Perception, 2009, volume 38, pages 859 – 877

doi:10.1068/ldmk-nak

Nakayama, Shimojo, and Ramachandran's 1990 paper

[Nakayama K, Shimojo S, Ramachandran V S, 1990 "Transparency: relation to depth, subjective contours, luminance, and neon color spreading" *Perception* **19** 497 – 513. Original paper reprinted in the appendix.]

Authors' update

Surfaces revisited

The target paper reviewed in this article was titled "Transparency: relation to depth, subjective contours, luminance, and neon colour spreading" coauthored by K Nakayama, S Shimojo, and V S Ramachandran, published in 1990. This paper, one of the first in a series on surface perception, examined how in untextured stereograms, local disparity and luminance contrast can drastically change surface quality, subjective contours, and the effect of neon colour spreading. When we began to conceive this and related work, the ascendant view on visual perception was derived from the pioneering studies of the response properties of visual neurons with microelectrodes, including those of Barlow (1953), Lettvin et al (1959), and Hubel and Wiesel (1959, 1962). All suggested that there are remarkable operations on the image by earliest stages of the visual pathway, which bestowed selectivity to colour, orientation, motion direction, spatial frequency, binocular disparity, etc. As such, it would seem that an understanding of vision would come through more systematic description of the properties of single neuron selectivities. This viewpoint was well summarised by Horace Barlow in his famous neuron doctrine paper (Barlow 1972), which emphasised the importance of analysing the image in successive stages of processing by neurons with specific classes of receptive fields. Later work altered this conception somewhat by seeing receptive fields as linear filters. Rather than detecting the presence of edges, bars, or otherwise perceptually identifiable elements in a scene, cells were seen as making measurements of an image. It was an 'image based' approach to vision, treating the basic operations of vision with no particular regard as to what aspects of scenes were being coded, whether a given cell's response corresponded to something about surfaces, edges, or objects in the real world.

Representing an older and radically different perspective were the views of perceptual psychologists, Irvin Rock and Richard Gregory among others. In the tradition of Helmholtz, they took a more psychological and cognitive view of visual processing, suggesting that visual perception was the result of inference, logic, or reasoning (Rock 1983), or that percepts were hypotheses (Gregory 1966/1997). Perception was a process that led to a real-world understanding of the image. This approach to vision suggested that vision was more akin to higher levels of thinking, not easily described in terms of individual neurons with their characteristic receptive fields.

Marr (1980), among others, realised that each of these two rival approaches was incomplete. He attempted a larger synthesis of vision, arguing that it must comprise distinct stages with very different properties. Analysis of the image could take place at early stages of the visual pathway, but later stages required processing that would appear to be more abstract. These stages were important to generate a viewpoint-dependent representation of surfaces in the world (the $2\frac{1}{2}$-D sketch), which then later stages matched with 3-D models of objects.

Our contribution in this and a series of related papers argues for the existence of an intermediate level of representation corresponding to visual surfaces, and to make an attempt to understand it. As evidence, we generated a variety of perceptual demonstrations relying heavily on binocularly presented stimuli to manipulate depth cues without changing the monocular image in obvious ways. Our goal was not to study the properties of stereopsis or binocular vision itself, but to use binocular disparity as a tool to manipulate depth conveniently. As such, our conclusions generalise to all aspects of perception and are not confined to the study of binocular vision.

Our efforts, while not duplicating Marr's efforts, were to argue for a similar kind of processing as that implied by the $2\frac{1}{2}$-D sketch. In particular, we argued that the level of surface representation was qualitatively different from an analysis of the image (low-level vision), yet it was also qualitatively different from the processes required for object recognition and visual praxis in the world (Goodale 1995; Ungerleider and Mishkin 1982). Moreover, this level of representation seems to be constructed in a mostly bottom–up fashion as a function of retinal stimulus. Thus, whereas it mimics what the Rock–Gregory type inferential account would predict, we argued that it is largely independent of higher-level cognitive inference. As such, we placed surface representation in-between the level of image processing and these higher stages, and claimed "... that higher functions require as an input a data format which explicitly represents the scene as a set of surfaces" (Nakayama et al 1995).

The primary discovery outlined in this paper is the fact that stereo disparity can dramatically suppress or enhance neon colour spreading. Figure 1a viewed with both eyes non-stereoscopically is the same as the neon colour spreading illusion (Redies and Spillmann 1981; Redies et al 1984). One can see that the colour of the red interior cross spills out into the dark surrounding region and is contained within a faint but distinct circular contour.

Our contribution was to add stereo disparity to this configuration and add it to just one tiny part of the figure, ie to the ends of the horizontal red cross (thus all the other edges have zero-disparity when fused). The results were dramatic. The whole phenomenon vanishes when these tiny ends are defined as in back, stereoscopically. We see an opaque surface in back, as when we are seeing a Japanese flag complete yet only partially visible behind a cruciform aperture (as depicted in figure 1b). With the ends stereoscopically in front, the phenomenon re-appears even more strongly, with an unmistakable transparent surface in front bounded by vivid subjective contours (as illustrated in figure 1c). The manipulation is local and minimal, but the effect is global and dramatic. We should point out that stereo disparity is not necessary to see these fluctuations. When viewed monocularly (when depth is more ambiguous), the same two sets of surface perception alternate with fluctuations of perceived depth.

This phenomenon and other ones that we discovered subsequently had a number of features that proved to be characteristic of surface processing. Perceived depth was found to have a profound qualitative effect on global scene appearance, not predicted by any model of vision derived from simple notions of relative disparity or receptive field properties. Such notions would predict just a simple reversal of depth, not changes in contours or material quality (from opacity to transparency).

The language of perceptual psychology, more related to ordinary common sense, had greater descriptive power in accounting for these demonstrations, albeit a posteriori. If there were a red transparent surface against an otherwise dark background in the real world, it would have the observed pattern of binocular stimulation. Likewise for the Japanese flag. If it were behind a cruciform aperture, the pattern of disparities presented would be the result.

A similar type of logical 'inference' explanation could also be applied to what we termed 'da Vinci stereopsis' (Nakayama and Shimojo 1990). Binocular images with no relative disparity give rise to a vivid perception of depth and subjective contours, a consequence of a few well-placed points seen only by one eye and not the other. From the perspective of binocular depth perception as understood from the existence of neurons selective to binocular disparity (Barlow et al 1967; Poggio and Fischer 1978), there was no ready explanation for this striking phenomenon. But from the perspective of the 'logic of perception' or Helmholtzian unconscious inference, the existence of an unpaired point in any image could have arisen plausibly only from the fact that an occluding surface covered points to one specific eye and not the other.

Another stereoscopic phenomenon that conformed to the 'logic of perception' was amodal completion. The basic idea is that, when two adjacent image regions of different depths have a common border, stereoscopic depth determines border ownership, it being always bestowed to the closer, occluding surface (Nakayama et al 1989). This leaves the farther surface essentially 'unbound' along this common contour, permitting amodal completion linking other unbounded surfaces. This predicts that amodal integration of slats containing image fragments (as in figure 2) will only occur when slats are in back. This figure gives a pictorial view of the stereo displays, where faces are more easily seen if horizontal segments are in back and likewise the vertical barber pole illusion is seen when the moving diagonal segments are also in back (Nakayama et al 1989; Shimojo et al 1989).

Yet, for da Vinci stereopsis the 'logic of perception' explanation seemed a bit odd because the primary information to produce depth was an 'eye-of-origin' signal, ie left versus right eye. Humans are very bad at reporting which eye was stimulated; and neurons very early in the visual pathway (as early as V2) lose their information about the eye of origin stimulated (Van Essen et al 1990). Thus, da Vinci stereopsis was compatible with 'reasoning', but only if we accepted the proposition that the neural substrate of this reasoning must be implemented very early, say as early as V1 or V2. As such, inferences were hypothesised to take place in exactly those same areas in which neurophysiologists were making their detailed measurements—to us an exciting prospect.

So it became obvious that different types of inference exist. The inferences implied by the term 'logic of perception' are clearly distinct from higher-order cognition and represent the autonomous processes of perception. Our next goal was to attempt to characterise the structure and origin of these inferences. Is there a special logic to them and where do they come from?

All of the perceived configurations in our demonstrations were consistent with image data. However, as we shall see, such consistency provides limited predictive power, because such images when presented to us are consistent with a number of different 3-D scenes in the real world. This is the well-known many-to-one problem of perceptual psychology as well as computational vision. Any new theory must predict what will be seen under such conditions of ambiguity.

Some authors have argued that it is the most probable events or structures in the environment that are decisive in determining what is perceived (Brunswik 1939; Purves et al 2001). This has been used to argue for an empirical theory of vision. While broadly in agreement with the role of experience of vision, we have identified the role of what is referred to as the likelihood term in the Bayes equation (Nakayama and Shimojo 1992). Given a particular surface configuration in the world, this term denotes the probability that a given image will be sampled. This term, rather than overall probability of a given surface configuration, is more predictive of what is seen. To illustrate this point, we constructed a stereogram, which consisted of red and white collinear rectangles where the contour between the red and white region was arranged to be in front, stereoscopically (Nakayama and Shimojo 1992). We identified two surface configurations that would be obviously consistent with the image data. First, would be a folded surface, which would be the result of linear depth interpolation (figure 3b). Second, would be that of a transparent red surface occluding a white bar behind (figure 3c). This too would be compatible with the image data as the white surface seen through the transparent surface would be in the back plane, whereas the edge of the red transparent surface would be in front.

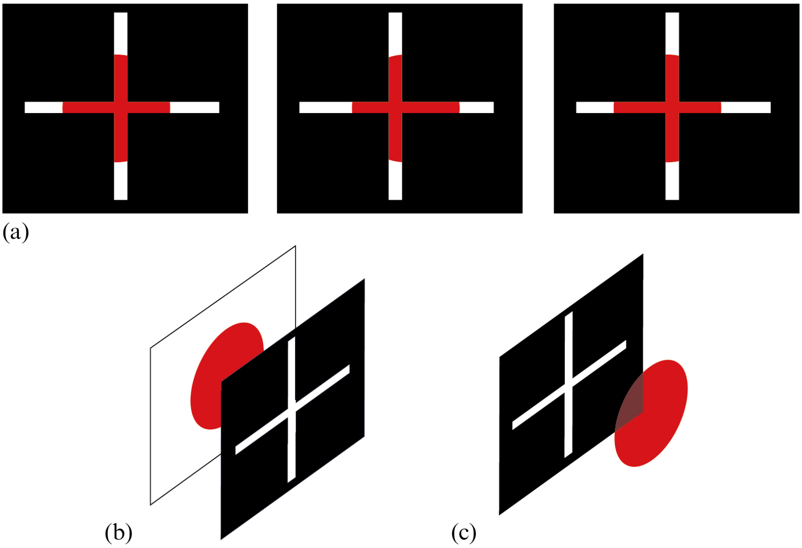

Figure 1. The basic stereo/filling-in demo by Nakayama et al (1990). (a) Stereogram. Both crossed and uncrossed fusers can experience crossed and uncrossed disparity cases side by side, as they fuse these stereo half-images. (b) and (c) illustrate typical surface perception in the uncrossed, and the crossed disparity case, respectively.

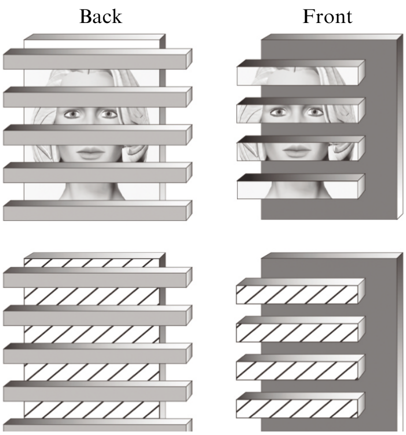

Figure 2. Effects of amodal completion in face recognition and the 'barber pole' illusion. When the horizontal fragments are seen in back, they are linked and the face is much more easy to identify (Nakayama et al 1989). The same situation applies to the barber pole illusion where fragments in back connect to form a big vertical occluded rectangle such that a vertical barber pole illusion is seen but not when the fragments are seen in front (Shimojo et al 1989).

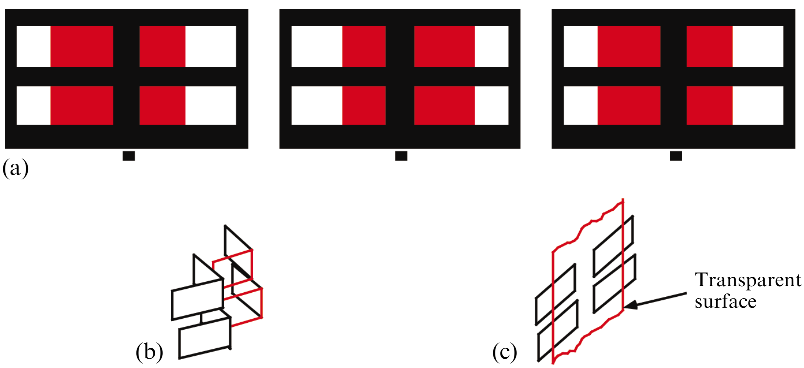

Figure 3. A stereo demonstration where the generic-view principle matters (Nakayama and Shimojo 1992). (a) Stereogram. Pay attention to the crossed disparity case. (b) An 'accidental' interpretation. (c) A more 'generic' interpretation.

What is actually seen? Will there be a similar ambiguity between these equally plausible scene interpretations, perhaps showing rivalry or some other multi-stable phenomenon? The verdict from the perspective of human perception was decisive. All observers see a vivid transparent surface in front, with colour spreading, as a dominant percept. The perception is stable and enduring. Why is this the case? Why is there such a bias to see transparency in this binocular pattern when there is an equally or even more plausible scene interpretation of opacity. Transparent surfaces are less probable surfaces in our environment, much less so than opaque surfaces. Yet, we see transparency not surface opacity. Thus one cannot make an easy argument based on the probabilities of various events or objects in scenes in the world.

With this in mind, we developed an alternative theory (Nakayama and Shimojo 1992). It came in part from J J Gibson's broad conception of ecological optics, outlining how the multiplicity of station points in space is accompanied by their own optic arrays (Gibson 1950, 1966). Thus the optic array is the sample of the visual environment from a particular point in space. Another contribution from Gibson was his emphasis on the mobile observer. During locomotion, this observer is confronted with many samples (optic arrays) of the visual scene.

With Gibson's notion of station point and optic array, it's an easy step to the well-articulated notion in computer vision, the accidental versus the generic viewpoint distinction. This states that, when confronted by an image, the observer should make the assumption that the scene is sampled from a generic not an accidental viewpoint. Examples of an accidental viewpoint are viewing a pencil along its collinear axis, viewing a plane along its extension, etc. Generic viewpoints are derived from all other station points. In the cases shown in figure 3, the stereogram presented is an accidental view of folded planar surfaces but a generic view of a red transparent surface occluding a white bar behind. To sample such a binocular image from folded surfaces requires that the observer be along just the right horizontal axis and sufficiently distant. As such, this candidate set of surfaces violates the generic-view assumption. From this it follows that we must see the transparent surface as the generic interpretation, which we indeed do.

It should be clear that these distinctions are relevant to currently popular notions of Bayesian statistics in interpreting scenes. As mentioned above, we argue that the likelihood term in the Bayes theorem is decisive in determining what we see, especially as it clearly trumps any prior term (such as prior probability distribution how common transparent surfaces are in the real world) insofar as we always see a transparent surface, never folded cards.

We have found that the generic-view principle has surprising power, making otherwise counterintuitive predictions in a wide variety of areas, accounting for the presence of subject contours (Albert and Hoffman 1994), perceived rigidity in moving displays (Kitazaki and Shimojo 1996), perception of 3-D round/curved shapes (Tse and Albert 1998) etc. We also argued (Nakayama et al 1995) that the generic-view principle by itself is enough to force border ownership assignments to surfaces in scenes. Otherwise, one has to make the assumption that surfaces from different objects tessellate perfectly from a given station viewpoint, which can hardly be the case.

We also argued (Nakayama and Shimojo 1992) that the generic-view principle is not simply an ad hoc assumption of the visual nervous system. Rather it is the result of associative learning, acquired as the mobile observer samples various optic arrays from many station points during locomotion. As such, the likelihood term is learned via associative pairing with the representation of real surfaces in the world.

More recent developments

What has happened since? Because we argued that surface processing must have been occurring in the early retinotopic cortex, physiological studies of the responses of neurons are obviously relevant. Following on the pioneering work of von der Heydt et al (1984) subjective contours as well as amodal completion phenomena were demonstrated in the response properties of V2 neurons using stereoscopic stimuli (Bakin et al 2000). Stereoscopically driven amodal completion was also shown to influence the direction of motion as predicted (Duncan et al 2000). In a series of remarkable studies, border ownership neurons have been shown to exist in V2 (Zhou et al 2000) and more recent work shows that border ownership defined monocularly and those defined binocularly are co-coded in populations of V2 neurons (Qiu and von der Heydt 2005). All these studies provide evidence that, indeed, early retinotopic cortex reflects surface processing. That these properties are evident from the earliest parts of a response further indicates that these processes are occurring in the vicinity of these areas, not determined at a higher level, then reflected by feedback or re-entry. Thus real-world occlusion constraints are implemented in the neural circuitry of early extrastriate cortex.

Computational models, emphasising the power of local intracortical networks also show that surface phenomena, border ownership, subjective contours, and amodal completion can plausibly arise from low-level inputs (Finkel and Sajda 1994; Zhaoping 2005). In computation for stereopsis, the eye-of-origin information of interocularly unpaired inputs is proved to be effective in quickly narrowing down global surface interpretations (Hayashi et al 2004; Jones and Malik 1992). Moreover, together with the occlusion constraints, the binocular system can systematically (un)suppress those unpaired regions (Hayashi et al 2004; Shimojo and Nakayama 1990).

In addition many psychophysical and psychological studies have been conducted to even more convincingly show the importance of surface processing. Such studies indicate that there are very few phenomena that are not influenced in some way by the coding of surfaces. We have room here to mention only a few. Stereopsis itself is strongly coupled with the coding of surfaces, as shown by the da Vinci stereopsis, and also in many other unsuspected ways (Anderson 1994, 1999). Already mentioned has been the coding of motion (Shimojo et al 1989), but many other studies have shown the importance of this in much greater detail with dramatic results (He and Nakayama 1994; Lorenceau and Shiffrar 1992; McDermott et al 2001; Watanabe 1997). Other studies have shown the importance of surfaces for visual search, indicating that we use surfaces rather than low-level features (He and Nakayama 1992). In fact, it has been claimed that surfaces are the first level to which we can direct our attention (He and Nakayama 1995) and that it is the earliest levels in the visual system that we have conscious access to (He and Nakayama 1994; Jackendoff 2007, page 98).

In this sense therefore we feel that our studies have been vindicated, that our original conception was fruitful, and that it has led either directly or indirectly to considerable progress in appreciating the importance of surface perception and how it might arise. Indeed, we are gratified by how effectively the concept of visual surface has been supported by a wide range of evidence.

Yet, we must acknowledge that work in this area is still in its infancy, and the framework that we have provided over a decade ago is as yet still provisional. As such, we address two issues that we feel are yet to be resolved.

Nature of mid-level surface processing, more than filtering and association?

Until now, in our discussion of transparency, we have described it simply as the end point, the result of associative learning (as summarised heuristically by the generic-view principle). Yet this would be a misleading omission, as the coding of transparency implies more. For example, whether a surface is seen as transparent determines whether, at a given location, a surface might appear at all. Thus, if all aspects of figure 1a are preserved, through the simple exchange of red and white the whole geometry of the perceived scene changes. There is a hole or aperture in the middle of the figure, surrounded by a new more extended transparent surface over the rest (Nakayama et al 1990, figure 13). Thus, perceived transparency, mediated via so-called Metelli's relations that govern luminance requirements for surface transparency perception, dictates the border ownership of the subjective contour. It now became assigned to an outer, larger surface. We wrote: "Rather than seeing transparency as a perceptual end-point, determined by seemingly more primitive processes, we interpret perceived transparency as much as a 'cause' as an 'effect'" (Nakayama et al 1990, page 497).

Subsequent research on stereopsis has supported the importance of explicit representations of transparency and opacity to dictate perceived scene layout. Consider two real-world situations: sticks that lie within a plane versus those which are arranged more randomly in depth. From the perspective of real-world knowledge, planar sticks could be embedded within a surface. The random sticks cannot. Surfaces are often opaque and can occlude, whereas a random set of sticks cannot. Subjective contours are strongly related to occlusion. Thus, planar sticks in front could lead to the perception of subjective contours, occluding regions behind. Random sticks cannot occlude and cannot give rise to subjective contours. Gillam and Nakayama's (2002) demonstration clearly supports this. In figure 4, we see that when sticks are in front and are planar, they form an occluder, leading to subjective contours. When the non-planar jumbled sticks are in front, no subjective contours are seen. A planar surface can occlude; a bunch of sticks cannot. In a similar vein, Takeichi et al (1992) demonstrated that subjective contours, formed as occluding edges, and local limited disparity information vigorously interact via explicit surface representations. Even a single dot with an uncrossed disparity is sufficient, when surrounded (even remotely) by subjective contours, to create an impression of a partly occluded surface (again, no such long-range surface formation is observed in the case of crossed disparity, thus depth symmetry). And when there is ambiguity in which subjective contours appear, very few local disparity points can be decisive.

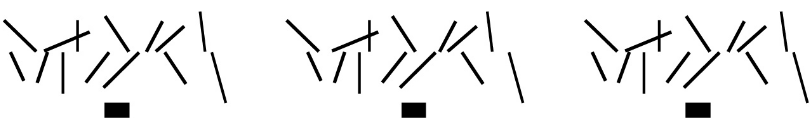

Figure 4. Stereoscopic demonstration which shows that a set of coplanar sticks can serve as an occluding surface to induce subjective contours but a set of random depth sticks cannot. When the coplanar sticks are seen in front, there is a clear subjective contour delineated in the upper and lower portion of the stereogram. This is not the case when the non-planar arrangement of sticks is in front (from Gillam and Nakayama 2002).

These and other examples reinforce the notion that surface processing requires explicit representations of surfaces in the world, through which qualities, such as contours, local depth and colour, affect each other to come up with a consistent layout of scenes. Are these high-level codings part of a necessary causal chain of 'reasoning' much as described by perceptual psychologists, or can we conceive of the process as more passive and mechanistic? Modeling of surface processing suggests that it's possible to be both mechanistic yet explicit about surface processing. Finkel and Sajda's (1994) and Zhaoping's (2005) models, for example, both use a surface property, that of border ownership, to further disambiguate the scene iteratively, using low-level units as that found in V2.

Integrity and autonomy of early surface processing?

Another major unresolved issue has to do with the level at which surfaces are determined. We have assumed as a working hypothesis that surfaces emerge very early in visual processing. This is supported by various psychophysical observations detailed above. It is further supported by strong neurophysiological signs of surface processing as early as V2. Yet, we cannot neglect the possibility of top–down knowledge, higher-level interpretations that can play some role in ambiguous situations. For example, it is possible to flip the face and vase in the famous Rubin's figure–ground demonstration just by thinking of seeing faces or vases, so there must be some obvious top–down influence, but it is limited. It can only provide a choice between several alternatives, not an infinite number of them. In addition, even very obviously plausible interpretations, such as that depicted in figure 3b, are not seen, indicating the critical role of lower-level more autonomous processes.

The key question is the degree to which bottom–up processes are adequate to parse the scene meaningfully or whether they are profoundly deficient. This latter view has been adopted by Borenstein and Ullman (2001), who argue that stored object knowledge is required, that a very large view-specific library of high-level image templates can provide a rough layout of surfaces for later refinement by lower-level mechanisms. According to them, this top–down feedback model is needed to correct the mistakes in categorising bottom–up image processing that surface processing requires.

The timing of single-unit responses is of possible relevance here. Distinctive neurophysiological signals indicating surface processing can occur at the same latency as the earliest cortical responses (Bakin et al 2000) or with a minimal delay of 25 ms (Zhou et al 2000; see also Nakayama 2005). This suggests that, if top–down processing does occur, autonomous bottom–up processing is also operative. How stored knowledge, if it is needed, gets reconciled with bottom–up information is not specified at present. We can only hope that sustained effort to resolve these questions will both illuminate the nature of surface processing itself and that it also may provide a model case for understanding functional cortical architecture and processing more generally.

Ken Nakayama
Department of Psychology, Harvard University, Cambridge, MA 02138, USA;
e-mail: ken@wjh.harvard.edu
Sinsuke Shimojo
California Institute of Technology, Pasadena, CA 91125, USA

References

Albert M K, Hoffman D D, 2000 "The generic-viewpoint assumption and illusory contours" *Perception* **29** 303–312

Anderson B L, 1994 "The role of partial occlusion in stereopsis" *Nature* **367** 365–368

Anderson B L, 1999 "Stereoscopic surface perception" *Neuron* **24** 919–928

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

Barlow H B, 1953 "Summation and inhibition in the frog's retina" *Journal of Physiology* **119** 69–88

Barlow H B, 1972 "Single units and sensation: A neuron doctrine for perceptual psychology?" *Perception* **1** 371–394

Barlow H B, Blakemore C, Pettigrew J D, 1967 "The neural mechanism of binocular depth discrimination" *Journal of Physiology* **193** 327–342

Borenstein E, Ullman S, 2001 "Class specific top–down segmentation" *Proceedings of the European Conference on Computer Vision* (Los Alamitos, CA: IEEE Computer Society Press) pp 110–112

Brunswik E, 1939 "The conceptual focus of some psychological systems" *Journal of Unified Science (Erkenntnis)* **8** 36–49

Duncan R O, Albright T D, Stoner G, 2000 "Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context" *Journal of Neuroscience* **20** 5885–5897

Finkel L H, Sajda O P, 1994 "Constructing visual perception" *American Scientist* **82** 224–237

Gibson J J, 1950 *The Perception of the Visual World* (Boston, MA: Houghton-Mifflin)

Gibson J J, 1966 *The Senses Considered as Perceptual Systems* (Boston, MA: Houghton-Mifflin)
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
Goodale M A, 1995 "The cortical organization of visual perception and visuomotor control", in *An Invitation to Cognitive Science: Visual Cognition* volume 2, Eds S M Kosslyn, D N Osherson (Cambridge, MA: MIT Press) pp 167 – 213
Gregory R L, 1966 *Eye and Brain: The Psychology of Seeing* (New York: McGraw Hill)
Hayashi R T, Maeda T S, Shimojo S, Tachi S, 2004 "An integrative model of binocular vision: a stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry" *Vision Research* **44** 2367 – 2380
He Z J, Nakayama K, 1992 "Surfaces vs. features in visual search" *Nature* **359** 231 – 233
He Z J, Nakayama K, 1994 "Surface shape not features determines apparent motion correspondence" *Vision Research* **34** 2125 – 2136
He Z J, Nakayama K, 1995 "Visual attention to surfaces in 3-D space" *Proceedings of the National Academy of Sciences* **92** 11155 – 11159
Heydt R von der, Peterhans E, Baumgartner G, 1984 "Illusory contours and cortical neuron responses" *Science* **224** 1260 – 1261
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, Wiesel T N, 1962 "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex" *Journal of Physiology (London)* **160** 106 – 154
Jackendoff R, 2007 *Language, Consciousness, Culture: Essays on Mental Structure* (Cambridge, MA: MIT Press)
Jones D G, Malik J G, 1992 "Computational framework for determining stereo correspondence from a set of linear spatial filters" *Image and Vision Computing* **10** 699 – 708
Lettvin J Y, Maturana H R, McCulloch W S, Pitts W H, 1959 "What the frog's eye tells the frog's brain" *Proceedings of the Institute of Radio Engineers* **47** 1940 – 1951
Lorenceau J, Shiffrar M, 1992 "The influence of terminators on motion integration across space" *Vision Research* **32** 263 – 273
McDermott J, Weiss Y, Adelson E H, 2001 "Beyond junctions: Nonlocal form constraints on motion interpretation" *Perception* **30** 905 – 923
Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)
Nakayama K, 2005 "Resolving border disputes in midlevel vision" *Neuron* **47** 5 – 8
Nakayama K, He Z J, Shimojo S, 1995 "Visual surface representation: a critical link between lower-level and higher-level vision", in *Visual Cognition* Eds S M Kosslyn, D N Osherson (Cambridge, MA: MIT Press) volume 2, pages 1 – 70
Nakayama K, Shimojo S, 1990 "Da Vinci stereopsis: Depth and subjective contours from unpaired monocular points" *Vision Research* **30** 1811 – 1825
Nakayama K, Shimojo S, 1992 "Experiencing and perceiving visual surfaces" *Science* **257** 1357 – 1363
Nakayama K, Shimojo S, He Z J, 1990 "Ttransparency: relation to depth, subjective contours, luminance, and neon color spreading" *Perception* **19** 497 – 513
Nakayama K, Shimojo S, Silverman G H, 1989 "Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects" *Perception* **18** 55 – 68
Poggio G F, Fischer B, 1977 "Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkeys" *Journal of Neurophysiology* **40** 1392 – 1405
Purves D, Lotto R B, Williams S M, Nundy S, Yang Z, 2001 "Why we see things the way we do: evidence for a wholly empirical strategy of vision" *Philosophical Transactions of the Royal Society of London, Section B* **356** 285 – 297
Qiu F T, Heydt R von der, 2005 "Figure and ground in the visual cortex; V2 combines stereoscopic cues with Gestalt rules" *Neuron* **47** 155 – 166
Redies C, Spillmann L, 1981 "The neon color effect in the Ehrenstein illusion" *Perception* **10** 667 – 681
Redies C, Spillmann L, Kunz K, 1984 "Colored neon flanks and line gap enhancement" *Vision Research* **24** 1301 – 1309
Rock I, 1983 *The Logic of Perception* (Cambridge, MA: MIT Press)
Rubin E, 1921 *Visuell wahrgenommene Figuren* (Copenhagen: Gyldendals)
Shimojo S, Nakayama K, 1990 "Real world occlusion constraints and binocular rivalry" *Vision Research* **30** 69 – 80
Shimojo S, Silverman G H, Nakayama K, 1989 "Occlusion and the solution to the aperture problem for motion" *Vision Research* **29** 619 – 626

Takeichi H, Watanabe T, Shimojo S, 1992 "Illusory occluding contours and surface formation by depth propagation" *Perception* **21** 177 – 184

Tse P U, Albert M K, 1998 "Amodal completion in the absence of image tangent discontinuities" *Perception* **27** 455 – 464

Ungerleider L G, Mishkin M, 1982 "Two cortical visual systems", in *Analysis of Visual Behavior* Eds D J Ingle, M A Goodale, R J W Mansfield (Cambridge, MA: MIT Press) pp 549 – 586

Van Essen D C, 1979 "Visual areas of the mammalian cerebral cortex" *Annual Reviews of Neuroscience* **2** 227 – 263

Watanabe T, 1997 "Velocity decomposition and surface decomposition—reciprocal interactions between motion and form processing" *Vision Research* **37** 2879 – 2889

Zhaoping L, 2005 "Border ownership from intra-cortical 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

Comments

Revisiting the relationship between transparency, subjective contours, luminance, and colour spreading

It is now 18 years since the publication of Nakayama et al's (1990) paper, and a cursory overview of my published research during this time period reveals the influence this work had in shaping my own intellectual pursuits. The core discovery described in this paper is the link between stereoscopic depth and neon colour spreading. Nakayama et al convincingly showed that the strength of colour spreading was modulated by relative depth, and in this and subsequent papers used stereoscopic stimuli as a tool for examining the processes and impact of 3-D surface-level representations in a variety of domains. Although the interplay between the perception of relative depth and luminance relationships was already present in Metelli's transparency research, this relationship became more vivid and dramatic in the stereoscopic stimuli used by Nakayama et al, and revealed that neon colour spreading was linked with surface-level computations. These discoveries highlighted some significant puzzles that had yet to be adequately addressed within extant theories of stereopsis, and also illuminated the underlying processes involved in neon colour spreading. Nakayama et al showed that the stereoscopic depth of image regions that lack localised disparity signals—ie regions whose stereoscopic depth was ambiguous—could be altered by varying the luminance relationships of surrounding features. This demonstration, as well as other stereoscopic phenomena observed when using sparsely textured or untextured stereograms revealed that any complete theory of stereopsis needs to explain how the disparity and photometric properties of the sparse contours present in these stereograms were used to derive the depth and surface properties of untextured regions that did not support disparity computations. These discoveries also raise questions about the generality of these effects. Are they restricted to untextured stereograms, or are they general processes used to derive surface properties from geometric and photometric image properties? Much of my own research during the past 15 years has attempted to answer these questions, and has revealed that geometric and photometric properties interact to determine the perception of transparency, depth, colour, and lightness in both stereoscopic and non-stereoscopic stimuli (eg Anderson 1997, 1999, 2003; Anderson and Julesz 1995; Anderson et al 2006; Anderson and Winawer 2005; Singh and Anderson 2002a, 2002b, 2006)

One of the most significant aspects of the phenomena reported in Nakayama et al's paper is their demonstration that relative depth has a dramatic and asymmetric role in constraining the way that contours and surfaces are interpolated [for a recent debate on this topic, see Anderson (2007a, 2007b) and Kellman et al (2007)]. Such effects reveal that relative depth is not simply a 'stage' on which surfaces, contours, or objects are placed; but, rather, that relative depth plays a significant and substantial role in determining the nature of the contours and surfaces that are constructed from the images. The fact that surface and contour appearances can be dramatically altered by simple inversions of relative depth places strong constraints on theory construction. What is the source of these asymmetries? One of the most ubiquitous arises from the geometry of occlusion: nearer surfaces occlude more distant surfaces, whereas more distant surfaces can only influence the appearance of nearer (opaque) surfaces through camouflage. And transparent surfaces—a form of partial occlusion—also introduce photometric asymmetries, as they can only reduce (or leave unaltered) the contrast of underlying surfaces. Both types of constraints play a significant role in the phenomena reported in Nakayama et al's paper. Occlusion geometry blocks the formation of illusory contours and colour spreading. The contrast variations along the 'arms' of the neon-cross stimulus or the Kanizsa variants of the neon-colour-spreading display give rise to colour spreading and percepts of transparency only when the contrast in the coloured segment is lower than in the arms. The fact that such effects are driven predominantly by luminance constraints is revealed by their red–green variant of the colour spreading. During the past 15 years, significant progress has been made in understanding precisely how 3-D geometric constraints interact with photometric constraints in producing percepts of occlusion, transparency, and colour spreading. Much of this work is indebted to some of the demonstrations that were presented in Nakayama et al's paper.

Since I was asked to provide a critique of their paper, I should note that there are a few things within the target paper that I think are not correct as stated. One source of disagreement arises in their assertion that transparency can block stereo capture. To make this point, Nakayama et al replace the textured 'mouths' of the Kanizsa inducing elements with chromatically homogeneous patches, which induces a percept of transparency without any accompanying percept of stereo capture. However, the reason for this is not because of transparency per se, as they claimed, but, rather, because the main cause of stereo capture has been removed. In the typical capture stimulus the disparity of the pie-shaped inducing elements is an integer multiple of the spacing in the wallpaper paper. In the typical capture stimulus, this causes the leftmost stripe within an inducing element to be matched with the leftmost stripe in the other half-image, giving rise to 'false' matches and stereo capture. This is missing in their transparency variant; if it is retained, stereo capture can again be observed (see figure 1).

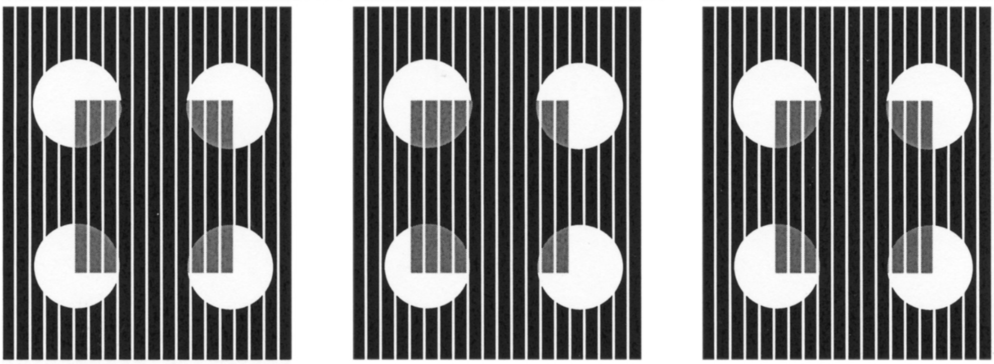

Figure 1. Transparency does not block stereo capture (left two images, divergent fusion; right two images, cross fusion). When the lines continue into the mouths of the pacmen inducing segments, both capture and transparency can be experienced.

A minor (and now pervasive) error that plagues the vision literature is the repeated use of the terms 'crossed' and 'uncrossed' to refer to relative disparities, when these terms refer to absolute disparities (ie disparities that are defined by the coordinate frame of the retinas, not by the disparities of other image features).

Finally, my greatest difference in viewpoint involves the theoretical framework that Nakayama et al presented in this and subsequent work on stereoscopic surface perception. However, to this end, I cannot offer any real complaint. Without these differences in viewpoint I would not have been inspired to develop many of the demonstrations and theoretical ideas that have emerged from my own work, which—if nothing else—has provided me with an occupation. I am therefore indebted to these theoretical differences for motivating my search for an alternative viewpoint. I should also note that this work, as well as other work that emerged from the Harvard vision lab during my postdoctoral period there, transformed my approach to vision science. It was there where I came to appreciate the powerful role that phenomenology can play in helping reveal the nature of visual processes, which I have attempted to employ at every possible opportunity. We are fortunate that—in many cases—we can literally see our data, which frees us to spend much of our time trying to understand our discoveries rather than worrying about whether they are simply methodological artifacts.

Barton L Anderson
University of New South Wales, Sydney, Australia; e-mail: barta@psych.usyd.edu.au

References

Anderson B L, 1997 "A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions" *Perception* **26** 419 – 454 (Special issue on contextual effects in lightness and colour perception)
Anderson B L, 1999 "Stereoscopic surface perception" *Neuron* **24** 919 – 928
Anderson B L, 2003 "The role of occlusion in the perception of depth, lightness, and opacity" *Psychological Review* **110** 785 – 801
Anderson B L, 2007a "The demise of the identity hypothesis, and the insufficiency and non-necessity of contour relatability in predicting object interpolation: Comment on Kellman, Garrigan, and Shipley (2005)" *Psychological Review* **114** 470 – 487
Anderson B L, 2007b "Filling-in theories of completion: Rejoinder to Kellman, Garrigan, Shipley, and Keane (2007) and Albert (2007)" *Psychological Review* **114** 509 – 527
Anderson B L, Julesz B, 1995 "A theoretical analysis of illusory contour formation in stereopsis" *Psychological Review* **102** 705 – 743
Anderson B L, Singh M, Meng J, 2006 "The perceived opacity of inhomogeneous surfaces and media" *Vision Research* **46** 1982 – 1995
Anderson B L, Winawer J, 2005 "Image segmentation and lightness perception" *Nature* **435** 79 – 83
Nakayama K, Shimojo S, Ramachandran V S, 1990 "Transparency: relation to depth, subjective contours, luminance, and neon color spreading" *Perception* **19** 497 – 513
Kellman P J, Garrigan P, Shipley T F, Keane B P, 2007 "Interpolation processes in object perception: Reply to Anderson (2007)" *Psychological Review* **114** 488 – 508
Singh M, Anderson B L, 2002a "Toward a perceptual theory of transparency" *Psychological Review* **109** 492 – 519
Singh M, Anderson B L, 2002b "Perceptual assignment of opacity to translucent surfaces: The role of image blur" *Perception* **31** 531 – 552
Singh M, Anderson B L, 2006 "Photometric determinants of perceived transparency" *Vision Research* **46** 879 – 894

Clear waters, murky waters: Why transparency perception is good for you and underconstrained

Nakayama et al (1990) mention that "transparent surfaces are fairly uncommon in our environment". One encounters mist and vapours much more often, and transparency perception could have evolved to identify objects through such atmospheric disturbances. However, there is one particular substance that has a transparent surface and is critical to our survival: water. Transparency perception could have evolved to help us determine whether water is clear and drinkable or murky and unhealthy.

Nakayama et al argued that human transparency perception does not take all information into account that is available to it. Their plate 1b shows a cross with a red inner part and green outer parts. When the red part is given a positive disparity, a red transparent disk is perceived in front of an opaque cross (see also Metzger 1955). If the cross were seen as uniformly green, as claimed by the authors, this would indeed be rather odd, because natural filters seem incapable of attenuating green in such a way that it can appear red. In principle, though, it is also possible that the cross is not perceived as uniform in colour. Underneath the red transparent surface, for example, the cross might be white (or even red), despite that elsewhere it looks green. In this case, the change in the colour of the cross would coincide exactly with the boundary of the red filter. The visual system does not seem to favour such accidental interpretations. However, boulders at a river's edge often have a different colour above and below the transparent surface of the water (eg because they are covered with dirt above and not below, or with moss below and not above) and such accidental situations may not be too very uncommon in the environment.

If there are no luminance, stereopsis, or motion cues, a mere colour difference can induce a percept of transparency. In this case, a violation of ecological validity by the use of opposite colours like red and green is possible only when these colours are highly desaturated (Chen and D'Zmura 1998; see also da Pos 1977). Yet, Nakayama et al were probably right when they concluded that transparency perception relies predominantly on luminance and much less on colour. The question that this conclusion raises is why things should be this way. Nakayama et al provided an integrative vision of the perception of transparency, depth, subjective contours, luminance, and neon colour spreading, and stated both its behavioural and neurophysiological implications. In all likelihood, this broad scope contributed to the success of the article by making it relevant to a wide range of very different studies in both psychology and neuroscience. We suggest that taking primate evolution into account might broaden this scope still further and offer a reason why transparency perception seems so little sensitive to colour cues.

We just argued that transparency perception could very well be important to our survival by helping us to determine whether water is potable or not. If this is true, transparency perception may have evolved early. Colour vision was already present in the earliest primates, but these animals were nocturnal and might not have used it much. Moreover, as it is still the case in nearly all mammals, their colour vision was dichromatic rather than trichromatic (Isbell 2006; Jacobs 1993). That is, early primates could discriminate yellow from blue, but not red from green and, consequently, they would not be sensitive to the fact that the colours in Nakayama et al's plate 1b are inconsistent with an ecologically valid form of transparency.

Suppose that the colours red and green in plate 1b were replaced by yellow and blue. A yellow filter and a blue cross absorb different wavelengths, and a blue cross, seen through a yellow filter, would normally look green (owing to subtractive colour mixing, assuming the filter absorbs at least some, but not all, of the light that the cross reflects). Since there is no green in the yellow-and-blue version of the figure, its colours are no more consistent with transparency than the colours of the original version. Lacking the ability to tell red from green, the dichromatic early primates would be incapable of detecting any ecological invalidity in the red-and-green figure. However, lacking the ability to detect the presence or absence of green, they would also be incapable of detecting any ecological invalidity in the yellow-and-blue version of the figure.

Quite possibly, therefore, transparency perception depends so little on colour because it was already in place before trichromacy started to develop. The late evolution of trichromacy might also go some way in explaining why colour has such a small effect on the perception of not just transparency, but motion and stereo depth as well.

Peter Kramer, Paola Bressan
Dipartimento di Psicologia Generale, Università di Padova, via Venezia 8, I 35131 Padua, Italy;
e-mail: paola.bressan@unipd.it

References

Chen V J, D'Zmura M, 1998 "Test of a convergence model for color transparency perception" *Perception* **27** 595 – 608

Isbell L A, 2006 "Snakes as agents of evolutionary change in primate brains" *Journal of Human Evolution* **51** 1 – 35

Jacobs G H, 1993 "The distribution and nature of colour vision among the mammals" *Biological Reviews* **68** 413 – 471

Metzger W, 1955 "Über Durchsichtigkeits-Erscheinungen (Vorläufige Mitteilung)" [About transparency phenomena (preliminary report)] *Rivista di Psicologia* **49** 187 – 189

Nakayama K, Shimojo S, Ramachandran V S, 1990 "Transparency: relation to depth, subjective contours, and neon color spreading" *Perception* **19** 497 – 513

Pos O da, 1977 "Contributo teorico sperimentale allo studio della percezione della trasparenza parziale con colori" *Atti dell'Istituto Veneto di Scienze Lettere ed Arti* **135** 47 – 70

From neon colour spreading to stratification and mechanisms

Who would have thought that a small coloured cross embedded in the centre of the humble Ehrenstein figure would give rise to a shift in theoretical outlook from single-neuron doctrine to linear filters (channels), cognitive inference, and stages of representation?

The neon effect had been described before by Varin (1971), van Tuijl (1975), and van Tuijl and de Weert (1979). Redies and I had also studied it (1981). I actually produced a pair of stereograms consisting of a matrix of 3×4 Ehrenstein figures, and viewed the illusory disks stereoscopically (figure 1). With crossed disparity (a), semitransparent 'umbrellas' could be seen in the fused image suspended between the tips of the Ehrenstein figures, whereas with uncrossed disparity (b) the same disks appeared sharply delineated and opaque.

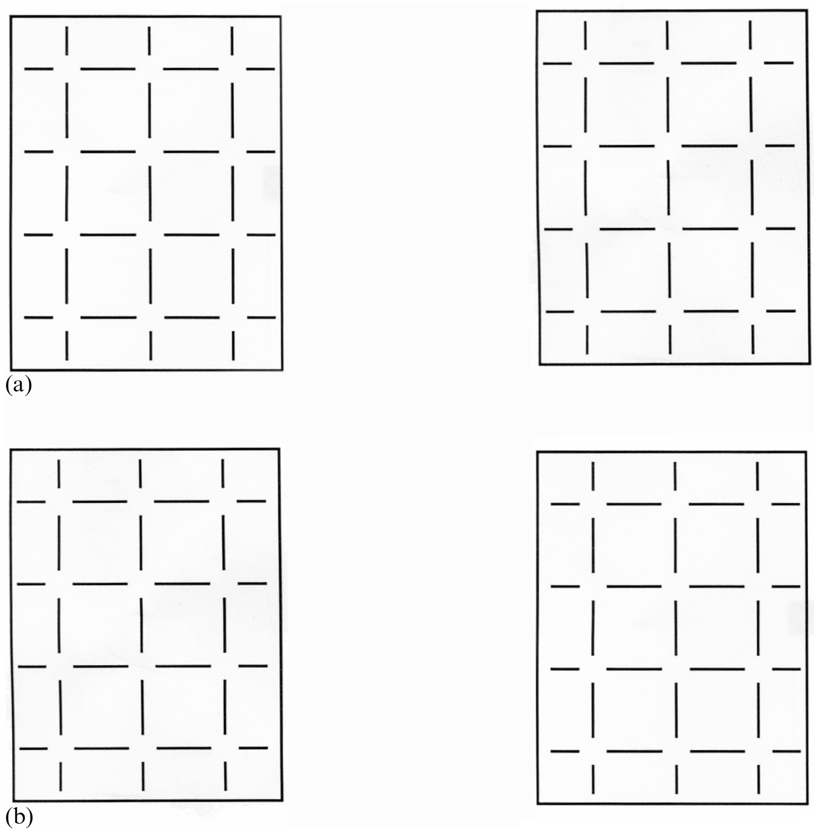

Figure 1. With free-fusing, semitransparent 'umbrellas' are seen in the upper pair of Ehrenstein patterns (a), whereas opaque disks are seen in the lower pair (b).

It took the foresight of Nakayama, Shimojo, and Ramachandran (1990) to extend neon colour spreading to depth, transparency, and surface representation. This kind of transparency was new, as it arose from the juxtaposition of four thin black or white lines with a collinear red cross in the centre of the figure, rather than the well-known superposition of stacked surfaces (Metelli 1974).

Figure 2 (from Kanizsa 1979, supplied by Todorović) demonstrates the relationship between illusory brightness, contours, and perceived depth emerging from surfaces without stereo-disparity cues. A wide vertical rectangle with dark holes in it intersects a narrow horizontal rectangle. The horizontal rectangle is typically seen as lying in front, but with some practice, can also be seen as lying behind the vertical rectangle. This is an example of stratification that can also be seen with one eye only.

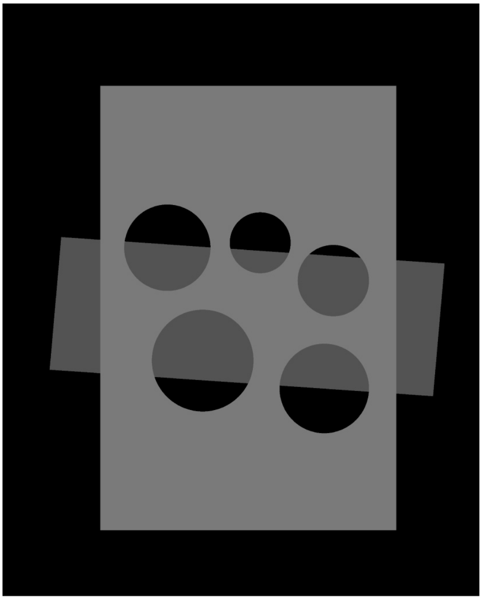

Figure 2. The horizontal rectangle appears semi-transparent and delineated by illusory contours if seen in front, but opaque and delineated by real contours if seen behind. Notice also the change in brightness from lighter to darker. (Courtesy of Dejan Todorović.)

If seen in front, illusory contours traverse the wider vertical rectangle. If seen behind (through the dark holes), the illusory contours become real contours. At the same time, the surface of the horizontal rectangle changes from lighter and transparent to darker and opaque.

In Nakayama's stereo rendition of the neon effect, the neon disk, when seen in front, similarly floats like a veil over the central gap of the Ehrenstein figure, but loses its neon-like colour and becomes opaque when seen behind.

The Zeitgeist for this kind of research was right. In 1987, at the Badenweiler Conference of the *Neurophysiological Foundations of Visual Perception*, Grossberg had ecstatically exclaimed: "I like neon". The percept was crucial to his idea of diffusion in the visual system (Grossberg and Mingolla 1985). In contrast, the findings of von der Heydt et al (1984), presented during that same meeting, had made it plausible that illusory contours were mediated by V2 neurons, receiving input from beyond the classical receptive field. Taken together with Livingstone and Hubel's (1987, 1988) landmark papers on the functional correlations of the parvocellular and magnocellular streams, these developments formed the background against which Nakayama first presented his groundbreaking ideas at the Tübingen Conference on *Visual Processing of Form and Motion* in 1988.

The subsequent claim of the authors that "higher (visual) functions require as an input a data format which explicitly represents the scene as a set of surfaces" (Nakayama et al 1995) was bold and seemed to run counter to the prevailing idea that the visual system codes uniform stimulus areas primarily in terms of its borders (stick figures). After all, neurons respond poorly to uniform stimuli, begging the question why we see surfaces at all instead of mere skeletons (Marr's 1982, primal sketch).

Today we have evidence that surfaces are preserved—as well as reconstructed—by eye movements and that in their absence (ie with strict fixation) the brain fills in a faded percept with brightness, colour, and texture propagated from the surround (Paradiso and Nakayama 1991; de Weerd et al 1998; Paradiso et al 2006). The neon effect in the Ehrenstein figure is an example of colour spreading out from the cross to the illusory contour (Bressan et al 1997). The subsequent stratification into a transparent overlay can be seen without stereo cues, but is dramatically enhanced by crossed disparity. Examples of lightness changes due to the decomposition of luminance differences into multiple layers and transparency can be found in the work of Anderson (1997) and

Anderson and Winawer (2005) among others [see the Special Issue on Lightness, and Colour Induction, Transparency, and Illumination—*Perception* 1997(4)].

The opacity of the neon disk obtained with uncrossed disparity follows from the loss of border ownership when the central area loses its 'figure' status and is seen as ground. Neurons in area V2 with a 'side-of-figure' selectivity have been found representing borders of 2-D figures (Zhou et al 2000). These same neurons are also activated by 3-D displays and thus are likely to mediate stereoscopic depth information (Qiu and von der Heydt 2005). Most relevant to the topic under consideration, neurons that assign border ownership according to what we see as transparent, have recently been identified in area V2 (Qiu and von der Heydt 2007).

Amodal completion, illusory contours, border ownership, and transparency have all been shown to have neurophysiological correlates at low levels within the visual system (Spillmann 2009). Functional MRI studies of neon colour spreading support the findings from single-cell recordings (Sasaki and Watanabe 2004). It thus seems that 18 years after Nakayama et al's seminal paper, we are well on our way to an understanding of the processes and mechanisms that underlie the perception of surfaces.

Lothar Spillmann
Department of Neurology, Neurocenter, University of Freiburg, Freiburg, Germany;
e-mail: lothar.spillmann@zfn-brain.uni-freiburg.de

References

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, Winawer I, 2005 "Image segmentation and lightness perception" *Nature* **443** 79–83

Bressan P, Mingolla E, Spillmann L, Watanabe T, 1997 "Neon color spreading" *Perception* **26** 1353–1366

Grossberg S, Mingolla E, 1985 "Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading" *Psychological Review* **92** 173–211

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

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

Livingstone M S, Hubel D H, 1987 "Psychophysical evidence for separate channels for the perception of form, color, movement, and depth" *Journal of Neuroscience* **7** 3416–3468

Livingstone M S, Hubel D H, 1988 "Segregation of form, color, movement, and depth: anatomy, physiology, and perception" *Science* **240** 740–749

Marr D, 1982 *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (San Francisco, CA: W H Freeman)

Metelli F, 1974 "The perception of transparency" *Scientific American* **239**(4) 90–98

Nakayama K, He Z J, Shimojo S, 1995 "Visual surface representation: a critical link between lower-level and higher-level vision", in *Visual Cognition* Eds S M Kosslyn, D N Osherson (Cambridge, MA: MIT Press) pp 1–70

Nakayama K, Shimojo S, Ramachandran V S, 1990 "Transparency: relation to depth, subjective contours, and neon color spreading" *Perception* **19** 497–513

Paradiso M A, Blau S, Huang X, MacEvoy S P, Rossi A F, Shalev G, 2006 "Lightness, filling-in, and the fundamental role of context in visual perception", in *Progress in Brain Research* volume 155, Eds S Martinez-Conde, S L Macknik, L M Martinez, J-M Alonso, P U Tse (Amsterdam: Elsevier) pp 109–123

Paradiso M A, Nakayama K, 1991 "Brightness perception and filling-in" *Vision Research* **31** 1221–1236

Qiu F T, Heydt R von der, 2005 "Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules" *Neuron* **47** 155–166

Qiu F T, Heydt R von der, 2007 "Neural representation of transparent overlay" *Nature Neuroscience* **10** 283–284

Redies C, Spillmann L, 1981 "The neon color effect in the Ehrenstein illusion" *Perception* **10** 667–681

Sasaki Y, Watanabe T, 2004 "The primary visual cortex fills in color" *Proceedings of the National Academy of Sciences of the USA* **101** 18251–18256

Spillmann L, 2009 "Phenomenology and neurophysiological correlations: Two approaches to perception research" *Vision Research* (in press)

Tuijl H F J M van, 1975 "A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours" *Acta Psychologica* **39** 441 – 445

Tuijl H F J M van, Weert C M M de, 1979 "Sensory conditions for the occurrence of the neon spreading illusion" *Perception* **8** 211 – 215

Varin D, 1971 "Renomeni di contrasto e diffusione cromatica nell'organizzazione spaziale del campo percettivo" *Rivista di Psicologia* **65** 101 – 128

Weerd P de, Desimone R, Ungerleider L G, 1998 "Perceptual filling-in: a parametric study" *Vision Research* **38** 2721 – 2734

Zhou H, Friedman H S, Heydt R von der, 2000 "Coding of border ownership in monkey visual cortex" *Journal of Neuroscience* **20** 6594 – 6611

Authors' response

Lothar Spillmann rightly brings up a critical turning point in the history of neon colour spreading reminding us of importance of Steve Grossberg's ecstatic outburst "I like neon". One of us (Nakayama) had never seen this demonstration and was thus enthralled and mystified. Nakayama had absolutely no clue why there should be neon despite its obvious phenomenological similarity to transparency and he and many others did not make this connection. Most of us trained in the rigours of psychophysics or neurophysiology were looking for some kind of low-level mechanistic interpretation to explain this amazing illusion.

But enter technology. Deluxe Paint, a software program for the Amiga computer, allowed the instant construction of a variety of geometrical images. Nakayama was aimlessly drawing geometric patterns on the screen and for some unknown reason created the stereogram shown in figure 1. When the edge of the red was in front, neon colour spreading was more vivid than ever. When in back, the neon vanished, gone! Bart Anderson is right. Phenomenology under the right circumstances is one of the most powerful and thrilling ways to study perception.

After this, we couldn't get transparency out of our minds. Neon colour spreading had to be related to transparency. Eventually, we began to accept the 'logic of perception', skeptical and untutored as we were regarding the ideas of Irvin Rock and other perceptual psychologists. And, as Spillmann indicates in his reproduction of the Kanizsa figure (his figure 2), we had numerous other demonstrations to show that it had to do with relative depth more generally. Stereopsis was just a method to create relative depth reliably.

Kramer and Bressan imply our neglect of the natural existence of transparent surfaces in the world. We were well aware of naturally occurring transparency: water, Egyptian see-through fabrics, etc and acknowledge that the prior probability of transparency is certainly non-zero (Nakayama and Shimojo 1992). But, the number of transparent surfaces is tiny compared to the number of opaque surfaces. So, reasoning from prior probability, ie environmental encounters, we should expect to see opaque surfaces (figure 3b), most of the time. Yet subjects only see transparent ones. Such an outcome is predicted by the generic view principle. In adopting this principle, we pinpoint the likelihood term of the Bayes theorem which we hypothesise to be learned from experience as an organism locomotes through the world and sees elementary surface configurations from a variety of viewpoints.

In closing we wish to salute this journal, *Perception*, and its illustrious founder, Richard Gregory, for being the muse of so much quirky creative research reflected in this special issue that we are honored to be part of.

Reference

Nakayama K, Shimojo S, 1992 "Experiencing and perceiving visual surfaces" *Science* **257** 1357 – 1363

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 38 2009

www.perceptionweb.com