Pergamon

PII: S0042-6989(96)00156-3

Vision Res., Vol. 37, No. 5, pp. 633–644, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Binocular Matching of Dissimilar Features in Phantom Stereopsis

LEI LIU,*‡ SCOTT B. STEVENSON,*† CLIFTON M. SCHOR*

Received 17 August 1995; in revised form 15 February 1996; in final form 23 May 1996

Previously we have demonstrated that quantitative depth perception can be elicited from a stereogram that lacks contrast defined binocular corresponding elements (phantom stereopsis). In this report, we use computer simulation to demonstrate that it is biologically plausible for some known binocular cortical cell types to combine non-conventional matching features. Therefore, binocular matching processes based on the responses of these cells could be a conventional one, namely, looking for similar response patterns in the two eyes. While all cell types we simulated gave identical disparity outputs to the conventional stereogram, they responded differently to the phantom stereogram. Processes other than low-level disparity detectors may have to be invoked in order to achieve a unique depth solution. © 1997 Elsevier Science Ltd. All rights reserved.

Stereopsis Binocular matching Binocular disparity Receptive fields

INTRODUCTION

Stereoscopic depth discrimination is perhaps the finest visual judgment that human observers are capable of making. The stimulus attribute that leads to stereoscopic depth perception is horizontal relative disparity. Individuals with normal binocular vision can detect relative disparity lower than 10" (Ogle, 1962; Westheimer, 1979). Up to a limit, the perceived depth rises linearly with increasing relative disparity (Ogle, 1950; Richards, 1971; Richards & Kaye, 1974; Tyler, 1983). In this sense, stereoscopic depth perception is quantitative.

The presence of two ocular images of the same object (corresponding elements) is crucial for quantitative stereopsis. Stimuli designed to study stereoscopic depth perception invariably contain features of similar shapes and similar contrasts in the two eyes. The computational study of stereopsis also focuses its attention on establishing one-to-one correspondence between the image elements in a complex stereogram (for example, Marr & Poggio, 1979; Pollard *et al.*, 1985). In the real world, however, the absence of true binocular correspondence is not rare. When the two eyes are used to observe the world, part of the background may be projected to one eye but not the other due to the occlusion of objects in the foreground. Leonardo da Vinci pointed out that when a small object was in front of a background, the two eyes would see different parts of the background (cited by Wheatstone, 1838). Consequently, the perception of depth due to partial occlusion is referred to as Da Vinci Stereopsis (Shimojo & Nakayama, 1990). Shimojo and Nakayama (1990) showed that whether unpaired monocular features could escape from binocular rivalry depended on whether they complied with the geometry of occlusion. Nakayama and Shimojo (1990) suggested that unpaired monocular features indicated the occurrence of an occluder. They even demonstrated that unpaired monocular dots could elicit a sense of an illusory occluder hovering in front of a background. However, the depth percept was only qualitative. One could see an occluder in front of the background, but could not estimate the amount of depth. Attempts to change the perceived depth by manipulating the unpaired elements has not been successful (Nakayama, personal communication).

Recently we showed that quantitative depth perception could be elicited by a stereogram in which there appeared to be no luminance defined corresponding features (Liu *et al.*, 1994). When cross-fused, the square brackets shown in Fig. 1(b) will give rise to an impression of a white rectangle (an occluder) floating in front of a larger, black rectangle. Our depth-matching experiments showed that the perceived depth of the white rectangle was linearly related to the width of the vertical bars of the brackets. We call this phenomenon "phantom stereopsis" because there are no conventional matching elements in the central region of the stimulus, where the white rectangle is vividly perceived in depth.

Does this phenomenon represent an anomaly to our current understanding of stereopsis, which includes such doctrine that corresponding local features are necessary for quantitative stereopsis? We have demonstrated that

*School of Optometry, University of California, Berkeley, CA 94720, U.S.A.

†Present address: College of Optometry, University of Houston, Houston, TX77204, U.S.A.

‡To whom all correspondence should be addressed at The Arlene Gordon Research Institute, The Lighthouse Inc., 111 East 59th Street, New York, NY 10022, U.S.A.

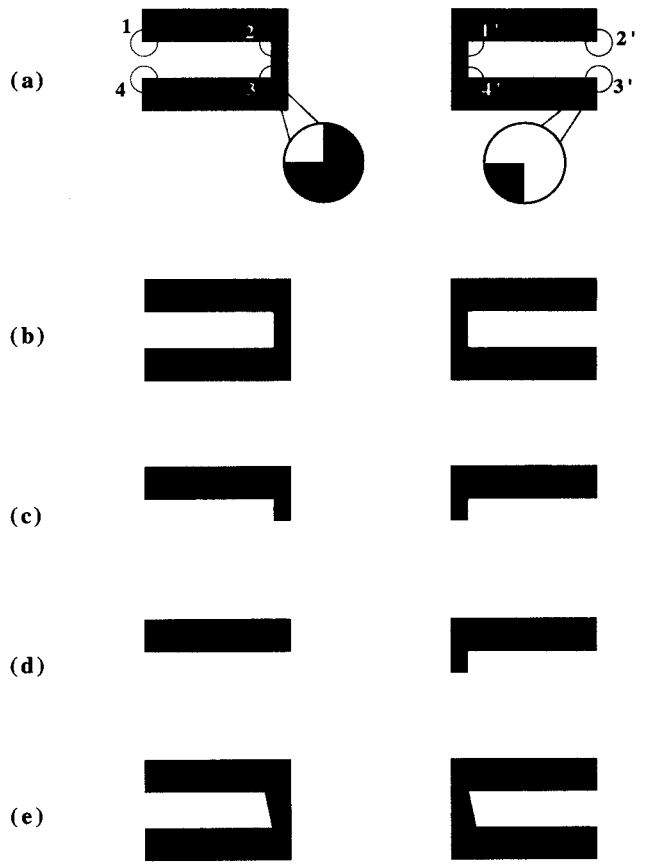

FIGURE 1. (a) Corner pairs in the phantom stereogram. These corner pairs, however, are by no means the conventional matching elements that are supposed to be similar in shape and contrast polarity. (b) Phantom stereogram. The vertical bars of the square brackets do not have binocular corresponding elements in the other eye. When cross-fused, this stereogram gives rise to an impression of a white rectangular occluder standing in front of a black rectangle on the background. (c) The lower half of the original stereogram is erased. One sees a white occluder whose upper half is defined by corner pairs 11′ and 22′ in (a) but whose lower half merges into the background. (d) Only corner pair 11′ in (a) is intact. Under binocular viewing, a well defined corner is seen in front of the background. (e) Corner pairs 22′ and 33′ in (a) are given more disparity than 11′ and 44′. A twisted white occluder is seen in front of the background. The upper right and lower left corners of this occluder are farther from the background than the other corners.

visually completed monocular features (the white rectangular regions defined by the square brackets) are not responsible for the quantitative depth perception. Gillam (1995) argued that although there were no corresponding vertical contours in the central portion of the phantom stereogram, there were corresponding horizontal edges in the display and these horizontal edges could define a rectangular region in depth. Gillam illustrated her argument by replacing the horizontal edges in Fig. 1(b) with horizontal lines. When cross-fused, this display did give an impression of two longer horizontal lines lying in the background and two shorter horizontal lines lying in the foreground. However, Gillam's line stimulus was not a correct representation of the original display because the corresponding terminations of the line drawings all have the same contrast polarity while the corresponding terminations of the horizontal edges in the phantom stereogram have opposite contrast (Liu *et al.*, 1995). For example, in Fig. 1(a), corner 3 in the left picture is a white corner in the second quadrant whereas corner 3′ in the right picture is a black corner in the third quadrant. The vertical edge of corner 3 is black on the left and white on the right. The opposite is true for the vertical edge of corner 3′. Features of opposite contrasts can not be fused and cannot elicit depth perception (Treisman, 1962). Therefore we think, at the level of the stimulus, that the stereogram shown in Fig. 1(b) does constitute an anomaly to our current understanding of stereoscopic depth perception.

However, we would have agreed with Gillam if her point were that binocular disparities could be extracted at the locations of the terminations of the horizontal edges (or the inner corners) on the phantom stereogram. To be specific, the perceived depth of the upper left corner of the white occluder is determined by the horizontal disparity between corner 1 in the left image and corner 1′ in the right image, as shown in Fig. 1(a). Similarly, the depth of the upper right corner is determined by corners 2 and 2′, and so on. Then the immediate question would be how to establish binocular correspondence between these corners. As we have mentioned above, the local features around the terminations of the horizontal edges are usually not considered as conventional matching features.

In this study we first demonstrated that despite their different shapes and contrast polarities, the corner pairs in Fig. 1(a) behave in a similar fashion to conventional local matching features in stimulating depth perception. We also performed computer simulations to illustrate how these unconventional local features might be binocularly combined by neural mechanisms which were known to exist in the human visual system.

DEMONSTRATIONS

Various stereograms are shown throughout this paper to demonstrate the phenomena under discussion. These stereograms should be observed with the eyes crossed, i.e., the image on the right should be seen by the left eye and the image on the left should be seen by the right eye.

The stereograms in Fig. 1(c–d) demonstrate that a pair of dissimilar corners in isolation could give rise to a depth percept, as would a pair of conventional matching features. In Fig. 1(c), the lower half of the patterns in Fig. 1(b) is erased but the corner pairs 11′ and 22′ are preserved. In this display, a black rectangle is perceived as being partially occluded by a white occluder in the foreground. The upper half of the occluder is sharp and clear but its lower half is not defined and appears to merge into a white background. In Fig. 1(d) three of the four inner corner pairs of the pictures in Fig. 1(b) are eliminated. Only corner pair 11′ is intact. A white corner is perceived as standing in front of the black rectangle in the display. The horizontal boundary of the white corner slopes to the background on the right and the vertical boundary of the corner extends a little downward and then merges into the background. These stereograms

demonstrate that individual corner pairs are capable of producing depth, independent of the global configuration.

We can also demonstrate that changing the disparity between the corresponding corners will change the perceived shape of the white occluder accordingly. In Fig. 1(e), corner pairs 22′ and 33′ have larger disparity than corner pairs 11′ and 44′. The upper right and the lower left corners of the perceived occluder should appear further from the background than the upper left and lower right corners. When cross-fused, the white rectangle in front does appear twisted. Therefore, corner pairs in the phantom stereogram can independently change the perceived shape of the occluder.

So far, the corners in the left and right pictures have been discussed as though they were conventional binocular features (e.g. lines or dots). However, they are not, as we have explained in the Introduction. How can these corners be binocularly combined? Is it possible that some neural mechanisms in the visual cortex could establish binocular correspondence between these drastically different corners and extract binocular disparity between them? In the following sections we utilize computer simulation to examine the possibility that binocular cells in the visual cortex can match these corners and thereby support stereoscopic depth perception.

COMPUTER SIMULATIONS

Methods

Physiological experiments have established that there are binocular cells in the visual cortex that can be driven by stimuli in the two eyes (Hubel & Wiesel, 1962, 1965). Among these binocular cells there are simple, complex and end-stopping cells that are tuned to different orientations and spatial frequencies. Could these cells extract disparity information from the stimulus shown in Fig. 1(b)? Our approach in searching for the answer is illustrated in Fig. 2. We first filter the left and right picture with an operator that simulates the behaviors of a given cell type. The filtered versions of the original stimuli can be considered as the activities on two arrays of monocular cells of the given type. A local cross-correlation analysis between left and right filtered images is then performed to produce a disparity map which, we assume, corresponds to our perception of the stimulus. A local pattern from the left filtered pictures is correlated with a neighboring region in the right filtered image. The regions are chosen to include the corners. The peak in this cross-correlation is taken to be the best match. In the example shown, a matching pattern is cropped from the left filtered image. It is correlated with a horizontal strip from the right filtered image. The location of the peak of the cross-correlation is found at (68, 81) which means that a left eye pattern centered at (41, 79) can find the best match at (68, 81) in the right eye image. Other best matching positions in the two images can be found in a similar way and they are indicated by the crosses. The positional difference between a pair of corresponding crosses represents the disparity at this location in the binocular field. For example, the coordinates of the two matching locations shown in Fig. 2 are (68, 81) and (41, 79). In the binocular field, the horizontal disparity at this location is therefore $68 - 41 = 27$ pixels.

This local analysis is conceptually in agreement with our knowledge about binocular cortical cells. Barlow *et al.* (1967) and Nikara *et al.* (1968) demonstrated that the receptive fields of binocular cortical cells might not fall exactly on the true corresponding points. This spatial offset made the cell selectively responsive to stimuli with a certain range of positional disparities. Our analysis simulates the behavior of such a binocular cell, where a cell with a given receptive field (the matching pattern) is activated by the best matching local features in the two half images.

When we later discuss the results of our computer simulation, we present the filtered stereograms to show the local patterns created by various filters. Cross-fusing these stereograms provides a chance to directly visualize the effects of various cell types. Small icons of the original stereogram and the filter used are shown at the upper left corner of each filtered stereogram. The results of local cross-correlation analysis are shown in disparity maps similar to the one at the bottom of Fig. 2. The numbers on a disparity map are the horizontal disparities at the locations indicated by the crosses.

Stimulus

Figure 3(a, b) illustrates the stereo stimuli used in our simulations. The dimensions in the figure are in pixels. The black areas have a luminance value of zero and the white areas a luminance value of one. Figure 3(a) is a conventional stereogram. The difference between the corresponding vertical bars is 30 pixels. This is the horizontal disparity of the central white rectangle. Figure 3(b) is a pair of square brackets whose vertical bars were also 30 pixels. According to the result of our depth matching experiment (Liu *et al.*, 1994), stimuli in Fig. 3(a and b) produced the same perceived depth. Any neural mechanism that might underlie the phantom stereopsis phenomena should produce the same disparity output to these two stereograms.

Equipment

All the simulations were conducted using MathWorks' Matlab software on a Power Macintosh 7100/66 computer and/or a Sun Sparc 10 workstation.

Results

Simple cells. The behavior of a simple cell can be approximated by filtering the stimulus pattern with a Gabor function. We first simulated vertically oriented simple cells. The receptive fields for these cells are described by the following formulae

$$S_v(x,y) = \exp(-(x^2+y^2)/\sigma^2) * \sin(2 * \pi * x/p) \quad (1)$$

where p is the spatial period in pixels and σ is the spatial constant of the Gaussian window in pixels.

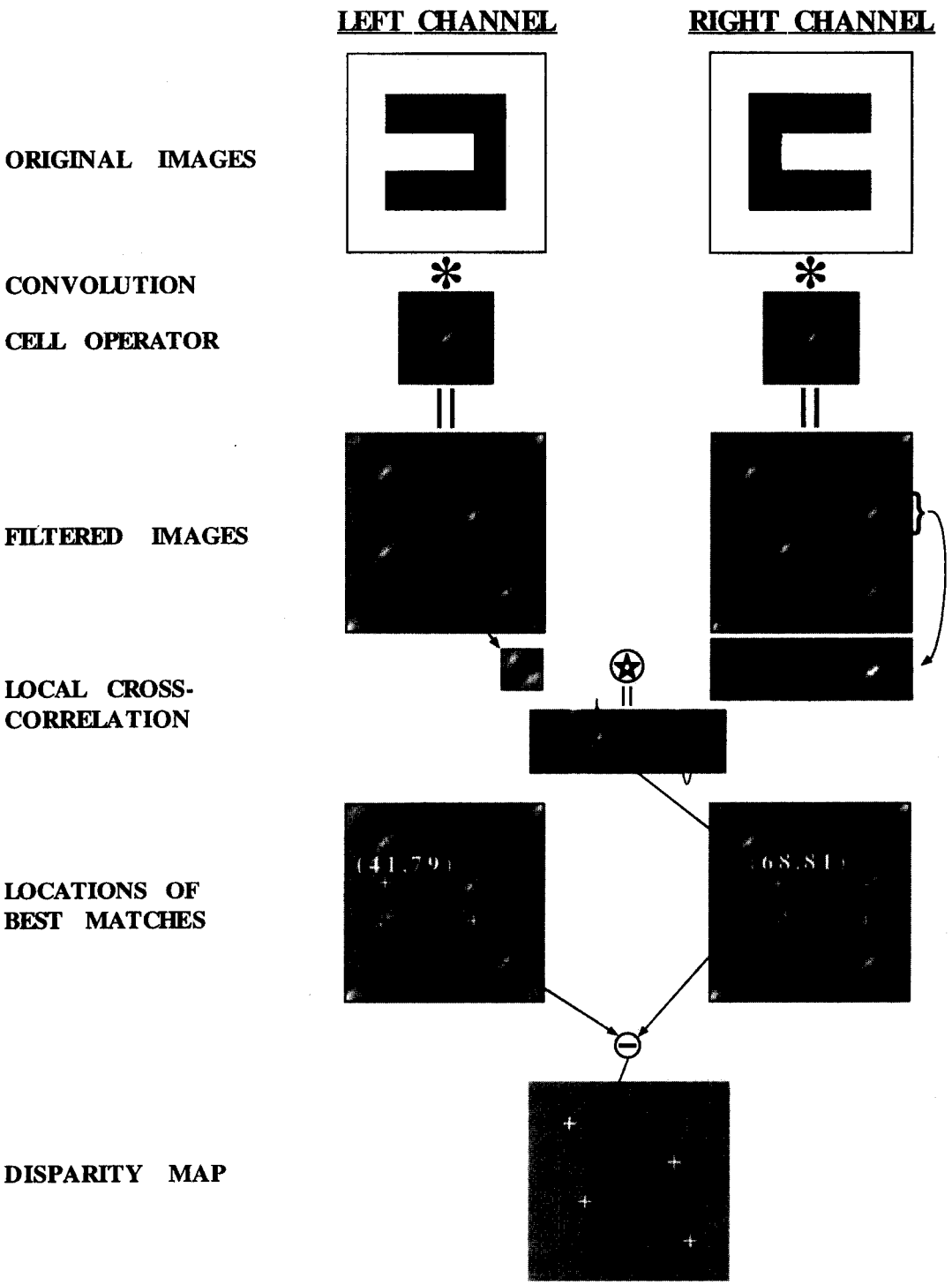

FIGURE 2. The scheme of computer simulation. The original stereograms are first convolved with an operator that simulates the known behavior of a cortical cell type. The filtered images then go through a cross-correlation analysis. A local region in the left image (a matching pattern) is cross-correlated with a horizontal stripe in the right image. The height of this stripe is slightly larger than the matching pattern. The peak of the cross-correlation is taken to be the best match. In the example shown, the maximum is found at (68, 81) which means that a left eye pattern centered at (41, 79) can find the best match at (68, 81) in the right eye image. The positional difference between these positions is the binocular disparity at that location. The disparity map produced by this analysis is composed of local symbols (crosses or dots) representing locations in the binocular field. The disparities (in pixels) at these locations are shown near the symbols.

Figure 3(c and d) shows the results of filtering the stereograms in Fig. 3(a and b) with a vertical Gabor function. The spatial period of the Gabors are 16 pixels and the parameter $\sigma = 11$ pixels. These Gabors contain about 1.6 cycles of the sinusoid at half-height and the bandwidth is 0.81 octaves. In this and the following

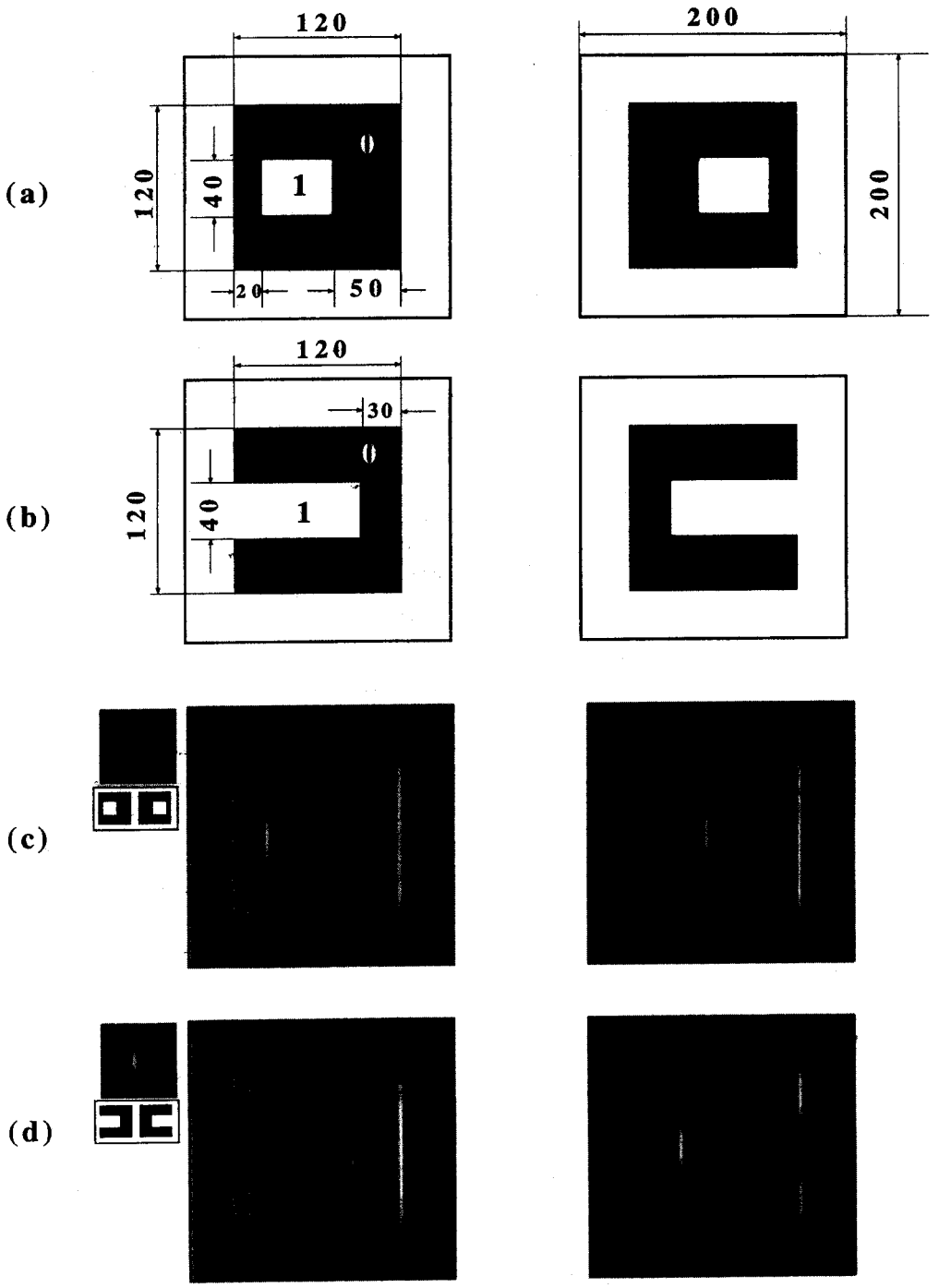

FIGURE 3. (a) Conventional stereogram used in computer simulation. (b) Phantom stereogram used in computer simulation. All dimensions are in pixels. (c) Conventional stereogram filtered with a vertical Gabor function. Stable depth perception can be reached when the stereopair is cross-fused. (d) Phantom stereogram filtered with a vertical Gabor function. The binocular percept is rivalrous.

simulations, the bandwidth of the operator is not critical, as long as strong inhibitory side-lobes are present. The filtered version of the conventional stereograms [Fig. 3(c)] gives rise to stable depth percepts under binocular viewing. The phantom stereogram filtered with a vertical Gabor [Fig. 3(d)], however, appears rivalrous. The depth positions of the short vertical bars in the middle are not well defined. They may obtain certain depth at first glance, then the depth fades away and they appear either a little in front of, or a little behind the background. The perceived lateral separation between these short bars also changes over time. If we examine this pair of images carefully, we find that the locations occupied by corresponding corners in the original picture are now occupied by distributions of opposite contrasts. This is due to the opposite contrast polarities of the edges around

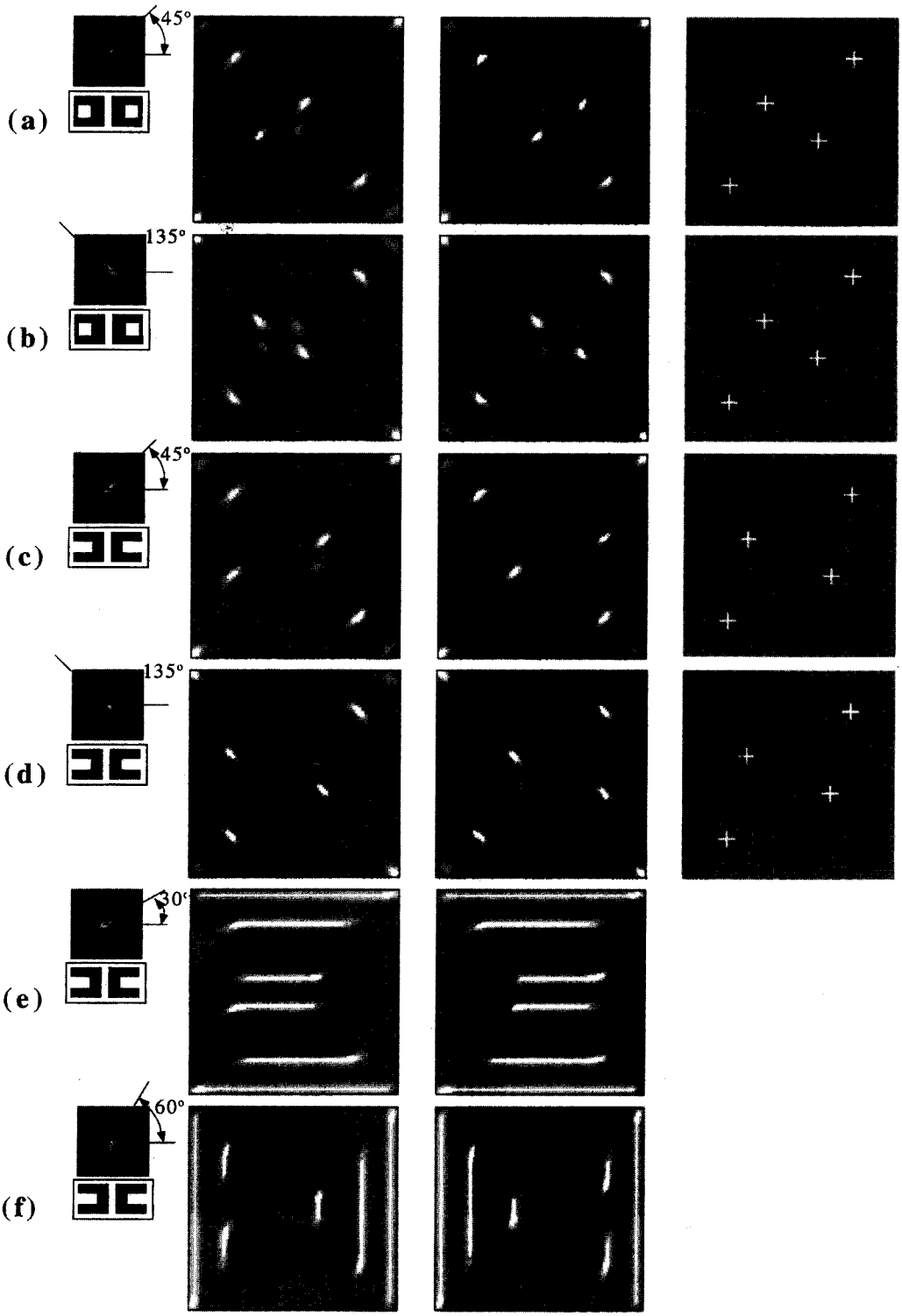

FIGURE 4. (a) Conventional stereogram filtered with a 45 deg Gabor. (b) Conventional stereogram filtered with a 135 deg Gabor. (c) Phantom stereogram filtered with a 45 deg Gabor. (d) Phantom stereogram filtered with a 135 deg Gabor. (e) Phantom stereogram filtered with a 30 deg Gabor. (f) Phantom stereogram filtered with a 60 deg Gabor. In these stereograms, local patterns which resemble the Gabor filter used occur at the locations of the corners in the original. Stable stereoscopic depth perception can be seen from these stereograms. The rectangular regions in the centers of (c) and (d) appear slightly twisted. The results of a correlation analysis on (a)–(d) are shown to the right of the stereograms.

the corners, as we have mentioned in the Introduction. Figure 3(d) demonstrates that vertically tuned binocular cells may have an adverse effect in processing binocular disparity from the phantom stereogram, although they are usually considered as the main carriers of binocular disparity information.

Next we analyzed a system composed of Gabor functions oriented at 45 and 135 deg. These Gabor functions are defined by

$$S_{45}(x,y) = \exp(-(x^2+y^2)/\sigma^2) * \cos(2 * \pi * (x-y)/2p) \quad (2)$$

$$S_{135}(x,y) = \exp(-(x^2+y^2)/\sigma^2) * \cos(2 * \pi * (x+y)/2p) \quad (3)$$

Figure 4(a–d) shows stereograms filtered by a 45 deg Gabor or a 135 deg Gabor function. All the filtered images have obvious local features at locations approximately coincident with the corners in the original pictures. These local features have similar distributions as the filters used, i.e., the distribution of a Gabor function. Depending on the shape of the corners, these local features can have either a bright central lobe or a dark central lobe.

Figure 4(a and b) show a conventional stereogram filtered by a 45 and a 135 deg Gabor function, respectively. The spatial period (p) of the Gabor was 16 pixels and σ was 11 pixels. The local features around the corresponding corners in the two eyes all have similar distributions. This is not surprising because the corresponding corners in the original pictures are identical in the conventional stereogram. When cross-fused, these pairs of pictures produce a flat square standing in front of the background with a Gabor-like pattern at each of the four corners. Figure 4(c and d) show the results of filtering the phantom stereogram with a 45 and a 135 deg Gabor, respectively. In these stereograms the four pairs of features at the corresponding inner corners are very similar, even though the unfiltered corners are very different. One can easily fuse these stereo pairs and see a rectangular shaped object cornered by four Gabor-like features standing in front of the background. Therefore, the dissimilar corners in the phantom stereogram can elicit similar responses from a cell whose receptive field resembles an oblique Gabor.

We conducted a cross-correlation analysis on the images in Fig. 4(a–d). The matching patterns were cropped from square areas $1.5*\sigma$ pixels in dimension, centered at the positions of the corners in the unfiltered images. The results are shown in the disparity maps on the far right of Fig. 4. Figure 4(a and b) illustrates that the four corners of the inner rectangles of the conventional stereogram all have a disparity of 30 pixels relative to the outer squares. This result does not change when a Gabor of a different spatial period is used. Thus, both 45 and 135 deg Gabor filters preserve the correct amount of disparity in a conventional stereogram. The four corners of the rectangular region in the phantom stereogram, on the other hand, have different disparity values and which corner has more disparity depends on what Gabor filter is used. For example, when the filter is a 45 deg Gabor [Fig. 4(c)], the upper-left and lower-right corners have 29 and 27 pixels of disparity and the same corners have 33 and 32 pixels of disparity when a 135 deg Gabor filter is used [Fig. 4(d)]. When Fig. 4(c) is cross-fused, the upper-left and lower-right corners appear closer to the background than the upper-right and lower-left corners and the rectangle appears twisted in depth. When we view the filtered version produced by a 135 deg Gabor [Fig. 4(d)], the rectangle also appears twisted, but in the opposite direction, that is, the upper-left and lower-right corners appear further from the background than the other two corners. The difference in the disparity among the four corners also increases with the spatial period of the Gabor filters. For a Gabor whose spatial period is 8 pixels, the largest disparity discrepancy is 2 pixels. For a Gabor of 16 pixels period, the largest discrepancy is 6 pixels.

The 45 and 135 deg Gabors happen to be balanced to both horizontal and vertical edges. They therefore create localized patterns around the corners. Gabors whose orientations differ from these will either emphasize horizontal or vertical edges and thereby create bar-shaped features. Figure 4(e and f) shows the phantom stereogram filtered with a 30 and a 60 deg Gabor, respectively. These filters are similar to those used in Fig. 4(a–d) except for their orientations. Stable stereoscopic depth perception can be reached by cross-fusing these stereograms. Compared to the stereogram filtered with a 45 deg Gabor [Fig. 4(c)], the distortion of the central rectangle is smaller with the 30 deg Gabor [Fig. 4(e)], but is more pronounced with the 60 deg Gabor [Fig. 4(f)]. When these Gabors are applied to a conventional stereogram (not shown), the perceived depth structure does not differ from that seen in Fig. 4(a or b). Therefore, our simulation shows that all orientation channels produce the same disparity outputs when similar features appear at the corresponding points (a conventional stereogram). When the stimulus is a phantom stereogram, however, binocular combination of dissimilar corners can be reached within each orientation channel, although the disparity outputs of different channels are different.

End-stopped cells. Another type of cell that we simulated was the so-called "end-stopped" cell. An end-stopped cell's response to a bar or a grating of the preferred orientation increases with the increasing length of the stimulus up to a point and then starts to decrease with further increase of the length. Hubel and Wiesel (1965) first found cells with these characteristics in areas 18 and 19 of the cat cortex and referred to them as hypercomplex cells. Later physiological studies have shown that many cells in area 17, both simple and complex cells, exhibited end-inhibition property (Rose, 1977; Orban *et al.*, 1979a, b). A recent systematic inquiry of this cell type (DeAngelis *et al.*, 1992) showed that the end-inhibition mechanism had broader, but otherwise similar orientation and spatial frequency tunings as the excitatory receptive field. End-inhibition was also shown to be insensitive to the spatial phase of the stimulus in the inhibition zone. Hubel and Wiesel (1965) suspected that these cells were specialized in detecting truly 2-dimensional features such as line interruptions, curvatures and vertices of a square. Dobbins *et al.* (1987, 1989) recently demonstrated through physiological experiments and computer simulations that end-stopping cells were

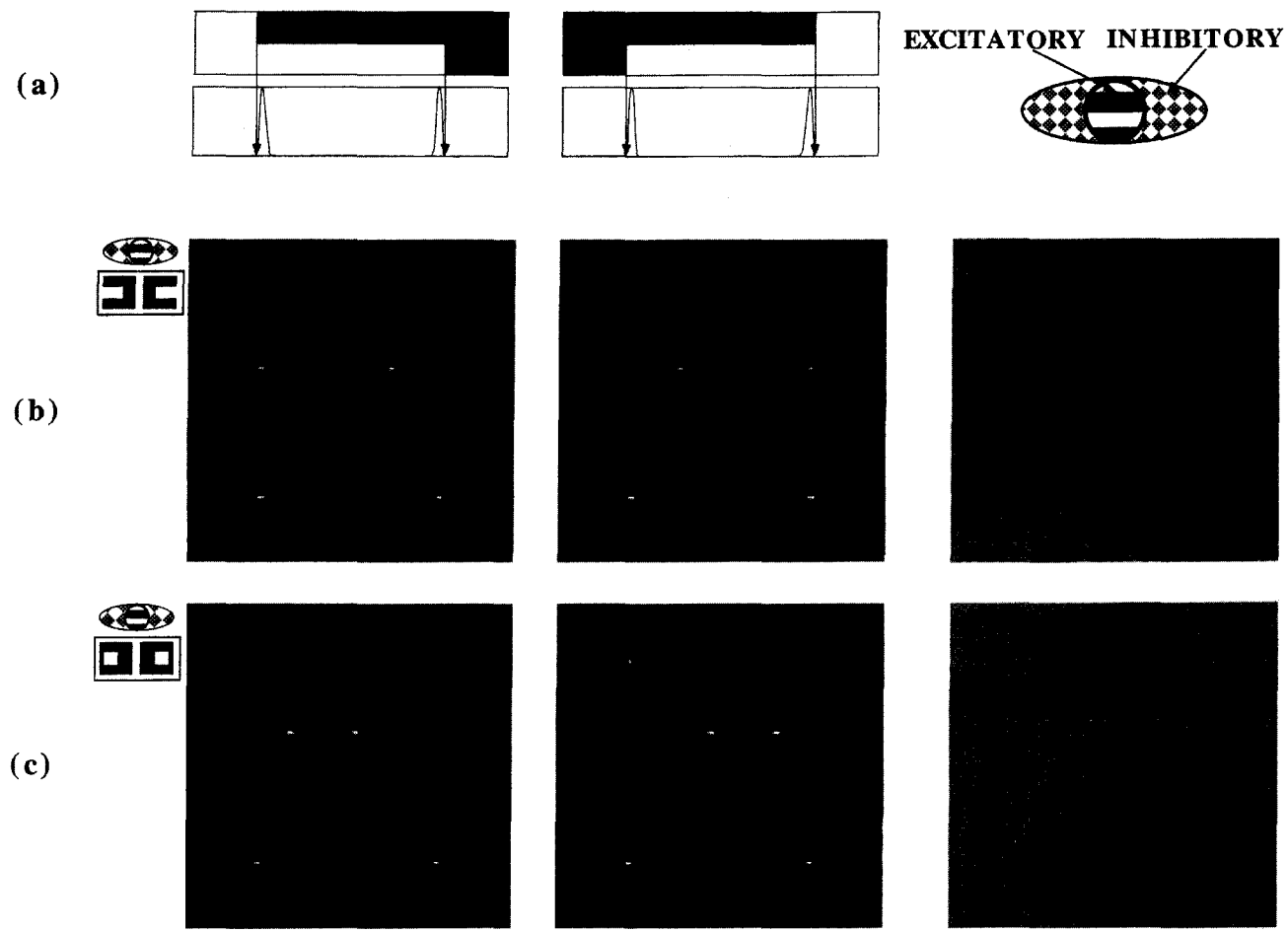

FIGURE 5. (a) End-stopped cells. The activity of the excitatory center is suppressed by that of an elongated inhibitory unit. The twin-peak curves are the responses of an end-stopped operator to the upper part of the phantom stereogram. The locations of the corners in the stimulus (indicated by arrows) are designated by zero-crossings. Notice that the end-stopped operator gives identical responses to dissimilar corners in the stimuli. (b) and (c) show the results of filtering a phantom stereogram and a conventional stereogram with 30 pixels disparity. The results of the correlation analysis are shown to the right of the stereograms.

curvature-selective. The computer simulation conducted by Heitger *et al.* (1992) showed that the local maxima in the activities created by end-stopped operators explicitly signal the locations of some true 2-D patterns such as line ends, corners and segments of strong curvature. In the current context, the matched features in phantom stereograms seem to be corners, so we constructed a model of an end-stopped cell and applied it to the stereograms shown in Fig. 3(a–b) to determine if end-stopped units might show disparity tuning for phantom stereograms.

The end-stopped cell we constructed is composed of an excitatory subunit whose activity is suppressed by an inhibitory subunit [Fig. 5(a)]. The excitatory subunit is a simple odd-symmetric Gabor function with horizontal optimal orientation. This excitatory subunit is located at the center of an elongated horizontal inhibitory subunit which is a compound of an even- and an odd-Gabor function. The outputs of the even- and odd-Gabors are combined quadratically (each one was squared and then added together) to represent the local energy. Therefore, the inhibitory subunit has the same optimal orientation, the same optimal spatial frequency but is not sensitive to spatial phase of the stimulus. The horizontal σ of the inhibitory subunit is twice as large as that of the excitatory subunit. We assume that the inhibition between the excitatory and the inhibitory subunits is subtractive in nature, which means that output of an end-stopped operator is the difference between the excitatory and the inhibitory subunits.

$$O_{es} = O_{ex} - \alpha * O_{in} \tag{4}$$

where $\alpha = 0.5$ is the inhibition coefficient. We also assume that the inhibition becomes effective only when the excitatory subunit has positive responses.

Figure 5(a) shows the upper part of the square bracket pattern in Fig. 1(b). It can be considered as a horizontal edge terminated differently on the two ends but it can also be considered as two different corners connected by a horizontal edge. The twin-peak curves are the responses of end-stopped operators to these stimuli. The end of an edge (or a corner) is signified by a zero-crossing in the output of an end-stopped operator. This simulation shows that the end-stopped cell we constructed gives identical responses to two dissimilar corners.

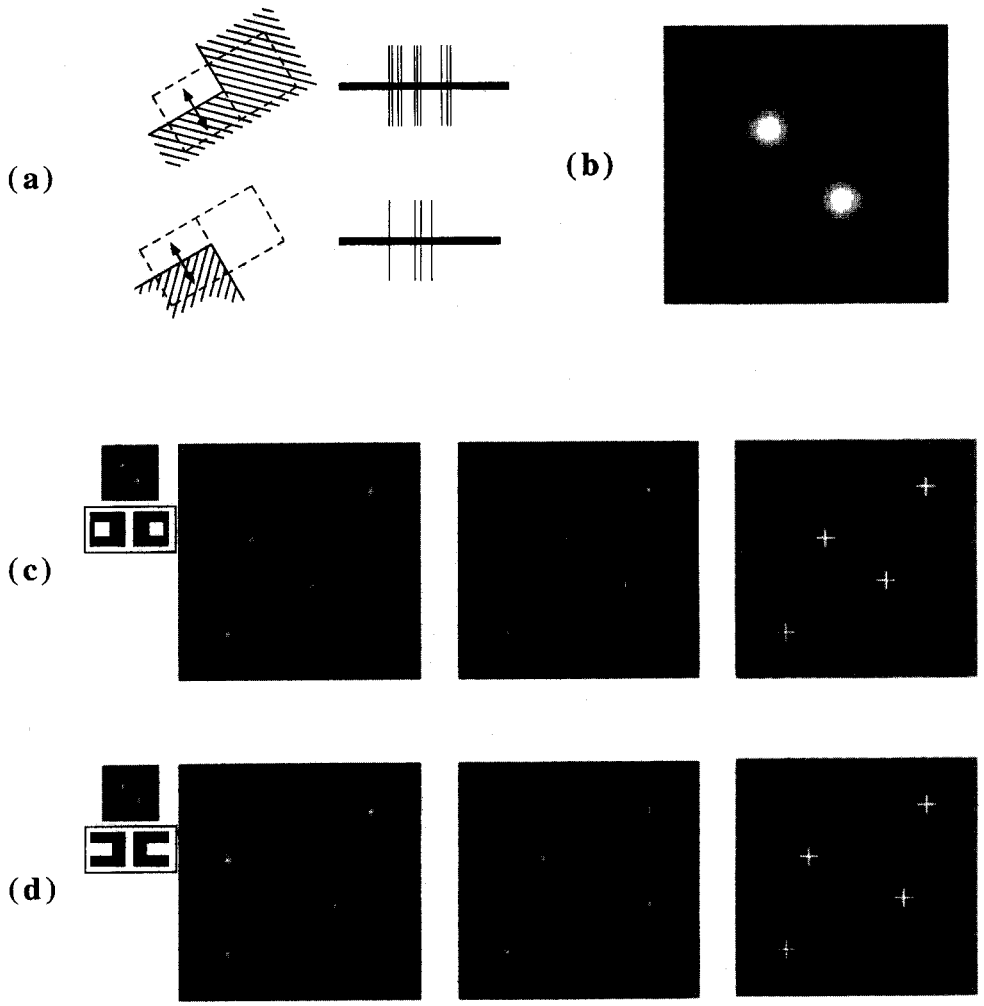

FIGURE 6. Quadrant cell. (a) Stimulus patterns used by Hubel and Wiesel (1965) and the responses of a hypercomplex cell. These corner patterns are similar to corner pair 33′ in Fig. 1(a). (b) The receptive field of a hypothetical quadrant cell which has four antagonistic quadrants. (c) and (d) show the results of filtering a conventional stereogram and a phantom stereogram with the quadrant cell. The results of a correlation analysis are shown to the right of the stereograms.

We filtered the stereograms in Fig. 3(a–b) with the end-stopped operators. Figure 5(b) is the filtered version of the phantom stereogram, while Fig. 5(c) is the filtered version of the conventional stereogram. Because the inhibition between excitatory and inhibitory subunits is high ($\alpha = 0.5$), the features in the filtered images are blobs, designating the corners in the unfiltered images. Figure 5(b) demonstrates that an end-stopped operator creates similar local response patterns on the phantom stereogram. The locations of these patterns coincide with the positions of the corners in the original stereogram. The filtered images in Fig. 5(b) can be binocularly fused and four sets of blobs are seen in front of the background. Similar percepts can be reached with the filtered images in Fig. 5(c).

The results of a correlation analysis are shown in the disparity maps to the right of the stereograms. All the blobs in the central portion of the filtered phantom stereogram have 30 pixels of disparity [Fig. 5(b)]. An identical amount of disparity was found between the central features in the filtered conventional stereogram ([Fig. 5(c)]. Therefore, the end-stopped operators show disparity tuning for the corner features in the phantom stereograms and they might be responsible for mediating the perception of quantitative depth in these stimuli.

Quadrant cells. In their exploration of the visual areas 18 and 19 of the cat cortex, Hubel and Wiesel (1965) described in detail the behavior of a hypercomplex cell in area 18 (pp. 244–247). Two of the stimuli to which this cell was responsive are shown in Fig. 6(a) (redrawn from Hubel and Wiesel's Fig. 8). One may notice the similarity between these corners and corners 3 and 3′ in our Fig. 1(a). Accordingly, we proposed an idealized quadrant cell whose receptive field is composed of four antagonistic quadrants, as is shown in Fig. 6(b). This receptive field is described by the following function

$$C(x,y) = \exp(-(x^2+y^2)/\sigma^2) * \sin(2 * \pi * x/p) * \sin(2 * \pi * y/p) \quad (5)$$

which is the product of a horizontal sinewave grating and a vertical sinewave grating, windowed by a gaussian

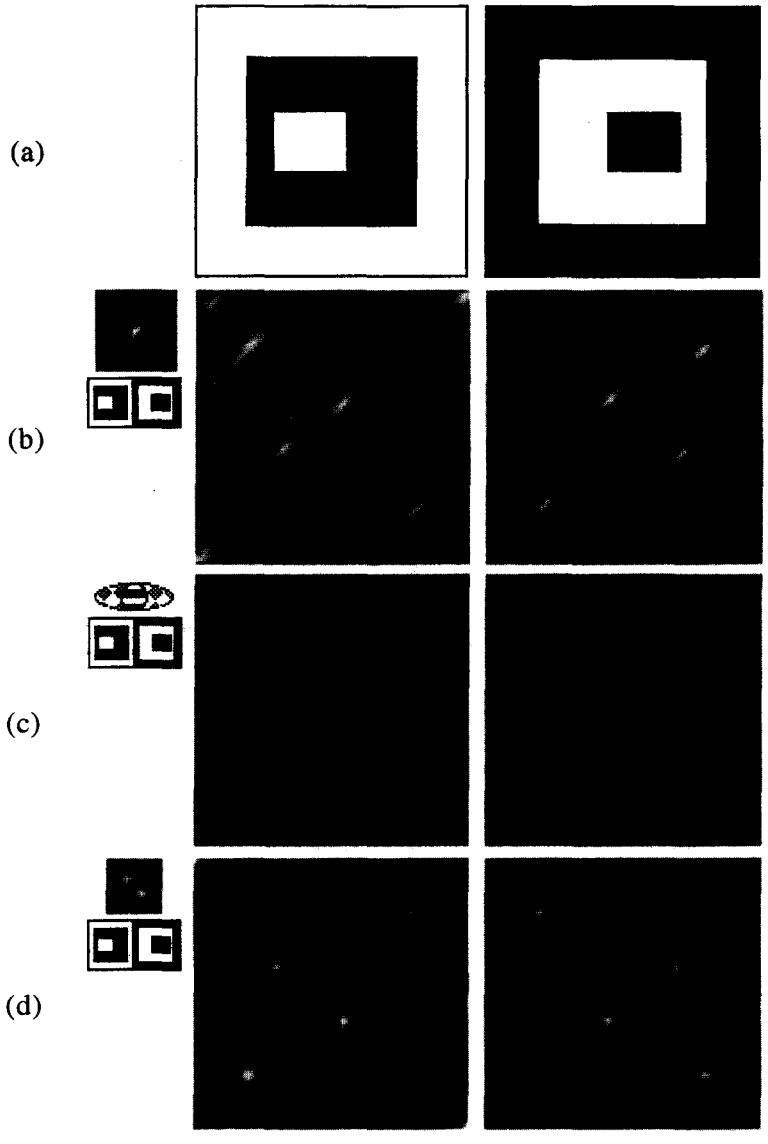

FIGURE 7. (a) Stereogram of opposite contrast. (b)–(d) are the results of filtering the stereogram in (a) with a 45 deg Gabor, an *end-stopped cell* and a quadrant cell, respectively. These stereograms appear rivalrous under binocular viewing.

function. It is easy to demonstrate that the tests conducted by Hubel and Wiesel on their edge-stopped cell (Hubel and Wiesel's Figs 5–11) will evoke qualitatively similar responses from this quadrant cell. The quadrant cell is selectively responsive to only some types of corners. It responds to either a white corner in the second quadrant or a dark corner in the third quadrant. A dark corner in the second quadrant or a white corner in the third quadrant will elicit strong inhibitory responses. Notice that a single corner, whether it is a dark corner or a white corner, is not the optimal stimulus for the quadrant cell. Nevertheless, two different corners can elicit moderate but similar responses from this cell. We also assume that this cell is binocular and its receptive fields in the two eyes have similar shapes. A binocular quadrant cell is especially suitable for the detection of phantom disparity because a pair of dissimilar corners in the phantom stereogram is an adequate stimulus for this cell. This cell can also be used to detect disparity between the similar features in the conventional stereogram.

We filtered the stereograms in Fig. 3(a, b) with this quadrant cell operator and the results are shown in Fig. 6(c, d). We see that local maxima and minima occur at the positions where corners of the unfiltered pictures are located. In the filtered version of the phantom stereogram [Fig. 6(d)], a local maximum in one picture corresponds to a local maximum in the other picture and a local minimum corresponds to a minimum, even though the corners in the unfiltered image are different in shape. Similar corresponding patterns also occur in the filtered version of conventional stereograms [Fig. 6(c)]. When cross-fused, both stereograms give a stable percept of a rectangular region in the center standing in front of the background. A correlation analysis is conducted on the filtered stereograms in Fig. 6(c, d). The matching patterns are again cropped from the locations of the corners in the

unfiltered images. The disparity maps on the far right show the result of this analysis. Correct amounts of disparity (30 pixels) are extracted in both conventional and phantom stereogram. This result does not change with the spatial extent of the quadrant cell.

Stereograms with opposite contrasts. Although the three types of cells we discussed so far can respond similarly to very different stimulus patterns, they still possess the feature selectivity necessary for stereopsis. We have shown that these cells can extract binocular disparity embedded in a conventional stereogram. We can also demonstrate that they cannot establish stable binocular correspondence between stereopairs of opposite contrasts.

Figure 7(a) is a stereogram with opposite contrast. Binocular rivalry occurs when one tries to fuse these pictures. Figure 7(b–d) are the results of filtering Fig. 7(a) with a 45 deg Gabor, an end-stopped cell and a quadrant cell, respectively. These pictures also appear rivalrous under binocular viewing. Although similar response patterns occur along a same horizontal line, they are not near each other's corresponding point. Therefore, their match subtends a very large disparity. For example, in Fig. 7(d), a white blob in the left eye always corresponds to a black blob in the right eye, or vice versa. Any attempts to eliminate disparity (e.g., with vergence eye movements) between a pair of similar blobs would invariably increase the disparity between other pairs of similar blobs. Therefore, a stable percept cannot be reached. In Fig. 7(b), the dots have only one polarity, but the double-dot in one eye would compete for a single dot in the other eye. Again, no stable percept can be reached.

DISCUSSION

In this paper, we demonstrated that operators with known characteristics of certain cortical cells can encode disparity from the dissimilar corners in the corresponding areas of the phantom stereogram. To these operators, the phantom stereogram is nothing but conventional. It is possible, therefore, that the quantitative depth perception observed with the phantom stereogram is determined by the outputs of these cells. Higher level depth processing mechanisms based on the heuristics of occlusion geometry need not be invoked. There are, however, some problems with this low level disparity detector explanation.

First, the cell types we discussed so far can establish binocular correspondence at isolated points in the phantom stereogram. The depth percepts based on the outputs of these cells would be isolated points in front of the background. The percept elicited by the stereogram, on the other hand, is a solid plane in depth. This percept would be achieved naturally from the point of view of occlusion heuristics because an opaque plane in the foreground is a necessary condition for partial occlusion to occur. If one tries to explain the percept with the outputs of the low level disparity detectors, he has to invent another neural mechanism which can generate a surface through the points in depth. Since there are points with different disparities in the binocular field (for example, there are four points with 0 pixel disparity and four points with 30 pixels disparity in most of the disparity map shown), one has to explain why this mechanism would make two planes, one through the four zero disparity points and one through the four disparate points but not an inversed 4-sided funnel. Would occlusion heuristics help to select a depth configuration which is most proper for the stimulus configuration?

Second, we noticed from the previous discussion that different cell types give different disparity responses to the phantom stereogram. The fact that we only perceive one stable, flat occluder at a definite depth from the phantom stereogram suggests that there must be a process that can integrate the outputs of different disparity detectors and derive a single solution. Before the demonstration of the phantom stereogram there was no need to consider such a process because, as we have shown in the previous sections, all operators produce virtually identical and correct disparity outputs in response to a conventional stereogram where similar matching features occupy the corresponding regions in the stimuli. What, then, is the nature of this process? Is it a simple average of the outputs of all the disparity detectors? Or is it a selection process guided again by the heuristics of occlusion geometry?

REFERENCES

Barlow, H., Blakemore, C. & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology, 193*, 327–342.

DeAngelis, G. C., Robson, J. G., Ohzawa, I. & Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology, 68*, 144–163.

Dobbins, A., Zucker, S. W. & Cynader, M. S. (1987). Endstopping in the visual cortex as a substrate for calculating curvature. *Nature, 329*, 438–441.

Dobbins, A., Zucker, S. W. & Cynader, M. S. (1989). Endstopping and curvature. *Vision Research, 29*, 1371–1387.

Gillam, B. (1995). Matching needed for stereopsis. *Nature, 373*, 202–203.

Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E. & Kubler, O. (1992). Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Research, 32*, 963–981.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106–154.

Hubel, D. H. & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology, 41*, 229–289.

Liu, L., Stevenson, S. B. & Schor, C. M. (1994). Quantitative stereoscopic depth without binocular correspondence. *Nature, 367*, 66–69.

Liu, L., Stevenson, S. B. & Schor, C. M. (1995). Matching needed for stereopsis—reply. *Nature, 373*, 203.

Marr, D. & Poggio, T. A. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society, London, B, 204*, 301–328.

Mayhew, J. E. W. & Frisby, J. P. (1981). Psychophysical and computational studies towards a theory of human stereopsis. *Artificial Intelligence, 17*, 349–385.

Nakayama, K. & Shimojo, S. (1990). da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research, 30*, 1811–1825.

Nikara, T., Bishop, P. O. & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Experimental Brain Research, 6*, 353–372.

Ogle, K. N. (1950). *Researches in binocular vision*. Philadelphia: Saunders.

Ogle, K. N. (1962). Spatial localization through binocular vision. In Davson, H., (Ed.), *The eye*. New York, NY: Academic Press.

Orban, G. A., Kato, H. & Bishop, P. O. (1979a) Dimensions and properties of end-zone inhibitory areas in receptive fields of hypercomplex cells in cat striate cortex. *Journal of Neurophysiology, 42*, 833–849.

Orban, G. A., Kato, H. & Bishop, P. O. (1979b) End-zone region in receptive fields of hypercomplex and other striate neurons in the cat. *Journal of Neurophysiology, 42*, 818–832.

Pollard, S. B., Mayhew, J. E. W. & Frisby, J. P. (1985). PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception, 14*, 449–470.

Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America, 61*, 410–414.

Richards, W. & Kaye, M. G. (1974). Local versus global stereopsis: two mechanisms? *Vision Research, 14*, 1345–1347.

Rose, D. (1977). Responses of single units in cat visual cortex to moving bars as a function of bar length. *Journal of Physiology, London, 271*, 1–23.

Shimojo, S. & Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry. *Vision Research, 30*, 69–80.

Treisman, A. (1962). Binocular rivalry and stereoscopic depth perception. *Quarterly Journal of Experimental Psychology, 14*, 23–37.

Tyler, C. W. (1983). Sensory processing of binocular disparity. In Schor, C. & Ciuffreda, K. (Eds), *Vergence eye movements*. Boston: Butterworths.

Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research, 36*, 585–597.

Wheatstone, C. (1838). Some remarkable phenomena of binocular vision–Part the first. *Philosophical Transactions of the Royal Society, 128*, 371–394.

Acknowledgement—This work was supported by NIH grant EY-08882.