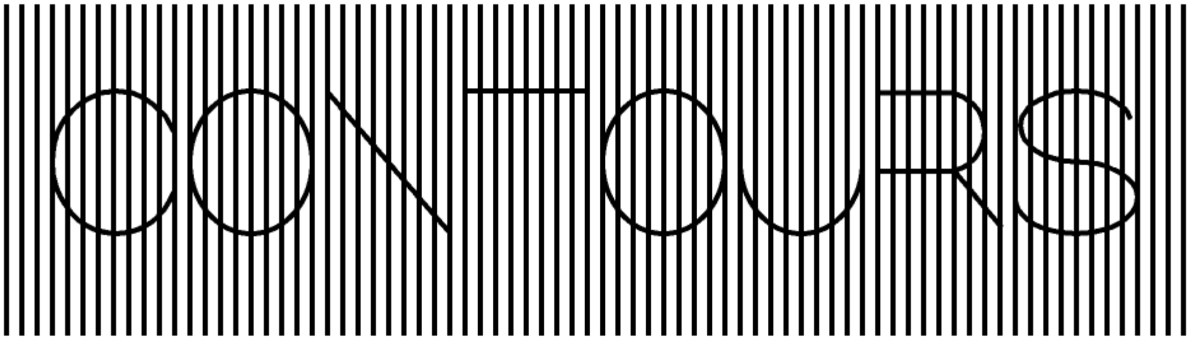
Lund University Cognitive Science

Jens Månsson

Occluding Contours: A Computational Model of Suppressive Mechanisms in Human Contour Perception



Lund University Cognitive Studies, 81
2000

OCCLUDING CONTOURS: A COMPUTATIONAL MODEL OF SUPPRESSIVE MECHANISMS IN HUMAN CONTOUR PERCEPTION

Jens Månsson

Lund University Cognitive Science
Kungshuset, Lundagård
S–222 22 Lund
Sweden

jens.mansson@lucs.lu.se

*Abstract:*A fundamental problem in vision is how to identify the occluding contours of objects and surfaces, given the ambiguity inherent in low-level visual input. A computational model is proposed for how occluding contours could be identified by making use of simple heuristics that reduce the ambiguity of individual features. In the striate cortex, a large majority of cells are selective for both contrast and orientation; i.e., they respond preferentially to simple features like contrast edges or lines. The heuristics we propose enhance or suppress the outputs of model striate-cortical cells, depending on the orientation and spatial distribution of stimuli present outside of the "classical" receptive field of these cells. In particular, the output of a cell is suppressed if the cell responds to a feature embedded in a texture, in which the "component features" are oriented in accordance with the orientation-selectivity of the cell. The model has been implemented and tested on natural as well as artificial grey-scale images. The model produces results that in several aspects are consistent with human contour/form perception. For example, it reproduces a number of known visual phenomena such as illusory contours, contour masking, pre-attentive pop-out (due to orientation-contrast), and it enhances contours that human observers often report perceiving as more salient.

1 INTRODUCTION

This paper addresses various questions related to human form perception, with a particular emphasis on how occluding contours might be processed in the visual cortex. An occluding contour can technically be defined as a contour that marks a discontinuity in depth (Marr, 1982). That is, if traced back to its source in the physical world, an occluding contour corresponds to the line on a surface where the view-line touches both the object and the background; or more formally where the view-line is tangent to the surface (fig. 1). Also considered here are contours that arise due to sharp changes in the orientation, or slant, of a surface; such as along edges/ridges. Although not always occluding, these contours are similar in that they define, or mark, abrupt changes in depth.

Occluding contours are interesting entities of early vision for a very simple reason: they mediate fundamentally important information about the 3-D structure of the physical environment. If accurately identified they can provide information on the position, orientation and extension of object and surface boundaries. This information in turn is crucial for a number of our visual abilities such as determining foreground-background relationships, segmenting the visual input into meaningful entities (objects), and recognising objects from shape, etc. Given how dependent we are on these abilities for solving even the simplest task, it is evident that accurate identification of occluding contours is an important key to effective and reliable visual scene analysis. Surely, the ability to identify (and the capability to use the information on) occluding contours gave our early ancestors an advantage over species who did not have it. And surely this ability will be an important component in the perceptual system of any artificial agent that is to interact with a "real world" environment.

In general, there is some kind of visual contrast along a depth-discontinuity; e.g. a contrast in luminance or colour, a difference in motion or binocular-disparity,

Lund University Cognitive Studies – LUCS 81. ISSN 1101–8453. 2000

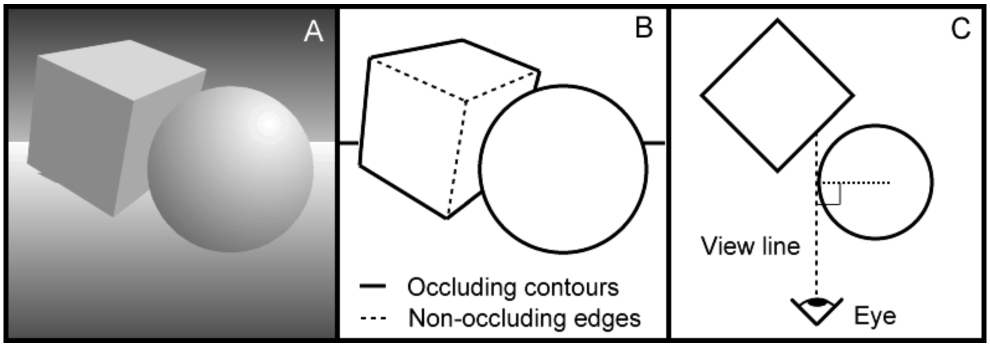


Figure 1: (A) A cube, partly occluded, by a sphere. (B) Occluding "outlines" (whole lines) and non-occluding edges (dashed). (C) Along an occluding contour, the view-line is orthogonal to the normal of the (occluding) surface.

or a discontinuity of pattern. Hence, the human visual system could potentially use a variety of different cues to identify occluding contours. Given that binocular-disparity and (relative) motion are particularly powerful cues to depth, our visual system most likely relies heavily on such information when available. However, even in the absence of such direct cues to depth (-discontinuities), we are often remarkably good at identifying occluding boundaries in a visual scene. Given, for example, a black and white photograph, we can usually rapidly identify the occluding contours of objects and surfaces, even if the scene or the objects in it have not been encountered before.

An interesting aspect of this ability is that we are usually not aware of the underlying process, or the computational difficulties that this process deals with. This suggests that the neural mechanism responsible for the identification of occluding contours mainly is a pre-attentive one, i.e. that it operates relatively autonomously from conscious influence. A growing body of neurophysiological, psychophysical and anatomical studies, described below, supports this view.

The main question this paper addresses is what operations these early low-level mechanisms perform on the visual input in order to produce useful representations of occluding contours; i.e. useful in the sense of assisting movement in, and manipulation of, the physical environment. In this paper, the discussion of possible mechanisms will be limited to visual input that is monocular, static and monochromatic. Of particular interest is the effect the local visual surround has in modulating how we perceive occluding contours. More precisely, how the arrangement and orientation of various low-level contrast features (e.g. edge and line segments) in the nearby surround could determine whether we perceive a visual structure as an occluding contour, or as a part of a surface texture/pattern.

2 COMPUTATIONAL CONSIDERATIONS

In a static monochromatic image, the only information that may reveal an occluding contour is the presence of some kind of luminance contrast along the contour; e.g. a contrast edge, a line or a pattern discontinuity. However, in images of natural scenes, contrast information may not only be found along the occluding boundaries of surfaces, but may also be found in the surfaces themselves, due to, for example, textures, patterns, shadow-lines and reflections (fig. 2). Hence, a major problem with identifying occluding contours is to discriminate contrast features that are caused by occluding contours from features that are produced by other physical structures and phenomena.

Figure 2: Surfaces often contain a variety of different contrast markings; due to for example texture (changes) and shadows.

Several factors make this discrimination difficult. First, the type of visual trace that exists along an occluding contour often changes from one point to another. That is, the type of feature, or local luminance pattern, that visually defines different parts of a contour, may change from, for example, a pattern discontinuity at one point, to a contrast edge or a line at some other point (fig. 3a). Such changes are caused by a wide variety of factors, such as variations along a contour in the texturing or reflection properties of a surface; or changes in the orientation of a surface, causing different amounts of light to be reflected into the eye. In other situations, the only physical evidence of an occluding contour may be a small disruption in a texture density, or a misalignment of the component features in a texture (fig 3b). From a computational perspective, this large variation in the type of feature/visual trace that can define a contour poses a delicate discrimination-problem to any visual system. That is, any useful discrimination strategy must not only "tolerate", or generalise over, many different features that may be present along an occluding contour, but must also be sensitive to features that are not caused by occluding structures, and discard the latter.

Further, there may be no visual trace at all along parts of an occluding contour. Such situations arise when, for example, there is no difference in the reflected luminance from (between) the occluded and the occluding surfaces, and there are no visible surface markings (see the middle left section of the sphere in fig. 3a). In such situations, the problem of identifying occluding contours is not so much a matter discrimination, but rather one of reconstruction, or "filling-in".

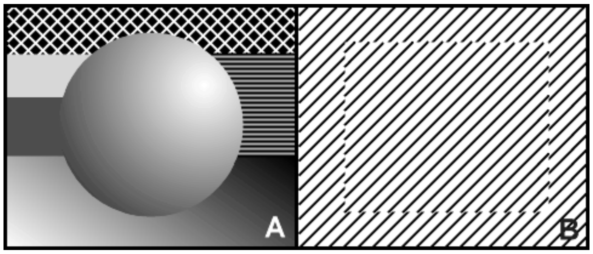


Figure 3: (A) Following the contour of the sphere around, the type (and polarity) of the contrast markings changes from point to point. (B) Rectangle defined only by a small pattern discontinuity.

Finally, problems may arise due to the loss of explicit depth information that occurs when a visual scene is projected onto a 2-D surface. Because the 3-D structure of a scene is compressed in the retinal projection, some contour (parts) may, for example, end up closer to contours that are caused by other objects than to contours originating from the same object. For the same reason, contour parts originating from different objects may also overlap each other in the retinal image. Hence, even if the individual parts of a number of contours have been correctly identified and/or filled-in, it may still be difficult to bind, or integrate, these parts appropriately into meaningful structures (objects etc.).

Considering these computational difficulties, it is remarkable how easily and rapidly we are able to pick out the occluding contours in a 2-D image, and how biased we are to perceive these structures as coherent entities, even if the visual information along them has a complex composition and/or is partly missing. Before turning to the question of how the human visual system handles these perceptual difficulties, it is first appropriate (in order to pose the proper questions) to briefly consider how visual input is represented at the cortical level.

3 REPRESENTATION OF VISUAL INPUT IN THE STRIATE-CORTEX

In the primary visual cortex (area V1) a large majority of cells are highly sensitive to visual stimuli that contain some oriented contrast. Depending on their response-properties to basic visual stimuli/features, these cells can be divided into three broad categories: simple, complex or hypercomplex (or end-stopped) cells (Hubel & Wiesel, 1962; Hubel, 1988). The receptive field of a typical simple cell is divided into two, three or more alternately excitatory and inhibitory sub regions, arranged in parallel bands along a common axis of orientation. Due to this receptive-field mapping, these cells respond strongly to stimuli such as contrast edges or lines of a particular orientation and polarity (i.e. contrast direction). Complex cells have slightly larger receptive fields and are not sensitive to the exact positioning of a stimulus/feature within their receptive field, but otherwise respond to similar stimuli as simple cells. The third category of cells, the hypercomplex or end-stopped cells, are also sensitive to oriented contrast patterns, but differ in one major aspect from the simple and complex cells. As the name suggests, the end-stopped cells only respond to features that terminate within their receptive fields (e.g. line- endings, corners). If the stimulus extends over their whole receptive field, the response is weakened or totally suppressed.

Another interesting but more global feature of the cortical organisation is that the topography of the retinal image, in general, is preserved in the striate-cortical representation (Kandel, 1991). That is, stimuli that are close together in the retinal image, will in general be represented by cells in area V1 that are situated near each other in the cortical tissue.

4 INHERENT AMBIGUITIES

How does our visual system identify occluding contours given 1) the computational difficulties discussed above, 2) the response properties of the cells in area V1, and 3) the fact that the retinal image is retinotopichally represented over the cortical surface? Because the receptive fields of the various simple and complex cells are relatively small compared to the whole visual field, it is evident that no individual cell can represent the presence of an occluding contour that spans a larger region of the visual field. Consequently, our visual system must, at some level of processing, integrate the responses from a potentially large number of simple/complex cells that may be firing along a contour line. However, in order for this integration mechanism to produce meaningful results, it should avoid integrating the responses from cells that fire due to causes other than occluding contours. That is, it should avoid integrating the responses from cells that fire to features caused by, for example, surface textures or shadow lines etc. But this situation creates somewhat of a paradox. Before it has identified an occluding contour, how can our visual system "know" which cells fire due to occluding contours, and which cells fire due to other causes? The problem is that, in general, the reason why any individual cell fires can not be unambiguously established by only considering the type of stimuli a cell is sensitive to. Consider, for example, a simple cell that responds optimally when a contrast edge is present within its receptive field. This cell will fire with equal strength whether the contrast edge is caused by an occluding contour, a shadow line, a reflection or some detail in a texture pattern. How then could this ambiguity be resolved?

One conceivable solution to this problem would be that some central higher-level process, which could integrate information from all over the visual field, simply tried out every possible combination of grouping the responses from the cells in V1 into contours, and then somehow determine the solution that seemed most plausible. This could involve comparing the results to stored representations of objects and scenes, consulting higher-level knowledge and experiences, and considering the context in which the stimuli was perceived.

Occasionally, such high-level processing might be needed to resolve certain perceptual ambiguities, but in general the perceptual process seems to be much faster and less accessible for conscious manipulation than such a scheme would suggest. Nor would it explain how we are able to identify occluding contours of unknown, or partly hidden, objects in unfamiliar contexts where no high-level knowledge or experience is relevant.

Moreover, leaving the disambiguation of low-level stimuli to such a late stage of processing leads to a combinatorial explosion in the number of ways there are to combine the responses from the cells in area V1 into different contour paths, even when the visual input is modestly complex. In other words, in its pure form the above scheme does not seem to account for how we identify occluding contours, but instead suggests that some of the response-ambiguity must be resolved at a much earlier stage, before or at the level where spatial integration takes place.

5 NEUROPHYSIOLOGY AND PSYCHOPHYSICS

A growing body of empirical evidence supports the view that important aspects of contour, or form, processing is carried out at a relatively early stage in the visual pathway. For example, Peterhans and von der Heydt (1993) described "contour neurones" in area V2 that respond not only to contrast edges or lines of a particular orientation, but also to pattern discontinuities (i.e. when the discontinuity is orientated in accordance with the cells orientation selectivity) and even to broken edges and lines (i.e. illusory contours). The fact that these cells seem to respond to discontinuities, invariant to the exact composition of the luminance pattern within their receptive field, strongly suggests that these cells are important for identifying occluding contours.

However, a number of other recent studies have shown that some form-related processing which could serve to facilitate the identification of occluding contours may be done as early as in area V1. Single cell recordings (Gilbert & Wiesel, 1990; Knierim & van Essen, 1992; Kapadia et al., 1995), and real-time optical imaging (Grinvald et al., 1994), have demonstrated that the firing rate of individual cells in area V1 is not exclusively determined by the stimulus present within a cells receptive field, but can be modulated (i.e. enhanced or suppressed) by stimuli located outside the receptive field.

More specifically, Kapadia et al. (1995) have shown that the firing rate of an individual complex cell, which in isolation responds to a bar of a certain orientation, can be enhanced if one or several other similarly oriented bars are positioned along the cell's axis of orientation, but outside its receptive field. They further showed that the enhancement effect decreased as the bars were i) separated along the common axis of orientation, ii) separated from co-linearity, or iii) separated in orientation (fig 4).

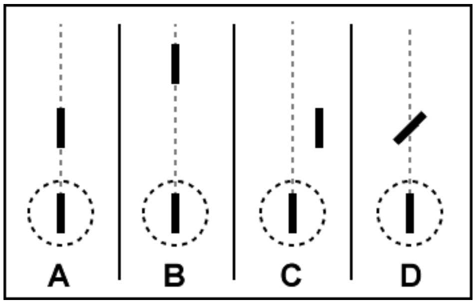


Figure 4: According to the study of Kapadia et al. (1995), the response of a cell was (A) enhanced when a bar, co-aligned with the cell's axis of orientation selectivity, was placed outside its receptive field (dashed circle). Further, the enhancement decreased if the bars were (B) separated along the common axis of orientation, (C) separated from co-linearity or (D) separated in orientation.

A related but suppressive effect has also been reported by Knierim and van Essen (1992), who have demonstrated that the firing rate of an individual cell can be significantly reduced if a number of bars that are oriented similarly to the preferred orientation of the cell are placed outside the receptive field (fig. 5). If the surrounding bars are oriented orthogonal to the central bar, the suppressive effect is reduced but still present.

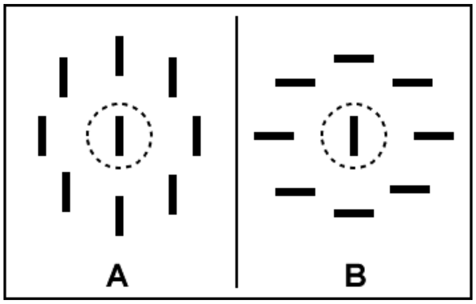


Figure 5: The response of a cell is suppressed more when (A) surrounded by similarly oriented features, than when (B) surronded by differently oriented ones (Knierim & van Essen, 1992).

Further, Kapadia et al. (1995) demonstrated that the suppression observed in a cell's response when a large number of randomly oriented bars are placed outside its receptive field can be considerably reduced, or even eliminated, if some of the surrounding bars are positioned along the cell's axis of orientation and are oriented in the same direction as the central bar (fig. 6).

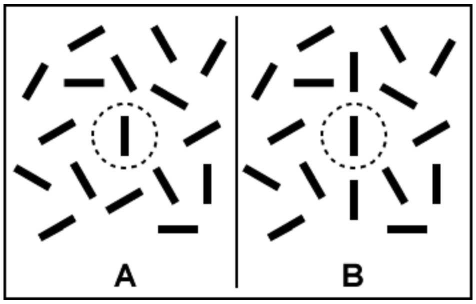


Figure 6: Kapadia et al. (1995) also observed a reduction in the suppression (caused by randomly oriented bars; A) in a cells response, if some of the bars were co-aligned with the orientation selectivity of the cell (B).

Similar findings have been reported in a number of psychophysical studies. The contrast detection threshold of a central low-contrast Gabor-patch can be increased or decreased depending on the position and orientation of surrounding Gabor-patches (Polat & Sagi, 1993, 1994) or gratings (Cannon & Fullenkamp, 1991). It has also been shown that a path of Gabor-patches, presented against a background of evenly distributed and randomly oriented patches, can be more easily detected when the relative angle between the adjacent elements in the path is less than +/-60**?** (Field et al., 1993), or the elements form a closed rather than open path (Kovacs & Julesz, 1993).

An interesting parallel, in this context, is how closely several of the above findings coincide with the Gestalt laws (Wertheimer, 1923; see also Rock & Palmer 1990) that were formulated to account for how we group low-level stimuli. Of particular note are the laws which postulate that we are perceptually biased to group together features that are arranged into smooth paths (good continuation), form closed curves (closure), and are close to each other (proximity); see fig. 7.

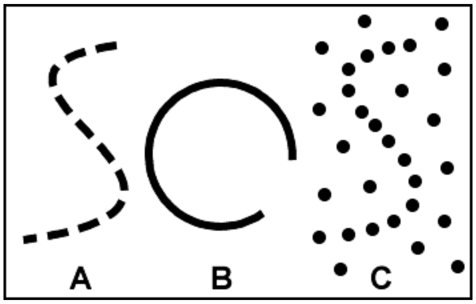


Figure 7: Illustration of the Gestalt (grouping) laws of (A) good continuation, (B) closure and (C) proximity.

It is not yet clear whether the effects (described above) arise within the striate cortex, or are produced by feedback connections from higher visual areas. Kapadia et al. (1995) have suggested that the long-range horizontal connections formed by pyramidal cells in the striate cortex could constitute the physiological substrate allowing spatial integration of information over several hypercolumns. These long-range connections enable the target cells to integrate information over regions well beyond the classical receptive field, but preferentially from cells having similar orientation tuning that are positioned along the target cells axis of orientation. However this may be, feedback connections from area V2 and other visual areas can not, of course, be ruled out. Nor can it be ruled out that different mechanisms may be responsible for different modulatory effects. Knierim and van Essen (1992) reported a time delay between the onset of the general (orientation independent) [7 ms] suppressive effect, and the orientation-dependent [18-20 ms] suppressive effect, which may indicate different origins.

6 POSSIBLE FUNCTIONAL SIGNIFICANCE

Kapadia et al. (1995) have suggested that the purpose of the selective enhancement in the firing rate of certain cells may be to make contours more salient, particularly when perceived against noisy and textured backgrounds. Given that the enhancement effect seems to be stronger for stimuli-configurations that consist of smoothly aligned features, and that occluding contours in general tend to produce such constellations in the retinal image, this interpretation seems highly plausible. The idea is appealing also because it is consistent with the computationally recognised need for mechanisms that can reduce the response-ambiguity of the simple/complex cells at an early stage of visual processing. Another interesting aspect of this interpretation is that, if correct, it might not only provide an explanation as to why we experience the Gestalt laws of good continuation, closure and proximity (i.e. to aid the identification of occluding contours), but it may also place the origin of these phenomena at a much earlier stage of visual processing than previously thought.

Regarding the suppressive effect, Knierim and van Essen (1992) have suggested that the observed difference in the suppression of a cell's response depending on the difference in orientation between the central and the surrounding stimuli may be important for texture segregation; and that it may be responsible for certain orientation-dependent pop-out phenomena such as our ability to quickly spot a single "V" embedded in a 2-D array of "T´s"; see for example Treisman and Gelade (1980). While basically agreeing that the suppression could be involved in both texture segregation and pop-out, a slightly different interpretation is here made on what the main functional significance of the suppression is. That is, we rather emphasise the possibility that the primary purpose of the orientation-dependent suppression -like possibly the corresponding enhancement effect- may be to aid the identification of occluding contours.

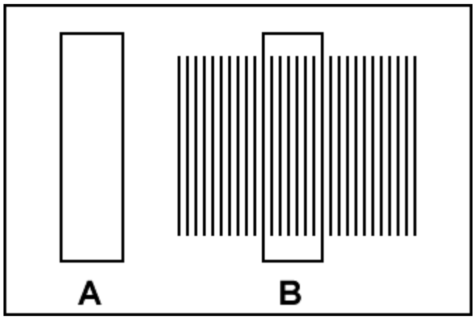


Figure 8: Example of contour "masking" (Modified after Kanizsa, 1979).

From the earlier discussion on the combinatorial explosion in the number of possible ways there are of grouping the responses from the simple/complex cells into contour paths, it is evident that our visual system somehow must constrain the grouping process. The observed enhancement in certain cells firing rates could be seen as such a constraint, as a way to guide higher-level integration processes. Letting a simple heuristic which prefers smoothly aligned features control the enhancement, seems a reasonable first approach to narrowing down the number of potential visual structures that may correspond to occluding contours.

However, because any heuristic by definition occasionally will be wrong, there needs to be an opposing, or complementary, mechanism that can balance or even override the effect of the enhancement. In many visual contexts, selective enhancement of co-aligned stimuli will not be a helpful strategy for identifying important occluding boundaries. Consider, for example, the fact that most surfaces in nature are heavily textured (e.g. fur, feathers, grass, leaves, rocks) and may produce regions with periodic or quasi-periodic patterns in the retinal image. Often, the components of such patterns consists of locally smoothly aligned features. Because such stimuli "fit the description" of co-alignment they would inappropriately be integrated into contours, unless some opposing system could counteract, or suppress, the integration mechanism.

Consider also a visual scene such as, a hungry lion lurking behind some high but possible-to-see-through grass; in such a context, a visual system would not serve its owner well if it enhanced every single straw of grass, but not the partially hidden outline of the lion. Clearly, not all occluding contours are equally important to us, but some deserves more attention than others. Preferably those that mark the peripheral boundaries of regions, objects and surfaces.

For these reasons, it seems that a more reliable representation of occluding boundaries, less "polluted" with nonsense contours, would be obtained if the integration of low-level stimuli into contours was suppressed within densely textured regions of the visual field; particularly if the features within such regions are periodically or quasi-periodically arranged, and are oriented in accordance with the axis along which the contour-integration is carried out.

Apart from computational considerations, ecological speculations, and the earlier reviewed physiological and psychophysical observations of suppressive effects, there is a rather compelling phenomenon referred to as contour masking (Kanizsa, 1979) which indicates that such a suppressive mechanism may control contour-integration in the human visual system. When the rectangle in figure 8a is viewed on its own, the vertical lines are clearly perceived as contours of the rectangle. However, when embedded into a texture such as in figure 8b, the vertical lines are no longer perceived as contours, but rather appear as if they are parts of a surface that seems to lie in front of the rectangle. What is perhaps the most interesting aspect of this phenomenon is that the experienced difference between the two viewing conditions seems to be entirely qualitative. That is, in fig 8b there seems to be no quantitative reduction in the perceived contrast of the lines, at least not large enough to cause the contours to vanish, but only a reduction in our inclination to perceive them as contours. This suggests that a higher-level representation of contours, or the process that integrates low-level stimuli into contours, is suppressed rather than the early representation of low-level stimuli per se. If this is the case, then the observed orientation-dependent suppression of V1 cells may be due to feedback connection from these higher visual areas where the contours are integrated/suppressed. This idea is consistent with the observed time delay between the onset of the general suppression and the orientation-dependent suppression in area V1 cells reported by Knierim and van Essen

(1992).

In the next section, a computational model based on these ideas is presented in which a layer of model contour neurones integrate oriented low-level stimuli according to a simple heuristic of co-alignment. To prevent "non-occluding" stimuli from being integrated into contours (i.e., being represented as contours), the model contour cells are suppressed depending on the magnitude and orientation of the stimuli in the near surrounds of their receptive fields. Although not intended as a quantitative description of the human visual system, the model is nevertheless consistent with several of the above described properties of both cell responses and psychophysical observations. Simulations with a computer implementation of the model do, for example, produce contour completion (e.g. illusory contours; Kanizsa, 1979), contour enhancement (i.e. increased saliency; Kapadia et al., 1995; Field et al., 1993), contour masking (Kanizsa, 1979) and orientation dependent pop-out (Treisman and Gelade, 1980), and it identifies occluding boundaries in natural images.

7 A COMPUTATIONAL MODEL

The model presented below is first and foremost a hypothetical functional model of how contours might be processed in the early stages of the human visual pathway. However, although function has been the main constraint, most design choices in the model architecture have been influenced by known response properties of various cell types and their inter-connections.

Figure 9 gives an overview of the model. On a coarse scale the model can be divided into two major levels of processing, roughly corresponding to the processing carried out by the simple and complex cells (Hubel & Wiesel, 1962) in area V1, and by the contour neurones/cells (Peterhans and von der Heydt 1993) in area V2.

At the first level, oriented contrast features in an input image are detected by a layer of model simple cells with anti-symmetric receptive fields (fig. 10a). To increase the spatial and orientation selectivity of these cells, all nearby simple cells (i.e. near in both the spatial and orientation domain) laterally inhibit each other. The responses from the simple cells are then fed into a layer of complex cells. Any given model complex cell pools the information from two simple cells that are separated by π rad (180°) in orientation tuning, and that are positioned at the same location in the visual/image field. Like the model simple cells, the model complex cells mutually suppress one another. However, the complex cells do so over a much larger distance than the simple cells (approximately 6 compared to 1 times the radius of a cell's receptive field) and they do so independently of orientation selectivity.

At the second level, a layer of model contour cells sum the outputs from the level 1 complex cells. The contour cells are also orientation selective, and any given cell only sums the outputs from complex cells with a particular orientation tuning. The receptive fields of these cells are (approximately 6 times) larger than the simple/complex cells, and are divided into two drop-shaped sub-receptive fields (fig 11). Only when there is sufficient activity from complex cells within both of a contour cells two half-fields does it become activated. To prevent stimuli within densely textured regions from being integrated into contours, all complex cells that respond to such stimuli are given a lesser weighting than those that respond to stimuli not embedded into textures. Finally, all nearby model contour cells with the same orientation selectivity inhibit each other along an axis orthogonal to the axis of their orientation selectivity.

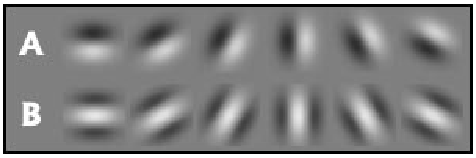


Figure 10: (A) Anti-symmetrical Gabor-filters used to detect oriented contrast features in the input image. (B) Symmetrical Gabor-filters. Not drawn to scale.

In the following sub-sections, a more thorough presentation of the various processing steps and their functional motivation is given. For technical and implementation details, the reader is directed to appendix A.

7.1 LEVEL 1

7.1.1 MODEL SIMPLE CELLS

The receptive fields of the simple cells are modelled with anti-symmetric Gabor-functions (i.e. the product of a sine and a Gaussian function). Symmetrical and anti-symmetrical elementary Gabor-signals (fig. 10) have been shown to correspond well with the receptive field-mappings of real simple cells (Marcelja, 1980). In order to capture contrast stimuli of different polarity and at all different orientations, 12 model cells each differing $\pi/6$ in orientation from the next are used to sample the image structure at each given position. The receptive field response is then half-rectified (i.e. negative values are ignored) and normalised for contrast by a divisive gain mechanism.

While cells with symmetrical and anti-symmetrical receptive fields respond optimally to different stimuli (i.e. a bar and edge respectively), each type of cell also responds to the optimal stimuli of the other type, although less so and at a slightly shifted position. From a computational point of view, using either kind of receptive field mapping is therefore sufficient to detect oriented contrast stimuli. In the computer implementation of the model, only one type of mapping was chosen to hold the computational cost down. The choice of anti-

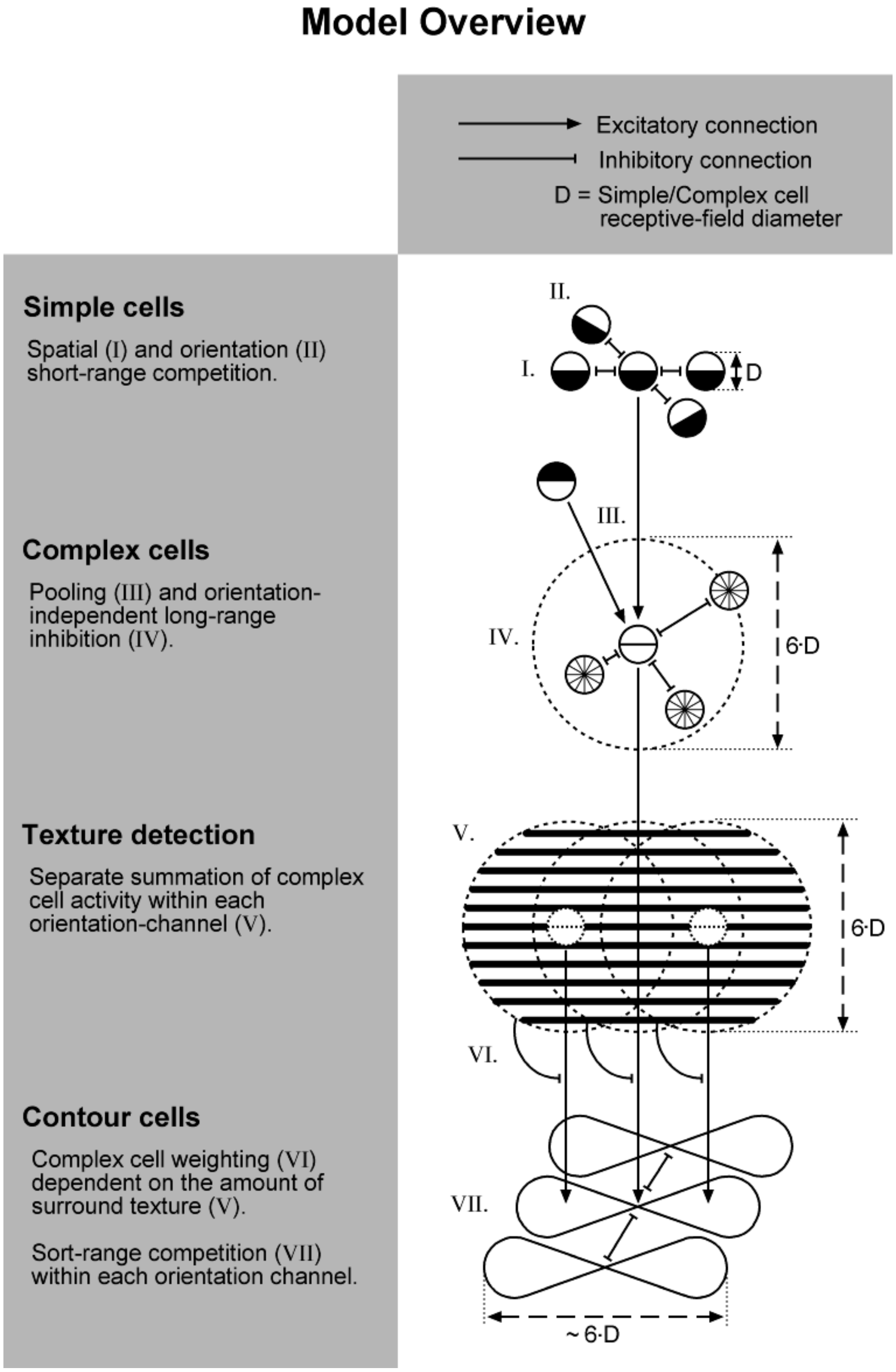


Figure 9: Model overview. (I) Spatial and (II) orientation short-range competition competition between similarly tuned simple cells. (III) Pooling of simple cell responses (by the complex cells). (IV) Mutual Long-range complex cell suppression (orientation independent). (V) "Texture" detection. (VI) Complex cell output weightied inversely proportional to the amount of texture-surround. (VII) Spatial integration of the complex cells' outputs along the common axis of orientation; and short-range (contour cell) competition.

symmetrical ones was arbitrary, except for the observation that some kind of contour completion phenomena seems to be stronger when the inducers are solid edges rather than thin lines (Kanizsa, 1979). However, in the human and primate brain, both cell types most likely contribute to the processing of form.

7.1.2 LATERAL INHIBITION

Because the model simple cells have partially overlapping receptive fields and because they are quite broadly tuned to orientation, any given stimulus will evoke activity in a number of cells nearby in both the spatial and orientation domain. Hence, the representation of an image will initially be somewhat blurred. In order to obtain a higher spatial and orientation acuity in the array of simple cells all near cells (i.e. near in either the spatial or the orientation domain, or both) laterally inhibit each other. Apart from sharpening the spatial and orientation selectivity of the model cells, this operation also has the effect of creating a relative activity enhancement in cells that respond to line-ends and corners, compared to those that respond to the interior parts of such stimuli (this mechanism is similar to the “end-cut” mechanism of Grossberg & Mingolla, 1985). In general, these “end-points” are the ones of interest for a contour completion mechanism (see also von der Heydt, 1995).

7.1.3 MODEL COMPLEX CELLS

While complex cells, like simple cells, are sensitive to stimuli having a particular orientation, many complex cells, unlike simple cells, fire independently of the polarity of a contrast stimulus (Livingstone et al., 1987). The intuitive observation that we easily can complete and integrate fragments differing in contrast into contours (see fig. 3a) suggests that the complex cells rather than the simple cells provide the main input to the neural mechanism responsible for form analysis.

In the current model, the responses of the complex cells are obtained by simply taking the absolute value of the difference between each two simple cells that share the same position and are separated in orientation selectivity by π rad (for a more sophisticated model of complex cell responses, see for example Heeger, 1991).

7.1.4 ORIENTATION-INDEPENDENT LONG-RANGE SUPPRESSION

Apart from what is present within its classical, or primary, receptive field, the response of a model complex cell is also determined by the degree of general activity within a larger region surrounding its receptive field. More precisely, the activity of the cell is suppressed proportionally to the squared and weighted sum over all orientations of the complex cell activity within a Gaussian envelope of approximately 6 times the radius of the complex cell.

The functional motivation for this suppression is twofold. First, it further enhances the activity of cells that are responding to edge- and line-ends, which are important for identifying texture borders and points were contours should be completed, or filled-in. Second, it creates an initial relative difference in the strength of activity in cells that respond to stimuli positioned at the periphery of textured, or otherwise crowded, regions, compared to those cells that are positioned at the interior of a texture-field. In general, such peripheral stimuli are statistically more likely to correspond to parts of surface/object borders.

Long-range suppression of a striate complex cells, induced by stimuli positioned outside the classical receptive field, have been observed in several studies (Knierim & van Essen, 1992; Kapadia et al., 1995; Grinvald et al., 1994).

7.2 LEVEL 2

7.2.1 MODEL CONTOUR CELLS

At the second level of processing, the suppressed outputs from the complex cells are integrated by a layer of model contour cells. Like the model simple and complex cells, the contour cells are selective to stimuli of a particular orientation. The contour cells, however, have considerably larger receptive fields, which allow them to integrate information from several complex cells along their axis of orientation. Another important feature of the model contour cells is that they are heavily suppressed by stimuli within textured regions, if the stimuli making up the texture are oriented similarly to the axis of orientation to which the cell is tuned.

7.2.2 SPATIAL INTEGRATION AND TEXTURE SUPPRESSION

The receptive field of each model contour cell is divided into two drop-shaped half-fields (fig. 11). Each half-field “hangs” down, along the axis of orientation-selectivity, from the centre of the receptive field, and reaches out to a distance of about 6 times that of the radius of the (primary) receptive field of a model simple/complex cell. Further, each sub-field separately sums the weighted outputs from all complex cells within its range that are selective to the same orientation as the contour cell. How the output from any particular complex cell is weighted is determined by two factors. First, the response is weighted by a factor that is determined by the spatial respectively angular distance from the contour cell’s centre, respectively, axis of orientation. The effect of this weighting is that only relatively co-aligned stimuli will become integrated. Second, the response of a complex cell is also weighted by an iso-orientation-measure,(see Appendix A for details) that is inversely proportional to the degree of activity of all other complex cells tuned to the same orientation within a larger region around the complex cell (6 times the diameter of a model complex cell’s primary receptive field). In other words, a model complex

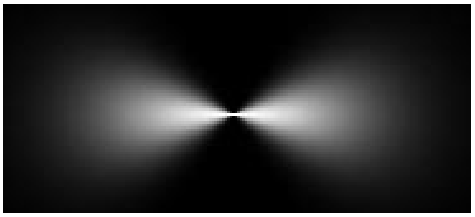


Figure 11: Receptive field of a model contour cell. Not drawn to scale.

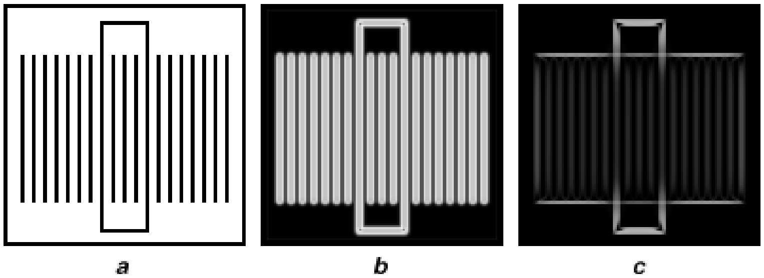


Figure 12: Example of contour “masking”. Redrawn from Kanizsa (1979). a) Input image. b) Initial “simple” cell activity. c) Output, i.e. “contour” cell activity.

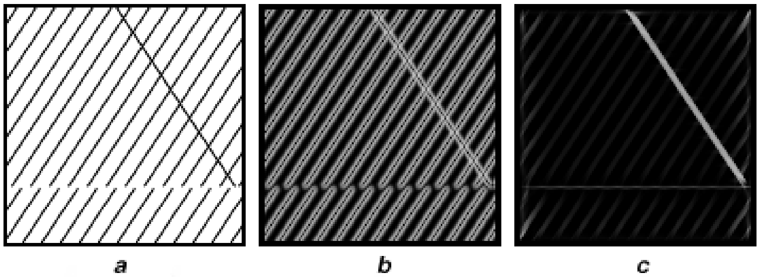


Figure 13: Partly “masked” and partly “illusory” triangle. Modified from von der Heydt (1995) who modified it from Galli and Zama (1931).

cell that responds to an isolated stimulus will be more heavily weighted than one that responds to a stimulus surrounded by other similarly oriented stimuli. Further, to prevent the contour-cells from becoming active at points in the image where there are no contours to fill-in or complete such as outside of corners and line-terminators, sufficient activity in both sub-receptive-fields is needed to make it respond. The contribution from each half-field is therefore integrated in a multiplicative fashion. The result of the integration is then passed through a threshold-function (an inverted Gaussian) that particularly compresses the higher response-interval, but also reduces the amount of noise in the lowest response-interval.

7.2.3 LATERAL INHIBITION

Because the contour-integration is performed on a relatively coarse scale (i.e. with relatively wide receptive fields), the positioning of the boundaries within the resulting representation will not be precise. Therefore, in order to better locate the spatial positions of the boundaries, all contour cells that are sensitive to the same orientation and lie near each other along an axis orthogonal to their axis of orientation-selectivity inhibit each others output.

8 SIMULATION RESULTS

The model presented in the previous section has been implemented as a computer program, and simulations have been run with images of both artificial and natural scenes. For all simulations presented here, the model parameters were set as described in appendix A, and all input-images were 128×128 pixels.

Apart from producing results that are consistent with humanly observed phenomena such as illusory contours (Kanizsa, 1979) and the Gestalt laws of good continuity, proximity and closure (Wertheimer, 1923; Rock & Palmer 1990), the model also reproduces various contour masking (Kanizsa, 1979) and orientation-dependent pop-out (Treisman and Gelade, 1980) phenomena. Further, some capacity for texture segregation has been observed, provided that the major components of the textures differ in their orientation by more than approximately 60°, or the textures have significantly different periodicities (densities).

In each of the examples below, the input image is depicted to the left, the initial model simple cell response in the middle and the model output to the right. High intensity in the middle and rightmost images corresponds to high activity in the model simple and model contour cells respectively. The intensity value at each point in these representations was obtained by pooling the activity in all orientation-channels (see appendix A, section 4). The simple cell representation is shown only for comparison. The images presented below are also available at: www.lucs.lu.se/people/jens.mansson/contours/index.html

8.1 ARTIFICIAL IMAGES

Contour Masking and Pop-Out Figure 12-14 shows examples of the contour masking effect, caused by the suppression of the contour-integration mechanism within regions containing densely positioned parallel

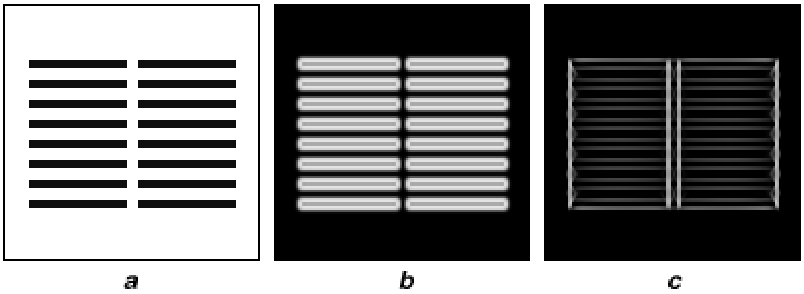


Figure 14: An illusory white bar in front of parallel horizontal lines.

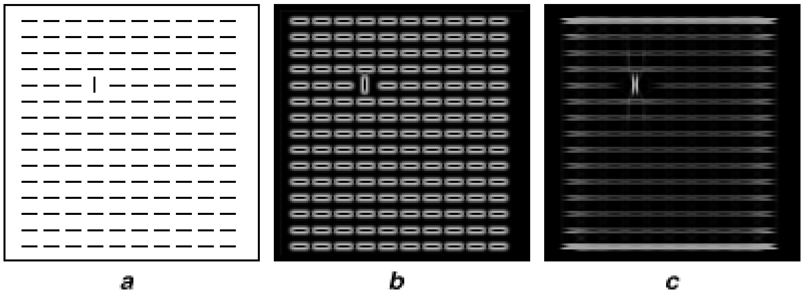


Figure 15: "Pop-out" of a single oriented line in an array of orthogonaly oriented lines. a) Attention is automatically drawn to the vertical bar. c) the output activity is stronger at the position of the vertical bar.

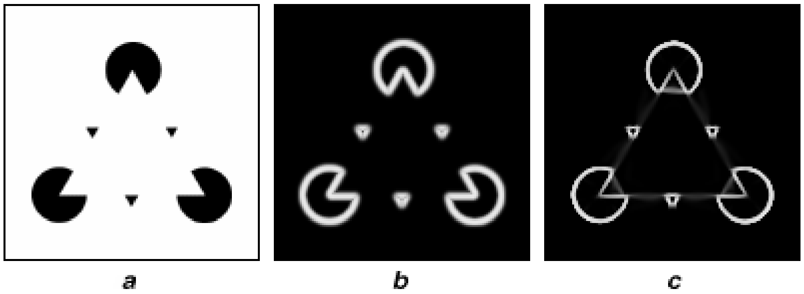


Figure 16: Kanizsa triangle. Modified from Kanizsa (1979). Note the illusory lines that have been formed between the black discs.

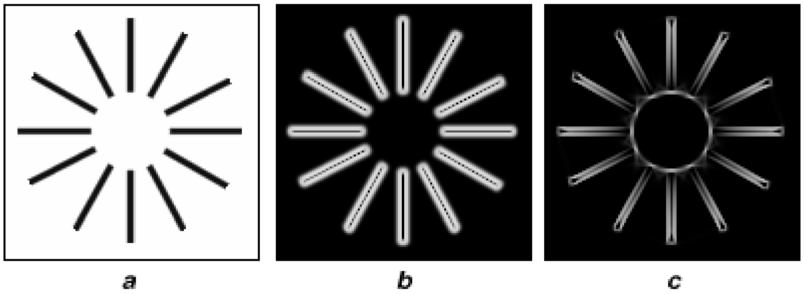


Figure 17: Illusory white disc covering the black radial lines.

lines, or other iso-oriented stimuli. Note also that in all of these three artificial images, illusory contours are formed at the ends of the lines, and the contour cell activity at these points is significantly higher (i.e. the illusory contours are more salient) than the activity along some of the actual intensity lines.

The suppressive mechanism that produces the above masking effect also makes a single oriented feature relatively more enhanced than the features in a surrounding array, if these are differently oriented (see figure 15). This could explain why attention is drawn to such parts of an image, and why the search for such stimuli is considerably faster than for stimuli that differ less, or in more than one of several possible aspects (orientation, colour, motion etc.) compared to surrounding features (see for example Treisman and Gelade, 1980).

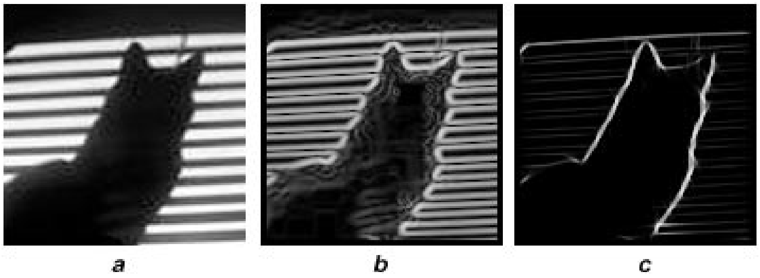


Figure 18: Shadow on a wall of a cat. The contour of the cat is both filled-in and enhanced, while the horizontal lines, on the wall in the background, are suppressed.

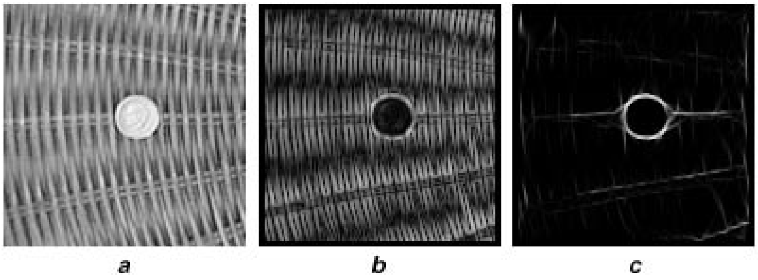


Figure 19: Coin on a table. Note how some of the horizontal lines, which barely are present in the simple cell representation, are strongly enhanced.

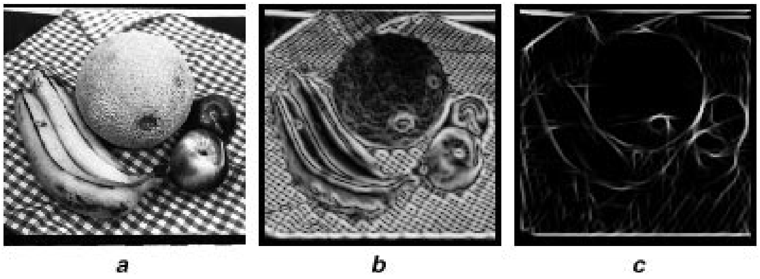


Figure 20: Compared to the simple cell representation (b), the contour representation (c) is much sparser, and almost entirely confined along the occluding contours of the fruit.

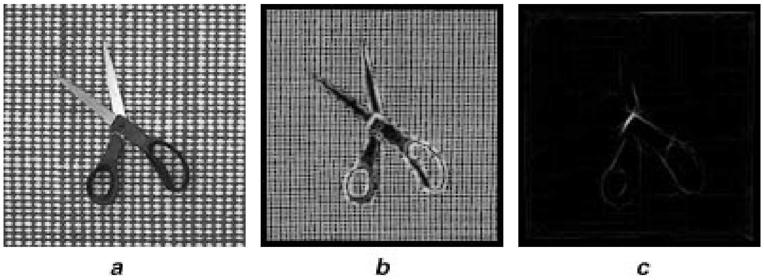


Figure 21: Pop-out of a pair of scissors on a carpet.

Illusory Contours As figure 16 and 17 show the model produces both straight and smoothly curved illusory contours at positions were human observers generally report perceiving these. These results are produced because a model contour cell integrates the information from the orientation-selective complex cells over relatively large region of the image; and hence can be activated even if there is no stimulus at the centre of its receptive field.

8.2 NATURAL IMAGES

The remaining examples demonstrate the model's performance on natural input. Figure 18-21 show how periodic textures are suppressed while leaving the majority of "real", or object, contours intact. Particularly note how the horizontal lines in figure 19c become relatively enhanced in the output representation, even though they are barely present in the initial simple cell representation (19b); and in figure 20, note how most of the con-

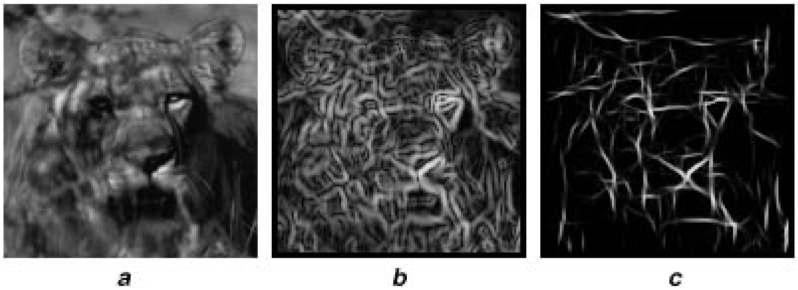


Figure 22: A lion resting in the shadow of a tree. In regions where there is model simple/complex cell activity in all different orientation channels (e.g. a lot of noise), most stimuli are suppressed.

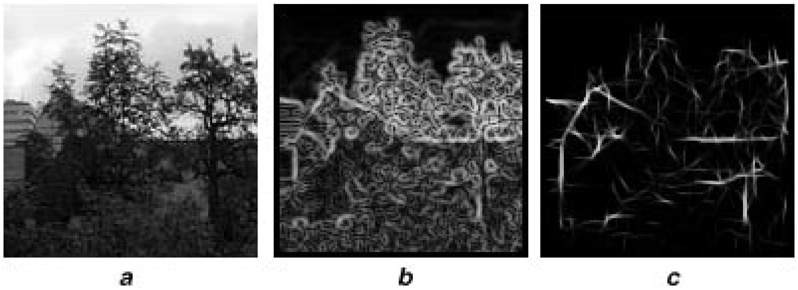


Figure 23: The skyline of a building behind some trees. Stimuli within crowded regions, such as the tree tops and bushes in front of the house, are suppressed in the output representation.

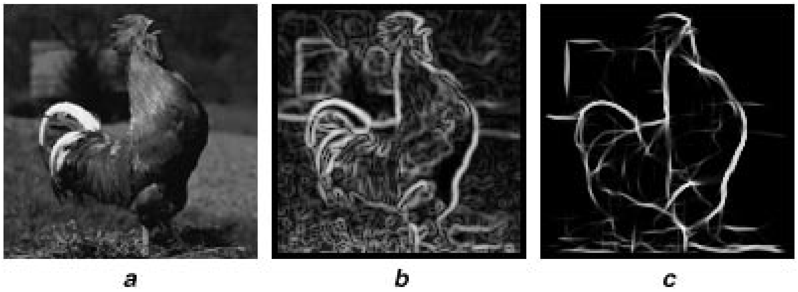


Figure 24: A rooster. Note how the majority of "false" contours in (b), caused by the feathers and grass, have been reduced in the final output representation (c).

tours of the fruit are left intact while the background table cloth pattern is suppressed. Also note in figure 22 and 23 how not only periodic iso-oriented textures are suppressed, but also random textures, or otherwise busy regions if there is sufficient activity in all orientation channels to drive the suppression mechanism.

9 DISCUSSION

9.1 RELATED WORK

The model presented in this paper shares several features with the models on contour perception suggested by Ullman (1976), Grossberg and Mingolla (1985), Gove et al. (1995), Heitger and von der Heydt (1993) and Yen and Finkel (1998). Although these models differ in various assumptions, for example in the proposed contour inducing elements, they all share the assumption that occluding contours, in general, produce relatively smoothly aligned features in an image. Hence, all the models locally constrain the spatial integration to features that are similarly oriented and relatively co-aligned. Due to this common feature, most of the models produce results that are more or less consistent with each other, and could account for why we perceive illusory contours and why we experience the Gestalt grouping laws of good continuity, closure and proximity.

However, in none of these models is the integration of low-level stimuli modulated by the contextual information available in the local surroundings, in the sense described in this paper. Therefore it seems unlikely that any of these models can account for phenomena like contour masking and orientation-dependent pop-out, considering that these phenomena seem to be highly context dependent. Further, given that textures often produce locally co-aligned visual stimuli, it is likely that these models will be relatively poor at discriminating such stimuli from actual occluding contours.

9.2 TEXTURE DISCRIMINATION

Because the model presented here is only intended as a functional model of how contour information might be processed early on in the human visual pathway, the individual processing steps described in the model can be only loosely mapped onto particular neurological structures. A particularly loose mapping is the one between the proposed contour-suppression-mechanism and a possible neural substrate that could implement it. In the current model, "textures" are crudely sensed by simply integrating the responses from a large number of model complex cells. From a computational point of view, this is most likely not the best procedure for detecting and discriminating between textures. An interesting question that therefore arises is what neural substrates other than the complex cells could provide information about texture.

One type, or category, of cells that seem particularly fit for this job are the "grating-cells" (von Heydt et al., 1992). These cells not only respond vigorously to gratings, but often fail altogether to respond to isolated bars or edges. Further, they are narrowly tuned to both orientation and spatial frequency, and have low contrast thresholds. According to von der Heydt et al. (1992), about 4% of all cell in area V1 and 1.6% of the cells in area V2 are of this type. An interesting property of these cells is that they not only respond to gratings of a particular frequency and orientation, but also to a number of other periodic, or quasi-periodic, patterns such as checkerboard patterns (when the diagonal rows are aligned with the preferred orientation of the cell), or patterns with "jittered" periodicity (e.g. lines separated by alternately small and large distances). von der Heydt et al. concluded that these cells do not perform a spatial-frequency analysis of the stimulus, but instead seem to be specialised for detecting periodic patterns. Considering the narrow tuning for both orientation and spatial frequency, as well as the low contrast-detection threshold, it clearly seems these cells are better fit than complex cells for performing discriminative texture detection. And hence, could provide more detailed/sophisticated information to a contour suppression mechanism. Whether this is the case will of course have to be shown in empirical studies.

9.3 POSSIBLE ROLE OF SPATIAL FREQUENCY

A final consideration, not yet either discussed nor modelled is the possible role the spatial frequency of the stimuli have on our perception of contours. Intuitively, it seems that the phenomena of contour masking (Kanizsa, 1975) can be reduced, or even eliminated, if the lines of the rectangle in figure 8 are made considerably thicker than the row of parallel lines (see fig. 25). This suggests that not only the periodicity and orientation of stimuli in the surround control contour-integration in the human visual system, but that the spatial frequency of the stimuli also control it. The output of the complex cells are maybe more suppressed (or less weighted by an integration mechanism) when the spatial frequency of surrounding stimuli is similar to the frequency that to which a cell is tuned.

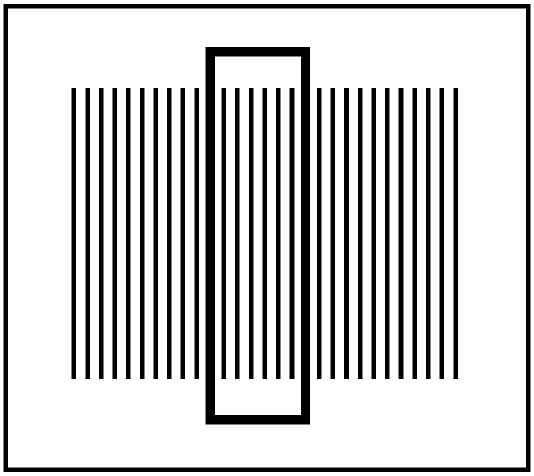


Figure 25: The contour masking effect is lost if the lines of the rectangle are made thicker .

10 SUMMARY

A computational model is proposed for how information on occluding contours might be processed in the early cortical visual areas (roughly V1 and V2). A central subsystem in the model is a mechanism which suppresses the integration of oriented low-level stimuli into contours, if these stimuli are embedded into a texture composed of similarly oriented stimuli/features. This operation is motivated by the fact that features in natural scenes which are situated inside patterned regions are more likely (from a statistical point of view) to have been produced by surface textures, than they are likely to have arisen due to occluding structures. A computer implementation of the model demonstrates results consistent with the percepts that are reported by human observers. The model does, for example, fill-in missing segments of contours (i.e., produce illusory contours; Kanizsa, 1979) and enhances weak ones (i.e., increase the saliency; Field et al., 1993). Further it reproduces the phenomena of contour masking (Kanizsa, 1979) and certain orientation-dependent pop-out effects (Treisman and Gelade, 1980). It also works well on natural images where noise and, particularly, ambiguous stimuli may present problems to models that do not consider the contextual information available in the local surround.

A TECHNICAL SPECIFICATION

A.1 SIMPLE CELLS

A.1.1 RECEPTIVE FIELDS

The receptive fields of the simple cells were modelled with 12 rotated copies each separated $\pi/6$ rad from the next $[\theta = n \cdot \pi/6 : n = 1 \ldots 12]$, of the following anti-symmetric Gabor-function:

$$\begin{aligned} G_{edge}(x,y) &= \sin(2\pi f[x - x_c])e^{\frac{1}{4}(\frac{x^2}{\sigma_{x1}^2} + \frac{y^2}{\sigma_{y1}^2})}, \\ f &= \frac{1}{2R}, \, \sigma_{x1} = \frac{R}{4}, \, \sigma_{y1} = \frac{R}{3.3} \end{aligned} \tag{1}$$

where f is the frequency, R is the radius (3.5 pixels in the implementation) of the cells receptive field, σ_{x1} and σ_{y1} are space constants and x_c the centre of the receptive field.

A.1.2 NORMALISATION AND HALF-RECTIFICATION

The response from a simple cell, positioned at (x,y) and tuned to orientation θ, is obtained by convolving the raw image, I, with the corresponding (Gabor) mask G_{edge}^{θ}. The result is then normalised for contrast and half-rectified.

$$S_{norm}^{\theta}(x,y) = \left\lfloor \frac{\sum_{s=x-R}^{x+R} \sum_{t=y-R}^{y+R} I(s,t) G_{edge}^{\theta}(s,t)}{\kappa + \sum_{s=x-R}^{x+R} \sum_{t=y-R}^{y+R} I(s,t) G_{edge}^{\theta}(s,t)} \right\rfloor, \tag{2}$$

$$\kappa = 0.02 I_{max} R^2$$

Here, κ is a threshold constant, which is determined by R (same as above) and the maximum possible intensity value, I_{max} , in the image representation (e.g. 256 for an 8 bit grey-scale coding). The floor-brackets denotes half-rectification.

A.1.3 LATERAL INHIBITION

The inhibited output of a model simple cell, $S_{inhib}^{\theta}(\ldots, \theta, x, y)$, tuned to orientation θ and positioned at (x,y), is given by:

$$S_{inhib}^{\theta}(S_{norm}, \theta, x, y) = \tag{3}$$
$$1 - e^{-\frac{1}{2}\left\lfloor \frac{A(S_{norm}, \theta, x, y)^2}{s_A^2} + \frac{B(S_{norm}, \theta, x, y)^2}{s_B^2} \right\rfloor}$$

The terms $A()$ and $B()$ provide the contribution from the orientation- and spatial-dependent inhibition respectively. s_A and s_B are saturation constants that control the contribution of $A()$ and $B()$ respectively (see below).

$$\begin{aligned} &A(S_{norm}, \theta, x, y) = \\ &\quad \sum_{n=-\Omega}^{\Omega} G_1(n \cdot \alpha, \sigma_o) \cdot \left[S_{norm}^{\theta}(x,y) - S_{norm}^{\theta + n \cdot \alpha}(x,y) \right], \\ &\quad n \neq 0, \end{aligned} \tag{4}$$

$$\Omega = 3, \, \alpha = \frac{\pi}{6}, \, \sigma_o = \frac{2\pi R}{3},$$

$$S_A = 0.2 \sum_{n=-\Omega}^{\Omega} G_1(n \cdot \alpha, \sigma_o), \; n \neq 0$$

Ω determines the angular range of the inhibition, and α is the minimum angular separation between two differently tuned cells. so σ_0 is a space constant that determines the shape of the Gaussian envelope provided by the function $G_1(r, \sigma)$ (this determines how much neighbouring cells contribute to the inhibition; see eq. 6).

$$\begin{aligned} &B(S_{norm}, \theta, x, y) = \\ &\quad \sum_{i=-R}^{R} \sum_{j=-R}^{R} G_1(\sqrt{i^2 + j^2}, \sigma_s) \cdot \\ &\quad \left(S_{norm}^{\theta}(x,y) - S_{norm}^{\theta}(x+i, y+j) \right), \\ &\quad i \neq j \neq 0 \end{aligned} \tag{5}$$

$$G_1(r, \sigma) = e^{-\frac{r^2}{\sigma^2}}, \, \sigma_s = \frac{R}{2}, \tag{6}$$

$$S_B = 0.2 \sum_{i=-R}^{R} \sum_{j=-R}^{R} G_1(\sqrt{i^2 + j^2}, \sigma_s), \; i \neq j \neq 0$$

Again, σ_s is a space constant that determines how fast the Gaussian envelope (eq. 6) falls off.

A.2 COMPLEX CELLS

A.2.1 POOLING

The initial complex cell response, $C^{\theta}(x,y)$, for a cell tuned to orientation θ:

$$C^{\theta}(x,y) = \left| S_{inhib}^{\theta}(x,y) - S_{inhib}^{\theta + \pi}(x,y) \right| \tag{7}$$

A.2.2 ORIENTATION-INDEPENDENT LONG-RANGE SUPPRESSION

The activity in a model complex cell after long-range suppression,$C_{lrs}^{\theta}(.,.,x,y)$, is given by:

$$C^{\theta}_{lrs}(C,\theta,x,y) = C^{\theta}(x,y)\cdot \tag{8}$$

$$\left[1-\frac{T}{2}\left(\sum_{i=-W}^{W}\sum_{j=-W}^{W} G_2(\sqrt{i^2+j^2},\sigma_{c1},\sigma_{c2})\cdot\right.\right.$$

$$\left.\left.\sum_{n=0}^{5}\left[C^{n\pi/6}(x+i,y+j)\right]^2, s_c\right)\right],$$

$$\text{if } \sqrt{i^2+j^2} \leq W$$

else

$$C^{\theta}_{lrs}(C,\theta,x,y) = 0$$

$$W = 6R, \sigma_{c1} = 4R, \sigma_{c2} = \frac{r}{1.2}, s_c = 16R$$

$$G_2(r,\sigma_{c1},\sigma_{c2}) = e^{\frac{-r^2}{\sigma_{c1}^2}} - e^{\frac{-r^2}{\sigma_{c2}^2}} \tag{9}$$

$$T(x,s) = 1 - e^{-\frac{x^2}{s^2}} \tag{10}$$

Complex cells beyond the distance W do not contribute to the suppression. The inner (rightmost) summation in equation 8 is summation over all orientations. The purpose of the squaring is to preferentially let high-contrast stimuli contribute to the suppression. The weight function, G_2, is the difference between two Gaussians with space constants σ_{c1} and σ_{c2}. The latter creates an inner region approximately the size of the receptive field of a model complex cell, with near zero values so that a given cell does not suppress itself. s_c is a saturation parameter for the threshold-function $T()$ (eq. 10).

A.3 CONTOUR CELLS

A.3.1 SUB-RECEPTIVE FIELDS

Depending on the distance, r, (respectively, the angular deviation, α-θ) from a contour cell's receptive-field centre (respectively axis of orientation), the two sub-receptive-fields $\underset{\leftarrow}{F}^{\theta}$ and$\underset{\rightarrow}{F}^{\theta}$ of a contour cell (tuned to orientation θ) weights the outputs from all complex cells tuned to orientation θ according to (borrowed from Heitger and von der Heydt, 1993):

$$\underset{\rightarrow}{F}^{\theta}(r,\alpha,\theta,\sigma_f) = \cos^{2n}(\alpha-\theta)\cdot e^{\frac{-r^2}{2\sigma_f^2}} \tag{11}$$

$$\text{if } -\pi/4 < \alpha-\theta < \pi/4,$$

else

$$\underset{\rightarrow}{F}^{\theta}(r,\alpha,\theta,\sigma_f) = 0$$

$$n = 4, \sigma_f = 3R,$$

$$\underset{\leftarrow}{F}^{\theta}(r,\alpha,\theta,\sigma_f) = \underset{\rightarrow}{F}^{\theta}(r,\alpha,\theta+\pi,\sigma_f)$$

A.3.2 ISO-ORIENTATED STIMULI DENSITY

The function $\tau^{\theta}_{iso}(x,y)$ is a measure of the amount of stimuli (with orientation θ) present within a region of radius W, centred at (x,y).

$$\tau^{\theta}_{iso}(C^{\theta}_{lrs},x,y) = \tag{12}$$

$$T\left(\sum_{i=-W}^{W}\sum_{j=-W}^{W} G_2(\sqrt{i^2+j^2},\sigma_{iso1},\sigma_{iso2})\cdot\right.$$

$$\left.T(C^{\theta}_{lrs}(x+i,y+j),s_{c2}),s_{iso}\right)$$

$$\text{if } \sqrt{i^2+j^2} \leq W,$$

else

$$\tau^{\theta}_{iso}(C^{\theta}_{lrs},x,y) = 0$$

$$\sigma_{iso1} = 4R, \sigma_{iso2} = \frac{R}{1.2}, s_{c2} = 0.15,$$

$$s_{iso} = 0.15\sum_{i=-W}^{W}\sum_{j=-W}^{W} G_2(\sqrt{i^2+j^2},\sigma_{iso1},\sigma_{iso2}$$

The purpose of letting the the complex cell response, $C^{\theta}_{lrs}(x,y)$, first pass through the (inner) threshold function, T, (in eq. 12) with the low saturation constant s_2, is to enhance weak responses and thereby emphasise the orientation of the stimuli and not the contrast-intensity. The function G_2 (see e.q 9) with the space constants s_{iso1} and s_{iso2}, determines how a stimuli at distance $\sqrt{i^2+j^2}$ from point (x,y) is weighted. s_c and s_{iso} are saturation constants for the threshold-function T (eq. 10).

A.3.3 SPATIAL INTEGRATION AND TEXTURE SUPPRESSION

The summed activity, $\underset{\rightarrow}{K}^{\theta}$, within a contour cell's sub-receptive field, $\underset{\rightarrow}{F}^{\theta}$, is given by:

$$\underset{\rightarrow}{K}^{\theta}(C^{\theta}_{lrs},\tau^{\theta}_{iso},x,y,\theta,\sigma_f) = \tag{13}$$

$$\sum_{i=-L}^{L}\sum_{j=-L}^{L} \underset{\rightarrow}{F}^{\theta}\left(\sqrt{i^2+j^2}, \arctan(\frac{j}{i})-\theta,\sigma_f\right)\cdot$$

$$C^{\theta}_{lrs}(x+i,y+j)e^{-k\tau^{\theta}_{iso}(x+i,y+j)}$$

$$\text{if } -\pi/4 < \arctan(\frac{j}{i})-\theta < \pi/4,$$

else

$$\underset{\rightarrow}{K}^{\theta}(C^{\theta}_{lrs},\tau^{\theta}_{iso},x,y,\theta,\sigma_f) = 0$$

$$L = 5R, k = 2.2, \sigma_f = 3R$$

Note in eq. 13 how both the sub-receptive field ($\underset{\rightarrow}{F}^{\theta}$) and the stimuli-density measure (τ^{θ}_{iso}) together determine how any given complex cell is weighted.

The combined responses, K^{θ}_{comb}, from the half-fields $\underset{\leftarrow}{K}^{\theta}$ and $\underset{\rightarrow}{K}^{\theta}$ is:

$$K^{\theta}_{comb}(x,y) = T\left(\sqrt{\underset{\leftarrow}{K}^{\theta} \cdot \underset{\rightarrow}{K}^{\theta}}, \sigma_K\right), \sigma_K = \frac{1}{3} \qquad (14)$$

T is the threshold-function (eq. 10), and σ_K is a constant that determines how early the threshold-function saturates (i.e. reaches its maximum value).

A.3.4 LATERAL INHIBITION

The final contour representation K^{θ} is obtained by convolving the combined representation K^{θ}_{comb} with a symmetric Gabor-function $G^{\theta}_{bar}(x,y)$ followed by half-rectification and filtering through the threshold-function T (eq. 10).

$$K^{\theta}(x,y) = T\left(\left\lfloor \sum_{i=x-R}^{x+R} \sum_{j=y-R}^{y+R} K^{\theta}_{comb}(i,j) G^{\theta}_{bar}(i,j) \right\rfloor, \right. \qquad (15)$$
$$\left. 0.1 \cdot \sum_{i=x-R}^{x+R} \sum_{j=y-R}^{y+R} \lfloor G^{\theta}_{bar}(i,j) \rfloor \right)$$

$$G_{bar}(x,y) = \cos(2\pi f_2 [x - x_c]) e^{-\left(\frac{x^2}{\sigma^2_{x2}} + \frac{y^2}{\sigma^2_{y2}}\right)}, \qquad (16)$$

$$f_2 = 0.75R, \sigma_{x2} = \frac{R}{3.5}, \sigma y2 = \frac{r}{3}$$

A.4 SIMULATION OUTPUT-REPRESENTATION

All output images (I_{out}) presented in section 8 were obtained by pooling the activity in all 6 orientation channels $[\theta = 0, \pi/6, 2\pi/6, ..., 5\pi/6]$ as shown below. Here, R^{θ} is the initial simple cell, or the final contour cell, representation.

$$I_{out}(x,y) = 1 - e^{-\sum_{n=0}^{5} R^{n\pi/6}(x,y)} \qquad (17)$$

REFERENCES

Cannon, M.W., Fullenkamp, S.C. 1991, “Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations.” *Vision Research, Vol. 31, No 11,* pp 1985-1998

Field, D.J., Hayes, A., Hess, R. F. 1993, “Contour integration by the human visual system: Evidence for a local “Association field”.” *Vision Research, Vol. 33, No 2,* pp 173-193

Galli, A., Zama, A., 1931, “Beiträge zur Theorie der Wahrnehmung.” *Z. Psychol.* 123:308-348

Gilbert, C., Wiesel, T. N., 1990, “The influence of contextual stimuli on the orientation of cells in the primary visual cortex of the cat.” *Vision Research, Vol. 30, No 11,* pp 1689-1701

Gove, A., Grossberg, S., Mingolla, E., 1995, “Brightness perception, illusory contours, and corticogeniculate feedback.” *Visual neuroscience, 12,* pp 1027-1052

Grinvald, A., Edmund, E., Frostig, R. D., Hildesheim, R. 1994, “Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of Macaque monkey primary visual cortex.” *The Journal of Neuroscience,* May, 14(5): 2545-2568

Grossberg, S., Mingolla, E., 1985, “Neural dynamics of form perception: Boundary completion, Illusory contours and Neon color spreading.” *Psychological Review, Vol. 92, No. 2,* 173-211

Heeger, D.J., 1991, ”Nonlinear model of neural responses in cat visual cortex”, In *Computational Models of Visual Processing.* eds. Landy M.S., Movshon J. A., MIT Press, Cambridge

Heitger, F., von der Heydt, R., 1993, “A computational model of neural contour processing: Figure-ground segregation and illusory contours”, In *Proceedings of the Fourth Int. Conf. On Computer Vision.* IEEE Computer Society Press

Hubel, D. H., Wiesel, T. N., 1962, “Receptive fields, binocuar interaction and functional architecture in the cats visual cortex.” *Journal of Physiology, 160,* pp 106-154.

Hubel, D. H. 1988, *Eye, brain and vision.* Scientific American Library.

Kandel, E. R., 1991, “Perception of motion, depth, and form.” In Kandel, E. R., Schwartz, J. H., Jessel, J. J. (eds.), 1991, *Principles of neural science.* Prentice-Hall International, London, UK.

Kanizsa, G. 1979, *Organization in vision: Essays on Gestalt perception.* Praeger Publishers, New York

Kapadia, M.K., Ito, M., Gilbert, C. D. 1995, “Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and V1 of alert monkeys.” *Neuron, Vol. 15,* pp 843-856

Knierim, J.J., van Essen, D.C. 1992, “Neuronal responses to static texture patterns in area V1 of the alert Macaque monkey.” *Journal of Neurophysiology, Vol. 67, No 4*

Kovacs, I., Julesz, B. 1993, “A closed curve is much more than an incomplete one. Effect of closure in figure-ground segmentation.” *Proc. Natl. Acad. Sci. USA, Vol. 90,* pp 7495-7497

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 (11),* pp 3416-3468

Marcelja, S., 1980, “Mathematical description of the responses of simple cells.” *J. Opt. Soc. Am., Vol. 70, no. 11*

Marr, D. 1982, *Vision.* W. H. Freeman and Company, New York

Peterhans, E., von der Heydt, R. 1993, “Functional organization of area V2 in the alert Macaque.” *European Journal of Neuroscience, Vol. 5*, pp. 509-524

Polat, U., Sagi, D. 1993, “Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments.” *Vision Research, Vol. 33, No 7*, pp 993-999

Polat, U., Sagi, D. 1994, “The architecture of perceptual spatial interactions.” *Vision Research, Vol. 34, No 1*, pp 73-78

Rock, I., Palmer, S., 1990, “The legacy of Gestalt psychology.” *Scientific American*, December

Treisman, A. M., Gelade, G., 1980, “A feature-integration theory of attention.” *Cognitive Psychology*, 12:97-136

Ullman, S., 1976, “Filling-in the gaps: The shape of subjective contours and a model for their generation.” *Biological Cybernetics, 25*, 1-6

von der Heydt, R. & Peterhans, E., 1989, “Mechanisms of contour perception in monkey visual cortex: I. Lines of pattern discontinuity.” *J. Neurosci.* 9:1731-1748

von der Heydt, R., Peterhans, E., Dursteler, M. R., 1992, “Periodic-pattern-selective cells in monkey visual cortex.” *The Journal of Neuroscience, April 1992, 12(4)*, pp 1416-1434

von der Heydt, R., 1995, “Form analysis in visual cortex.” In Gazzaniga, M. S. (ed.) *The Cognitive neurosciences*, Cambridge, Ma:MIT Press. pp 383-400

Wertheimer, M., 1923, “Untersuchungen zur Lehre von der Gestalt”. II. *Psychologische Forschung* 4:301-350.

Yen, S-C., Finkel, L. H., 1998, “Extraction of peceptually salient contours by striate cortical networks”. *Vision Research, Vol. 38 , No 5*, pp 719-741