Depth propagation across an illusory surface

Shigeaki Nishina

ATR Computational Neuroscience Laboratories, 2-2-2 Hikaridai, Seika, Soraku, Kyoto 6190288, Japan

Arash Yazdanbakhsh

Harvard Medical School, 220 Longwood Avenue, Boston 02115, Massachusetts, USA

Takeo Watanabe

Boston University, 64 Cummington, Boston 02215, Massachusetts, USA

Mitsuo Kawato

ATR Computational Neuroscience Laboratories, 2-2-2 Hikaridai, Seika, Soraku, Kyoto 6190288, Japan

Received June 12, 2006; revised November 13, 2006; accepted November 20, 2006;
posted November 22, 2006 (Doc. ID 71749); published March 14, 2007

We investigate the spatiotemporal dynamics of depth filling in on an illusory surface by measuring the temporal asynchrony of perceived depth between an illusory neon-colored surface and real contours. We temporally modulated the horizontal disparity at vertical edges of the illusory surface and measured the perceptual delay for the interpolated surface's depth under two different boundary conditions: disparity given at both sides, or disparity given at one side and a free boundary at the other side. The results showed that the amount of the delay depends on the spatial distance between the measured point and the edges where disparity was physically given. Importantly, the observed delay as a function of spatial distance was clearly different under the two boundary conditions. We found that this difference can be fairly well explained by a model based on a diffusion equation under different boundary conditions. These results support the existence of locally represented depth information and an interpolation process based on mutual interaction of this information.

OCIS codes: 330.1400, 330.5510, 330.4060.

1. INTRODUCTION

While there are multiple depth cues in a natural scene, a random-dot stereogram shows that binocular disparity can provide sufficient information for one to perceive a surface.[1] Even a random-dot stereogram with an extremely sparse distribution of dots (5%) can induce a smooth and unique surface perception. This indicates that the visual system fills in depth (or surface media) between given local binocular disparity signals to reconstruct a unique surface perception.

To understand the mechanism of filling in, it is essential to examine the dynamic nature of the filling-in process. Dynamic aspects of filling in have been empirically examined. Davey *et al.* investigated the temporal properties of the brightness changes due to the Craik–O'Brien–Cornsweet effect (COCE).[2] They used a COCE pattern with a continuously reversing edge contrast and showed that changes in perceived brightness depend on the spreading rate of the pattern and its distance from the edge. Paradiso and Nakayama investigated the temporal dynamics of brightness filling in by using a visual masking paradigm.[3] Their results were consistent with the hypothesis that brightness signals are generated at the borders of the target stimuli and propagate inward at some rate.[4,5] The temporal properties of brightness perception that influence filling in have also been examined in terms of contextual modulation,[6,7] using visual patterns that cause brightness assimilation or contrast. While the results of these studies suggest that filling in occurs as a result of gradual spreading from a border in the case of a brightness filling in, the temporal progression of filling in was not directly tested with a time course.

Recently, Nishina *et al.* directly measured delays in depth interpolation.[8] In this study, a horizontal bar with uniform luminance was presented, and the disparity of its vertical edges on left and right ends was periodically changed. Although no depth information is available except at the vertical edges, this setting introduced a vivid perception of depth motion at the vertical edges but also induced depth motion perception throughout the entire bar. The subjects were asked to match the phase of the periodical depth motion of the horizontal bar with a probe, which was moving in depth at the same frequency and amplitude as the ends of the bar. It was found that the temporal delay, calculated from the measured phase differences, increased monotonically as a function of the distance between the probe and the ends of the bar.

A computational model[9] was formulated based on a diffusion equation, which postulates that the depth of the center of the object is calculated using the depth information provided at both ends. The model assumes that depth information propagates over a surface in the same man-

1084-7529/07/040905-6/$15.00

ner that heat is conducted along a material from two heat sources at its ends. The dynamics of this physical phenomenon has been well studied. When the temperature of the heat sources changes, the temperature of the material also changes with a temporal delay that depends on the distance from the sources. Psychophysical data examining depth perception fits this model well, with a quadratic relationship between delay and spatial distance when depth information is provided at the ends of an object.

The current study aimed to further empirically test the validity of the diffusion model of illusory depth perception. This model predicts a greater delay in surface formation when depth information is available only at one end (one-side condition) than when information is provided at both ends. The model further predicts that there will be a linear relationship between delay and spatial distance under the one-side condition, instead of the quadric relationship that experimental and computational investigations have shown under double-side condition.[9]

In the present study, disparity information was presented independently at one end and at both ends of an illusory surface to test the validity of the diffusion model. In contrast to the luminance-defined surface used in our previous work,[8] we created an illusory neon-colored surface induced by four circles with colored sections (Fig. 1), because it is difficult to have disparity information at a single side with a luminance-defined surface. In the illusory neon-colored surface, the edge at one side of the surface can be eliminated by removing two circles at the side [Fig. 1(b)]. A neon-colored surface was used rather than an ordinary Kanizsa pattern,[10] since the percept of the neon-colored surface is much more stable and this made the task easier, which made the results more reliable.

(a)

(b)

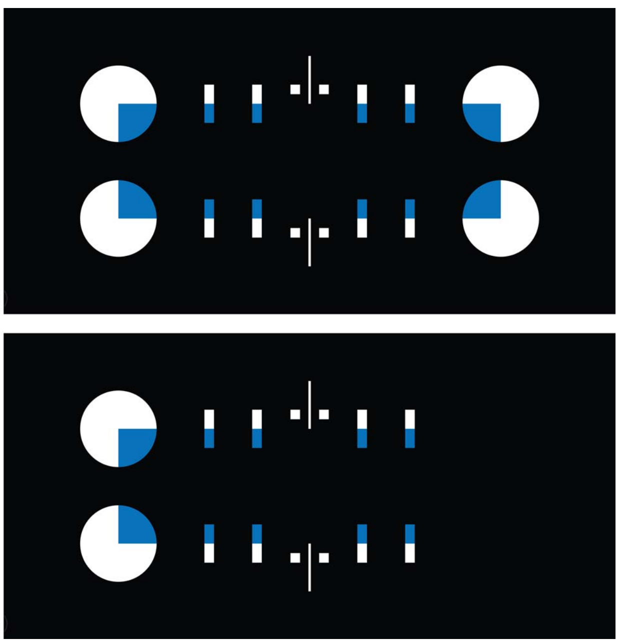

Fig. 1. (Color online) Typical spatial configurations of the stimuli are shown. Gray areas were presented in dark blue in the actual experiment. In the double side condition, (a) four circles with blue sections were presented as inducers at the four corners of the illusory surface. In the single-side condition, (b) only two circles were presented at one side. Small rectangles with blue and white colors were presented as supplementary inducers. The four square dots around the center were fixation dots, and the two vertical lines at the center were comparison stimuli.

2. EXPERIMENT

A. Apparatus and Subjects

Two 21″ CRT monitors controlled by a Macintosh Computer running MATLAB and Psychophysics Toolbox (http://www.psychtoolbox.org/) were used to present binocular images. The observers viewed the images through a haploscope that was configured so that the view distance and the vergence were correctly matched. Subjects were seated in a chair and responded with a button press on a keyboard. The screen refresh rate was 70 Hz with a resolution of 1600×1200 pixels. The view distance was 0.8 m, and the pixel size was 1.13 arc min. A chin rest was used to maintain head position. Three volunteer subjects, who were naïve about the purpose of the experiment, and one of the authors participated in the experiment. Two of those four subjects and another naïve subject participated in a control experiment described in the following sections.

B. Stimuli

We adopted a neon-color filling-in pattern that has four circular inducers, white pacmen with blue wedges at their mouths [Fig. 1(a)].[11,12] When this pattern is statically presented, a vivid bluish area is perceived, surrounded by the pacmen. This percept is very robust and even naïve observers easily perceive a filled-in surface. In this experiment, we added horizontal disparity at the vertical edges of the borders between the white areas and blue areas of inducers, so that the observers perceived the filled-in surface as being closer to the observer than the inducers (Fig. 2).[13] The radius of the circles was 1.13°, and the length of the subjective vertical edges were 3.78°. It is important to note that the horizontal edges of the circlelike inducers and the rectangular inducers do not have disparity information. In another condition, only two circles were presented [Fig. 1(b)]. The disparity of the vertical edges was temporarily varied by two-dimensionally moving the edges presented on both monitors according to a sinu-

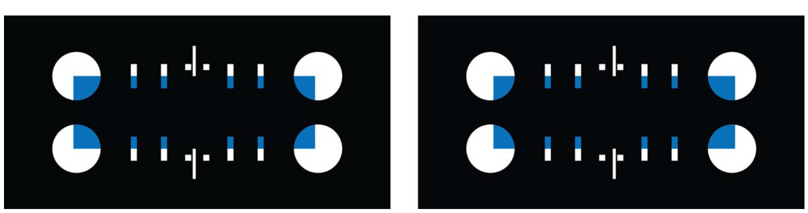

Fig. 2. (Color online) Static example of the binocular stimulus for cross fusers. Although these two images differ only at the vertical edges in the circle (and the vertical edges at the center), the entire filled-in surface is perceived as popping out in depth.

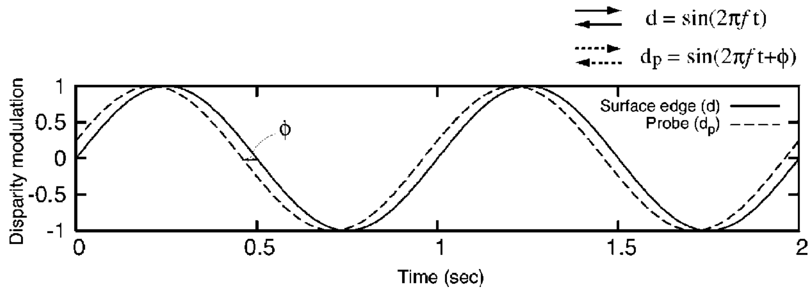

Fig. 3. Disparities of the vertical edges in the circles and the probe stimulus were updated at every display frame according to these equations. Both were sinusoidally modulated with the same amplitude and frequency, but they could have different phases.

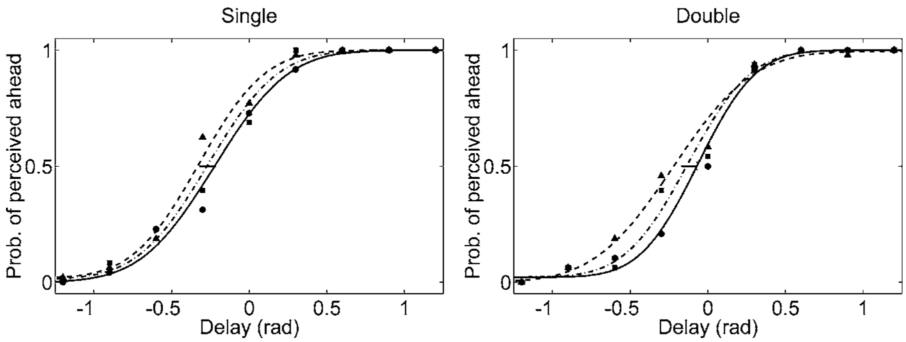

Fig. 4. Psychometric curves show the ratio of ahead response as functions of phase difference. The circles, squares, and triangles represent the different horizontal lengths of the stimuli, 4.0°, 5.0°, and 6.0°, respectively. Curves are sigmoid functions fit to the data.

soidal function of 1.0 Hz. The amplitude of the disparity oscillation was 5.7 arc min, which was twice the maximum displacement in each image. Since the mean amplitude was equal to the amplitude, the disparity was always crossed, and the surface was always perceived as being closer to the observer than the inducers. Two vertical lines presented around the center were used as comparison stimuli. They were presented just above and below the filled-in surface without any overlap. The lines were moving in depth in the same frequency and amplitude as the edges of the inducers, but their phase was set to different values (Fig. 3). The phase difference was 0.0, ±0.3, ±0.6, ±0.9, or ±1.2 in radians. Four small fixation dots were presented beside the vertical lines. The subjects were asked to gaze at one of the four dots during the task, yet subjects were allowed to move their eyes from one dot to another in order to keep the filled-in surface from fading due to adaptation at the edges. The depth of the fixation points was the same as the depth of the inducers. Eight additional small rectangular inducers were presented between the fixation dots and the primary inducers. The border of blue and white areas of these inducers were aligned to the horizontal edges of the filled-in surface, such that they enhanced the perception of neon-color filling in but had little effect on depth filling in.

C. Procedure

The task was to compare the depth motion of the probe and the center of the filled-in surface. The phase difference between the probe and the vertical edges of the inducers was randomized across trials.) In each trial, the subjects were asked to compare the depth motion and answer whether the probe was moving ahead or behind the center of the surface. There were two conditions, the double-side condition [Fig. 1(a)], and the single-side condition [Fig. 1(b)]. Four circlelike inducers were presented in the double-side condition, and only two of them on the left or right side were presented in the single-side condition. Each subject completed three sessions for each condition, for a total of six sessions, which required approximately 3 h. Each session consisted of 84 trials. The order of the two conditions was counterbalanced among subjects. Trial types in a session varied by the horizontal distance between the inducers and the probe (4.0°, 5.0°, or 6.0°) and the side of the inducers in the sessions of the single-side condition (left or right), with the order of presentation being pseudorandomized. The eight rectangular inducers were equally presented in both sessions. The depth motion was continuously presented in a trial until a response was made. The subjects responded by pressing one of two keys on a keyboard. A control experiment was conducted to determine whether the illusory surface was actually mediating the depth information. In this control experiment, two real contours were presented at the same location as the illusory vertical edges instead of the inducers, and the subject was asked to compare the depth of the real contours with the stimuli presented around the center.

D. Results

The averaged results are shown in Fig. 4. The responses as function of the delay were fit to sigmoid functions to obtain points of subjective equality. Small deviations and steep slopes indicate that the subjects made consistent responses. The phase delays at 50% of the points were aggregated and are shown in Fig. 5. We found that there was a qualitative difference between the responses to the two boundary conditions, and this difference was well explained by two simple equations. Here the goodness of fit is assessed by the R^2 statistic. With a reasonable assumption of no delay at zero distance, the relationship between the delay and the distance was considered to be nonlinear in the double-side condition, and the averaged result fitted well with a quadratic function ($R^2=0.902$). On the other hand, the relationship in the single-side condition was nearly linear ($R^2=0.988$). Those data are fitted poorly with the opposite functions ($R^2=0.579$ and 0.0850, respectively). The light gray lines in Fig. 5 are the result of the control experiment. There is no clear delay or advance at all without an illusory surface, indicating the reported delay with the illusory surface is not attributed to the temporal difference between foveal and peripheral locations. In the following section, we show how our model explains this qualitative difference between the two boundary conditions.

3. COMPUTATIONAL MODEL

In our previous study,[9] we showed that a model based on the heat conduction equation can well predict temporal properties of depth interpolation. Here, we applied the model to the psychophysical results of the experiment described in the previous section that differentiates the two different boundary conditions.

The visual system reconstructs the ambiguous depth of the central area using an unambiguous depth available at the borders. This process can be considered as a propagation of depth information. We consider a heat conduction equation as a model of this propagation process. The depth D behaves as temperature in the equation. Suppose the time parameter τ [s] and the space parameter λ [m] are constant and do not depend on the oscillation frequency of the stimulus.

$$\tau\frac{\partial D}{\partial t} = \lambda^2\frac{\partial^2 D}{\partial x^2}. \quad (1)$$

This equation signifies that the depth information diffuses over space. Under a static condition, the model reconstructs a surface with uniform depth after a certain amount of time. To clarify the temporal characteristics of

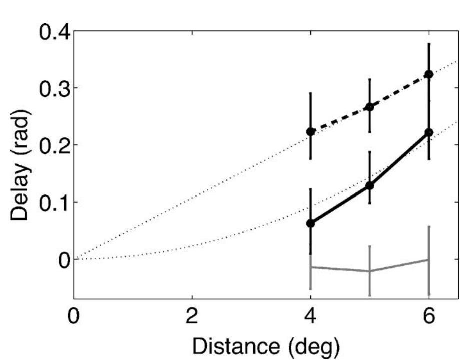

Fig. 5. Phases at which the filled-in surface and the probe were perceived in phase were shown for each distance between the bar's ends and the probe. The dots connected with the black solid line represent the double-side condition, and the dots connected with the thick dashed line represent the single-side condition. The dotted curves are the model predictions. The gray lines at the bottom are the result of a control experiment, where no illusory surface was perceived.

the model, let us consider the sinusoidally varying depth of the ends. Rewriting the previous expression using the following equation:

$$\frac{\lambda^2}{\tau} = \frac{v^2}{2\omega}, \quad (2)$$

where v [m/s] is the propagation velocity, $\omega=2\pi f$ [s^2] is the angular frequency at the ends when $f\,[s^{-1}]$ is the frequency. Now we can calculate the velocity:

$$v = \sqrt{\frac{2\lambda^2\omega}{\tau}}. \quad (3)$$

Using the velocity and the angular frequency, we can rewrite the heat conduction equation as

$$\frac{\partial D}{\partial t} = \frac{v^2}{2\omega}\frac{\partial^2 D}{\partial x^2}. \quad (4)$$

Then we obtain

$$D_{\text{single}}(x,t) = C\exp\left(\frac{\omega}{v}x\right)\sin\left(\omega t - \frac{\omega}{v}x + \theta\right)$$

$$= C\exp\left(\frac{\omega}{v}x\right)\sin\left\{\omega\left(t - \frac{x}{v}\right) + \theta\right\}, \quad (5)$$

where C is an arbitrary constant. This equation reads that when oscillating depth information at a point propagates toward a one-dimensional direction, the phase of the oscillation is delayed according to the distance. The delay as a function of the distance L can be written as

$$\Phi \sim \frac{\omega}{v}L. \quad (6)$$

The propagation speed is v. This equation corresponds to the single-side condition in the experiment and predicts linear relationship between the delay of the oscillation and spatial distance.

The delay under the double-side condition is the addition of the above and one obtained by reversing the sign of the spatial variable of it:

$$D_{\text{double}}(x,t) = D^{+}_{\text{single}}(x,t) + D^{-}_{\text{single}}(x,t),$$

$$D^{+}_{\text{single}}(x,t) = c\exp\left(\frac{\omega}{v}x\right)\sin\left(\omega t - \frac{\omega}{v}x + \theta\right),$$

$$D^{-}_{\text{single}}(x,t) = D^{+}(-x,t) = c\exp\left(-\frac{\omega}{v}x\right)\sin\left(\omega t + \frac{\omega}{v}x + \theta\right). \quad (7)$$

We can then calculate the depth at the center of the horizontal bar ($x=0$) and at the ends ($x=L$):

$$D(0,t) = 2c \sin(\omega t + \theta),$$

$$\begin{aligned} D(L,t) &= D(-L,t) \\ &= c \exp\left(\frac{\omega}{v}L\right)\sin\left(\omega t - \frac{\omega}{v}L + \theta\right) \\ &\quad + c \exp\left(-\frac{\omega}{v}L\right)\sin\left(\omega t + \frac{\omega}{v}L + \theta\right) \\ &= Ac \sin(\omega t + \theta - \Phi), \end{aligned} \tag{8}$$

where

$$A(L) = \sqrt{\exp\left(2\frac{\omega}{v}L\right) + \exp\left(-2\frac{\omega}{v}L\right) + 2\left\{1 - 2\sin^2\left(\frac{\omega}{v}L\right)\right\}},$$

$$\Phi = \arctan\left[\frac{\left\{\exp\left(\frac{\omega}{v}L\right) - \exp\left(-\frac{\omega}{v}L\right)\right\}\sin\left(\frac{\omega}{v}L\right)}{\left\{\exp\left(\frac{\omega}{v}L\right) + \exp\left(-\frac{\omega}{v}L\right)\right\}\cos\left(\frac{\omega}{v}L\right)}\right]. \tag{9}$$

Suppose that $(\omega/v)L$ is much smaller than 1. Then the phase delay is well approximated as a concave quadratic function:

$$A(L) \sim 2 + o(L^2),$$

$$\Phi \sim \left(\frac{\omega}{v}L\right)^2. \tag{10}$$

Contrary to the single-side condition, the delay increases nonlinearly with spatial distance in the double-side condition. It should also be noted that the amplitude of oscillation at the center is smaller than that of the edges, but this difference is very small (small order of the square of the bar length).

In Fig. 5, the dotted lines represent model behaviors calculated by fitting the obtained model functions to the psychophysical results. The black solid line displays responses to the double-side condition, and the thick dashed line displays responses to the single-side condition. The different slopes of the two boundary conditions are inherent in the model itself. The model has only one free parameter v, which corresponds to the propagation speed. The psychophysical results could be fit by the model using the same value $v=117.2$ [deg/s] for both boundary conditions, which is also roughly equivalent to the propagation speed reported in the previous study ($v=95.3$ [deg/s] under a comparable condition).[9] The model predicted the data of both conditions simultaneously, indicating the existence of qualitatively different process between two boundary conditions.

4. DISCUSSION

The current study demonstrated that the different results in depth filling in from two different boundary conditions are well predicted by a computational model based on a diffusion equation. To measure a temporal delay for depth interpolation, we asked the subjects to compare phases of the interpolated surface and a comparison stimulus. The results showed that the delay in response is dependent on the spatial distance, which was a replication of our previous results obtained using a luminance-defined surface.[8,9] The results also verified our model's prediction that there would be different spatial dependencies under the two different boundary conditions, one with disparity edges on both sides and the other only on a single side. While the time to propagate the depth is linear with disparity on a single edge, it is quadratic with disparity on both edges. The dependency of the delay on spatial distance is qualitatively different between the two boundary conditions, and this difference agrees extremely well with the model predictions.

The close agreement between the human performance and the proposed model based on the diffusion equation clarifies some fundamental properties of depth perception. First, depth is represented in a spatially localized form and cannot be responsively updated according to a change of input because of its inertial dynamics. In addition, the interaction of the spatially localized representation would be based on short-range connections, at least when the local input is not available and interpolation is necessary. The model assumes spatially uniform distribution of neurons that represent depths on the surface, which is not the case in the visual cortex. In our previous experiment,[8] we examined the effect of vertical location of the stimulus on depth perception and found that propagation speed is faster when a surface is presented peripherally. This is thought to be attributed to the different density of neurons between foveal and peripheral areas. If the propagation speed between two arbitrary neurons is constant, overall speed is faster if mediating neurons are spatially sparse. We can consider such physiologically more plausible assumptions to improve the model and to provide a better understanding of the dynamical mechanism of filling in.

In conclusion, a temporal property of human depth interpolation was highly predicted by a diffusion model based on the local connection of locally represented depth information and iterative processing. These findings strongly support a propagation mechanism based on neural spreading.

ACKNOWLEDGMENT

This research was supported in part by the National Institute of Information and Communications Technology (NICT).

Shigeaki Nishina's current address is the Department of Psychology, Boston University, 64 Cummington Street, Boston, Massachusetts 02215. His e-mail address is nishina@bu.edu.

REFERENCES

1. B. Julesz, *Foundations of Cyclopean Perception* (U. of Chicago Press, 197l).
2. M. P. Davey, T. Maddess, and M. V. Srinivasan, “The spatiotemporal properties of the Craik–O’Brien–Cornsweet effect are consistent with ‘filling-in’,” Vision Res. **38**, 2037–2046 (1998).
3. M. A. Paradiso and K. Nakayama, “Brightness perception and filling-in,” Vision Res. **31**, 1221–1236 (1991).
4. S. Grossberg and E. Mingolla, “Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading,” Psychol. Rev. **92**, 173–211 (1985).
5. K. F. Arrington, “The temporal dynamics of brightness filling-in,” Vision Res. **34**, 3371–3387 (1994).
6. R. L. De Valois, M. A. Webster, K. K. De Valois, and B. Lingelbach, “Temporal properties of brightness and color induction,” Vision Res. **26**, 887–897 (1986).
7. A. F. Rossi and M. A. Paradiso, “Temporal limits of brightness induction and mechanisms of brightness perception,” Vision Res. **36**, 1391–1398 (1996).
8. S. Nishina, M. Okada, and M. Kawato, “Spatio-temporal dynamics of depth propagation on uniform region,” Vision Res. **43**, 2493–2503 (2003).
9. S. Nishina and M. Kawato, “A computational model of spatio-temporal dynamics in depth filling-in,” Neural Networks **17**, 159–163 (2004).
10. G. Kanizsa, “Margini quasi-percettivi in campi con stimulazione omogenea,” Rivista di psicologia **49**, 7–30 (1955).
11. Y. Sasaki and T. Watanabe, “The primary visual cortex fills in color,” Proc. Natl. Acad. Sci. U.S.A. **101**, 18251–18256 (2004).
12. H. F. v. Tuijl, “A new visual illusion: neonlike color spreading and complementary color induction between subjective contours,” Appl. Phys. (Berlin) **39**, 441–445 (1975).
13. K. Nakayama, S. Shimojo, and V. S. Ramachandran, “Transparency: relation to depth, subjective contours, luminance, and neon color spreading,” Perception **19**, 497–513 (1990).