Global *yet* Early Processing of Visual Surfaces

Yukiyasu Kamitani[a*] and Shinsuke Shimojo[b,c]

[a] Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave. KS-454, Boston, MA 02215, USA.
[b] Computation and Neural Systems, Division of Biology, California Institute of Technology, MC 139-74, Pasadena, CA 91125, USA.
[c] Human and Information Science Laboratory, NTT Communication Science Laboratories, Atsugi, Kanagawa 243-0198, Japan.

* To whom correspondence should be addressed:
Phone: 617-667-5273; Fax: 617-975-5322
Email: ykamitani@hms.harvard.edu

Running Head: Surface Representation

To appear in *The Visual Neurosciences* (L.M. Chalupa and J. S. Werner (Eds), MIT Press, 2003)

Introduction

From the ecological optics viewpoint (Gibson 1950), the input to the eyes is basically ambient light, i.e., a lattice of lights that reflects geometry and perspectives between the observer and the environment. Since the lights are reflected from objects' surfaces, the lattice may be considered a summary of the environmental surfaces which is created lawfully with regard to the optical rules. More importantly, the representation of visual surfaces is an essential step for that of visual objects, which are the materials the observer's action is aimed at and performed with. Thus, the task the visual system faces is to retrieve from the retinal input the information about surfaces and objects that are ecologically relevant. It is primarily for this reason why understanding of visual surface representation is critical to our understanding of the entirety of visual processing.

In spite of its functional significance, it is only recently that surface representation began to be a subject of neurophysiological studies. This may be partly because the traditional view of visual processing derived from electrophysiological studies of single visual neurons is at odds with the fundamental nature of perceptual surfaces. In this chapter, we first illustrate the essential characteristics of perceptual surfaces, emphasizing the following points:

a) Local information given to limited areas of the visual field can lead to the global percept of surfaces, including the three-dimensional surface layout.

b) Surface representation is early in visual processing, in the sense that it is not mediated by conscious/cognitive processes, and even precedes other perceptual processes such as motion perception.

These characteristics may sound contradictory, but only if one takes the physiologically-driven, classical notion of visual processing as a hierarchical, unidirectional relay of signals through local feature detectors, whose size and complexity increase with the level in the hierarchy. According to the traditional view, global perception would be achieved only at high/late stages of visual processing, involving even problem-solving processes. We attempt to reconcile these two seemingly incompatible characteristics of perceptual surfaces in the light of recent physiological evidence for dynamic and global processing in the early visual cortex.

From Local Inputs to Global Percept of Surfaces

In classical illusions such as the Kanizsa figure (Figure 1A; Kanizsa 1979) and the Varin configuration for neon-color spreading (Figure 1B; Varin 1971), we can observe completed surfaces even though the physical information is spatially limited. In the Kanizsa configuration (Figure 1A), in addition to illusory contours, the Pacman-shaped stimuli create an illusory, rectangular surface where the entire area appears slightly darker than the background. In the Varin configuration, the color in the wedge-shaped portions appears to fill the square region. The term "filling in" refers to the situation where a property such as brightness and color propagates beyond the region of physical stimulation to form a clear percept of a delineated surface. Filling-in phenomena are important because they provide a

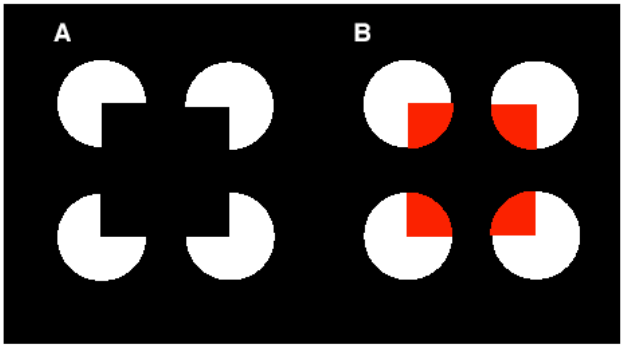

Figure 1. (A) Kaniza square. In addition to illusory contours forming a square, the Pacman-shaped stimuli give rise to the percept of an illusory surface in which the entire area appears slightly darker than the background, and amodally completed white disks. (B) Varin configuration for neon color filling-in. By adding colored wedge portions to A, a transparent, colored surface, and white disks seen through it are induced.

psychophysical paradigm to explore, as well as give insights into, the mechanism that integrates local inputs to form a global surface representation.

What is noteworthy is that filling-in effects often (though not always) come with an impression of a surface in front of another (or others). An example may be seen in a white disk that appears to be partly occluded by the illusory rectangular surface in the Kanizsa configuration (Figure 1A). This is sometimes called “amodal” completion because the occluded surface is perceived, but not locally “visible” in the literal sense. In the same Kanizsa configuration, the rectangular illusory surface can be considered “modally” completed because it is perceived as an entirely visible, occluding surface. Another representative example of modal completion may be seen in the illusory colored surface in the Varin configuration (Figure 1B). In this case, the colored portions are seen as parts of a single, semi-transparent, colored surface through which the back surface is also visible, thus allowing visibility of two surfaces along the same line of sight. Thus in short, filling-in and multiple surface layouts characterize visual surface perception from spatially sparse inputs.

Furthermore, it has been shown that minor changes in local visual inputs often lead to a drastic, global change in surface perception. Figure 2 is an example demonstrating the effect of local disparity (Nakayama, Shimojo, and Silverman 1989). The left-middle pair and the middle-right pair, which can fuse, in Figure 2 contain opposite signs of local disparity, while the retinally stimulated areas are largely identical. “Three disks through one window” or “one disk through three windows” can be perceived depending on the disparity, i.e., which pair is fused. Thus, the relationship between the modal (occludding) and amodal (occluded) can be reversed by a local change in disparity. Nakayama, Shimojo, and Silverman (1989) also showed that global surface layouts defined by local sdisparity influences recognition of a face occluded by stripes. The performance was better when the face was seen behind the stripes than in front, indicating that amodal completion of the face facilitated recognition.

Nakayama, Shimojo, and Ramachandran (1990) studied the effect of local disparity on color filling-in using the Varin and other configurations. When the colored portions were

Figure 2. Effect of local disparity information on global surface layouts. The observer is expected to see either three disks through one window, or one disk and three windows, depending on the direction of disparity. The converging fuser should fuse the left and middle images to see three disks through one window, and the middle and right images to see one disk through three windows. The diverging fuser should fuse the opposite pairs. Reproduced from Nakayama, Shimojo, and Silverman (1989).

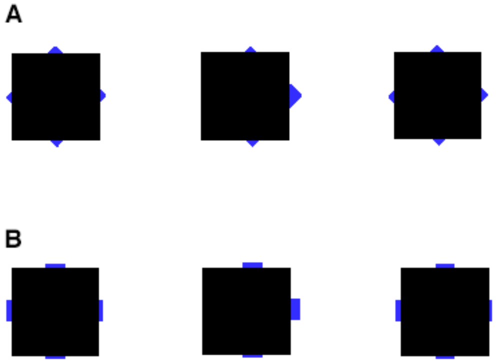

Figure3. Effect of local edges on global surface completion. The converging fuser should fuse the left and middle images; The diverging fuser should fuse the middle and right images. The difference in the small colored regions leads to remarkably different filled-in surfaces: a diamond in A, and a cross in B. Note that one of the colored portions is given only for one eye (the colored region on the left of the middle image is missing). Global filled-in surfaces can be formed even in the absence of binocularly matched inputs. By courtesy of Ken Namayama (Copyright 1994).

defined as front (by a crossed disparity), color filling-in, subjective contours, and transparency were all together enhanced. On the other hand, when they were defined as behind, all these became amodal, thus suppressed.

Figure 3 presents examples of color filling-in derived from very limited, colored areas (the original was developed by Ken Nakayama; an unpublished observation). Here, colored areas, as well as disparity and collinearity cues, are given very locally and sparsely. Note also that one of the colored areas is even unpaired between the left- and right-eye images (Nakayama and Shimojo 1990). Yet, a microscopic difference in edge orientation alone gives rise to a global difference in the completed surface (compare the top and the bottom stereograms when fused). Whereas collinearity itself may be considered a global property, the information that defines this property is edge location and orientation which are given only very locally.

Local changes in luminance related cues, such as contrast at edges (Nakayama, Shimojo, and Ramachandran 1990) and background luminance (Anderson 1999) are also known to be critical to determine global surface properties. It should be noted that the local factors determining global surface properties, such as edge orientation, contrast at edges, disparity, etc., are typical features that are detected in the early visual cortex (such as area V1 and V2). Hence, the global aspects of surface perception do not exclude the critical role of local feature detection in the early visual cortex. As we will discuss later, one of the possible mechanisms underlying global surface representation is a propagation-like process starting from local features, which in effect can fill in a big gap in space to establish a surface representation. In such a mechanism, a local feature detector is "local" in the sense that it is activated by an isolated stimulus presented in a limited area (classical receptive field), but also "global" in the sense that the activity can be modulated by global context outside the receptive field (Gilbert, Hirsch, and Wiesel, 1990). Likewise, a local feature is "local" in that it is given in a limited retinal area, but can also be "global" in that it can modulate the activity of distant local feature detectors.

Surface Formation in Early Visual Processing

Surface completion may be considered as an example of problem solving or reasoning. As an alternative, it may be achieved by bottom-up mechanisms based on local feature detection, at the very early level of cortical processing. In the case of subjective contours (and surface), for example, top-down, cognitive theories (Gregory 1972; Rock and Anson 1979) and bottom-up mechanistic theories (Grossberg and Mingolla 1985; Peterhans and von der Heydt 1989) have been proposed.

Recent evidence indicates that the process of surface completion precedes other cognitive/perceptual processes. For example, Davis and Driver (1994) demonstrated that Kanizsa subjective figures can be detected without focal attention, at parallel (thus early) stages of the human visual system. He and Nakayama (1994a) showed that visual search occurs only after the stage of amodal completion. Furthermore, a clinical study suggested that in a parietally damaged patient who suffered from "extinction" (a symptom in which a visual stimulus is neglected when there is another in the intact visual field), the condition was less severe when bilateral stimuli formed a common surface, such as an illusory Kanizsa figure and a surface completed behind an occluder (Mattingley, Davis, and Driver 1997). Likewise, the perceived direction of ambiguous motion, both smooth (Shimojo, Silverman and Nakayama 1989; Figure 4) and apparent (Shimojo and Nakayama 1990; He and Nakayama 1994b), is affected by amodally completed surfaces, indicating the precedence of

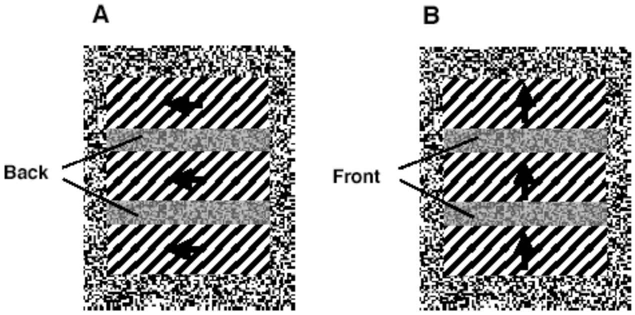

Figure 4. Effect of amodal completion on the barber pole illusion. In the barber pole illusion, the movement of stripes (moving orthogonal to their orientation) is seen through a window. They appear to move vertically (horizontally) through a vertically (horizontally) elongated frame. (A) If the panels separating the diagonal stripes are seen in back by adding disparity, the stripes, seen in three horizontally elongated regions, appear to move horizontally. (B) If the panels are put in front, the diagonal stripes are completed behind them, forming a vertically elongated region. Then, the movement of the stripes is seen as vertical. Reproduced from Shimojo, Silverman, and Nakayama (1989).

surface completion over, at least, some types of motion processing. For example, it is known that the perceived direction of a drifting grating is ambiguous when presented in a rectangular window, with the direction along the longer axis being dominant (the barber pole illusion; Wallach 1935). This dominance is preserved even when the window is partly occluded such that the longer axis of the visible (non-occluded) areas is different from that of the entire window (Shimojo, Silverman and Nakayama 1989; Figure 4B).

These results demonstrate that visual surface representation is established at a level earlier than cognitive and some perceptual processes, as a result of pre-attentive or unconscious operations. On the other hand, it has been argued that surface perception can be regarded as a process finding a statistically optimal solution to the inverse optics problem with regard to the real-world constraint (Poggio, Torre and Koch 1985; Nakayama and Shimojo 1992), thus analogous to cognitive inference at the functional level. It should be noted, however, that the "early" and "inference-like" aspects are not necessarily mutually exclusive. The real issue is whether the "inference-like" process is implemented at an early level, not mediated by conscious/cognitive processes, or at a late level, mediated by conscious/cognitive processes. The evidence above seems to support the former.

Aftereffects Induced by Perceptual Surfaces

Aftereffects, visual sensations that persist after prolonged viewing of stimuli, have been widely used to characterize mechanisms underlying visual perception. The percept is thought to reflect adaptation of neural subunits which respond to the adapting stimuli (see also Webster, Chapter XX, this volume). Aftereffects can be highly specific with respect to the features of adapting stimuli, such as orientation and spatial frequency, corresponding to the feature selectivity found in single neuronal responses in the visual system. Thus, aftereffects have been called "psychologist's microelectrode" (Frisby 1979), which could probe into the relationship between visual perception and underlying neural mechanisms. In this and the following sections, we highlight the global and early nature of visual surface representation in the light of its role in the formation of aftereffects.

The orientation-contingent color aftereffect (the McCollough effect; McCollough 1965) is one of the striking examples of highly specific sensory adaptation. In the paradigm of the McCollough effect, an observer views alternating horizontal and vertical stripes with different colors, such as red vertical stripes alternating with green horizontal stripes, for several minutes. After adaptation, a test stimulus made of achromatic, horizontal or vertical stripes are perceived to be tinged in the color complementary to that of the adapting stripes with the same orientation. Watanabe, Zimmerman and Cavanagh (1992) replaced the test stripes with a grid pattern that appears to be overlapping horizontal and vertical stripes due to its appropriate luminance combination for perceptual transparency and surface segregation. They found that the orientation-contingent color aftereffect is perceived in the subjective, overlapping stripes. Watanabe (1995) also showed that a test pattern made of largely occluded, but perceptually completed, stripes can elicit the McCollough effect. These results demonstrate that the McCollough effect, which is generally thought to involve early visual processes specific to orientation and color (Stromeyer 1978), can be mediated by, or preceded by, the global processing of perceptual surfaces.

In the example described above, visual patterns inducing subjective surfaces were used as test stimuli on which the effects of adaptation were perceived. Is it possible to adapt to perceptually filled-in surfaces, resulting in aftereffects observed in the region that is not retinally stimulated during adaptation? It is known that prolonged viewing of a

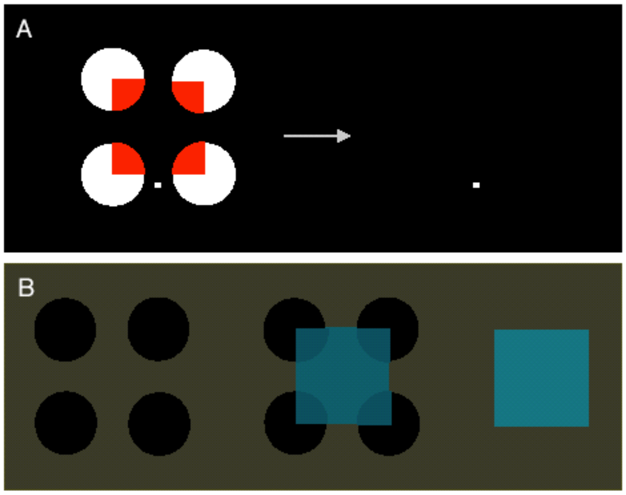

Figure 5. Afterimage formed by adaptation to a perceptually filled-in surface. (A) After adapting to the Varin configuration for neon color filling-in (left image; fixating to the white dot for about 30 s), the afterimage is observed on a blank screen (right). (B) Typical afterimages. During the observation period, the afterimage changes its appearance dynamically: The afterimages for the disks (left), for the filled-in surface (right), and for both (middle) appear to alternate several times. Adapted from Shimojo, Kamitani, and Nishida (2001).

grating leads to a decrease in apparent contrast, or the elevation of threshold in a test grating with a similar orientation and spatial frequency (Blakemore and Cambell 1969). Instead of a fully visible grating, Weisstein (1970) used a grating partially occluded by an object as an adapting stimulus, and measured the apparent contrast of a small grating patch within the region where the occluding object had been presented during adaptation. The apparent contrast of the grating patch was significantly lower compared to those seen after adaptation to a blank screen or to the object only. It suggests that the representation of amodally completed gratings can undergo adaptation in a similar manner to that of non-occluded gratings.

Afterimages of Filled-in Surfaces

Afterimages are often distinguished from other aftereffects (tilt, motion, size, etc.) in that they do not require a particular test stimulus to observe the effect. Afterimages are modulation of the first-order visual features, luminance and color, which are defined at each point of an image. Thus, a homogeneous background, having first-order features (luminance/color), is sufficient in order to observe their modulation. On the other hand, other aftereffects are modulation of higher-order visual features, such as orientation and motion, which are defined by the *relations* among more than one point, thus require a patterned (higher-order) test stimulus to observe their modulation. Hence, afterimages have been considered to reflect the most primitive, point-by-point visual processing, and their origin has been believed to arise from either bleaching of photochemical pigments or neural adaptation in other retinal cells (Craik 1940; Brindley 1962; Craik 1966; Virsu and Laurinen 1977).

Can afterimages, which have been thought to be concerned with primitive, point-by-point processing, be created by adaptation to perceptually completed, global surfaces? Shimojo, Kamitani, and Nishida (2001) showed that perceptually filled-in surfaces can give rise to afterimages, using color filling-in displays such as the Varin configuration (Figure 1B, 5A). After prolonged adaptation, the adapting stimulus was replaced with a dark homogeneous background (Figure 5A). The observers perceived not only the afterimages of the local inducers (Pacmen/wedges or disks), but also a global afterimage of the perceptually filled-in surface (Figure 5B). The global afterimage cannot be attributed to the general fuzziness and leaky edges of the afterimage, because adaptation to the inner wedge portions alone (without the outer Pacmen) leads to the four parts corresponding to the wedges separately visible in the afterimage. The global afterimage is distinct from conventional afterimages in that it is visible at a portion that has not been retinally stimulated, but corresponds to a perceptually filled-in surface.

The observation described above, however, does not necessarily imply that the global afterimage arises from the adaptation of the neural mechanism producing the perceptually filled-in surface. It is possible that the global afterimage originates from local afterimages of the inducers: The color of the global afterimage may be merely a result of the ordinary filling-in mechanism that may equally treat real stimuli and signals due to local adaptation. Shimojo, Kamitani, and Nishida (2001) performed a series of experiments to demonstrate that the

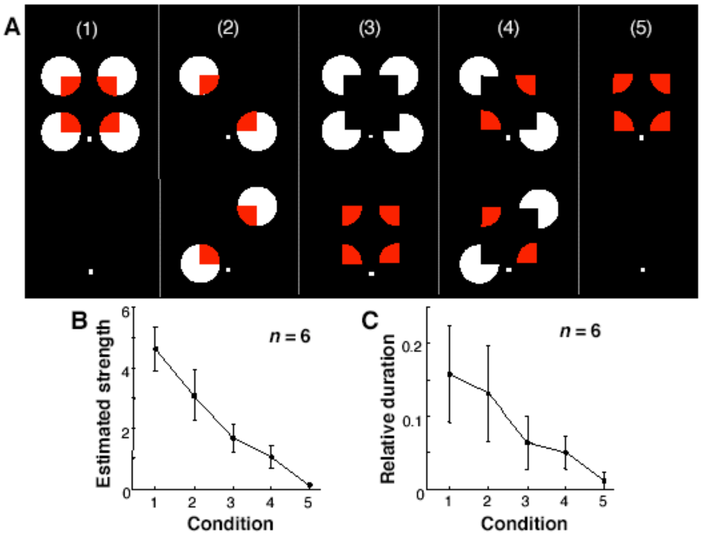

Figure 6. Surface vs. element adaptation I. (A) The inducer for the color filling-in was divided into complementary pairs (except condition 5, where only the wedge portions were shown), and they were alternated (667 ms for each frame) for adaptation. The strength of perceptual filling-in during adaptation was varied (decreasing approximately in this order), while the total duration of adaptation to each portion of the inducer was held constant (except condition 5). (B) Estimated strength of the global afterimage for the conditions depicted in A. The data for six subjects are pooled (error bar, standard error of the mean). The estimated value 10 corresponds to the strength of the global afterimage for the complete Vairn figure (no alternation or decomposition), which had been observed before this experiment. (C) Duration of the global afterimage relative to the total test period. Reprinted from Shimojo, Kamitani, and Nishida (2001).

global afterimage is indeed due to adaptation of the representation of the filled-in surface (surface adaptation hypothesis), as opposed to local adaptation followed by an ordinary filling-in process (element adaptation hypothesis).

First, the time course of the afterimage was analyzed. During the test period, different types of afterimage appeared and disappeared several times in different time courses. Subjects were asked to monitor the visibility of two types of afterimage separately using two buttons: the local afterimage, which corresponded to elements of the inducer (Pacmen, wedges, disks, or their combinations), and the global afterimage, which extended out of them toward the central portion to form a color-filled rectangle (Figure 5B). They pressed one button while the local afterimage was visible, and another while the global afterimage was visible. The results show that there were significant time periods during which the subjects reported visibility of the global, but not the local, afterimages (average, 20% of the total test period of 20 s). Furthermore, the likelihood analysis for the global afterimage in the presence or absence of the local afterimage revealed that the global afterimage tended to be more visible when the local afterimage was *not* visible than when it was visible. These observations indicate that the visibility of the local afterimage is not a necessary, or even a favorable condition for that of the global afterimage, in disagreement with the element adaptation hypothesis.

One prediction of the surface adaptation hypothesis is that since the global afterimage is due to the perceptually filled-in surface during adaptation, the strength of the global afterimage should be correlated with that of the perceptual filling-in during adaptation. The element adaptation hypothesis, on the other hand, predicts that the strength of the global afterimage should be determined solely upon that of local afterimages, and thus remain constant as long as the local afterimages are the same.

In another experiment, the strength of perceptual filling-in during adaptation was manipulated by alternating two frames (667 ms each) composed of complementary parts of the inducer as shown in Figure 6A (except condition 5 where only the colored wedge portions turned on and off). Since the strength of perceptual filling-in decreases approximately in the order of the conditions shown in Figure 6A, the surface adaptation hypothesis would predict that the strength of global afterimage also decreases in the same order. On the other hand, since the total duration of adaptation was equal across portions of the stimulus in all the conditions except condition 5, the strength of the local afterimage should be the same across conditions 1 to 4. Thus, the element adaptation hypothesis would predict that the relative strength of the global afterimage would be approximately the same across conditions 1 to 4 (and perhaps weaker in condition 5). As shown in Figure 6B and C, both the estimated strength and the visible duration of the global afterimage decreased according to the order of

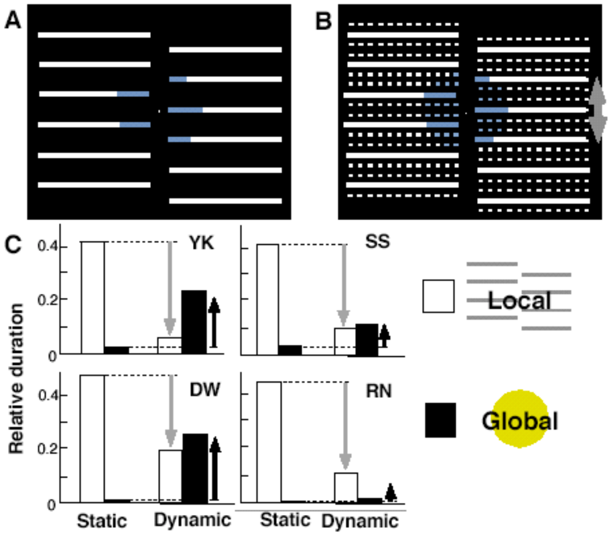

Figure 7. Surface vs. element adaptation II. (A) Static condition. The line segments were presented sparsely and statically. (B) Dynamic condition. The line segments underwent vertical apparent motion while their portions in the fixed disk-shaped area were colored blue. The broken lines indicate positions for displacement, and were invisible in the actual stimulus. The moving inducer produces more vivid filling-in but less local adaptation, compared to the static condition. (C) The duration of visible afterimages relative to the total observation time is plotted for the static and the dynamic conditions in each subject. The white and black bars indicate local and global afterimages, respectively. Reprinted from Shimojo, Kamitani, and Nishida (2001).

strength of perceptual filling-in, supporting the surface adaptation hypothesis, but not the element adaptation hypothesis.

In the last experiment, a new type of dynamic stimulus for color filling-in was employed to further dissociate the predictions of these hypotheses. In the static condition (Figure 7A), line segments were placed sparsely and statically, so that the impression of color filling-in would be minimal, while the local afterimage of the line segments formed strongly. In the dynamic condition (Figure 7B), the line segments were displaced up and down to create an impression of motion, while the disk-shaped area within which the line segments were colored blue was fixed. It appeared as though a set of white line segments moved up and down behind a semi-transparent, stationary colored disk. This condition was designed to enhance the impression of color filling-in maximally during adaptation, while minimizing the local afterimage of line segments by the constant displacement. The surface adaptation hypothesis predicts that the duration of the local afterimage would be reduced whereas that of the global afterimages would be increased, relative to the static condition. The element adaptation hypothesis, on the other hand, predicts that both local and global afterimages would be attenuated in the dynamic condition, since the duration of the global afterimage should strictly depend upon that of the local afterimage. The results of the visible duration of the local and global afterimages clearly indicate the enhancement of the global afterimage in spite of the attenuation in the local afterimage, consistent only with the surface adaptation hypothesis.

These results support the idea that afterimages can arise from the adaptation of the neural mechanism representing perceptually filled-in surfaces. Since unlike other aftereffects, afterimages are concerned with most primitive visual attributes, it may suggest that filled-in surfaces are represented at some very early stage of visual processing. In fact, Shimojo, Kamitani and Nishida (2001) also reported that interocular transfer of adaptation did not occur to induce the global afterimage in the unadapted eye, indicating that the adaptation occurred in monocular units, which exist only in the primary visual cortex or earlier (Hubel and Wiesel 1962). In addition, when one eye was suppressed by pressure-blinding the retina after adaptation, the global as well as the local afterimages became invisible. Although these observations suggest contributions of retinal adaptation to the global afterimage, it is unlikely that retinal processing explicitly represents filled-in, global surfaces, since each retinal cell deals with inputs only from a small area and lacks extensive long-range connections with other cells and projections from higher visual areas, which visual cortical neurons have. Therefore, the earliest level of visual cortical processing, which can handle both local primitive features and global connections, would need to be involved in the representation of perceptual surfaces, and the adaptation that gives rise to the global afterimage. In the next section, we will discuss possible neural correlates in the early visual cortex.

Neural Substrates for Perceptual Surfaces

It is challenging to explain the global aspects of surface perception in terms of topographically organized, small receptive fields of visual neurons (Barlow 1953; Kuffler 1953; Hubel and Wiesel 1962). This is the case not only for "illusory" surfaces, which we have focused on in the previous sections, but also for "real" homogenous surfaces whose edges and surface attributes are physically available, because visual neurons generally respond better to variation in luminance or color, than to their absolute values (homogeneous surface attributes)**.** It is now widely accepted, however, that the response of visual neurons can be substantially and selectively modulated by stimuli presented outside the classical receptive field (Allman, Miezin and McGuinness 1985; see also Lamme, Chapter XX, this volume). Surface perception may be one of the global perceptual phenomena that require such long-range interactions beyond classical receptive fields (Spillmann and Werner 1996, see also Spillmann and Ehrenstein, this volume).

Neural responses correlated with global surface perception have been demonstrated by manipulating the global context determining perceived surface attributes while holding the local features inside the receptive field constant. Rossi, Rittenhouse, and Paradiso (1996) studied the activity of neurons whose receptive field fell within a gray square surrounded by a background with changing luminance. This display produces illusory modulation of brightness of the central gray square: The brightness negatively correlates with the background luminance (brightness induction; Hering 1964). They found that in many neurons in the primary visual cortex, the activity was modulated by the background luminance in accordance with the perceived brightness. Lamme (1995) showed that the activity of neurons in the primary visual cortex depends on the figure-ground context, i.e., which region appears as a surface in front of the other. A significantly larger response was observed when the receptive field was in the figure region than when it was in the background, while the local feature within the receptive field was constant. Zhou, Friedman and von der Heydt (2000) (see also von der Heydt, Chapter xx, this volume) reported another type of figure-ground selectivity of early visual cortical neurons. A local light-dark edge, for instance, could be the left side of a dark square or the right side of a light square. They found that neural responses to the same edge can be modulated by the side to which the border belongs to ("border ownership"), while the context to determine border ownership was provided outside the receptive field.

Neural responses to illusory or completed contours/lines that emerge in association with surface perception have also been found in early visual areas. Cells in area V1 as well as V2 are known to be responsive to illusory contours, such as those perceived in the Kanizsa figure (e.g., von der Heydt, Peterhans, and Baumgartner 1984; Lee and Nguyen 2001; see also von der Heydt, Chapter xx, this volume). Sugita (1999) demonstrated that cells in the primary visual cortex respond to a bar occluded by a small patch (surface) when a disparity is given outside the receptive field such that the patch is seen in front. Bakin, Nakayama, and Gilbert (2000) also found that responses in the early visual areas (more in V2 than V1) are correlated with modally and amodally completed lines, while varying depth cues for surface segregation outside the receptive field.

Most of these electrophysiological data are consistent with the view that perceptual surfaces are topographically represented: The response of each cell reflects the perceived surface attributes at the point/area corresponding to its receptive field. The notion of topographical representation is often criticized for assuming "redundant processes of painting an internal screen" (Dennett 1991). It is logically possible that perceptual surfaces are represented in a "symbolic" way, without being mediated by topographical representation. For instance, a neuron that is selectively activated by images of faces can be thought to symbolically represent faces, but such representation may not require a map of activity similar to faces formed in a visual cortical area. The coding of "border ownership" described above (Zhou, Friedman and von der Heydt 2000) may be regarded as a semi-symbolic representation in early visual areas, since it enables a determination of surface attributes and configuration by tracing the activity corresponding to the border, without looking at the whole two-dimensional map of activity.

It has been argued, however, that in order to determine surface attributes from local ambiguous information given within receptive fields, the brain may make effective use of the topographically organized circuits in early visual areas (Pessoa, Thompson, and Noe 1998). A solution to the determination of surface attributes may be achieved by propagation of activity from cells receiving critical information such as edges, through synaptic cascades of local connections. This propagation mechanism may be considered more consistent with the topographical, as opposed to the symbolic representation. A recent human fMRI study provides further support for topographical representation of perceptual surfaces. When a color filling-in stimulus similar to the Varin figure was presented, clear activation was produced in the cortical region in the primary visual cortex corresponding to the filled-in area in the visual field (Sasaki et al. 2001). Thus, perceptually filled-in surfaces could be represented by "filling-in" in the cortical map. This finding is highly consistent with the demonstration of global afterimages induced by filled-in surfaces, which we described in the previous section in detail. A similar implication comes from a totally different angle and method. Kamitani and Shimojo (1999) showed that by transcranial magnetic stimulation (TMS) of the human occipital cortex, a hole, or scotoma, can be created in a flashed, large-field visual pattern, in accordance with the anatomy of early cortical maps (Figure 8). They also found distortion of the scotoma in grating patterns (Figure

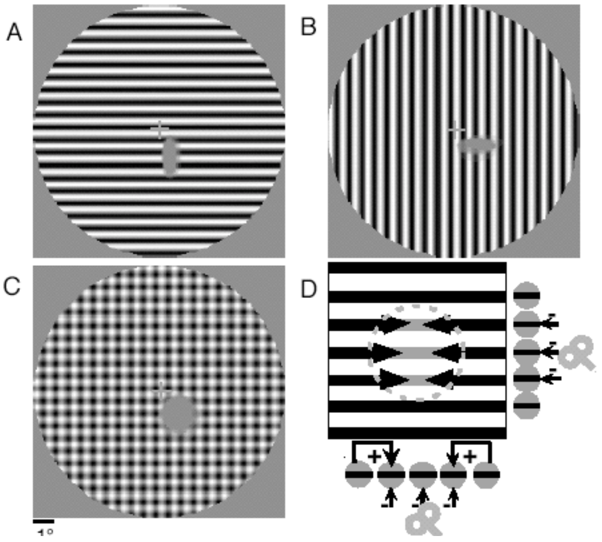

Figure 8. (A-C) Reconstructed percepts of TMS-induced scotomas seen on flashed visual patterns: a horizontal grating (A), a vertical grating (B), and a grid pattern (C). The visual pattern was presented for 40 ms, and the magnetic stimulation was delayed 106 ms. (D) Hypothetical mechanism for the anisotropic distortion of the scotoma. A horizontal grating is shown with horizontally tuned neural units aligned vertically (right) and horizontally (bottom). The central gray circle represents the visual space corresponding to the cortical area directly affected by the magnetic stimulation. Long-range facilitatory connections among collinearly aligned units (+ in the bottom row) mask inhibition by inputs caused by TMS (–); Thus, the suppressed region appears compressed along the stripes. Reprinted from Kamitani and Shimojo (1999).

8A and B), which can be interpreted in terms of the local connection among orientation-detection units observed in the primary visual cortex (Kapadia et al. 1995). The distortion may reflect a completion process that operates against the inhibitory effect of TMS, by means of propagation of activity through local synaptic connections (Figure 8D).

While it remains to be further studied whether, and to what degree, perceptual surfaces are topographically represented, functional characteristics of early visual cortical areas seem to be ideal as a medium to handle global as well as local information required for the processing of visual surfaces. It has been proposed that horizontal intrinsic connections within areas (Gilbert, Hirsch and Wiesel 1990), and recurrent inputs from higher areas (Lamme and Roelfsema 2000) result in dynamic changes in tuning in early visual areas, including the emergence of the global properties correlated with surface perception (Spillmann and Werner 1996). It should be noted that recurrent inputs from higher areas are not necessarily derived from conscious/cognitive commands. They can be fast, automatic processes, not mediated by conscious/cognitive processes. Furthermore, representation at early visual areas would remain "early" even after modulation by recurrent inputs, as long as it is available for later perceptual, or cognitive processes. Although contributions of possible symbolic representation at higher areas, as well as global processing at subcortical levels (e.g., Pöppel 1986) should not be excluded, the global and early aspects of visual surfaces discussed in the previous sections could be best explained by the global processing of early visual areas.

Conclusions

In this chapter, we have provided a framework to integrate psychophysical and physiological findings around the key concept of visual surface representation. We have shown that in the light of recent findings, the classical view of visual processing as a serial relay of signals through local feature detectors with progressively increasing size and complexity needs to be fundamentally reconsidered. The classical view assumes certain analogy among the notions, such as local/global and early/late: Local processing is early, and global processing is late. However, we have demonstrated that visual surface representation is achieved by global *yet* early processing.

We have also seen that the dichotomies, local vs. global and early vs. late, themselves also need to be carefully applied to describe psychophysical phenomena and physiological processes involved in surface perception . For instance, a "global" surface representation can be created by "local" retinal inputs, and the activity of a neuron with a "local" receptive field is modulated by "global" context. "Early" perceptual solutions look similar to "late" cognitive solutions based on inference, and recurrent processing in the visual cortex makes the distinction of "early" and "late" stages less meaningful. Such complication derives from the fact that that surface representation is beyond the scope of the classical view of visual processing as sequential, local-to-global feature detection.

References

Allman, J., Miezin, F., and McGuinness, E., 1985. Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons, *Annu. Rev. Neurosci.* 8:407-430.

Anderson, B. L., 1999. Stereoscopic surface perception, *Neuron* 24:919-928.

Bakin, J. S., Nakayama, K., and Gilbert, C. D., 2000. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations, *J. Neurosci.* 20:8188-8198.

Barlow, H. B., 1953. Summation and inhibition in the frog's retina, *J. Physiol.* 119:69-88.

Blakemore, C. and Cambell, F. W., 1969. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images, *J. Physiol.* 203:237-260.

Brindley, G. S., 1962. Two new properties of foveal afteimages and a photochemical hypothesis to explain them, *J. Physiol.* 164:168-179.

Craik, K. J. W., 1940. Origin of visual after-images, *Nature* 145:512.

Craik, K. J. W., ed 1966. *The Nature of Psychology: A Selection of Papers, Essays and Other Writings by the Late K. J. W. Craik,* Cambridge: Cambridge University Press.

Davis, G. and Driver, J., 1994. Parallel detection of Kanizsa subjective figures in the human visual system, Nature 371:791-793.

Dennett, D. C., 1991. *Consciousness explained*, Boston: Little, Brown and Company.

Frisby, J. P., 1979. *Seeing: Illusion, Brain, and Mind,* Oxford: Oxford University Press.

Gibson, J. J., 1950. *The Perception of the Visual World*, Boston: Houghton Mifflin Company.

Gilbert, C. D., Hirsch, J. A., and Wiesel, T. N., 1990. Lateral interactions in visual cortex, *Cold Spring Harb. Symp. Quant. Biol.* 55:663-677.

Gregory, R. L., 1972. Cognitive contours, *Nature* 238:51-52.

Groberg, S. and Mingolla, E., 1985. Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading, *Psychol. Rev.* 92:173-211.

He, Z. J. and Nakayama, K., 1994a. Perceiving textures: Beyond filtering, *Vision Res.* 34:151-162.

He, Z. J. and Nakayama, K., 1994b. Perceived surface shape not features determines correspondence strength in apparent motion, *Vision Res.* 34:2125-2135.

Hering, E., 1964. *Outlines of a Theory of the Light Sense,* Cambridge, MA: Harvard University Press.

Hubel, D. H. and Wiesel, T. N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visaul cortex, *J. Physiol.* 160:106-154.

Kamitani, Y. and Shimojo, S., 1999. Manifestation of scotomas created by transcranial magnetic stimulation of human visual cortex, *Nature Neuroscience* 2:767-771.

Kanizsa, G., 1979. *Organization in Vision: Essays on Gestalt Perception,* New York: Praeger.

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

Kuffler, S. W., 1953. Discharge patterns and functional organization of mammalian retina, *J. Neurophysiol.* 16:37-68.

Lamme, V. A., 1995. The neurophysiology of figure-ground segregation in primary visual cortex, *J. Neurosci.* 15:1605-1615.

Lamme, V. A. and Roelfsema, P. R., 2000. The distinct modes of vision offered by feedforward and recurrent processing, *Trends Neurosci.* 23:571-579.

Lee, T. S. and Nguyen, M., 2001. Dynamics of subjective contour formation in the early visual cortex. *Proc. Natl. Acad. Sci. USA* 98:1907-1911.

Mattingley, J. B., Davis, G., and Driver, J., 1997. Preattentive filling-in of visual surfaces in parietal extinction, *Science* 275:671-674.

McCollough, C., 1965. Color adaptation of edge-detectors in the human visual system, *Science* 149:1115-1116.

Nakayama, K. and Shimojo, S., 1990. DaVinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Res* 30:1811-1825.

Nakayama, K. and Shimojo, S., 1992. Experiencing and perceiving visual surfaces, *Science* 257:1357-1363.

Nakayama, K., Shimojo, S., and Silverman, G. H., 1989. Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects, *Perception* 18:55-68.

Nakayama, K., Shimojo, S., and Ramachandran, V. S., 1990. Transparency: Relation to depth, subjective contours, luminance, and neon color spreading, *Perception* 19:497-513.

Pessoa, L., Thompson, E., and Noe, A., 1998. Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception, *Behav. Brain Sci.* 21:723-748.

Peterhans, E. and von der Heydt, R., 1989. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps, *J. Neurosci.* 9:1749-1763.

Poggio, T., Torre, V., and Koch, C., 1985. Computational vision and regularization theory, *Nature* 317:314-319.

Pöppel, E., 1986. Long-range colour-generating interactions across the retina, *Nature* 320:523-525.

Rock, I. and Anson, R., 1979. Illusory contours as the solution to a problem, *Perception* 8:665-681.

Rossi, A. F., Rittenhouse, C. D., and Paradiso, M. A., 1996. The representation of brightness in primary visual cortex, *Science* 273:1104-1107.

Sasaki, Y., Watanabe, T., Dale, A. M., and Tootell, R. B., 2001. V1 involvement for color filling-irevealed by human fMRI, *Society for Neuroscience Abstracts*:12.11.

Shimojo, S. and Nakayama, K., 1990. Amodal representation of occluded surfaces: Role of invisible stimuli in apparent motion correspondence, *Perception* 19:285-299.

Shimojo, S., Silverman, G. H., and Nakayama, K., 1989. Occlusion and the solution to the aperture problem for motion, *Vision Res.* 29:619-626.

Shimojo, S., Kamitani, Y., and Nishida, S., 2001. Afterimage of perceptually filled-in surface, *Science* 293:1677-1680.

Spillmann, L. and Werner, J. S., 1996. Long-range interactions in visual perception, *Trends Neurosci.* 19:428-434.

Stromeyer, C. F. I., 1978. Form-colour aftereffects in human vision, In: *Handbook of Sensory Physiology: VIII. Perception* (R. Held, H. W. Leibowitz, and H. L. Teuber, eds.), Berlin: Springer-Verlag, pp. 97-142.

Sugita, Y., 1999. Grouping of image fragments in primary visual cortex, *Nature* 401:269-272.

Varin, D., 1971. Fenomini di contrasto e diffusione cromatica nell'organizzazione spaziale del campo percettivo, *Rivista di Psicologia* 65:101-128.

Virsu, V. and Laurinen, P., 1977. Long-lasting afterimages caused by neural adaptation, *Vision Res.* 17:853-860.

von der Heydt, R., Peterhans, E., and Baumgartner, G., 1984. Illusory contours and cortical neuron responses, *Science* 224:1260-1262.

Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung* 20:325-380.

Watanabe, T., 1995. Orientation and color processing for partially occluded objects, *Vision Res.* 35:647-655.

Watanabe, T., Zimmerman, G. L., and Cavanagh, P., 1992. Orientation-contingent color aftereffects mediated by subjective transparent structures, *Percept. Psychophys.* 52:161-166.

Weisstein, N., 1970. Neural symbolic activity: A psychophysical measure, *Science* 168:1489-1491.

Zhou, H., Friedman, H. S., and von der Heydt, R., 2000. Coding of border ownership in monkey visual cortex, *J. Neurosci.* 20:6594-6611.