Vision Research 46 (2006) 171–178

Illusory volumes in human stereo perception

Hiroshi Ishikawa [a,*], Davi Geiger [b]

[a] *Department of Information and Biological Sciences, Nagoya City University, Nagoya 467-8501, Japan*

[b] *Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, USA*

Abstract

Any complete theory of human stereopsis must model not only how the correspondences between locations in the two views are determined and the depths are recovered from their disparity, but also how the ambiguity arising from such factors as noise, periodicity, and large regions of constant intensity are resolved and missing data are interpolated. In investigating this process of recovering surface structure from sparse disparity information, using stereo pairs with sparse identifiable features, we made an observation that contradicts all extant models. It suggests the inadequacy of retinotopic representation in modeling surface perception in this stage. We also suggest a possible alternative theory, which is a minimization of the modulus of Gaussian curvature.

Keywords: Stereopsis; Depth interpolation; Computational model; Statistical modeling

1. Introduction

The human brain can perceive depth by fusing two slightly different views from the left and right eyes. In the computational model of the process of stereopsis, the correspondences between locations in the two images formed on the retinae are determined, and the depths are recovered from their disparity. However, this can be done only approximately. Ambiguity arising from such factors as noise, periodicity, and large regions of constant intensity makes it impossible in general to identify all locations in the two images with certainty. Thus, any convincing model of stereopsis must detail how ambiguities are resolved and missing data are interpolated. Models (Marr & Poggio, 1976, 1979; Grimson, 1981; Poggio & Poggio, 1984; Pollard, Mayhew, & Frisby, 1985; Gillam & Borsting, 1988; Ayache, 1991; Belhumeur & Mumford, 1992; Jones & Malik, 1992; Faugeras, 1993; Geiger, Ladendorf, & Yuille, 1995; Belhumeur, 1996) have generally used as criterion some form of smoothness in terms of dense information such as the depth and its derivatives. To investigate this disambiguation process in human vision, we examined disparity interpolation, the process of recovering surface structure from sparse disparity information in a pair of visual images, using stereo pairs with sparse identifiable features. We obtained results, reported in the next section, that no current model is capable of explaining. We explain why this is so in section 3. In the last section, we discuss the implications and also suggest a possible alternative theory, which is a minimization of the modulus of Gaussian curvature.

* Corresponding author. Tel.: +81 52 872 5191; fax: +81 52 872 3495.
E-mail address: hi@nsc.nagoya-cu.ac.jp (H. Ishikawa).
doi:10.1016/j.visres.2005.06.028

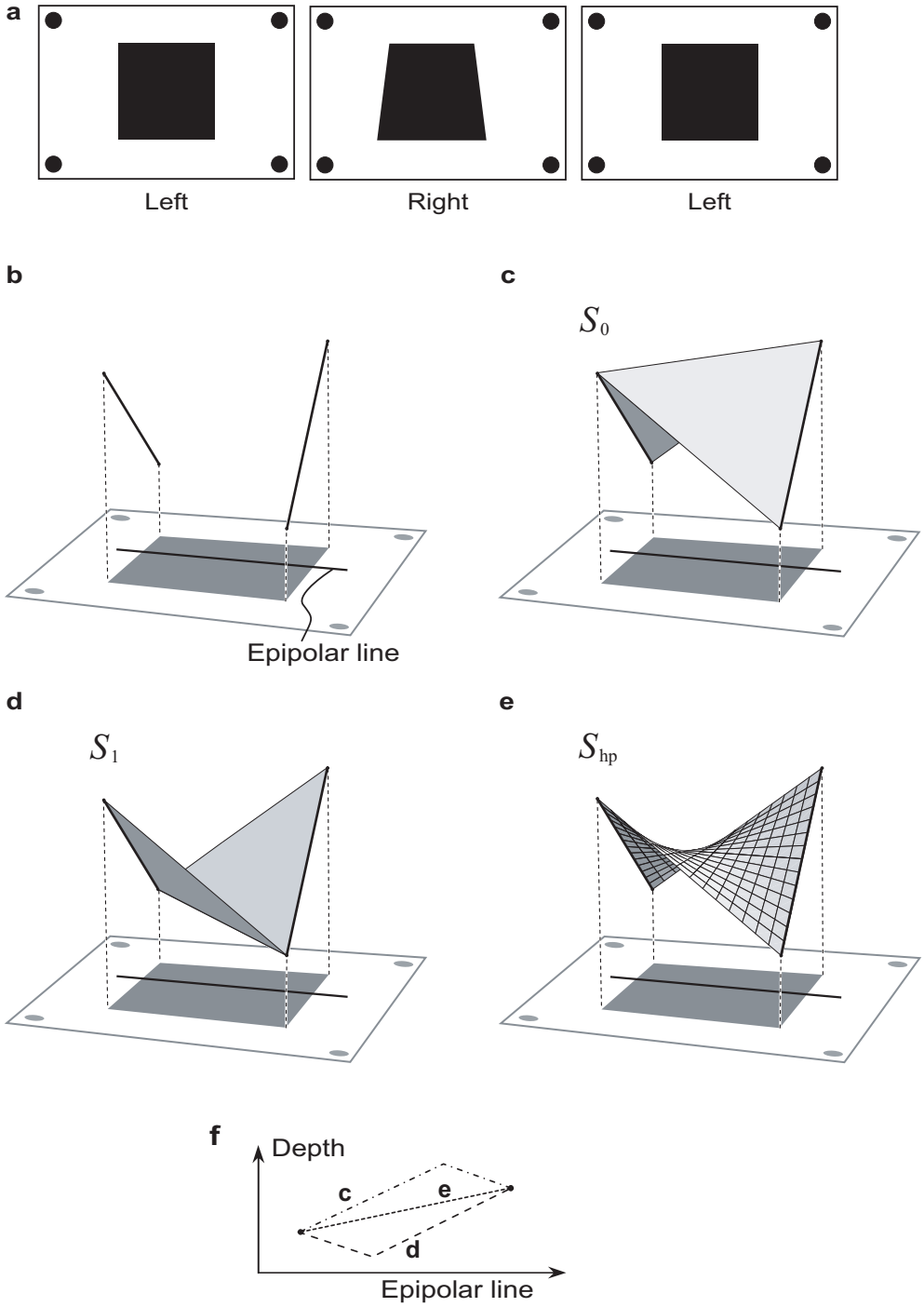

Fig. 1. The stereogram and possible surfaces. We used a stereogram of textureless surface patches with explicit luminance contours in order to investigate how the disambiguation process in human stereopsis recovers surface structure from sparse disparity information in a pair of visual images.
(a) A stereo pair. When the right images are cross-fused (or the left two images are fused divergently,) a three-dimensional surface is perceived.
(b) The thick lines represent the disparity values unambiguously obtainable from local feature. Because the matching is so ambiguous elsewhere, the brain is forced to guess the rest, revealing its expectation. Possible surfaces that agree with this boundary condition are **(**c) , (d), and (e).
(c), (d) The human brain tend to perceive either of these two. However, no current theory can explain this observation.
(e) Algorithms that seek to minimize gradient give a "soap bubble" surface like the one shown here. Models in which the prior probability distribution on epipolar lines are independent also give this solution.
(f) The cross sections for solutions above on an epipolar line are shown.

2. Methods and Results

Seven observers naïve to the purpose of the experiment with normal or corrected to normal vision viewed the stereoscopic images in Fig. 1a and Fig. 5a. The images were presented on a CRT monitor (NEC 98Mate Display 17) at a viewing distance of 1.5m through liquid crystal shutter goggles (model PLAY3DPC by I-O Data Device, Inc. of Kanazawa, Japan.) The shutter goggles switch between opaque and transparent at 100Hz, synchronized to the monitor so that alternate frames can be presented to the left and right eyes, allowing stereoscopic displays. Images contained the black shape shown in the figures, the height of which was 10cm on the monitor surface. Four of the observers first viewed the image in Fig. 1a, and then Fig. 5a; the rest viewed the images in the reverse order. In each viewing, the observer was asked to describe what was perceived after 15 seconds; and then was asked to choose from the three pictures in Fig. 1c-e (when Fig. 1a is shown) or Fig. 5c-e. There was no discrepancy between what they described and what they chose. A few stereo pairs of color pictures were shown to each viewer prior to the experiment in order to ascertain that the observer is capable of binocular stereo perception.

Only one of the observers reported the percept of a saddle-type shape (Fig. 1e). Other six viewers reported the percept of either convex (Fig. 1c) or concave (Fig. 1d) shape. One reported the percept of both of the convex and concave shapes.

Viewer	#1	#2	#3	#4	#5	#6	#7
Fig.1a	convex	saddle	concave	concave	convex	convex	both
Fig. 5a	concave	saddle	concave	concave	convex	convex	convex
Which first?	Fig.1a	Fig.1a	Fig.1a	Fig.1a	Fig.5a	Fig.5a	Fig.5a

3. Analysis

Fig. 1a shows one of the stereograms of textureless surface patches with explicit luminance contours. In these displays, there are very few features that can be depended upon when matching the points. The only distinguishing feature is the intensity edges on the circumference of the shape, where the discontinuous change in luminance occurs. There are no other cues that are ordinarily present, such as surface shade and partial occlusions (Gillam & Borsting, 1988; Nakayama & Shimojo, 1990; Anderson, 1994; Malik, 1996). In other words, the images are stereoscopic silhouettes or boundary contours (Fatih & Ramakant, 1993; Norman & Raines, 2002). Matching the edges gives the depth information illustrated in Fig. 1b. Everywhere else, each location in one image can perfectly match to a variety of locations in the other. This corresponds to the fact that any perfectly black surface spanning the two segments in Fig. 1b looks exactly the same. Nevertheless, the perception human observers report is much less ambiguous. As reported in the previous section, most observers who viewed the stereogram reported the percept of one of the two surfaces shown in Fig. 1c and d, which we call S_0 and S_1. This result is in stark contrast to the smooth surface S_{hp} (Fig. 1e) that is predicted by most extant computational models of stereopsis.

3.1. One-dimensional Models

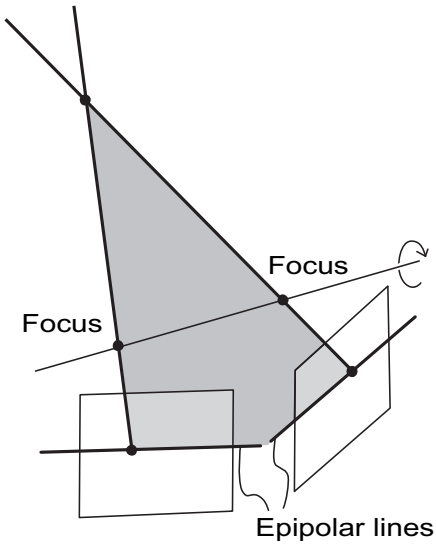

Fig. 2. The geometry of stereopsis. Only points on the corresponding epipolar lines can match to each other.

First of all, any 1D interpolating model would predict the ruled surface S_{hp}. The three-dimensional geometry of image formation dictates the possible pairs of points in the image that can match each other (Fig. 2). A point in a 3D scene and the two focal points determine a plane in the space. The projecting rays from the point through the focal points onto the retinae must lie on this plane. Thus, when the correspondence is not known, it can at least be said that a feature on one image can match only those locations on the other image that lie on the plane determined by the point and the two foci. Such possible matching points form a line called the epipolar line. Imagine a plane rotating around the line connecting the two foci: it sweeps the retinae, defining a set of corresponding epipolar lines. Geometrically, only points on the corresponding epipolar lines can match to each other. Thus, in theory, stereopsis can be a 1D process that matches the locations on the two images line by line. One may be lead to postulate that the interpolation is also done one-dimensionally. However, the experiment shows that is not what is done in human perception. If the interpolation is done one-dimensionally on each epipolar line as Fig. 1f shows, the perceived surfaces S_0 and S_1 have forms that cannot be readily explained. Since the sole depth data given on each epipolar line are at the two endpoints, the only reasonable 1D interpolation is to connect the two points by a straight line, as shown as **e** in Fig. 1f. As a whole, the lines give the smooth surface S_{hp}. Thus, theories that only use one-dimensional information do not predict the surfaces seen by most human observers.

3.2. Gradient Minimization Models

In some computational models of stereopsis, epipolar lines are not independent. It would be useful to have an interaction between the matching on different epipolar lines even just for the sake of robustness in the presence of noise. Most current theories model the matching by a depth surface that gives a dense map of the depth at each point in the view. Mathematically, a depth surface S is typically represented by specifying the value of the depth $d^S_{i,j}$ at each of dense sample points, which usually are laid out as an equally-spaced grid $X = \{(x_i, y_j)\}$. In such models, distinguishing features such as intensity edges can give a strong evidence of matches, determining the depth value $d^S_{i,j}$ at some of the sample points. Ever since Marr & Poggio (1976, 1979) and Pollard et al. (1985), most computational models of stereopsis have used a weak smoothing scheme that in effect predicts a surface S that minimizes the total change in depth:

$$E(S) = \sum_X \{(d^S_{i+1,j} - d^S_{i,j})^2 + (d^S_{i,j+1} - d^S_{i,j})^2\}, \quad (1)$$

which approximates the total depth gradient $\int\int\left|\nabla d^S\right|^2$. Here, the sum evaluates how much the depth value changes from one sample point to the next. One can see that the value $E(S)$ is minimum when the surface is flat and the depths $d^S_{i,j}$ at all points are equal, in which case $E(S)=0$. When some data points have definite depth values that come from the matching, which is usually the case, these models predict a depth surface S that minimizes $E(S)$ under the constraint that it has definite values where they are known. Or, the data from the feature matching are evaluated as another "energy" function and the sum of the two is minimized. This can be considered as giving a probability to each possible surface. If the surface has the depth value that is strongly supported by the matching, it would have a higher probability; other than that, the surface has higher probability when the sum (1) is smaller. In the Bayesian formulation of stereopsis (Szeliski, 1989; Belhumeur & Mumford, 1992; Belhumeur, 1996), this "energy" corresponds to a negative logarithm of the *prior probability distribution*, which gives an *a priori* probability for possible surfaces. It represents the model's idea of what surfaces are more likely in the absence of data. In the case of the image pair in Fig. 1a, the edges determine the depth at the two intensity edges that can be matched, as illustrated in Fig. 1b. At other sample points, however, there is not enough data to decide what depth to give to the point. This is why the model must have some disambiguating process.

How would such models react to the stereo pair in Fig. 1a? The answer is that all current theories predict a surface similar to $S_{\rm hp}$, rather than the most perceived surfaces S_0 and S_1. This is because the gradient modulus $\left|\nabla d^S\right|$, at all points, is larger for S_0 and S_1 than for $S_{\rm hp}$. This can be easily seen by simple calculation.

Let $2l$ be the side of the square and $2h$ the height (the difference of the maximum and the minimum depth) of the surface. We set up a coordinate system where the four corners of the square have the coordinates $(x,y)=(\pm l,\pm l)$ (Fig. 3a). Of the definite depths determined by matching the intensity edges, we assume that the two corners (l,l) and $(-l,-l)$ have the depth h and the other two have the depth $-h$. Thus, the boundary condition is the two line segments, shown as the thick line segments in Figs. 3b, c, and d, determined by the equations

$$x=l,\ d=\frac{h}{l}y,\ -l\le y\le l$$

and

$$x=-l,\ d=-\frac{h}{l}y,\ -l\le y\le l\,.$$

Then, the depth and the depth gradient for the surfaces S_0 and S_1 are as follows (Fig. 3b, c):

$$S_0:\ d^{S_0}(x,y)=\begin{cases}\frac{h}{l}(x-y)+h & (x\le y)\\ \frac{h}{l}(-x+y)+h & (x\ge y)\end{cases}\qquad \nabla d^{S_0}(x,y)=\begin{cases}(\frac{h}{l},-\frac{h}{l}) & (x<y)\\ (-\frac{h}{l},\frac{h}{l}) & (x>y)\end{cases}$$

$$S_1:\ d^{S_1}(x,y)=\begin{cases}\frac{h}{l}(x+y)-h & (x\ge -y)\\ \frac{h}{l}(-x-y)-h & (x\le -y)\end{cases}\qquad \nabla d^{S_1}(x,y)=\begin{cases}(\frac{h}{l},\frac{h}{l}) & (x>-y)\\ (-\frac{h}{l},-\frac{h}{l}) & (x<-y)\end{cases}$$

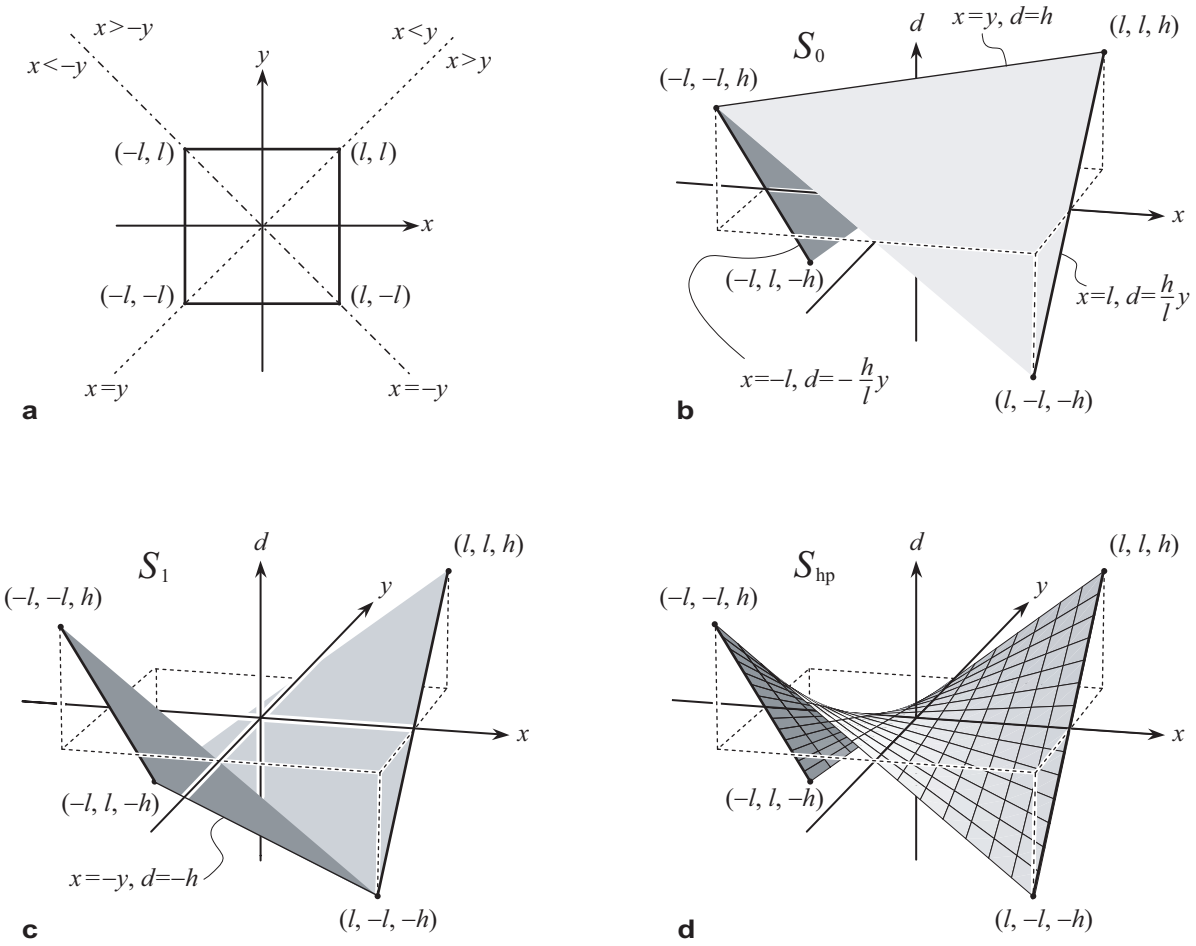

Fig. 3. The coordinate system for showing that gradient minimization is not enough.
(a) The possible surfaces are described by specifying the depth at each point inside a square. We define the coordinate system so that the four corners of the square are at $(x, y) = (\pm l, \pm l)$. The two diagonal lines $x = y$ and $x = -y$ separate the areas where different planes represent the surfaces in S_0 and S_1.
(b) The boundary condition is shown as the two thick line segments, determined by the equations $x = l,\ d = (h/l)y,\ -l \le y \le l$ and $x = -l,\ d = -(h/l)y,\ -l \le y \le l$. The surface S_0 consists of two planes defined on the two areas separated by the line $x = y$: one plane is defined by $d^{S_0} = (h/l)(x-y)+h$ in the area defined by $x \le y$ whereas the other is defined by $d^{S_0} = (h/l)(-x+y)+h$ where $x \ge y$. The two planes coincide on the line $x = y$.
(c) Similarly, the surface S_1 consists of the two planes defined by $d^{S_1} = (h/l)(x+y) - h$ (where $x \ge -y$) and $d^{S_1} = (h/l)(-x-y) - h$ (where $x \le -y$), meeting on the line $x = -y$.
(d) The surface S_{hp} is defined by $d^{S_{\text{hp}}} = (h/l^2)xy$ everywhere on the square.

Thus, we obtain $\sqrt{2}h/l$ as the gradient modulus for S_0 and S_1 everywhere on the square, except on the diagonal where it is not defined, i.e., $x = y$ for S_0 and $x = -y$ for S_1. On the other hand, the depth and its gradient for S_{hp} (Fig. 3d) at point (x, y) are defined by

$$S_{\text{hp}}: \quad d^{S_{\text{hp}}} = \frac{h}{l^2}xy, \qquad \nabla d^{S_{\text{hp}}} = (\frac{h}{l^2}y, \frac{h}{l^2}x)$$

Thus the gradient modulus for S_{hp} at point (x, y) is $\frac{h}{l^2}\sqrt{x^2 + y^2}$, which is smaller than $\sqrt{2}\frac{h}{l}$ wherever $x^2 + y^2$ is smaller than $2l^2$, which is the case inside the square. In fact, this observation rules out not only the energy (1) but also any energy that is the sum of an increasing function of the gradient modulus, which is to say most models.

3.3. Convex Models

To rule out the rest of the current prior models (and more), we can consider a functional of the form

$$E(S) = \sum_X f(\delta d^S), \tag{2}$$

where δd^S represents the derivative of some order of the depth function. For instance, the first-order case is the gradient such as in (1). The derivative δd^S, which in general is a vector, can be of any order, or a combination of several derivatives of different orders. Then, for a real number u between 0 and 1, we define a surface S_u that interpolates the two surfaces:

$$d^{S_u} = (1-u)d^{S_0} + ud^{S_1} \quad (0 \le u \le 1).$$

We assume that f is a *convex* function of the derivative. In general, a function $f(x)$ that has the property

$$f((1-u)x_0 + ux_1) \le (1-u)f(x_0) + uf(x_1), \qquad (0 \le u \le 1)$$

is said to be convex (see Fig. 4a.) If f is convex, then

$$E(S_u) \le (1-u)E(S_0) + uE(S_1) \le \max\{E(S_0), E(S_1)\}$$

implies that any linear interpolation of the two surfaces has the value of $E(S)$ that is at least as small as the larger of the values for the two surfaces. Moreover, if the energy is symmetric with respect to the sign of depth, it would give $E(S_0) = E(S_1)$; and if the energy is strictly convex, the extremes S_0 and S_1 would be maxima among all the interpolated surfaces, not minima. All theories of which the authors are aware satisfy the latter two conditions. We conclude that the most perceived surfaces are not predicted by any theory that uses the minimization of the energy function of the form (2) with convex f for disambiguation. Most current theories employ a convex energy functional as their prior, when seen in this representation. The minimization problem of the continuous version of (1) (called the Dirichlet integral) has the surface S_{hp} as the solution.

3.4. Non-Convex Models

Note that δd^S has a boundary condition that its total sum in each order is constant. In order to minimize a sum $f(x) + f(y)$ of a convex function $f(x)$ while keeping $x + y$ constant, the value should be distributed as much as possible. Thus convex energy functions such as (1) tend to round the corners and smooth the surface. What, then, about functions that are not convex? More recent theories of stereopsis use sophisticated priors that model

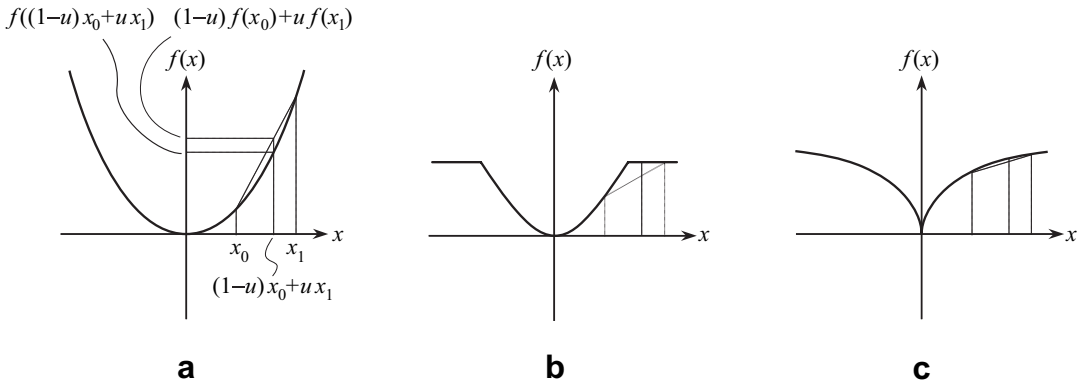

Fig. 4. Convex and non-convex functions
(a) A convex function. Note that the line connecting $(x_0, f(x_0))$ and $(x_1, f(x_1))$ is above the graph of the function. This is because the convexity requires $f((1-u)x_0 + ux_1) \le (1-u)f(x_0) + uf(x_1)$ for any u $(0 \le u \le 1)$.
(b) A convex function with a cut-off value. Because of the cut off, the function is not convex, as can be seen where the line comes below the graph.
(c) A function that is concave on both sides of the origin.

discontinuity in depth and slope. One such model (Belhumeur, 1996) minimizes the second derivative of depth, except for certain locus where it gives up and allows discontinuities in slope, or a crease, making it non-convex. In effect, it uses a function $f(x)$ such as shown in Fig. 4b, which is still convex in low-modulus region but with a cut-off value beyond which the function value stays constant. This model actually can predict S_0 and S_1, with right parameter values, since both surfaces have zero second derivatives except at the crease, where the curvature can be as high as needed without any impact on the functional more than the cut-off value.

However, this model fails to predict the outcome on the other stereogram, shown in Fig. 5a. Most observers reported a percept of one of the surfaces in either Fig. 5c or d. Assume that the non-convex energy above predicts the outcome. That there are no creases in the two solutions indicates that the curvature stays below the cut-off value everywhere. Since the function is convex in this domain, and because any interpolation of the two solutions would also have no creases, the same argument as the convex case applies. It follows that any interpolation of the two surfaces would have lower function value than the higher of the two.

Going even further, we can think of using almost concave functions, such as shown in Fig. 4c. This is akin to minimizing $\sqrt{x} + \sqrt{y}$ while keeping $x + y$ constant, and tend to concentrate the value at fewer variables. If we use such a function $f(x)$ in (2) with a second-order δd^S, it would try to concentrate the second derivative at fewer points, and always predict a creased, piecewise-flat surface, never a smooth surface as shown in Fig. 5c or d.

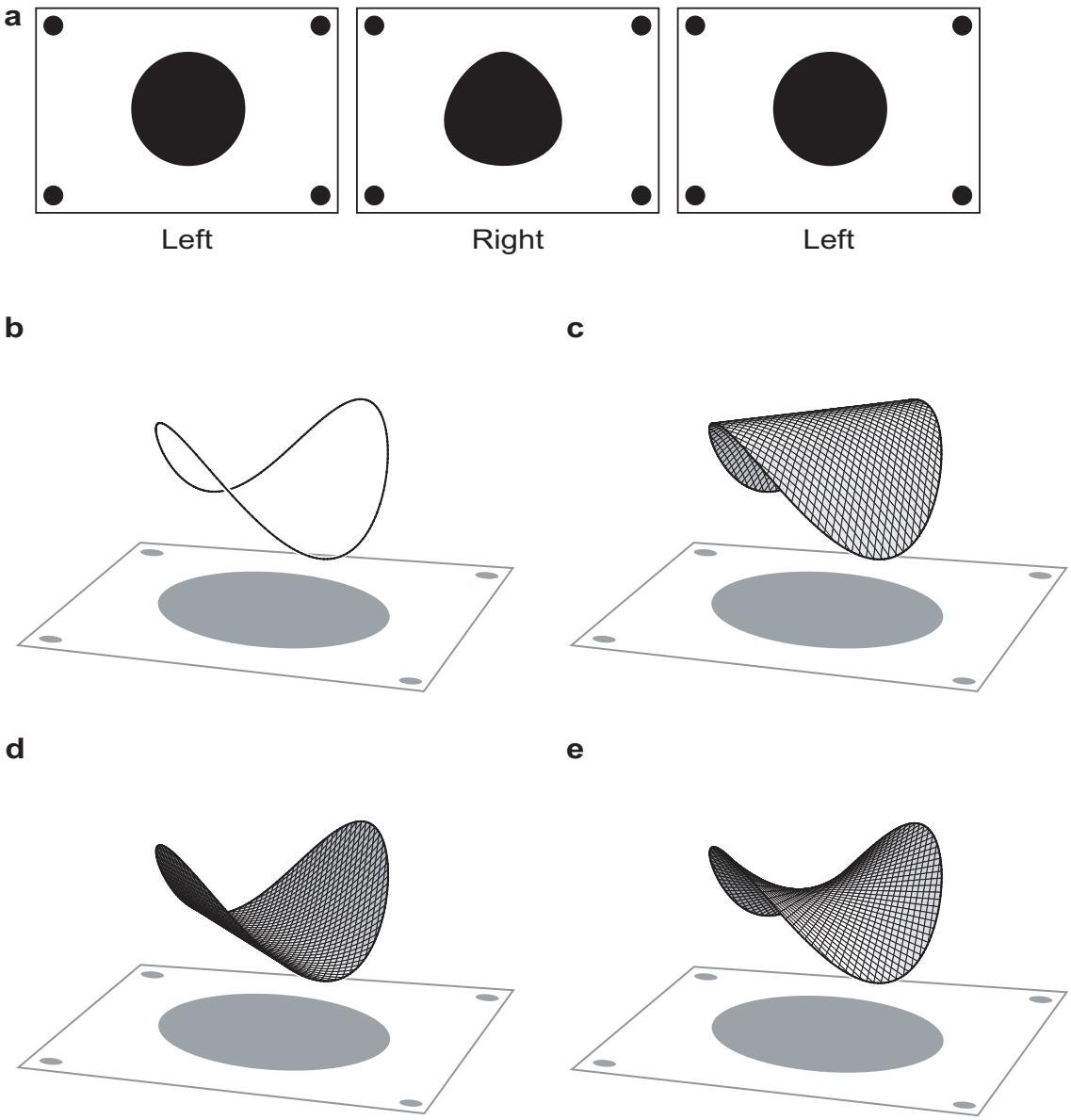

Fig. 5. Another stereogram further rules out possible theories.
(a) Another stereo pair.
(b) The unambiguous wire frame.
(c), (d), and (e) are all possible surfaces that agree with the boundary condition. The human brain tend to perceive either (c) or (d). Even algorithms that use non-convex functionals to allow discontinuities in depth and slope cannot have the solution (c) and (d) without having (e) as a better solution.

4. Discussion

Thus, it would seem difficult to produce a computational model that correctly predicts these observations using the kind of energy-minimizing scheme as currently used. The experiment shows a clear tendency towards the opposite direction than such theories predict, leading us to the conclusion that the current computational theories of stereopsis are not adequate for explaining the disambiguation by the human brain.

Is it possible that minimum disparity gradient or some similar models are a good model of stereopsis except when overridden by a strong prior preference for some special features? For instance, in the case of Fig. 1a, the observers' percepts might be biased towards S_0 and S_1 and away from S_{hp} because of the linear contours and sharp corners of the black square, since normally straight boundary contour edges derive from polygonal objects while curved contour edges derive from curved surfaces. However, the second experiment shows that even in the case where there are no straight edges, the percepts tend to be those of the extreme surfaces, rather than the hyperbolic paraboloid. Three of the observers were shown the round shape in Fig. 5a first (thus no bias because of the other pair) and still reported the percept of either convex or concave shape. Note that the same computational models are excluded by this experiment alone, by the same argument as above. Thus, even if there is a bias towards linear surfaces for linear contours and sharp corners, it is not enough to explain the observation, nor does it change our conclusion.

Also, it has been demonstrated (Mamassian & Landy 1998) that human perception prefers elliptic (egg shaped) to hyperbolic (saddle shaped). Since the prediction of current theories is hyperbolic, the observed departure from it may be because of this bias. However, note that all the surfaces that are preferred are parabolic, i.e., neither elliptic nor hyperbolic. This is remarkable since the parabolic case constitutes a set of measure zero in the space of all possible local shapes. Because of this, it is hard to argue that any tendency or bias toward elliptic brought the percept exactly to that rare position.

However, the fact that the percepts are parabolic suggests another possibility. Namely, the Gaussian curvature of the four preferred surfaces in the two experiments is zero everywhere it is defined. Zero Gaussian curvature is a characteristic of parabolic points. Thus, minimizing the total sum of the absolute value or square of Gaussian curvature, for example, would predict all four surfaces as minima. Such surfaces are developable, meaning they can be made by rolling and bending a piece of paper. In other words, one possibility is that the human vision system tries to fit a paper on the boundary wire frame (the sparse frame that represent definite depth data shown in Fig. 1b and Fig. 5b in the case of the experiments). Such a functional would be neither convex nor concave, and nonlinear, which means that the solutions depend on the starting location; that makes the analysis of such a problem nontrivial, although a study of such a functional is surely a possible direction for future research.

Note, also, that in both of the examples the two surfaces perceived by the human brain are the front and back of the convex hull of the boundary wire frame. A set in a space is called convex when any line segment that connects two of its points is also contained in it. The convex hull of a set of points is the minimal convex set containing all the points. In the case of Fig. 1b, the convex hull is the tetrahedron defined by the four endpoints of the line segments with definite depth data.

Interestingly, there is a strong connection between the convex hull and the zero Gaussian curvature: a point on the surface of the convex hull of a set (such as the boundary wire frame), if it does not belong to the original set, has zero Gaussian curvature wherever it is defined.

(We could not find any proof in the literature. Since it is short enough, we include it in the appendix.) Thus we at least know that the minimization problem of the Gaussian curvature has solutions: we can take the convex-hull of the wire frame and take its front and back surfaces.

We conclude this paper with a speculation. The convex-hull interpretation suggests that the percept of these surfaces arise from the later stage of the human visual process, where the information leaves the retinotopic realm and is represented as more complex constructs that express qualitative scene geometry (Anderson, 1999). Specifically, we hypothesize that an illusory volume is induced by the matched intensity contour, which in turn induce the percept of the surfaces. The volume defined by the convex hull has in a sense the simplest 3D shape that is compatible with the data, much in the way the Kanizsa triangle (Kanizsa, 1979) is the simplest 2D shape that explains incomplete contour information. Indeed, we conjecture that the processes that give rise to the illusory contours and the stereo surfaces in the cases like these experiments are one and the same.

Appendix

Here, we give a proof of the proposition that is mentioned in the Discussion.

Proposition. *Let A be a set in the three-dimensional Euclidean space, B its convex hull, and p a point in $\partial B \setminus A$, where ∂B denotes the boundary of B. Assume that a neighborhood of p in ∂B is a smooth surface. Then the Gaussian curvature of ∂B at p is zero.*

Proof : Since p is in B and not in A, there are finite number of distinct points $q_1, \cdots, q_n$ $(n \geq 2)$ in A and positive numbers $a_1, \cdots, a_n$ such that $p = \sum_{i=1}^{n} a_i q_i$, $\sum_{i=1}^{n} a_i = 1$. Also, since p is on the boundary ∂B of a convex set B, all points of B are in the same half space H whose boundary ∂H is the tangent plane of ∂B at p. Since p is on the plane ∂H and all q_i's are in H, it follows that all q_i's are on ∂H because a_i's are all positive. Consider the convex hull C of $\{q_1, \cdots, q_n\}$. Then $C \subset \partial B$ since $C \subset B \cap \partial H$. Since C is not a point, the plane ∂H is tangent to the surface ∂B around p along at least a line segment. Thus the Gaussian curvature of ∂B at p is zero. □

References

1. Anderson, B.L. (1994). The role of partial occlusion in stereopsis. *Nature*, *367*, 365–368.
2. Anderson, B.L. (1999). Stereoscopic Surface Perception. *Neuron*, *24*, 919–928.
3. Ayache, N. (1991). *Artificial Vision for Mobile Robots*. Cambridge, MA: MIT Press.
4. Belhumeur, P. N. (1996). A Bayesian approach to binocular stereopsis. *Int. J Comput. Vision*, *19*, 237–262.
5. Belhumeur, P. N. & Mumford, D. (1992). A Bayesian treatment of the stereo correspondence problem using half-occluded regions. In: *Proc. IEEE Conf. on Comput. Vision and Pattern Recognition (CVPR '92)*. 506–512.
6. Faugeras, O. (1993). *Three-Dimensional Computer Vision*. Cambridge, MA: MIT Press.

7. Geiger, D., Ladendorf, B., & Yuille, A. (1995). Occlusions and binocular stereo. *Int. J Comput. Vision*, *14*, 211–226.
8. Gillam, B. & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, *17*, 603–608.
9. Grimson, W. E. L. (1981). *From Images to Surfaces*. Cambridge, MA: MIT Press.
10. Jones, J. & Malik, J. (1992). *Image Vision Comput*., *10*, 699–708.
11. Kanizsa, G. (1979). *Organization in Vision*. New York: Praeger.
12. Malik, J. (1996). On Binocularly viewed occlusion junctions. In: *Fourth European Conference on Computer Vision, vol.1.* Springer-Verlag, 167–174.
13. Mamassian P. & Landy M.S. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research*, *38*, 2817-2832.
14. Marr, D. & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, *194*, 283–287.
15. Marr, D. & Poggio, T. (1979). A computational theory of human stereo vision. *Proc. R. Soc. Lond. B*, *204*, 301–328.
16. Nakayama, K. & Shimojo, S. (1990). Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research*, *30*, 1811–1825.
17. Norman, J. F., & Raines, S. R. (2002). The perception and discrimination of local 3-D surface structure from deforming and disparate boundary contours. *Perception & Psychophysics*, *64*, 1145–1159.
18. Pollard, S. B., Mayhew, J. E. W., & Frisby, J. P. (1985). PMF: A stereo correspondence algorithm using a disparity gradient. *Perception*, *14*, 449–470.
19. Poggio, G., & Poggio, T. (1984). The Analysis of Stereopsis. *Annu. Rev. Neurosci.*, *7*, 379–412.
20. Szeliski, R. (1989). *A Bayesian modelling of uncertainty in low-level vision*. Boston, MA: Kluwer Academic Press.
21. Ulupinar, F. & Nevatia, R. (1993) *Perception of 3-D Surfaces from 2-D Contours*. *IEEE Transactions on Pattern Analysis and Machine Intelligence 15*, 3–18.