Dissertation for the Degree of Doctor of Philosophy in Psychology presented at Uppsala University in 2003

ABSTRACT

Poom, L. 2003. Binding Three Kinds of Vision. *Acta Universitatis Upsaliensis*. Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 121. 60 pp. Uppsala. ISBN 91-554-5518-2.

Pictorial cues, together with motion and stereoscopic depth fields, can be used for perception and constitute 'three kinds' of vision. Edges in images are important features and can be created in either of these attributes. Are local edge and global shape detection processes attribute-specific? Three visual phenomena, believed to be due to low-level visual processes, were used as probes to address these issues. (1) Tilt illusions (misperceived orientation of a bar caused by an inducing grating) were used to investigate possible binding of edges across attributes. Double dissociation of tilt repulsion illusions (obtained with small orientation differences between inducer and bar) and attraction illusions (obtained with large orientation differences) suggest different mechanisms for their origins. Repulsion effects are believed to be due to processes in striate cortex and attraction because of higher level processing. The double dissociation was reproduced irrespective of the attributes used to create the inducing grating and the test-bar, suggesting that the detection and binding of edges across attributes take place in striate cortex. (2) Luminance-based illusory contour perception is another phenomenon believed to be mediated by processes in early visual cortical areas. Illusory contours can be cued by other attributes as well. Detection facilitation of a near-threshold luminous line occurred when it was superimposed on illusory contours irrespective of the attributes used as inducers. The result suggests attribute-independent activation of edge detectors, responding to real as well as illusory contours. (3) The performance in detecting snake-like shapes composed of aligned oriented elements embedded in randomly oriented noise elements was similar irrespective of the attributes used to create the elements. Performance when the attributes alternated along the path was superior to that predicted with an independent channel model. These results are discussed in terms of binding across attributes by feed-forward activation of orientation selective attribute-invariant cells (conjunction cells) in early stages of processing and contextual modulation and binding across visual space mediated by lateral and/or feedback signals from higher areas (dynamic binding).

Key words: Attribute-invariance, binding, contours, edge linking, gestalt-laws, illusory contour, lateral interactions, neuronal mechanisms, tilt effect, vision.

Leo Poom, Department of Psychology, Uppsala University, Box 1225, SE-751 42 Uppsala, Sweden

ISSN 0282-7492
ISBN 91-554-5518-2
Printed in Sweden by Kopieringshuset, Uppsala, 2003.

This thesis is based on the following three studies:

Study 1. Poom, L. (2000). Inter-attribute tilt effects and orientation analysis in the visual brain. *Vision Research*, 40, 2711-2722.

Study 2. Poom, L. (2001). Visual summation of luminance lines and illusory contours induced by pictorial, motion, and disparity cues. *Vision Research*, 41, 3805-3816.

Study 3. Poom, L. (2002). Visual binding of luminance, motion, and disparity edges. *Vision Research*, 42, 2577-2591.

Acknowledgements

First of all I wish to express my gratitude to Erik Börjesson, my supervisor, for supporting me and for many creative discussions during my years as a graduate student. I thank the 'human encyclopedia' Henrik Olsson for many stimulating brainstorming sessions, as well as for his comments on a preliminary version of this thesis, which resulted in a much improved final product. I am also in dept to Sture Eriksson who opened my eyes for the vision sciences when I was an undergraduate student. I also thank Claes von Hofsten for assisting with valuable comments on this thesis and Leslie Shaps for a fast proof-reading process. I would like to thank the administrative and technical staff at the Department of Psychology at Uppsala University for providing an environment that has made this thesis possible. Special thanks to our instrument maker, Lars-Erik Larsson, for the technical constructions. Last but not least, I thank my friend Gary Wife for providing insights and basic facts about life outside the lab, and for not beating me up when I smashed his car.

Leo Poom
January, 2003

Contents

A brief view on the problem 7

Inverse optics and edges in images 11

Seeing edges by 'three kinds of vision' 12

Pictorial edges *12*

Temporal edges *16*

Cyclopean edges *18*

Binding edges across attributes and space 20

Tilt illusions (Study 1) 22

Probing binding with tilt illusions *24*

Illusory contours (Study 2) 29

Probing binding with illusory contours *31*

Seeing 'snakes' in noise (Study 3) 35

Probing binding with 'snakes' *38*

Binding the results 44

A new view on vision 46

Brunswik, Bayes, and binding *46*

Binding by conjunction cells *47*

Dynamic binding and computational modeling *49*

Edge detection precedes surface 'filling-in' *50*

Concluding remarks *52*

References 54

"Illusions have the same aesthetic appeal to a psychologist that Drosophila mutants such as Antennepedia have to the developmental biologist. They are bizarre and beautiful but, hopefully, they will also allow us to peak into the black box and tell us something about the underlying mechanisms."

(Ramachandran, 1992, p. 89)

A brief view of the problem

"Segmentation is one of the most difficult tasks in image processing" (Gonzalez & Woods, 1993, p. 413)

When we open our eyes, the visual impressions seem to impinge on us with no delay. No time consuming processes are required and no intermediate processes mediating the visual experience are revealed to us. Seeing is such an everyday experience that it may be taken-for-granted until one attempts to explain it. One fundamental task for the visual system is figure-ground segmentation where it is decided what image parts belong to the same object. This task has been shown to be extremely difficult, which stands in sharp contrast to our visual impressions of what constitutes figure and ground in normal viewing conditions. How is this computationally difficult problem solved in the brain? Visualizing an image in some way to prevent the normal interpreting processes to operate can highlight the problem. Transforming an image from the normal representation in luminance levels to a landscape representation can do this (Morgan, 1996). Figure 1a is easily seen as two trees with branches and leaves, and the depth relationships are easily recognized because the image mimics a typical retinal image resulting from natural viewing conditions. The brightness or luminance levels from Figure 1a are drawn as heights of hills across the image in Figure 1b preserving the information[1]. Nevertheless, inspection of Figure 1b shows that the visual system is unable to use this landscape representation to make a figure-ground separation and recognize the trees.

Fragments of information in images are somehow bound together in the visual system resulting in perception of shapes and patterns. There are several attributes that can be used as information-bearing media in this process. Attributes, as used throughout the text, refer to image surface properties. The most widely recognized among these are the *pictorial* ones (luminance-, color-, and texture-fields). Artists use spatial modulations of these attributes to mimic retinal images resulting in illusions of depth and shape on canvas. Another attribute is the *motion-field* in visual images, resulting from the relative motion between the eye and environmental layout. Furthermore, we are equipped with two eyes providing us with two partly overlapping, slightly different images of the surroundings. Still, we see only one fused cyclopean image from which we obtain stereoscopic depth perception. This *cyclopean depth-field* provides the final attribute.

Figure-ground segmentation carried by pictorial fields requires integration of image point in one eye across the same retinal image. Segmentation carried by the motion-field is temporally defined because it requires integration of image points across time. Finally, figure-ground segmentation carried by the cyclopean field requires integration of image points between the eyes. Spatial and temporal modulations of these attributes (reviewed

[1] The depths of the cavities may represent the corresponding photoreceptor activity across a population of receptors in the eye in that the photoreceptors are inhibited by light and stimulated by darkness.

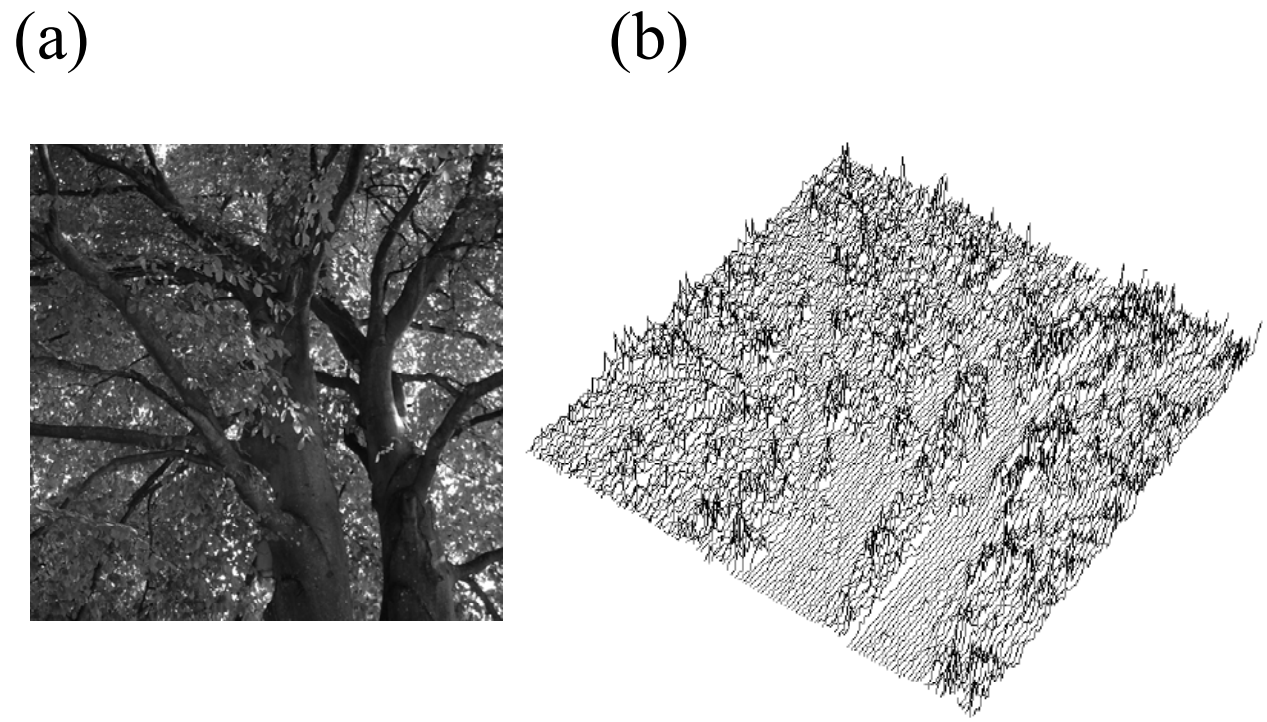

Figure 1. Although all information from the image in **(a)** is preserved in the landscape representation in **(b)** it is hard to recognize the original image by inspection because the representation has been transformed in a way that prevents the visual system from using it. In addition, the image landscape-plot is a snapshot of the time varying individual receptor activities caused by relative motion between observer and the environmental layout. Further, the eyes are involved in jittering motions about every 100 ms, even during gaze fixation, so the individual receptor activities may show strong fluctuations.

by Regan, 2000) create our visual world, and the three qualitatively different modes of integration constitute the 'three kinds of vision'.

Visual segmentation can be achieved by a single attribute but is usually accomplished by several attributes in conjunction. However, as noticed by Regan (2000), if independent attribute-specific channels are used for figure-ground segmentation, this can lead to a problem that is not so obvious:

> "*Spatial information about any given object can be encoded in terms of any one of five kinds of spatial contrast, and the early processing of any one of these different kinds of contrast is to some extent independent of processing of any other kinds of contrast. Nevertheless, in everyday life we usually see only one version of any given object rather than several versions of the same object.*" (p. 376)

This means that information from numerous attributes and from various locations in the image is integrated. In what way information is combined is still unknown.

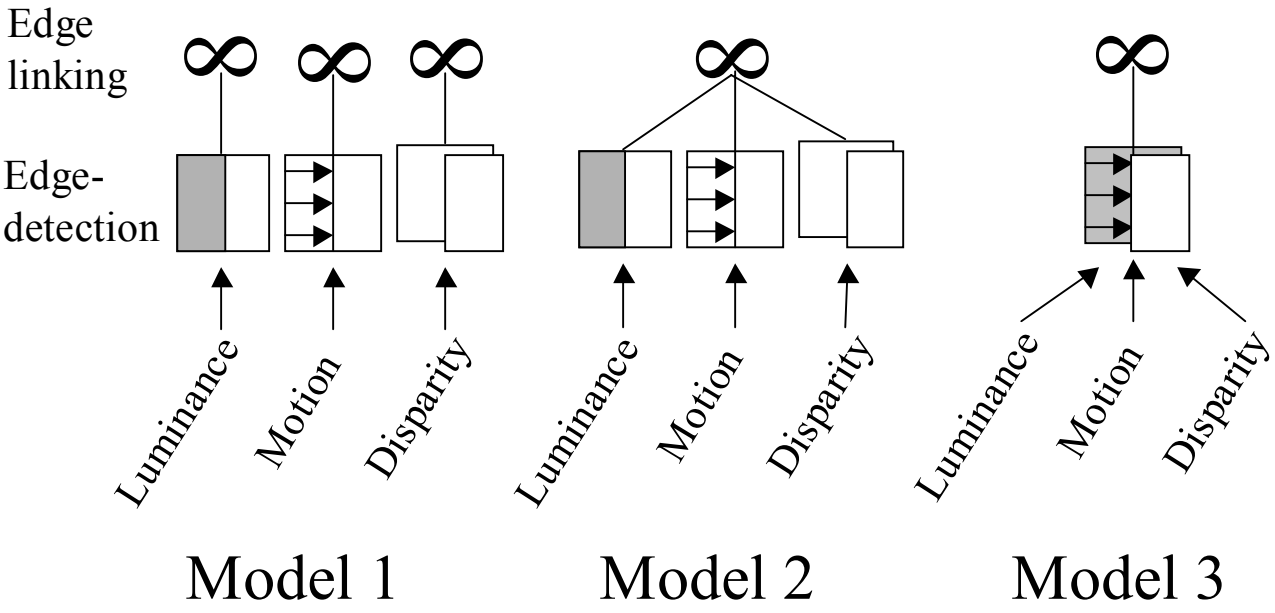

Figure 2. Three hypothetical models for form perception are depicted. Local edge segments are detected at the first stage and linked together in the second stage. Signals may go both up-steams and down-streams in a recursive process, allowing for contextual modulation at the edge detection stage. Model 1: Edge linking from three attributes (luminance, motion, and cyclopean vision or disparity) is processed along separate pathways. Model 2: Local edges are detected with attribute-specific detectors, but edge linking is attribute-invariant. Model 3: Attribute-invariance occurs already at the edge detecting stage.

From a neurophysiological perspective, the activities of a large number of cells widely distributed in the brain have to be assembled. How this is accomplished has been labeled the *binding problem* (von der Malsburg, 1995). Binding may be achieved in two ways. One way is by convergence of multiple signals on common target cells. Another way is to enhance the signals to be bound by increasing their salience relative to other unbound signals. The first kind of binding has been labeled *'binding by conjunction cells'* and the second has been labeled '*dynamic binding*' (Singer, 1999). The questions posed in this thesis address on what stage binding across visual space and binding of the three kinds of vision (i.e. luminance, motion, and cyclopean edges) take place. Another question is whether binding by conjunction cells or dynamic binding is used to bind signals across attributes and across space, respectively. Shortly, does binding across attributes precede or follow binding across space, and what are the principles used for the binding?

Figure 2 illustrates three possible two-stage models for binding across attributes and space. Edges from various attributes are detected at the first stage, either by attribute-selective (Model 1 and 2) or attribute-invariant edge detectors (Model 3). At a second stage, grouping occurs, either by attribute-selective channels (Model 1), or by a single grouping process that operates across attribute-selective edge detectors (Model 2) or on

attribute-invariant edge detectors (Model 3). The grouping processes may be recursive allowing signals go in either direction between the edge detection and linking stages.

Different levels of understanding are required to probe the functioning of the visual system (Marr, 1982). First, analyses of images typically encountered in the environment (natural images) are required to gain knowledge of what image properties reflect useful information about the surroundings. Brunswik (1947; 1955) was probably first to emphasize the importance of analyzing natural stimuli to understand perception. Gibson (1950; 1979) adapted similar ideas in the widely recognized ecological approach to perception. Second, computational models are required to understand the transition from image to output. Typically, human performance in an experimental task is compared with the model performance in a similar task. The model performance should be close to human performance, or else the model is dismissed. Finally, neurophysiological investigations may be required to differentiate among possible implementations of a specific model. During the last decades, this strategy has lead to an impressive gain of understanding as to what actually goes on during our daily experiences of seeing.

Initially, it was widely believed that visual subsystems break down the visual scene into basic components such as brightness, color, form, depth, motion, and texture and then combine these components to create our visual perception of the world. Based on computational analysis, Marr (1982) formulated the principle of modular design whereby different components of the visual image are processed more or less independently:

> *"This principle is important because if a process is not designed in this way, a small change in one place has consequences in many other places. As a result, the process as a whole is extremely difficult to debug or improve, whether by a human designer or in the course of natural evolution, because a small change to improve one part has to be accompanied by many simultaneous, compensatory changes elsewhere."* (p. 102)

In an influential paper Livingstone and Hubel (1987) found support for this view by correlating a large number of psychophysical, anatomical, and physiological studies. They demonstrated that the visual system is composed of several subsystems:

> *"Introspection suggests that visual perception can be subdivided into several subprocesses. If asked to list these, most people would include form, color, movement, depth, and perhaps texture. The intuitive impression that vision is multipartite, that it comprises several systems, has been supported by centuries of human psychophysics and by some recent anatomical and physiological studies in primates."* (p. 3416)

It is true that such image attributes as color, movement, depth, and texture may convey information about different properties of the world. On the other hand, different attributes can be used for signaling the same message. For example, the information carried by luminance-edges may convey the same information to the perceiver as texture-, color-,

motion-, and stereoscopic depth-edges: namely about the figure-ground relationships and the shapes of objects.

To what extent are edges from the different attributes independently processed? Before going to the investigations addressing these issues, it might be helpful to discuss what edges in images are good for to perceivers, and to inspect the attributes used to mediate the edges in the experiments.

Inverse optics and edges in images

"The true heart of visual perception is the inference from the structure of an image about the structure of the real world outside." (Marr, 1982, p. 68)

Following the laws of optics, a specific three-dimensional configuration in the world projected on a two-dimensional surface, such as the image projected on the film in a camera, results in a uniquely specified image. Contrary to the purpose of man-made optical systems such as the camera, the visual system deals with the reverse problem. Namely, to decide what configuration of objects and surfaces in the physical world created the image, not what the image will be like given the physical configuration. This process, therefore, may be labeled reverse or *inverse optics*. The interpretation of an image projected on any two-dimensional surface such as a sensory surface or the film in a camera is ambiguous. This is because the direction of a light ray is specified (given the anatomy of the eye and its refraction properties), but not the distance to its source or the distance to a surface reflecting it (although accommodation might be used to estimate the distances to nearby light sources). What image points belong to the same object are not specified from the information contained in images only. Luckily, a large number of physical laws regulate the world and impose constraints that reduce the ambiguity and ultimately might result in a uniquely specified interpretation or lead to the most probable interpretation of the image in impoverished viewing conditions. Regularities in the world result in regularities in images, which in turn may lead to corresponding regularities in the wiring of brain structures during development. This means that the required knowledge about physical properties of the world is hardwired into the neural tissue, requiring no explicit knowledge.

A number of physical constraints can be used to reduce the ambiguities in the solution to the inverse optics problem. One constraint is that a given point on a surface has a unique position in space at any one time; another is that matter is cohesive and separated into objects (Marr, 1982). The surfaces of objects are typically smooth in the sense that the distance variation that is due to roughness or cracks is small compared with the overall distance to the viewer. In other words, the distance to a surface shows small variation, except at object boundaries, resulting in smooth or small cyclopean depth variations within the outlines of physical objects in the visual field. By the same cause, the motion field arising from relative motion will be locally continuous within the object outline and typically differs from the motion of other objects in the image and the background. Similarly, it is likely that figure and ground surfaces differ in other

properties because of different surface orientations, light scattering properties, color, texture, etc. Therefore, the rims of objects in images may be indicated by discontinuities in the pictorial, motion, and cyclopean fields[2].

Ideas from information theory and computational image analysis highlight that the information in images is concentrated along contours, especially where the contour changes orientation most rapidly (Attneave, 1954). This is further demonstrated by the simple observation that lines alone are efficiently used to create recognizable and expressive drawings. Several computational models of shape perception emphasize the importance of points in images in which luminance contours change orientation (curvature extrema) or intersect (Barrow & Tennenbaum, 1981; Hoffman & Richards, 1984; Kanade, 1980) and psychophysical investigations support such models (Biederman, 1988). In addition, lines and edges presented for the eyes cause immense activation of cells in the retina (Kuffler, 1953) and in the early visual areas in the brain (Hubel & Wiesel, 1962).

Seeing edges by 'three kinds of vision'

Luminance edges are important cues to figure-ground segmentation and perception of the surrounding layout. In addition, edges in visual images may be created by contrasts in any other attribute. Nevertheless, edges created with pictorial stimuli (luminance, color, and texture) are abundantly used in studies on edge perception and figure-ground segmentation as compared with studies investigating how the visual system deals with motion or cyclopean edges. This may be because the questions posed often concern how segmentation is carried out in the first place, and the easiness whereby pictorial stimuli can be created and used as probes compared with motion and cyclopean defined stimuli. Consequently, the extent to which attributes are integrated is unknown. The abundant research using pictorial cues as probes is also reflected in the following sections where more emphasis is given to pictorial sections than to the motion and cyclopean sections.

Pictorial edges

It is common experience that spatial modulations of the pictorial attributes in images may result in figurative and interpretable pictures and it seems trivial that luminance contrast can be used to visualize objects on images. However, a luminance-defined contour in an image may be the result of a shadow, textured painting or the rim of an object. Many animals use this ambiguity to camouflage themselves with texture on their skin that mimic the surrounding texture. All the sources to luminance-edges are muddled up in

[2] The physical orientation in three-dimensional space is fully described by the orientation in the frontal plane (tilt) and the orientation in depth (slant). In addition to the perceived tilt, perceived slant can be conveyed by single attributes: for example, impressive slant percepts can be induced by cyclopean edges and surfaces. In this thesis the investigations are restricted to binding of edges in frontal planes.

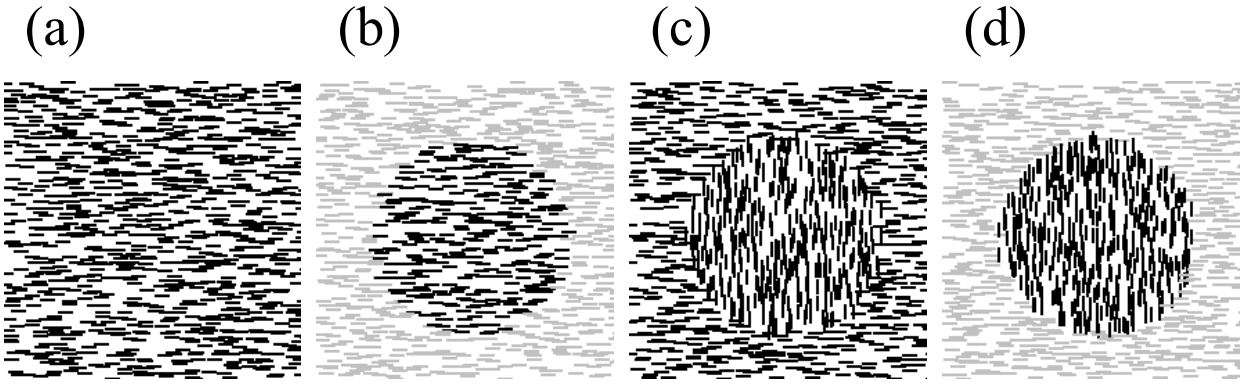

Figure 3. (a) A perfect camouflage is created by uniform texture. **(b)** The disc is revealed by luminance contrast. **(c)** The disc is revealed by texture orientation contrast. **(d)** Luminance and texture contrasts are superimposed.

natural images but the visual system can almost always make sense of the retinal input and separate these factors (Marr, 1982). Edges in other attributes may help the visual system to overcome difficulties in figure-ground segmentation created by shadows or surface coloring. Figure 3 illustrates figure-ground segmentation by luminance and texture. No shape can be seen in Figure 3a because the image is covered by uniform texture resulting in a perfect camouflage; luminance contrast reveals the disc shown in Figure 3b; texture contrast is used in Figure 3c; and luminance and texture contrast is used in conjunction in Figure 3d.

From neurophysiological investigations, it has been known for a long time that cells in the eye and primary visual cortex in the back of the brain (area V1) are activated when luminance edges stimulate the photoreceptors in the eye (Hubel & Wiesel, 1962). The *receptive field* of a neuron in the visual pathway, as originally defined by Hartline (1938), has become an essential concept for visual neuroscientists. It consists of all the photoreceptors in the eye that influence the activity of a specific neuron in the visual pathway.

The receptive fields of ganglion cells located in the retina in the eye have circular symmetric center-surround organized receptive field structures that integrate information across small portions of the image (Kuffler, 1953). About one million of these are distributed across each retina performing a massive parallel image analysis. Some of these cells are activated when receptors located at their receptive field center are stimulated with light and inhibited when the light strikes neighboring receptors; others show an opposite receptive field structure. Furthermore, the response may be sustained or transient. It has been shown that these cells encode the information in images to optimize the transmission between the eye and the lateral geniculate nucleus (LGN) given the limited capacity of the optic nerve (Dan et al., 1996).

Cells in primary visual cortex (area V1) in the brain then reformulate this information to selectively represent local primitives, including edge orientation, motion directions,

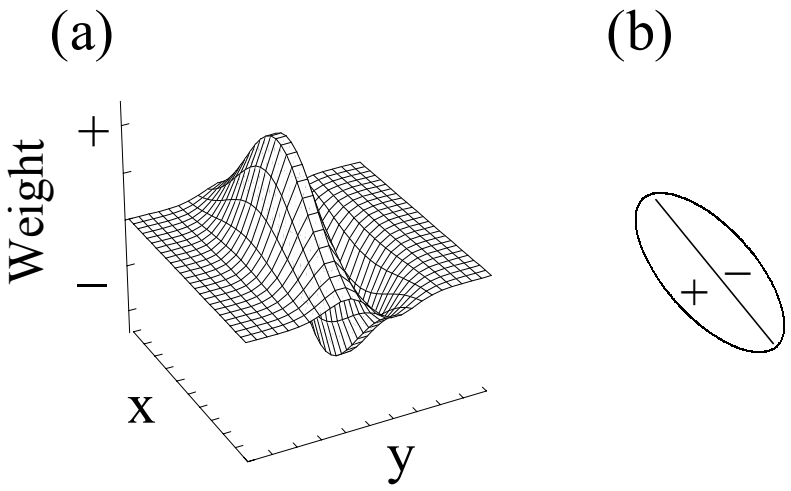

Figure 4. (a) Gabor model of a typical simple cell receptive field structure. Light stimulating spots of the receptive field in the x-y plane is weighted by the value of the Gabor function at the same spot before summation occurs across the whole field. A homogeneous stimulation will not produce any response because the positive and negatively weighted luminance values will cancel. **(b)** A simpler way of describing the simple cell receptive field structure.

and spatial frequencies of gratings[3].

The existence of different cell types has been proposed and is based on differences in their receptive field structures. The receptive fields of simple cells in V1 are relatively small, elongated, and consist of inhibitory and excitatory regions, or sub-fields oriented along the main axis. Such a receptive field structure is typically modeled mathematically by a Gabor wavelet, a two-dimensional Gaussian multiplied with a one-dimensional sinusoid (Figure 4). The response properties of simple cells, their orientation, and spatial frequency selectivity and phase selectivity (they respond when the edge contrast has the appropriate sign) are explained as a spatial summation across their receptive fields. Stimuli with brightness variations that match the receptive field layout will produce maximal response. No response will occur if the average brightness levels across the inhibitory and excitatory sub-fields are identical. Feedforward signals are probably used to accomplish orientation tuning of cortical simple cells (Ferster & Miller, 2000).

Complex cells are also orientation selective but, contrary to simple cells, they are phase invariant and may also be selective for line-ends. Simple and complex cells are retinotopically organized so that near/far locations on the image, as well as the retina in the eyes, are represented by near/far positioned cells on the cortical surface.

Nowadays it is widely recognized that cell responses may be modulated by stimulation of areas in the visual field that surround the receptive fields and by attention. This extended receptive field structure has been labeled non-classical receptive field (nCRF) as

[3] A grating is a number of separate parallel bars whose luminance may vary with sharp contrasts or sinusoidally in a direction perpendicular to the bars. The spatial frequency is the number of bars per degree of visual angle and is the inverse of the bar width.

opposed to the classical receptive field (CRF). These modulations are probably caused by lateral interactions between neurons located in the same visual area as well as feedback connections from other areas in the brain (for a review, see Gilbert, 1998).

Based on the early neurophysiological results, it has been claimed that the task of low-level vision is to analyze images in terms of spatial frequencies and orientations in the luminance domain (e.g., Campbell & Robson, 1968). In technical terms, such analysis is reminiscent of making a local piecewise Fourier analysis of the image. However, a Fourier-based visual processing scheme cannot explain how we can perceive features that are not directly created with luminance variations. Luminance and color are *first-order* image attributes and can be conveyed by a single-point source in the image. *Second-order* attributes require observations of more than one point and include texture, motion, and cyclopean fields. Technically, second-order edges (e.g., those forming the disc in Figure 3c) are supposed to be unnoticed by linear filters such as the Gabor filter in Figure 4 (Chubb & Sperling, 1988).

Although the division between first- and second-order image properties was originally used to describe motion signals (Cavanagh & Mather, 1989), they can be used to describe any image property, including edges carried by first- and second-order attributes, i.e. first- and second-order edges. Figure 5 shows typical stimuli frequently used to study second-order vision consisting of a fine grain, or high spatial frequency texture pattern (the carrier) whose contrast is modulated by a low spatial frequency pattern (the envelope). Further, a simple cell receptive field aligned with the second-order grating is outlined in the middle in the two panels of Figure 5. The cell will not respond because the mean luminance is equal on both its sides. The contrast of the carrier at a specific location is determined by the value of the envelope at the same location (similarly, radio-wave transmission may use amplitude modulation of a carrier signal). Because the mean luminance on the high contrast locations is the same as the mean luminance on the low contrast locations, there are no spatial frequency or orientation signals in the luminance domain corresponding to the frequency and orientation of the envelope. In other words, contrary to our visual abilities, a Fourier-based mechanism would not ‘see’ the envelope pattern!

It has been debated whether first- and second-order form is processed along a common pathway or if there are separate pathways. At present, there is considerable evidence, both neurophysiological and psychophysical, for separate pathways processing contrast-modulated texture and luminance-defined patterns (e.g., Zhou & Baker, 1993; McGraw et al., 1999; Wenderoth et al., 2001). Moreover, first- and second-order properties (local variations in contrast and texture) are uncorrelated in natural images, which may be the ecological basis for separate first- and second-order systems based on luminance contrast modulations (Schofield, 2000). Although the first-order pathway is blind to second-order gratings, the second-order pathway signals the orientation of both first-order luminance gratings and second-order contrast modulations (Zhou & Baker, 1993).

In addition to the contrast-modulated patterns, edges and gratings may be created in other second-order attributes, such as motion fields and cyclopean fields. Although

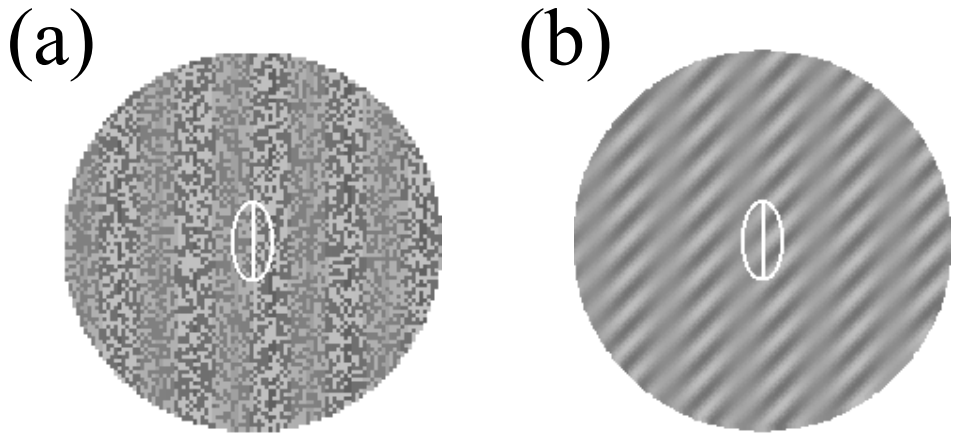

Figure 5. Vertically oriented second-order envelope gratings consisting of a contrast-modulated **(a)** random-dot pattern, and **(b)** a 45 deg tilted high frequency grating. First-order edge detectors or simple cells will not respond to the envelope patterns because the mean luminance is the same on both the inhibitory and excitatory regions of their receptive fields as outlined in the middle.

motion, stereoscopic depth and texture patterns are all second-order attributes, they require different strategies for detection. Detection of texture requires comparison and integration of first-order attributes (e.g., luminance) across several locations in the same image. Motion signal detection requires integration of first-order values across time, and cyclopean signals require integration of corresponding image points between the two eyes.

Temporal edges

The motion field is a second-order attribute that is efficiently used by the visual system to recover depth relationships and figure-ground segmentation in dynamic images. This is recognized by anyone who has observed that totally camouflaged animals become visible as soon as they move. The perceptual use of motion as an information bearing media has attracted much attention since demonstrations of grouping-from-motion phenomena (Wertheimer, 1923; Johansson, 1950), perception of egocentric motion (Gibson, 1950), and the kinetic depth effect whereby three-dimensional structure is seen from two-dimensional motion patterns (Wallach & O'Connell, 1953).

In demonstrations of shape-from-motion perception the information provided by motion has to be isolated from other cues to the shape. This is usually done by computer-generated motion sequences, where each frame consists of randomly distributed texture, i.e. a random-dot kinematogram. Such movie sequences typically simulate some object in relative motion to the observer. The shape is totally camouflaged by the texture when stationary, but can be seen when set in motion. Stated as an inverse optics problem: Given the two-dimensional motion pattern in the image, what is the environmental three-dimensional structure, and what is the relative motion in three-dimensional space causing the motion pattern on the image?

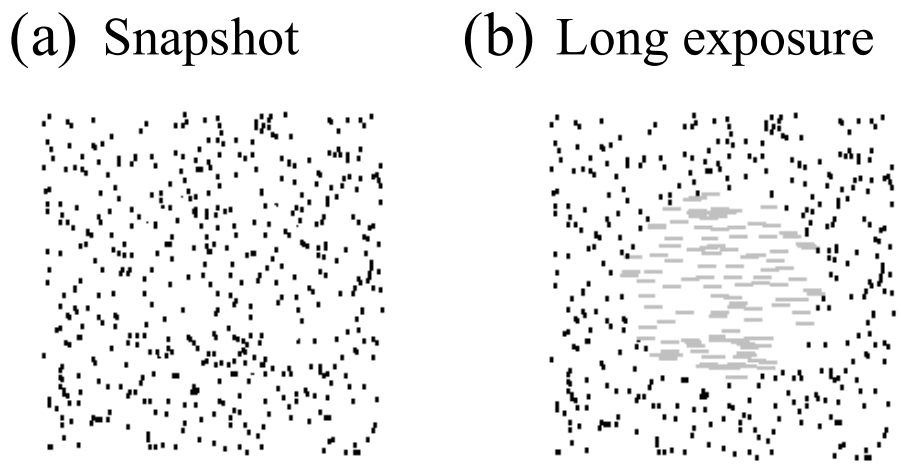

Figure 6. (a) One image taken from a random-dot kinematogram, corresponding to a snapshot of the motion sequence. The shape is efficiently camouflaged in the random dot pattern. **(b)** The streamlines are created from moving dots resulting from a long exposure time during the motion sequence. Contrast in the motion field is efficiently used to estimate figure-ground segmentation and reveal the disc.

The relative motion is used to estimate the depth-order, i.e. what image points belong to the figure and what image points belong to the ground (Gibson et al., 1969; Kaplan, 1969; Yonas et al., 1987). Figure 6a shows an example of a snapshot (one frame) of a random-dot kinematogram, and Figure 6b shows a long exposure during a motion sequence resulting in motion strikes. No disc can be seen from any single frame in such a motion sequence, but is immediately perceptually distinct when the motion sequence begins.

Sharp edges are seen even when the random-dot kinematograms are sparsely textured. This suggests that there are mechanisms in the visual system that signal sharp edges created by motion contrast. Consequently, a subset of orientation selective cells in cortical area V2 in monkeys are tuned to the same orientation for both luminance edges, texture edges and edges from motion (Leventhal, 1998; Marcar et al., 2000). Furthermore, cells in the optic tectum of birds (superior colliculus in mammals) (Frost et al., 1990) and mechanisms in insects (Kern et al., 1997) signal edges in random-dot motion kinematograms, and even signal whether the edge is occluding or disoccluding.

The neural architecture responsible for the sensing of motion-defined edges is unknown. Still, it has been suggested that feedback processes may be involved in that neural responses to motion-defined edges are delayed relative to luminance edges (Leventhal et al., 1998), and human perception of motion-defined edges depends on motion integration, which is thought to take place in areas beyond V2 (Sinha, 2001). The response delay to motion-defined edges may also be because time, by definition, is required for the detection of such edges. Thus, the existence of feed-forward mechanisms for the detection of edges created by motion cannot be excluded (Nakayama & Loomis, 1974).

There are attributes used by the visual system other than motion that require time for its definition. Figure-ground segmentation can be achieved when different regions flicker at different frequencies (Wong & Weisstein, 1984). It can also be achieved when local components of one region have phase locked temporal oscillations (e.g., motion or luminance variations) and components in other regions in the visual field oscillate in random relative phases (Alais, et al., 1998). The label temporal field is therefore more general than the label motion field when describing attributes that require time for their definition.

Cyclopean edges

The overlapping parts of the images in the two eyes differ slightly because of the different perspectives that are caused by the separation between the eyes (Figure 7). This can easily be seen if each eye is closed alternately, whereby near objects will appear to shift position in the images. However, when we look with both eyes, the two images melt together resulting in a single cyclopean image. For over a century, it has been known that the positional binocular *disparity* is useful for seeing depth relationships (Wheatstone, 1838). The disparity becomes smaller with increasing viewing distance, making stereovision useful at relatively short distances. Artificial computer generated stereoimages are widely used by researchers to probe the functioning of stereovision. One problem when creating artificial stereopairs is that the images must be presented separately to each eye and fused perceptually to one single image. Such a presentation may require some sort of viewing apparatus, a stereoscope[4].

Stated as an inverse optics problem, given the two images in the eyes, what is the cause of their differences? Figure 7a shows a simulation of the images projected on the retina of two eyes, separated by an interocular-distance. The eyes are focusing on the rim of a circular area located in front of a quadratic area textured with randomly distributed dots. Because the circular shape and the background are separated in depth, the position of the circle is shifted in opposite horizontal directions in each image (horizontal disparity), covering different portions of the background. Further, the perspective causes a vertical difference as seen in Figure 7b (vertical disparity). Because this difference is typically small, it is normally ignored when artificial stereopairs are created (Figure 7c).

In typical natural images the image features to be matched between the eyes consist of monocularly visible objects, or other unique features, offering a single solution for the matching (such as the disc in Figure 7c). Julesz (1960) was first to show that stereopsis

[4] It is possible to see three-dimensional images from stereograms without any stereoscopes, though it may require some practice. To ‘cross-fuse’ the stereograms presented here fixate a pen-tip halfway between the stereo-images and the nose so that the stereopairs appear as three separate images. Concentrate on the middle image where the apparent three-dimensional image appears. The stereograms can also be viewed by putting the focus behind the images (diverging the eyes or uncrossed fusion), again resulting in three perceived images where the stereoscopic depth may be seen in the middle one, this time with reversed depth relationships.

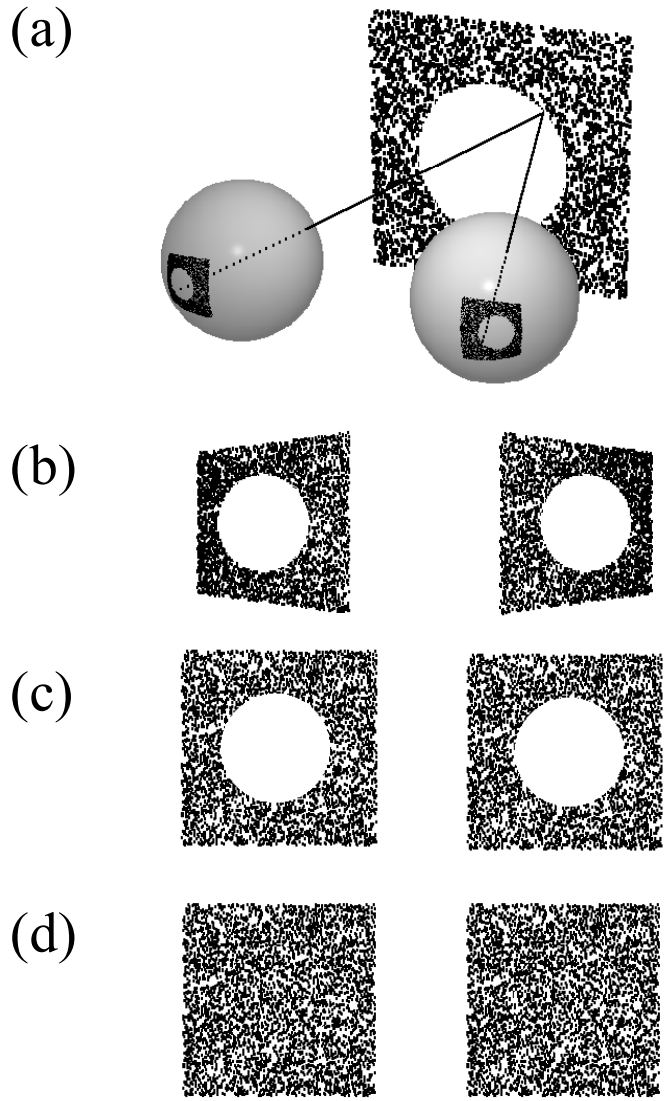

Figure 7. (a) Two 'naked' eyes focus on the rim of a circular shape hovering in front of a randomly textured square. The images projected in each eye differ because of the different perspectives that are caused by the separation between the eyes. **(b)** Because of the different perspectives, the depth separation between the circular white area and the textured square causes a shift of the background/circular shape in opposite positions between the eyes. In addition, the shapes in the two images differ. The large vertical deformations in these images arise because the viewing distance used in the simulations is small. **(c)** Such deformations are typically ignored in artificial stereoimages, leaving only the horizontal shift between the features belonging to the foreground and background. **(d)** When the circular shape is covered with the same texture as the background, it is impossible to see the circle from any image alone, but it can still be seen when the stereoimages are fused.

can still be used to recover depth relationships when the monocularly visible unique features are camouflaged by using the same texture, e.g., dots, for the figure as for the ground (Figure 7d). This discovery raised a problem: because no monocular forms are required for matching the right and left eye images, how does the visual system know which dot in the left eye is the right match for another dot in the right eye? Because the same texture elements are used across the images, there are many possible matches of dots between the eyes. Still, the visual system solves this problem, which has been at the focus of computational models of stereopsis (e.g., Marr, 1982). Random-dot stereograms are widely used as probes when trying to isolate stereovision processes in that no information about structure and figure-ground relationships is available without appropriate binocular fusion of the stereopairs. The luminance texture acts as a carrier of disparity signals that, in turn, can be used to carry signals about the depth relationships.

Image points at occluded surfaces near the rim of occluding surfaces are visible to only one eye and therefore lack positional disparity values (referred to as half-occluded areas). Therefore, in such events the binocular correspondence breaks down at such edges. Nevertheless, these image points are always perceived as belonging to the partially occluded 'far' surface, and it has been shown that the stereovision process actively uses these areas (Anderson & Nakayama, 1994). What is more, homogenous surface regions, such as the white region within the disc in Figure 7c, are perceived at the same depth as the edges of the disc, even though disparity values are missing.

As for edges created by motion-contrast, edges created by depth-contrast in random-dot stereograms are perceived as distinct even when the texture is sparse. This observation suggests the existence of mechanisms tuned to signal sharp edges though there are no such edges physically present. Anderson and Nakayama (1994) suggested the existence of receptive fields sensitive to the breakdown of binocular correspondence that occurred at occluding edges. Neurons selectively sensitive to the orientation of cyclopean edges, and the near-far relations, created by random-dot stereograms have been found in cortical area V2 in monkeys. These neurons are also selectively sensitive to the same orientation of luminance edges (von der Heydt et al., 2000).

Positional disparity of second-order features (e.g., texture edges) (Ramachandran et al., 1973) and temporally defined edges (Poom, 2002) can also be used to recover cyclopean depth, although the texture carrying the edge signals is totally uncorrelated between the eyes. Therefore the label cyclopean field is more appropriate than the label disparity-field since depth can be recovered although disparity values may be missing in the visual image.

Binding edges across attributes and space

The previous sections showed that information for visual segmentation and perception of form is carried by pictorial, temporal, and cyclopean image properties. In natural viewing conditions it is likely that figure-ground segmentation is signaled by several attributes in conjunction. For instance, figure and ground may differ in pictorial attributes such as reflectance properties, resulting in edges in the luminance field. In addition, figure and

ground surfaces are typically separated in depth, resulting in edges in the cyclopean field during binocular viewing. Furthermore, images in natural viewing conditions are rarely static and the relative motion between figure and ground results in edges in the motion-field in the corresponding images. Moreover, the visual system acts like a pattern-seeking machine binding image fragments across visual space. It tries to interpret images in impoverished viewing conditions and may add missing pieces of information to support some probable interpretation. Sometimes this strategy may lead us astray, but mostly it provides perceivers with accurate information. Isolated fragments of images may be treated differently by the visual system depending on other distant fragments in the image. This means that the perception of image fragments depends on the surrounding context. Both the fragments and the context may be conveyed by several attributes.

Early in the twentieth century the gestaltists formulated rules describing contextual modulations of visual stimuli, including *similarity, proximity, common faith* (motion), *good continuation,* and *closure* (Wertheimer, 1923). The last two rules are vital for the present purposes where perceptual binding of edges across space and attributes is examined. It is important to note that the perceptual grouping from common motion (common fate) and the grouping of aligned lines and edges created by motion contrast (good continuation) are distinct phenomena, although both are carried by motion signals.

The gestalt laws are merely descriptions of visual phenomena; they do not offer any deeper understanding of perceptual grouping in terms of possible processes. The concept of isomorphism, according to which there is a match between brain states and perception, was born when Köhler (1947) tried to provide a field theory of brain functioning based on concepts from physics. Although the originally proposed ideas were wrong, the idea of isomorphism lives on and gains support from knowledge gathered by psychophysical experiments, computer simulations of computational models, and tools now available in the study of the brain. This knowledge supports the commonly used two-stage models of perceptual grouping in which local stimulus features are detected at a first stage with grouping processes operating on these features at a second stage and also feedback to the first stage to modulate the processing of the local features (Figure 2). However, these models seldom address integration of attributes in the grouping process.

The recurrent excitatory connections in the visual areas V1/V2 found in primates primarily connect cells having collinear preferred orientations in the luminance domain (Gilbert, 1992; Malach et al., 1993; Das & Gilbert, 1995; Bosking et al., 1997). This probably explains why orientation selective cells are more responsive when stimulated with edges that are aligned with other edge elements outside their classical receptive field (Kapadia et al., 1995). Further, parallel neurophysiological and psychophysical experiments have shown that single cell responses are enhanced by flanking elements outside the receptive field of the cell, a stimulus arrangement that was also found to improve the observer's detection of the target line (Kapaida et al., 1995). It has been shown that cells with non-overlapping classical receptive fields but aligned orientation preferences synchronize their firing in the 40-60 Hz register when both are stimulated by a single moving bar, even when there is a gap in the bar somewhere between the cells receptive fields. No such synchrony is observed when the bar elements sweep over the

receptive fields independently. The horizontal connections between orientation selective neurons have been suggested to underlie the 40-60 Hz synchronized oscillations of neural responses and may be the neural underpinning of the gestalt laws such as good continuation (for a review of the neural synchrony hypothesis, see Singer & Gray, 1995).

An issue that has not received much attention is whether the gestalt laws are attribute-specific. Although information about the shape of objects is carried by several attributes simultaneously, we see only one version of any given object rather than several versions of it (Regan, 2000). This observation suggests that information from different attributes is integrated somehow or that only one channel at a time is allowed to operate. As suggested in Figure 2, grouping processes may operate on attribute-invariant feature detectors, perform grouping across attribute-specific feature detectors, or they might be completely attribute-specific.

Visual illusions provide important tools that can be used to investigate interactions across attributes. For instance, the tilt illusion can be used to selectively study early and late processes and is used in Study 1 to probe these issues. Investigations of perceptual grouping of edge fragments, e.g., by the rules of good continuation, offer an important probe to study cross-attribute binding processes. One typical well-known visual phenomenon based on the good-continuation rule is illusory contour formation, which is used in Study 2. Finally, a path detection paradigm is used in Study 3 to investigate binding across attributes and space.

Tilt illusions (Study 1)

Tilt illusions, i.e. misperceived orientation of vertical gratings, lines or bars surrounded by a tilted inducing grating offer important clues about how and on what level the visual system performs edge orientation analysis. The tilt illusion resembles the rod-and-frame illusion produced by replacing the inducing grating with a tilted square. A large tilted frame produces a sensation of self-tilting that is caused by visuo-vestibular interactions, which cause the physical vertical to be misperceived (Ebenholtz & Benzschawel, 1977). This effect can be avoided by presenting small displays, isolating visually based tilt illusions (Spinelli et al., 1995). Here, the tilt illusion induced by a small display is used as a probe to investigate edge orientation detection across attributes. The orientation of a test grating or line is misperceived when surrounded by an inducing grating having a different orientation (simultaneous tilt illusion, Gibson, 1937) or after prolonged adaptation to the inducing grating (successive tilt illusion, Gibson & Radner, 1937).

The sign of the illusion depends on the orientation difference between the test and inducing grating (O'Toole & Wenderoth, 1977). An S-shaped graph appears when the tilt illusion is plotted against the orientation of the inducer. A *repulsion* effect, whereby test gratings appear to be rotated away from the inducing grating, occur when the orientation difference is below 50 deg and the inducing grating is adjacent to the test grating. Figure 8a shows an example with a 15 deg inducer tilt and a vertically oriented test grating. An orientation *attraction* effect occurs when the orientation difference is between 50 deg and 90 deg. The orientation of the test grating seems to be oriented toward the inducing

grating. Importantly, the repulsion and attraction can be selectively influenced by various manipulations of the inducing grating or other contextual features. Examples of such manipulations are displayed in Figures 8b-d.

Common mechanisms are believed to underlie the simultaneous and successive tilt effect (Blakemore, Carpenter & Georgeson, 1970; Magnussen & Johnsen, 1986; Tolhurst & Thompson, 1975; Wenderoth, O'Connor, & Johnson, 1986; Wenderoth & Johnstone, 1988a,b), but there is evidence for that the attraction and repulsion effects have different origins (Wenderoth & Johnstone, 1987, 1988a,b). Certain manipulations of stimuli variables lead to a degradation of the repulsion but not the attraction, whereas other manipulations have the opposite effect.

The following dissociation studies, providing evidence for different origins of the tilt repulsion and attraction illusions, have been performed with first-order luminance defined gratings:

1. A visible frame surrounding the inducing grating selectively decreases the attraction (the frame effect, Figure 8b) (Köhler & Wallach, 1944; Wenderoth & Johnstone, 1988a).
2. The size of repulsion, but not attraction, decreases with spatial gap between the test grating and the inducing grating and selectively abolishes the repulsion effect with spatial separations of 1 deg (the gap effect, Figure 8c) (Virsu & Taskinen, 1975; Wenderoth & Johnstone, 1988a; Wenderoth, van der Zwan, &.Williams, 1993).
3. The same outcome is obtained when the spatial frequency or the size of the bars differs between the test and the inducing grating (the size effect, Figure 8d) (Georgeson, 1973; Ware & Mitchell, 1974; Wenderoth & Johnstone 1988a).
4. The tilt repulsion effect is selectively reduced when the radius of the outside diameter of the inducing area is decreased (Wenderoth & Johnstone, 1988a).
5. The repulsion reaches a peak at 25 ms and levels out to reach an asymptote at 100ms, whereas the attraction is much slower and levels out to an asymptote at exposures near 400ms (Wenderoth & Johnstone, 1988b).
6. Finally, when concurrent images are presented to the left and right eye, the perception switches apparently spontaneously between them. Binocular rivalry during adaptation to a grating tilted 10 deg from the vertical does not affect the repulsion tilt after-effect (Wade & Wenderoth, 1978) but when the orientation difference is 75 deg, the attraction tilt after-effect is reduced (van der Zwan & Wenderoth, 1994).

The double dissociations indicate that different processing strategies are activated mediating the tilt attraction and repulsion. Taken together, the tilt repulsion depends on orientation differences in the immediate neighborhood of the test pattern, whereas the attraction can be induced and modulated at large separations. This, together with the much faster time course of the repulsion compared with attraction, suggests that the repulsion is mediated by fast and local feedforward signals in the brain followed by subsequent large scale lateral and/or recurrent processing as a cause to the orientation attraction. The attraction reduction and preserved repulsion with rivalry during adaptation also suggest an earlier basis for repulsion than attraction because the binocular rivalry has

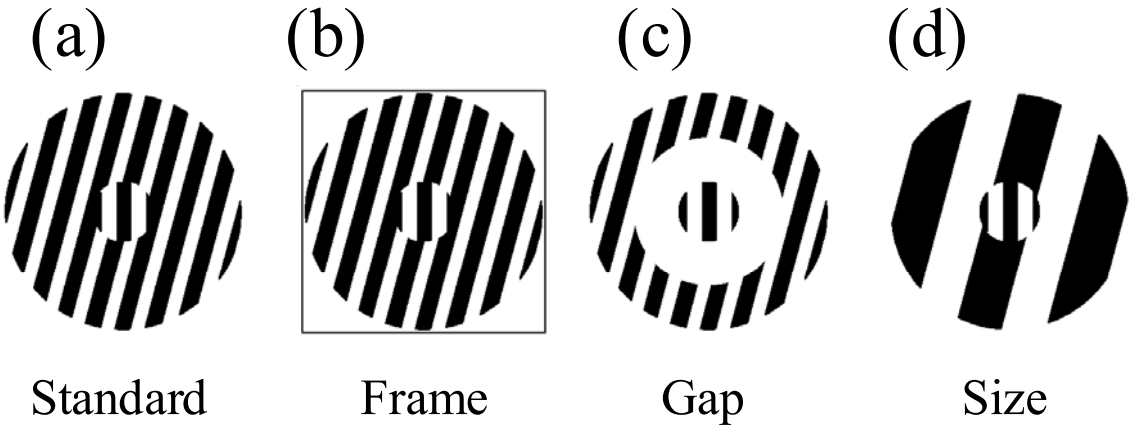

Figure 8. The stimuli used in the four experiments are shown here with gratings created by luminance contrasts. **(a)** The standard stimulus setting. **(b)** Attraction but not repulsion is selectively reduced when the inducing grating is surrounded by a frame of reference. **(c)** Repulsion but not attraction is selectively reduced when a gap separates the inducing grating and the test figure, or **(d)** when the spatial frequency (size) of the inducing and test grating differs.

to take place; after that, the information from each eye has interacted. Accordingly, line segments presented outside the classical receptive field of V1 cells in cats influence the orientation preference of these cells (Gilbert & Wiesel, 1990).

Probing binding with tilt illusions

Study 1 (Poom, 2000) was inspired by the results and conclusions reached in the earlier double dissociation experiments on the tilt illusions obtained with first-order luminance defined gratings. Can the tilt repulsion and the tilt attraction be obtained with gratings created by motion and cyclopean edges? If so, a possible double dissociation, similar to that obtained with luminance defined gratings, would provide evidence for low-level local orientation analysis for motion- and cyclopean-defined edges combined with other long-range spatial interactions. Furthermore, by using different attributes for the inducing grating and the test bar (inter-attribute conditions), it was possible to determine the location for possible interactions or binding between edge orientations of different attributes.

Four experiments were performed to investigate these issues. The idea was to measure possible frame- (Figure 8b), gap- (Figure 8c), and size-effects (Figure 8d) by comparing tilt illusions obtained in these conditions with the illusion obtained in a standard condition (Figure 8a) across the attributes. Edges in luminance, motion, and the cyclopean field were used to create the inducing gratings and a test bar. The same attribute was used to create the inducer and test bar in intra-attribute conditions while different attributes were used in inter-attribute conditions. Each observer was presented a randomized order of Experiments 1 to 4. Experiment 1 provided a standard condition (displayed in Figure 8a) used to compare the frame, gap, and size effects obtained in Experiments 2 to 4 (Figure

8b,c,d). The tilt illusion was measured with the method of adjustment by letting observers adjust a test grating to the apparent vertical position. The compensation for repulsion effects resulted in orientation settings toward the inducer orientation, whereas compensation for attraction resulted in settings away from the inducer orientation. The results were averaged across 25 participants.

The first experiment showed that both tilt attraction and repulsion effects occur irrespective of what attribute created the edges in the inducing grating and the probe. Figure 9 shows the characteristic S-shapes obtained by plotting the tilt illusion against the inducer tilt in intra- and inter-attribute conditions. Both effects were observed even when different attributes were used to create the inducing and test grating (inter-attribute conditions). In intra-attribute conditions the strength of the repulsion illusion was strongest for luminance gratings followed by the motion-defined grating. The cyclopean grating did not seem to induce any tilt repulsion effect for a 15 deg inducer tilt, which may have been obscured by the large variance. It was observed, however, that the cyclopean test bar and the inducing grating were difficult to separate when their orientation difference was small, resulting in a test setting away from the inducer. Nevertheless, repulsion effects were found when cyclopean gratings were used in inter-attribute conditions.

Note that in inter-attribute conditions the repulsion split in two categories of illusion strength: about 2 deg (filled symbols) and 0.5 deg (open symbols). From luminance-based studies, it has been shown that reducing the salience (visibility) of a test grating increases the repulsion effect while reducing the salience of the inducing grating decreases the repulsion effect (Berkeley et al., 1994). This dependence of the ratio between inducing grating and the test grating on the repulsion strength suggests an explanation of the two categories obtained in inter-attribute repulsion conditions. Large inter-attribute repulsion was obtained for inducing attributes that generated larger repulsion effects than the test attribute in the corresponding intra-attribute conditions (Lum/Mot, Lum/Dis, Mot/Dis). Small inter-attribute repulsion was obtained in reversed conditions (Mot/lum, Dis/lum, Dis/mot).

The results from Experiments 2 to 4 are shown in Figure 10. In Experiment 2 it was determined if the frame effect was preserved when other attributes than luminance was used to create the edges in the inducer grating and the probe. The frame was created by a thin luminous square used in all attribute conditions. The results showed that, irrespective of the attributes used to create the inducer and test gratings, the repulsion was spared but the attraction was reduced by the presence of the frame. The results obtained from Experiment 3 showed that the gap effect, whereby repulsion is selectively reduced when luminance gratings are used, occurs regardless of the attributes used to create the inducing and test gratings. In line with results obtained with luminance gratings, the results from Experiment 4 showed that the tilt repulsion illusion is sensitive to differences in spatial period between the inducing grating and the test grating irrespective of the attributes used to create the gratings. No repulsion illusion occurred in any attribute condition when the width of the stripes in the inducing grating was four times as wide as the width of the

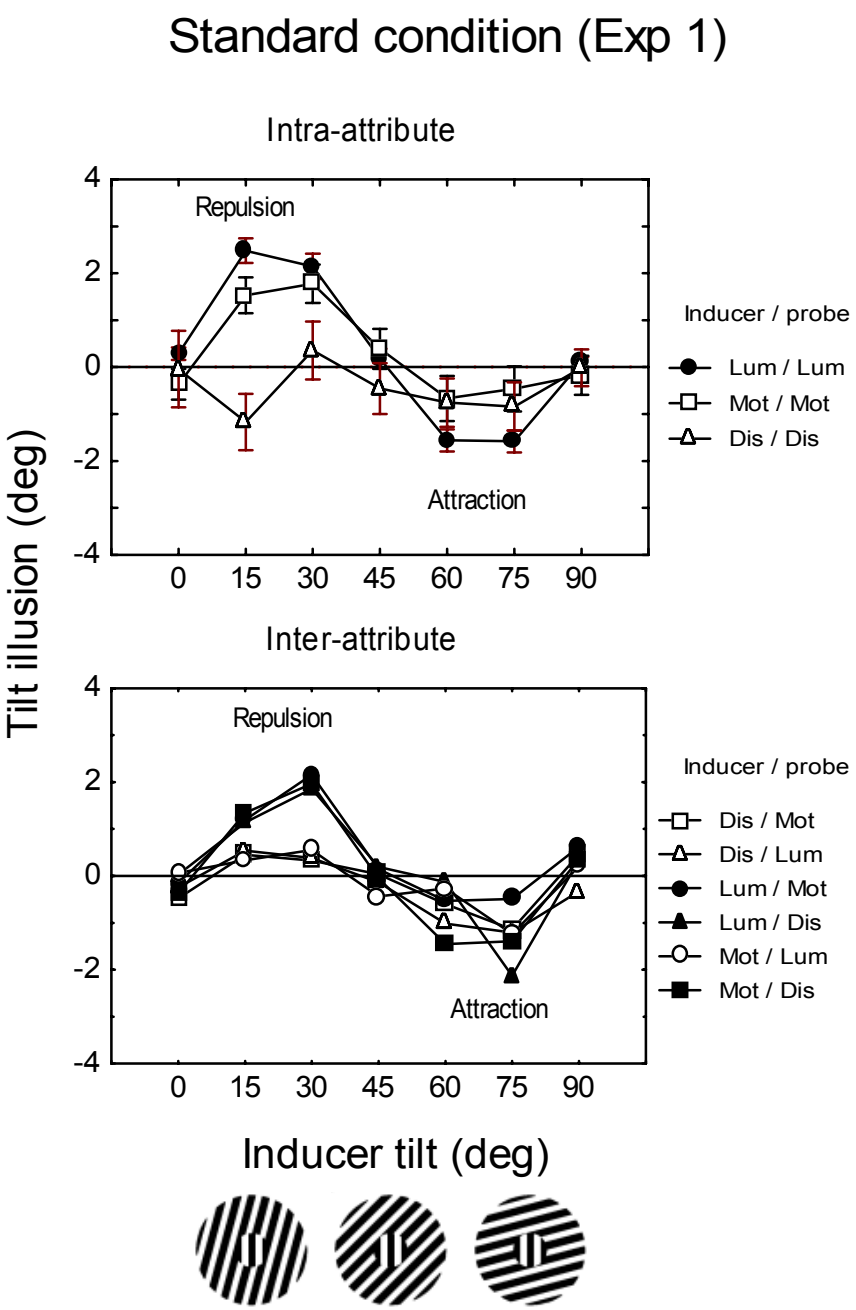

Figure 9. The direction of the misperceived tilt of a central vertical test grating depends on the orientation of the surrounding grating, resulting in S-shaped curves. Observers' settings of the test gratings are in the same direction as the inducer to compensate for the repulsion of orientation and in the opposite direction to compensate for the attraction effect. The insets below the graphs show 15, 45, and 75 deg inducer tilt using luminance-defined gratings. Top panel shows the results from intra-attribute conditions. Whiskers show +/- 1 SE. Bottom panel shows the results from inter-attribute conditions.

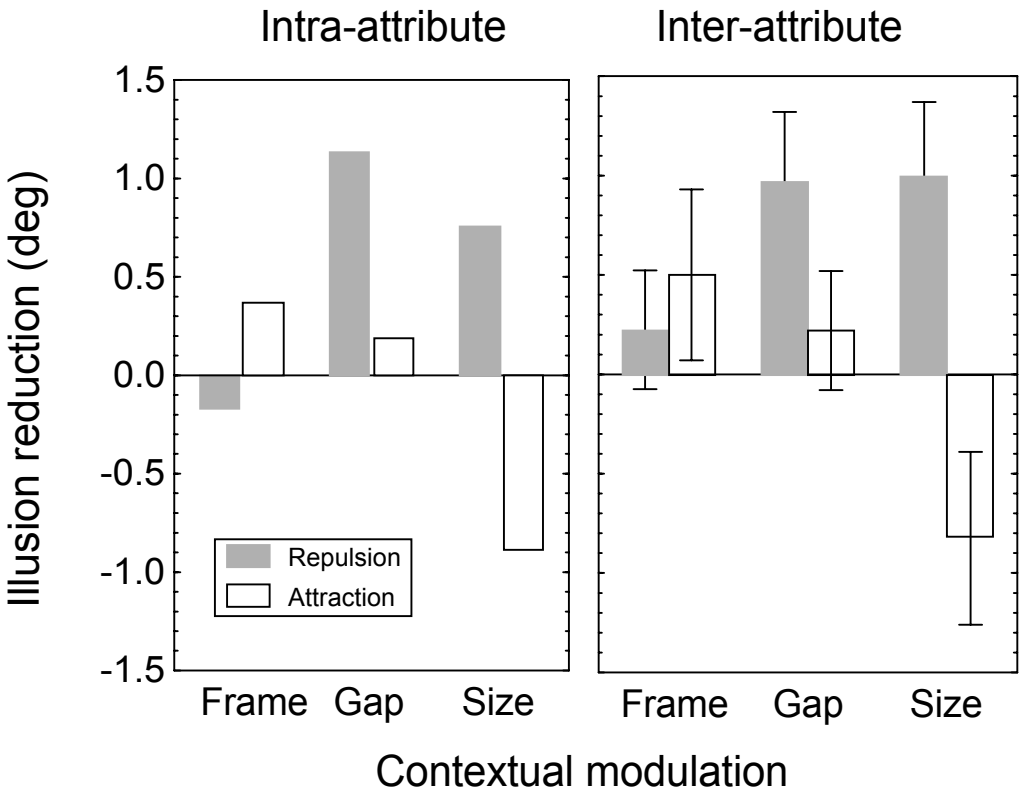

Figure 10. The results from the tilt illusion experiments pooled across 15 and 30 deg inducer tilt in repulsion conditions and 60 and 75 deg in attraction conditions and averaged across 25 observers. The three intra- and six inter-attribute conditions are pooled and shown in separate panels. The tilt attraction is preferably reduced when a frame surrounds the inducer. When a gap separates the inducing grating and the test bar or they differ in size, the repulsion is reduced. The attraction is increased when the size of the test bar and the inducer differ, as indicated by the negative reduction. The 95% confidence intervals are shown in the inter-attribute conditions.

probe. The attraction effect was even attenuated compared with when the same width is used for the inducing grating and the probe.

In addition to the evidence for early luminance-based processing of orientation, the attribute-invariant orientation-repulsion illusion provides evidence for early mechanisms detecting the orientation of edges and gratings created by motion and cyclopean contrasts. The inter-attribute repulsion illusion provides evidence for early interactions between the attributes in the analysis of edge orientations. No matter what attributes create the edges in the inducing and test gratings, the tilt repulsion illusion has a local character and the tilt attraction illusion has a global character.

It has been claimed that orientation attraction effects have to be mediated by processes in extrastriate regions in the visual brain where there are neurons whose response properties are affected by global stimulus properties (Wenderoth & Johnstone, 1987). However, such changes can be mediated by interconnected distant neurons in the same area as well as feedback from extrastriate areas. Even if the neural underpinning of the

orientation attraction and repulsion effects can be traced to the same cortical site, the dissociation experiments provide evidence for different processing strategies.

One study has examined tilt attraction and repulsion induced by second-order contrast modulations of high frequency carrier gratings (as shown in Figure 5b). Partial pooling occurred across orthogonal carrier orientations for the tilt repulsion and complete pooling occurred across orthogonal carrier orientations for the tilt attraction effect (Wenderoth et al., 2001). These results suggest that partial pooling across carrier orientations occurs at the level of V1 followed by additional pooling at higher levels. Nevertheless the partial low-level pooling of carrier orientations for second-order contrast-modulated gratings is in line with the results shown here. This provides evidence for pooling across different types of second-order gratings, created by cyclopean- and motion-contrast, and first-order luminance gratings.

Other previous studies have found interactions between neural edge-detecting mechanisms with cross adaptation techniques. It is known that adaptation to a random-dot stereo grating with no monocularly visible contours induces tilt aftereffects and that this affects transfer from gratings produced by cyclopean contrast to dot-density contrast (Tyler, 1975). Berkley, DeBruyn, and Orban (1994) found tilt aftereffects in a cross-adaptation procedure where illusory, motion, and luminance defined contours were combined in all pair combinations used during the adaptation period and the subsequent test. Only repulsion effects were investigated in that the orientation difference between the adaptation and test gratings was restricted to maximum 45 deg. Similar interactions have been found between color and luminance contours (Cavanagh, 1989).

Rivest et al. (1997) found that orientation discrimination of luminance, color, and motion-defined edges improved, regardless of whether the same or different attributes were used to create the edges in the learning and test phase. Because the improvement was local and unspecific to the attributes used, the authors suggested that the training changed the sensitivity of attribute-invariant orientation selective cells. Results from an edge localization paradigm further suggest that luminance, color, texture, and motion edges are integrated at a common site before the decision stage where the edge location is decided (Rivest & Cavanagh, 1995). However, Banton and Levi (1993) found evidence for independent processing between motion and luminance-defined edges in a Vernier or alignment acuity paradigm. The discrepancy between the results from different experimental paradigms might result from separate labels used by the brain for orientation, position, and position contrast (Regan, 2000). Whereas position contrast might be coded independently for different attributes, position and orientation might be pooled across attributes.

The results from Study 1 together with the cross-attribute tilt after-effects and the cross-attribute learning of orientation discrimination suggest the existence of mechanisms that sense and pool edge orientation from different attributes. This corresponds to early binding (Model 3 in Figure 2). Although long-range interactions in the form of orientation attraction was studied it, was beyond the scope of Study 1 to investigate long-range grouping of edges across space, the second stage in the models displayed in Figure 2. Illusory contours may be used to probe such processes.

Illusory contours (Study 2)

"It is evident that no theory of human perception would be acceptable if it could not handle an illusory contour configuration, nor would any artificial intelligence effort at real-world pattern recognition" (Meyer & Petry, 1987, p. 4).

Illusory contours have attracted considerable attention since they were first demonstrated (Schumann, 1900; Kanizsa, 1955). They are defined as impressions of contours where no physical contrast exists in the image and are strongly related to the concept of gestalt laws, specially the law of good continuation in which separate but aligned lines or contours are seen as linked together. Illusory contours are usually induced with spatially separate but aligned physically existing luminance contours, such as the traditionally used pacman figures resulting in the traditional Kanizsa square shown in Figure 11a (Kanizsa, 1955).

Illusory contours can be seen as long as the inducing edges are aligned and meet at obtuse angles (Figure 11b), but most people do not see illusory contours when this constraint is violated (Figure 11c). This has been labeled the relatability constraint in illusory contour formation (Kellman & Shippley, 1991) and mimics the gestalt law of good continuation. Illusory contours may also traverse and bend in depth when induced by elements slanted and aligned in depth in the cyclopean domain. This suggests that the relatability constraint should be generalized to three-dimensional space and provides evidence for a cortical site for illusory contour formation.

Illusory contours occur quite often in natural viewing conditions when illumination is low and pieces of the outlines of objects are missing. Mechanisms for generating illusory contours may have evolved to fill in such gaps in physically incomplete stimuli (Dresp, 1997).

Initially, it was claimed that perceived illusory contours was the result of a problem solving cognitive strategy employed by the visual system (Gregory, 1972). Since then, several studies have shown that real and illusory contours are involved in similar interactions thought to involve low level, mainly uncontrollable, processes. For example, both real and illusory contours are influenced by the orientation of nearby contours resulting in tilt illusions. The tilt repulsion effect is reduced when a spatial gap separates the inducer and test grating and a frame surrounding the inducer selectively reduces the tilt attraction effect. Both the gap-effect and frame-effects are obtained when illusory contours are used to create the inducer and probe edges (van der Zwan & Wenderoth, 1995). That insects respond to illusory contours further strengthens the evidence that cognitive contributions to illusory contour formation are nonessential (Horridge, 1992).

Subthreshold summation techniques provide additional evidence that automatic processing in humans mediates illusory contour perception. Such techniques were originally used to measure summation between real contours (Kulikowski & King-Smith, 1977). The contrast detection threshold at which a thin near threshold target is detected is

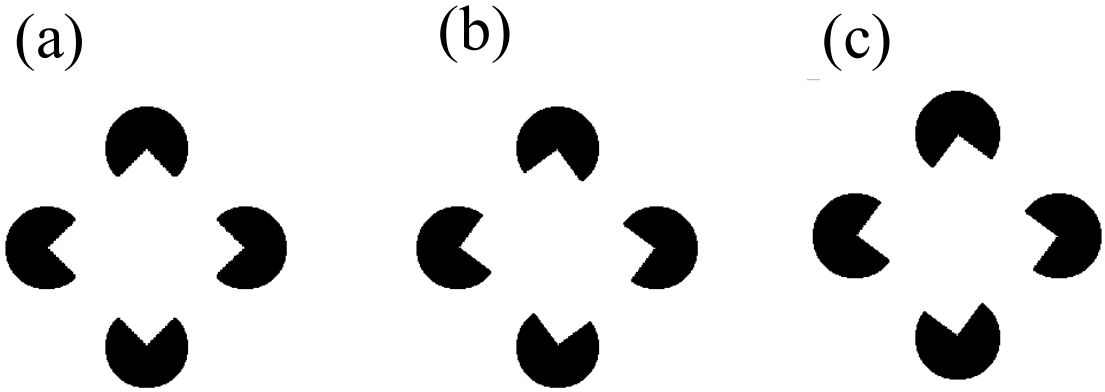

Figure 11. (a) The Kanizsa illusory contour. A square is typically seen occluding the black discs as long as the gaps in the discs are aligned along smooth contours as in **(a)** and **(b)** but not in **(c),** where it is difficult to see an occluding figure although it is still possible.

lowered when the target is superimposed and aligned with a subthreshold line. Similarly, the detection threshold at which a near threshold line is detected is reduced when the line is presented superimposed on an illusory contour (Dresp & Bonnet, 1995; Dresp & Grossberg, 1997; 1999). The interpretation of such spatial facilitation is that the target line and the illusory contour activate common neural elements. When the line is superimposed on an illusory contour, neural elements sum the signals arising from the line and the illusory contour, making the summed activity increase beyond the activity provoked by the sub-threshold line alone, which pushes the signal above threshold.

Neurophysiological studies show that luminance induced illusory contours and real luminance edges activate common neurons in cortical area V2 in monkeys (von der Heydt, Peterhans, & Baumgartner, 1984; Peterhans & von der Heydt, 1989), as do disparity induced illusory contours (Bakin et al., 2000). Others have shown orientation selective responses in cortical area V1 from illusory contours cued by abutting gratings in monkeys (Grosof et al., 1993) and cats (Sheth et al., 1996). Lee and Nguyen (2001) trained monkeys to attend to the location of Kanizsa-like illusory contours while activity was recorded in cortical area V1. They found responses to illusory contours in V1 that were delayed relative to the responses evoked from real contours and also relative to the response to illusory contours in area V2. Response to illusory contours in V2 preceded that of V1 with about 100 ms, suggesting that feedback signals from higher levels in the brain may participate in the illusory contour formation. Accordingly, lesions in the inferior temporal (IT) cortex impair the ability for monkeys to see illusory contours (Huxlin et al., 2000).

In addition to mammals, birds and even insects respond to illusory contours as if they were real, and in all animals tested illusory contours activate neurons in early visual brain areas (for a review on illusory contours in animals, see Nieder, 2002). Furthermore, the functional importance of illusory contour formation is strengthened because mechanisms signaling such contours exist in independently evolved visual systems.

Illusory contours can be perceived between spatially separate but aligned edges created by motion contrast (Kellman & Cohen, 1984; Prazdny, 1986; Poom, 2001a) and spatially separate but aligned edges created by random texture stereograms resulting in a purely cyclopean illusory contour formation (Julesz & Frisby, 1975; Mustillo & Fox, 1986; Poom, 2001a). The stereogram in Figure 12a shows illusory contours induced by luminance contrast (this can be seen without any fusion of the stereogram). The inducing elements in the stereogram in Figure 12a are also displayed at a stereoscopic depth separate from the background, so that the luminance and cyclopean contrasts of the inducing elements are superimposed. When the stereopictures are fused the illusory contours become more distinct and the square is seen as hovering in front of the background. Fusion can be accomplished by using crossed fusion of the leftmost stereopair or uncrossed fusion of the rightmost stereopair (see footnote 4). The hovering square and the illusory contours can still be seen when the stereoscopic contrast is isolated as seen by fusing the stereopictures in Figure 12b. Poom (2001a) showed that illusory contours could be observed between two aligned inducing edges defined by different attributes as demonstrated by the stereopictures in Figure 12c.

Even though the same three-dimensional illusory form can be induced by different attributes (Carman & Welch, 1992), no conclusion can be reached about whether a common grouping mechanism exists or if there are separate similar attribute-specific mechanisms. That illusory contours are seen between inducing edges of different attributes provided evidence for a single attribute-invariant gestalt forming mechanism (Poom, 2001a).

Nevertheless, two mechanisms of spatial facilitation with collinear edges have been found using the near threshold summation technique, one that operates in the luminance domain and another that operates in the domain of color contrast (Dresp & Grossberg, 1999). Is this kind of near threshold summation attribute-specific across other attributes as well?

Probing binding with illusory contours

The purpose of Study 2 (Poom, 2001b) was to determine whether detection of a near threshold thin line was facilitated when the line was superimposed on illusory contours induced by different attributes. Such facilitation has previously been reported for luminance-induced illusory contours and would provide evidence for activation of common neural elements by illusory contours irrespective of the attribute used to induce the illusory contour.

Kinetically defined inducing edges were created by relative motion in random-dot kinematograms. The appearance was of a stationary surface that partly occluded an oscillating background. Random-dot stereograms were created that displayed vertically or horizontally aligned cyclopean defined inducing elements. The appearance of the binocularly fused stereoimages was a horizontally or a vertically oriented band with illusory contours completing the gap between the inducing elements. The vertically oriented target line was presented in the same depth plane as the edges of the occluding

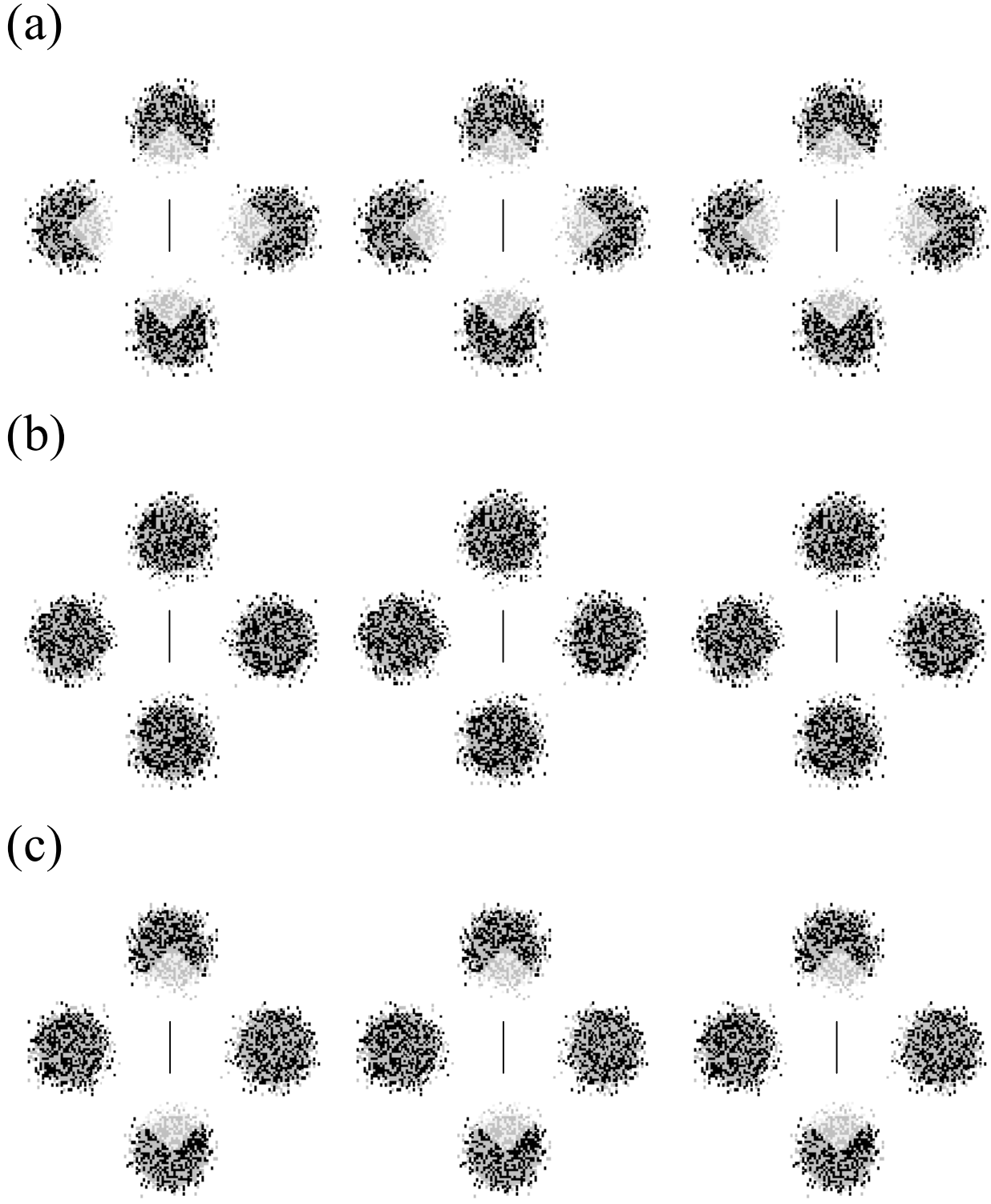

Figure 12. (a) Illusory contours cued by luminance contrasts. Stereoscopic depth contrast superimposed on the luminance contrast strengthens the appearance of the illusory contours (may be accomplished by cross fusing the leftmost stereopairs or by uncrossed fusion of the rightmost stereopairs). **(b)** Illusory contours can be cued by stereoscopic depth contrast alone. **(c)** Inter-attribute illusory contours are cued between inducing elements created by luminance and stereoscopic depth contrasts. The vertical lines may aid fusion of the stereopairs.

surface. Perceptual summation between the near threshold luminance defined target line and the illusory fractions of the contours would be expected if these contours and the line have a common neural representation. A vertical thin near threshold luminous target line was presented on the gray background between either the left or right pairs of inducing elements and the edges of the inducers were either vertically (Figure 13a) or horizontally aligned (Figure 13b). The test line was superimposed on the illusory contour only when the inducers were vertically aligned.

A two-alternative forced-choice technique was used with fixed stimuli, where the luminance contrast between the test line and the background was set to various fixed values. The observers' task was to indicate on which side the test line appeared on each trial. Four inducing elements were positioned at the corners of a square with 2.8 deg side length. The radius of the inducers was 0.85 deg and the gap between pairs of inducers was 1.1 deg of visual angle.

Illusory contours produced by luminance-defined inducers were used in Experiment 1 together with inducing edges created by offset concentric gratings where the inducing contour itself was illusory (Figure 13a and b). It has been suggested that illusory contours created by such inducers and those created with luminance-defined inducers may be mediated by different mechanisms (e.g., Halpern, 1981). Motion-defined inducers in Experiment 2 and cyclopean inducers in Experiment 3 produced the illusory contours. Threshold was defined as 75% correct responses and was calculated as the intersections between the Weibull psychometric function fitted to the data and the threshold level.

The thresholds were lower when the vertical target line was superimposed on the illusory fraction of the contour than if the inducing elements, and hence the illusory contour, were horizontally aligned. This result was obtained for three observers, regardless of whether luminance, grating, motion, or cyclopean edges created the inducing elements (Figure 14, Exp 1 – 3).

To what degree did the positional cues, offered by the inducing edges, facilitate the line detection? It might be that the effects in the previous experiments were due to the positional cues and hence no summation between the illusory contour and the target line occurred. In Experiment 4 the illusory contour inducing elements were replaced by positional cues that did not elicit any illusory contour perception. These cues were composed of arrows pointing to the location of the ends of the target line (Figure 13c) and lines enclosed in outlined circles (not shown here). The lines enclosed by the circle were either vertically or horizontally aligned; no illusory contour was seen between these elements. Possible threshold reduction was calculated as the difference in luminance levels for the target line required to reach 75% correct responses in these two conditions. The results showed that it is unlikely that the positional cues offered by these edges facilitated the detection of the target line in that other positional cues that did not elicit any illusory contour perception were unable to reduce the threshold (Figure 14, Exp 4).

If the target line and one of the inducing edges are both located within the sampling area of single neurons, receptive fields activate the same orientation selective cells. This concordant activation might account for the summation. In experiment 5 the target line

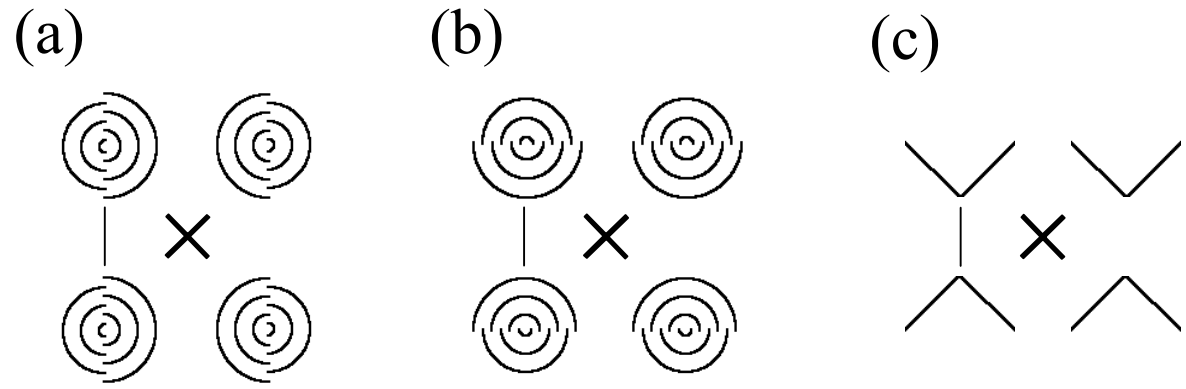

Figure 13. The experimental conditions shown with illusory contours cued by offset gratings, with the test line displayed to the left of the fixation crosses. The test line was presented with near threshold contrast in the actual presentations where the task was to judge whether the test line appeared to the left or to the right of the cross. **(a)** The line is superimposed on the illusory contour. **(b)** The same situation as in **(a)** but with horizontally oriented illusory contours. **(c)** Positional cues offered by the arrowheads.

was shortened and the distance between the inducing elements was increased, introducing a gap of 1 deg between the target line and the inducing elements. Otherwise, the same stimulus settings as used in experiments 1 to 4 were used again. Summation between the illusory contour and near threshold line was found though there was a gap between the target line and the inducing edges. No facilitation was found with the positional cues (Figure 14, Exp 5; results are averaged across five observers)

Taken together, the Dresp and Bonnet (1995) results were replicated using the luminance induced illusory contours. In addition, illusory contours cued by the offset gratings, motion contrast, and cyclopean contrast summate with the near threshold line. An explanation of these results, based on previous models of near threshold summation, is that common neural elements are activated by the illusory contour and the near threshold target line. If so, these results provide evidence that common neural elements summate luminous lines and illusory contours irrespective of the attributes used to induce the illusory contour. Taken together with the results from Study 1 on cross-attribute tilt illusions, these results suggest that edges and their orientations from different attributes activate common edge detectors in visual area V1 and/or V2. Furthermore, it is likely that the edge linking process resulting in illusory contours operates on these attribute-invariant edge detectors (Model 3 in Figure 2).

The final experiment provides good evidence against simple summation within single filters, with classical receptive fields covering both the inducing edges and the test line. The spatial gap of 1 deg between the test line and inducing edges is wide enough to severely reduce summation within the small single classical receptive fields in cortical area V1/V2. However, it cannot be ruled out that such process did occur within larger classical receptive field structures elsewhere. One way to provide additional evidence for a linking process that connects the activity of aligned sub-units is to use curved contours.

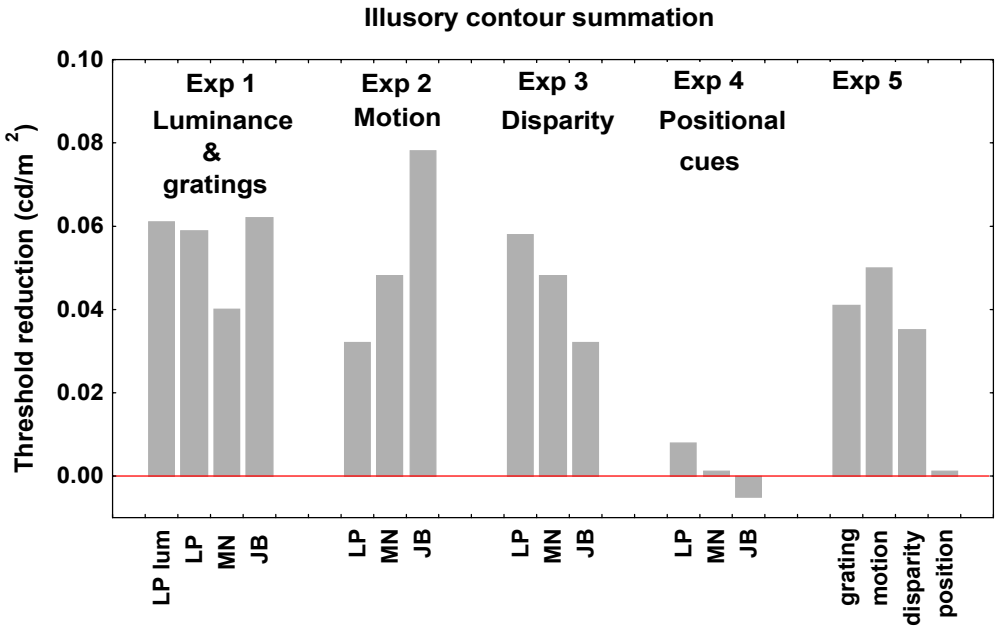

Figure 14. The results from the illusory contour summation experiments. The threshold reduction was calculated as the difference between performance when targets were and were not superimposed on the illusory contours. The threshold was estimated as the luminance required to attain 75% correct detections. The results are shown from Experiments 1 to 4 from three observers (LP, MN, and JB) and pooled across six observers in Experiment 5. Illusory contours cued by luminance (LP only) and grating edges were used in Experiment 1, motion contrast in Experiment2, and cyclopean edges in Experiment 3. Positional cues were used in Experiment 4. A gap separated the inducing elements from the target line in Experiment 5. The detection thresholds are reduced only when the target line is superimposed on the illusory contour.

The detection of randomly curved contours by single filters mapped to each possible contour would lead to a combinatorial explosion of required filters (Hess & Field, 1999). Representations of curved contours by distributed cell activity (dynamic binding) would reduce the amount of required cells.

A visual phenomenon related to the gestalt law of good continuation and illusory contour formation that can be used for this purpose is the perceptual 'pop out' of aligned path elements. The gestalt law of good continuation guides the perception of aligned path elements. Such elements are seen as belonging together forming 'snake like' paths that are distinct from the surroundings.

Seeing 'snakes' in noise (Study 3)

Psychophysical studies demonstrate that human sensitivity to an edge element is higher when it is aligned with flanking edge elements (Polat & Sagi, 1993; Kapadia et al., 1995). This is a phenomenon that is reflected by the gestalt law of good continuation, where separate aligned elements forming snake like patterns are seen as linked together and more distinct than a path composed of elements orthogonally oriented to the path forming

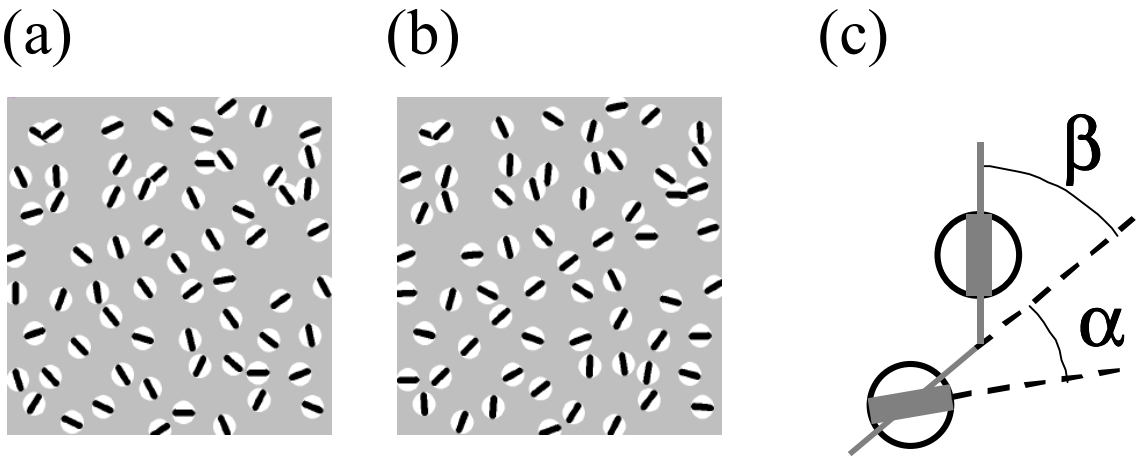

Figure 15. The same path embedded in noise is shown in **(a)** with path elements orthogonal to the path (ladder) and **(b)** aligned with the path (snake). Path detection is typically easier when the elements are aligned. **(c)** Two parameters are shown that guide visual path formation: α is the orientation of the elements relative to the path and β is the path angle or path curvature. β=10 deg and α=90 deg in **(a)**; β=10 deg and α=0 deg in **(b)**.

a 'ladder-like' path (Figure 15a, b). The path curvature (β in Figure 15c) is an important parameter limiting the distinctness of the path as demonstrated in Figure 16.

Moreover, aligned line elements that form a closed path (a snake biting its tail) are easier to detect when embedded in noise elements than an open path (the closure effect, Kovacs & Julesz, 1993). The gestalt law of closure, in addition to good continuation, guides the detection of closed paths.

Contextual modulation of local features seems to be a general strategy employed by the visual system and affects both perceived orientation and detectability. When it comes to figure-ground segmentation, the visual system has to be sensitive enough to detect contours embedded in image contrast noise. At the same time, the system has to be selective for the relevant edge segments that belong to object contours and stable so that the noise is not enhanced. In neural network jargon this is recognized as the sensitivity/stability dilemma (Li, 1998).

Many properties of contextual modulation of the classical receptive field properties in visual area V1 correspond to well-known properties of perceptual organization and visual Gestalt laws. A target line imbedded in a field of other lines with the same orientation loses its perceptual salience and 'blends in' with its background compared with when it is presented alone in an empty field. The response to a line segment presented within the classical receptive field of a V1 cell is similarly modulated by flanking line segments. The response is typically greater in the 'pop out' condition than in the 'blend in' condition. When the orientations of the lines in the field differ from that of the target, the single target line seems to 'pop out' from the background elements. On the contrary, when the target line is embedded in randomly oriented elements, it seems to blend into the back-

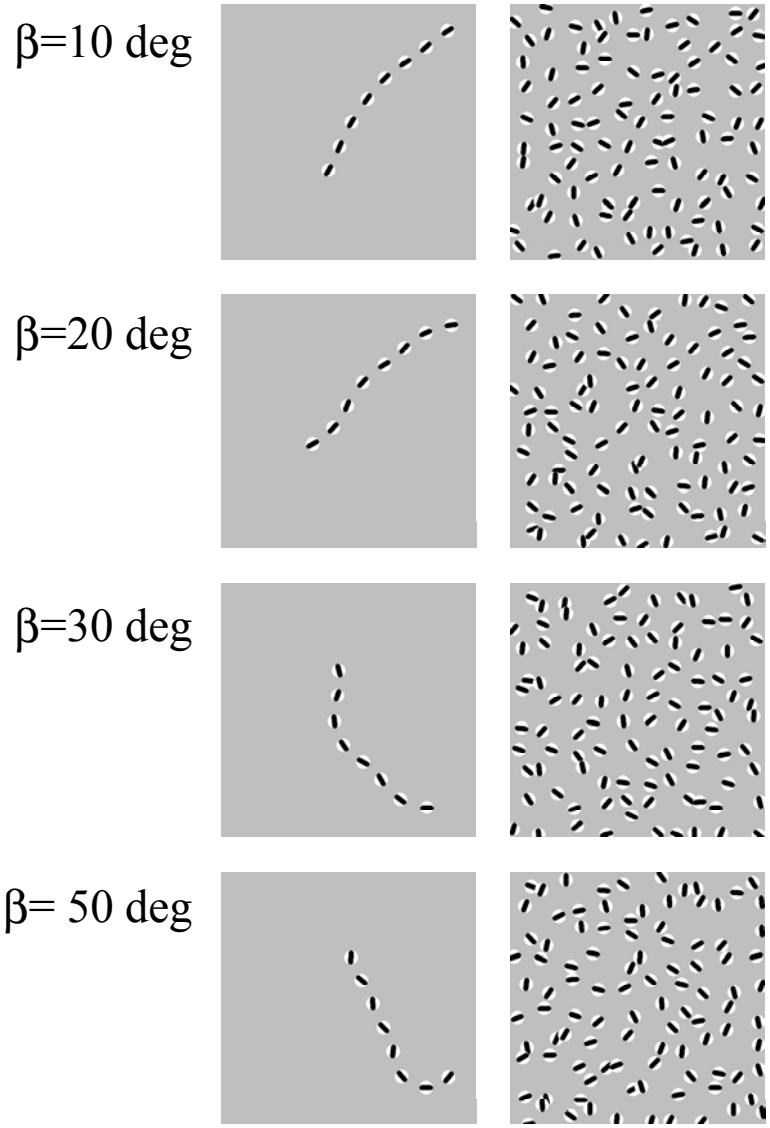

Figure 16. Snake paths composed of oriented micropatterns created by luminance contrast. The left panels show the 'naked' paths and the right panels show the same paths embedded in randomly oriented noise elements. The path angles are displayed. The paths embedded in noise can be detected for path angles of up to about 30 deg but detection is very difficult when path angle is 50 deg.

-ground elements; however, when several line elements are aligned with the target, the chain of elements 'pop out' from the background.

The perceptual binding or linking of edge segments depends on their relative distance, angle, and alignment, encompassing the Gestalt rule of good continuation (Field, Hayes, & Hess, 1993). Psychophysical investigations have also pinned down similar rules guiding the formation of illusory contour perception (Kellman & Shippley, 1991). Similarly, edge co-occurrences in natural images follow the same rules (Geisler, et al.,

2001). In cortical area V1 there are cells that reflect these properties of perceptual salience. The response to a line segment presented within the receptive field of a V1 neuron is enhanced when other line segments are aligned with the target, although presented outside the classical receptive field (Kapadia et al., 1995).

Probing binding with 'snakes'

The purpose of Study 3 (Poom, 2002) was to further investigate binding of different attributes across space by using a path detection paradigm. Randomly curved paths that exclude single filter detection were used in Experiments 1 and 3. Experiment 2 investigated whether the closure effect is attribute-specific. Experiment 3 investigated the path analogue to inter-attribute illusory contour formation (Poom, 2001a) demonstrated in Figure 12c.

Some texture is required to carry the motion and cyclopean signals. Therefore, oriented micropatterns composed of dots randomly distributed within a circular window were used to create both the path elements and the noise elements (Figure 17). Bars within the windows were defined by luminance contrast (bright dots created the bar against the background, filled with black dots), motion contrast (the bar dots were stationary and the background dots were oscillating back and forth horizontally with a small amplitude), or stereoscopic depth contrast (the bar dots were stereoscopically separated in depth from and located in front of the background dots).

The locations of the target micropatterns belonging to the snake pattern was first calculated by the computer according to predetermined path parameters, then the locations of the noise elements were quasi-randomly selected by the computer to minimize overlap with other elements. The locations of successive path elements were at the mid-section of line segments that were joined as a chain. The length of each line segment was equal to the mean distance of the noise elements, with a small pertuberation added. One endpoint of first line segment was confined to be 3 deg of visual angle from the display center and the other endpoint was located in a direction toward the center of the display. The orientation of the following line segment was randomly assigned a positive or negative sign of the path angle (β in Figure 15c). The orientations of the noise micropatterns were randomly distributed. Path detection across different path angles was measured in Experiments 1 and 3. The orientation relative to the path of individual path elements was randomly varied within a restricted range (α-range) when closed paths (as displayed in Figure 18) were used in Experiment 2.

The results in the multi-attribute conditions were compared with probability summation of hypothetical independent attribute-specific channels. The following description of probability summation is adapted from Tyler and Chen (2000). Consider two human observers, LP and EB. Their task is to detect snakes on a lawn. The two observers may be considered as independent detectors with respect to their visual performance in detecting snakes. Sometimes their judgments are uncertain because their visual resolution has a limit and the snakes are camouflaged. However, when LP's or EB's confidence of seeing a snake reaches a specific criterion, either LP, EB, or both will

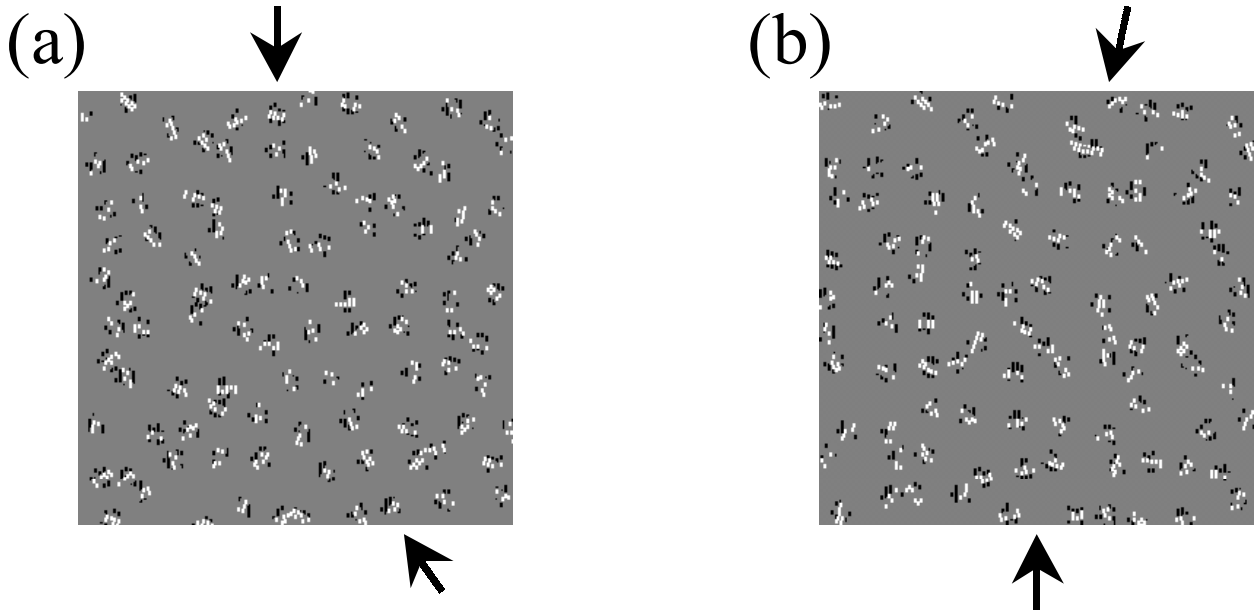

Figure 17. Examples of the luminance-defined stimuli used in the first experiment. Some of the micropatterns are arranged along snake-like paths whose endpoints are marked with the arrows. **(a)** Micropatterns aligned with the path (snake) and **(b)** orthogonal to the path (ladder).

report that a snake is spotted. The receiver of the messages can use the two observers' reports to improve the detection performance, compared with if only one of the observers was available. The probability of detecting a snake is higher when any one of the observer's spots a snake than if one has to rely on either LP or EB alone. Thus, in probability summation one considers the response from any single member of a group of independent observers or detectors.

If the probability that observer LP reports a snake is P_{LP}, then the probability that no snake is reported by the same observer is $(1-P_{LP})$, and the probability that neither LP nor EB report any snake, assuming that they are independent, is: $(1-P_{LP})(1-P_{EB})$. The probability P that at least one of the observers LP and EB report a snake is complementary to the event that no one reports a snake, that is: $P = 1-(1-P_{LP})(1-P_{EB})$. This means that after N snake finding sessions on average $P{\cdot}N$ snakes will be reported. Similarly, the output from independent path detecting channels can be used to improve performance compared with when a single channel is used. The same reasoning can be applied to any number N of independent observers or channels. The corresponding probability $P_{1,2,\ldots,N}$ that at least one of these channels signals target present is:

$$P_{1,2,\ldots,N} = 1-(1-P_1)(1-P_2)\ldots(1-P_N) \qquad (1)$$

The probabilities in the example above have a range between $P=0$ (no snakes detected) and $P=1$ (100% of the snakes detected). When the task is to decide which of two lawns, or presentations, contains a snake, and the snake is always present in one of the two lawns, then the range is between $P=0.5$ (50% correct when guessing) and $P=1$. In this case the P values in equation 2 have to be changed by substituting P with $2P'-1$ in equation (1). The resulting equation is then:

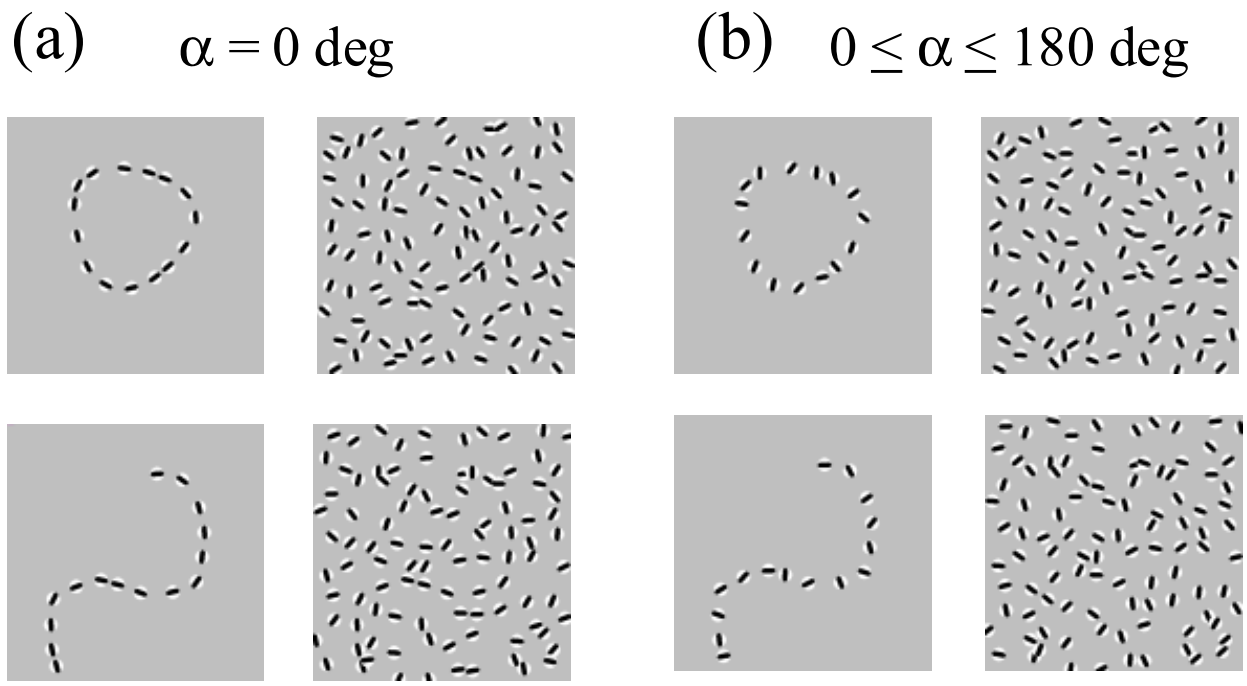

Figure 18. Illustration of the closure effect with luminance-defined elements. The same paths are presented alone in the left panels and embedded in noise in the right panels. The path elements are aligned along the paths in **(a)** and randomly oriented, i.e. α-range = 180 deg in **(b)**.

$$P'_{1,2,\ldots,N} = 1 - \frac{2^N}{2}(1-P'_1)(1-P'_2)\ldots(1-P'_N) \qquad (2)$$

From the experiments, the results from conjunction conditions were compared with predictions made from probability summation of hypothetical independent channels, as calculated from the results obtained in the corresponding single attribute conditions.

The first experiment addressed two questions: Can edge linking be accomplished with edge elements created by other attributes than luminance and color, and is path detection facilitated if edges of several attributes are superimposed on the same path elements?

In Experiment 1 the attributes used to create the path elements served as the independent variable while path angle (β) was varied in order to obtain an estimate of the threshold. The path is a straight line when β=0 deg. The path elements were oriented along the path, i.e. the element orientation (α) was zero (snakes). In another condition with β=20 deg for AM and 30 deg for LP the path elements were orthogonally oriented to the path (α=90 deg, ladders). The percent correct path detection was measured with a 2-interval forced choice procedure (2IFC). The results presented in Figures 19 are the estimated 75% correct thresholds based on the best fit of a 3-parameter (slope, asymptote and location) Weibull psychometric function to the data. Figure 19a shows the estimated path curvature thresholds of paths composed of eight elements.

The results showed that paths could be detected irrespective of the attribute used to create the edge elements, and performance levels were similar across the attributes

(Figure 19a). These findings may indicate that the particular contrast levels used for the different attributes were accidentally similar or that the particular contrast levels were not a limiting factor for achieving similar saliency across the attributes. When path elements were orthogonally oriented to the path, performance fell to near chance levels for both observers. This establishes that the linking is critically dependent on the alignment of the path elements and that the positional cues were very weak, or absent. The performances in single attribute conditions were degraded identically with increasing path angle. Finally, there was no facilitation whatsoever when edges of several attributes with the same orientations were superimposed on the same locations. This means that if independent channels are used in the path formation processes, no use is made in the decision stage as offered by probability summation.

Good continuation and closure have been proposed as two gestalt laws of perceptual organization. It has been demonstrated that closed paths are easier to detect than open paths when the path elements are created by luminance contrast, the 'closure effect' (Figure 18a). It was further investigated in Experiment 2 if the rules limiting path formation were the same across the attributes concerning the closure effect. Further, as in Experiment 1, it was investigated if path detection facilitation occurs when edges of several attributes were superimposed along the paths. The closed paths were formed by adding a sinusoid with amplitude 10% of the radius of the path and three periods along the path. The distance between successive path elements was slightly varied randomly along the path. This created closed paths with path elements irregularly positioned along the path (Figure 18). Open S-shaped paths were created by initially cutting closed paths at a randomly chosen position and then unfolding these closed paths. The range of random element orientation around perfect path alignment, the α-range, was across conditions while the percent correct responses were measured with a 2IFC procedure. When the α-range was zero, all the path elements were aligned with the path (Figure 18a).

Figure 19b shows the estimated α-range thresholds for open and closed paths (composed of 12 elements) required to obtain 75% correct detections. The results from Experiment 2 indicated that the closure effect was attribute-invariant. As in Experiment, 1 it was shown that no facilitation occurred when several attributes were superimposed (Figure 19b). The results from Experiments 1 and 2 are accounted for by binding of attributes at either the first or the second stage in the two-stage model of edge linking (Model 2 and Model 3 in Figure 2). Late binding across attributes may occur at the linking stage or at the decision stage. However, if binding across attributes occurs at the decision stage so that attribute-specific channels mediate linking (Model 1 in Figure 2), the possibility of facilitation offered by probability summation is not used.

Can linking occur between edge elements created by different attributes? This question was addressed in Experiment 3 by using four aligned edge elements of one attribute interleaved with four elements of another attribute along the path. If $\overline{}$ is used as a label for edge elements created by one attribute and $\sim$ is used as a label for the other attribute, a path can be drawn as:

— ~ — ~ — ~ — ~ (1)

The performance in such inter-attribute conditions in Experiment 3 was compared with performances in conditions where the edges in one of the attributes were turned off. This resulted in four edge elements in one attribute (say ~) interleaved with four unoriented micropatterns which may be labeled with ? as shown below:

? ~ ? ~ ? ~ ? ~ (2)

It was argued that paths (1) and (2) would be indistinguishable in the eyes of a single attribute-specific channel provided with input from ~ elements only. The edges created in the nonpreferred attribute in path (1) would be invisible, providing noise at most as in case (2). As an analogy, consider a path composed of alternating luminance and color defined edge elements. A color-blind person would only be able to distinguish the luminance-defined edges interleaved with 'noise' elements carrying no visible edges.

Figure 19c shows the accuracy in percent correct path detection for two observers averaged across three path angles (β=10, 20, and 30 deg). The element orientation (α) was zero. Performance was much better than predicted by probability summation of independent channels, as calculated from the results obtained in the single attribute conditions (Lum, Mot, Dis) when the attributes were interleaved along the path (Lum-Dis, Lum-Mot, Mot-Dis).

The same attribute was used along the paths in Experiment 1, whereas two attributes were interleaved in the inter-attribute conditions in Experiment 3. If the linking process is completely attribute-invariant, the performance in Experiment 1 should not exceed the performance in corresponding inter-attribute conditions in Experiment 3. Only L.P. (see Figure 19) participated in both experiments. The performance for L.P. averaged across path angles β=10, 20, and 30 deg was 88%, 84%, and 83% correct detection of luminance, motion, and cyclopean paths, respectively, compared with 77%, 77%, and 74% correct path detections of luminance/motion, luminance/cyclopean, and motion/cyclopean inter-attribute conditions, respectively. So, the path linking process is not completely blind to the attributes used to create the path elements, which indicates the existence of an attribute-specific component in the path linking process.

Briefly, path detection facilitation by using several attributes occurred only when the different attributes alternated along the path (Figure 19c), not when they were superimposed on the same locations (Figure 19a). These conditions are analogous to the illusory contour formation illustrated in Figure 12a, where the inducing elements were defined by several attributes in conjunction, and in Figure 12c, where illusory contours formed between edge elements created by different attributes. The facilitation found in Experiment 3 was superior to the performance predicted by probability summation of independent channels calculated from the results in the corresponding single attribute conditions, providing evidence for attribute-invariance (although not complete) in the

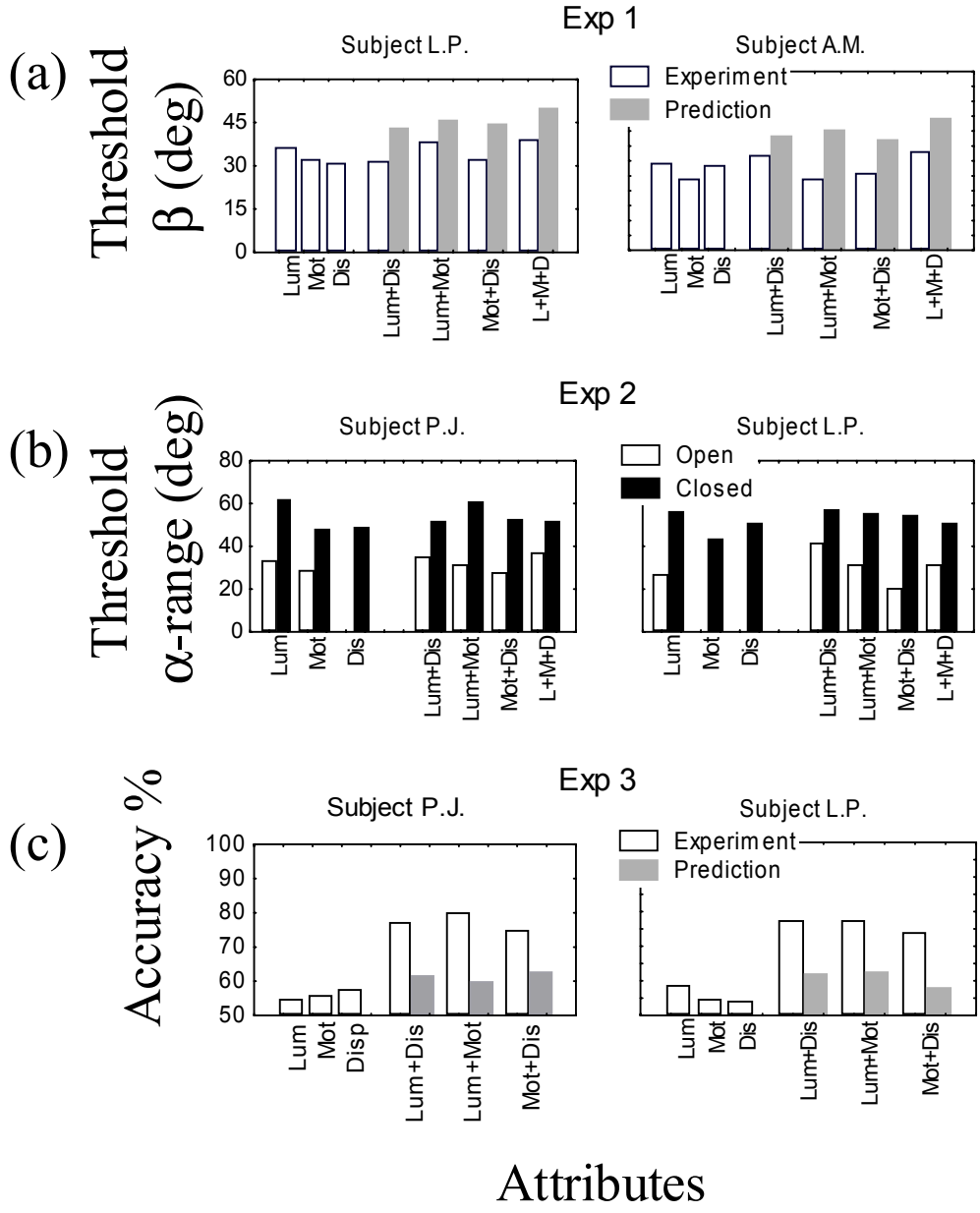

Figure 19. (a) The estimated 75% path detection thresholds, expressed in path angle (β) from Experiment 1, for two observers. Luminance- (Lum), motion- (Mot), and disparity- (Dis) defined path elements were used alone or in conjunction. The predictions are based on probability summation calculated from the proportion of correct estimates obtained in the single attribute conditions. **(b)** The estimated 75% thresholds from Experiment 2 expressed as α-range for two observers obtained for closed and open paths. The asymptote of the psychometric function fitted to the data was below 75% in some open path conditions. **(c)** The results from Experiment 3. The accuracy for two observers when single attributes were interleaved with noise elements compared with when two different attributes alternated along the path in inter-attribute conditions. The predictions are based on probability summation calculated from the proportion of correct estimates obtained in the single attribute conditions.

path-linking process as displayed in Models 2 and 3 in Figure 2. Taken together with the results from Experiments 1 and 2, these results provide evidence for the existence of attribute-invariance at the first stage in the figure-ground segmentation process as illustrated by Model 3 in Figure 2.

Binding the results

Spatial modulations of pictorial, temporal, and cyclopean fields in the visual image may mediate the perceived shape of an object. The existence of attribute-invariance in visual perception is suggested by the simple observation that only one version of a shape is perceived as opposed to multiple copies arising from the separate attributes (Regan, 2000). Accordingly, shape recognition processes in the inferior temporal cortex (IT) in the brain are known to be attribute-invariant (Sary et al., 1993; Kourtzi et al., 2002). Late binding may cause such invariance such that perception is a result of independent attribute-specific grouping processes whose outputs are bound at some subsequent stage, e.g., IT (Model 1 in Figure 2). Alternatively, attribute-invariance may occur at the first stage in which local image features such as edges are detected, e.g., V1 or V2 (Model 3 in Figure 2). An intermediate alternative is that one grouping process pools the responses across attribute-selective feature detectors (Model 2 in Figure 2).

Tilt illusions, illusory contours, and edge linking were used in three studies to probe spatial integration of edges across attributes and visual space. They all involve interactions between edges that are spatially separate in the visual field and have widely been used to probe visual processes with luminance-defined stimuli. The conclusion reached from luminance-based studies has been that all three phenomena result from neural processes taking place in cortical areas V1 and/or V2. When a subset of these experiments were repeated with inducing elements created by luminance, motion, and cyclopean contrasts, the results were similar across all attributes tested and mimicked those obtained with luminance-defined stimuli. Therefore, the conclusions reached from the earlier luminance-based studies are widened to include temporal- and cyclopean-defined stimuli.

Study 1 used an orientation illusion, the tilt effect, as a probe to investigate binding across attributes and orientation interactions across space. The double dissociation between tilt-repulsion and tilt-attraction effects in the luminance domain (where tilt repulsion effects can be abolished without any significant reduction of the attraction effects and vice versa) has been taken as evidence for different underlying neural processes causing these effects (Wenderoth & Johnstone, 1987, 1988a,b). It is believed that repulsion arises from local lateral inhibition between orientation and spatial frequency (size) selective cells like those in visual areas V1 and/or V2 whose response characteristics match those of the repulsion effect. Attraction is believed to be mediated by processes at higher levels in the visual brain or by large-scale interactions in V1/V2. Attraction and repulsion effects, as well as the double dissociation were found in Study 1 irrespective of the attributes used to create the inducing grating and the target grating, and even when different attributes were used for the inducer and target. Although these results

provide evidence for early interactions between the processing of edge orientation from different attributes, they may be due to interactions between attribute-invariant orientation detectors or interactions between different populations of attribute-specific detectors. Both these explanations may also account for the previously reported tilt repulsion effect induced by cross-adaptation with real and illusory contours (Berkeley et al., 1994).

Study 2 used illusory contours to investigate binding across attributes and space. Several studies have shown that thin subthreshold luminous lines summate with illusory contours cued by luminance edges and become detectable when superimposed on such contours (Dresp & Grossberg, 1997). Such summation provides psychophysical evidence that the perception of real and illusory contours are at least partly mediated by a common neural substrate. Complementary evidence for this idea comes from the existence of neurons in visual area V2 that respond to both real and Kanizsa-like illusory contours with the same orientation preference (Peterhans & von der Heydt, 1989). Near threshold summation was found in Study 2 irrespective of the attributes used to induce the illusory contours. As position cues were unable to account for this facilitation, it was concluded that summation occurred between the illusory fraction of the contours and the luminous target line, probably by activation of common neural elements. The summation provides evidence for attribute-invariant edge signaling mechanisms in that the luminous target line summed with the illusory contours induced by luminance, grating, motion, and cyclopean edges in the different experiments. It furthermore provides support for attribute-invariant long-range processes in illusory contour formation in that summation was found when a gap of 1 deg separated the inducing elements and the line. This makes it unlikely that the effect was accomplished with summation within single classical receptive fields (binding by conjunction cells) but involves long distance interactions (dynamic binding).

Study 3 demonstrated that the rules limiting path formation were the same across the attributes tested and that the closure effect was attribute-invariant. No facilitation occurred when several attributes were superimposed when creating the path elements as compared with when only one attribute was used. However, path detection was facilitated when two attributes alternated along the path composed of eight elements in comparison with when only one attribute was used with four elements along the path. The facilitation could not be accounted for by probability summation of independent channels as derived from the data obtained in the single attribute conditions. These results, together with those from Study 2, provide evidence for the existence of attribute-invariant edge linking mechanisms. However, performance was superior in conditions when all path elements were the same as compared with when two attributes alternated along the path (one observer). This result indicates that the path formation process has some access to which attributes created the path elements. Furthermore, the random curvature of the target paths makes it unlikely that single filters tuned to each path mediated the performance. Such a process would result in a combinatorial explosion of the number of filters necessary (Hess & Field, 1999).

Some psychophysical studies have found local attribute-invariant edge detecting mechanisms (e.g., Rivest & Cavanagh, 1995; Rivest et al., 1997), whereas others have

found long-range attribute-invariant grouping and comparator processes. For example, inter-attribute apparent motion is seen between two spatially separated successively flashed features even if the features are created by different attributes (Cavanagh et al., 1989). Moreover, human observers can compare the mean orientation, orientation difference, separation, and mean location of two widely separated bars created by different attributes, ignoring stimuli between the bars (Kohly & Regan, 2002). Finally, inter-attribute illusory contours are observed between inducing edges created by different attributes (Poom, 2001a).

Three main conclusions emerge from the three studies summarized in this thesis. First, local edge elements created by different attributes are detected at an early level of processing in the visual system (Studies 1 and 2). Second, edge detectors exist that sense local edges irrespective of whether they are created by luminance, motion, or cyclopean contrasts (binding by convergence) though attribute-specific edge detecting mechanisms cannot be excluded (Study 2). Third, the spatial grouping process results in attribute-invariant enhancement of local edge elements (dynamic binding) (Studies 2 and 3). In the following section, these results are discussed in terms of binding by convergence across attributes and attribute-invariant contextual modulations resulting in dynamic binding.

A new view on vision

Most studies investigating visual binding of edges across space, computational modeling of such processes, and edge statistics in natural images are based on first-order luminance edges. The studies presented here clearly demonstrate that the rules limiting binding of edges are attribute-invariant across the attributes tested and that such binding probably occurs at an early level of visual processing. Recent developments in natural image analysis and the Bayesian approach to understand perception as a probabilistic inverse optics problem may provide a link in understanding the cause of attribute-invariance at an early level of visual processing. Therefore, a short overview of the conclusions reached from these areas is provided, and the consequences drawn from these studies on binding across attributes and space are discussed.

Brunswik, Bayes, and binding

Within the Bayesian approach to perception, it is assumed that the primary goal of an observer is to estimate the conditional probabilities of the actual environmental layout given the visual image. Bayes theorem states that this probability (the posterior probability) is proportional to the product of two factors, namely, the probability of a specific environmental state (the prior probability) and the probability of the retinal image given that environmental state (the likelihood). By multiplying these factors, it is possible to determine the most likely environmental layout given the retinal image and solve the inverse optics problem. The prior probability can be interpreted as the observer's knowledge (implicit or explicit) of the world and the physical laws that regulate it; the

likelihood can be interpreted as knowledge of projective geometry and probabilities of specific viewpoints.

Statistical analysis of natural images has been used to find regularities that reflect properties of the environmental layout. For example, Brunswik and Kamiya (1953) emphasized the importance of the statistical properties of natural images to understand perception by showing that closely spaced parallel edges in a retinal image are likely to originate from the same object. This finding provided a functional basis for the gestalt rule of proximity. However, it is not until recent developments of powerful statistical methods and computers that the statistics of natural images has been widely explored. For example, the gestalt law of good continuation has a statistical basis in natural images (Geisler et al., 2001) and the arrangements of edges across the whole visual field follow a simple rule of co-circularity (Sigman et al., 2001). Specially, the high conditional probability (likelihood) that edge co-occurrences occur given that pairs of edge segments belong to the same or to different objects provides a functional reason for the gestalt law of good continuation (Geisler et al., 2001).

So far, natural image analysis of edge co-occurrences has been restricted to luminance edges, but one might suspect that similar co-occurrences exist between motion and cyclopean edges as they also signal figure-ground relations. How then is the information across attributes represented as neural activity?

Binding by conjunction cells

A question that has not received much attention until recently is why do the classical receptive fields look like they do? Attneave (1954) and Barlow (1961) proposed that information theory (developed by Shannon, 1948) might provide a tool to establish a link between the information contained in natural images and neural processing. Since then, powerful computers and statistical tools have been used for such analyses (for a review, see Simoncelli & Olshausen, 2001). From the perspective of information theory, the goal of visual processing is to transform the raw sensory input to an efficient representation by eliminating redundant information.

Based on simulation results showing that relatively few simulated simple cells were activated by natural images, Field (1987) proposed a sparseness constraint whereby the primary visual areas try to minimize the activities of these cells while preserving as much information as possible. The basic idea is to remove redundancies and represent images by using components (basis functions) of the visual scene that occur independently of each other for large numbers of natural scenes. Such representation is sparse in the sense that only a few of the total number of basis functions are needed to represent a particular image. Accordingly, Vinje and Gallant (2000) found evidence that the activities of visual neurons were much sparser during presentations of natural images than artificial non-natural images.

Olshausen and Field (1996, 1997) trained a neural network to maximize the sparsity of the representation of natural images. The basis functions that emerged after training on hundred thousands of image patches taken from natural images, starting from initially

random conditions, provided a good fit of the classical receptive field properties of simple cells in cortical area V1. Similarly applied sparseness constraints have been used to model the receptive fields of complex cells (Hyvärinen & Hoyer, 2000). It is reasonable that the basis functions are tuned to edges and lines of various length and sizes because natural images can be described and recognized by a small collection of such edges and lines by drawings on canvas.

It seems likely that the visual system compresses information carried by other attributes than luminance. As in the luminance domain, this would result in orientation and spatial frequency selective basis functions (receptive fields or filters) operating on, for example, color, texture, cyclopean, and temporal image properties. Psychophysical studies on humans have shown that orientation and spatial frequency tuned filters exist for color, texture, motion, and cyclopean edges and gratings. These filters are similar to those for luminance-defined edges and gratings (Tyler, 1974; Regan, 1989; Hamstra & Regan, 1995; Gray & Regan, 1998; Kwan & Regan, 1998, Hogervorst et al., 2000). Luminance-defined filters have a more narrow width than the texture, motion, and cyclopean filters. This might be because the processing of individual texture elements must occur before second-order attributes can be conveyed. The second-order filters may require input from first-order filters, which may explain why the second-order filters pass lower spatial frequencies than purely luminance-based filters (Regan, 2000).

Studies 1 and 2 in this thesis provide psychophysical evidence for binding of superimposed luminance, kinetic, and cyclopean edge segments by conjunction cells in humans. Further psychophysical evidence for the existence of attribute-invariant filters comes from demonstrations of perceptual pooling of luminance and cyclopean edges (Regan, 2000), luminance and color edges (Rivest et al., 1997; Suzuki & Cavanagh, 1998; Suzuki & Rivest, 1998), and luminance, color, and motion edges (Rivest et al, 1997).

Some neurophysiological studies provide evidence for the existence of conjunction cells that respond to edge segments across attributes and thereby support early attribute-invariance. A subset of V2 monkey neurons shows selective responses to the same orientation of both cyclopean and luminance-defined edges (von der Heydt et al., 2000) and motion and luminance edges (Marcar et al., 2000). Leventhal et al. (1998) found that many neurons in area V2 in monkey brain responded to edges with the same preferred orientation irrespective of whether luminance, texture, or motion was used to create the edges; only a few such cells were found in area V1.

In conclusion, orientation and spatial frequency selective filters operating in the luminance domain provide a sparse code of natural images. Similar filters operating on other attributes may have the same functioning principle. Attribute-invariant filters, or conjunction cells, compress the information in images and bind the attributes in a single process. Such early binding increases the sparseness of the image representation in that only one filter is required across several attributes.

Dynamic binding and computational modeling

Provided that natural images are compressed in the best possible way across attributes, how is figure-ground segmentation achieved by the outputs from cells with local orientation selective receptive fields? One way would be to let the responses from orientation selective cells converge on subsequent curvature sensitive conjunction cells. However, this leads to a combinatorial explosion of the required number of such conjunction cells when detection of randomly curved contours is required (Hess & Field, 1999). A better alternative is dynamic binding whereby curved contours are represented at the population level of cell responses. Facilitated activation along smooth contours can be achieved by the wiring structure of excitatory and inhibitory connections between orientation selective cells.

The contextual modulation of responses from orientation selective neurons corresponds well to the statistical regularities between edge elements in natural scenes, as well as to human performance in grouping experiments. This observation suggests that the primary visual cortex develops to detect regularities arising from natural images. Modulations of connectivity may be mediated by Hebbian learning, which may be summarized with the widely used slogan, 'cells that fire together wire together'. Correlated image features in the long run result in increased excitatory interactions between cells responding to these features (the basis functions). Because the features can be defined by different attributes and these attributes work in conjunction in natural images, it should be expected that attribute-invariant feature detectors develop that are highly interconnected.

Neurophysiological studies show that similar contextual modulations of the same magnitude occur in primary visual cortex regardless of the stimulus attribute used and of the number of attributes used to define the figure (Lamme, 1995; Zipser, Lamme, & Schiller, 1996). Recall that this mimic the results obtained from the path linking experiments in Study 3 in which no facilitation was found by adding contrasts of several attributes to the path elements.

Two-stage computational models are commonly used to describe figure-ground segmentation. Local features, such as edges and their orientations, are detected at the first stage. Contextual modulations of these signals occur at the second stage by lateral connections within the same area or feedback signals from other areas (e.g., Grossberg & Mingolla, 1985; Li, 1998). In such models the connection strength between two edge detectors decreases as the angle between their orientation selectivity increases, their alignment decreases, and the distance between their receptive-fields increases. Connections may also be inhibitory between units with overlapping or adjacent receptive fields having different orientation selectivity and between units having spatially separate receptive fields that signal non-collinear edge segments.

The excitatory connection patterns in these models mimic the patterns of connectivity found between orientation selective cells in the brain. The connectivity also embodies the gestalt rule of good continuation and the association field limiting edge linking as illustrated in Figure 20 (Field et al., 1993), as well as the relatability constraint limiting illusory contour formation illustrated in Figure 11 (Kellman and Shippley, 1991). Thus,

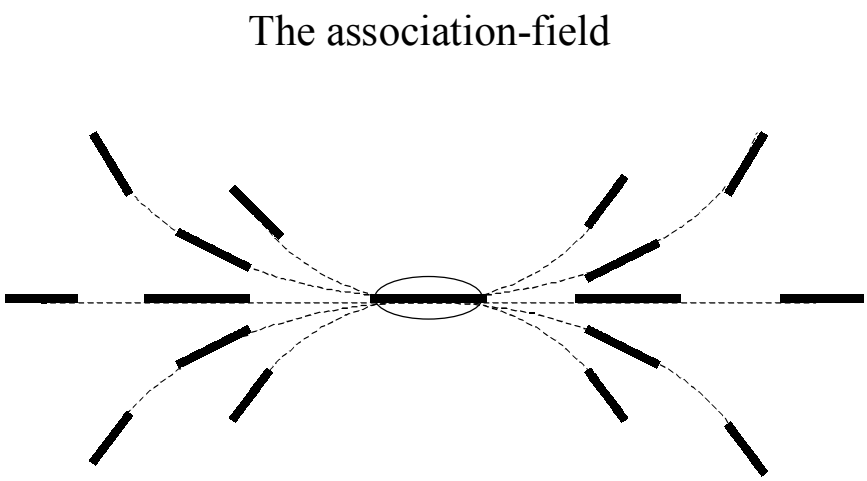

Figure 20. The rules limiting perceptual edge linking can be illustrated with an association field (Field, et al., 1993). If both lobes are stimulated with lines or edges oriented as shown by the line segments, they facilitate the detection of the line segment in the middle or give rise to the perception of an illusory contour. The lengths of the line segments indicate the strengths of facilitation between the middlemost edge segment and the flanking ones. Similar connection structure has been found between orientation selective cells in primary visual cortex, and in the statistics of edge co-occurrences in natural images.

the resulting connection structure results in enhancement of activity that is caused by smooth collinear edges in images and accounts for the distinctiveness through which collinear edges and illusory contours are perceived.

Figure-ground segmentation models have almost exclusively been formulated to account for contextual modulation of luminance edges. However, most figure-ground segmentation models can be extended to other attributes by assuming several attribute-selective channels using the same mechanism (Model 1 in Figure 2). Another option is to make the first-stage edge detectors attribute-invariant (Model 3 in Figure 2) or by letting the linking process operate on several attribute-selective edge detectors (Model 2 in Figure 2).

The results from Studies 1 and 2 supply evidence for the existence of attribute-invariant edge detectors. Finally, the results from Experiments 2 and 3 provide evidence for attribute-invariant edge linking processes (although the existence of attribute-specific processes cannot be excluded). Taken together the results furnish evidence for the existence of early attribute-invariance in both edge detection and grouping processes as described by Model 3 in Figure 2.

Edge detection precedes surface 'filling-in'

At the expense of creating a sparse code, representating edges, gratings, and lines, representations of homogenous regions in images are lost. Still, there is evidence for an active neural 'filling-in' process operating in early visual areas. The phenomena of perceptual filling-in occur when someone perceives a physically local blank region in

space as having the same properties as the surrounding regions, such as brightness, color, and texture. The filling-in process is typically exemplified with the fact that we do not see a hole in the visual field during monocular viewing resulting from the blind spot; instead, this region appears to have the same brightness, texture, color, and depth as the surrounding area. Similar filling-in of surface properties is found by patients with retinal scotomas.

When Ramachandran and Gregory (1991) presented displays with twinkling noise and a 1.5 x 1.5 deg gray square at 6 deg eccentricity from a fixation point, the square faded away in about 5 sec and appeared to be filled in from the outside to the inside by the twinkling dots. The gray square appeared like an 'artificial scotoma' that was filled-in by the surrounding surface property. Further, when a smaller disc enclosed by a larger one of different color is stabilized on the retina, it becomes invisible and the color of the larger disc spreads over it (Krauskopf, 1963).

The recently discovered 'water color' illusion in which a colored line flanking a darker line appears to spread its color onto enclosed areas is an additional example of surface filling-in (Pinna, Brelstaff, & Spillman, 2001). Inspection of the stereoimages in Figure 7c and Figure 12 also provides evidence that edges guide surface formation. The depth of the whole interior region inside the uniform disc in Figure 7c or the illusory squares in Figure 12 lacks disparity signals but acquires the same depth value as the inducing edges.

Phantom edges, or contours (Rogers-Ramachandran & Ramachandran, 1998) in which adjacent fields of black and white spots flicker in counter-phase on an intermediate gray background provide evidence for a dissociation between surface and edge detection processes. Phantom contours are seen for flickering frequencies of 15 Hz though the adjacent fields are seen as identical. The surface characteristics (the out of phase flicker) can only be perceived for much lower frequencies (about 7 Hz).

Several studies on brightness, color, and texture perception have reported that visual detection of edges and boundary formation in these attributes precedes the filling-in of surface properties on both sides of the boundary. It seems that across studies boundary detection is about twice as fast as surface filling-in (Lamme et al., 1999). When the brightness of a homogenous disc is gradually changed, the alteration is perceived as beginning at the border of the disc and spreading inwards (Paradiso & Nakayama, 1991; Paradiso & Hahn, 1996). Similar results have been obtained with texture surfaces and borders (Caputo, 1998).

Neural correlates of surface filling-in have recently been studied. For instance, contextual response modulations of neurons in area V1 at the population response level occur where neural boundary signals precede filling-in signals of enclosed surfaces (Lamme et al., 1999). Generally, the response of many cortical neurons consists of a transient peak response followed by a much slower decay. The initial transient peak is insensitive to contextual modulations, but the following decay is sensitive to such modulations. These two parts of the response seem to reflect separate contributions of the fast feedforward and possibly lateral transport of signals and the delayed feedback transport from higher-level areas in the brain (Lamme & Roelfsema, 2000). Contextual modulation within enclosed areas of the activities of V1 cells does occur after a delay of

about 80-100 ms, as compared with the initial response peak at 30 ms after stimulus onset (Maunsell & Gibson, 1992; Lee & Nguyen, 2001). Remarkably, this quick peak coincides well with the 25 ms for the peak of the tilt repulsion illusion to occur, whereas the tilt attraction has longer temporal dynamics (Wenderoth & Johnstone, 1988b). Recollect that the tilt repulsion illusion had a short spatial range and was believed to depend on inhibitory couplings between orientation selective cells in the brain.

The feedforward flow of information is highly retinotopic in the first visual areas and accounts for the small size classical receptive field (cRF) properties of V1 cells. The feedback signals may account for contextual modulation typical for the non-classical receptive fields (for a review, see Gilbert, 1998). In line with this view, extrastriate lesions eliminate the surface signals but not the edge signals (Lamme et al., 1998). These results suggest that different attributes are filled-in after the edges are detected and binding of edges has created extended contours across space, irrespective of the attributes used to create the edges. Neural correlates of such filling-in have been found in brain areas V2 and V3 in alert monkeys (De Weerd et al., 1995). Remarkably, the contextual modulation of V2 and V3 cells (De Weerd, 1998) and V1 cells within closed contours is not present during anesthesia (Lamme et al., 1998) or when higher areas are inactivated by cooling (Hupé et al., 1998). Supported by such results it has been suggested that the low-level visual areas V1 and V2 act as 'active blackboards', binding signals that have been processed in higher visual areas and mediating awareness (Lamme & Roelfsema, 2000; Bullier, 2001; Pascal-Leone & Walsh, 2001).

Concluding remarks

Physical laws regulate the natural world and the optics of light. These laws have been utilized by living creatures to develop eyes and the accompanying neural system required for interpreting the images that are perceived through the lenses of the eye and recorded in the brain. First-order attributes (e.g., luminance and color) and second-order attributes (e.g., texture) carry visual information that requires integration of image points across an image. These attributes can be conveyed by single static images and are therefore used by artists to provide illusions of depth and structure on canvas. The temporal and cyclopean fields are other second-order attributes that are a bit more complicated to demonstrate. The former attribute requires that image points are integrated across time and the latter attribute requires integration of image points across the two eyes. In this thesis, these three qualitatively different kinds of integration of spatial information have been labeled the 'three kinds of vision'. Nevertheless, the experimental results show that a single processing stream exists that mediates our unified visual experience of figure-ground segmentation, although attribute-specific channels are not excluded. Conjunction cells form a sparse code and subsequent dynamic binding accounts for the visual grouping of collinear edge segments. Surface properties (color, luminance, depth, motion, etc.) are thereafter filled-in within enclosed areas.

The raw first-order luminance (as displayed by the landscape representation in Figure 1b) and color fields provide the necessary input to subsequent representation of second-

order attributes. At this level, it may be possible to detect edges in both first- and second-order attributes (first stages in the models shown in Figure 2). Early combination across attributes may be beneficial because noise or ambiguities in one attribute may be overcome with other attributes, segmentation may speed up, and sharp contrast in any attribute may be used for filling-in surface properties of other attributes (Møller & Hurlbert, 1997). Regularities in images, both across spatial locations and attributes, may provide information about regularities in the world, and the organization of the brain reflects the organization of the physical world with which the brain interacts. If regularities across attributes and spatial locations co-occur, which is expected in natural images, attribute-invariant processes should develop as shown by the three studies presented in this thesis.

References

Alais, D., Blake, R., & Lee, S. H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience*, 1, 160-164.

Anderson, B. L., & Nakayama, K. (1994). Toward a general theory of stereopsis: binocular matching, occluding contours, and fusion. *Psychological Review*, 101, 414-445.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61, 183-193.

Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, 20, 8188-8198.

Banton, T., & Levi, D. M. (1993). Spatial localization of motion-defined and luminance-defined contours. *Vision Research,* 33, 2225-2237.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In *Sensory Communication* (ed. W. A. Rosenblith), pp. 217-234. Cambridge, MA: MIT Press.

Barrow, H. G., & Tennenbaum, J. M. (1981). Interpreting line drawings three-dimensional surfaces. *Artificial Intelligence*, 17, 75-116.

Berkeley, M. A., DeBruyn, B., & Orban, G. (1994). Illusory, motion, and luminance-defined contours interact in the human visual system. *Vision Research,* 34, 209-216.

Biederman, I. (1988). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.

Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature,* 228, 37-39.

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in three shrew striate cortex. *Journal of Neuroscience*, 17, 2112-2127.

Brunswik, E. (1947). *Systematic and representative design of Psychological Experiments*. Berkeley: University of California Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.

Brunswik, E., & Kamiya, J. (1953). Ecological cue validity of 'proximity' and other Gestalt factors. *American Journal of Psychology*, 66, 193-217.

Bullier, J. (2001). Feedback connections and conscious vision. *Trends in Cognitive Sciences*, 5, 369-370.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, London*, 197, 551-566.

Caputo, G. (1998). Texture brightness filling-in. *Vision Research*, 38, 841-851.

Carman, G. J., & Welch, L. (1992). Three-dimensional illusory contours and surfaces. *Nature*, 360, 585-587.

Cavanagh, P., Arguin, M., & von Grunau, M. (1989). Interattribute apparent motion, *Vision Research*, 29, 1197-1204.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, 4, 103-129.

Cavanagh, P. (1989). Multiple analyses of orientation in the visual system. In *Neural Mechanisms of Visual Perception,* (Eds: Lam, D. M. K. & Gilbert, C.), pp. 261-279. Woodlands, Tex,: Portfolio Publishing. 713.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America, A,* 5, 1986-2007.

Dan, Y., Attick, J. J., & Reid, R. C. (1996). Efficient Coding of Natural Scenes in the Lateral Geniculate Nucleus: Experimental Test of a Computational Theory. *Journal of Neuroscience*, 16, 3351-3362.

Das, A., & Gilbert, C. D. (1995). Long-range horizontal connections and their role in cortical reorganization revealed by optical recording in cat primary visual cortex. *Nature*, 375, 780-784.

De Weerd, P. (1998). Linking spread of neural activity and filling in: a few more arguments in favor. *Behavioral and Brain Sciences*, 21, 754-755.

De Weerd, P., Gattass, R., Desimone, R., & Ungerleider, L. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature*, 377, 731-734.
Dresp, B. (1997). On 'illusory' contours and their functional significance. *Current Psychology of Cognition*, 16, 489-518.
Dresp, B., & Bonnet, C. (1995). Subthreshold summation with illusory contours. *Vision Research*, 35, 1071-1078.
Dresp, B., & Grossberg, S. (1997). Contour integration across polarities and spatial gaps: From local contrast filtering to global grouping. *Vision Research*, 37, 913-924.
Dresp, B., & Grossberg, S. (1999). Spatial facilitation by color and luminance edges: boundary, surface, and attentional factors. *Vision Research*, 39, 3431-3443.
Ebenholtz, S. M., & Benzschawel, T. L. (1977). The rod and frame effect and induced head tilt as a function of observation distance. *Perception and Psychophysics*, 22, 531-538.
Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual Review of Neuroscience*, 23, 441-471.
Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4, 2379-2394.
Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integrations by the human visual system: evidence for a local 'association' field. *Vision Research*, 33, 173-193.
Frost, B. J., Wylie, D. R., & Wang, Y. C. (1990). The processing of object and self-motion in the tectofugal and accessory optic pathways of birds. *Vision Research*, 30, 1677-1688.
Georgeson, M. A. (1973). Spatial frequency selectivity of a visual tilt illusion. *Nature*, 245, 43-45.
Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurences in natural images predicts contour grouping performance. *Vision Research*, 41, 711-724.
Gibson, J. J. (1937). Adaptation, aftereffect and contrast in the perception of tilted lines: II. Simultaneous contrast and areal restriction of the aftereffect. *Journal of Experimental Psychology*, 20, 553-569.
Gibson, J. J. (1950). *The Perception of the Visual world*. Boston: Houghton Mifflin.
Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
Gibson, J. J., Kaplan, G., Reynolds, H., & Wheeler, K. (1969). The change from visible to invisible. *Perception and Psychophysics*, 5, 113-116.
Gibson, J. J., & Radner, M. (1937). Adaptation, aftereffect and contrast in the perception of tilted lines: I. Quantitative studies. *Journal of Experimental Psychology*, 20, 453-467.
Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron*, 9, 1-13.
Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Reviews,* 78, 467-485.
Gilbert, C. D, & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research*, 30, 1689-1701.
Gonzalez, R. C., & Woods, R. E. (1993). *Digital image processing*, New York: Addison Wesley Publiching Company.
Gray, R., & Regan, D. (1998). Spatial frequency discrimination and detection characteristics for gratings defined by orientation texture. *Vision Research*, 38, 2601-2617.
Gregory, R. L. (1972). Cognitive contours. *Nature*, 238, 51-52.
Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons signal 'illusory' contours. *Nature*, 365, 550-552.
Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon colour spreading. *Psychological Review*, 92, 173-211.
Halpern, D. F. (1981). The determinants of illusory contour perception. *Perception*, 10, 199-213.
Hamstra, S. J., & Regan, D. (1995). Orientation discrimination in cyclopean vision. *Vision Research*, 35, 365-374.
Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to the illumination of the retina. *American Journal of Physiology*, 121, 400-421.
Hess, R. F., & Field, D. (1999). Integration of contours: new insights. *Trends in Cognitive Sciences*, 3, 480-486.

Hoffman, D., & Richards, W. (1984). Parts of recognition. *Cognition*, 18, 65-96.
Hogervorst, M. A., Bradshaw, M. F., & Eagle, R. A. (2000). Spatial frequency tuning for 3D corrugations from motion parallax. *Vision Research,* 40, 2149-2158.
Horridge, G. A., Zhang, S. W., & Ocarroll, D. (1992). Insect perception of illusory contours. *Philosophical Transactions of the Royal Society of London, B*, 337, 59-64.
Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106-154.
Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, 394, 784-787.
Huxlin, K. R., Saunders, R. C., Marchionini, D., Pham, H.-A., & Merigan, W. H. (2000). Perceptual deficits after lesions of inferotemporal cortex in macaques. *Cerebral Cortex,* 10, 671-683.
Hyvärinen, A., & Hoyer, P. O. (2000). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Research*, 41, 2413-2423.
Johansson, G. (1953). *Configurations in event Perception*. Uppsala: Alqvist & Wiksell.
Julesz, B. (1960). Binocular depth perception of computer generated patterns. *Bell System Technical Journal,* 39, 1125-1161.
Julesz, B., & Frisby, J. P. (1975). Some new subjective contours in random line stereograms. *Perception*, 4, 145-150.
Kanade, T. (1980). A theory of origami world. *Artificial intelligence*, 13, 279-311.
Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimulazione omogenea. *Rivista di Psicologia*, 49, 7-30 (Also translated in: S. Petry and G. E. Meyer Eds. *The perception of illusory contours*, pp. 40-49, New York/Berlin: Springer Verlag, 1987).
Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, 15, 843-856.
Kaplan, G. A. (1969). Kinetic disruption of optical texture. *Perception and Psychophysics*, 6, 193-198.
Kellman, P. J., & Cohen, M. H. (1984). Kinetic subjective contours. *Perception and Psychophysics*, 35, 237-244.
Kellman, P. J., & Shippley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, 23, 141-221.
Kern, R., Egelhaf, M., & Srinivasan, M. V. (1997). Edge detection by landing honeybees: behavioural analysis and model simulations of the underlying mechanisms. *Vision Research*, 37, 2103-2117.
Kohly, R. P., & Regan, D. (2002). Fast long-distance interactions in the early processing of motion-defined form and of combinations of motion-defined, luminance-defined, and cyclopean form. *Vision Research*, 42, 969-980.
Kourtzi, Z., Bülthoff, H.H., Erb, M., & Grodd, W. (2002). Object-selective responses in area MT/MST. *Nature Neuroscience*, 5, 17-18.
Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one. *Proceedings of the National Society of Sciences*, 90, 7495-7497.
Krauskopf, J. (1963). Effects on retinal image stabilization on the appearance of heterochromatic targets. *Journal of the Optical Society of America*, 53, 741-744.
Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16, 37-68.
Kulikowski, J. J., & King-Smth, P. E. (1977). Spatial arrangement of line, edge, and grating detectors revealed by subthreshold summation. *Vision Research*, 13, 1455-1478.
Kwan, L., & Regan, D. (1998). Orientation tuned filters for texture-defined form. *Vision Research*, 38, 3849-3855.
Köhler, W. (1947). *Gestalt Psychology*. New York: Liveright.
Köhler, W., & Wallach, H. (1944). Figural aftereffects: an investigation of visual processes. *Proceedings of the American Philosophical Society*, 88, 269-357.

Lamme, V. A. F. (1995). The neurophysiology of figure ground segregation in primary visual cortex. *Journal of Neuroscience*, 15, 1605-1615.
Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekrejse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, 9, 406-413.
Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23, 571-579.
Lamme, V. A. F., Zipser, K., & Spekreijse, H. (1998). Figure-ground activity in primary visual cortex is supressed by anesthesia. *Proceedings of the National Academy of Science,* 95, 3263-3268.
Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in early visual cortex. *Proceeding of the National Academy of Sciences, USA*, 98, 1907-1911.
Leventhal, A. G., Wang, Y. Schmolensky, M. T., & Zhou, Y. (1998). Neural correlates of boundary perception. *Visual Neuroscience*, 15, 1107-1118.
Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10, 903-940.
Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7, 3416-3468.
Magnussen, S., & Johnsen, T. (1986). Temporal aspects of temporal adaptation. A study of the tilt aftereffect. *Vision Research*, 26, 661-671.
Malach, R., Amir, Y., Harel, M., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocyting injections in primate striate cortex. *Proceedings of the National Academy of Sciences USA*, 90, 10469-10473.
Marcar, V. L., Raiguel, S. E., Xiao, D., & Orban, G. A. (2000). Processing of kinetically defined boundaries in areas V1 and V2 of the macaque monkey. *Journal of Neurophysiology*, 84, 2786-2798.
Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and Company.
Maunsell, J. H. R., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, 4, 1332-1334.
McGraw, P. V., Levi, D. M., & Whitaker, D. (1999). Spatial characteristics of the second-order visual pathway reveled by positional adaptation. *Nature Neuroscience*, 2, 479-484.
Meyer, G. E., & Petry, S. (1987). Top-down and bottom-up: The illusory contour as a microcosm of issues in perception. In *The Perception of Illusory Contours*, (Eds: Petry, S., & Meyer, G. E.), pp. 3-20. New York: Springer-Verlag.
Morgan, M. J. (1996). Visual illusions. In *Unsolved Mysteries of the Mind*. (Ed. V. Bruce), pp. 29-58. Erlbaum: UK.
Møller, P., & Hurlbert, A. C. (1997). Motion edges and regions guide image segmentation by colour. *Proceedings of the Royal Society of London. B*., 264, 1571-1577.
Mustillo, P., & Fox, R. (1986). The perception of illusory contours in the hypercyclopean domain. *Perception and Psychophysics*, 40, 362-363.
Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity sensitive neurons, and space perception: a hypothesis. *Perception,* 3, 63-80.
Nieder, A. (2002). Seeing more than meets the eye. *Journal of Comparative Physiology A*, 188. 249-260.
Olshausen, B. A., & Field,D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 607-609.
Olshausen, B. A., & Field,D. J. (1997). Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Research*, 37, 3311-3325.
O'Toole, B. I., & Wenderoth, P. (1977). The tilt illusion: Repulsion and attraction effects in the oblique meridian. *Vision Research*, 17, 367-374.
Paradiso, M. A., & Hahn, S. (1996). Filling-in percepts produced by luminance modulation. *Vision Research*, 36, 2657-2663.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, 31, 1221-1236.
Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area nessesary for visual perception. *Science*, 292, 510-512.
Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience*, 9, 1749-1763.
Pinna, B., Brelstaff, G., & Spillman, L. (2001). Surface color from boundaries: a new 'watercolor' illusion. *Vision Research*, 41, 2669-2676.
Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33, 993-999.
Poom, L. (2000). Inter-attribute tilt effects and orientation analysis in the visual brain. *Vision Research*, 40, 2711-2722.
Poom, L. (2001a). Visual inter-attribute contour completion. *Perception*, 30, 855-865.
Poom, L. (2001b) Visual summation of luminance lines and illusory contours induced by pictorial, motion, and disparity cues. *Vision Research*, 41, 3805-3816.
Poom, L. (2002). Visual binding of luminance, motion and disparity edges. *Vision Research*, 42, 2577-2591.
Poom, L. (2002). Seeing stereoscopic depth from disparity between kinetic edges. *Perception*. 31, 1439-1448.
Prazdny, K. (1986). Illusory contours from inducers defined solely by spatiotemporal correlation. *Perception and Psychophysics*, 39, 175-178.
Ramachandran, V. S. (1992). Perception: A biological perspective. In *Neural Networks for Vision and Image Processing*. (Eds: G. A. Carpenter and S. Grossberg), pp. 45-91. Cambridge: MIT Press.
Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling-in of artificially induced scotomas in human vision. *Nature*, 350, 699-702.
Ramachandran, V. S., Rao, V. M., & Vidyasagar, T. R., (1973). The role of contours in stereopsis. *Nature* 242, 412-414.
Regan, D. (1989). Orientation discrimination for objects defined by relative motion and objects defined by luminance contrast. *Vision Research*, 29, 1389-1400.
Regan, D. (2000). *Human Perception of Objects*. Sunderland, Massachusetts: Sinauer Associates.
Rivest, J., Boutet, I., & Intriligator, J. (1997). Perceptual learning of orientation discrimination by more than one attribute. *Vision Research*, 37, 273-281.
Rivest, J., & Cavanagh, P. (1995). Localization of contours defined by more than one attribute. *Vision Research,* 36, 53-66.
Rivest, J., Intriligator, J., Warner, J., & Suzuki, S. (1997). Color and luminance combine at a common neural site for shape distorsions. *Investigative Opthalmology and Visual Science (ARVO abstracts),* 38, S1000.
Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research*, 38, 71-77.
Sary, G., Vogels, R., & Orban, G. A., (1993). Cue -invariant shape selectivity of macaque inferiortemporal neurons. *Science*, 260, 995-997.
Schofield, A. J. (2000). What does second-order vision see in an image? *Perception*, 29, 1071-1086.
Schumann, F. (1900). Beiträge zur analyse gesichtswahrnehmungen: Erste abhandlung einige beobeachtungen über die zusammenfassung von gesichteindrücken zu einheiten. *Zeitscrift für Psychologie und Physiologie der Sinnesorgane*, 23, 1-32 [Also translated in: *The perception of Illusory Contours*, (Eds: S. Petry and G. E. Meyer), pp. 21-34. New York/Berlin: Springer-Verlag, 1987].
Shannon, C. (1948). The mathematical theory of communication. *Bell Systems Technical Journal*, 27, 379-423.
Sheth, B. R., Sharma, J., Rao, S. C., & Sur, M. (1996). Orientation maps of subjective contours in visual cortex. *Science*, 274, 2110-2115.
Sigman, M., cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: natural scenes

and Gestalt rules. *Proceedings of the National Academy of Science*, 98, 1935-1940.
Simoncelli, E. P., & Olshausen, B. A. (2001). Neural image statistics and neural representation. *Annual Review of Neuroscience*, 24, 1193-1216.
Singer, W. (1999). Neural synchrony: a versatile code for the definition of relations? *Neuron*, 24, 49-65.
Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555-586.
Sinha, P. (2001). Role of motion integration in contour perception. *Vision Research*, 41, 705-710.
Spinelli, D., Antonucci, G., Daini, R., Fanzon, D., & Zoccolotti, P. (1995). Modulation of the rod-and-frame illusion by additional external stimuli. *Perception*, 24, 1105-1118.
Suzuki, S., & Cavanagh, P. (1998). A shape contrast aftereffect for briefly presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance,* 24, 1315-1341.
Suzuki, S., & Rivest, J. (1998) Interactions among "aspect ratio channels." *Investigative Opthalmology and Visual Science (ARVO abstracts), 38, S1000.*
Tolhurst, D. J., & Thompson, P. G. (1975). Orientation illusions and aftereffects: inhibition between channels. *Vision Research*, 15, 967-972.
Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature*, 251, 140-142.
Tyler, C. W. (1975). Stereoscopic tilt and size aftereffects. *Perception*, 4, 187-192.
Tyler, C. W., & Chen, C. -C. (2000). Signal detection theory in the 2AFC paradigm: attention, channel uncertainty and probability summation. *Vision Research*, 40, 3121-3144.
van der Zwan, R., & Wenderoth, P. (1994). Psychophysical evidence for area V2 involvement in the reduction of subjective contour tilt aftereffects by binocular rivalry. *Visual Neuroscience*, 11, 823-830.
van der Zwan, R., & Wenderoth, P. (1995). Mechanisms of purely subjective contour tilt effects. *Vision Research*, 35, 2547-2557.
Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation on primary visual cortex during natural vision. *Science*, 287, 1273-1276.
Virsu, V., & Taskinen, H. (1975). Central inhibitory interactions in human vision. *Experimental Brain Research*, 23, 65-74.
von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and neuron responses. *Science*, 224, 1260-1262.
von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, 40, 1955-1967.
von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology*, 5, 520-526.
Wade, N. J., & Wenderoth, P. (1978). The influence of color and contrast rivalry on the magnitude of the tilt aftereffect. *Vision Research*, 18, 827-835.
Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, 45, 205-217.
Ware, C., & Mitchell, D. (1974). The spatial frequency selectivity of the tilt aftereffect. *Vision Research*, 14, 735-737.
Wenderoth, P., Clifford, C. W. G., & Wyatt, A.M. (2001). Hierarchy of spatial interactions in the processing of contrast-defined contours. *Journal of the Optical Society of America, A*, 18, 2190-2196.
Wenderoth, P., & Johnstone, S. (1987). Possible neural substrates for orientation analysis and perception. *Perception*, 16, 693-709.
Wenderoth, P., & Johnstone, S. (1988a). The different mechanisms of the direct and indirect tilt illusion. *Vision Research*, 28, 301-312.
Wenderoth, P., & Johnstone, S. (1988b). The differential effects of brief exposures andsurrounding contours on direct and indirect tilt illusions. *Perception*, 17, 177-189.
Wenderoth, P., O'Connor, T., & Johnson, M. (1986). The tilt illusion as a function of the relative and absolute lengths of test and inducing lines. *Perception and Psychophysics*, 39, 339-345.
Wenderoth, P., van der Zwan, R., & Williams, M. (1993). Direct evidence for competition between local and global mechanisms of two-dimensional orientation illusions. *Perception*, 22, 273-286.

Wertheimer, M. (1923). Undersuchungen zür lehre von der gestalt. In *A source book of Gestalt psychology*, (Ed: W. D. Ellis), pp 71-88. London: Routledge and Kegan Paul.
Wheatstone, C. (1838). On some remarkable, and hitherto unresolved, phenomena of binocular vision. *Philosophical Transactions of the Royal society of London*, 2, 271-394.
Wong, E., & Weisstein, N. (1984). Flicker induces depth: Spatial and temporal factors in the perceptual segregation of flickering and nonflickering regions in depth. *Perception and Psychophysics*, 35, 229-236.
Yonas, A., Craton, L. C., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception and Psychophysics*, 41, 53-59.
Zhou,Y-X., & Baker, C, L. (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science*, 261, 98-101.
Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, 16, 7376-7389.