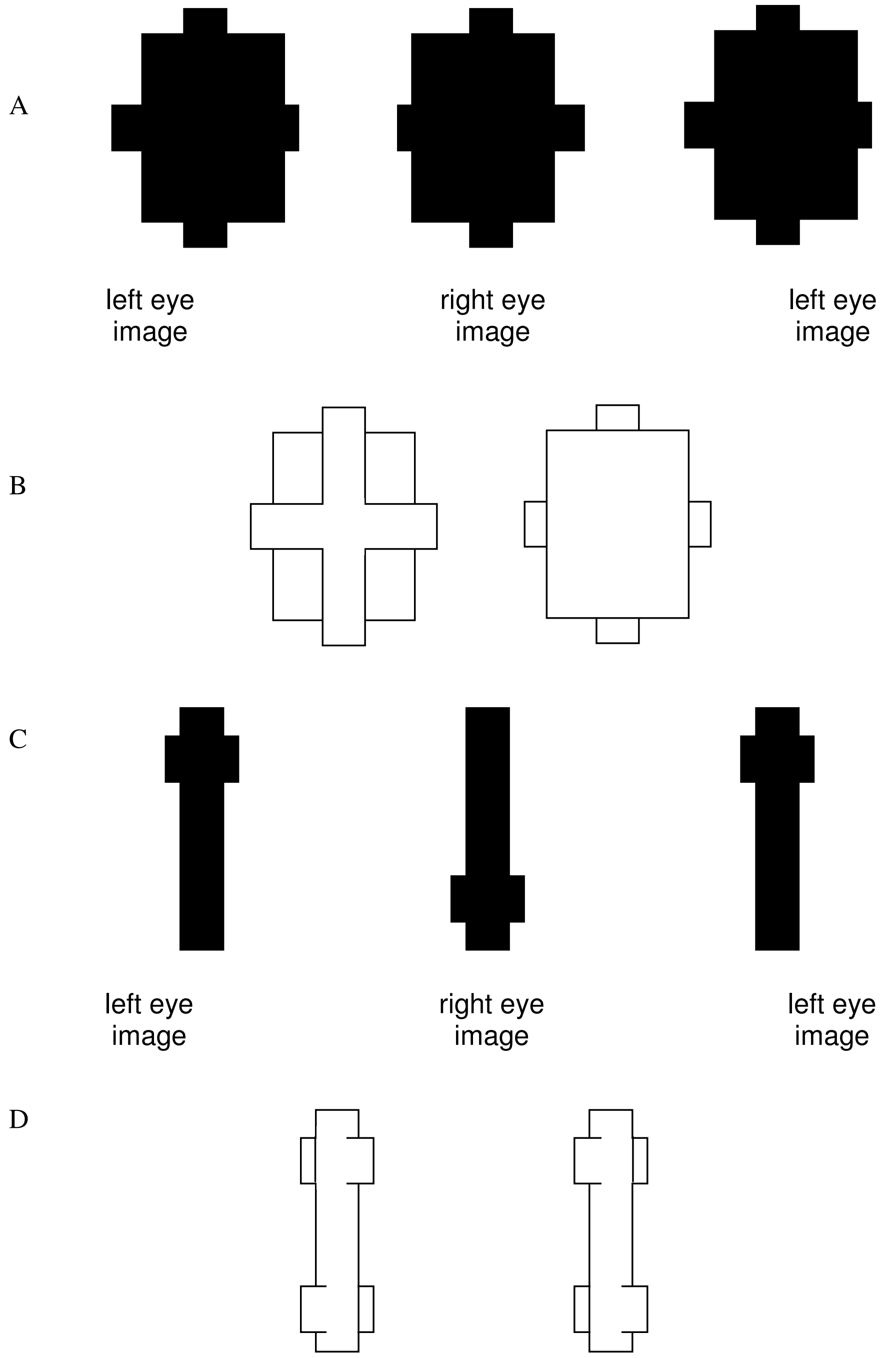


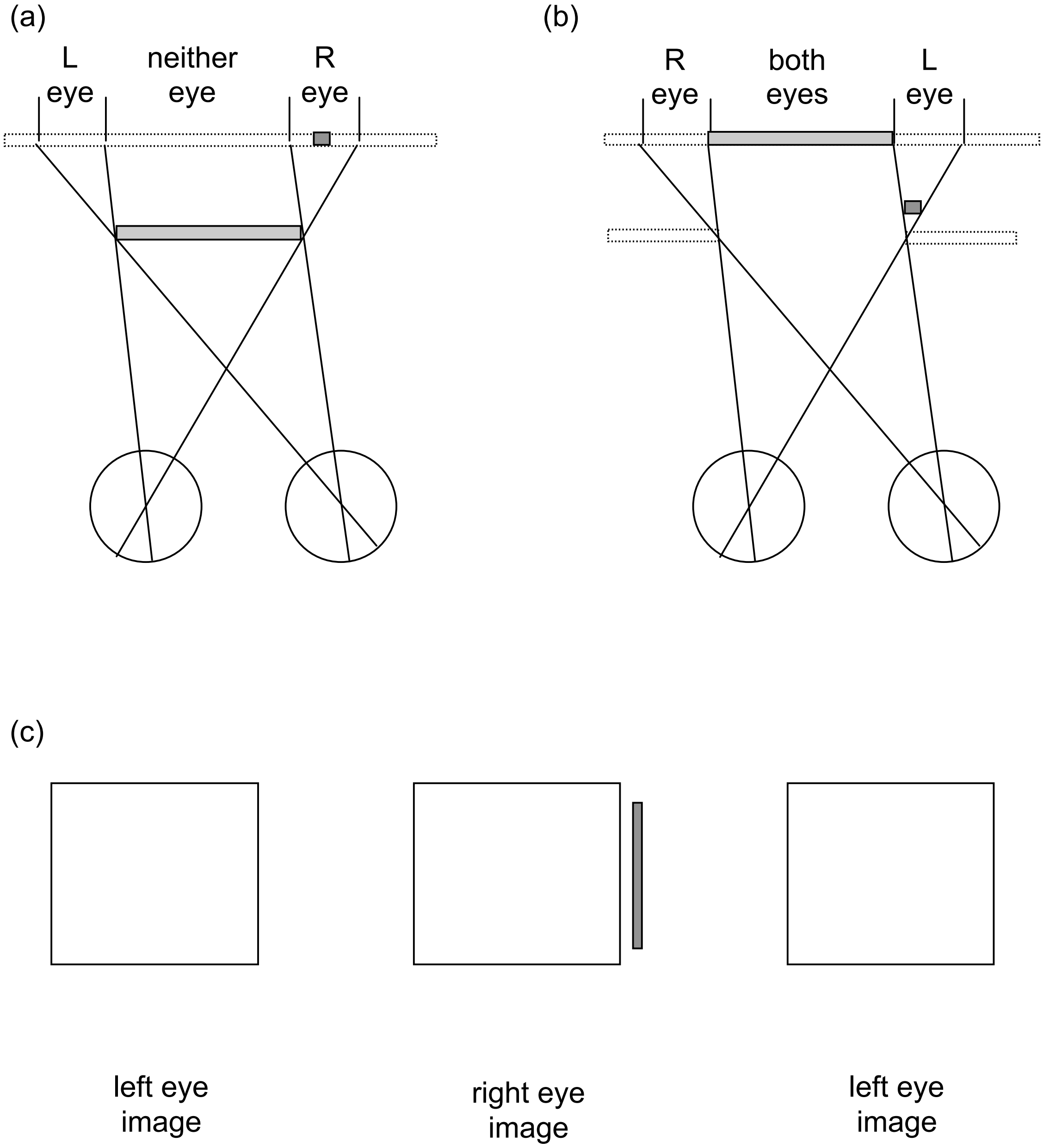
Figure 1. (a) An example image showing one eye's view of a background when occluded by a 'picket fence' of foreground grey objects. (b) shows how each eye sees different regions of the background behind the occluder, depending on the separation between occluder and background. This example shows geometry for the situation where each portion of background is only seen by one eye.

Figure 2



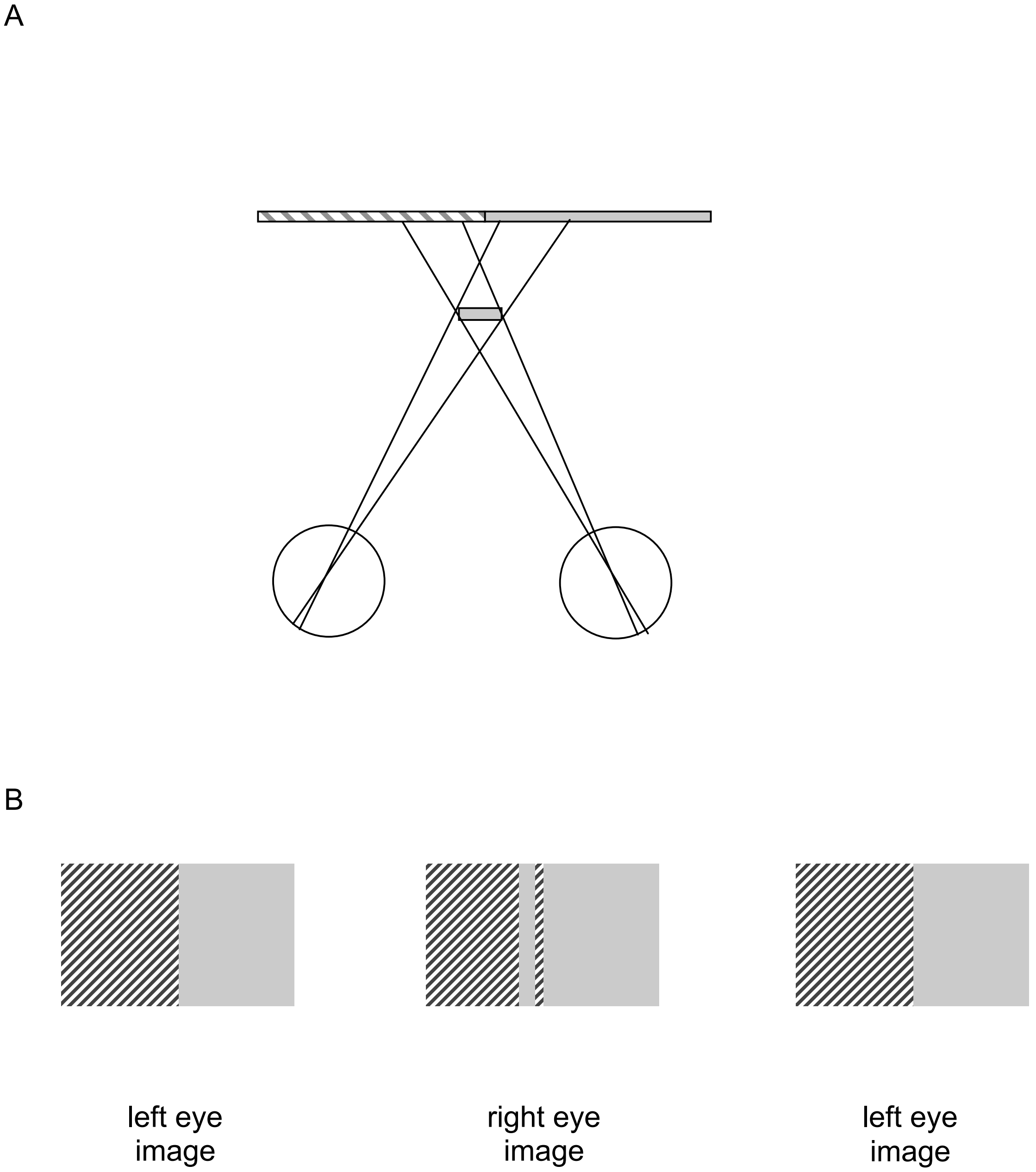
Figures illustrating the powerful effects that half-occlusions can deliver. Panel A shows right and left eye views of a stereogram. When the left and middle are cross-fused (or middle and right for uncrossed fusion) a percept like that in the left of panel B occurs. When the middle and right are cross-fused, a percept like that in the right of panel B occurs. Notice that very clear contours are induced. When the left and middle figure in panel C are cross-fused, a percept like that in the left of panel D occurs, with the monocularly viewed ‘wings’ delivering a percept of a surface slanting in depth. The slant goes in the opposite direction when the middle and right figures in panel C are cross-fused. Figure is adapted, with permission, from figures 2 and 6, Ehrenstein, W. H., & Gillam, B. J. (1998) Early demonstrations of subjective contours, amodal completion, and depth from half-occlusions: ‘‘Stereoscopic experiments with silhouettes’’ by Adolf von Szily (1921). Perception, 27, 1407–1416.

Figure 3



The geometry of occlusion. (a) A foreground rectangle (shown as light grey for illustrative purposes only) sits in front of a continuous background (shown by dotted lines because it is not patterned or coloured). Lines of sight from each eye are drawn to illustrate that, in each case, parts of the background are visible to one eye, but not to the other. A single bar is visible to the right eye only. (b) A background rectangle is viewed through a hole in an invisible foreground object. A single bar is visible to the left eye only. (c) The stimuli used by Nakayama & Shimojo (1990). Cross-fusing the right and middle figures (or uncross fusing the left and middle figures) results in their 'valid' case: the monocular bar appears behind the rectangle. Cross-fusing the left and middle figures gives their 'invalid' case.   Figure is adapted, with permission, from figure 1, Nakayama, K., & Shimojo, S. (1990).

Figure 4



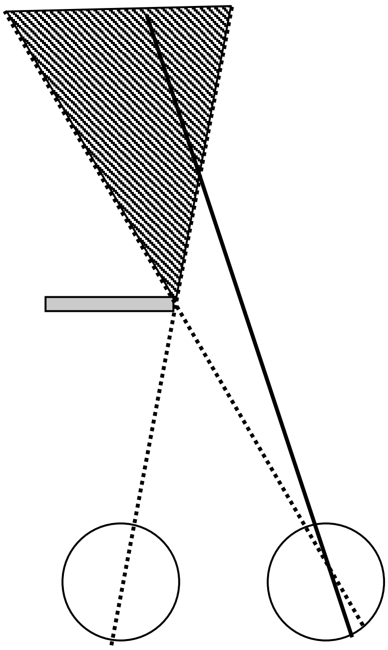
left eye image

right eye image

left eye image

(A) A top-down view showing an example of monocular camouflage. The small foreground target is camouflaged in the left eye's view (in B below), and therefore only visible in the right eyes view (B, centre panel). If the middle and right panel of B are cross-fused (or left and middle uncross-fused), the small monocular bar appears in front of the background.

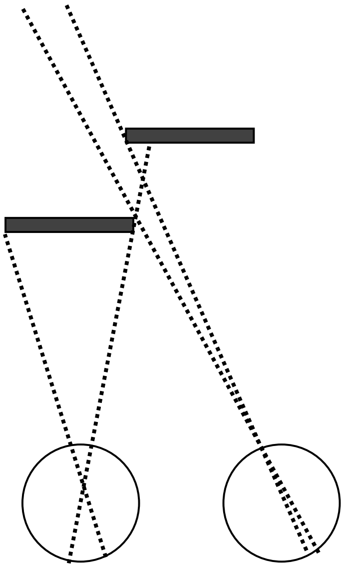
Figure 5



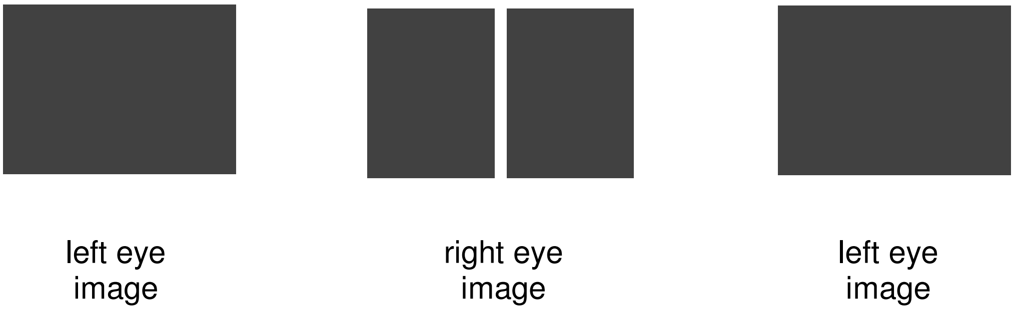
An illustration of the depth constraint zone, shown by the striped region. This figure shows that a monocular point or line (the solid line in the right eye) is consistent with an object at any depth within the zone, along the points line of sight. Figure adapted, with permission, from figure 3, Nakayama, K., & Shimojo, S. (1990). da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. Vision Research, 30, 1811–1825.

Figure 6

(a)

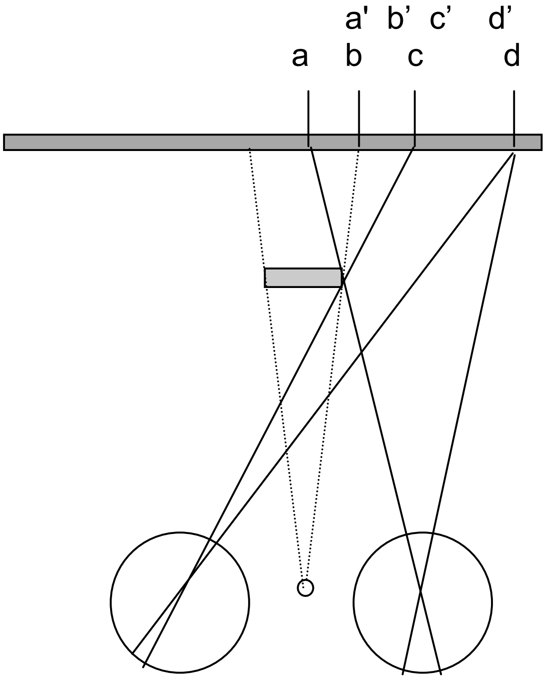


(b)



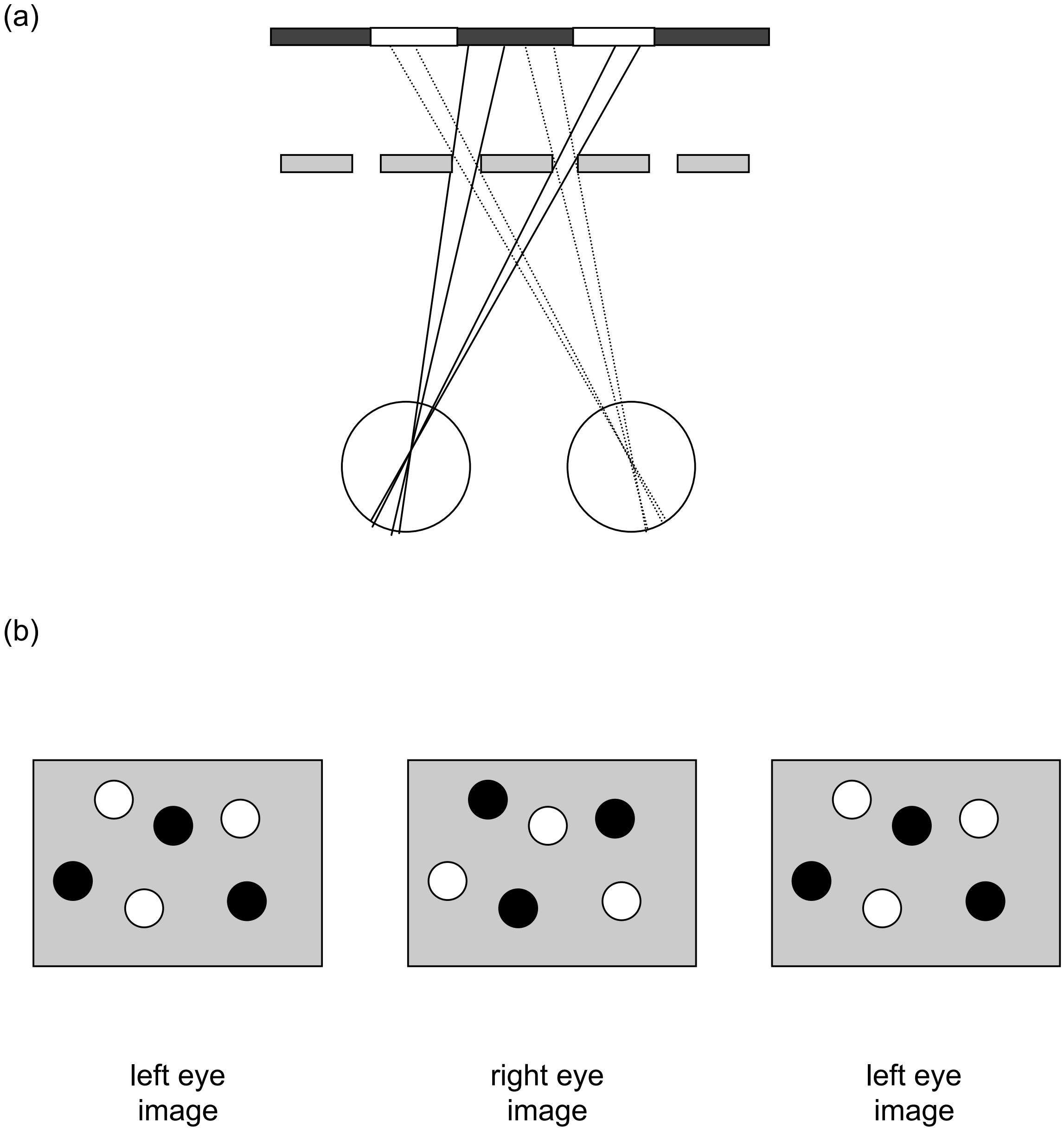
(A) illustrates the geometrical situation consistent with monocular gap stereopsis (Gillam et al, 1999). The right eye can see a bright background through a gap between two dark foreground objects. For the left eye, there is no gap as it is occluded by the foremost object. (B) Stereograms that illustrate depth in this situation. If the middle and right figure are cross-fused the left side should appear closer than the right. If left and middle are cross-fused, the right side will appear closer. Figure adapted, with permission, from figure 1a, Gillam, B; Blackburn, S; Nakayama, K (1999) Stereopsis based on monocular gaps: Metrical encoding of depth and slant without matching contours. Vision Research, 39, 493-502.

Figure 7



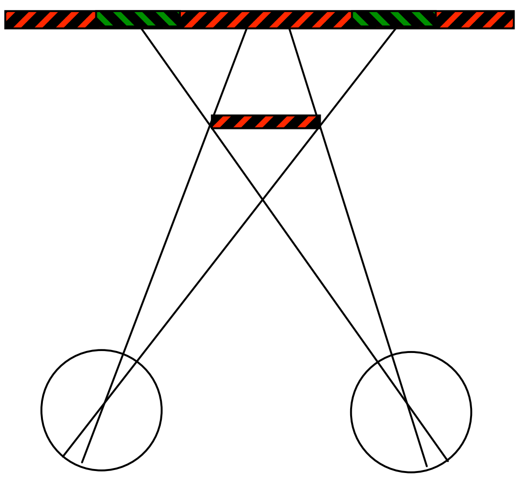
When both eyes view a scene with a foreground object, the background region from a-d is visible to one eye or the other, or both. This is a larger region than could be seen as if from a single eye at a central location, midway between the eyes (it would see b-d). Figure adapted, with permission, from figure 2, Mapp A.P. and Ono, H. (1999) Wondering about the wandering cyclopean eye. Vision Research, 39, 2381-2386.

Figure 8

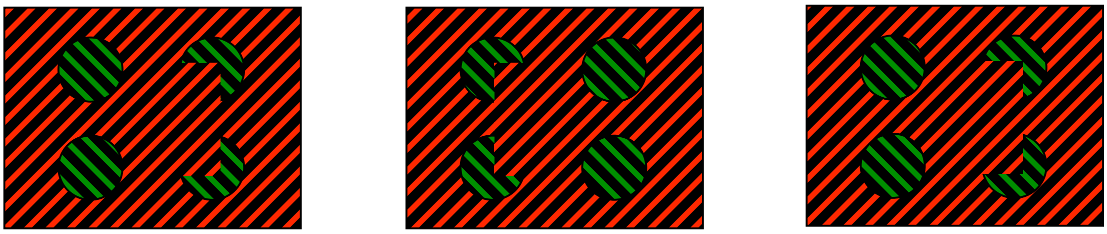


(a) Geometry consistent with stimuli used in the sieve effect. The observer views a black and white background through a series of holes in a foreground occluder (the sieve). (b) Stereograms to illustrate the effect. The regions within the circular patch may appear behind the grey rectangle, whichever pair of images are fused. Figure adapted, with permission, from Howard, I. P. (1995). Depth from binocular rivalry without spatial disparity. Perception, 24, 67–74.

Figure 9
(a)



(b)



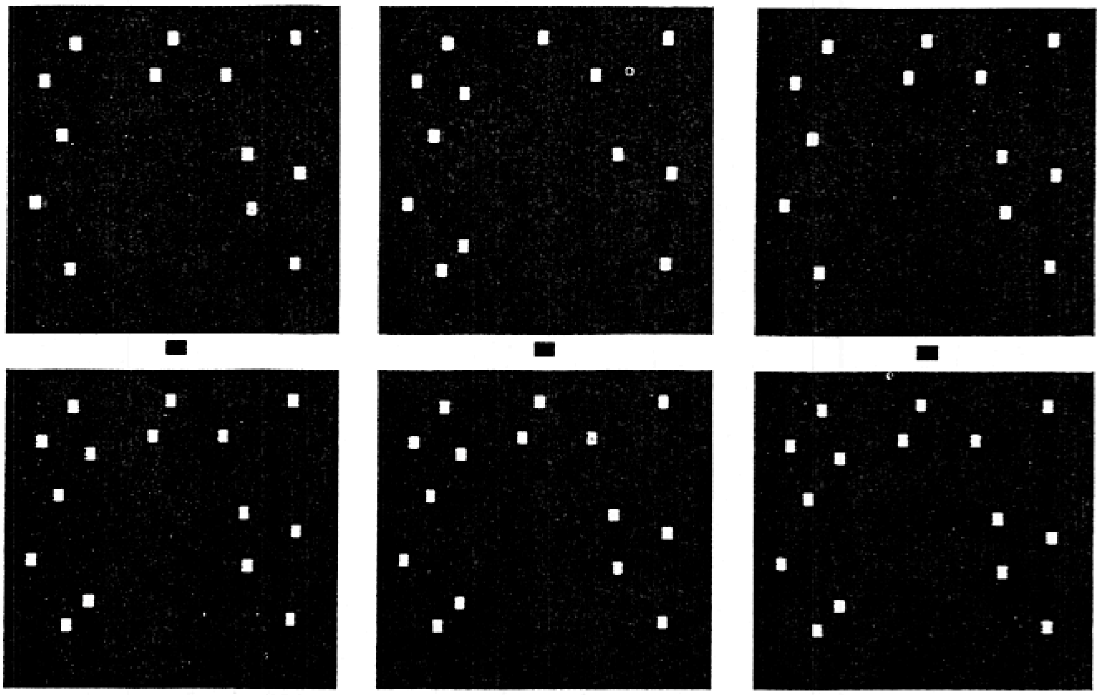
left eye image

right eye image

left eye image

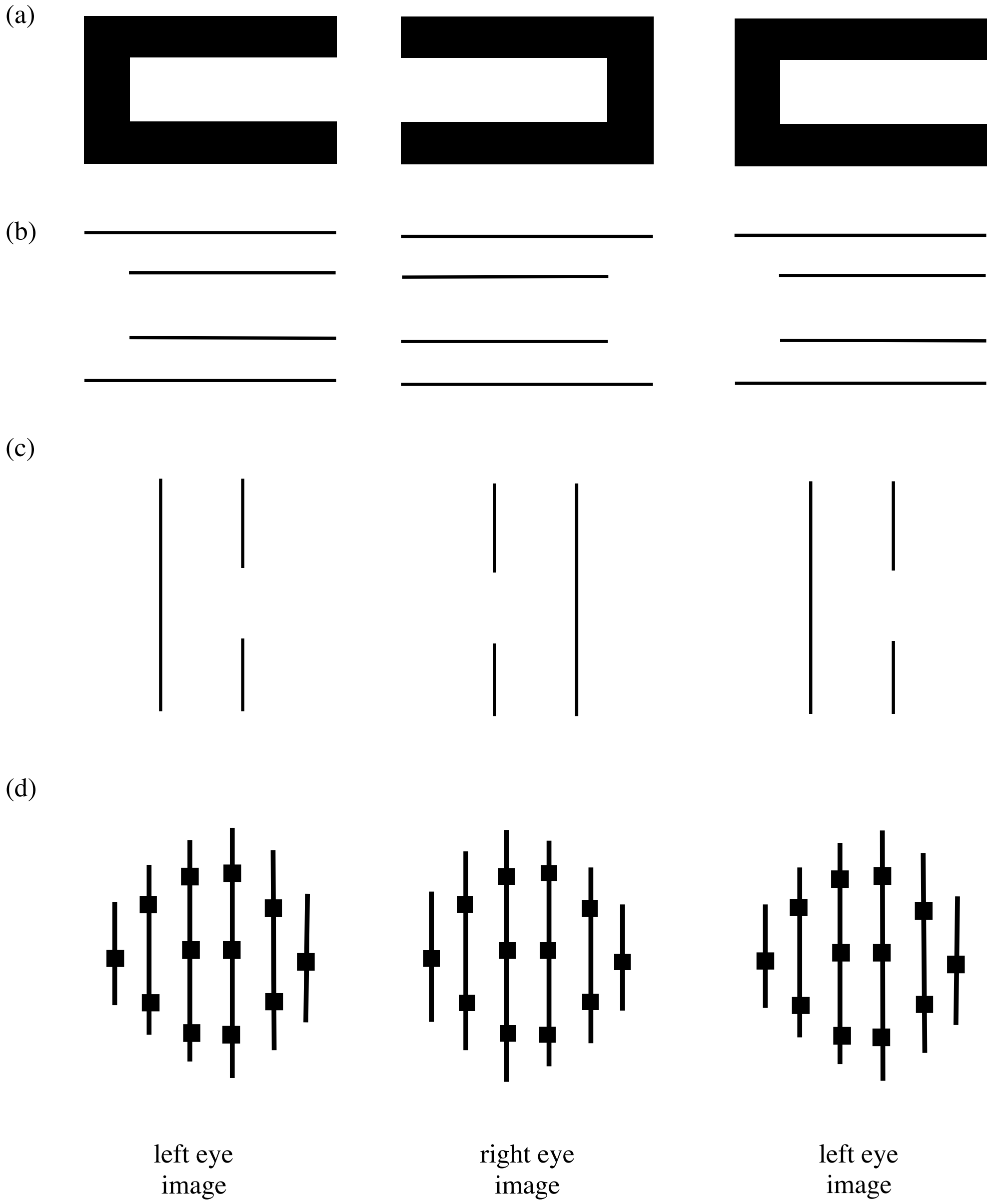
(a) A top-down view showing a foreground occluder, with the same pattern as most of the background. A pair of circular regions with a different pattern are each only partially visible in one or other eye. (b) Stereopairs consistent with the above geometry. If right and middle panels are cross-fused the pattern is stable and a foreground occluder perceived. If the other pair are cross fused there is no such stability and rivalry occurs. After van Bogaerts, Ooi & He, 2008, with permission.

Figure 10



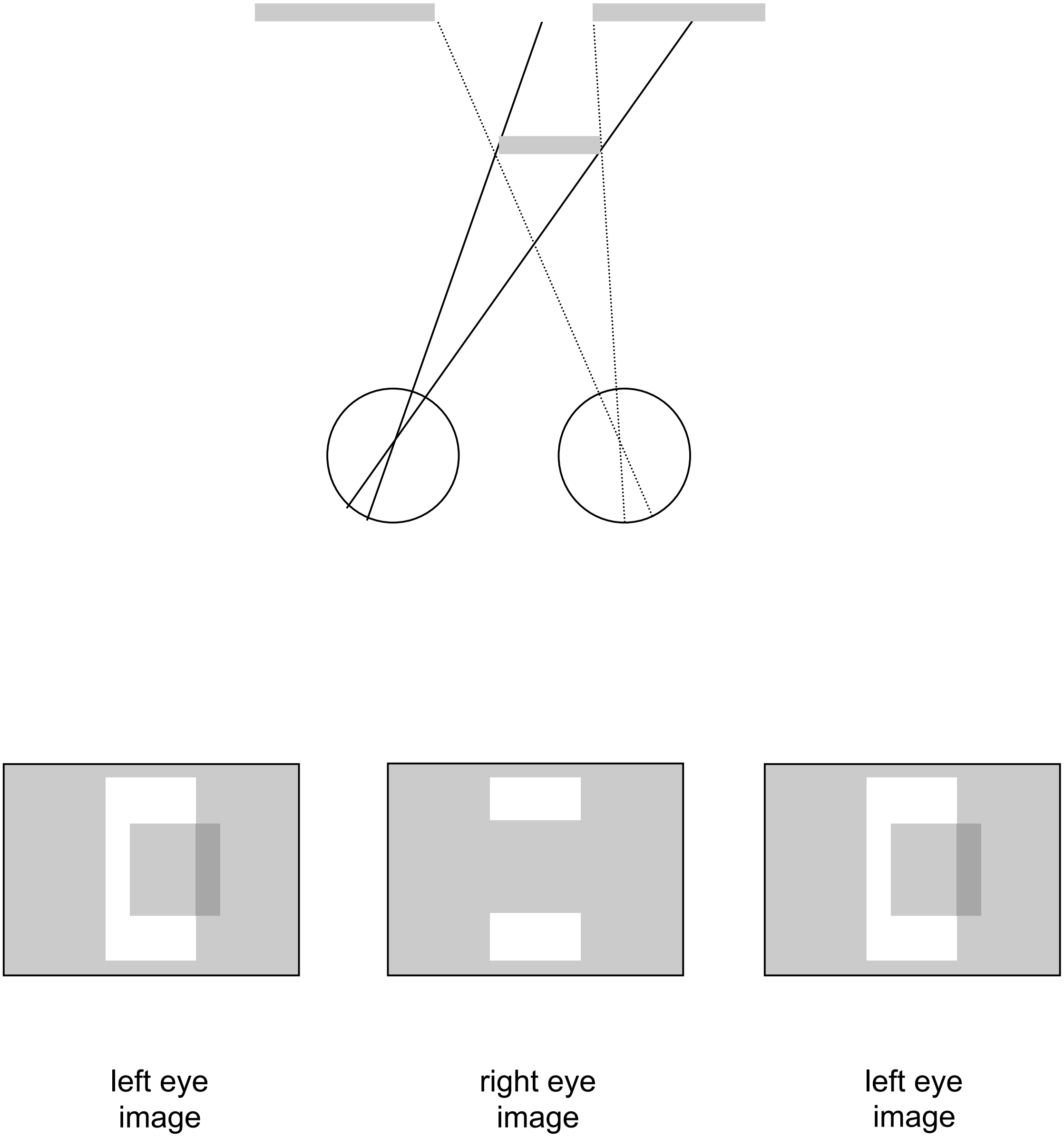
Reproduced from Nakayama & Shimojo (19950), with permission from the authors. In the upper stereo pairs two points are missing from each of the left and right eye views. A clear phantom occluding surface is perceived when the left and middle panels are cross-fused (or middle and right panels divergently fused). In the lower panels all points are visible in left and right eye views and no occluding contour is seen.

Figure 11



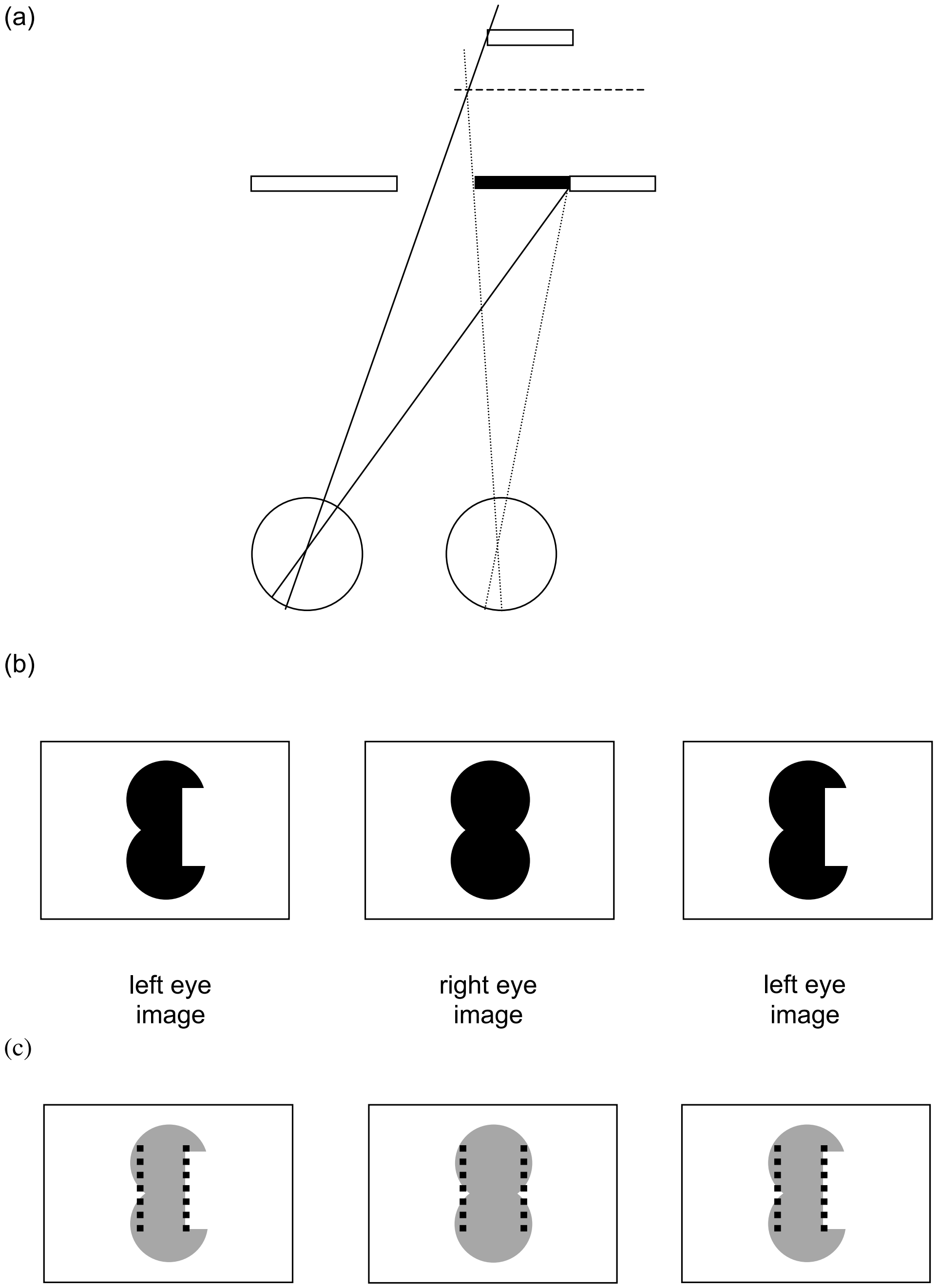
Stereograms illustrating several different stimuli that achieve phantom stereopsis. (a) The central white patch appears in front of the black background when middle and right panels are cross-fused (left and middle for uncrossed-fusion). Figure adapted, with permission, from figure 1, Liu, L; Stevenson, SB; Schor, CM (1994) Quantitative Stereoscopic Depth Without Binocular Correspondence, Nature, 367, 66-69. (b) Gillam used these stereopairs to demonstrate that stereoscopically matchable features do exist in (a). Figure adapted, with permission, from figure 1c, Gillam, B (1995) Matching Needed For Stereopsis, Nature, 373, 202-203. (c) Stimulus used to illustrate that phantom stereopsis can occur when matchable features have bene removed. When the middle and right panels are cross-fused a white phantom rectangle is seem in depth. Figure adapted, with permission, from figure 2, Gillam, B., & Nakayama, K. (1999). Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. Vision Research, 39, 109–112. (d) Vertical offsets of bars can result in phantom stereopsis. When the middle and right panel are cross-fused the pattern appears as if behind a phantom window that occludes the pattern at the edges. Figure adapted, with permission, from figure 1, Anderson, B. L. (1994) The role of partial occlusion in stereopsis. Nature, 367, 365-368.

Figure 12



(a) Viewing geometry consistent with monocular transparency. (b) Stereograms used to illustrate this effect. When the middle and right images are cross-fused a grey square appears in front of the background. When the left and middle images are cross-fused a grey square appears behind the partly transparent white rectangle. Figure adapted from figure 5, Howard, IP; Duke, PA (2003) Monocular transparency generates quantitative depth. Vision Research, 43, 2615-2621.

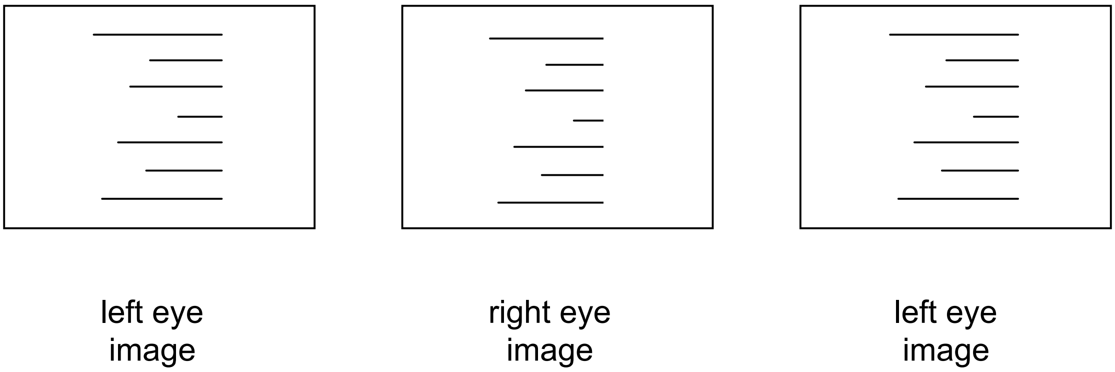
Figure 13



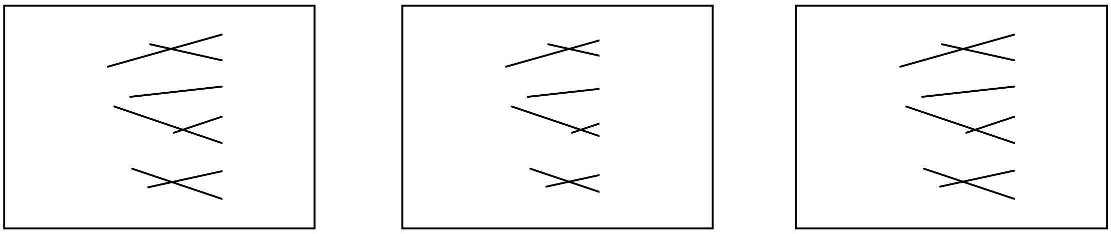
(a) Geometry consistent with the intrusion stereogram, where a white rectangle appears to intrude into the black background, seen behind a window formed by the outline of the black part of the figure. (b) Stereograms used to depict depth from intrusion. If the middle and right images are cross-fused the white rectangle appears behind. If the left and middle images are cross-fused, a white rectangle appears in front of the black object. (c) If some coarse mechanism were to extract edges (shown by dotted vertical lines), these could be used by a binocular disparity mechanism to obtain depth. Figure partially adapted, with permission, from figure 2, Cook, M., & Gillam, B. (2004). Depth of monocular elements in a binocular scene: The conditions for da Vinci stereopsis. Journal of Experimental Psychology. Human Perception and Performance, 30, 92–103.

Figure 14

(a)



(b)



(a) Line stereograms, after Gillam & Grove (2004), with the lines shorter by a fixed amount in the right eye view. When middle and right stereo pairs are cross-fused, an apparent foreground occluder is visible on the right of the pattern. (b) Oblique line stereograms, after Grove, Byrne & Gillam (2005). As before, the lines are cut, as if by a vertical occluder, in the right eye image and the resulting foreground occlusion in the fused image is stronger than that for the horizontal lines in (a).