

# Surface Construction by a 2-D Differentiation–Integration Process: A Neurocomputational Model for Perceived Border Ownership, Depth, and Lightness in Kanizsa Figures

Naoki Kogo  
Katholieke Universiteit Leuven

Christoph Strecha  
EPFL/IC/ISIM/CVLab and Katholieke Universiteit Leuven

Luc Van Gool  
Katholieke Universiteit Leuven and ETH,  
Computer Vision Laboratory (BIWI)

Johan Wagemans  
Katholieke Universiteit Leuven

Human visual perception is a fundamentally relational process: Lightness perception depends on luminance ratios, and depth perception depends on occlusion (difference of depth) cues. Neurons in low-level visual cortex are sensitive to the difference (but not the value itself) of signals, and these differences have to be used to reconstruct the input. This process can be regarded as a 2-dimensional differentiation and integration process: First, differentiated signals for depth and lightness are created at an earlier stage of visual processing and then 2-dimensionally integrated at a later stage to construct surfaces. The subjective filling in of physically missing parts of input images (completion) can be explained as a property that emerges from this surface construction process. This approach is implemented in a computational model, called DISC (Differentiation–Integration for Surface Completion). In the DISC model, border ownership (the depth order at borderlines) is computed based on local occlusion cues (L- and T-junctions) and the distribution of borderlines. Two-dimensional integration is then applied to construct surfaces in the depth domain, and lightness values are in turn modified by these depth measurements. Illusory percepts emerge through the surface-construction process with the development of illusory border ownership and through the interaction between depth and lightness perception. The DISC model not only produces a central surface with the correctly modified lightness values of the original Kanizsa figure but also responds to variations of this figure such that it can distinguish between illusory and nonillusory configurations in a manner that is consistent with human perception.

*Keywords:* illusory contours, surface completion, depth perception, lightness/brightness perception, neural computation

When there is a part missing in an image, our visual system is often capable of filling in the gap immediately. How does this completion phenomenon work? The conventional explanation is *contour completion and filling in* (see Pessoa, Thompson, & Noë, 1998, for review). However, as we point out later, this

concept needs to be articulated further. We propose that the completion phenomenon can be viewed in terms of a surface construction process that is based on the two-dimensional (2-D) integration of “difference” signals. After introducing this article’s principal ideas, we report the implementation of this new approach in an elaborate and detailed model, called DISC (Differentiation–Integration for Surface Completion), which computes depth, lightness, and brightness and their interactions. We also report the results of testing the DISC model on the Kanizsa figure, the prototypical example of completion phenomena, and on a wide range of its variations. We show how the DISC model can lead to the emergence of completion and discuss the advantages of this approach in comparison with the contour-completion-and-filling-in approach. Finally, we discuss the potential application of the principles and the model to a broader range of perceptual phenomena.

Before introducing the model however, the next section provides a more general overview of the nature of the completion phenomena we explain here.

## How Do We “Complete” in Vision?

The fact that physically missing parts of images are filled in human vision is called *completion*. It has been assumed commonly

---

Naoki Kogo, Laboratory of Experimental Psychology, Katholieke Universiteit Leuven, Leuven, Belgium, and ESAT-PSI, Katholieke Universiteit Leuven, Heverlee, Belgium; Christoph Strecha, EPFL/IC/ISIM/CVLab, Lausanne, Switzerland, and ESAT-PSI, Katholieke Universiteit Leuven; Luc Van Gool, ESAT-PSI, Katholieke Universiteit Leuven and ETH, Computer Vision Laboratory (BIWI), Zürich, Switzerland; Johan Wagemans, Laboratory of Experimental Psychology, Katholieke Universiteit Leuven.

Naoki Kogo was supported by Fellowships F/01/043, F/02/104, DB/04/047, and DB/06/008 from the Research Fund of the University of Leuven. This research was supported by Research Fund of the University of Leuven Grant GOA 2005/03, European Science Foundation CNCC-Boundaries Grant G.0688.06, and Methusalem program Grant METH/08/02 to Johan Wagemans. We would like to thank Lee de Wit for improving the quality of the writing.

Correspondence concerning this article should be addressed to Johan Wagemans, University of Leuven, Laboratory of Experimental Psychology, Tiensestraat 102, bus 3711, B-3000 Leuven, Belgium. E-mail: johan.wagemans@psy.kuleuven.be

that the mechanism underlying this phenomenon is the interpolation of the missing parts of the contours, followed by the filling in of the contents of the enclosed areas (see Pessoa et al., 1998, for review). Imagine that we are looking at a picture in which one part is damaged. When we see that a part is missing in an image, it seems natural to assume that we make the interpolation of the contours first and then fill in the information of the enclosed areas such as colors and textures later. Indeed, in an art class, where we are challenged to draw a giraffe, for example, it is natural for us to first draw the lines that show the shape and then fill in the interior. The view of contour completion and filling in is well accepted, probably because of its analogy to this tendency of our thought/drawing process. However, this concept has to be treated with caution, especially when the underlying neural mechanisms for completion are investigated.

Before going further in this discussion, the terms *contour*, *borderline*, and *edge* need to be clarified first. According to *The Concise Oxford Dictionary of Current English* (8th ed.; Allen, 1990), *borderline* is “a line marking a boundary” (p. 128) and *contour* is “an outline, esp. representing or bounding the shape or form of something” (p. 249). In other words, a contour is (a) the borderline of (b) a shape or form. Thus, for a borderline to be called a contour, it has to satisfy both conditions. When there is a sudden change of surface properties such as color or texture, that location can be represented by a borderline. Once the borderline is used to represent the shape of the surface on one side of the borderline, the borderline can be called the contour of the surface. These two terms, however, tend to be used without being cautious about their difference. It should be noted that the concept of contour is thus related to the edge of the surface. Note, however, that contour lines or borderlines do not physically exist in the image: There are no actual drawn lines at the location. Lines are an artificial method of representation developed to satisfy a need of schematically describing a shape or pointing out the location where two surfaces are clearly separated. Edges, on the other hand, are physical parts of objects. (The existence of an edge also indicates the existence of a figure, as contours do.) Another important difference between contours (and borderlines) and edges is that the term *edge* incorporates relatively smooth as well as sharp step-wise endings of objects. Contours or borderlines, on the other hand, are bound to describe only surfaces with step-wise changes that can be represented by lines.

Once a region is marked as *figure*, the region gains ownership of the borderline. The term “border ownership” (BOWN) is closely related to “contour” in the sense that BOWN indicates that the borderline is a part of that figure, and the contour represents the shape of that figure. The concept “border ownership” makes the distinction with “borderline” explicit (borderline and its ownership). As shown later, the computation of the BOWN in fact plays the key role in the DISC model.

If we return to our example in the art class, we can note that when we draw the line as a contour, it is likely that we are able to do so because we are already aware of the surfaces to be enclosed by the lines. In other words, these lines are actually not mere lines, but they are used to represent the enclosed surface (the figure) that is already in our mind. On the other hand, when lines are drawn where parts are missing simply because of the discontinuity, this should be called *borderline completion*. In terms of the mechanisms, this difference is essential. The question is, when there is a

part missing that needs to be completed, does our visual system also draw lines first, and does it then fill in and become aware of the surface as a result? Or does it become aware of the figure first, such that the perception of these subjective contours merely reflects the perception of this subjective surface? In terms of neural mechanisms, the question is translated to whether the neurons complete the borderlines first wherever missing parts are found and then construct the surfaces within, or do they somehow reflect the global configuration, such as the existence of surfaces, first and start the completion process accordingly. This distinction plays a key role for a computational model to be able to distinguish illusory and nonillusory variations of the Kanizsa figure.

Before discussing these issues further, we examine the nature of completion phenomena in the context of the representative and extensively studied Kanizsa figure and its variations in the next section. We then review previous computational models that detect the collinear alignments of contours in order to complete the gaps in Kanizsa figures. We point out, however, that this approach has its own difficulties. We indicate that another cue (i.e., occlusion) plays an essential role in causing the Kanizsa illusions. We then explain why occlusion cues should be considered as 2-D differentiated signals and why, to create the depth map, it is most natural to apply 2-D integration to the signals. We show that this view leads to a surface construction model within the differentiation–integration framework. We end this first section with a brief discussion of the further advantages and plausibility of this approach.

### Clues From the Variations of the Kanizsa Figure

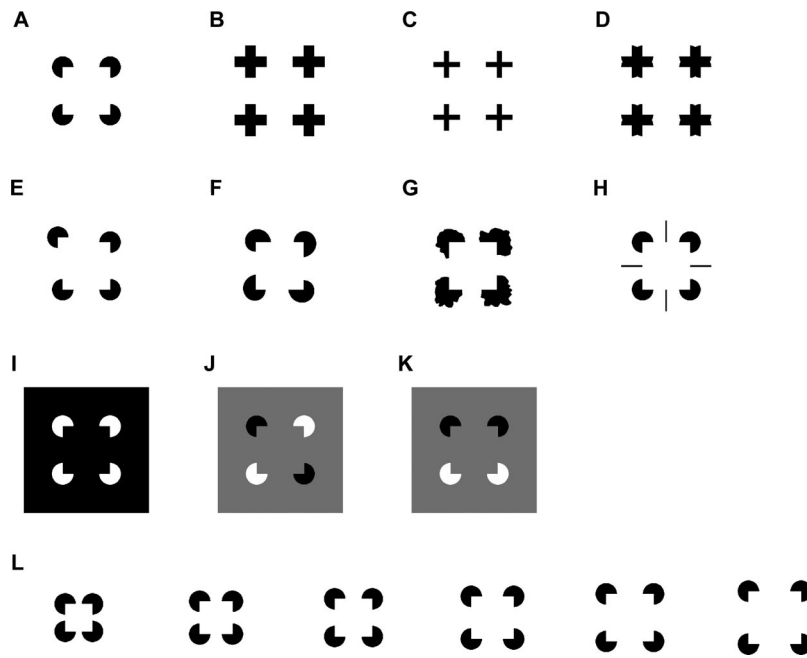
The illusory percepts invoked by the Kanizsa figure were first reported more than 50 years ago (Kanizsa, 1955). Since this time, the Kanizsa figure has been used as an important tool to investigate the mechanisms of visual perception. In this figure, observers often report the perception of contours surrounding the central area as well as a relative brightness change in the central area. Both the illusory contour perception and the illusory brightness perception are essential components of the Kanizsa figure. The fact that such a simple figure can create such striking illusions must signify some fundamental mechanisms of signal processing in the visual system. It should be noted that these illusions are not specific to the configuration of Pac Man shapes with missing segments typically employed in the Kanizsa figure. They can, for example, also be evoked in a range of figures with arbitrary shapes. This indicates that the completion and the perception of illusory surfaces are generic properties of the visual system. The original configuration was employed merely to highlight how this illusory phenomenon could manifest itself in the highly simplified configurations.

The Kanizsa figure has inspired a large number of studies in psychophysics, neuroscience, and computer modeling (for a review, see Leshner, 1995). Attempts to explain this phenomenon have explored figural cues such as collinearity, occlusion, depth, good Gestalt, and familiarity of shape. These studies have focused not only on the original Kanizsa configuration but also on a comparison between this and its variants. The investigation of the variations to the original figure, especially the investigation of essential differences between illusory and nonillusory figures, provides key clues to the fundamental mechanisms behind these illusions. Because our goal is to answer how and in what condi-

tions the completion phenomenon emerges, it is essential to investigate theories/computer models in terms of whether they explain not only the illusions in the original Kanizsa figure but also the various perceptions of the variation figures. As pointed out later, however, this aspect of investigation has been neglected or followed only to a limited extent. In the following section, we discuss the critical figural properties that can be obtained from the analysis of the variations to the prototypical figure.

In Figures 1A–1K, the original Kanizsa square (1A) as well as its variations (1B–1K) are shown. It indicates that replacing the Pac Man objects by various shapes can create strikingly different effects: In some figures, the illusions disappear, whereas in others, the illusions seem to occur in various degrees. Replacing them by crosses, as in Figures 1B and 1C, for instance, almost causes the disappearance of the illusions. Figure 1D is almost identical to Figure 1B but is further modified by changing the ending of the crosses so that the shapes of all the side corners are exactly the same as the ones of the Pac Man shapes, and yet the illusions do not reappear. When one of the Pac Man objects is misaligned, as shown in Figure 1E, we perceive an incomplete square with three corners (bottom right) as well as a small brighter area within the corner of the misaligned Pac Man (top left). Between these two areas, we tend to see a rather ambiguous gradual modification of

the brightness without clear contour perception. Figure 1F is made in such a way that the “good continuation” (or “good Gestalt”) of the contours of the individual surrounding objects is reduced compared with the original Kanizsa figure (by smooth extension of the contours, the side junctions of the object do not meet each other). Figure 1G uses objects with an arbitrary complex form. In both cases, they still generate the illusions. They indicate, therefore, that the surrounding objects themselves do not need to evoke a clear perception of their own contours behind the central area (although the surfaces are perceived to continue, their contours remain underspecified). Figure 1G also shows that the surrounding objects do not need to have a simple, familiar, or regular shape. In Figure 1H, short straight lines are added to the original Kanizsa figure so that the line endings are aligned with the positions of the illusory contours. Both evoke strong illusions. Figure 1I is exactly the same as Figure 1A but with opposite contrast polarity (i.e., white objects on a black background). In this figure, the illusions are evoked again, but the central area is now perceived to be darker than the surrounding, indicating that the illusory brightness is sensitive to the polarity of the contrast. Figures 1J and 1K are also similar to Figure 1A but configured with an equal number of white and black objects on a midgray background. In these figures, the brightness in the central area appears to be the same as the



*Figure 1.* The Kanizsa figure (A) and its variations (B–K) used in this article and the effect of support ratio (L). The original Kanizsa figure creates brightness as well as subjective contour illusions. In some figures (B, C, D), the illusions are suppressed. E indicates that the central area shows a gradual change of the illusory brightness when objects are not aligned. The surrounding objects without the “good continuation” of the contours (F) or with an arbitrary complex form (G) still create the illusions. The straight lines which are added to the original Kanizsa figure (H) evoke strong illusions. In I, a configuration with opposite contrast polarity, the illusory brightness now consists of a darkening of the central area. In the cases of J and K, where an equal number of black and white objects are placed on a midgray background, the illusory brightness is suppressed, whereas the subjective contours remain. In the series of figures in L, the distances between the Pac Man shapes are changed and, hence, the support ratio (Shipley & Kellman, 1992) is decreased from left to right. Both the strength of the illusory brightness and the clearness of the illusory contours seem to increase as the support ratio increases.

surrounding, but the contours of the square are still perceived: an example of the persisting illusory contours without the illusory brightness.

In addition, Figure 1L shows that the strength of the illusions changes when the distance between the Pac Man objects changes (i.e., the support ratio effect, Shipley & Kellman, 1992). In all the figures shown in Figure 1L, the illusions still exist. It appears, however, that both the brightness of the central surfaces and the clearness of the illusory edges become weaker from left to right.

It is important to note that the nonillusory figures (Figures 1B, 1C, and 1D) indicate that the mere collinearity of the edges of the surrounding objects is not the key factor in generating the illusions. In addition, the illusions are shown to persist when clear contours are not observed (Figure 1E). Collinearity (or cocurvilinearity in general) of borderlines defined by a luminance difference is, therefore, neither a sufficient nor a necessary condition (although it might enhance the effects when the illusions occur). Furthermore, Figure 1D indicates that intersections between straight and curved edges (as in Figure 1A), as opposed to two straight edges (as in Figure 1B), are not the cause of the illusions either.

### Collinearity-Based Models

Most of the many attempts to mimic the Kanizsa illusory phenomenon in neurocomputational models have been inspired by the borderline-completion scheme driven by the collinear alignment of the contours of the Pac Man shapes (e.g., Grossberg, 1994; Grossberg & Mingolla, 1985b; Grossberg & Pessoa, 1998; Heitger, von der Heydt, Peterhans, Rosenthaler, & Kubler, 1998; Peterhans & Heitger, 2001; Peterhans, von der Heydt, & Baumgartner, 1986; Ullman, 1977). These models aimed to complete the gap between these aligned contours by introducing convolution filters or algorithms specifically designed for this purpose. Note that there are two illusory phenomena in the Kanizsa figure, as mentioned earlier: In addition to the illusory contours, the perceived brightness of the central area differs from that of the surrounding. Hence, mere borderline-completion models are, by design, incomplete. To reproduce both phenomena, borderline-completion models must also incorporate a filling-in process that constructs the surface in the central area, leading to what is called the *borderline-completion-and-filling-in approach*. Among the models mentioned above, the filling-in process has been realized only by the Boundary Contour System/Feature Contour System (BCS/FCS) model (Grossberg & Mingolla, 1985b) and by the FACADE theory (Grossberg, 1994). Other models only reported the reproduction of the borderline-completion phenomenon.

As shown in the previous section, however, the collinear pairs of line segments along the central square are not always completed in the variation figures. For these models to successfully distinguish the responses to the illusory and nonillusory figures, two possible solutions can be implemented: (A) to elaborate the filters or the algorithms in such a way that they can somehow distinguish, on the basis of the geometric properties of the objects, the differences between the variation figures, or (B) after completing the collinearly aligned line segments, to determine which completed lines should be perceived at the higher level and then to eliminate the others from the visualization process. Because most of these collinear contour-completion models have not been tested on the variation figures, one cannot tell decisively if and how any of the

models would succeed in reproducing the responses to them. If classified, however, most models have implemented Approach A. The exceptions are the BCS-FCS model (Grossberg & Mingolla, 1985b) and the FACADE theory (Grossberg, 1994), where both approaches are implemented. We discuss these two approaches in more detail by considering how these models by Grossberg's group might deal with the Kanizsa variations.

These models are designed to first create borderline completions of the collinearly aligned contours of the inducers by using so-called "bipole cells." The completion is achieved only if the orientation-competition algorithm associated with this process allows certain contours to extend (Approach A; see Grossberg & Mingolla, 1985a; Grossberg & Mingolla, 1985b). In addition, in contrast to other models, these models go further and fill in the area enclosed by these contours with the contrast values by using anisotropic diffusion (Neumann, Pessoa, & Mingolla, 1998). (The diffusion is called *anisotropic* because it is blocked or resisted at boundaries; see Mumford & Shah, 1989; Perona & Malik, 1990; Proesmans & Van Gool, 1999.) In these models, only when the contrast differences remain at the contour, the contour is visualized later on and the other contours are then eliminated from visualization (Approach B). These models successfully create the surface with modified brightness in the center of the original Kanizsa figure. It is important to note, however, that the collinear contours in examples such as Figures 1B, 1C, and 1D do not evoke illusory contour in our perception, whereas these contour pairs are physically exactly the same as the ones in the original Kanizsa figure. Their algorithms therefore need to incorporate some means of avoiding the creation of faulty illusory contours in these examples.

One possibility is that they do not complete the contours in these nonillusory figures at the early stage because of some property of the orientation competition algorithms (Approach A). In Figure 1A, the original Kanizsa figure, the straight contour may supersede the curved contour in this competition at the side corners of the Pac Man shapes, and hence, a borderline is extended from the straight contours, whereas in Figure 1B, it is possible that the competition of two straight contours at the side corners of each object prevents either of them from extending. However, if the width of the cross objects in Figure 1B is reduced as shown in Figure 1C (narrow crosses), the strength of the straight contours is different, and in such a case, it is not clear if the orientation competition algorithm would still successfully avoid the extension of the borderlines while evoking the extension seen in the original Kanizsa figure. Furthermore, consider Figure 1D. Its contours are exactly the same as in the original Kanizsa figure: The long straight contours in the middle are intersected by the curved contours at the side corners. The only difference is the existence of the same L-junctions oriented in other directions. To have borderline completion in the original Kanizsa figure but not in Figure 1D, the orientation competition has to be set so that it is in favor of the extension of the borderlines in the former but not in the latter. Considering the subtle differences of the contours in Figure 1A and 1D, this would be a quite difficult task (especially because this has to happen in a wide range of object sizes).

The second possibility for the models to succeed in reproducing proper responses in these examples is that the contour pairs in Figure 1B, 1C, and 1D are completed at the early stage of the model, just as it occurs in the Kanizsa figure, but they are then eliminated at the later stage (Approach B). Such an account could



work as follows. The completion of the collinearly aligned paired contours in these figures would create rectangular patches in the image at the early stage of the algorithm. The filling in of the contrast values in these patches (by the diffusion process implemented in their model) would create brightness differences and hence would preserve at least some of the borderlines along the patches, according to their rule mentioned previously. The responses to these three figures predicted from this approach, therefore, may not correspond to our perception (i.e., no or only little illusory perception). These arguments illustrate the difficulties of the models with the borderline-completion approach in general. These models are indeed able to reproduce the Kanizsa illusions, but many would encounter difficulties in avoiding creating illusions in the nonillusory figures.

These figures are therefore important because they clearly indicate that the existence of the collinearly aligned contours is neither a necessary nor a sufficient condition for the creation of the illusory surface. The Kanizsa illusions probably arise on the basis of other, more fundamental factors. What is the mechanism, then, that differentiates the responses to these variation figures?

### Occlusion Cues as Essential Triggers of the Kanizsa Illusions

The variation figures in Figure 1 suggest that algorithms based on borderline completion of the collinearly aligned contours of inducers may encounter some difficulties in reproducing responses congruent with human perception. These difficulties stem from the fact that the collinearly aligned borderlines do not always result in the completion. It is, therefore, highly possible that another more fundamental factor is determining the illusions. Careful comparison of our perception of the Kanizsa variation figures in Figure 1 indicates that whenever we perceive the central square to be occluding the surrounding objects (or segmented as the figure closest to the viewer) we perceive the illusions, and vice versa (Coren, 1972). This observation raises the possibility that what constitutes the core of the illusions is an attempt of the visual system to determine the depth order of the objects in the image.

In fact, there are several lines of evidence indicating that the perceived depth order determines the creation of the Kanizsa illusions. With an isoluminant Kanizsa image, the illusory figure is not evoked (Brussell, Stober, & Bodinger, 1977). This can be explained by the fact that depth perception is severely suppressed with isoluminant images, as has been convincingly demonstrated by de Weert (1979, 1983). Various studies using rating (Bradley & Dumais, 1984; Halpern, 1981), matching (Coren & Porac, 1983), and nulling (Gregory & Harris, 1974; Whitmore, Lawson, & Kozora, 1976) methods to measure depth perception have in fact indicated a strong association of the Kanizsa illusions with the perception of the central area being closer to the viewer. The fact that depth nulling (Gregory & Harris, 1974; Whitmore et al., 1976) can cancel the illusions supports this view. Moreover, a visual agnosia patient was reported to perceive the illusory figure only when it was presented with stereo disparity (Stevens, 1983), suggesting that this patient lacked neural mechanisms to detect pictorial depth cues and, as a result, could not see the illusions on the basis of pictorial cues alone. Gillam and Nakayama (2002) showed empirically that the perceived depth relationships determine the illusory contours and they suggested a schematic model. Finally, a

functional brain-imaging study (Mendola, Dale, Fischl, Liu, & Tootell, 1999) indicated that the brain areas that are active when participants see Kanizsa figures overlap with those that are active during depth-recognition tasks (e.g., V3A, V4v, V7, and V8). The recovery of three-dimensional (3-D) information from the 2-D signals in the retina is such a vital and fundamental function in the visual system that it may be an automatic mechanism triggered even by very simple pictorial stimuli. The evidence summarized above indicates that this process is playing a key role in creating the Kanizsa illusions.

Once it is realized that the detection of depth order is an essential trigger of the Kanizsa illusions, the perceived illusory brightness in the central area of the Kanizsa figure may be explained by the psychophysically well-known effect that depth perception has on the perception of photometric parameters (i.e., lightness, brightness, and perceived illuminance; for empirical reports, see Adelson, 1993; Dalby, Saillant, & Wooten, 1995; Gilchrist, 1977, 1980; Knill & Kersten, 1991; Schirillo, Reeves, & Arend, 1990; and Wishart, Frisby, & Buckley, 1997; for theoretical discussions, see Adelson & Pentland, 1996; Barrow & Tenenbaum, 1978; and Bergstrom, 1977). In some studies (Doshier, Sperling, & Wurst, 1986; Mallot, 1997), the notion of proximity luminance covariance has been developed to indicate the influence of perceived viewing distance to the perception of luminance and vice versa. Furthermore, some reports (Gilchrist, 1977, 1980; Schirillo et al., 1990) have clearly indicated that lightness perception is influenced by the 3-D interpretation. This means that the perceived lightness can be described as a function of the perceived depth.

### Border-Ownership Computation and Differentiation–Integration as a Mechanism

The hypothesis that depth perception is the fundamental key in creating the Kanizsa illusions, in conjunction with the known influence of perceived depth order on perceived lightness, provides a framework for the DISC model: The computed perceived depth is linked to lightness perception to reproduce the final subjective perception of the image. To realize this framework, we need to introduce an important concept of the DISC model: the differentiation–integration approach. In this section, it is explained why border ownership is important and how it can be regarded as the differentiated signal of depth that in turn leads to the differentiation–integration approach.

First, it should be noted that some models (Geiger, Pao, & Rubin, 1998; Gillam & Nakayama, 2002; Kumaran, Geiger, & Gurvits, 1996; Sajda & Finkel, 1992; Sajda & Finkel, 1995; Williams & Jacobs, 1997) have shown robust responses to the Kanizsa figure, including some (Geiger et al., 1998; Williams & Jacobs, 1997) that showed correct responses to the nonillusory figures. What they have in common is that these models are directed at computing the depth order of image regions (see Gillam & Nakayama, 2002, for a good example of this). To determine the depth order, these models compute which side of the borderline is closer to the viewer. FACADE theory (Grossberg, 1994, 1997; Grossberg & Yazdanbakhsh, 2005; Kelly & Grossberg, 2000), another model that shows robust responses, also implements depth computation by filling in the depth values within the regions

enclosed by the completed borderlines to create surfaces with different depth.

The region that is perceived as closer to the viewer can also be considered as the occluding surface in contrast to the more distant region that is considered as the occluded surface. Once the occluding surface is determined, the line that divides the two regions is not a mere borderline anymore. It is now a contour of the occluding surface. This condition can also be described in terms of border ownership: By determining the depth relationship, the ownership of the borderline is assigned to the occluding surface. The process that detects borderlines and the process that determines the border ownership should be clearly distinguished. The former merely indicates the existence of sudden changes, whereas the latter reflects the polarity (the direction) of the changes. It is possible that the reason behind the robustness of these models is that they are tuned not only to detecting and completing borderlines but instead to computing the border ownerships of visual images.

Note that BOWN signals are given in 2-D space, indicating the distribution of depth differences at each location. Note also that the BOWN signals preserve the polarity of the depth difference. This information is essential in constructing the depth map of the image by assigning depth values to all the areas in the surrounding space. An important notion proposed in this article is that, mathematically, this process can be regarded as 2-D integration of differentiated signals: BOWN signals are considered differentiated signals, and surface construction is a 2-D integration process. The surface completion by this 2-D differentiation–integration approach gives the fundamental design of the DISC model reported in this article. In the lightness domain, essentially the same principle can be applied. On the basis of the information from luminance contrast, differentiated signals of lightness can be constructed, and their 2-D integration results in a lightness map. In the rest of this section, we support the plausibility of this concept by explaining how it can be linked to the well-known relativistic behavior of neurons in the visual cortex as well as to the Gestalt psychologist's relational viewpoint.

It has been shown that neurons in primary visual cortex respond to the borderlines of objects (Hubel & Wiesel, 1959). This discovery has inspired the conception that neurons in visual cortex are borderline detectors. However, if neurons are merely borderline detectors, all the information of the surfaces (i.e., the interior between the borderlines) would be lost, and it could not be explained then how the visual system can perceive interior surfaces. If, for instance, an image consists of only two halves, one being black and the other being white, and if neurons merely respond to the borderlines in the middle, the information about the color of the surfaces on both sides would be lost. Hubel expressed this paradox as follows: "Many people, including myself, still have trouble accepting the idea that the interior of a form . . . does not itself excite cells . . . . [O]ur awareness of the interior . . . depends only on cells sensitive to the borders . . ." (Hubel, 1988, p. 87).

To solve this paradox, it is important to remember that Hubel and Wiesel (1962) actually reported two different types of neurons: line detectors and edge detectors. Whereas line detectors responded maximally when a line-like object was placed in the middle of their receptive field, the optimal stimulus for edge detectors was the edge of a surface with a particular transition

of the contrast (light-to-dark or dark-to-light surfaces; i.e., the polarity of the contrast). The existence of edge detectors is quite important in terms of the aforementioned paradox: If some neurons are sensitive to contrast polarities, the visual cortex as a whole is capable of reconstructing the surfaces attached to borderlines. Although this example is about the polarity of the contrast, this can be generalized: The visual cortex is capable of detecting not only the existence of differences (of specific properties, such as luminance, depth, or texture) but also the polarity of the differences at the borderlines, and hence, it is capable of reproducing the information of the interiors. In fact, recent neurophysiological studies showed that neural activities of some neurons in V2 and V4 reflect the ownership of the borderlines (Qiu & von der Heydt, 2005; Zhou, Friedman, & von der Heydt, 2000), indicating not only the existence of the depth difference but also that the side of the area that owns the border is closer to the viewer in the depth domain. Note that differentiated signals preserve the direction for the differentiation so that the integration process can be performed properly. The aforementioned neural activities that preserve the direction of the difference correspond well to the nature of the differentiated signals, thus highlighting how these responses could be considered as representing input information in the form of the differentiated signals (see the following).

In terms of perceptual phenomena, on the other hand, Gestalt psychologists have emphasized the fundamentally relational nature of human vision (Koffka, 1935): Human vision is sensitive to the relative values of input signals, as opposed to the intensities of the input per se. If, as described previously, the local differences of the input signals are detected at the early stage of the visual system and the visual system tries to reconstruct the entire image on the basis of this information, the macroscopic properties emerging through this process would indeed be relational. This way of explaining the relativistic nature of our vision has been discussed theoretically (Arend, 1973; Gilchrist, Delman, & Jacobsen, 1983; Retinex theory; Land & McCann, 1971; Ross & Pessoa, 2000; Wallach, 1948; for a formalism of the concept, see also Arend & Goldstein, 1987) and has been supported experimentally (Arend, 1973; Gilchrist, 1979; Hung, Ramsden, Chen, & Roe, 2001; Krauskopf, 1963; Whittle & Challands, 1969). In fact, Land and McCann's (1971) Retinex theory (see also Horn, 1974) succeeded, to a certain extent, to reproduce the perceived color constancy by detecting the luminance ratios, by eliminating the illuminance components (low-luminance ratio signals) from them, and by spatially integrating the remaining luminance ratios. Furthermore, the integration process accumulates the local information and the final value at each point reflects the values from the entire space. In this way, the process would be able to reveal the macroscopic properties reflecting the local properties in the entire space. This may, in fact, correspond to the observation that the visual system is capable of reflecting the luminance ratio of surfaces that are remote from each other (Arend, 1973; Land & McCann, 1971; Whittle & Challands, 1969).

By combining the relativistic behavior of the neurons that neurophysiologists reported and the relational nature of our perception defended by the Gestalt psychologists, we can offer the following important insight: The activities of the neurons that are sensitive to the polarity of the difference can be interpreted as differentiated signals derived from the input properties, and the visual system

(re)constructs the images by integrating these differentiated signals. That is, when neurons appear to respond to borderlines, their responses actually carry the information about the interior between the borderlines implicitly in the form of differentiated signals. By applying the spatial integration process, the information of the interior becomes explicit. Our relational perception emerges through this process.

### A Surface-Construction Model Based on the Differentiation–Integration Approach

In the previous two sections, we explained how we reached the conclusion that (a) depth perception and its influence on lightness perception is the fundamental factor in creating the illusions and that (b) surface construction by the 2-D differentiation–integration approach is the most plausible underlying mechanism. Because the input is given as 2-D information at the retina, the differentiation and integration processes should be 2-D as well. As a result, the DISC model aims at surface construction. In this section, we compare the DISC model with other surface construction models as well as with other theories that have taken the differentiation–integration approach, we point out the novelties and the advantages of our approach, and we conclude this section.

The responses of the models directed at surface construction mentioned earlier (Geiger et al., 1998; Kumaran et al., 1996; Sajda & Finkel, 1992, 1995; Williams & Jacobs, 1997) were surfaces segmented on the basis of depth differences. This segmentation of surfaces in the depth domain is not, however, the complete reproduction of our subjective visual perception of the input images, which is the goal of the DISC model. The model, then, has to create as a final output a perception of the visual image in the lightness and brightness as well as in the depth domains. These earlier surface-construction models, in fact, do not explain the brightness change in the central area nor explain the sensitivity of this phenomenon to contrast polarity (see Figure 1A vs. Figures 1I, 1J, and 1K). As described later, in the DISC model, the computed depth is linked to the computation of the final lightness perception. In this way, the DISC model is able to reproduce the lightness and brightness perception of the images as influenced by depth perception. In addition, perceived surfaces can have gradual changes in depth and brightness without having clear contours (see Figure 1E for an example). It is very unlikely that the surface-completion models that create segmented regions surrounded by contours (mentioned earlier in this paragraph) are able to reproduce the illusions where clear contours are not perceived. The 2-D (leaky) integration process implemented in the DISC model can incorporate gradual depth and lightness changes in surfaces and flat and segmented surfaces (patch works) are not necessarily assumed.

Whereas the surface-completion models based on the depth computation mentioned above (Geiger, Pao, & Rubin, 1998; Gillingam & Nakayama, 2002; Kumaran, Geiger, & Gurvits, 1996; Sajda & Finkel, 1992, 1995; Williams & Jacobs, 1997) lacked a way to reproduce the perception in the photometric domains, the theories and models based on the differentiation–integration approach mentioned previously (Arend, 1973; Arend & Goldstein, 1987; Gilchrist, Delman, & Jacobsen, 1983; Land & McCann, 1971; Ross & Pessoa, 2000; Wallach, 1948) dealt with only the photometric

values but not the depth values. They used the relative values of photometric parameters (luminance at retina) to reconstruct perception of the image. It is important to note that the DISC model is different in this regard. The DISC model is designed to integrate the relative depth (BOWN) signals between the objects to generate a 3-D reconstruction of the image, in addition to the integration of the luminance ratios to generate a lightness map of the image. In other words, it aims at a surface reconstruction in the geometric domain as well as in the photometric domain using this differentiation–integration approach.

In summary, although there have been numerous attempts to create models to mimic the Kanizsa illusions, there seems to be no single model that is capable of realizing all three of the following aspects at once: (a) reproducing the two characteristics of the phenomenon (illusory contours and illusory brightness); (b) producing appropriately varying responses to the variation figures, most importantly including nonillusory figures, with consistent algorithms; and (c) not only segmenting out the surfaces in the depth domain but also reproducing our subjective visual perceptions of the figures in the photometric domain. We have therefore developed a neurocomputational model—the DISC model—that takes a new approach which realizes all of these aspects in an integrated way. As a first step in this model, L- and T-junctions are detected as occlusion cues, and then BOWN signals are computed at each point of the borderlines by horizontal interactions between the occlusion cues and the borderlines. These BOWN signals are considered as differentiated signals in the depth domain, and 2-D integration is then applied. The same principles are also applied to the lightness domain. The interactions of these domains, depth and lightness, play a key role in creating the subjective perception. The design of the DISC model reflects the above analysis of the well-known characteristics of the Kanizsa illusions, relies on the overall biological structure of the visual system, and applies neurophysiologically plausible algorithms. As a result, the model consistently reproduces the perception of the original Kanizsa figure as well as all of its variations. We are convinced that these figures illustrate some fundamental aspects of signal processing in human visual cortex and reveal the basic mechanisms of the computation of depth based on occlusion cues, of the interaction of perceived depth and lightness, and of visual completion. By analyzing the responses to these figures, we believe the DISC model brings to light the biological mechanisms underlying subjective visual perception more generally. The wide range of applications of this differentiation–integration approach will be published elsewhere (Kogo, Galli, Van Gool, & Wagemans, 2010; Kogo, Van Gool, & Wagemans, 2010); here we use it as a core component of a more elaborate model for the Kanizsa illusions.

The next section describes the model in more detail. It first detects occlusion cues, computes the border ownership at each point, and constructs surfaces with perceived depth from them; then, finally, the depth values are linked to the perceived lightness values to reproduce the Kanizsa illusions. The subsequent section shows some results obtained with the model, which is then followed by a more general discussion to conclude the article. Detailed mathematical descriptions of the model are given in the Appendix.

### The Differentiation–Integration for Surface Completion (DISC) Model

#### Principles

The DISC model, which is summarized in Figure 2 and described in more detail in the following sections, is constructed in two channels: the photometric channel and the geometric channel. The model starts from the input signals given as luminance values. In the photometric channel, to reflect the ratio but not the difference of the luminance between the two areas, the differentiation is performed on the logarithm of the input luminance following the formalism of the Retinex theory (Horn, 1974; Land & McCann, 1971). The result is called *luminance ratio map* in this article. Taking the exponential of its integration creates the *primary lightness map*. In the geometric channel, the model determines the depth order of the objects in the image. This is achieved by first detecting occlusion cues and then constructing the relative depth. The relative depth is signaled by the BOWN map that is constructed by reflecting the distributions of the occlusion cues (L- and T-junctions) and the borderlines in the image. These signals indicate which side of the borderline is the owner area (Zhaoping, 2005).

The construction of the difference maps (in luminance ratio and in depth) can be described by Equation 1 that defines the gradient ( $\nabla$ ) of (log) luminance and depth.

$$\nabla L_0 = \left( \frac{\partial}{\partial x} \log(I_0), \frac{\partial}{\partial y} \log(I_0) \right)$$

$$\nabla D = \left( \frac{\partial}{\partial x} D, \frac{\partial}{\partial y} D \right). \tag{1}$$

The letters indicate the following:  $L_0$  = (log) luminance,  $I_0$  = input luminance,  $D$  = depth. The construction of surfaces can be described as the integration of these differentiated signals (see Equation 2).

$$L_p = \exp\left(\alpha \cdot \int_w \nabla L_0 d\vec{r} + C_L\right)$$

$$D = \int_w \nabla D d\vec{r} + C_D \tag{2}$$

$$d\vec{r} = dx \cdot \vec{i} + dy \cdot \vec{j},$$

where  $L_p$  = primary lightness;  $w$  is an integration pathway;  $\alpha$ ,  $C_D$ , and  $C_L$  are constants;  $r$  with an arrow on top indicates a unit vector in space;  $i$  and  $j$  with arrows on top indicate unit vectors in the  $x$  and  $y$  directions, respectively. Equations 1 and 2 are used to formulate the concept of the differentiation–integration approach in mathematical terms. In the real computation in the model,

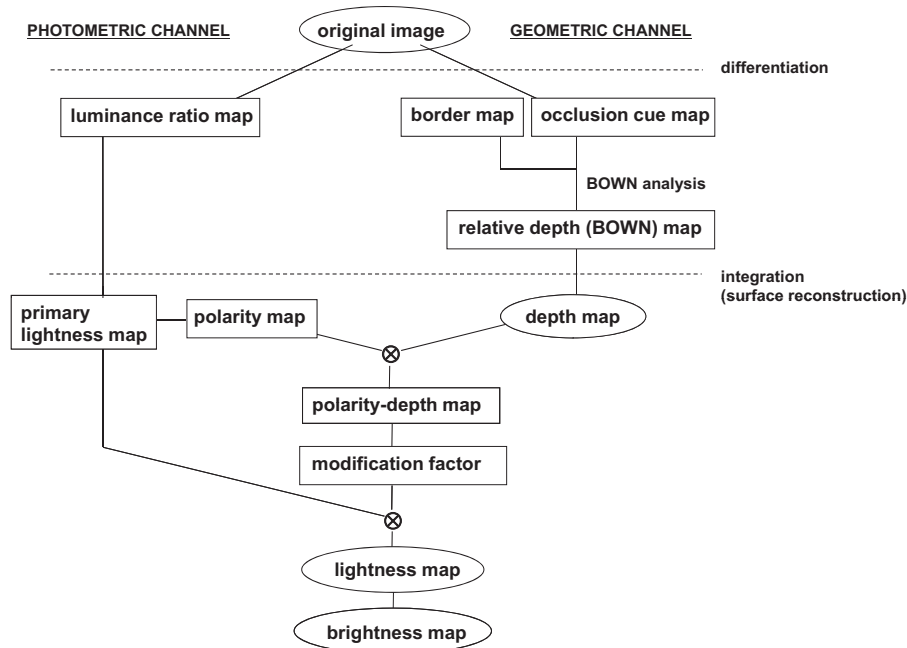


Figure 2. Basic structure of the DISC model. The model is divided into photometric (left) and geometric (right) channels first. In both channels, relative value maps (luminance ratio map and border-ownership [BOWN] map) are created. The luminance ratio map is created by detecting differences between the values in the logarithm of the input luminance, whereas the BOWN map is created by detecting local occlusions cues (L- and T-junctions) and by their interactions with borderlines. The relative value maps are integrated to reconstruct surfaces in the primary lightness map and the primary depth map. The primary depth map is further processed in the feedback process to create the final depth map. From the primary lightness map, the polarity map is created, and the product of it and the depth map results in a polarity-depth map ( $D_p$ ). This map is used to create a modification factor, the product of which with the primary lightness map results in the lightness map.



elaboration of the algorithms is necessary to realize this concept (i.e., the differentiated map in the depth domain is constructed by the computation of the border ownership and the computation of the integration is different from mathematical integration to achieve more biologically plausible algorithms). We describe the specific elaboration of these principles in more detail in the following paragraphs.

The final lightness map is created after further reflecting the depth perception. This is done by constructing a modification factor based on the depth map, which is multiplied to the primary lightness map (see Equation 3). In the most general form of the DISC model, the brightness map is created by multiplying the lightness with the perceived illuminance of the entire scene. In the figures used in this article, however, the illuminance is assumed to be constant. Hence, the brightness map is a simple product of the lightness map and a constant value representing the perceived illuminance.

$$L = L_p \cdot M(D_p)$$

$$B = \gamma \cdot L, \quad (3)$$

where  $M$  = modification factor,  $B$  = brightness, and  $\gamma$  = constant (replaced by perceived illuminance in a general form). Because the subjective perception of lightness is sensitive to contrast polarity (compare Figures 1A and 1I), the sign of the contrast between the objects and the background has to be implemented in the modification factor. The polarity-depth map,  $D_p$ , is therefore used to construct the modification factor (see details later in this section). Equation 3 indicates the relationships between the lightness domain and the depth domain as well as the brightness domain.

L- and T-junctions are considered to be occlusion cues that play a key role in this model. Note that the differentiated signals in general preserve the directions of the differences (the directionality of the changes). The junction signals initiate the BOWN computation as the depth difference is implied by the signals. L-junctions divide areas into two areas: areas  $S_1$  and  $S_2$  in Figure 3A. In the model, it is tentatively assumed that the area on the concave side with the narrower angle (less than  $180^\circ$ ) of a junction is the occluding area ( $S_1$  at the center junction,  $L$ ), and the other one with the wider angle (more than  $180^\circ$ ) is the occluded area ( $S_2$ ). However, this is only an implication from individual L-junctions. To determine the depth order of the two abutting surfaces, all existing

junctions must be considered. In fact, there are two other junctions in each Pac Man shape (the side junctions,  $L_2$  and  $L_3$ ), which act as opposite cues. At the stage of the BOWN computation, the conflicting information from the center junction ( $L_1$ ) and the two side junctions ( $L_2$  and  $L_3$ ) contribute to it competitively: The center junction contributes to the signal that indicates that the surface in the central area is higher, whereas the side junctions contribute to the signal that indicates that the surface inside the Pac Man shape is higher. The BOWN computation reflects the global distribution of junctions and borderlines. When a T-junction is detected, our perception of occlusion works in the opposite way. As shown in Figure 3B, the areas on the side of the stem of the junction ( $S_{S1}$  and  $S_{S2}$ ) are considered to be the occluded ones, and the area above the top of the junction ( $S_T$ ) is the occluding one. The detected T-junctions also contribute to the BOWN computation accordingly.

In the following sections, the details of the algorithms are explained. Mathematical descriptions of the algorithms are given in the Appendix.

### Occlusion-Cue Detections and Border-Ownership Computation

Borderlines are detected by detecting luminance differences pixel by pixel. When a borderline changes its direction, the bending point is detected. This point is assigned as an L-junction. Even when a junction is constituted by two one-pixel-long borderlines, it is detected as a junction. The small bending points along the circular edges of the Pac Man figures are therefore also considered L-junctions. If the borderline is separated into two directions, the point is assigned as a T-junction (see Appendix for details). Amplitudes of all junction signals have a value of one (binary signals).

The BOWN map is created by the algorithm described in the Appendix and is summarized here only schematically (see Figure 4). In this article, a BOWN signal is indicated as an arrow with a side fin always on its left side, as shown in the top of Figure 4A. The arrow corresponds to the orientation of the borderline, and the side of the fin indicates the side of the ownership (with its length indicating the strength of the signal; see Zhaoping, 2005). Note that two  $180^\circ$  opposing ownerships are possible at each point on the borderlines (e.g.,  $0^\circ$  and  $180^\circ$ ; see the bottom of Figure 4A). The BOWN computation in the DISC model takes account of the distributions of junctions and borderlines in the entire space so that the macroscopic properties of the image are reflected. The idea is that the individual BOWN signal at each point on the borderline is compared with junction signals first and the resulted BOWN signals are compared with each other in the next iteration process. In both processes, the same principle is applied: If the pair agrees that the same side is the owner of the borderline, the BOWN signal is enhanced, and if it disagrees, it is inhibited. More specifically, if a concave side of an L-junction matches (or does not match) the owner side of a BOWN signal, they are considered to be in agreement (or in disagreement; see Figure 4B). Similarly, in the next step, if a pair of BOWNs is oriented so that the sides of their owners correspond with each other, the signals are considered to be in agreement, and if not, they are considered to be in disagreement (see Figure 4C). The amount of enhancement and inhibition decreases as a function of the increasing distance between them. After the iteration, the BOWN signal on one side becomes larger than the other (see Figure 4D, quantitative

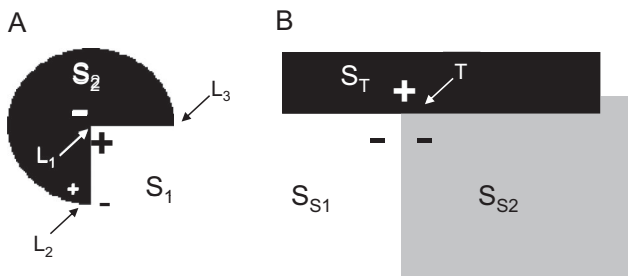
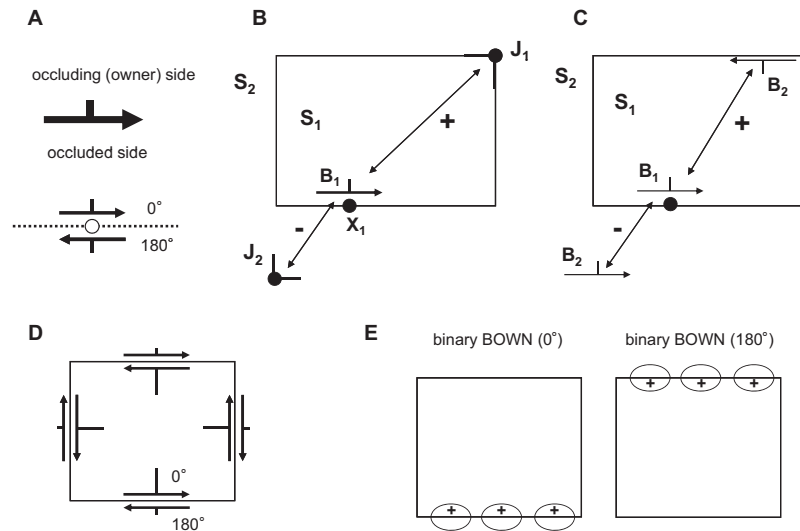


Figure 3. Occlusions hinted at by L- and T-junctions. In case of L-junctions (A), the central junction ( $L_1$ ) indicates the inside surface ( $S_1$ ) as the occluding area, whereas the surface on the other side of the junction ( $S_2$ ) is occluded. The side junctions ( $L_2$ ,  $L_3$ ) suggest otherwise. In case of T-junctions (B), the areas  $S_{S1}$  and  $S_{S2}$  are interpreted as being occluded.



*Figure 4.* Border-ownership (BOWN) computation. A BOWN signal is indicated as an arrow with a side fin on its left side (A, top). The arrow corresponds to the orientation of the borderline, and the side of the fin indicates the side of the ownership. At each point along borderlines, there are two possibilities of ownership (A, bottom). First, the BOWN signal ( $B_1$ ) at a point  $X_1$  is compared with junction signals (B). If a concave side of an L-junction (e.g.,  $J_1$ ) matches with the owner side of  $B_1$ , it is considered to be in agreement, and  $B_1$  is enhanced. If it does not match, it is considered in disagreement, and  $B_1$  is inhibited. In the next iteration process (C), the resulting BOWN signals are compared with each other. If a pair of BOWNs is oriented so that their owner sides correspond to each other (e.g.,  $B_1$  and  $B_2$ ), the signals are considered to be in agreement. The BOWN signal is again enhanced in that case. After the iteration, the BOWN signal on one side becomes larger than the other (D). The BOWN signals of the opposing directions are compared and a value of +1 is given to the winning side at each point (E).

BOWN map). The BOWN signals of the opposing directions are compared, and the value of 1.0 is given to the one that is stronger (binary BOWN map). This binary BOWN is indicated as an oval on top of the borderline point, and the owner side is indicated by a + sign in this article (see Figure 4E).

In addition to the BOWN computation at the points on the borderlines, there is an extra step implemented in the model to compute the BOWN at the locations where a physical borderline does not exist (free-space BOWN). In Figure 5, BOWN signals of a rectangle (A) and two rectangles forming T-junctions (B) are shown. In the case of a rectangle, two BOWNs attached to an L-junction ( $J_1$ ) indicate that the inside of the rectangle is the owner (both are in agreement; see the inset). On the other hand, at the T-junction,  $J_2$ , the BOWNs on the top part of the T indicate that the area above is the owner, whereas the BOWNs on the stem indicate that one side (the left in this case) is the owner. As is shown later, the same condition can emerge at L-junctions with Kanizsa-type illusory figures. This condition is shown in Figure 5C as L-junction  $J_3$ . This condition of the BOWNs suggests the existence of an occluding surface above, and hence  $J_3$  becomes a T-junction. In such a case, when this illusory T-junction condition is detected, the free-space BOWN computation is initiated. Once again, the interaction between junctions and BOWN signals are computed, except that this time, all points in the entire space (all potential borderlines between pixels) are considered, and the junction signals are from the illusory T-junctions (see Figure 5D). At each pixel, there are four potential borderlines surrounding it. At the bottom of a pixel, for example, two opposing BOWNs are possible ( $B_1$  and  $B_2$ ). They are either enhanced or inhibited according to the

interaction with the illusory T-junctions. This process results in the illusory BOWNs in the free space. The final BOWN map is a combination of this illusory binary BOWN map and the previously computed physical BOWN map (wherever the BOWN signal exists, no matter whether it is the physical BOWN or an illusory BOWN, the BOWN signal is assigned as the final BOWN signal at the location; see Equation A3 in the Appendix). This BOWN map is used for integration at the next stage.

## Surface Reconstruction

In the DISC model, the integration of the BOWN map results in the depth map, and the integration of the luminance-ratio map results in the lightness map. This integration process makes the macroscopic properties explicit. To determine the exact method for these integrations, it was important to first analyze the characteristics of surface reconstruction in the visual system. In Figure 1E (with a misaligned Pac Man) the perceived central surface seems to continue from the central area near the three aligned Pac Man shapes toward the misaligned Pac Man as a gradual change of brightness. The surface reconstruction by spatial integration as it occurs in visual perception, therefore, should differ from mathematical integration in this regard. When a pulse-like signal is mathematically integrated, for instance, a step-wise signal is created (causal integration). Vision, on the other hand, seems to perform a gradually decaying integration. We developed an algorithm to mimic this biological characteristic of surface construc-

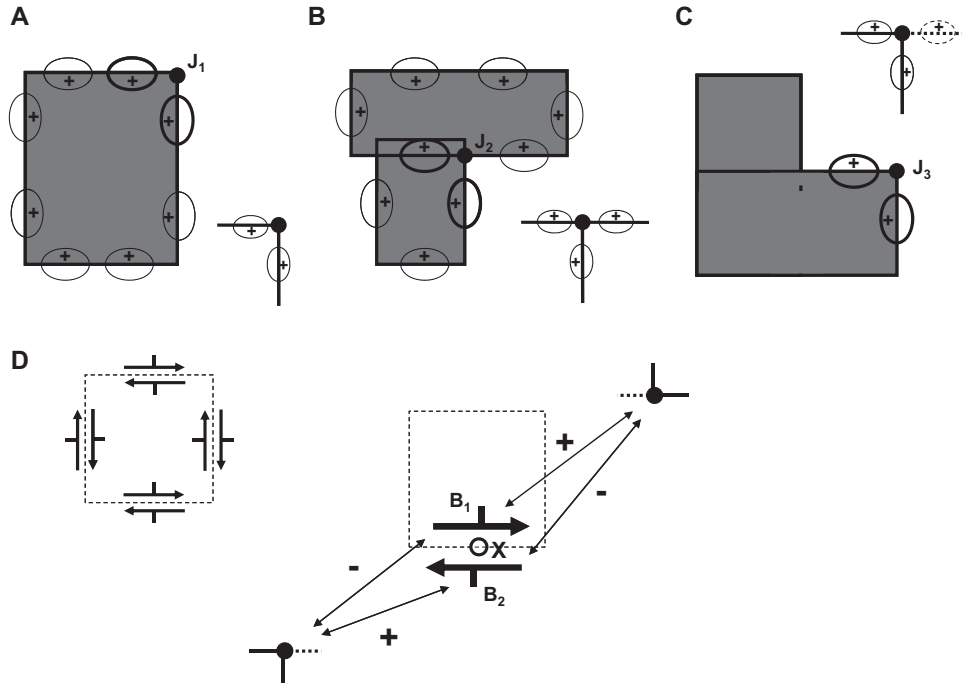


Figure 5. Illusory T-junction and free-space border-ownership (BOWN) computation. A: BOWN computation with a rectangle figure results in consistent BOWN signals. For example, at an L-junction,  $J_1$ , the attached two BOWN signals both indicate that the inside of the rectangle is the owner side. B: When a T-junction exists, the top part of the T and one side of the stem are the owners for these two separate areas. In such a case, the BOWN signals attached to the T-junction are not “consistent” in the sense that they do not indicate the same side as the owner. This makes sense, as it is defined by a T-junction. C: The same condition can appear on L-junctions after BOWN computation. In such a case, the BOWN signals attached to the L-junction do not make sense unless the junction is assumed to be an illusory T-junction. D: As one half of the top part is missing, the existence of the illusory T-junction triggers the free-space BOWN computation. For this process, first, all four borders surrounding each pixel in the entire space are considered. The BOWN computations as above are repeated between the borders of pixels and the illusory T-junctions.

tion by the use of the so-called “leaky integration” described in the following.

Leaky integration is “a deliberately imperfect integration” (Claerbout, 1992, p. 66), typically found when the step change of voltage decays because of the leakage of capacitors in electric circuits. As shown schematically in Figure 6A, if a differentiation map indicates the existence of a step change in the input, the leaky integration of these signals creates, instead of step functions, gradually decaying signals. It is applied to the BOWN map and the luminance-ratio map two-dimensionally to create the depth map and the (primary) lightness map, respectively. Taking the example of depth computation of a figure with a small rectangle inside of a large rectangle, this process is explained in Figure 6B. In the model, all integration pathways start from a point where a BOWN signal exists. When the integration pathway,  $p_1$ , for example, starts at the edge of the large rectangle ( $a$ ), the integration would first result in one step up because of the BOWN signal at  $a$  in the direction of the integration. Furthermore, the integration goes up another step when it crosses the edge of the small rectangle ( $b$ ). Because this is leaky integration, the value decays in distance and reaches, for instance, at point  $c$ . When the pathway goes further, it crosses BOWN signals of opposite direction (points  $d$  and  $e$ ),

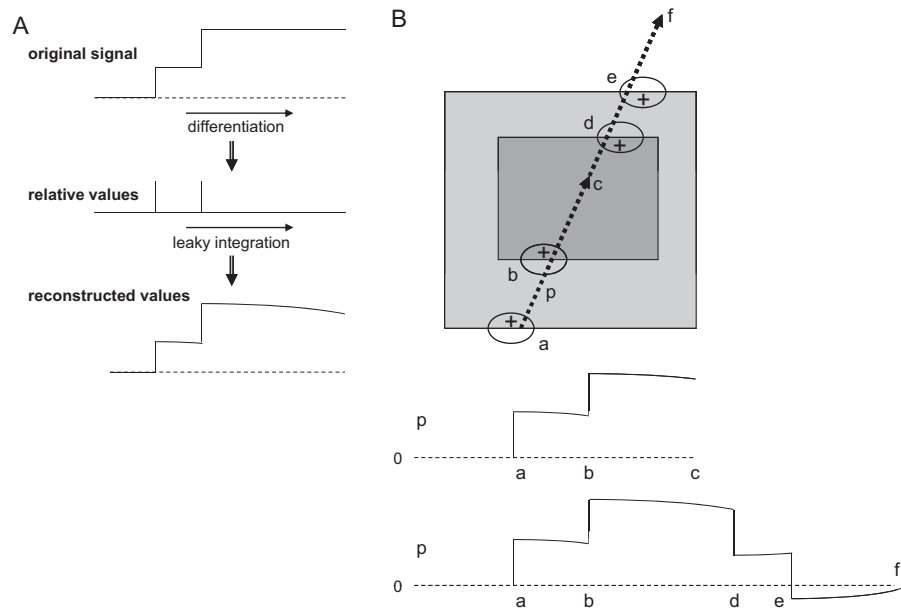
and hence, the integration values step down before reaching a point outside of the rectangles ( $f$ ). The depth map is yielded by averaging the integration values from all pathways at each point (see Appendix for details). The accumulation of integration values at each local BOWN signal along the pathways results in depth values that are influenced from the entire space (the accumulative effect).

To obtain the lightness map, the logarithm of the luminance is first differentiated (see Equation 1). This corresponds to detecting the ratio, not the contrast, of the lightness (the luminance ratio map,  $\nabla L_0$ ). Applying the same process of integration described above creates the luminance-ratio-sensitive lightness map (the primary lightness map,  $L_p$ ).

$$D = Q\{BOWN\} + C_D$$

$$L_p = \exp(\alpha \cdot Q\{\nabla L_0\} + C_L). \tag{4}$$

Here,  $Q$  indicates the application of the leaky integration algorithm to the values inside the brackets. This equation set is equivalent to Equation 2 except that the integration is now leaky integration and the differentiated depth is now written explicitly as BOWN (see the Surface Reconstruction section in the Appendix for details).



*Figure 6.* Schematic drawings of the leaky integration process. A: The basic property of the leaky integration explained one-dimensionally. A computation of a relative value of a certain aspect of an input image (e.g., border ownership or luminance ratio) for a certain direction (left to right in this example) would result in signals as shown in the middle. Application of the leaky integration to the same direction creates surfaces with assigned integrated values that decay in distance (bottom). B: Expected border-ownership signals ( $0^\circ$ ) for a large rectangle and a small rectangle within should indicate that the borderlines are always owned by the inside area of the rectangles. With the example integration pathway shown here, the integration value steps up at the borderline of the large rectangle at the bottom and steps up further at the borderline of the small rectangle at its bottom.

## Contrast Polarity

As described in the preceding, the basic design of the DISC model is to construct the modification factor,  $M$ , based on the depth map and then to modify the photometric structure provided by the primary lightness map to obtain the final lightness map (see Equation 3). The construction of the modification factor, however, has to be done with caution. By comparing the two Kanizsa figures with opposite contrast polarities (see Figures 1A and 1I), an important aspect of these illusions can be observed: The central area in the black-Pac-Man-figures-on-white-background (black-on-white) configuration is perceived as brighter than the surrounding area, whereas in the white-Pac-Man-figures-on-black-background (white-on-black) configuration, it is perceived as darker. In addition, the variations of the Kanizsa figure shown in Figures 1J and 1K with midgray background and equal numbers of white and gray Pac Man shapes (midgray Kanizsa figures) do not yield the illusory brightness in the central area. The Kanizsa illusions are therefore sensitive to contrast polarities.

Figure 7A shows schematic drawings of the predicted depth map of the figures with these three configurations. Because the polarity and the strength of the contrast are irrelevant for the geometric channels of the model, the depth maps of these figures are exactly the same. Therefore, if the modification factor simply reflected the depth values, it would increase the brightness of the central areas in all three figures. As explained

above, this is incorrect, so the modification factor has to be sensitive to contrast polarity. In other words, the influence of depth perception is to enhance the contrast, not the intensity itself. Hence, it is necessary to create a polarity (of contrast) map from the result of integrating the luminance ratio map.

The predicted polarity maps of these three figures are shown in Figure 7B (they can only have values of 1,  $-1$ , or 0). The product of the depth map and this polarity map is called the *polarity-depth map*,  $D_p$ , shown in Figure 7C. Comparing the black-on-white and white-on-black figures, their polarity-depth maps (see Figures 7Ca and b) are now the mirror images for each other. In the midgray figure, on the other hand, because the background is midgray and there are equal numbers of white and gray Pac Man shapes, only negligible differences would be created between the central area and the surrounding area after integration. This results in assigning the value of zero in the central and surrounding areas in the polarity map, as shown in Figure 7Bc. The polarity-depth map of this figure, therefore, also obtains a zero value in these areas. In sum, with the black-on-white and white-on-black figures, the polarity-depth maps would result in the central areas having positive or negative values, respectively. The polarity-depth map of the midgray Kanizsa, on the other hand, would have the value of zero in this area. Hence, if these values in the polarity-depth map are reflected in the modification factor, it would result in the correct brightness maps for these figures.



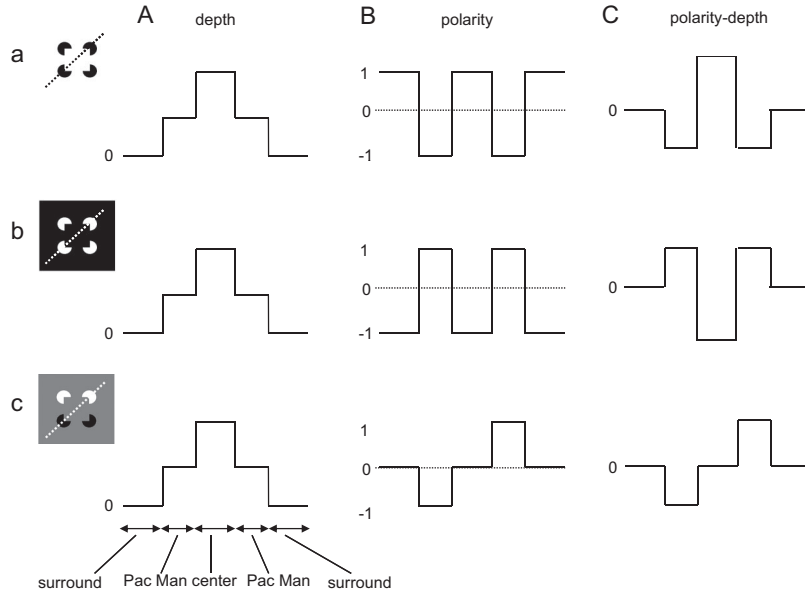


Figure 7. Schematic representation of the depth (A), polarity (B), and polarity-depth (C) maps in the figures with different contrast polarity configurations. Profiles are expressed one-dimensionally, as if the figures are cut through by the dotted lines shown in the insets on the left. In all three figures, the model gives the same depth maps (A). Because of the difference of the contrast polarity, however, different polarity maps are created (B). Reflecting the differences of the polarity, the polarity-depth maps (the product of the depth map and the polarity map) are also different in these three figures (C).

**Anchoring and Construction of the Modification Factor and Brightness Map**

As indicated in Equation 2, the integration process introduces an offset value (the constants,  $C_D$  and  $C_L$ ). This additional degree of freedom is inevitable in the differentiation-integration approach (as it is in mathematical integration). This means that additional information is necessary to determine the offset values. In our view, this is exactly equivalent to the anchoring problem that has been known in lightness-perception research. Whereas the perception of lightness of adjacent areas depends on the luminance ratio between the areas (i.e., their relative values, not their absolute values), in the end, the visual system also has to assign perceived lightness values (absolute values) to these areas. Different rules of assigning these values (i.e., to solve the anchoring problem) have been investigated in the literature on lightness perception (for a review, see Gilchrist et al., 1999). In the DISC model, the anchoring problem is dealt with at the stage of creating the depth map and the primary lightness map. The detailed arguments are described elsewhere (Kogo, Van Gool, & Wagemans, 2010). In the primary lightness map, the highest value rule is applied in such a way that the area with the highest primary lightness is anchored. In the depth map, the area rule is applied in such a way that the area with the largest surface area is anchored. The polarity-depth map is computed from the anchored depth map, and this is implemented in the modification factor. The anchoring algorithm ensures that the areas that are anchored in both the primary lightness and the polarity-depth maps obtain a pre-fixed value in the lightness map (corresponding to perceived white), whereas the areas that are not anchored in either map receive values that are different from this fixed value (see Equation 5; see Appendix for details).

$$D_p = \underline{D}_0 \cdot P$$

$$L = \underline{L}_{p1} \cdot \underline{M}(D_p)_1, \tag{5}$$

where,  $D$  and  $P$  are the final depth and the polarity maps, respectively;  $L_p$  is the primary lightness map;  $L$  is the final lightness map; and  $D_p$  is the polarity-depth map. Underlining indicates that the map is anchored to the value indicated on the right of the line. At the point where the value of  $M$  is 1.0 (no modification), the lightness value does not receive a modification from the primary value. In case the primary lightness value is 1.0 (corresponding to white), this simply leads to the value of white (1.0) in the final lightness map. If the  $M$  value is more (or less) than 1.0 at the point, the final lightness value becomes higher (or darker, respectively) than white. (The perception of an achromatic color lighter than white is often called “super-white”; Gilchrist et al., 1999.)

$M$  is a function that determines how the computation of lightness is influenced by the depth interpretation of the image: In other words, this is the link from depth to lightness perception. There is not enough information available in the psychophysical or perceptual literature to determine the exact form of  $M$ . Phenomenologically speaking, it should be a monotonically increasing function. We therefore considered a simple linear function, as indicated in Equation 6, where  $\beta$  is a constant. In the future, when more precise information of the depth-lightness linkage becomes available, the form of this function can be modified.

$$\underline{M}_1 = \beta \cdot D_p + 1. \tag{6}$$

The final output of the model, brightness, is computed from the lightness map multiplied by a constant value (Equation 7:  $B$  = the brightness map and  $\gamma$  = the constant). The constant is set to 1.0 in

this article. In a general form, this value should reflect the perceived illuminance. It is assumed, however, that the perceived illuminance for the figures used in this article is constant in the whole space.

$$B = \gamma \cdot L. \quad (7)$$

The DISC model hence gives outputs for lightness, brightness, and depth. Lightness is the perceived reflectance. The computation of this value can therefore be directly compared with psychophysical measurements of the perceived lightness values elicited from the Kanizsa figure and its variants.

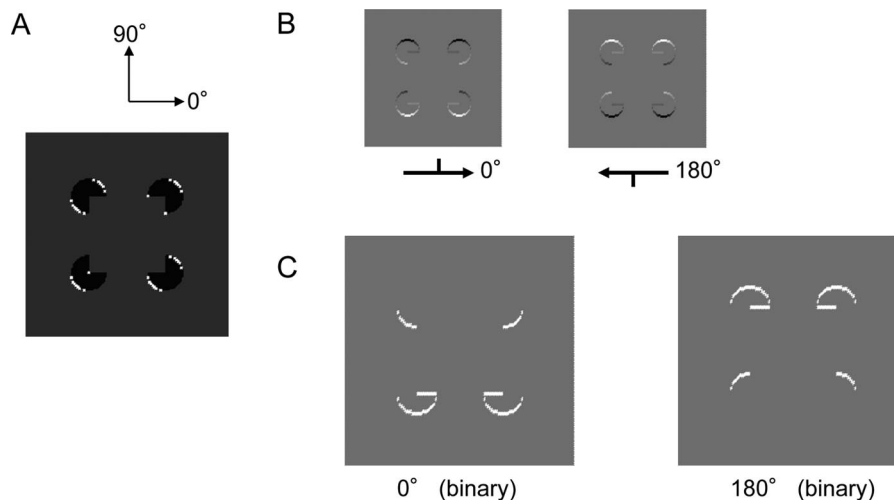
## Results

### Junctions and BOWN Maps

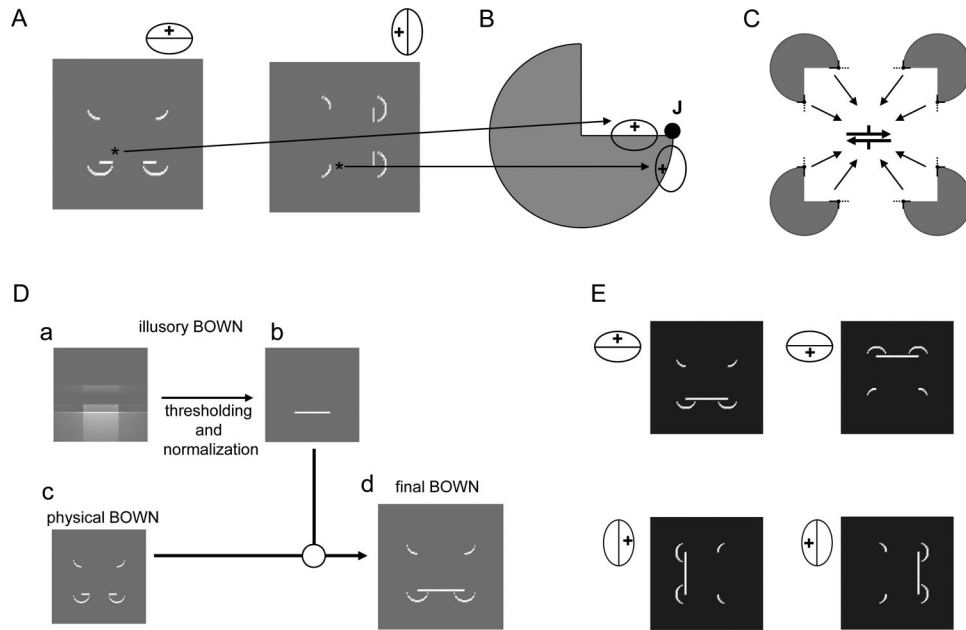
The results of L-junction detection for the original Kanizsa figure are shown in Figure 8. Figure 8A shows the signals for the junctions that consist of  $0^\circ$  and  $90^\circ$  edges. The central and the side junctions of the Pac Man shapes as well as all small junctions along the curved contours of these figures are detected. By the interactions of the junction signals with the borderline signals and the proceeding iteration process, the quantitative BOWN signals develop (see Figure 8B). The BOWN signals of the opposing directions are then compared, and the binary BOWN map is created (see Figure 8C). This result of the BOWN computation indicates that the central area is the owner of the borderlines surrounding it. The BOWN signals at this stage are called *physical BOWNS* because they indicate the ownership of the physically existing borderlines. Next, BOWN signals in the entire space, even

where physical borderlines do not exist (*illusory BOWNS*), are computed.

The result of the above BOWN computation creates the *illusory T-junction* condition described earlier (see Figure 9A). For example, the BOWN signals attached to the junction,  $J$ , of the Pac Man shown in Figure 9B indicate that the straight contour of the Pac Man is owned by the area above it (outside of the Pac Man) and the curved contour attached to  $J$  is owned by the area on the left side of it (inside of the Pac Man). At all side junctions of the Pac Man shapes, this illusory T-junction condition emerges. At the following stage of the BOWN computation, the free-space BOWN computation, these illusory T-junctions influence the development of the BOWN signals in the entire space (see Figure 9C). The result shows that the BOWN signals become strongest at the location within the gap of the straight borderlines of the Pac Man shapes because only in these locations do all illusory T-junctions consistently contribute to develop the BOWN signals that indicate the ownership by the central area (see Figure 9Da). The contributions by the illusory T-junctions to the remaining area are not consistent because these locations are on the top side of some of the T-junctions but also on the stem side of the other T-junctions. The BOWN signals in these locations are therefore weaker. After the application of a threshold, only these strong illusory BOWN signals within the gaps of the Pac Man shapes remain and are normalized (illusory BOWNS; see Figure 9Db). The BOWN signals on the physically existing borderlines computed by the preceding process (physical BOWNS; see Figure 9Dc) and the illusory BOWN signals are combined to create the final BOWN signals (see Figures 9Dd and 9E with all four directions).



*Figure 8.* The results of L-junction detection (A) and the quantitative border-ownership (BOWN) map (B) and the binary BOWN map (C) from the original Kanizsa figure. Junction signals are indicated as white dots superimposed on the original figure. Only the L-junction signals constructed from the borderlines with the directions of  $0^\circ$  and  $90^\circ$  are shown in A. Note that the central junction of the bottom-left Pac Man and the side junctions of the top-right Pac Man are detected as well as small bending points along the curved contours of the Pac Man shapes. In B, quantitative BOWN maps of two opposing directions,  $0^\circ$  (top) and  $180^\circ$  (bottom) are shown at the top. They are compared, and then a winner-take-all process is applied (binary BOWN map; C). The junction signals at corners of individual pixels. Therefore, they are marked in the four pixels surrounding each junction in the plot in A. The borderlines are detected between pixels. The BOWN signals in B (and the rest of the figures) are therefore marked in the two pixels on both sides of the borderline.



*Figure 9.* The appearance of illusory T-junction conditions in Kanizsa figures. After border-ownership (BOWN) computation of the Kanizsa figure (A), the two BOWN signals attached to the side junction, *J*, of the Pac Man shapes show an inconsistency (B). One indicates that the area above (outside of the Pac Man) is the owner, whereas the other indicates that its left side (inside the Pac Man) is the owner. This condition can only happen in case of a T-junction. Hence, the junction is considered an illusory T-junction in the model. In the free-space BOWN computation, the illusory T-junctions contribute to develop the BOWN signals in the entire space, even where there are no borderlines (C). Because only within the gap between the straight borderlines of the Pac Man shapes do all illusory T-junctions consistently contribute to develop the BOWN signals that indicate the ownership by the central area; the BOWN signals become strongest at the location (Da). After the application of a threshold, these strong illusory BOWN signals remain and are normalized (illusory BOWNs; Db). The BOWN signals on the physically existing borderlines (physical BOWNs; Dc), and the illusory BOWN signals are combined to create the final BOWN signals (Dd). As a result, the final BOWN signals indicate that the central area owns the surrounding straight borderlines, whereas the areas within the Pac Man shapes own the curved borderlines of the Pac Man shapes (E; the final BOWNs of all four directions are shown).

## Surface Reconstruction and the Depth and Lightness Maps

The application of the integration algorithm to the BOWN map results in surfaces with different depth (depth map; see Figure 10A). Both the Pac Man shapes and the central square in the Kanizsa figure gain heights in the depth map compared to the background, whereas the central square is in turn higher than the Pac Man shapes, as the BOWN map indicates that the central area is the owner of the borderlines between them. Applying the integration to the luminance ratio map results in a primary lightness map with nearly flat surfaces created in the areas segmented by its original colors (black and white) in the input (see Figure 10D). The polarity map for this figure, computed from the primary lightness map, gives a value of one to the center and the surrounding areas and a value of minus one to the areas inside the Pac Man shapes (see Figure 10B). The polarity-depth map ( $D_p$ ; see Figure 10C) is created by multiplying the depth map (anchored to zero; see Figure 10A) with the polarity map. Reflecting the contrast polarity, the areas within the Pac Man shapes gain negative values.

The modification factor is constructed from the polarity-depth map following Equation 6. It results in higher values in

the central square compared with the surrounding area (see Figure 10E), which has a value of 1.0. The lightness values (see Figure 10F) are computed by multiplying the primary lightness map with this modification factor following Equation 5. The lightness map for the original Kanizsa figure obtains higher values in the central square compared with the surrounding area because of the influence of depth perception enhancing the contrast between them; this corresponds to the brightness component of the Kanizsa illusions.

The responses of this model (the depth and the lightness maps) to different support ratio (see Figure 1L) are shown in Figure 11. The central area is segmented from the surrounding area more clearly and its lightness value becomes larger with a larger support ratio.

## Responses to the Variation Figures: 1. Four Crosses Figure

The responses of this model to the four crosses figure (see Figure 1B) are shown in Figure 12. In clear contrast to the original Kanizsa figure, the borderlines surrounding the central area are now owned by the individual cross objects (see Figure 12A), and there are no illusory

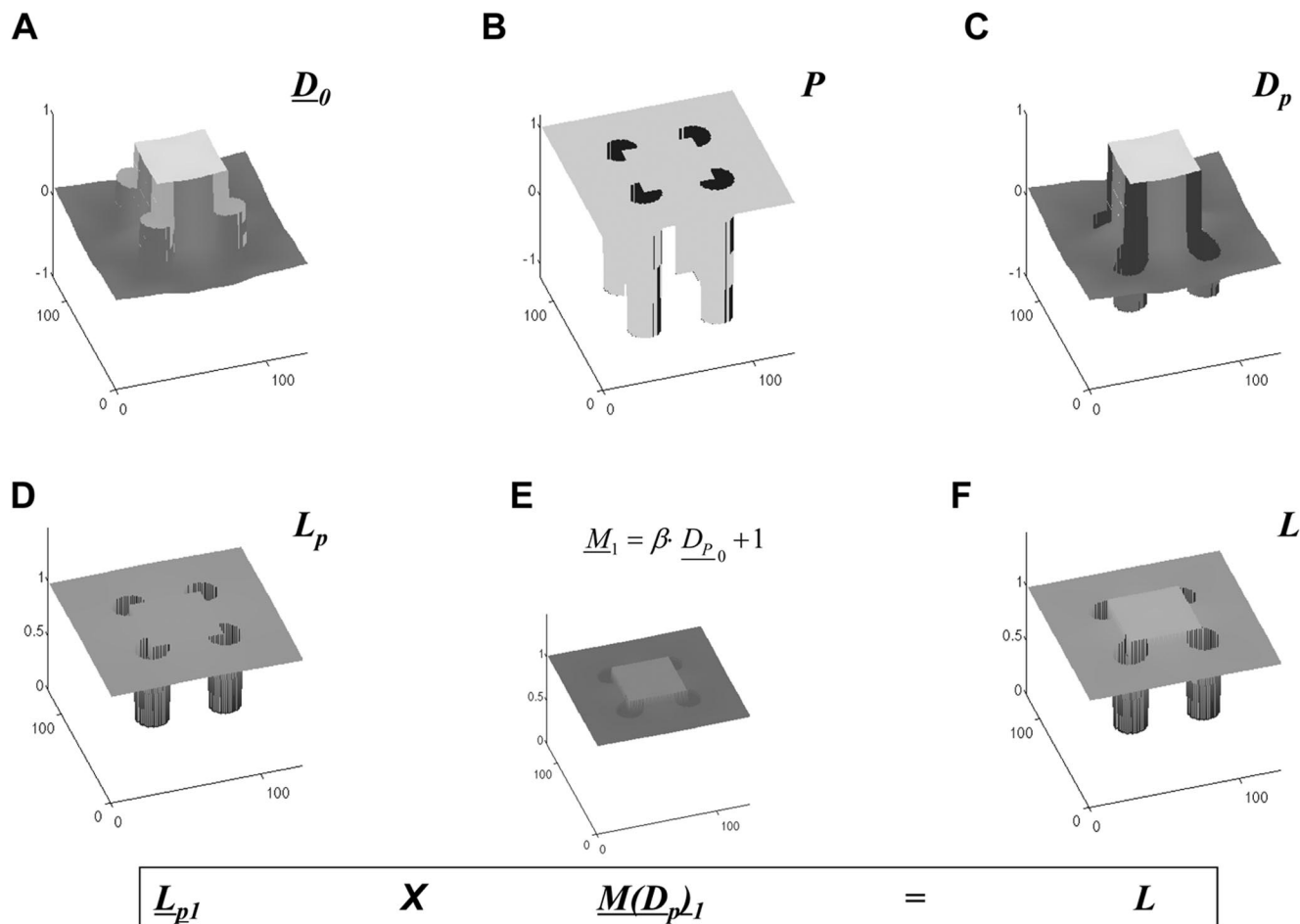


Figure 10. The final depth map (anchored to zero; A) and the polarity map (B) from the Kanizsa figure are multiplied to create the polarity-depth map (C). The modification factor is computed from the polarity-depth map (E). The product of the primary lightness map (D) and the modification factor results in the lightness map (F). Because of the influence of the polarity depth, the central square area obtains a higher value than the surrounding area.

T-junction conditions, and hence, no illusory BOWNs develop. In the depth map (see Figure 12C), the difference between the central square and the surrounding area is virtually nonexistent. The area within each cross object, on the other hand, gains significant height in the depth map because of the border ownership by this area. As a result, no central surface develops. Accordingly, the depth map and the polarity-depth map (see Figure 12D) result in a nearly flat surface in the area corresponding to the entire white area of the input. Consequently, the lightness map (see Figure 12E) shows no clear difference between the central and surrounding areas, in accordance with our perception.

### Responses to the Variation Figures: 2. Figures With Supporting Side Lines

Adding supporting side lines to the original Kanizsa figure, our perception of the central occluding square becomes stronger (see Figure 1H, side-line figure). The BOWN computation process for this figure correctly indicates this occlusion. As shown in the inset of Figure 13A, the end point of the side line toward the center is owned by the central area, whereas the opposite end and the longitudinal

borderlines along the side line are owned by the line itself (i.e., the side line is considered as an object). This BOWN computation creates the illusory T-junction conditions at the side junctions of the Pac Man shapes. In the final BOWN map, just as in the original Kanizsa figure, therefore, the illusory BOWN signals surrounding the central square areas develop (see Figure 13A).

As in the Kanizsa figure, the depth map develops the central surfaces that are accompanied with sharp edges (see Figure 13C), which in turn enhance the lightness values in the areas. The resulting lightness map of this figure is shown in Figure 13E. These results correspond well to our perception of these figures.

### Responses to the Variation Figures: 3. Misaligned Pac Man Figure

The responses of the model to the Kanizsa figure with one of the Pac Man shapes being misaligned (*misaligned Pac Man figure*; Figure 1E) are as follows. For the three aligned Pac Man shapes, BOWN signals develop just as in the original Kanizsa figure. The free-space BOWN computation results in strong



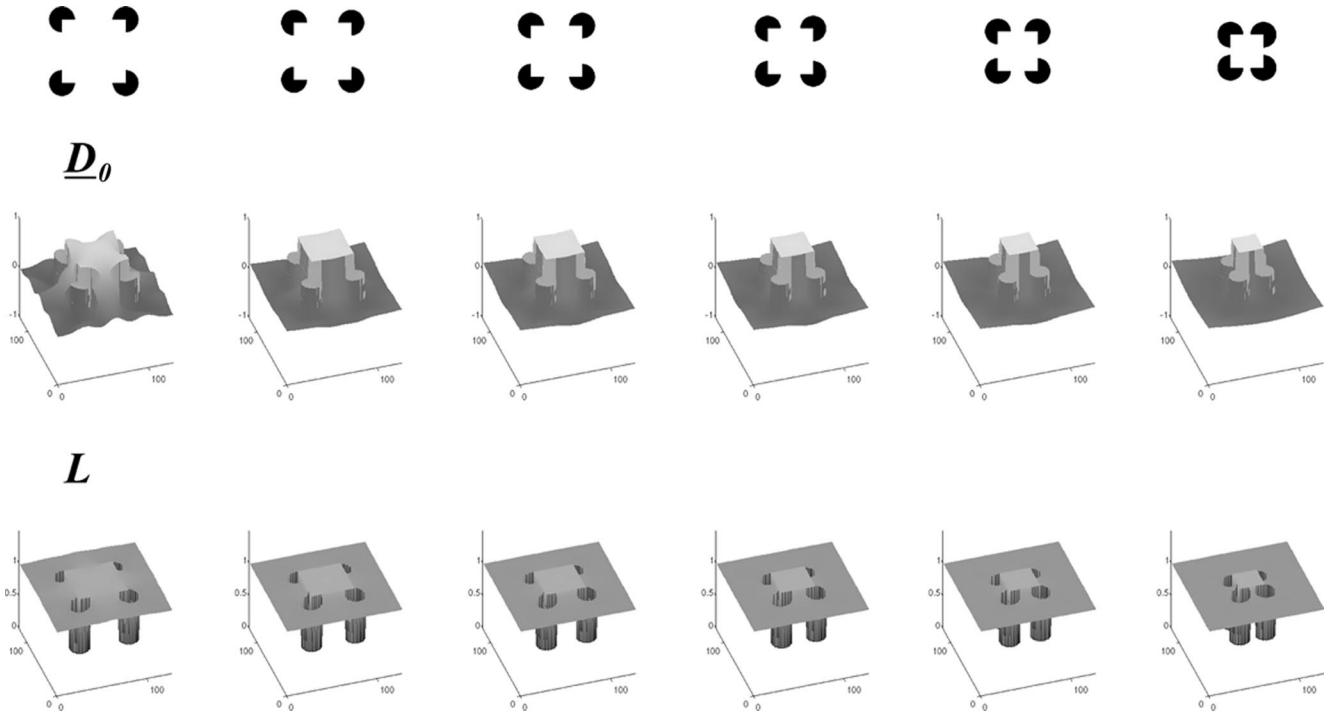


Figure 11. The depth maps (middle) and the lightness maps (bottom) of the Kanizsa figures with variable support ratio.

enough illusory BOWN signals only between these Pac Man shapes but not between them and the misaligned Pac Man shape. The depth map which results from the integration of this BOWN map is shown in Figure 14D. It develops the central surface with high values near the three aligned Pac Man shapes. From there toward the misaligned Pac Man, a smooth gradual decay of the depth values is created. The values go up slightly again near the misaligned Pac Man. Accordingly, the lightness map reflects the depth profile in the gradual lightness change in the area between the aligned and misaligned Pac Man shapes, which again corresponds to our perception of the figure (see Figure 14E).

#### Responses to the Variation Figures: 4. Effects of Contrast Polarity

In the Kanizsa figure with white objects on a black background (see Figure 11), the central square is perceived as darker than the surround. The depth map of this figure is exactly the same as those with the opposite contrast polarity, because the algorithms in the geometric channel (see Figure 2) do not reflect the contrast polarity. The polarity-depth maps of these figures created by the model, on the other hand, are complete mirror images of those with the opposite polarity (compare Figure 15A with Figure 10C). As a result, the lightness map gains less in the central area than in the surrounding area, which corresponds to our perception of this figure (see the side view of the lightness map in Figure 15C; note the slightly lower value in the center).

#### Responses to the Variation Figures: 5. Figures With Gray Background

When the input figure consists of an equal number of white and black objects on a midgray background, as in Figure 1J (midgray [diagonal] Kanizsa) and in Figure 1K (midgray [horizontal] Kanizsa), the model creates exactly the same depth maps as in the original Kanizsa figure (see Figure 16A, using the example from Figure 1J) because of the contrast polarity insensitivity of the depth computation algorithm. The polarity-depth map of this figure (see Figure 16B), on the other hand, does not develop values in the central area because the signals of opposite polarities cancel each other out and the value of zero is assigned in the polarity map (see Figure 16B inset). As a result, the lightness values in the central area are the same as in the surrounding area (see Figure 16C).

#### Responses to the Variation Figures: 6. Summary

All the lightness maps of the model to the original Kanizsa figure and its variation figures are shown in Figure 17. The narrow four cross figure and the four crosses with round corner figure (see Figures 1C and 1D) result in virtually no increase of the lightness value compared with the surrounding area (see Figures 17C and 17D, respectively) as the four crosses figure (see Figure 17B). Figures 1F and 1G create higher values in the central square compared with the surrounding areas (see Figures 17F and 17G, respectively) as the other illusory figures described above (see Figures 17A and 17H). These results all match our perception of these figures.

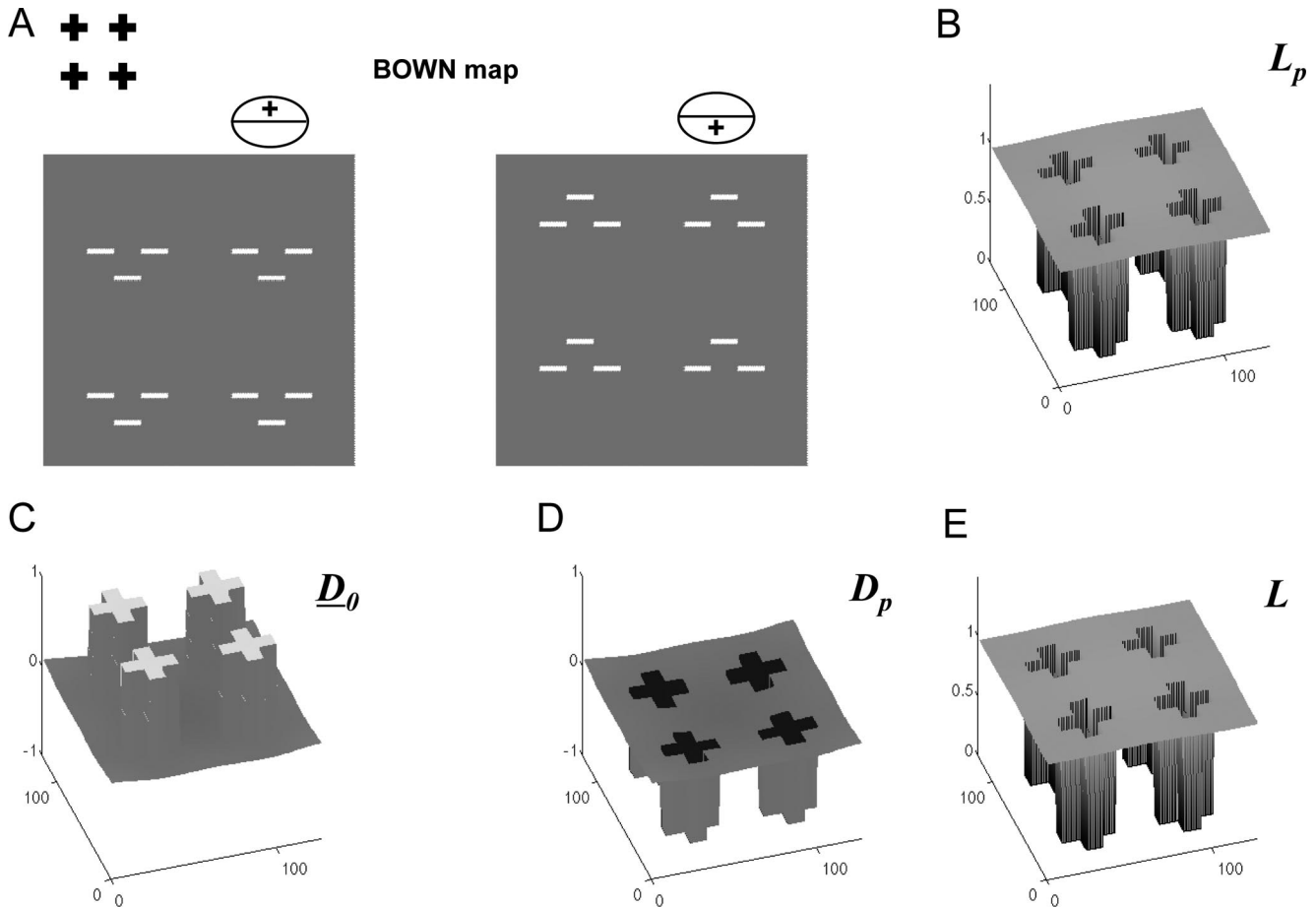


Figure 12. The maps of border ownership (BOWN; A), primary lightness (B), depth (C), polarity depth (D), and lightness (E) from the four-crosses figure (A, inset). Note that the BOWN signals of the borderlines surrounding the central area indicate that the cross objects are the owners of the borderlines. As a result, the depth map (C) indicates that only the cross objects are figures segmented out from the background. Because the depth map does not create differences between the central square and the surrounding area, the computed lightness map obtains about the same values in these two areas.

### Discussion

The general form of the Kanizsa type illusions is the perception of an occluding surface with a more arbitrary or complex shape in the middle surrounded by inducer objects. The Kanizsa figure is thought to indicate some of the visual system's fundamental mechanisms. To fully understand these mechanisms, it is critical to analyze not only the original Kanizsa figure but also its variations. An inspection of Figure 1 highlights that the illusions disappear or become effectively suppressed in some cases. Clearly, the issue is not simply the completion evident in the original Kanizsa figure. It is about how to and how not to complete the figures, depending on their contexts. If a model reproduces the Kanizsa illusions in the original figure but fails to reproduce the diversity of responses evoked by different variations, it lacks psychological plausibility. However, in contrast to the present model, most previous reports of computational models did not show their responses to variation figures or did so only to a very limited extent. In this respect, our model is the most extensive one so far.

To summarize, the DISC model is based on the following principles: (a) The fundamental triggers of the illusion are the depth (occlusion) cues, (b) the occlusion cues imply the difference of depth at each location, (c) border ownership is computed based on the occlusion cues and interactions over the entire space and is considered as 2-D differentiated signals, (d) the 2-D spatial integration of these differentiated signals create (illusory) surfaces in the depth map, and (e) the depth–lightness linkage creates the perceived lightness of the image. By implementing these principles, the DISC model successfully reproduces the perception of the variation figures and the two aspects of the Kanizsa illusion—the illusory contour and the illusory brightness in the central area—are explained by the creation of surfaces with sharp edges in the depth domain and its influence on the lightness domain. The robustness of the model supports the plausibility of this approach. In the following sections, we first discuss the implication of the approach taken in the DISC model and then discuss the connection with the empirical evidence. We then discuss the generalization of our view

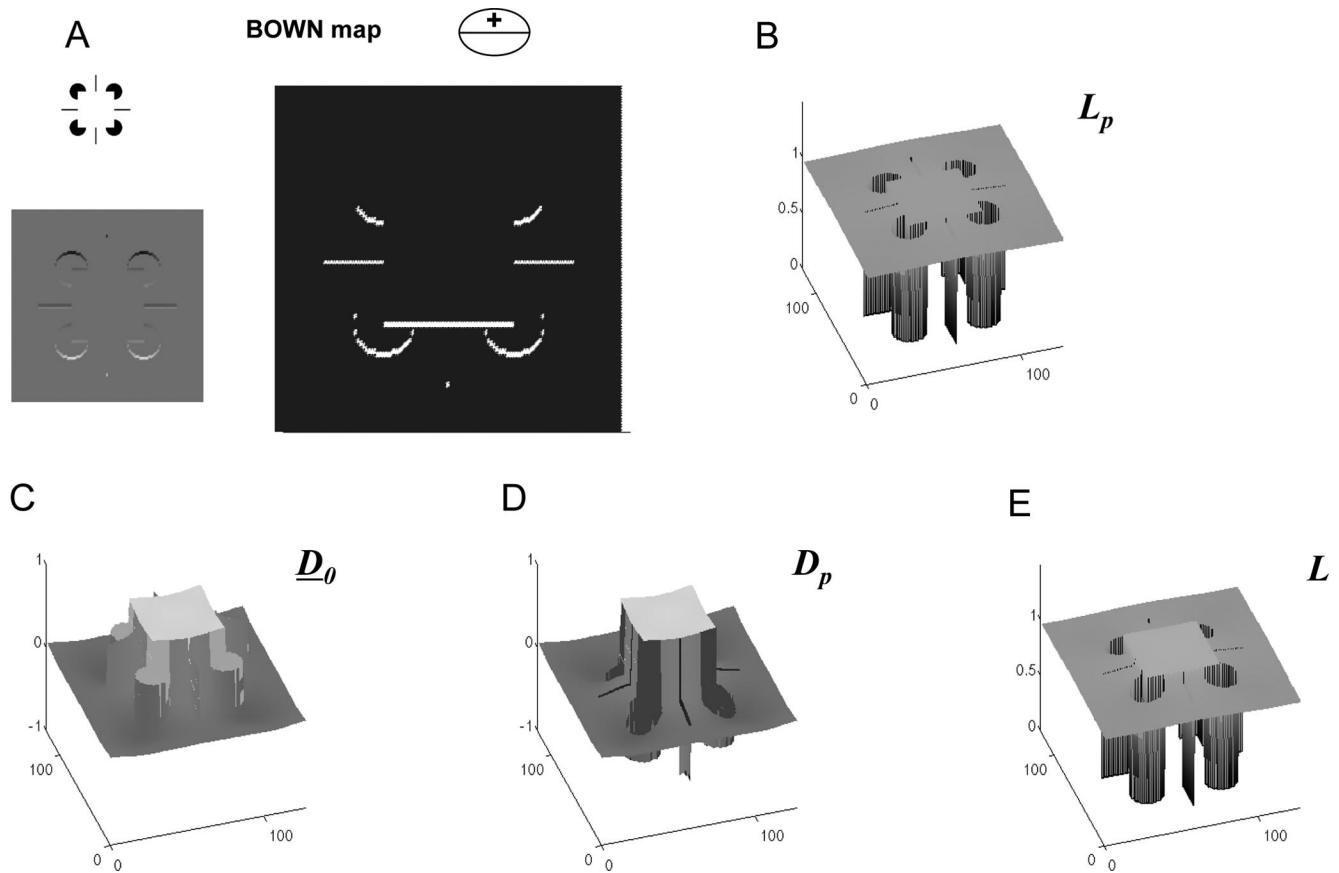


Figure 13. The maps of border ownership (BOWN; A), primary lightness (B), depth (C), polarity depth (D), and lightness (E) from the side-lined Kanizsa figure (A, inset top). The quantitative BOWN map before free-space BOWN computation is triggered is shown as an inset in A (bottom).

of Kanizsa-type illusions and discuss the general application of it to a wider range of perceptual phenomena.

### What “Illusory Contours” Really Are

The collinear alignment between contour segments of inducers (e.g., Pac Man shapes) is something we immediately notice when we see the figure. This is probably why we are inclined to argue that it is the key factor causing the illusions. This view, however, is not necessarily well founded. In fact, it is explicitly contradicted by counterexamples (nonillusory variants) in which the same collinear contour segments do not, in and of themselves, evoke illusory percepts. Because our perception of the figure is the final outcome of a variety of signal-processing mechanisms, phenomenal observations regarding our perception itself have to be interpreted with great caution. The fact that we notice the collinearity could just as well be the result, not the cause, of the perception of the figure after more fundamental processes of signal processing. Although it is still possible that collinearity of the contour segments contributes to the subjective perception of the figure, it cannot be the single satisfactory condition for the creation of the illusions. This highlights the difficulties with borderline-completion approaches described in the introduction. By regarding the computation of depth as the primary cause of the illusions, on

the other hand, the DISC model in fact develops BOWN signals from the occlusion cues. BOWN signals differ from mere borderline signals and indicate the occluding and occluded sides of a borderline (i.e., they indicate the polarity of the difference). This difference, whether it deals with the existence of changes (borderlines) or the polarity of changes (BOWNs), is crucial in distinguishing the responses to illusory and nonillusory figures. After the first (junction-BOWN) and the second (BOWN-BOWN) interaction, the BOWN computation results in different BOWN signals on the straight borderlines along the central area in the illusory and nonillusory figures (compare Figures 9A and 12A). This result reflects essential differences between these figures in terms of their geometrical properties. Whereas the surrounding objects in the four crosses figure are symmetric, the objects in the Kanizsa figure are asymmetric: The overall configuration of the Kanizsa figure is thus skewed inward. More specifically, there are more concave components than convex components viewed from the central area, in the illusory figure. In the nonillusory figure, there is no such bias, and concave and convex components are distributed equally in the individual objects (crosses). In the model, the side from which the curvature is seen as concave is more likely to be an occluding side (and hence creating a figure with convex shape on the side). Therefore, the skewed shapes of

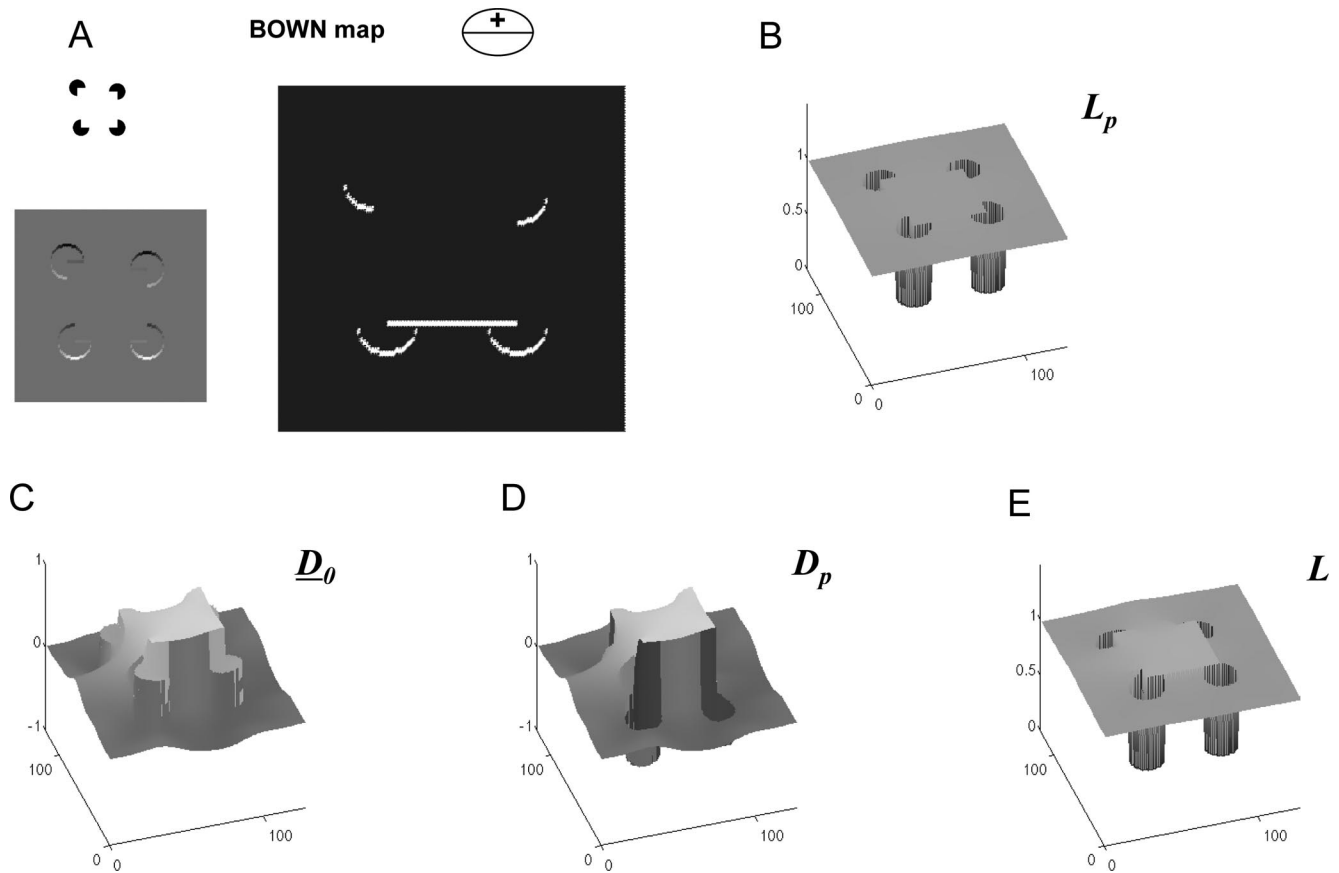


Figure 14. The maps of border ownership (BOWN; A), primary lightness (B), depth (C), polarity depth (D), and lightness (E) from the misaligned Kanizsa figure (A, inset). The quantitative BOWN map before free-space BOWN computation is triggered is shown as an inset in A (bottom).

the surrounding objects in the illusory figure give rise to a bias in the BOWN computation in the illusory figures, which favors the emergence of stronger BOWN signals for the central.

The final BOWN computation process by the DISC model is schematically shown in Figure 18. At individual junctions of the cross object (A), BOWN signals attached to the junction always indicate that the inside of the cross is the occluding object. The integration of BOWNs therefore creates surfaces only within the crosses, and they develop as individual objects in the depth map. With illusory figures, on the other hand, inconsistent BOWN signals are created at the side junctions (B). This condition occurs when the junction is a T-junction but does not occur when it is an L-junction as given physically. By taking these L-junctions with the inconsistent BOWN signals as illusory T-junctions, the DISC model repeats the algorithm of BOWN computation in the free space: interaction between (potential) BOWN signals in the entire space and the illusory T-junctions. As a result, the illusory BOWNs are created in the free space.

In other words, the BOWN computations determine the occluding surface by exploiting the global configuration of the image. The result should therefore be considered in terms of the completion of BOWN signals, as opposed to the completion of borderlines. Through the BOWN computation, which depends on the distribution of junctions and borderlines in the entire space, and the

integration process, which has an accumulation effect (the final integration value is the result of the accumulative sum of the values along the integration pathway), the algorithm reflects the global properties of the image.

Note that the BOWN computation is implemented in the DISC model because it is tuned to construct surfaces. The surface-construction approach is advantageous even when clear illusory contours are not observed. Note that only if the edge of the surface is completely sharp (i.e., a step-wise edge) can one draw a contour line there. A Kanizsa figure with misaligned Pac Man shapes (see Figure 1F) indicates, however, that the illusory central surface still persists without completely sharp edges. In fact, Stanley and Rubin (2003) clearly showed that subjects perceived smoother edges with the misaligned Kanizsa figure, whereas the salient region (the illusory surface) remained present. In contrast to the modal completion seen in the subjective contours of the Kanizsa figure, there is a related phenomenon, called *amodal completion*, which does not involve articulate contours either. The term *amodal completion* indicates perception with no clear sensory qualities (i.e., amodally completed contours are more believed to be there without being seen as such; see Michotte, Thines, & Crabbe, 1964, for the notions of modal and amodal completions; see Wagemans, van Lier, & Scholl, 2006, for a recent review). A typical example can be observed in Figure 1G. In this figure, the surrounding four black



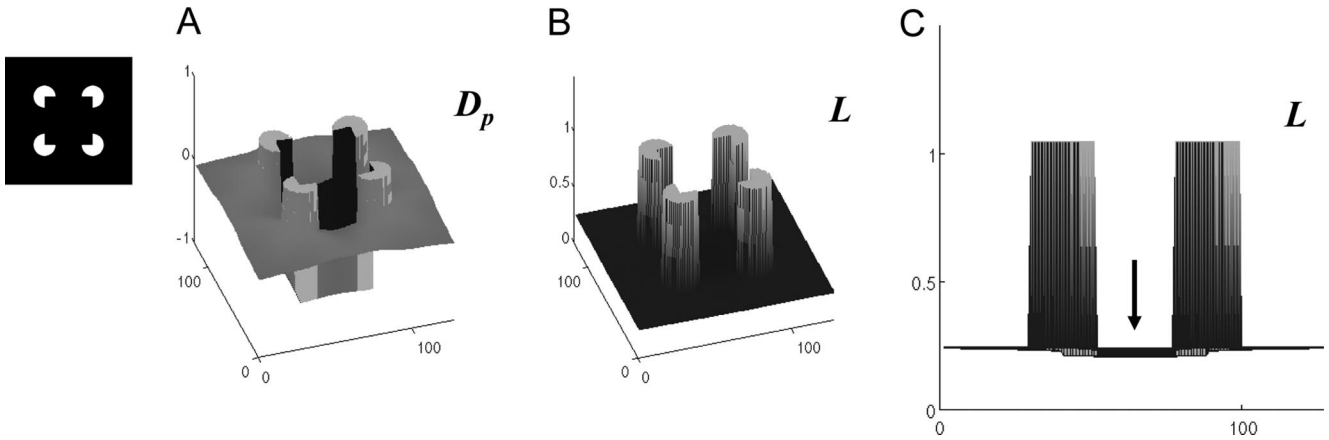


Figure 15. The Kanizsa figure with a white-on-black configuration (A, insets) creates a polarity-depth map (A) that is a mirror image of the one from black-on-white configuration because of the reversed polarity maps of the figure. This is reflected in the modification factor and, hence, in the lightness maps (B). As a result, the central square obtains a smaller lightness value than the surrounding area, which is shown in the side view of the lightness map (C, arrow).

objects with the arbitrary shapes are perceived to be occluded by the central square. Although it is not possible to perceive explicit, definitive shapes of the contours of the surrounding objects, we nevertheless perceive the continuation of the surface behind the square (see also van Lier, 1999). One can even imagine an object with a gradual change of color, such as a 2-D Gaussian function, behind a square in which case no (physical or subjective) contours of the object are perceived at all. The surface construction ap-

proach, in general terms, corresponds well with these observations. The aim of the completion in the visual system is probably not to draw lines per se but to construct surfaces in the ambiguous regions. The edges of the surfaces can be smooth as well as sharp, and the illusory contour phenomenon (with rectangular sharp edges) results only when particular conditions are met.

Another point that should be discussed here is that illusory contour and the illusory brightness have often been considered to be caused by

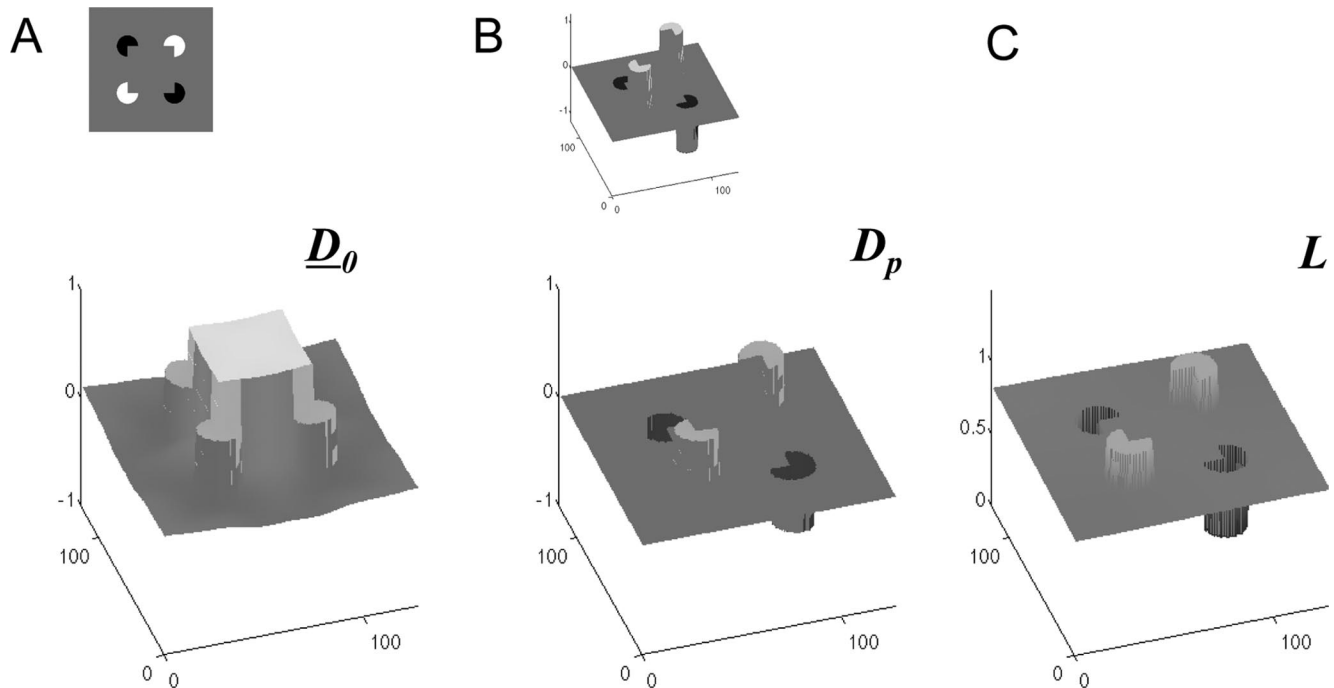


Figure 16. The depth (A), the polarity depth (B), and the lightness (C) maps from the midgray (diagonal) figure. Inset of B: The polarity map. No difference of lightness occurs in the central square compared with the surrounding areas.

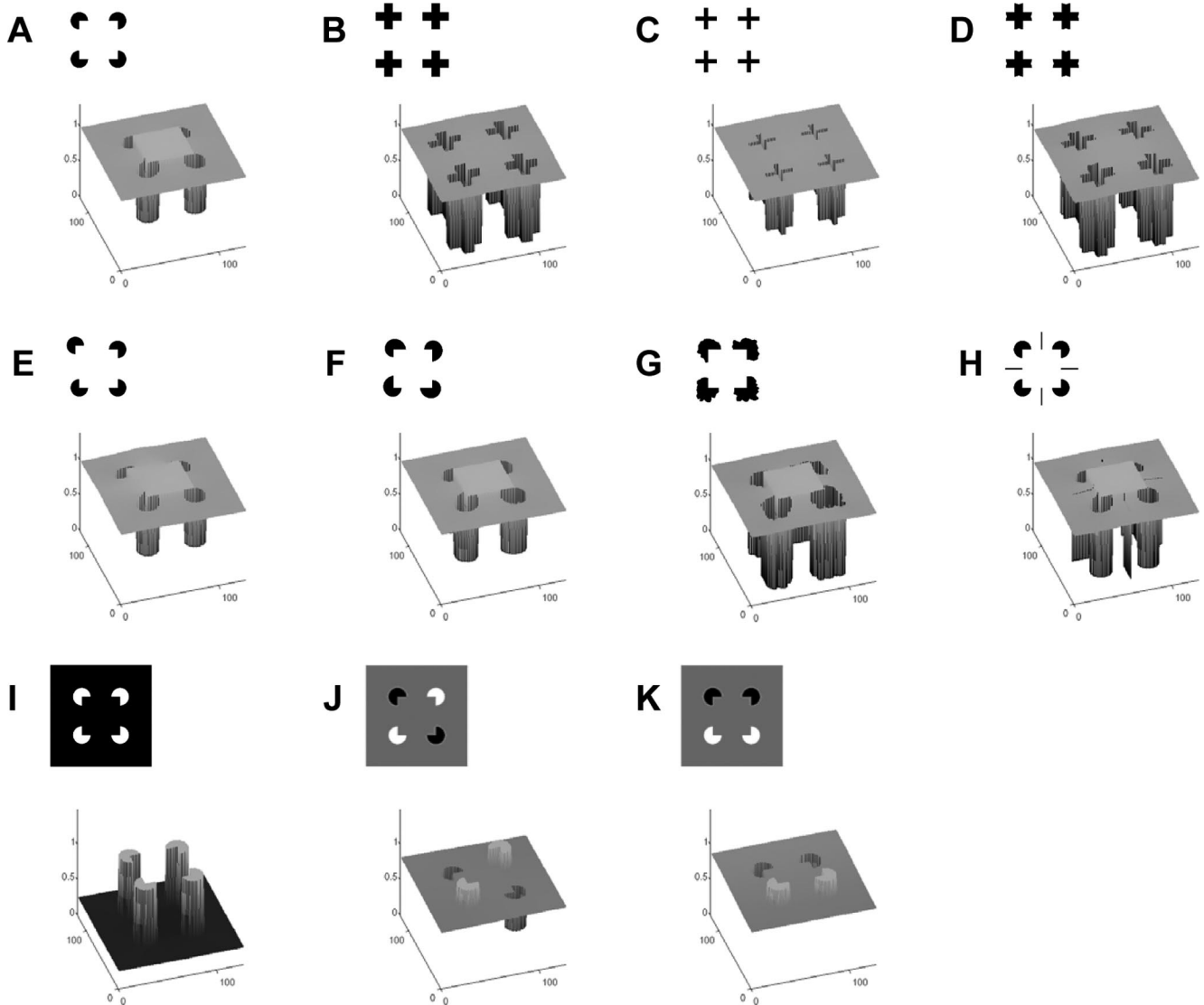


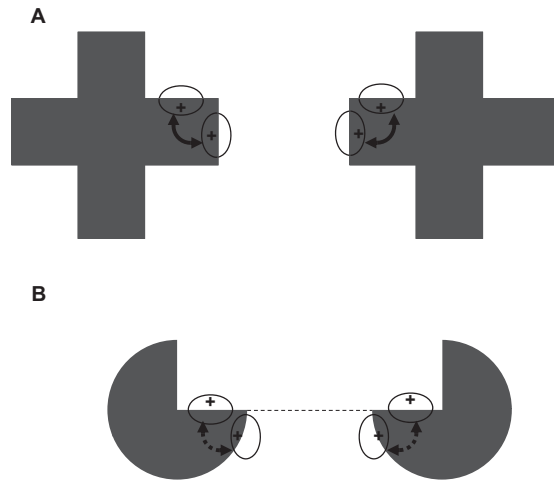
Figure 17. Lightness maps from all the figures used in this article.

two separate mechanisms (Dresp & Bonnet, 1991; Dresp, Lorenceau, & Bonnet, 1990; see also Rogers-Ramachandran & Ramachandran, 1998). The BCS/FCS model (Grossberg & Mingolla, 1985b) and the FACADE theory (Grossberg, 1994) actually realize this view: After borderline completion is done, signals that carry certain properties of a surface are filled in within the enclosure of the completed borderlines through a process of anisotropic diffusion. Note that the DISC model also has some two-stage processes. However, they are (a) the depth surface completion and (b) the depth-lightness linkage. In this framework, the suggested two-stage processes mentioned above can be reinterpreted as follows. The edges of the surfaces are perceived first but in terms of the depth domain, and this computation has no photometric qualities. Only after the depth-lightness linkage, the illusion is transformed into photometric terms, and this leads to the illusory brightness. Once the two-step process of the creation of the illusion is understood in this way (the perception of the figures in Figures 1J and 1K), the examples that show illusory contour but not

the illusory brightness can be interpreted as follows. Because in the depth domain, a central surface that is closer to the viewer is created, and because this central square in the depth domain possesses sharp edges surrounding it, the illusory contour phenomenon is observed in these figures. The second process (the brightness modification), on the other hand, does not occur because the overall contrast between the objects and the central area is nullified in these figures. In other words, illusory contours are perceived only in the depth domain in these figures. In fact, one cannot tell the achromatic color (photometric term) of the illusory contours in these figures.

### Neurophysiological Counterparts of the Differentiation-Integration Mechanism

The DISC model constructs a BOWN map based on local occlusion cues (junctions). It has been shown that some neurons in visual cortex are tuned to respond to junctions. L-junction detec-



*Figure 18.* Comparison of borderline-ownership (BOWN) signals in non-illusory (A) and illusory (B) figures. Only two bottom objects are shown for each figure. Comparison of the BOWN signals attached to the side junction of the objects (curved double arrows) indicates the clear difference between these two figures. With the four-crosses figure (A), the BOWN signals consistently indicate that the inside of the objects are the owners of the borderlines. The integration of the BOWN map therefore simply results in the four independent objects in the depth map. With the Kanizsa figure (B), on the other hand, the illusory T-junction condition emerges. As a result, free-space BOWN signals are developed.

tors have been found in V2 (Hegd  & Van Essen, 2003) as well as in the intermediate level of the ventral (object-recognition) pathway, V4 (Pasupathy & Connor, 1999). L-junction-sensitive neurons in V4 are unimodally tuned (i.e., they are sensitive to L-junctions with a particular angle in a particular orientation; Pasupathy & Connor, 1999). This suggests that the neural system is capable of preserving the direction of changes in depth based on L-junction signals. Note that these preserved directionalities of changes are critically required to construct the differentiated signals in the differentiation–integration approach of the DISC model.

Comparing this approach with the borderline-completion-and-filling-in approach brings up an intriguing neurophysiological issue, namely, whether the neurons in visual cortex that respond at the locations of subjective contours are mere borderline detectors extending and completing contours or whether they represent differentiated signals representing the central subjective surface. It has been shown that some neurons in V2 (and some in V1) are active at locations of subjective contours (Ffytche & Zeki, 1996; Grosf, Shapley, & Hawken, 1993; Lee & Nguyen, 2001; Peterhans & von der Heydt, 1989; Sheth, Sharma, Rao, & Sur, 1996; von der Heydt, Peterhans, & Baumgartner, 1984). This response property of neurons in low-level visual cortex has led to the idea of borderline completion by these neurons, and this idea has been implemented in some computational models (as “bipole cells,” Grossberg, 1994; Grossberg & Mingolla, 1985a, 1985b; as “grouping channel,” Heitger et al., 1998; Peterhans et al., 1986). If, however, the borderline-completion approach requires further elaboration to explain the Kanizsa illusions, as discussed in the introduction, how should these low-level neural activities be interpreted? The existence of two types of neurons, borderline detectors and edge detectors, reported by Hubel and Wiesel (1962), is of importance here. The fundamental difference between them is that

edge detectors not only respond to borderlines but are also sensitive to the polarity of the contrast. Furthermore, some neurons in V2 and V4 reflect BOWNs (Zhou et al., 2000). Qiu and von der Heydt (2005) indeed showed that the border ownership, as signaled from 2-D image cues, is often consistent with surfaces being closer to the viewer, as signaled by 3-D stereo cues. Therefore, these neurons are very likely the ones that signal the polarity of the difference in the depth domain. In other words, among the neurons which show responses at borderlines, some of them may indicate not only the existence of changes in the luminance or depth domains but also the polarity of changes in these domains. The polarity of change can be signaled in terms of which of the two abutting surfaces possesses a higher luminance intensity (in the photometric domain) or BOWN (in the depth domain), which in turn defines the properties of the surfaces in these domains. An important question is then whether the subjective-contour-sensitive neurons mentioned previously should be categorized as borderline detectors or edge detectors. Although they can be regarded as borderline detectors to group collinearly aligned contour segments, it is equally possible that the activities of these neurons represent the existence of the subjectively perceived central surface that occludes the surrounding objects, but only in a differentiated form, and hence, they are the BOWN sensitive neurons.

The activities of these differentiator type neurons at the locations of the illusory contours may or may not be linked to the feedback process of the computation of depth in the DISC model. In the feedback scheme, there is no depth difference at the location of the illusory contours in the original Kanizsa figure at first. Only when the depth perception of the image is refined by the feedback iteration do the illusory edge signals start to develop. In other words, the first development of the surface triggers the further development of their edges. Correspondingly, it is possible that the activities of these subjective-contour-sensitive neurons are enhanced when the central surface is reconstructed at the higher level and top-down feedback connections to these neurons are activated (Lamme & Roelfsema, 2000; Lamme, Super, & Spekreijse, 1998). Stanley and Rubin (2003) argued, based on their fMRI and psychophysical data, that the visual system first detects the “salient regions” at the lateral occipital complex (LOC), which, through the top-down feedback loop to V1/V2 regions, modifies the contour-sensitive activities at the level that, in turn, refines the perception of the surface with sharp edges. The involvement of feedback processes is also suggested by Lee and Nguyen’s (2001) finding that the illusory responses in V2 neurons actually occur earlier than those in V1 neurons. However, the interactions of neurons through the meshwork of the horizontal connections at the lower level visual cortex also should not be disregarded. In fact, the BOWN computation in the DISC model determines the final BOWN signals after interactions between local signals reflecting the distributions of the junctions and the borderlines. It is possible that this corresponds to the interactions of neurons by the horizontal connections. In fact, it has been shown that when the Kanizsa figure is presented so that the horizontal illusory contours cross the vertical meridian of the visual field, the illusory perception is significantly impaired (Pillow & Rubin, 2002), which suggests that the interactions among neurons in V1/V2 areas (where the communication between the left and the right hemisphere of the visual field is not established) are essential in creating the illusions. Indeed, it is entirely plausible that the first bottom-up processing extracts a crude representation of the shape of the surface that is refined by an iterative feedback process. This iterative process could be especially required in textured

surfaces. The possible involvement of the feedback connections in the subjective-contour-sensitive neural activities for the Kanizsa figure and their relationships with signal processing through the horizontal connections is, in our view, a quite vital component to fully understand the computational mechanisms involved. They can probably be shown clearly only by future neurophysiological experiments, for instance, making use of techniques to block top-down influence such as temporary cooling (e.g., Hupé et al., 1998, 2001).

Related to the issue of whether or not top-down effects are involved in the Kanizsa illusions, but somewhat more general, is the question as to where and how the surface creation (or filling in) happens. The current evidence on this issue is inconclusive (e.g., Cornelissen, Wade, Vladusich, Dougherty, & Wandell, 2006; Sasaki & Watanabe, 2004; see also Komatsu, 2006, for review), and at this point, we can only speculate on the basis of the limited available information. The BOWN-sensitive neurons in V1, V2, and V4 mentioned previously (Zhou et al., 2000) showed intriguing properties. When the C-shaped figure was presented, the sensitivity of these neurons showed quantitative differences in these areas. In V1, none of the BOWN-sensitive neurons showed responses to the C-shape, consistent with our perception, whereas the percentage was somewhat higher in V2 (17%) and much higher in V4 (43%). This suggests that, whereas BOWNs of simple figures such as disks and rectangles are already established at the lower level, for the more complicated shapes, the higher level neural processes are necessary. This interpretation is consistent with the DISC model. We hypothesized that BOWN signals carry the depth information implicitly as a form of differentiated signals, and at the higher level, the information becomes explicit when the integration of the BOWN signals creates surfaces. Mendola et al. (1999), Murray et al. (2002), and Stanley and Rubin (2003) showed, using fMRI, that areas in LOC are more sensitive to Kanizsa-type figures compared with nonillusory figures. Huxlin, Saunders, Marchionini, Pham, and Merigan (2000) showed that a lesion of area IT in monkeys (equivalent to LOC in human) caused an impairment of illusory perception in the Kanizsa figure. Huxlin et al.'s work recently has been complemented by human neuropsychology highlighting that the perception of a patient with visual form agnosia following a lesion to LOC is uninfluenced by Kanizsa-like figures (de-Wit, Kentridge, & Milner, 2009). The area is therefore a candidate location for the 2-D integration of BOWN signals. Certainly it is possible that BOWN computation completes at V4 or higher, whereas the creation of the surfaces (based on the BOWN signals) is performed at LOC. It is even possible that, especially with complex shapes, feedback iteration is necessary to establish the perceived depth order (Kogo, Galli, et al., 2010). We agree that there are still many unresolved issues regarding the neural implementation of the algorithms in the DISC model, but at least what we propose is not contradicted by the currently available evidence. Further research into the neural mechanisms of surface filling in is clearly needed.

In sum, we hypothesized that some neurons in the low/intermediate-level visual cortex act as differentiators that preserve the information of the relationship between the neighboring surfaces and that, at the higher level, the visual system is able to reconstruct the information about the surface properties by integrating them. This approach proved to be quite powerful in reproducing our perception which supports its psychological plausibility. Because it is possible that it also happens in the neuronal architecture of the brain, the process of the integration of differentiated signals can be viewed as a way of bridging the gap between the microscopic properties (corre-

sponding to the individual low/intermediate-level neural activities) and the macroscopic properties (corresponding to the neural activities representing the emergent higher level properties). It is possible that the reported neurophysiological responses at the location of subjective contours reflect the detected surfaces in the form of differentiated signals, probably via mechanisms to detect macroscopic properties, such as the interactions by the horizontal connections and the top-down feedback projections.

### Novelty, Current Limitations, and Possible Further Extensions of the DISC Model

In this article, we focused on fundamental issues such as the BOWN computation, surface construction, and the depth–lightness interaction to explain the Kanizsa illusions. Based on the principles and the framework of the model, further developments of the DISC model could be considered as the next step to address some of the limitations of the choices that are made in the current model. In this section, some possible improvements of the model in future research are discussed. Before presenting these discussions, however, we first summarize the difference between the model and the previous models to clarify the novelty of the DISC model.

Many models that showed correct responses to the original Kanizsa figure were based on the borderline-completion scheme driven by the collinear alignment of the contours (Grossberg, 1994; Grossberg & Mingolla, 1985b; Grossberg & Pessoa, 1998; Heitger et al., 1998; Peterhans & Heitger, 2001; Peterhans et al., 1986; Ullman, 1977). As noted earlier, there are two illusory phenomena in the Kanizsa figure: the illusory contours and the illusory brightness of the central area. Hence, to reproduce both phenomena, borderline-completion models must also incorporate a filling-in process to construct surfaces. This process was missing in most of the models except the BCS-FCS model (Grossberg & Mingolla, 1985a, 1985b) and the FACADE theory (Grossberg, 1994). The surface-completion models with the depth computation (Geiger et al., 1998; Kumaran et al., 1996; Sajda & Finkel, 1992; Sajda & Finkel, 1995; Williams & Jacobs, 1997), on the other hand, create illusory surfaces with surrounding (illusory) contours in depth domain, but they do not incorporate the process to reproduce the illusory brightness perception. On the contrary, the models and theories that implemented the differentiation–integration approach (Arend, 1973; Arend & Goldstein, 1987; Gilchrist et al., 1983; Horn, 1974; Land & McCann, 1971; Ross & Pessoa, 2000; Wallach, 1948) applied the approach only to the lightness computation. The depth perception and hence the influence of the depth to the lightness perception are not implemented in these models. In other words, all these models reproduce the Kanizsa illusory phenomena only partially.

The exception is the BCS-FCS model (Grossberg & Mingolla, 1985a, 1985b) and the FACADE theory (Grossberg, 1994, 1997). FACADE theory has been developed over decades, building further on the BCS-FCS model (Grossberg & Mingolla, 1985a, 1985b), and the model is capable of reproducing a wide range of perceptual phenomena. The model is complex. However, the basis of this complex model is the “bipole cell” that detects the collinear alignment of borderlines. In this model, the depth order of surfaces becomes evident when the filling-in process is performed later and the surface properties are captured in the enclosed borderlines. This points out the major difference between this approach and

ours. In the FACADE model, to create a surface, an enclosed borderline is created and then the surface properties are filled in. The filling-in process is performed by diffusion, and hence, the enclosed borderline is needed to trap the feature inside. Our approach is different in the sense that not the borderlines but the *ownership* of the borderlines (BOWN) is completed, and it is regarded as the differentiated signal of a surface. The integration of the BOWN signals therefore makes the surface explicit. In other words, the surface feature is not something that is brought in from another separate channel; rather, the BOWN signal itself carries the information of the surface implicitly.

Because the DISC model completes the BOWN signals but not the borderlines, it is able to distinguish between the illusory Kanizsa figure and the nonillusory variation (the four crosses figure). The collinearity detection and the borderline-completion approach in the FACADE theory, on the other hand, may encounter difficulties in distinguishing the illusory and the nonillusory figures, as pointed out earlier. Note that the four crosses figure is a classic, well-known example of nonillusory variations that even Kanizsa himself pointed out. For this kind of clear difference in perception by such a simple modification of the shape of the inducers, there should be a simple, straightforward explanation. The DISC model offers the explanation that the difference in the configurations of the surrounding objects, namely the inwardly skewed shapes (Pac Men shapes) versus the symmetric shapes (four crosses), causes a difference in the BOWN computation after the global interaction and that this difference essentially determines whether the central area becomes the owner of the surrounding borderlines or not. It seems that this indicates the advantage of constructing the BOWN signals instead of completing borderline signals.

The DISC model operates all the way from borderline and junction detections, BOWN computation, anchoring, and depth-lightness interaction, with the same differentiation-integration approach applied in both the lightness and the depth computation, and as a result, it is capable of reproducing the known properties of the completion phenomena, such as the illusory contours, the illusory brightness, and the depth stratification. In this way, it is able to explain how the illusory perception depends on factors such as the contrast polarity and the depth perception, and it is able to reproduce the various perceptions of the illusory and the nonillusory figures. As mentioned, however, further improvements to the model are possible, which are discussed in the next paragraphs.

The first issue to be discussed concerns the size of the junctions and the resolution at which they are detected. In the DISC model, the junctions play an important role. In fact, N. Rubin (2001) showed that the illusions disappear when the side junctions of the Pac Man shapes are masked. In addition, a slight modification of an L-junction to a T- or X-junction by extending one or two of its arms drastically changes our perception of the image. Furthermore, Pasupathy and Connor (1999) showed that some neurons in V4 are tuned to a junction of a specific orientation and a specific angle of the opening (see also Hegd  & Van Essen, 2003, for V2 neurons responding to junctions). Therefore, it is possible that the specific attributes of junctions indeed play some role in our processing of images. Whereas, in the DISC model, the amplitudes of the junctions are all normalized (binary signals), it is possible that the size of junctions has some perceptual meaning (McDermott, 2004). Through the BOWN computation of the model, however, if a

junction point occurs in the context of two straight borderlines that extend spatially without being interrupted by inconsistent junctions, the BOWN signal from the junction (indicating the inside area being the owner) would spread along the borderlines. Therefore, it is also possible that the size effect is the result of (the spread of) BOWN signals.

There is, however, an issue regarding the resolution at which the junctions are detected. It should be pointed out that, by detecting L-junctions of all sizes in the DISC model, the L-junction detection essentially works as concavity detection because curved borderlines are in fact treated as sequences of small junctions. In other words, the essential part of the BOWN computation in the DISC model is that it reflects the concavity of the borderlines. However, the size of the junction could have a particular meaning in our visual system as follows. Let us first consider a line-like object. A line, in a strict mathematical sense, does not exist in the real world. When we say "line" in our daily life, as well as in perception research, we actually mean a rectangle with extremely narrow width. It is important to note that, because of the size (width) of the receptive field of neurons in the retina, LGN, and lower visual cortex, when this line-like object becomes thin, the responses of the neurons are reduced. This response reduction is probably even stronger with neurons that respond to junctions. It is quite likely that when the size of a junction becomes very small, the response of neurons will eventually drop below threshold. In both articles mentioned previously (Hegd  & Van Essen, 2003; Pasupathy & Connor, 1999), the size of the junction has not been changed to smaller than the size of their receptive fields. To our knowledge, therefore, the information regarding how small a junction can become before neurons stop responding is missing. In computational models, the smallest width a line-like object can have is one pixel. In the current DISC model, this is the limit of the resolution as well. In other words, if a line-like object is as thin as one pixel, the borderlines surrounding the object still are detected and the corners are considered junctions. In this way, the BOWN computation indicates that the inside of the borderline is the owner, and the depth computation indicates that it is indeed an object. However, it may be possible that objects are treated differently in the visual system when they are very thin, in effect, distinguishing lines from surfaces. For simplification purposes, we did not implement this surface-line discrimination in the current DISC model. However, how a line is dealt with in real neural computations is an intriguing issue, and further investigation is necessary. As a future development of the model, an implementation of multiscale convolution filters that are elongated and oriented in multiple directions will be considered. In that implementation, the surface-line discrimination could be incorporated by setting the lower limit of the junction detection. In this way, the model will become more robust, not only in detecting junctions with various angles and sizes, even with various degrees of noise, but also in treating as a separate category the line-like objects that may have special perceptual meanings.

A second limitation of the current DISC model is that the role of T- and X-junctions in the perception of transparency and neon spreading has not been dealt with. These perceptions are based on the difference of the combination of the contrast polarities and the (T- and X-) junctions (Anderson, 1997, but for further discussion, see Anderson, 2001; Grossberg, 2001; Howe, 2001; Ross & Pessoa, 2000; Todorovic, 1997, 2001). Grossberg and Yazdanbakhsh



(2005) showed how the combination of the contrast polarities and the (T- and X-) junctions can be reflected into their model. Clearly, in these phenomena, the two borderlines that constitute the junctions are perceived to be located at different depths. In their model, the “border pruning” process separates these borderlines: The “stronger” borderline is kept intact while the other is “pruned.” After the filling-in process, the area enclosed by the intact borderline becomes closer to the viewer, while the area enclosed by the pruned borderline becomes less close to the viewer. To reproduce these phenomena in the DISC model, on the other hand, two things have to be considered. The BOWN-computation process has to give correct signals on the borderlines of the transparent, neon-colored, and opaque surfaces. Furthermore, in the process of creating the depth map, the model has to be able to assign two different depth values at the same location where the transparent or the neon-colored surface is perceived. Note that the DISC model computes the polarity-depth map where the information of depth and contrast polarity is combined. This could be the place where the combination of the contrast polarity at the junctions can determine the transparency or neon-color perception. Further investigation is necessary to create these different perceptions of surfaces in the model.

Finally, the model should ultimately be able to compute both lightness (perceived reflectance) and (perceived) illuminance from the luminance input in general (underconstrained inverse optics problem). In the Kanizsa figure, we hypothesized that the illuminance is perceived to be equal over the entire image. Therefore, the computation in the current model is focused only on lightness perception. In more general conditions where illuminance varies over a surface, however, the computation of illuminance is essential to determine lightness. This has been shown quite clearly in the famous checkerboard image (Adelson, 1995). In that figure, the illuminance of the surfaces in the shadow and under the light needs to be known to reproduce the perceived lightness of the surfaces. Such computation of the illuminance may be possible on the basis of the value of the luminance and its gradient. In Retinex theory, a gradual change of luminance is considered to be due to illuminance and is eliminated by thresholding the low-gradient values to compute the lightness perception. In other words, the perceived illuminance may be computed by detecting the gradually changing (low-gradient) luminance. However, the computation of illuminance and lightness is a complex issue. Further investigation is necessary to find out how the illuminance computation can be done and how lightness perception under uneven illuminance can be reproduced in the DISC model.

### Generalizations and Applications of the DISC Model

After all these analyses and arguments, one might ask what the Kanizsa illusion really is. From the DISC model, we can derive a plausible answer to this question. Our model explains the creation of the illusion by depth perception and by its influence on lightness perception. Once the central square is perceived as a surface higher than the surround in the depth map, the area becomes brighter (in the black-on-white configuration) or darker (in the white-on-black configuration). This phenomenon is quite similar to lightness perception of figures with center-surround organization, where the central area is perceived as brighter or darker depending on whether the reflectance of the surrounding area is darker or lighter

than the central area, respectively. In other words, the contrast between the central figure and the surrounding background is always enhanced. In Kanizsa figures, once the central area is perceived as a separate surface (hence perceived as a “figure”), this contrast enhancement occurs.

Prazdny (1983) pointed out the remaining illusion in the midgray variation of Kanizsa figure (see Figures 1J and 1K) and argued that the Kanizsa illusion is not a simultaneous contrast effect. The argument made above yields a different view. It states that one aspect of the Kanizsa illusion, the brightness illusion, is indeed the simultaneous contrast effect. It also states, however, that the brightness illusion happens because the central area is segmented because of the depth difference between the area and the background. The depth difference influences the lightness value by enhancing the contrast between them. With the midgray variation figures, then, the brightness illusion does not happen simply because there is no net contrast to enhance. The remaining Kanizsa illusion in the figure that Prazdny pointed out is the perception of the segmented central area in the depth domain that causes the other aspect of the Kanizsa illusion, the perception of illusory contours. If this comparison is correct, it may indicate that the Kanizsa illusion is a center-surround effect of contrast enhancement based on the illusory segmentation of the figure from the background, as opposed to the more conventional figure-ground segmentation defined by physically existing contours.

Completion is one of the fundamentally important operations of human vision, of which the Kanizsa figure is only a representative. The phenomenon of completion, or filling in (a term generally used for subjective surface construction), is observed in many instances, suggesting that this is a function that is regularly at work in the visual system. Images with physically missing parts occur frequently because of occlusions, the blind spot, or retinal vasculatures and also for patients with pathological scotomas. We are able to fill in the missing information quite naturally and quickly. The phenomenon is in fact experimentally observed in conditions of artificial retinal stabilization (Yarbus, 1967) or retinal stabilization in peripheral vision (Riggs, Ratliff, Cornsweet, & Cornsweet, 1953). The identification of the mechanisms underlying this phenomenon is therefore essential in understanding human vision. The current DISC model established a new surface-completion approach with differentiation-integration to reproduce the phenomenon. The thorough investigation with the variations of the Kanizsa figure indicated the plausibility of this approach. Furthermore, the differentiation-integration realizes the relationality of our perception (Koffka, 1935; Wallach, 1948). In addition, the influence of depth perception very likely underlies many known perceptual phenomena (Coren, 1972). The model also offers a framework to understand the interactions between depth, lightness, and brightness. We believe, therefore, that the DISC model can be applied to a wide range of perceptual phenomena. As described in the following, we have been working on a few applications of the model which will be developed further in the near future.

In the process of integration, one degree of freedom is inevitably introduced (i.e., the offset value, as indicated in Equation 2,  $C_L$  and  $C_D$ ). We believe that this constraint is the key to explaining the anchoring phenomenon (Gilchrist et al., 1999). In the current model, the area rule and the highest value rule are applied to the depth map and to the (primary) lightness map, respectively, as explained earlier. It is known that the anchoring phenomenon in lightness perception is

known to deviate from the highest luminance rule (to anchor the area with the highest luminance to perceived white) when the surface area of the lighter area is smaller than the one of the darker area (Li & Gilchrist, 1999). The DISC model is able to reproduce this phenomenon, reflecting the figure–ground segmentation of these areas in the depth computation and the anchoring rules mentioned above to determine the offset values. That is, when a surface area becomes smaller, it is more likely to be perceived as a figure (i.e., higher in the depth map than the rest of the image; Kogo, Van Gool, & Wagemans, 2010). This smaller area therefore achieves enhancement of the contrast that causes, in the above mentioned configuration, the super-white perception of the area (lightness perception of the area being lighter than white).

The figure–ground segmentation as a result of the computation in the depth domain plays a key role in the DISC model. The face-or-vase illusion (E. Rubin, 1921) is a well known example of the multistable perception of human vision where the figure–ground relationship of the face and the vase areas keep changing over time. We implemented a stochastic property into the signals from the occlusion detectors and BOWN signals that, in turn, influences the computation of the depth of separated surfaces (Kogo, Galli, et al., 2010). Furthermore, a positive feedback system was implemented in which the depth determination at the higher level gives a strong bias to the occlusion and BOWN signals. In addition, properties of adaptation and its recovery were added to the signals at the higher level. By doing so, this model showed the alteration of the borderline ownership by the surfaces with known psychophysical characteristics. Furthermore, it succeeded in mimicking the prolonged alteration by intermittent presentation of the stimuli as reported by Leopold, Wilke, Maier, and Logothetis (2002).

This approach of feedback can be further generalized into a dynamic feedback model: The development of the surfaces influences via a top-down connection the lower level computation and the 2-D integration process of these lower level signals, in turn, helps to develop the surface further. This suggests a context-sensitive dynamic feedback system where the response properties at the low level are constantly regulated by top-down projection to detect the macroscopic properties more clearly. We are developing a model in which the BOWN signals not only reflect the horizontal interactions but also reflect the global configuration of the image via the top-down projections and in which the figure–ground status of an area determined after the surface construction in the depth domain gives a strong bias to the BOWN computation.

In summary, we believe that the approaches taken in the DISC model—surface completion using differentiation–integration and depth–lightness linkage—are closely related to the fundamental mechanisms of the visual system by which subjectivities are brought into perception. We developed the details of the theories and the algorithms by testing them with the variations of the Kanizsa figures. The model is now being investigated further to test its applicability to a wider range of perceptual phenomena.

## References

- Adelson, E. H. (1993, December 24). Perceptual organization and the judgment of brightness. *Science*, *262*, 2042–2044.
- Adelson, E. H. (1995). Checkershadow illusion. Retrieved from [http://web.mit.edu/persci/people/adelson/checkershadow\\_illusion.html](http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html)
- Adelson, E. H., & Pentland, A. P. (1996). The perception of shading and reflectance. In D. C. Knill & W. Richard (Eds.), *Perception as Bayesian inference* (pp. 409–423). New York, NY: Cambridge University Press.
- Allen, R. E. (Ed.). (1990). *The concise Oxford dictionary of current English* (8th ed.). Oxford, England: Clarendon Press.
- Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, *26*, 419–453.
- Anderson, B. L. (2001). Contrasting theories of White's illusion. *Perception*, *30*, 1499–1501.
- Arend, L. E. (1973). Spatial differential and integral operations in human vision: Implications of stabilized retinal image fading. *Psychological Review*, *80*, 374–395.
- Arend, L. E., & Goldstein, R. (1987). Lightness models, gradient illusions, and curl. *Perception & Psychophysics*, *42*, 65–80.
- Barrow, H. G., & Tenenbaum, J. (1978). Recovering intrinsic scene characteristics from images. In A. R. Hanson & E. M. Riseman (Eds.), *Computer vision systems* (pp. 3–26). Orlando, FL: Academic Press.
- Bergstrom, S. S. (1977). Common and relative components of reflected light as information about illumination, color, and 3-dimensional form of objects. *Scandinavian Journal of Psychology*, *18*, 180–186.
- Bradley, D. R., & Dumais, S. T. (1984). The effects of illumination level and retinal size on the depth stratification of subjective contour figures. *Perception*, *13*, 155–164.
- Brussell, E. M., Stober, S. R., & Bodinger, D. M. (1977). Sensory information and subjective contour. *American Journal of Psychology*, *90*, 145–156.
- Clairbourn, J. (1992). *Earth soundings analysis: Processing versus inversion (PVI)*. Oxford, England: Blackwell.
- Coren, S. (1972). Subjective contours and apparent depth. *Psychological Review*, *79*, 359–367.
- Coren, S., & Porac, C. (1983). Subjective contours and apparent depth: A direct test. *Perception & Psychophysics*, *33*, 197–200.
- Cornelissen, F. W., Wade, A. R., Vladusich, T., Dougherty, R. F., & Wandell, B. A. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *Journal of Neuroscience*, *26*, 3634–3641.
- Dalby, T. A., Saillant, M. L., & Wooten, B. R. (1995). The relation of lightness and stereoscopic depth in a simple viewing situation. *Perception & Psychophysics*, *57*, 318–332.
- de Weert, C. M. M. (1979). Color contours and stereopsis. *Vision Research*, *19*, 555–564.
- de Weert, C. M. M. (1983). The role of colors in the formation of subjective contours. *Psychological Research*, *45*, 117–134.
- de-Wit, L. H., Kentridge, R. W., & Milner, A. D. (2009). Shape processing area LO and illusory contours. *Perception*, *38*, 1260–1263.
- Dosher, B. A., Sperling, G., & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research*, *26*, 973–990.
- Dresp, B., & Bonnet, C. (1991). Psychophysical evidence for low-level processing of illusory contours and surfaces in the Kanizsa square. *Vision Research*, *31*, 1813–1817.
- Dresp, B., Lorenceau, J., & Bonnet, C. (1990). Apparent brightness enhancement in the Kanizsa square with and without illusory contour formation. *Perception*, *19*, 483–489.
- Ffytche, D. H., & Zeki, S. (1996). Brain activity related to the perception of illusory contours. *NeuroImage*, *3*, 104–108.
- Geiger, D., Pao, H., & Rubin, N. (1998). Salient and multiple illusory surfaces. *Computer Vision and Pattern Recognition*, 1998, 118–124.
- Gilchrist, A. (1977, January 14). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*, 185–187.
- Gilchrist, A. (1979). Perception of surface blacks and whites. *Scientific American*, *240*, 112–122.
- Gilchrist, A. (1980). When does perceived lightness depend on perceived spatial arrangement. *Perception & Psychophysics*, *28*, 527–538.

- Gilchrist, A., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, *33*, 425–436.
- Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X. J., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*, 795–834.
- Gillam, B., & Nakayama, K. (2002). Subjective contours at line terminations depend on scene layout analysis, not image processing. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 43–53.
- Gregory, R. L., & Harris, J. P. (1974). Illusory contours and stereo depth. *Perception & Psychophysics*, *15*, 411–416.
- Grosf, D. H., Shapley, R. M., & Hawken, M. J. (1993, October 7). Macaque-V1 neurons can signal illusory contours. *Nature*, *365*, 550–552.
- Grossberg, S. (1994). 3-D vision and figure ground separation by visual cortex. *Perception & Psychophysics*, *55*, 48–120.
- Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, *104*, 618–658.
- Grossberg, S. (2001). Theory versus speculation in visual perception. *Perception*, *30*, 1505–1507.
- Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*, 173–211.
- Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics*, *38*, 141–171.
- Grossberg, S., & Pessoa, L. (1998). Texture segregation, surface representation and figure-ground separation. *Vision Research*, *38*, 2657–2684.
- Grossberg, S., & Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research*, *45*, 1725–1743.
- Halpern, D. F. (1981). The determinants of illusory-contour perception. *Perception*, *10*, 199–213.
- Hegd , J., & Van Essen, D. C. (2003). Strategies of shape representation in macaque visual area V2. *Visual Neuroscience*, *20*, 313–328.
- Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & Kubler, O. (1998). Simulation of neural contour mechanisms: Representing anomalous contours. *Image and Vision Computing*, *16*, 407–421.
- Horn, B. (1974). Determining lightness from image. *Computer Graphics and Image Processing*, *3*, 277–299.
- Howe, P. D. (2001). A comment on the Anderson (1997), the Todorovic (1997), and the Ross and Pessoa (2000) explanations of White’s effect. *Perception*, *30*, 1023–1026.
- Hubel, D. H. (1988). *Eye, brain, and vision* (Vol. 22). New York, NY: Freeman.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat’s striate cortex. *Journal of Physiology*, *148*, 574–591.
- 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.
- Hung, C. P., Ramsden, B. M., Chen, L. M., & Roe, A. W. (2001). Building surfaces from borders in Areas 17 and 18 of the cat. *Vision Research*, *41*, 1389–1407.
- Hup , J. M., James, A. C., Girard, P., Lomber, S. G., Payne, B. R., & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of Neurophysiology*, *85*, 134–145.
- Hup , J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998, August 20). 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.
- Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea [Quasi-perceptual borderlines in homogeneously stimulated fields]. *Rivista di Psicologia*, *49*, 7–30.
- Kelly, F., & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Perception & Psychophysics*, *62*, 1596–1618.
- Knill, D. C., & Kersten, D. (1991, May 16). Apparent surface curvature affects lightness perception. *Nature*, *351*, 228–230.
- Koffka, K. (1935). *Principles of Gestalt psychology*. New York, NY: Harcourt, Brace & World.
- Kogo, N., Galli, A., Van Gool, L., & Wagemans, J. (2010). *Switching dynamics of border ownership: A stochastic model for multi-stable perception*. Manuscript submitted for publication.
- Kogo, N., Van Gool, L., & Wagemans, J. (2010). *Linking depth to lightness and anchoring within the differentiation-integration formalism*. Manuscript submitted for publication.
- Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*, 220–231.
- Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterochromatic targets. *Journal of the Optical Society of America*, *53*, 741–744.
- Kumaran, K., Geiger, D., & Gurvits, L. (1996). Illusory surface perception and visual organization. *Network-Computation in Neural Systems*, *7*, 33–60.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579.
- Lamme, V. A., Super, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*, 529–535.
- Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, *61*, 1–11.
- Lee, T. S., & Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 1907–1911.
- Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*, 605–609.
- Leshner, G. W. (1995). Illusory contours: Toward a neurally based perceptual theory. *Psychonomic Bulletin & Review*, *2*, 279–321.
- Li, X. J., & Gilchrist, A. L. (1999). Relative area and relative luminance combine to anchor surface lightness values. *Perception & Psychophysics*, *61*, 771–785.
- Mallot, H. A. (1997). Spatial scale in stereo and shape from shading: Image input, mechanisms, and tasks. *Perception*, *26*, 1137–1146.
- McDermott, J. (2004). Psychophysics with junctions in real images. *Perception*, *33*, 1101–1127.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. H. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 8560–8572.
- Michotte, A., Thines, G., & Crabbe, G. (1964). *Amodal completion of perceptual structures*. Louvain, France: Universitaires de Louvain.
- Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational-problems. *Communications on Pure and Applied Mathematics*, *42*, 577–685.
- Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: Combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 5055–5073.
- Neumann, H., Pessoa, L., & Mingolla, E. (1998). A neural architecture of brightness perception: Nonlinear contrast detection and geometry-driven diffusion. *Image and Vision Computing*, *16*, 423–446.



- Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*, 2490–2502.
- Perona, P., & Malik, J. (1990). Scale-space and edge-detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 629–639.
- Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences*, *21*, 723–748.
- Peterhans, E., & Heitger, F. (2001). Simulation of neuronal responses defining depth order and contrast polarity at illusory contours in monkey area V2. *Journal of Computational Neuroscience*, *10*, 195–211.
- 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.
- Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In J. D. Pettigrew, K. J. Sanderson, & W. R. Levick (Eds.), *Visual neuroscience* (pp. 343–351). Cambridge, England: Cambridge University Press.
- Pillow, J., & Rubin, N. (2002). Perceptual completion across the vertical meridian and the role of early visual cortex. *Neuron*, *33*, 805–813.
- Prazdny, K. (1983). Illusory contours are not caused by simultaneous brightness contrast. *Perception & Psychophysics*, *34*, 403–404.
- Proesmans, M., & Van Gool, L. (1999). Grouping based on coupled diffusion maps. *Shape, Contour and Grouping in Computer Vision*, *1681*, 196–213.
- Qiu, F. T., & von der Heydt, R. (2005). Figure and ground in the visual cortex: V2 combines stereoscopic cues with gestalt rules. *Neuron*, *47*, 155–166.
- Riggs, L. A., Ratliff, F., Cornsweet, J. C., & Cornsweet, T. N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America*, *43*, 495–501.
- Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research*, *38*, 71–77.
- Ross, W. D., & Pessoa, L. (2000). Lightness from contrast: A selective integration model. *Perception & Psychophysics*, *62*, 1160–1181.
- Rubin, E. (1921). *Visuell wahrgenommene Figuren* [Visually perceived figures]. Copenhagen, Denmark: Glydenalske bogahandel.
- Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception*, *30*, 339–366.
- Sajda, P., & Finkel, L. H. (1992). Simulating biological vision with hybrid neural networks. *Simulation*, *59*, 47–55.
- Sajda, P., & Finkel, L. H. (1995). Intermediate-level visual representations and the construction of surface perception. *Journal of Cognitive Neuroscience*, *7*, 267–291.
- Sasaki, Y., & Watanabe, T. (2004). The primary visual cortex fills in color. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 18251–18256.
- Schirillo, J., Reeves, A., & Arend, L. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception & Psychophysics*, *48*, 82–90.
- Sheth, B. R., Sharma, J., Rao, S. C., & Sur, M. (1996, December 20). Orientation maps of subjective contours in visual cortex. *Science*, *274*, 2110–2115.
- Shipley, T. F., & Kellman, P. J. (1992). Strength of visual interpolation depends on the ratio of physically specified to total edge length. *Perception & Psychophysics*, *52*, 97–106.
- Stanley, D. A., & Rubin, N. (2003). fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron*, *37*, 323–331.
- Stevens, K. A. (1983). Evidence relating subjective contours and interpretations involving interposition. *Perception*, *12*, 491–500.
- Todorović, D. (1997). Lightness and junctions. *Perception*, *26*, 379–394.
- Todorović, D. (2001). Lightness, junctions, and depth. *Perception*, *30*, 1502–1505.
- Tomasi, C. (1998). Bilateral filtering for gray and color images. *Proceedings of the 1998 IEEE International Conferences on Computer Vision*, 836–846.
- Ullman, S. (1977). Filling-in the gaps: The shape of subjective contours and a model for their generation. *Biological Cybernetics*, *25*, 1–6.
- van Lier, R. (1999). Investigating global effects in visual occlusion: From a partly occluded square to the back of a tree-trunk. *Acta Psychologica*, *102*, 203–220.
- von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984, June 15). Illusory contours and cortical neuron responses. *Science*, *224*, 1260–1262.
- Wagemans, J., van Lier, R., & Scholl, B. J. (2006). Introduction to Michotte's heritage in perception and cognition research. *Acta Psychologica*, *123*, 1–19.
- Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, *38*, 310–324.
- Whitmore, C. L. G., Lawson, R. B., & Kozora, C. E. (1976). Subjective contours in stereoscopic space. *Perception & Psychophysics*, *19*, 211–213.
- Whittle, P., & Challands, P. D. (1969). The effect of background luminance on the brightness of flashes. *Vision Research*, *9*, 1095–1110.
- Williams, L. R., & Jacobs, D. W. (1997). Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation*, *9*, 837–858.
- Wishart, K. A., Frisby, J. P., & Buckley, D. (1997). The role of 3-D surface slope in a lightness/brightness effect. *Vision Research*, *37*, 467–473.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.
- Zhaoping, L. (2005). Border ownership from intracortical interactions in visual area V2. *Neuron*, *47*, 143–153.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*, 6594–6611.

## Appendix

### Mathematical Details and Algorithms

This Appendix provides the specific mathematical descriptions and some details of the algorithms used by the DISC model. The Matlab (Mathworks) code for the DISC model is available at [www.gestaltrevision.be/sources/DISC](http://www.gestaltrevision.be/sources/DISC)

#### Borderline and Junction Detections

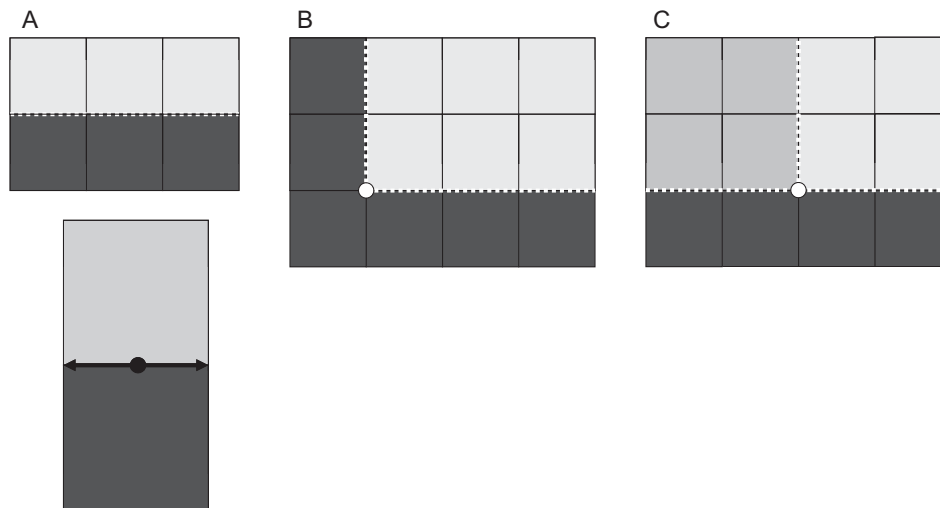
Borderlines are detected by detecting luminance differences pixel by pixel (see Figure A1A, top). Detected borderlines, therefore, constitute line segments with the length of one pixel, and each line segment is represented by a point at the center (see Figure A1A, bottom, black disk). To detect junctions, *bending points* of borderlines are detected. When a borderline changes its direction, the bending point is assigned as L-junction (see Figure A1B). The small bending points along the circular edges of objects are therefore also considered L-junctions. If the borderline is separated into two directions, the point is assigned as a T-junction (see Figure A1C). Amplitudes of all junction signals have values of one (binary signals). Note that the detected borderlines are always either horizontal or vertical line segments as the result of the pixel-by-pixel analysis, and the junctions always consist of the horizontal and vertical lines. All angles of the borderline segments and BOWN signals are measured counterclockwise relative to the  $x$ -axis in this article.

#### BOWN Computation

The BOWN-computation algorithm is modified from previous work by Zhaoping (2005). In this article, a BOWN signal is indicated

as an arrow with a side fin always on its left side, as shown in Figure 4A, top. The arrow corresponds to the orientation of the borderline, and the side of the fin indicates the side of the ownership. The length of the side fin indicates the strength of the BOWN signal. Each BOWN signal represents the BOWN value of the one-pixel-long line segment detected by the borderline-detection algorithm described above. Without considering the macroscopic properties, two  $180^\circ$  opposing ownerships are equally possible at each point (e.g.,  $0^\circ$  and  $180^\circ$  BOWNs in Figure 4A, bottom). Both signals have a value of zero before the computation starts.

The individual BOWN signals at each point on the borderline are compared with junction signals first. In Figure A2A, all possible geometrical relationships between L-junctions and BOWN are shown. Only if a junction is located on the side of the ownership, as indicated by the BOWN, and if the junction is directed so that it is seen from the point of the BOWN signal as concave (see Figure A2Aa) or collinear (see Figure A2Ae), it is considered “in agreement.” The junctions of opposite directions (see Figures A2Ab and A2Ac) are considered “in disagreement” with the BOWN signal. All other cases (see Figures A2Ad and A2Af) are ignored. When a T-junction exists, the judgment is essentially the same: If the owner side as indicated by the T-junction (above the top part) matches the owner side as indicated by the BOWN signal, it is considered “in agreement” (see Figure A2Ba, concave, and Figure A2Be, collinear), the T-junctions with opposite directions are considered “in disagreement” (see Figures A2Bb and A2Bc), and others are ignored



*Figure A1.* Borderlines are detected by detecting luminance differences pixel by pixel (A, top). Detected borderlines therefore constitute line segments with the length of one pixel, and each line segment is represented by a point at the center (A, bottom, black disk). Junctions are detected by finding a point where two borderlines with different angles meet (bending point). This point is assigned as L-junction (B, white disk). If the borderline is separated into two directions the point is assigned as T-junction (C, white disk).

(Appendix continues)



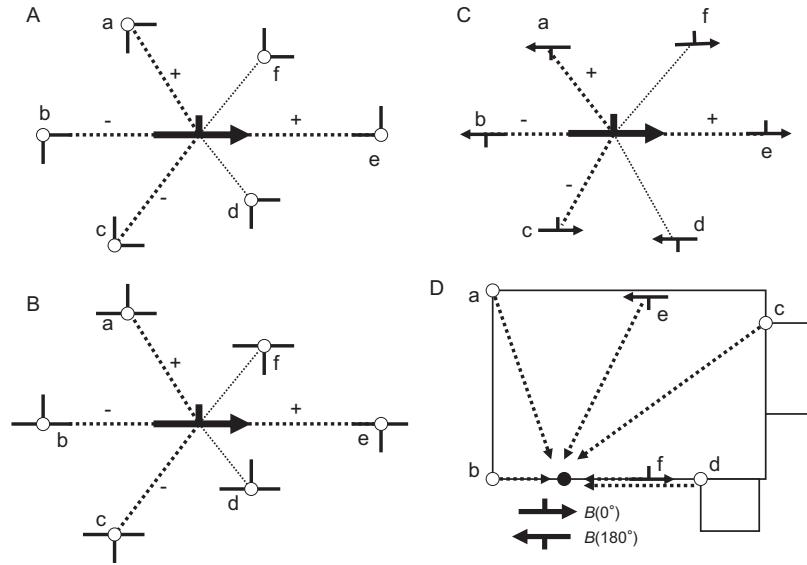


Figure A2. Border-ownership (BOWN) computation. Individual BOWN signals at each point on the borderlines are compared with junction signals first (A). If an L-junction is located on the side of the ownership, and if the junction is oriented so that it is seen as concave relative to the point of BOWN (Aa), or if it is aligned collinearly to the BOWN with the concave side matching with the direction of BOWN (Ae), it is considered in agreement, and the BOWN signal is enhanced. L-junctions on the opposite side and facing to the BOWN (Ac) or those that are collinear to the BOWN but indicating the opposite direction for the concave side (Ab) are considered in disagreement. All other cases are ignored (thin dotted lines). BOWN signals are enhanced or inhibited accordingly. Next, the BOWN signals are compared with each other (B). If their relationship is concave (Ba), convex (Bd), or collinear (Be), it is considered in agreement, and the BOWN signals are enhanced (thick dotted lines). If the relationship indicates the opposite, it is considered in disagreement, and the BOWN signals are inhibited. With T-junctions, the same principle is applied (C). An example figure (D) contains junction and BOWN signals in agreement with a BOWN signal at the location of the black disk (0°): concave L- (a), collinear L- (b), concave T- (c), convex T- (d) junctions and concave (e), collinear (f), and convex (g) BOWNs.

(see Figures A2Bd and A2Bf). Junction signals in agreement and in disagreement contribute to the BOWN signal as follows.

$$b_0(i, \theta) = \sum_{n+}^{N+} \exp(-d_{ni+}/\sigma_B) - \sum_{n-}^{N-} \exp(-d_{ni-}/\sigma_B), \tag{A1}$$

where  $b_0$  is the BOWN value for the direction of  $\theta$  at a point  $i$ ;  $d_{ni+}$  ( $d_{ni-}$ ) is the distance between a junction in agreement,  $n+$  (in disagreement,  $n-$ ), and the point  $i$ ; and  $N+$  ( $N-$ ) is the total number of L-junctions in agreement (in disagreement) with  $b_0$ .  $\sigma_B$ , a space constant determining how quickly the enhancement and the inhibition decay by the distance. This computation determines the BOWN values at each point at the first stage of BOWN computation (junction vs. BOWN).

At the second stage, the BOWN signals are compared with each other iteratively (see Figure A2C). Note that after the first iteration (junction vs. BOWN), one side of the BOWN signal becomes stronger than the other side at each point. Among all the possibilities of the relationship between two BOWN signals, two of them (see Figure A2Ca for concave and Figure A2Ce for collinear) are considered “in agreement,” and the other two (see Figures A2Cb and A2Cc) are considered “in disagreement,” whereas the rest are ignored (see Figures A2Cd and A2Cf). The enhancement and inhibition of the BOWN

signals is made according to Equation A1 (but the comparison is made between BOWN signals this time). In Figure A2D, an example figure is shown which contains junction and BOWN signals in agreement with the 0° BOWN signal at the location of the black disk,  $B(0^\circ)$ . They are concave L- (a), collinear L- (b), concave T- (c), collinear T- (d) junctions and concave (e) and collinear (f) BOWNs.

After the iteration, the BOWN signal on one side is larger than the other (quantitative BOWN map). The BOWN signals of the opposing directions are compared and a value of +1 is given to the winner side at each point (see Equation A2, binary BOWN map).

$$B(i, \theta) = \frac{\text{sign} \cdot \overline{b(i, \theta) - b(i, \theta + 180)}}{1}, \tag{A2}$$

$$\text{sign} = 1 \text{ if } b(i, \theta) > b(i, \theta + 180),$$

$$\text{sign} = 0 \text{ otherwise.}$$

Here, the upper line indicates the normalization of the value.

When illusory T-junctions are detected, as explained in the body of the article and Figure 5, the additional free-space BOWN computation is initiated (otherwise, the following process is skipped because it does not create any additional signals without the illusory T-junctions). The process is the interaction of the potential BOWNs around each pixel (see Figure 5D) and the illusory T-junctions, according to the same rule shown in Figure A2B. Here, L-junctions with BOWN signals attached with

the direction indicated in Figure 5C are considered the illusory T-junctions. Only the junctions of a certain size or larger are considered for this analysis. The process creates BOWN signals in the entire space with various amplitudes. However, BOWN signals with significant amplitudes develop only near the gap between the surrounding objects in the illusory figures. The results of subtraction of the individual BOWN signals with their opposing signals, therefore, are thresholded, and only the remaining signals are normalized and considered as the illusory BOWN signals (see Equation A3, top). The final binary BOWN signal ( $B_{fin}$ ) is created by combining the resulting illusory BOWN signals ( $B_{ill}$ ) with the original BOWNs ( $B_{phys}$ ) and is used for the integration at the next stage (see Equation A3, bottom).

$$\begin{aligned} B_{ill}(i, \theta) &= 1 \text{ if } b_{ill}(i, \theta) - b_{ill}(i, \theta + 180) \geq Th, \\ B_{ill}(i, \theta) &= 0 \text{ if } b_{ill}(i, \theta) - b_{ill}(i, \theta + 180) < Th, \quad (A3) \\ B_{fin} &= B_{phys} | B_{ill}. \end{aligned}$$

Here,  $Th$  is the threshold value and the vertical line indicates the logical odds ratio.  $B_{ill}$  and  $B_{phys}$  are both binary. With the logical odds ratio operator in Equation A3, therefore, the  $B_{fin}$  value at any location where the BOWN signal exists in either  $B_{ill}$  or  $B_{phys}$  becomes 1.0.

The differences between our algorithm and the one by Zhaoping (2005) are as follows. In the DISC model, the BOWN computation starts with the comparison of the BOWN signals with the junction signals. Zhaoping's algorithm did not consider junction signals, and the computation was only made by comparing the BOWN signals. The contribution of all signals that enhance the BOWN signals are weighted equally, whereas Zhaoping's algorithm introduced parameters for differentiated weights depending on the classification of the relationship between the pair (see Figures A2A, A2B, and A2C). Furthermore, the free-space BOWN computation is performed to compute illusory BOWNs. In the DISC model, the final BOWN value is binary to be used for the integration process. In other words, the model is independent of how the BOWN values are computed as far as it determines the winning side correctly.

### Surface Reconstruction

The surface construction is realized by "leaky" integration of BOWN signals (for the depth map) or the luminance ratio signals (for the lightness map). The use of the leaky integrator means that the causal integrator in Equation 2 in the text is replaced by the following.

$$F = \iint e^{-(u-x, v-y)} \cdot f(x, y) \partial x \partial y = e * f, \quad (A4)$$

where  $e$  should decay monotonically in one direction and be zero in the other direction. This is achieved by using a Gaussian function that is cut in half and by taking only the part where  $x$  is larger than zero. The integration is applied in both the depth and the lightness domains.

$$\begin{aligned} D_i(x, y) &= e(x, y, i) * B(i) \\ L_{log}^i(x, y) &= e(x, y, i) * C(i), \quad (A5) \end{aligned}$$

where  $e$  is the half Gaussian filter. The asterisk (\*) indicates convolution.  $D_i$  is the result of the leaky integration of the BOWN signal at a line segment  $i$ .  $L_{log}^i$  (log-lightness) is the result of the integration of the luminance ratio  $C$  at the line segment  $i$ .

In contrast to the conventional 2-D integration, the results of the integration could differ depending on the integration pathway. The contributions of different integration pathways to the integration value at a certain point are therefore measured as follows. Consider a one-pixel-length line segment of a borderline,  $i$ , (see Figure A3, between points  $a$  and  $b$ ). As explained above, the BOWN signal (valued + 1) of one line segment is represented by the central point. The contribution of the BOWN signal from this line segment to an integration value at a pixel point ( $p_1$ ) is, therefore, approximated by the integration value from the central point  $c$ . If the distance between  $p_1$  and  $c$  is  $r$ , the leaky integration value at  $p$  by this line segment,  $D_i$ , is

$$D_i = S \cdot \exp(-r^2/2\sigma_1^2), \quad (A6)$$

where  $S = +1$  if the direction of the integration pathway matches BOWN and  $S = -1$  if the direction of the integration pathway does not match BOWN.

However, if the straight line between the two points (integration pathway) crosses other BOWN line segments ( $p_2$ ), the integration value from the crossing is added. This results in a step up of the integration value or a step down according to the relationship of the direction of the integration pathway and the orientation of BOWN signals as in Equation A6. The  $D_i$  values from all line segments are summed. The total integration values,  $d$  (for depth map) and  $L_{log}$  (for log-lightness map) at each point are therefore computed from Equation A7.

$$\begin{aligned} d(x, y) &= \sum_i^N D_i(x, y) \\ L_{log}(x, y) &= \sum_i^N L_{log}^i(x, y). \quad (A7) \end{aligned}$$

Here,  $D_i$  and  $L_{log}^i$  are the individual integration values for the depth and the log-lightness from a line segment  $i$  and  $N$  is the total number of the line segments.

(Appendix continues)

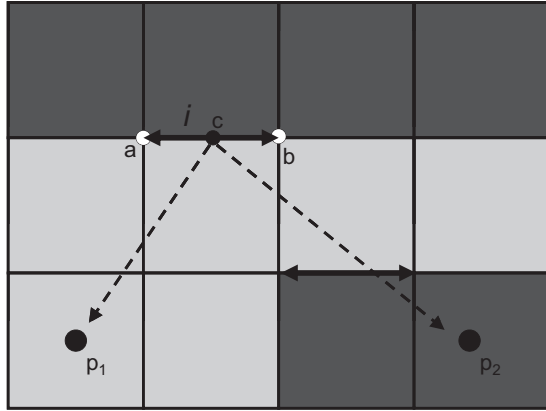


Figure A3. To measure the integration value, first the contribution from individual one-pixel-long line segments of borderlines to the integration value at a pixel is considered. These line segments already have border-ownership (BOWN) values to an assigned direction (+1) computed by the previous process. Each line segment (from  $a$  to  $b$ ) is represented by the central point ( $c$ ), and the leaky integration values from point  $c$  to point  $p_1$  are calculated. If the direction of the integration agrees with the direction of the BOWN, the value is positive, and if it disagrees, it is negative. If, in addition, the integration pathway crosses other BOWN signals before reaching another point ( $p_2$ ), the integration value steps up further or steps down according to the match or mismatch of the directions between the integration the BOWN. The values from all the points on the line segment are then summed and averaged.

The exponential of the log-lightness map is taken and anchored to create the primary lightness map. According to the anchoring rules described in the text, the largest area in the depth map and the area with the highest value in the primary lightness map are anchored, respectively. Furthermore, the depth map is normalized (the highest depth value becomes one). The largest area of the depth map is detected by depth slicing the map into 64 levels and by detecting the level that has the highest number of pixels belonging to it. The primary lightness map is anchored to the value of 1.0.

$$D = \frac{\overline{D_0}}{D_0}$$

$$L_{exp} = \exp(\alpha \cdot L_{log}) \quad (A8)$$

$$L_p = \frac{L_{exp0}}{L_{exp0}} + 1.$$

Here (and in all other equations), the underlining and the value on the side indicate that the side number is anchored to the value. The line above  $D_0$  indicates the normalization;  $L_{exp}$  is the exponential of the log-lightness map;  $\alpha$  is a constant; and  $D$  and  $L_p$  are the final depth map and the primary lightness map, respectively.

### Smoothing by Bilateral Filter

The integrated maps are smoothed by an iterative application of the bilateral filter (Tomasi, 1998). The principle of this method is to combine a spatial (Gaussian) smoothing filter with another smoothing filter in the amplitude domain. The combined filter is effective to smoothen out small changes (noise), while preserving larger changes as found at edges. This process can be described as follows:

$$h(X) = k^{-1}(X) \cdot \iint I(\xi) \cdot e(\xi, X) \cdot s(I(X), I(\xi)) \cdot d\xi, \quad (A9)$$

where  $h$  is the result of the filtering;  $X$  and  $\xi$  are the space coordinates;  $k$  is a normalization factor;  $I$  is the integration signal before filtering; and  $e$  and  $s$  are filters for the spatial smoothing and the amplitude smoothing, respectively. Both of them take the form of a Gaussian filter:

$$e(\xi, X) = \exp(-1/2 \cdot (d(\xi, X)/\sigma_e)^2)$$

$$s(\xi, X) = \exp(-1/2 \cdot (\|I(\xi) - I(X)\|/\sigma_s)^2), \quad (A10)$$

where  $d$  is the Euclidian distance between  $X$  and  $\xi$ . To iterate the process, the first result,  $h$ , of the bilateral filter is fed back to the process as a renewed input  $I$ . By iterating the process, the illusory edges of the surface become sharper and the surface flatter.

### Modification Factor

To create the polarity-depth map ( $D_p$ ), first the polarity map,  $P$ , is created from the log-lightness map. The log-lightness map is first thresholded, and any remaining positive signals are set to value one, any remaining negative signals are set to minus one, and finally, any areas with signals below threshold are set to zero (see Equation A11). The product of the polarity map,  $P$ , and the anchored depth map,  $D$ , is the polarity-depth map,  $D_p$ .

$$P(x, y) = 1 \text{ if } \alpha \cdot L_{log}(x, y) > th,$$

$$P(x, y) = -1 \text{ if } \alpha \cdot L_{log}(x, y) < -th,$$

$$P(x, y) = 0 \text{ otherwise.}$$

$$D_{p0} = \underline{D_0} \cdot P \quad (A11)$$

$th$  is the threshold value (to eliminate near zero signals).

The modification factor,  $M$ , is constructed as a linear function of polarity-depth and is used to modify the photometric structure provided by the primary lightness map,  $L_p$ , to obtain the lightness map,  $L$ .

$$\underline{M_1} = \beta \cdot \underline{D_{p0}} + 1$$

$$L(x, y) = \underline{L_p(x, y)}_1 \cdot \underline{M(D_p(x, y))}_1. \quad (A12)$$

Because the polarity-depth map is anchored to zero,  $M$  is, in effect, anchored to one following Equation A12 (top). The primary lightness map is anchored to the value of one and so is the lightness map as the result of Equation A12 (bottom).

### Parameter Settings

In this section, all the parameters used in the model and their values are described. All input figures have a size of  $128 \times 128$  pixels. The amplitude range (of the input "luminance") is from 0 to 1, the value for the white areas is set to 0.925, and the value for the black areas is set to 0.025. The diameter of the Pac Man shapes is 23 pixels, and the distance between the centers of the Pac Man

shapes is 48 pixels (support ratio 0.48; Shipley & Kellman, 1992). The side lines of Figure 1H have a line width of 2 pixels.

The space constant,  $\sigma_B$  (see Equation A1), determines the decay of the enhancement of the BOWN values by distance and is set as follows.

$$\sigma_B = 30.$$

The number of iterations for the BOWN computation is set to 20.

When the illusory T-junction condition appears, only the junctions with a certain size or larger are considered. They are chosen if the total length of the attached straight borderlines is larger than 10 pixels. The results of the free-space BOWN computation are thresholded so that only significantly large BOWNs are considered. The threshold value is set to 0.7.

The leaky integration also has a constant that determines the decay of the value in distance,  $\sigma_I$  (Equation A5, sigma of Gaussian) and is set as follows:  $\sigma_I = 60$ .

The inclusion of the bilateral filter introduces two parameters (see Equation A9, sigma of Gaussian). The sigma values for spatial smoothing ( $\sigma_e$ ) are set to 1.0 and 20 for the depth and the log-lightness maps, respectively. The sigma values for amplitude-wise smoothing ( $\sigma_s$ ) are set to 0.07 and 200, respectively. In addition, the size of a window within which the filtering is per-

formed is set to 5 for the depth map and 21 for the log-lightness map. The sigma values for spatial smoothing ( $\sigma_e$ ) are set to be large relative to the size of the window. The spatial smoothing therefore is constrained only by the window size in effect. The number of iterations for the bilateral filtering is set to 50.

The threshold value to eliminate near zero signals in creating the polarity map ( $th$  in Equation A10) is set to 0.1. Finally, the constant values  $\alpha$  in Equation A7 and  $\beta$  in Equation A11 are set as follows:  $\alpha = 0.0002$ ,  $\beta = 0.1$ .

Note that the final junction signals and the BOWN signals are binary and hence qualitative (their amplitudes have no meanings). Therefore, a wide range of parameter settings (e.g., for  $\sigma_B$  and  $\sigma_I$ ) work just as well as the above-mentioned settings, as far as junctions are detected properly and correct ownership of borderlines is obtained. Furthermore, as shown in Figure 2, the model is constructed with modules with well-defined functions. The parameters therefore can be determined separately, which also facilitates the process.

Received April 5, 2009

Revision received December 31, 2009

Accepted January 2, 2010 ■