The world is filled with objects and materials that interact with light, generating the image structure the visual system uses to recover scene structure. Our experience reflects certain aspects of the organization that exists in the material world, which allow us to perform the behaviors needed to survive and reproduce. The world is filled with "stuff" distributed in space and time, and vision provides some of the most important information about the behaviorally relevant properties of the physical world. The issue confronting vision scientists is to understand what information is, how it is extracted from retinal images, and how it is used to guide behavior.

There are two complementary problems that arise in recovering scene properties from images: segmentation (or decomposition), and synthesis (or grouping and/or interpolation), which both have associated sub-problems. The light reflected from materials is a conflated mixture of surface optics (color, lightness, translucency, specularity, etc.), the illumination field (the distribution of primary light sources and light reflects from different materials), and three-dimensional (3D) shape. These different sources of image variation must be sorted appropriately to recover scene structure. Another segmentation problem involves the problem of segmenting images into distinct objects and materials. This is not merely a problem of identifying "edges" in an image, since any luminance or chromatic discontinuity can arise in a variety of different ways. Image discontinuities can be generated by changes in pigmentation, depth discontinuities, 3D folds or corners of objects, changes in illumination, or specular reflections, which all contribute different kinds of information about scene structure. Thus, something much more than the mere detection of edges is needed to understand how objects are segmented from their surroundings. The complementary grouping and interpolation problem involves understanding how the visual system fills in missing information. The partial occlusion and camouflage of surfaces in natural scenes generates fragmented image data, which must be somehow unified to generate coherent representations of surfaces and materials.

In this chapter, I describe some of the background and recent progress into how the visual system sorts information into a stratified representation of surfaces in depth, with a focus on the problem of image segmentation. Although much research in vision is pursued with a "divide and conquer" strategy, I will argue that many of these ostensibly distinct problems are intimately coupled, and argue that a full understanding of one domain requires understanding how sources of information for one kind of property are distinguished from another.

THE PERCEPTUAL ORGANIZATION OF DEPTH

One of the general areas of vision research is depth perception, which refers to the transformation of 2D images into an experience of 3D scenes. One vivid and extensively studied source of depth information is stereopsis, which is based on the information provided by the slight differences in viewpoint generated by our frontally placed eyes. These different viewpoints generate binocular parallax, which is the apparent difference in position of an object when viewed from different positions. In order for the visual system to extract parallax, it must determine what counts as the "same object" in the two eyes' views, which is known as the correspondence problem. One of the main ongoing areas of research in stereopsis focuses on understanding how the visual system determines binocular correspondence and computes the positional shifts of common world features (binocular disparity) (Anderson & Nakayama, 1994; Kumano, Tanabe, & Fujita, 2008; Schreiber, Tweed, & Schor, 2006; Tanabe, Umeda, & Fujita, 2004; van Ee & Anderson, 2001). The prevailing assumption is that the "problem" of stereoscopic depth perception is essentially solved by the computation of disparity, which is predicated on the assumption that there is a simple one-to-one relationship between disparity and perceived depth. But it is not this simple.
The complexity arises because disparity and perceived depth do not always have a simple one-to-one relationship. Part of the problem arises as a consequence of the “stuff” that the brain uses to establish correspondence. There is currently an abundance of data demonstrating that disparity is computed by measuring the positional shifts of some local measure of image contrast. In its most general form, “contrast” refers to a normalized luminance difference, such as that generated by a local edge. Unfortunately, there is currently no single measure of image contrast that adequately captures perceived contrast, and it remains unclear precisely how disparity is computed from binocular image contrast (or what aspects of contrast are used as primitives for matching). Nonetheless, the mere fact that some form of luminance difference is used to compute disparity is sufficient to demonstrate that the relationship between disparity and depth can be a one-to-many mapping, which implies that there is more to the perceptual organization of stereoscopic depth than the computation of binocular disparity.

Consider a simple luminance discontinuity that is present in both eyes’ views, and assume that the visual system has correctly identified the discontinuity as arising from the same source in the two images (i.e., it is binocularly matched). This local discontinuity, or “edge,” would give rise to a single disparity, and hence be assigned a single depth. But what if this edge was generated by an occluding contour? In such contexts, one side of the edge (the occluder) is nearer than the other (the occluded background), a fact that is not captured by the single-value of disparity generated by the edge. For the visual system to recover scene geometry in the vicinity of occluding edges, it must map a single value of disparity onto two values of depth (Anderson, 2003b; Anderson, Singh, & Fleming, 2002; Fleming & Anderson, 2003).

This simple geometric fact implies that disparity does not always map onto depth (perceived or physical) in a simple one-to-one manner. This, in turn, implies that the “problem” of stereoscopic depth perception cannot be reduced to establishing a map of local disparity estimates, despite the fact that a large body of work in this field continues to embrace this view. This view has been driven in part by Julesz’s seminal invention of the computer-generated random-dot stereogram (RDS). In RDSs, depth can be manipulated by interocularly shifting the positions of individual dots, which results in a corresponding shift in perceived depth of those dots. But RDSs are special cases, whose rich texture obscures some of the complexities of the relationship between disparity and depth that arise in natural scenes. Understanding how the disparity of a local contrast relationship is translated into perceived depth requires a broader understanding of the different ways that contrast can be generated by the world, and how these different sources of image contrast are mapped onto physical and perceived depth.

The relationship between contrast and perceived depth depends on the kinds of surfaces and material properties that generate the pattern of local contrasts in the two eyes. One source of image contrast is local surface texture or variations in reflectance or pigmenta
tion along an opaque surface. In such conditions, the relationship between disparity and depth can be largely considered a simple one-to-one mapping, as is exemplified by the depth experienced in RDSs. However, the same is not true for something as simple as a luminance discontinuity (“edge”); how depth is assigned at an edge requires understanding what generated the edge in the world (Anderson, 2003b). If the edge is an occluding contour, then the occluding edge must be assigned a different depth than the occluded side of the edge. If the edge is a boundary of a transparent overlay (a form of partial occlusion), then depth must be split in two ways: at the location of the edge’s boundary, in the same manner as an occluding contour; and underneath the transparent surface, into a set of overlapping surfaces. The former decomposition cuts the image like a cookie cutter, splitting the image into pieces like a jigsaw puzzle; the latter decomposition splits the image into layers, like the layers of an onion. All of this may transpire in an image region containing only a single depth estimate or disparity value, which means that the full range of depth experienced in these regions must arise from other sources of information.

Although the geometry of occlusion introduces a source of complexity in interpreting local disparity signals, it also imposes a significant, inviolable constraint on how depth is organized in the neighborhood of a local contrast signal. I have previously dubbed this contain the “contrast depth asymmetry principle” (or CDAP; see Anderson, 2003b). This constraint is most easily understood by considering local luminance discontinuities. It expresses something quite simple and seemingly innocuous: in a world of opaque surfaces, the depth assignment of the individual luminances that define a luminance discontinuity (i.e., a local “edge”) are constrained such that they must appear at least as distant as the depth of the edge, or one side of the edge can appear more distant (as an occluded region). This seemingly benign constraint can explain a host of phenomena that receives no coherent explanation within conventional theories of binocular vision or depth perception. For example, the CDAP explains a large variety of “depth spreading” asymmetries in stereopsis. For
example, the inducing elements of a stereoscopic Kanizsa figures can be transformed from appearing as disconnected disks that are partially occluded by an illusory figure, into four illusory “portholes,” in which the interior of the disks appear as a unified background surface or void (figure 46.1). Takeichi, Watanabe, and Shimojo (1992) reported a related effect, wherein a few small dots that are placed inside and in front of a Kanizsa figure appear as isolated dots floating in front of the illusory figure, but transform the that figure to a vivid illusory hole when the depth of the dots is inverted, dragging all of the surround back in depth with the dots (figure 46.2). Such effects can be seen in even simpler displays: if a few dots are stereoscopically placed in front of the edges of a computer monitor, they appear to float in empty space; but if their depth is inverted, the background on which they are placed recede in depth and appear behind the boundaries of the monitor.

All of these phenomena involve asymmetries in how depth is assigned to local image contrast, or local “edges.” Although this idea may seem completely trivial and obvious when expressed so simply, it is far from universally understood or appreciated. Indeed, there are a large number of papers that purport to manipulate the depth of targets in ways that are simply impossible because they violate the constraints imposed by the CDAP. For example, some early studies of simultaneous contrast attempted to compare the effects of stereoscopically placing targets in front or behind their inducing fields (surrounds). The former is possible; the latter is not. Similar attempts to place targets behind their adjacent inducing fields were made with other forms of lightness illusions, such as White’s effect, which is also precluded by the CDAP (Spehar, Gilchrist, & Arend, 1995). Similar errors have occurred in visual search experiments (Davis & Driver, 1998), the role of familiarity in figure-ground displays (Peterson & Gibson, 1994), and host of other experiments that attempted to use binocular disparity a tool for inverting depth. Indeed, figure schematics have even been produced that are both physically impossible, and
non-representative of what is actually perceived in the displays they are intended to depict.

In general, the CDAP captures a constraint on how local image contrast can be mapped onto possible world events. Our discussion has focused on the role of the CDAP in explaining a variety of asymmetries in how surface properties are organized in depth in displays in which disparity relationships are simply inverted. In such contexts, the CDAP imposes the constraint that both sides of an edge must appear at least as distant in depth as the edge. But this simple statement fails to capture the full range of ways in which local edges (or local contrast more generally) can be generated, and the complexity of the percepts that can be evoked by simply inverting depth information. A deeper understanding of this constraint requires considering a broader class of world events, and how they relate to the local image context that they generate.

**SCISSION FROM DEPTH**

The preceding section articulated a constraint on how the brain maps local image structure (i.e., local image contrast) into a representation of surfaces in the world. The fact that our perceptual experience mirrors this constraint indicates that the visual system "understands" this constraint imposed by occlusion. Occlusion can, however, be a matter of degree; surfaces can vary in their opacity or "hiding power," which occurs in conditions of transparency. This fact introduces a new set of ambiguities in mapping local image structure (contrast) onto a representation of world properties.

Consider again the possible interpretations of a local image contrast (like an edge) with an associated depth signal (such as binocular disparity). For opaque surfaces, the edge could have been generated by a variation in surface reflectance, an illumination boundary, a fold in the 3D geometry, or an occluding and occluded surface. The problem is further complicated if the effects of transparent surfaces or media are considered. In this case, either side of the edge could be a transparent surface. It is also possible that both sides of the edge are partially obscured by a transparency overlay. In such cases, multiple depths would have to be assigned to one or both sides of the local image contrast to recover scene geometry (Anderson, 2003b).

How do these considerations impact on the interpretive constraints imposed by the CDAP? Transparency requires that depth partitioned into multiple depths along the same line of sight, a process historically described as scission (Koffka, 1935). There are three possible ways that this can happen: Either side of the edge may be split into a transparent surface overlying another surface, or both sides of the edge are interpreted as containing multiple layers. In the former case, one side of the edge is still treated as an occluding surface, but one that does not completely obscure the more distant, underlying layer. The latter case is more interesting, and provides a refinement of the CDAP as it has been expressed to this point. This refinement arises from considering how a local contrast (such as an edge) would be affected by an underlying transparent surface. Transparent overlays (surfaces or media) affect underlying surfaces in a number of ways; they reduce the overall amount of light coming from the underlying surface, and they typically add an additional source of reflected light from the pigments within the media or filter (Adelson & Anandan, 1990; Anderson, 1997; Beck, 1985; Metelli, 1970, 1974a, 1974b). The former effect is essentially identical to a change in illumination; the latter, additive effect reduces the contrast of the underlying surface. Note, however, that transparent surfaces cannot reverse the sign (or polarity) of contrast generated by underlying surface structure; they can only reduce or preserve the contrast that arises in a given direction.

These facts provide additional insights into the mapping between local image contrast and scene geometry. The constraint imposed by the CDAP on the interpretation of local image contrast applies to its polarity, or sign, not to its magnitude (or "strength"). In its unrefined form, it states that the component luminances along both sides of an edge must appear at least as distant as the depth of the edge. In its more refined form, when transparency is also considered, it states that some of the luminance on either side of the edge must appear at least as distant as the depth of the edge in a manner that preserves the polarity of the edge. The polarity constraints on transparency have been described in detail by a variety of authors, but the particular way in which polarity constraints are coupled to perceived depth is described more precisely in the CDAP.

The existence of transparency also induces a general source of ambiguity in the interpretation of all images. Consider a photograph of an arbitrary scene. The structure in the image is defined by the pattern of local contrasts that collectively form the image. How does the visual system determine whether a given image, or image region, is in plain view or partially obscured by a transparent surface or medium? Consider figure 46.3. The dark and light gray bars in the surround appear as portions of a surface in plain view; only the central ellipse appears transparent. Yet it is theoretically possible that the image of the surround was generated by a higher contrast set of stripes that is partially obscured by a contrast—reducing transparent layer. Yet this
percept is never experienced, which implies that some additional constraint is needed to explain when transparent scene interpretations are invoked.

I proposed that the visual system imposes a visual form of "Occam's razor" when computing transparency (Anderson, 1999; 2003b; Anderson & Winawer, 2005, 2008). In particular, I argued that transparency is only inferred when there is visual evidence of a perturbation in local luminance and/or contrast along continuous contours or textures. This constraint was dubbed the "transmittance anchoring principle," or TAP. The TAP states that the highest contrast region of a scene serves as an "anchor" point against which contrast variations are assessed; reductions in contrast along continuous contours or textures, which preserve contrast polarity, can induce percepts of transparency. The intuition motivating this principle is that changes in illumination preserve contrast, whereas transparent surfaces generically reduce contrast (the only exception being for transparent surfaces that reflect no discernible light).

These two principles—the TAP and CDAP—can provide a unified explanation of how the visual system partitions images into layered representations of surfaces in depth in a broad range of stimuli. Two examples of their predictive power are presented in figures 46.4 and 46.5 (from Anderson, 1999). Figure 46.4 depicts a sine wave grating placed within a diamond shaped aperture boundary. Binocular disparity is introduced by phase-shifting the grating relative to the aperture boundary in the two eyes. When the two images on the top right are cross-fused (or two left fused divergently), the grating appears in a single depth plane as a single surface. However, when disparity is inverted (cross fusing the two top left images, or divergently fusing the two top right images), the percept is dramatically transformed: the image now appears as uniform white diamond behind a series of fuzzy black stripes which vary in opacity. A similar effect can be observed in the image on the white background in the middle of figure 46.4, but the attribution of lightness to the two layers is inverted: a uniform black diamond now appears behind a series of fuzzy white stripes which vary in opacity. A similar transformation can be experienced when the two left columns in figure 46.5 are cross fused: dark clouds in front of light disks (top), and light clouds in front of dark disks on the bottom. Like the grating images, the luminance modulations within the near transparent surface appear to vary in opacity. Importantly, no clear percept of transparency or layers

THE PERCEPTUAL ORGANIZATION OF DEPTH, LIGHTNESS, COLOR, AND OPACITY 657
is experienced when the surround color falls between the luminance values of the texture inside the aperture; the intermediate luminance of the surround causes the contrast polarity along the aperture-texture boundary to reverse, which violates the conditions for coherent scission (bottom of figure 46.4).

There are a number of aspects of the perceptual organization of these displays that need to be explained. The two depth configurations use exactly the same images as input; they simply exchange which image the left and right eye sees. Thus, the disparity relationships in the two depth configurations are simply inverted, yet the perceptual organization of the texture is dramatically different in the two cases. Any cogent theory of these effects must explain: (1) what causes the dramatic transformation in the appearance of the texture when depth is inverted, and (2) the specific pattern of lightness, transparency, and depth that is experienced.

Consider the case in which the texture appears behind the aperture boundaries. The CDAP requires that the light and dark “sides” of a local texture must appear at least as distant as its disparity defined depth to account for its polarity. In this depth configuration, the texture is the most distant contrast in the image, so it appears in the same far depth behind the aperture boundary for both the grating and cloud textures, as conventional theories of stereopsis would predict. The contrast variations along the aperture boundaries occur in the near surface, and hence do not induce scission within the more distant texture (recall that the magnitude and polarity constraints apply to the more distant image contrast). However, when the texture is given a near disparity, the aperture boundary is now the far contrast in the scene. The CDAP requires that there must be a luminance relationship at this far depth that causes the edge to have its particular sign (polarity). When the surround is dark (top rows of figures 46.4 and 46.5), this means that there must be something at, or more distant than, the depth of the aperture boundary that is lighter than the surround, which accounts for the polarity of the contour. The converse holds for light surrounds: there must be something darker within the aperture, at this more distant depth, that is darker than the aperture boundary.

The constraint imposed by the CDAP is, however, purely local, and only requires that depth be assigned in the immediate vicinity of the contrast generated by the relatively distant contour. But the transformations experienced in figures 46.4 and 46.5 are global: the entire texture appears to split into two layers, and the distant layer appears to take on the lightness defined by either the most extreme light values in the texture.
representations of surfaces in depth. Binocular disparity was used as a source of local information about depth ordering, which can have a profound effect on how global depth and surface structure is experienced. Although the CDAP and the TAP were developed to explain how local depth determines the layout and organization of surfaces in images containing local depth information (like binocular disparity), their utility extends beyond images in which local depth is locally specified. More generally, these principles articulate general contingencies between perceived depth and surface properties, whether the depth is given or inferred.

Since the seminal work of Metelli, it has been known that there are basic geometric and photometric constraints that must be satisfied for transparency to be experienced (Metelli, 1970, 1974a, 1974b, 1985; Metelli, Da Pos, & Cavedon, 1985). Metelli’s own work focused on forms of transparency that can be understood with a simple physical devise called an episcotister—a rapidly rotating disk that had open sectors through which an underlying, two-tone background is visible (similar to figure 46.3). For such simple stimuli, it is possible to generate an algebraic solution to the transmittance and reflectance of the transparent layer (the fan blade). This solution was only possible because of the simple physical conditions under consideration, where the transparent surface had a uniform lightness and transmittance, and the underlying surfaces had a uniform reflectance. Nonetheless, some basic constraints on the conditions for perceived transparency can be derived from this special case. Geometric constraints involve the continuity of the underlying and overlying (transparent) layers, and photometric constraints involve the changes in the luminance and contrast between regions in plain view and regions of transparency.

The geometric and photometric constraints embodied in the CDAP and TAP suggest that it should be possible to generate layered image representations in non-steroscopic stimuli. Consider figure 46.6. The textures within the boundaries of the chess pieces are identical in the top and bottom, but appear dramatically different: The top image appears as a series of white chess pieces visible through black smoke, and the bottom appears as a series of black chess pieces visible through light fog. These images were created by changing the overall luminance and contrast range of a single “seed” texture (see figure 46.7). The target regions were created by increasing the contrast of the noise texture so that it spanned the full luminance range of the monitor. The luminance range of the surrounds was restricted to span from black to mid-gray on the top, and mid-gray to white on the bottom (note that the

DEPTH FROM SCISSION

The preceding sections focused on how the visual system uses local depth signals to induce layered
lightest and darkest components of the surround occur in the same locations in the top and bottom images; only their absolute intensities differ). The continuity of the texture within the chess pieces and their surrounds is assured by using the same seed texture for both the targets and surrounds. Moreover, by choosing an appropriate luminance range for the light and dark surrounds, the contrast polarity around the edges of the chess pieces all had a consistent, but opposite sign in the top and bottom images: dark-light in the top image, and light-dark in the bottom (referenced from surround to chess pieces). Only the magnitude of contrast along the surround—target borders varies in both of the images. Thus, the geometric and photometric constraints for transparency are satisfied in these images; the issue is explaining the particular pattern of depth, lightness, and opacity that they evoke.

First, note that the lightest regions in the top image’s chess pieces appear as the highest contrast regions and appear in plain view (e.g., the top of the left Bishop in the top), whereas the darkest regions within the bottom chess pieces appear as the highest contrast regions and appear in plain view (the top of the rightmost Rook or King in the bottom image). This is consistent with the TAP, and explains why the chess pieces appear to have the specific lightness that they do (white on top, black on the bottom). Second, note that the lowest-contrast regions along the boundaries of the chess pieces appear the most opaque (i.e., occluded), which is consistent with the scaling of opacity by the strength of the edge in plain view (but see below for some complexities with this computation). Third, the chess pieces of both surrounds appear to lie on an approximately mid-gray background, which are the regions where the contrast
of chess pieces appears strongest against the two surrounds. Thus, the CDAP and TAP can explain which regions are perceived in plain view, the depth and lightness associated with the two layers, and the relative opacity of the transparent surface or medium.

One complexity that was avoided in the previous analysis is the definition of local contrast used to determine how to apply the CDAP and TAP to these images. In the stereo images depicted in figures 46.4 and 46.5, the particular form of normalization used to define contrast could be largely avoided because the surrounds used in these images were a fixed, uniform luminance. Thus, the ordering of the contrast variations along the target—surround boundary was determined solely by the luminance differences between the targets and the surround; any normalizing factor would generate the same order of “edge strengths” (i.e., edge contrast) as those defined by the luminance difference. The same is not true for the images depicted in figure 46.6 because both the luminance within the targets (chess pieces) and the luminance in the surround varies. The concept of “local contrast” is currently ill defined because there is no definition of contrast that can capture perceived contrast in arbitrary images. This ambiguity has generated some controversy as to how to interpret the contrast constraints of the CDAP in these images. Although there is still no definition of contrast that can account for the perceived edge contrast in these images, the CDAP and TAP are still nonetheless predictive of the percepts in these images if perceived contrast is used to determine the regions in plain view.

COUPLED COMPUTATIONS OF DEPTH AND SURFACE QUALITIES

The phenomena and organizational principles articulated in the preceding reveal that there are systematic constraints on how local image data is organized into percepts of surfaces in depth. One of the core insights is that depth per se is not something that is seen. Rather, some thing (or “stuff”) is seen at as having a particular depth. In order to understand the relationship between how local image structure provides information about the layout of surfaces and materials—in short, the world—it is necessary to understand how different world properties constrain the local image data, and the extent to which visual processes embody these constraints when inferring scene structure from the images.

The CDAP expresses a constraint on the organization of depth imposed by the geometry of occlusion, which can have large transformative effects on perceived lightness and color. This suggests that scission may play a general role in lightness and color perception. This idea has not, however, been widely embraced. Indeed, a number of authors have argued that the transformations in perceived lightness that emerge in figures 46.4–46.6 are of a different kind (i.e., driven by different mechanisms) than other forms of contextual effects in
perceived lightness and color. The issue at stake concerns whether (and when) the process of layered image decompositions contributes to our experience of surface reflectance. At one extreme, some authors (Albert, 2007; Kingdom, 2008) have suggested that the lightness transformations experienced in figure 46.6 are simply a type of figure-ground reversal. The logic of this view is that the regions that appear as the occluded regions that appear in plain view in the two images occur on complimentary places in the two images, i.e., on the two “sides” of the luminance gradient. The link between the transformations in perceived surface reflectance and occlusion was one of the primary inspirations behind the development of these displays, and forms the basis of the CDAP and TAP. Nonetheless, there are a number of problems with the assertion that these effects can be reduced to a description as a figure-ground reversal. First, such accounts provide no explanation of the fact that the textures in these images contain a continuous distribution of luminance values that are decomposed into a percept of overlying layers. There are no “edges” in these images where the putative figure-ground reversal occurs; the regions that appear in plain view are perceptual outcomes that require explanation. Any proposed explanation must articulate where the occluding (transparent) layer begins, and why the luminance variations are experienced as variations in the opacity of a transparent surface, rather than variations in surface reflectance, 3D shape, or any other source in luminance variation.

A second problem with figure-ground “explanations” is their failure to appreciate the full range of stimuli that elicit percepts of scission and its transformative effects on perceived lightness and color (Anderson, 2003a; Anderson & Winawer, 2008; Wollschläger & Anderson, 2009). Some examples are presented in the images in figure 46.8. The central targets in the top and bottom images of all three columns are identical, yet appear dramatically different (particularly when played as movies); see (Wollschläger & Anderson, 2009). For example, the random textures of the central patch in the third column are composed of the exact same distribution of chromaticities, but the top image appears green, and the bottom image appears magenta. When played as dynamic movies, these regions generate a vivid percept of multiple layers: a random noise pattern identical to the surround, and a uniform, saturated surface, which appears as approximately the same color as the most saturated element (complementary to the surround) within the target. Similar effects can be seen for the lightness version of these effects in the first column: the top image appears blackish, and the bottom appears whitish. Neither pair of images contains image structures to which local figure-ground computations can be applied.

Whereas some authors have suggested that the effects of scission on lightness and color are special cases, other authors have argued that scission plays a crucial and general role in a variety of well-known induction phenomena. Recent work by Ekroll, Faul, and colleagues have argued that the phenomena of simultaneous contrast, “ganut expansion,” and “crispening” are all versions of the same effect, in which scission plays a crucial role (Ekroll & Faul, 2009; Ekroll,
Faul, & Niederee, 2004; Ekroll, Faul, Niederee, & Richter, 2002; Ekroll, Faul, & Wendt, 2011; Faul & Ekroll, 2002; Faul, Ekroll, & Wendt, 2008). Simultaneous contrast is well known, and refers to the changes in perceived lightness or color of targets placed on surrounds with a different lightness or color. Gamut expansion refers to the fact that the perceived range of chromaticities appears greater on a uniform background that is close to the chromaticity of the targets, than when the same targets are placed on a variegated surround. "Crispening" refers to the fact that the perceived difference between two targets is greatest when the target just changes from being an increment to a decrement relative to the chromaticity or lightness of its surround. In an elegant series of papers, these authors have shown that all of these phenomena can be modeled as the consequence of two sources of induction: a crispening component (which they attribute to scission), and a multiplicative scaling component (von Kries scaling), which treats the effect of context as a multiplicative transformation. In contradistinction to those suggesting that the scission experienced in textured displays may represent a special form of induction, the results by Ekroll, Faul, and colleagues suggest that layered image decompositions play a general and potentially ubiquitous role in computations of lightness and color; just how ubiquitous is an active open question of ongoing research (see, e.g., (Anderson, Khang, & Kim, 2011)).

CONCLUSION

The problem of depth perception has often been divorced from other problems of surface and material perception, such as perceived lightness and color. The phenomena and organizational principles described in the preceding chapter suggest that this is a false division, and that any understanding of depth perception requires an understanding of what is assigned a given depth. I have shown that simple inversions of depth can cause dramatic transformations in the perceptual organization of depth and material properties, such as their lightness, color, and opacity. I have also argued that these asymmetries can be understood with a few general principles derived from constraints on image interpretation that arise from two sources: the geometry of occlusion, and the photometric constraints imposed by transparent overlays. Future work is needed to understand how other material properties, such as translucency and gloss, shape the structure in the image and constrain how depth and material qualities are computed.

REFERENCES


