

The shading cue in context

Johan Wagemans[¶]

University of Leuven (K U Leuven), Laboratory of Experimental Psychology, Tiensestraat 102 bus 3711 3000 Leuven; present address: Tiensestraat 102 bus 3711 3000 Leuven, Belgium;
e-mail: johan.wagemans@psy.kuleuven.be

Andrea J van Doorn

Delft University of Technology, Industrial Design, Landbergstraat 15, 2628 CE Delft, The Netherlands;
e-mail: a.j.vandoorn@tudelft.nl

Jan J Koenderink

University of Leuven (K U Leuven), Laboratory of Experimental Psychology, Tiensestraat 102 bus 3711 3000 Leuven, Belgium. Also at Delft University of Technology, EEMCS and The Flemish Academic Centre for Science and the Arts; e-mail: j.j.koenderink@tudelft.nl

Received 28 July 2010, in revised form 29 November; published online 21 December 2010

Abstract. The shading cue is supposed to be a major factor in monocular stereopsis. However, the hypothesis is hardly corroborated by available data. For instance, the conventional stimulus used in perception research, which involves a circular disk with monotonic luminance gradient on a uniform surround, is theoretically ‘explained’ by any quadric surface, including spherical caps or cups (the conventional response categories), cylindrical ruts or ridges, and saddle surfaces. Whereas cylindrical ruts or ridges are reported when the outline is changed from circular to square, saddle surfaces are never reported. We introduce a method that allows us to differentiate between such possible responses. We report observations on a number of variations of the conventional stimulus, including variations of shape and quality of the boundary, and contexts that allow the observer to infer illumination direction. We find strong and expected influences of outline shape, but, perhaps surprisingly, we fail to find any influence of context, and only partial influence of outline quality. Moreover, we report appreciable differences within the generic population. We trace some of the idiosyncrasies (as compared to shape from shading algorithms) of the human observer to generic properties of the environment, in particular the fact that many objects are limited in size and elliptically convex over most of their boundaries.

1 Introduction

The shading cue (Horn and Brooks 1989; Luckiesh 1916; Metzger 1975; Turhan 1935) is one of the generic pictorial, or monocular, depth cues. The pictorial cues enable monocular three-dimensional spatial vision (Palmer 1999), and pictorial spatial vision (where binocular disparity merely serves to reveal the picture surface as flat) (Koenderink et al 1994). Three-dimensional spatial vision on the basis of monocular cues is known as ‘monocular stereopsis’. The shading cue is part of the optical interface of animals of many genera (Metzger 1975; Riedl 1984), including *homo*, playing a key role in camouflage and foraging. For instance, animals living on relatively featureless flattish terrain tend to be dark dorsally, light ventrally (Metzger 1975). Such pigmentation implements a ‘countershading’ that tends to optically ‘flatten’, and thus ‘dematerialize’ them, an important goal of camouflage. Newly hatched chicks peck at circular disks filled with linear luminance gradients in their visual fields (Hershberger 1970; Hess 1950; Metzger 1975), yielding a heightened probability to aim pecking activity at graminoid seed grains, thus promoting foraging success (Riedl 1984). In both examples the direction of the luminance gradient is important. Things tend to appear ‘object-like’ (‘animal’, ‘grain’, etc), that is *convex*, if they are light on top, dark at bottom, a polarization that can be traced to the predominantly tendency of natural illumination to be directed top down (Metzger 1975; Riedl 1984; Süffert 1932; Thayer 1909). Illumination from above derives

[¶] Corresponding author.

from both direct sunlight (the sun generally appearing above the horizon) and overcast skies (the zenith being the brightest patch in the scene) (Minnaert 1993). It is a common understanding in the literature that the ‘light from above assumption’ is a crucial part of the optical interfaces of the majority of genera (Riedl 1984), including man (Brewster 1832; Rittenhouse 1786).

The visual arts have exploited the shading cue since the earliest times, although only in recent times (middle ages in Western Europe) in a rational, explicit manner. In the nineteenth-century art academies, shading (or *chiaroscuro*) became an important aspect of the curriculum, easily at a par with linear perspective. Students would spend years in the cast room, patiently shading their drawings of plaster casts of classical sculptures. The shading cue was supposed to yield relief to otherwise flattish, cartoon-like drawings, the renaissance distinction of *rilievo* (or the ‘reception of the light’, also called *colorire*) and *disegno* (Baxandall 1972).

In the first half of the twentieth century, the shading cue was widely researched by the then mainly phenomenologically oriented psychology of perception of continental Europe, especially the Gestaltists of the Graz and Berlin schools. Most of our present understanding of the shading cue is due to this literature. Only much later, 1970s and ’80s, the topic reemerged in the Anglo-Saxon literature of psychology (Ramachandran 1988a, 1988b). Even more recently, 1980s till present, the topic emerged in computer vision (Forsyth and Ponce 2002; Horn and Brooks 1989; Zhang et al 1999), with the advent of formal, instead of mainly phenomenological, accounts. The earliest formal developments actually go back to lunar astronomy of the 1950s (van Diggelen 1951), but these appear to have remained ineffective with respect to perceptual studies. Much of the interest has centered on the ‘light from above’ prior (Adams 2007, 2008; Adams et al 2004; Kleffner and Ramachandran 1992; Mamassian and Goutcher 2001; Ramachandran 1988a; 1988b; Sun and Perona 1998) and the inherent ambiguity of the cue (Hill and Bruce 1994).

In the current literature of the psychology of visual perception, formal accounts of the level of physico-mathematical detail as common in machine vision do not seem to play a major role. Conversely, the current literature of machine vision reveals little understanding of the achievements of psychology. This actually makes some sense, either way, because the connection is rather less immediate than it is often made out to be. The relations are important for the present paper, which is why we start with a more formal discussion, although our contribution is mainly of an empirical, investigative nature.

2 The shading cue, a formal account

The conventional stimulus of biological vision research, with respect to the shading cue, is a circular disk filled with a linear luminance gradient, usually in a uniform surrounding field, often of roughly the average luminance of the disk (figure 1). It is frequently implied that this largely exhausts the possibilities, apart from having more of the same, and that this stimulus may evoke one of two alternative impressions, namely convexity or concavity. Although this is rarely explicitly acknowledged, in many experiments the stimuli are of this very type, and the acceptable responses are limited to these alternatives, whereas the conclusions are of a very general nature. Thus, we believe this characterization to be a fair one, although there are exceptions, of course.

Although not often explicitly formulated, the conventional stimulus is evidently an attempt to isolate the shading cue proper. The choice of the circular aperture serves to effectively localize the cue; the degree of localization can be controlled through the choice of diameter. Since shading requires a finite area for its definition, a rotationally symmetric aperture of limited size is the obvious choice. The choice of a linear gradient serves to define

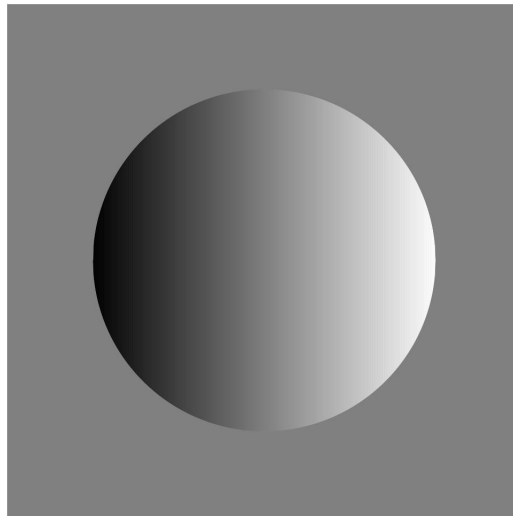


Figure 1. A conventional stimulus configuration in shape from shading psychophysics. The relevant feature is supposed to be the linear gradient (changing from black to white from left to right). Here, the gradient is put in a circular disk, superimposed on a uniform background of the average grey level, without encircling the circular shape in any other way. This is typical for many studies.

the purely local structure of the shading. At any generic location (eg not an extremum) of an arbitrary shading pattern the structure can arbitrarily be well approximated by such a linear gradient if the aperture is appropriately restricted. The structure of the shading is then fully described by the spatial gradient of the illuminance, a vectorial point property. It seems likely that the original choice for the conventional stimulus, sometime during the nineteenth century, was motivated by such considerations. It can easily be formalized in mathematical terms.

In a formal view, the linear gradient may thus be understood as a linear approximation to arbitrary smooth luminance distributions; in that respect the choice is indeed a natural one. The implication is that one understands the cue to be a local one; this is part of the linear approximation. In biological terms one assumes the existence of receptive field structures dedicated to the shading cue. In this respect, the conventional stimulus could also be called a minimal stimulus because the linear gradient is an abstraction of the first-order differential structure at a single point, rather than the pattern of illuminance over an extended surface patch.

A general shading pattern will have the illuminance gradient changing from point to point. It can be understood, at least in the formal sense, as a linear superposition of local samples. In a reductionist framework it then makes sense to study the local case first, because it embodies the information that is provided by a single local gradient detector, which many have posited to be the relevant information for the perception of local curvature. (In this respect, the choice is similar to that of sine-wave gratings as a basis for the description of arbitrary illuminance patterns.) Psychophysical research on more complicated cases involving shading requires very different methods from the ones discussed in this paper (Koenderink et al 1996).

One conceptual problem with the choice of the conventional stimulus is the nature of the boundary of the aperture, in this case a sharp, circular edge. This introduces ambiguity because the edge can be interpreted in a number of different ways as a depth or shape cue. We consider this issue in this paper.

Other conceptual problems have to do with the physics of illuminated, scattering, curved surfaces. If the conventional stimulus leads to monocular stereopsis, then the observer apparently made a number of implicit assumptions concerning the pictorial scene. The physics is complicated, but the conventional interpretation in psychology is in terms of Lambert's cosine law. This ignores many factors, such as the location and nature of the source, the physics of surface scattering, the effects of multiple scattering, and the effects of vignetting. Only when Lambert's law dominates is there a one-to-one relation between the radiance at the pupil of the observer's eye and the spatial attitudes of surface elements in the scene. Thus the conventional interpretation is indeed a natural one. It is also the natural setting for the simplest machine vision algorithms involving shading.

Physical optics yields a simple model for shading in the form of Lambert's cosine law. This ignores a number of physical effects that are typically important in real settings, though (Forsyth and Zisserman 1991; Koenderink and van Doorn 1983). The illumination of a surface element is proportional to the cosine of the angle subtended by its surface normal and the direction towards the light source as seen from the surface. This relation has been known since the eighteenth century (Bouguer 1729; Lambert 1760). In itself this relation is not important in vision, because observers sample radiance, not irradiance, whereas the relation between the two is often a complicated one. For typical surfaces the radiance depends both upon the direction of illumination and the viewing direction. Only for Lambertian surfaces does the viewing direction not matter and does one have a strict linear relation between the radiance received by the eye and the irradiance of the surface. Fortunately, many diffusely scattering natural surfaces (like paper) are approximately Lambertian. Given a smooth Lambertian surface, the luminance will thus co-vary with the direction of the normal, the direction towards the source being generally fixed. Over small stretches the change will be approximately a linear gradient. The luminance gradient thus reveals a change of the surface normal, that is to say, curvature, or shape. Not any aspect of surface curvature will be thus 'revealed', though, because rotating the normal about the light direction will not lead to any change of the cosine, and it consequently fails to imprint itself on the shading. Apparently the shading cue is inherently ambiguous (see figure 2).

Understanding the nature of this ambiguity of the shading cue is of obvious importance. It is a complicated issue, though, and we will approach it in a number of steps.

Consider a uniform patch in the visual field, and assume it to be due to an illuminated surface of constant albedo (say, white paper or plaster), illuminated with a homogeneous, unidirectional beam. (Sunlight is an example; the technical term is collimated beam.) This is perhaps the simplest example of 'shading'. What inferences are possible? This is an instructive example. The magnitude of the luminance is clearly irrelevant; vision has sufficient 'constancies' to ensure that donning sunglasses is not going to change the perception of the geometry of the scene in front of you all that much. Thus, the relevant 'observable' is simply the absence of a luminance gradient. It reveals the absence of surface normal variations with respect to the (assumed a priori unknown) direction of the beam, so the possible inferences are an arbitrary beam direction illuminating a surface that subtends a fixed slant with that direction. Such surfaces include cones of rotation with axes coinciding with the beam direction. These are evidently non-generic, though, because there is no reason why the scene should be 'tuned' to the beam direction. Hence, a reasonable shape from shading algorithm will discard such (infinite) possibilities offhand. One ends up with planes. The scene could be any plane, of any spatial attitude, illuminated from any direction. Thus, you have a rather strong shape inference (a plane), although most of the scene geometry remains in doubt.

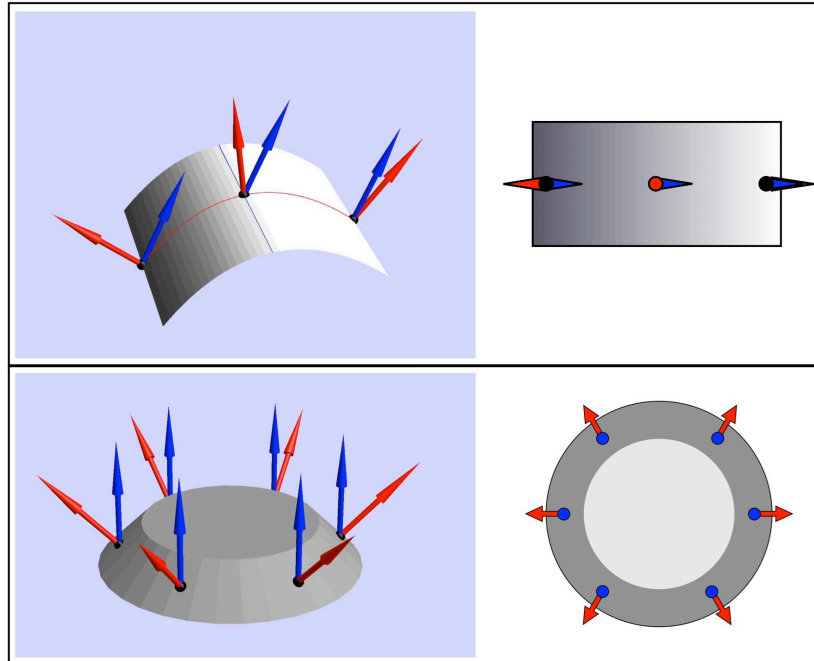


Figure 2. In the top row, a surface strip that curves along the direction of light flow. The blue arrows point towards the light source; the red arrows are surface normals. The angle subtended by the normals and the light direction decreases gradually from left to right in the picture because the normal turns due to the curvature. Due to Lambert's cosine law, the surface illumination increases gradually from left to right, hence the curvature is revealed by the shading. In the bottom row, the strip is curved transverse to the direction of light flow. The normals turn around the light direction, subtending a constant angle with it. Due to Lambert's cosine law, the surface illumination is constant around the strip. (This is not shown in the 3D rendering.) Thus, shading fails to reveal the surface curvature in this case. This is the basic ambiguity of 'shape from shading'. In the left column we show the 3D scene, in the right column the shading.

Next, consider a linear gradient. It remains the case that the absolute luminance has to be irrelevant; thus the 'observable' is the relative luminance gradient, a contrast. It is clear a priori that the spatial attitude of the surface element will remain unspecified, and so will the slant of the beam direction with respect to the surface. The relevant 'illumination direction' is the tilt (that is, the component of the beam direction at right angles to the surface normal direction); we call it the 'light flow direction', adopting the jargon of visual artists. One simple ambiguity has to do with the slant of the beam with respect to the surface. The gradient magnitude depends both on the curvature of the surface and on the slant of the beam with respect to the surface, more curvature and less slant leading to a greater gradient. Thus, you obtain a whole family of equivalent inferences. Another way of putting it is to say that shading does not reveal the depth of relief (see [figure 3](#)).⁽¹⁾

⁽¹⁾ Consider a shape $z(x, y) = \frac{1}{2}(ax^2 + by^2)$ in Cartesian coordinates, viewed from the direction of the Z-axis. The direction of the Z-axis is also the direction of the outward surface normal at the origin. More generally, the normal is $\mathbf{n}(x, y) = \{-ax, -by, 1\} / \sqrt{1 + a^2x^2 + b^2y^2}$. (The necessary differential geometrical background is summarized in Lipschutz 1969.)

Let the direction towards the source be $\mathbf{i}(t, s) = \{\cos(t) \sin(s), \sin(t) \sin(s), \cos(s)\}$ where t denotes the tilt, s the slant. Then the illuminance $I(x, y)$ equals $E\mathbf{n}(x, y) \cdot \mathbf{i}(t, s)$, by Lambert's cosine law. Here E is the normal illuminance caused by the beam, and the dot denotes the scalar (or dot) product. Thus one has $I(x, y) = E[-a \cos(t) \sin(s)x - b \sin(t) \sin(s)y + \cos(s)] / \sqrt{1 + a^2x^2 + b^2y^2}$. The irradiance at the origin is $I_0 = I(0, 0) = E \cos(s)$, which is just Lambert's cosine law. The gradient at the origin $x = y = 0$ is $\mathbf{g} = \text{grad } I(0, 0) = -E \sin(s)\{a \cos(t), b \sin(t)\}$ (by direct differentiation with respect to x and y). Perhaps a perceptually more relevant entity is the gradient contrast, which is $\mathbf{c} = \mathbf{g}/I_0$. One has

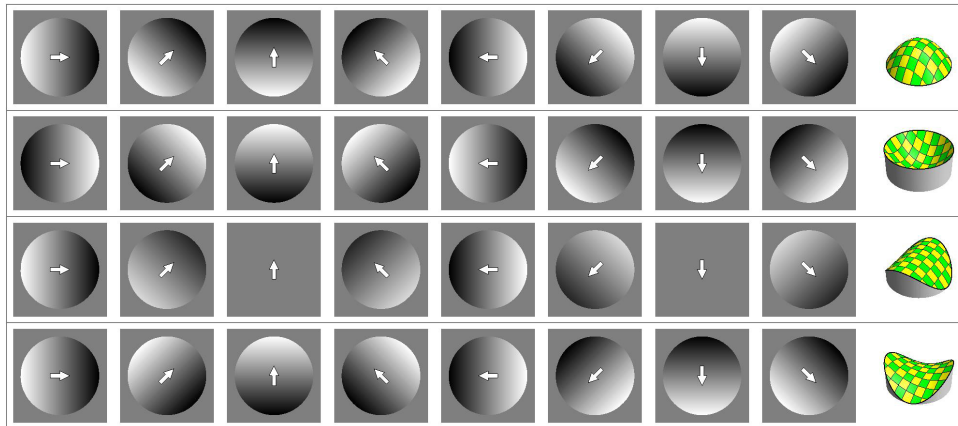


Figure 3. The basic shading ambiguity. This shows only the effect of the tilt; there is an additional ambiguity (known as the ‘bas-relief ambiguity’ in computer vision) due to the slant. The rows show the shading of the surfaces on the right for a variety of illumination directions (the tilt), as indicated by the arrows. Notice that the conventional stimulus allows valid interpretations of any of the surfaces in the right column. The ‘cap’ and ‘cup’ interpretations represent only part of the conventional response categories.

The tilt is very important; the slant mainly effects the contrast. If the curvature is orthogonal to the light flow direction, like a cylindrical gutter illuminated along its axis, this will fail to generate a gradient. Only curvature along the light flow direction, like a cylindrical ridge illuminated transverse to its axis, can lead to luminance modulations. This is an issue in cartography, where shading fails to reveal valleys and mountain ridges running along the (virtual) illumination direction. If the results are not acceptable, cartographers will often (arbitrarily) change the tilt locally. The general rule is simple enough: curvature in the direction of illuminance flow generates shading, whereas curvature orthogonal to it does not.

The simple considerations discussed above have far-reaching consequences. If you observe a luminance gradient, you thereby observe a curvature of the surface in the direction of the flow of illumination. However, the directions of principal curvature of the surface could be anything, and you need to infer not only their orientations but two independent principal curvatures. The upshot is that anything goes shape-wise. The surface could be ‘convex’ or ‘concave’ (implication being ‘umbilical’, that is to say, like the inside or outside of a spherical shell), but equally well cylindrical or saddle shaped. Thus the conventional response categories are artificially limited to two instances out of a continuum of principled possibilities. This is illustrated in [figure 3](#).

$\mathbf{c} = -\tan(s)\{a \cos(t), b \sin(t)\}$. Thus the direction of the gradient is given by the angle φ such that one has $\tan(\varphi) = b/a \tan(t)$.

This completely summarizes the fundamental theory of the conventional stimulus. It relates the surface shape (the ratio of principal curvatures b/a) and the tilt (t , reckoned with respect to a principal direction) to the direction of the linear gradient. When t is known and φ observed, the shape (b/a) is determined. If t is not known (the generic case) the shape and tilt are confounded. This type of shape ambiguity was considered by Freeman (1994) in global settings. Even in case t is known the magnitudes of a and b are confounded with the slant s : this is the ‘bas-relief ambiguity’ of machine vision.

For an umbilical $b/a = +1$, thus $\varphi = t$ (a cup), or $\varphi = t + 180^\circ$ (a cap), whereas for a symmetrical saddle $b/a = -1$, thus $\varphi = t + 90^\circ$, or $\varphi = t + 270^\circ$. Thus the inference should be a saddle if the gradient is at right angles to the light flow. When $\mathbf{c} = \{0, 0\}$, $s > 0$, one possible inference is $t = 0^\circ$, $a = 0$, a cylinder illuminated along its axis.

We conclude that the conventional response categories—‘cap’ or ‘cup’, both ‘umbilics’ in the terminology of the geometry of smoothly curved surfaces—by no means exhaust the actually relevant response categories. A large part of the literature (virtually all that uses the conventional stimulus) suffers from this constraint. One wonders how this came to be.

Here we meet our first research target: *Why do human observers limit possible shapes to umbilicals?* Are human observers somehow unable to see saddle shapes? There is indeed some historical indication for that. Leon Battista Alberti (1435) was an Italian intellectual who wrote an important treatise on painting, which contains an ‘exhaustive’ list of surface shapes. He writes (book I, paragraph 8):

We have now to treat of other qualities which rest like a skin over all the surface of the plane. These are divided into three sorts. Some planes are flat, others are hollowed out, and others are swollen outward and are spherical. To these a fourth may be added which is composed of any two of the above. The flat plane is that which a straight ruler will touch in every part if drawn over it. The surface of the water is very similar to this. The spherical plane is similar to the exterior of a sphere. We say the sphere is a round body, continuous in every part; any part on the extremity of that body is equidistant from its centre. The hollowed plane is within and under the outermost extremities of the spherical plane as in the interior of an egg shell. The compound plane is in one part flat and in another hollowed or spherical like those on the interior of reeds or on the exterior of columns.

Thus, Alberti lists the convexities and concavities, along with non-generic possibilities like cylinders and planes (having prior probability zero in the space of shapes), but he completely fails to list the (generic!) saddle shapes (figure 4). Alberti’s list remained unchallenged for centuries. The complete inventory of local surface shapes is due to Carl Friedrich Gauß and dates from the early-nineteenth century (Gauß 1827) (figure 5). Here, we focus our research target a bit tighter: Why are saddle shapes (figure 6) apparently ignored in human visual awareness?

A well-known observation is that the same linear gradient leads to different shape experience when rendered within different outlines. The same gradient that looks like a spherical shell in one case may look like a cylinder in another case. Given the ambiguities, it is not that surprising that perceptions may vary. What (perhaps) *is* surprising is that many observers have strong convictions of various kinds. Apparently, observers use more cues than just shading, and since it is not possible to present a ‘pure gradient’, boundaries will be present, and will be used as additional cues (see figure 7).

Consider how a ‘hard’ outline, like that used in the conventional stimuli, may appear (figure 8):

- as an ‘occluding contour’ as when looking at a sphere;
- as a ‘dihedral edge’ as when looking at an internal edge in a cube;
- dihedral edges may also appear as occluding contours (artist call these ‘cutting edges’);
- as the boundary of a surface patch, eg as when looking into a spherical cup (we will refer to it as a ‘flag edge’).

In case the outline has vertices (like a square), the interpretation may well change at a vertex. All such interpretations are generic, and it seems impossible to put a prior probability distribution on them. However, it certainly seems the case that some interpretations are more likely than others, given certain contexts. Some things seem unlikely, though—for instance, a change of interpretation along a smooth stretch of outline.

Simple contextual changes near the outline will load the priors on the various possible interpretations differently. Thus, one expects such modifications to change the experiences in the case of a single gradient; to investigate this is our second research objective.

Of course, one expects frequent disagreements between the visual awareness of different observers in experiments like these. Perhaps unfortunately, the literature is not provide a

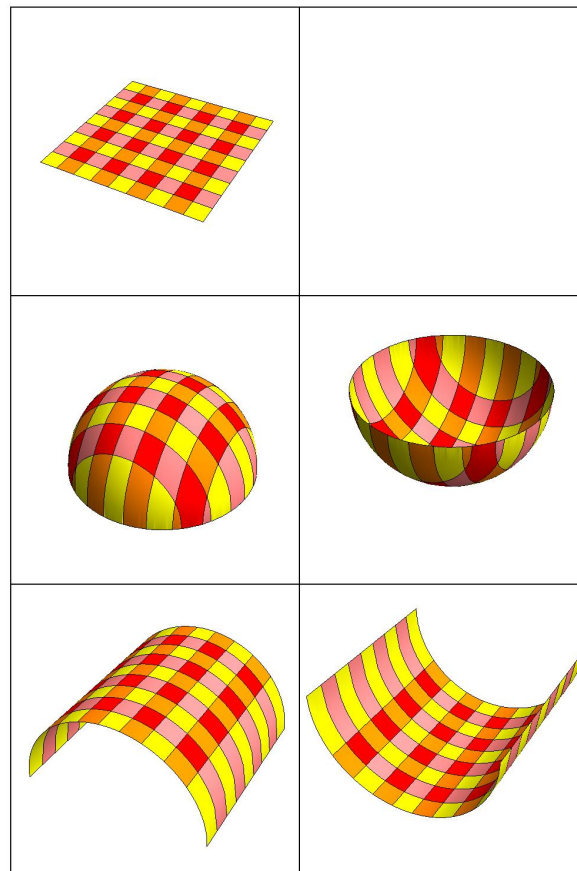


Figure 4. The inventory of local surface shapes from Alberti's *De Pictura*. Note the lack of hyperbolic (saddle) surfaces.

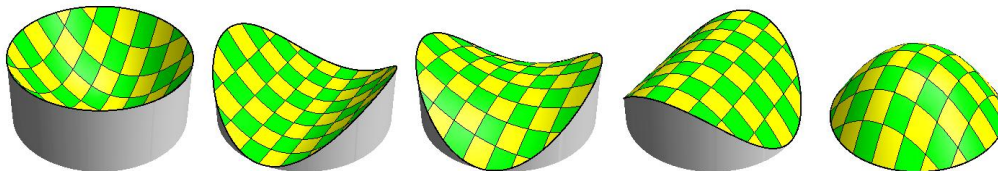


Figure 5. The full set of local surface shapes can be naturally parameterized by a finite length segment. The classification is due to Gauß. From left to right, you have a concave umbilic, a concave cylinder, a symmetrical saddle, a convex cylinder, and a convex umbilic. It is a continuous family; thus you have to imagine the interpolated shapes. The region between the cylinders comprises hyperbolic (saddle) shapes, whereas the outer regions are elliptic (like the inside or outside of egg shells). This rectifies and completes Alberti's list illustrated in figure 4.

rich source of data on that issue. But even a cursory investigation reveals major differences between generic observers. For instance, we estimate that at least one out of five persons fails to see any shading induced relief at all, even when confronted with the standard stimuli from the literature. This became evident when recruiting observers for the present task. Note that we did not continue testing these observers in the experiment reported below. Thus, the third research objective is to obtain an initial impression of the type of differences encountered in the generic population.

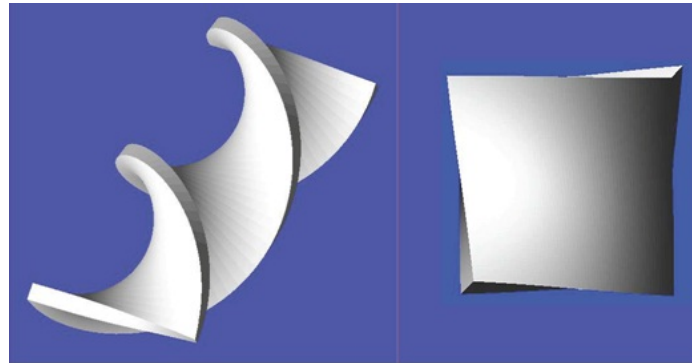


Figure 6. Left: a helicoidal (twisted) surface. Right: a square piece of it rendered as a ‘twisted thick plate’, illuminated from above. Note that the luminance gradient is at right angles to the illumination flow direction, something that is impossible with spheroidal surfaces. It is still just a linear gradient, though; thus saddles are possible interpretations of the conventional stimulus (figure 1), although they are never reported.

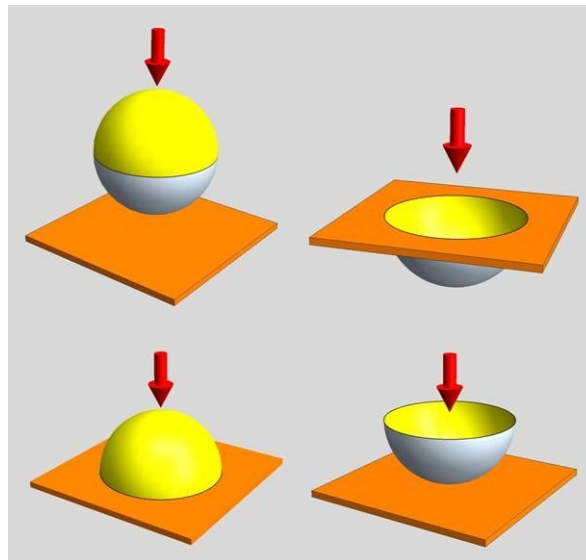


Figure 7. Some, though by no means all, possible interpretations of the circular outline of the conventional stimulus (see footnote 1). Suppose the interpretation is ‘cap’ (the left column), then the outline is often interpreted as an occluding contour (top) or a dihedral edge (bottom). In case the interpretation is ‘cup’ (the right column), the outline is often seen as a dihedral edge (top) or a flag edge (bottom). In the latter case the interpretation as an occluding contour does not work. Thus, the cup and cap interpretations are not symmetrical with respect to the possible interpretation of the circular outline as a depth cue.

3 Experiment

3.1 Stimuli

Our aim was to investigate the awareness of surface shape due to a single linear gradient of given extent in the presence of varying contexts. We considered different types of context, changing background, nature of outline, and indication of illuminance flow direction.

As contexts we used:

- a uniform background of contrasting color, suggesting a distant backdrop (figure 9, upper left);
- a uniform background of the average luminance, suggesting a substrate of the same material, the object most likely being ‘part’ of it (figure 9, upper right);

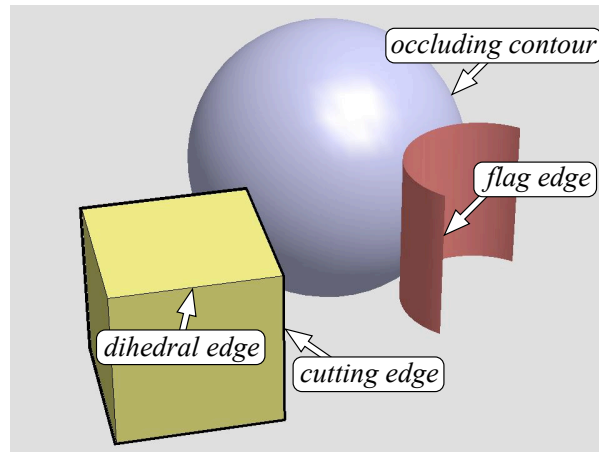


Figure 8. Examples of occluding contours (the sphere), dihedral edges (internal edges of the cube), and cutting edges (external edges of the cube; both cutting edges and occluding contours), and flag edges (the edge of the hemi-cylindrical surface on the right).

- a uniform background of the average luminance, suggesting a substrate of the same material, the object most likely being part of it, but ‘thickened’ and apparently illuminated, so as to reveal the direction of illumination (figure 9, lower left);
- a uniform background of the average luminance with an aperture, suggesting a substrate of the same material, ‘thickened’ and apparently illuminated, so as to reveal the direction of illumination. The object is ‘seen through the aperture’ and appears on the blue backdrop (figure 9, lower right).

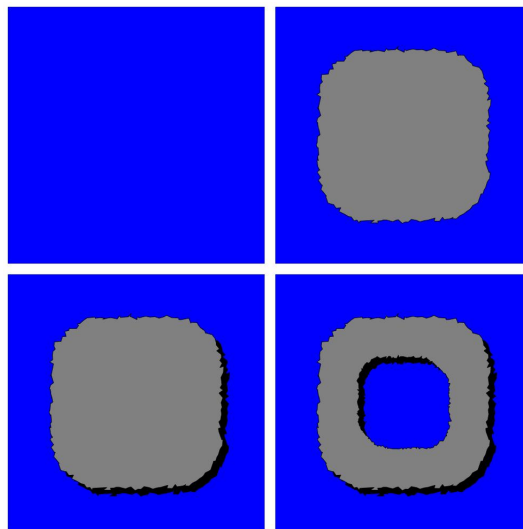


Figure 9. The four contexts used in the experiment. At top left, the ‘blue sky’ backdrop, it should appear totally unrelated to the stimuli proper, thus favoring cutting edges and occluding contours. At top right, the ‘cartoon’ background. It relates to the stimuli (same material) and thus favors dihedral edges. At bottom left, a ‘thick pedestal’. Like the cartoon background, it favors dihedral edges, in addition it visually specifies the illumination flow direction. At bottom right, the ‘window’. Here, the illumination flow direction is specified, but the stimuli appear (as seen through the aperture) in the ‘blue sky’.

These contexts can be combined with manipulations of the outline of the stimulus proper, the region (circular or rectangular as the case may be) containing the linear gradient. We implemented the following cases (figure 10):

- circles and squares (lined up with the gradient) without further embellishment (not illustrated);
- circles with a thin concentric annulus, modulated in gray level so as to suggest the bevel of a concave thick cup (figure 10, left);
- squares with two beveled sides of different gray level, so as to suggest a thick concave cylinder (figure 10, center);
- squares with two skewed beveled sides of different gray level, so as to suggest a thick twisted plate (figure 10, right).

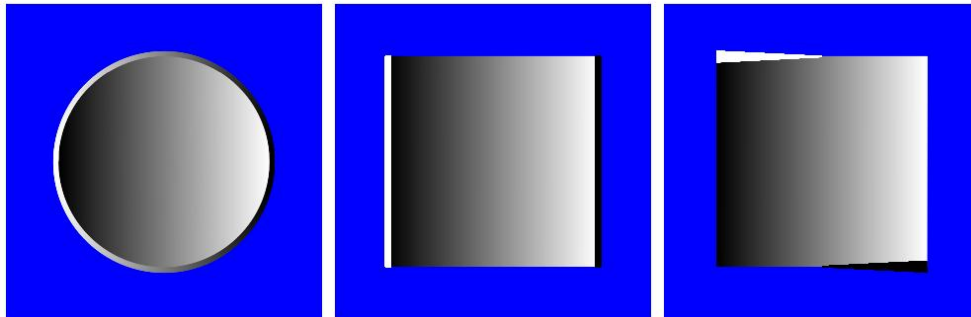


Figure 10. Three types of outline. From left to right, a concave ‘circular thick plate’, a concave ‘cylindrical thick plate’, and a ‘twisted thick plate’. In all cases, the luminance gradient is the same, yet some observers are expected to have the compelling awareness of (from left to right) a concave umbilic, a concave cylinder, and a saddle. The modulated outline yields a ‘minimum context’ of a special kind; it is intimately connected to (part of) the object (not a mere background).

In cases where the illumination direction was ‘visually specified’, we used a ‘drop shadow’.⁽²⁾ In the absence of a drop shadow, the illumination direction remains optically unspecified.

This leads to a large number of combinations [especially taking illumination directions, probe orientations (see below), and stimulus orientations into account], and combining and intermixing all these proved rather time consuming.⁽³⁾

3.2 Principle of the measurement

Although we are primarily interested in the pictorial relief (shape) evoked by the stimuli, it is not trivial to measure this. Remember that we consider surface shape in general, a two degrees of freedom (eg the ratio of the principal curvatures and the orientation of the direction of largest principal curvature) set, much more intricate than the classical convex–concave dichotomy.

Since the absolute distance and the spatial attitude of the pictorial surface are not specified through the shading (although fronto-parallelity is suggested by a circular outline, etc), we avoid the use of depth estimates of pairwise depth comparisons. Instead, we used a configuration of three probe points, one bisecting the segment subtended by the other two in the visual field (figure 11). The task then was to judge whether the center point in visual space is before, inline, or behind the midpoint of the segment in pictorial space. This reveals the sign of curvature in the direction of the segment. Repeating this for various orientations (at

⁽²⁾ On drop shadows, see: http://en.wikipedia.org/wiki/Drop_shadow.

⁽³⁾ For the majority of conditions (types of outlines x types of context), four illumination directions were used, in combination with three probe orientations, and these were repeated three times. For the cylinder cases, two orientations were used, in combination with six illumination directions and four probe orientations, and no repeats were used then. This yields $8 \times (4 \times 3 \times 3)$ and $3 \times (2 \times 6 \times 4) = 432$ trials.

45° increments) of the segment allows us to classify responses as convex elliptical, concave elliptical, convex cylindrical, concave cylindrical, and hyperbolic (saddle shaped).

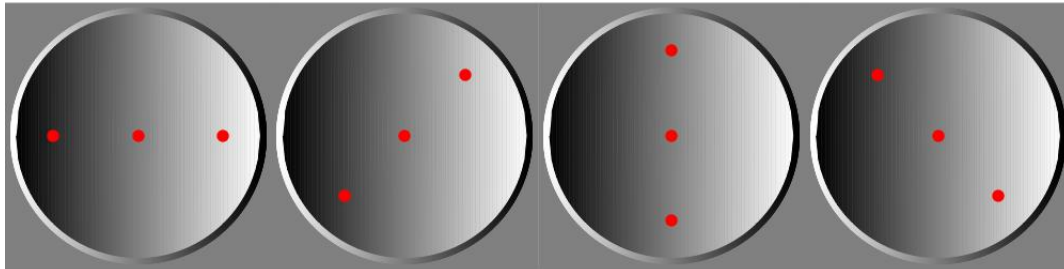


Figure 11. Example of the probe configuration, in this case for the thick concave circular plate. The three collinear points appear in any of four orientations, spaced by 45° increments. The task is to judge whether the point in the middle is in front, inline, or behind the linear segment (in pictorial space) of the two outermost points.

3.3 Methods

Observers were members of the different laboratories who volunteered to participate ($N = 12$). Two of them were authors; the remaining were naive regarding the details of the methods and the goals of the study. In addition, some others were given a few practice trials, but they were not tested formally because they appeared unable to reach satisfactory monocular stereopsis at all.

A session contained 432 presentations (see footnote 3); observers took about an hour to complete it. Each presentation started with an initial period of 2 s in which only the stimulus was presented. Observers were instructed to attend to their awareness of surface shape. In case they succeeded, they had apparently achieved monocular stereopsis. This period was immediately followed by a period of 0.5 s in which the probe configuration was superimposed on the picture. Observers were instructed to decide on the depth relation of the probe dots, that is, to decide on the depth of the center dot in relation to the segment (in pictorial space) defined by the two outermost ones. Then, both probe and stimulus disappeared, and the observers were free to take their time to indicate their response by selecting the appropriate radio button.⁽⁴⁾ The radio buttons were a triple, marked as ‘farther’, ‘closer’, and ‘in line’. Their response time was recorded, although we did not analyze it in detail. (It was not the measure of primary interest since the instructions did not emphasize speed of responding.) It was typically about two seconds (median). After responding, they could trigger the next presentation, and so forth till the conclusion of the session.

The pictures were presented on the LCD screen of a Macintosh notebook, subtending 37° of visual angle. Viewing distance was 50 cm. The stimulus window subtended 15°. The room was darkened, but the observers were fully aware of the screen, and thus the fact that they were looking at pictures rather than a physical scene. Viewing was binocular. The sequence of presentations was randomized over each session.

After conclusion of the session, responses were sorted and combined in subsets pertaining to single pictures, thus only differing by the probe orientation. Thus, we obtained ‘convex’, ‘concave’, and ‘flat’ responses for each of four probe orientations (differing by multiples of 45°), relative to a fiducial orientation for the stimulus. Notice that flat is a distinct category (cylinder axes or asymptotic directions in the case of saddles). Thus, the task cannot be

⁽⁴⁾A radio button, or option button, is a type of graphical user interface element that allows the user to choose only one of a predefined set of options. They were named after the physical buttons used on older car radios to select preset stations—when one of the buttons was pressed, other buttons would pop out, leaving the pressed button the only button in the ‘pushed in’ position.

formulated as a two-alternative one. Since stimuli themselves were presented in various orientations, again differing by multiples of 45° and including the horizontal and vertical, we obtained multiple responses for each relative orientation.

3.4 Analysis of the results

The data allow for the investigation of interobserver consistency, the influence of context, illumination direction, and stimulus shape.

3.4.1 Interobserver variability. The interobserver variability is quite low in the case of the conventional stimulus (see [figure 1](#)) presented in various contexts; the main variation is in an idiosyncratic tendency to fail to reach monocular stereopsis. This is evident from a tendency towards essentially random responses. In [figure 12](#) we plot the responses in barycentric coordinates with all convex, all concave, or all flat responses as vertices; thus the center of the triangle represents equal amounts of convex, concave, and flat responses. Although responses cluster near the convex vertex, there is a clear tendency towards the center of the triangle.

In all other cases the interobserver variability is very striking. These cases are discussed below.

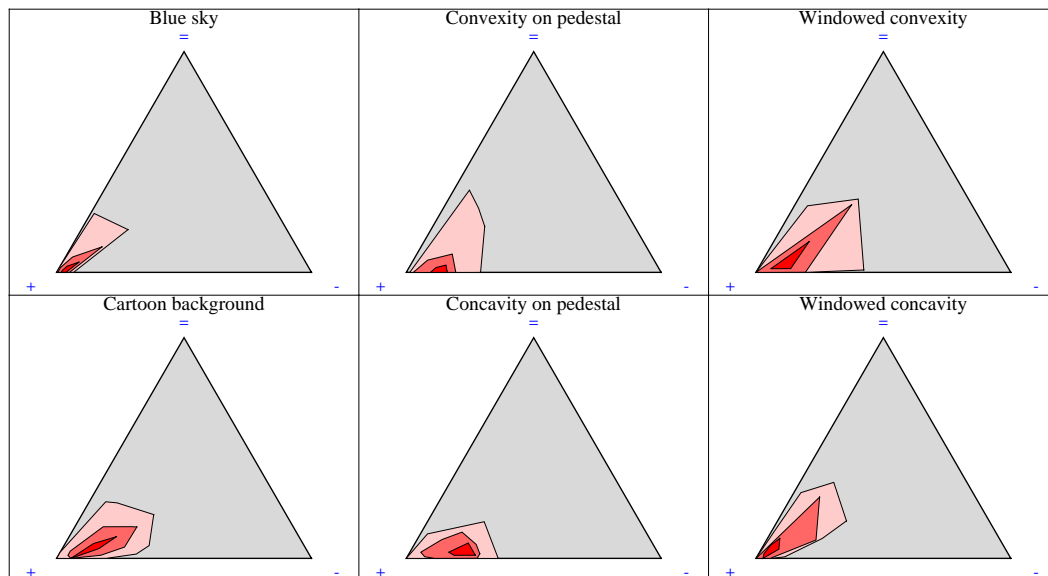


Figure 12. These are results for the conventional stimulus, the circular disk with linear gradient. Since a response is either convex (+), concave (-), or flat (=), it can be conveniently represented as a point in a triangle [using Möbius (1827) barycentric coordinates].⁽⁵⁾ We summarize the distributions through median and quartile regions (using linear interpolation between successive convex hulls). The leftmost column has the blue sky context on top, the cartoon background at bottom. The center column has the pedestal context, the right-hand one the windowed context. In the latter two cases, there is a distinction between nominally convex and concave cases; in the former cases, such a distinction cannot be made.

3.4.2 Influence of context and illumination flow direction. A major finding of this study is implicit in the representation of the data in [figure 12](#). These data are for the conventional stimulus, a circular disk filled with a linear luminance gradient. Irrespective of the nature of the background (blue sky, cartoon background, illuminated pedestal, or aperture), the responses cluster on the convex vertex. Even the visually compelling illumination direction in the case of the illuminated pedestal is ineffective in yielding a convex–concave distinction.

Although there may be statistically significant differences between these cases, they are evidently very minor. The preference for convex is perhaps surprising in view of the fact that all light directions occurred equally frequently.

3.4.3 *Influence of stimulus shape.* We have used both circular and square outlines. Does the shape of the outline make a difference on the pictorial relief? An answer to this question is implicit in the data presented in [figure 13](#). The result is not clear cut, though, because of rather strong idiosyncratic variations.

In the case of the circular disk, there are no significant differences between the responses for the various orientations of the sampling array. In the case of the square, a number of observers respond differently, though. A fraction (10–20%) of the observers experience the linear gradient inside a square as essentially a flat surface (all bars predominantly yellow in [figure 13](#) upper and center panel). Some observers (about half) evidently experience a convex cylinder. This is true even if the cylinder is actually concave, but this need not surprise us: we already know that observers often ignore context.

Thus, many observers experience a cylindrical instead of a spherical surface if the boundary shape is changed from circular to square. This is by no means a fixed rule, though; for roughly equally as many observers the pictorial relief simply flattens out (the yellow in the upper and center panels).

3.4.4 *Influence of boundary modulation.* We consider three cases: the thick concave disk, the thick concave cylinder, and the thick twisted plate. We expected the first two to be similar and the latter case to be qualitatively different from these. We start with the thick concave disk ([figure 14](#)).

Whereas the illuminated pedestal context is ineffective in revealing concavity, the boundary modulation evidently is, at least for many observers. For a few observers, the experience of a concavity is absolutely compelling, though for many it evidently is not.

The case of the thick concave cylinder (bottom panel of [figure 13](#)) is very similar. Half of the observers have the compelling experience of a concave cylinder. The others have a mixed response. (One should remember here that we selected only observers who actually obtained stereopsis.)

The case of the thick twisted plate is especially interesting given the historical context (Alberti's 'exhaustive' list of local surface shapes that remained unchallenged till Gauß's work). See [figure 15](#).

Only one observer (JW) responds in a way that clearly reflects the 'twist'. Apparently this observer experiences a saddle shape. There is a wide variability in the responses of the other observers. Some (three or four) appear to have the awareness of a convex cylinder, alternating with flatness; the others show mixed responses.

4 Conclusions

We may draw a few compelling conclusions from these data.

Appreciable differences exist within the population in the ability to achieve monocular stereopsis. In the interpretation of the data one should take notice of the fact that we did a quick, informal screening before setting observers to the task. We discarded about one in five offhand, those who were apparently unable to achieve stereopsis in any case. Even in the ones we set to the task, the results vary. For some the modulations of the boundary led to compelling experiences of the type one expects from the analysis of the optics, but for many these were apparently ignored.

Context turned out to be ignored by all observers. This is true for the nature of the substrate (blue sky, cartoon background, illuminated thick pedestal) as well as for the visual indication of illuminance flow direction (thick pedestal, window). This is a rather striking

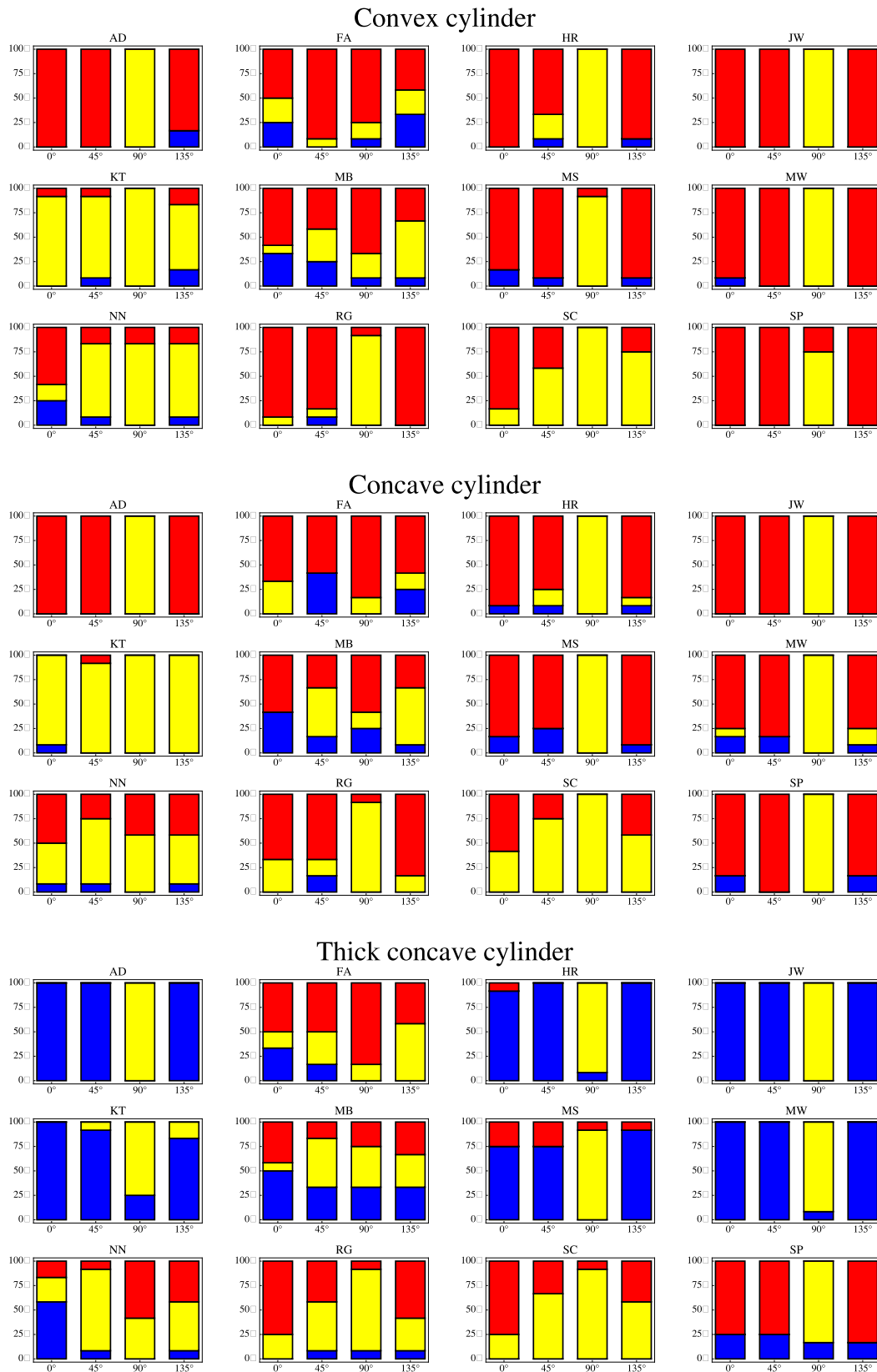


Figure 13. For the case of the cylinder we have split the responses with respect to the probe orientation. The orientation of 90° corresponds to the cylinder axis; the orientation of 0° thus should have the (absolute) largest curvature. The color code is: convex → Red(R), concave → Blue(B), flat → Yellow(Y). Thus, the ‘vertical response’ would be RRYR for the convex cylinder and BBYB for the concave cylinder and the thick concave plate [as, by a fortunate accident, exemplified by the first observer (AD)] in each panel. All observers are (arbitrarily) shown in alphabetical order of their initials.

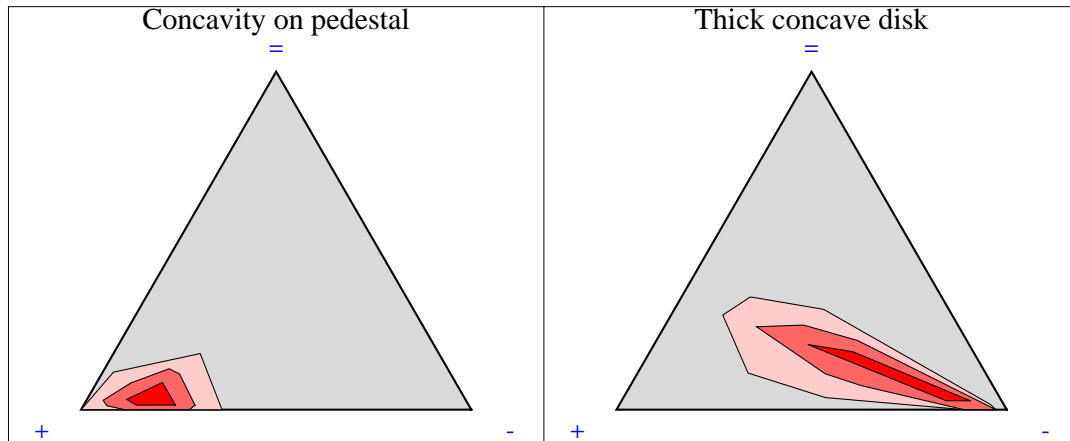


Figure 14. At left, the responses for a concavity in the pedestal context; at right, the responses for the thick concave disk in the pedestal context.

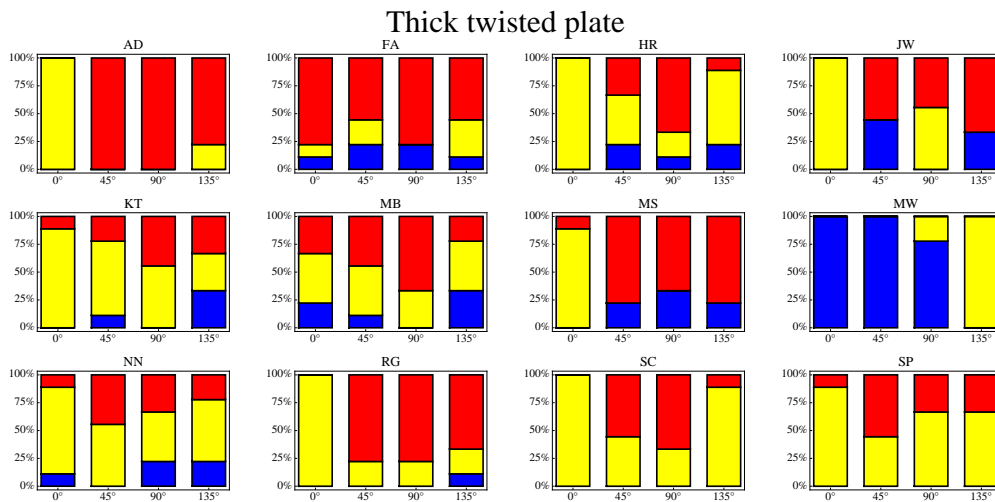


Figure 15. Again, we have split the responses with respect to the probe orientation. The orientations of 0° and 90° correspond to the asymptotic (flat) directions; the orientations of $\pm 45^\circ$ thus should have opposite curvatures. The color code is again: convex \rightarrow Red, concave \rightarrow Blue, flat \rightarrow Yellow. Thus the ‘veridical response’ would be Y(R/B)Y(B/R), although we cannot distinguish between R/B and B/R (unfortunately). The response of the fourth observer in the top row (JW) comes close to this, except for the tendency to respond “convex” for the 90° asymptotic direction.

result, since most demonstrations of context seem convincing from a phenomenological point of view.

The shape of the boundary evidently determines the shape of the pictorial relief. While the circular outline invariably led to spherical (almost always convex) impressions, the square outline led to cylindrical impressions in many observers and led to uncertainty in the responses of the others (a tendency to flatness instead of sphericity) (Cate and Behrmann 2010; Hayakawa et al 1994; Humphrey et al 1996).

Boundary modulations are very effective for a fraction of the observers. In the case of the thick concave disk and the thick concave cylinder, the boundary modulations led to compelling experiences of a spherical concave shell and a concave cylinder for some observers, the boundary information apparently overriding the (strong) tendency to experience convexity. Others largely ignored this information, though traces can often be spotted in their pattern of responses (Cate and Behrmann 2010; Humphrey et al 1996).

Human observers have a strong bias towards convexity as opposed to concavity (Langer and Bühlhoff 2001).

Human observers ignore the possibility of saddle-shaped surfaces. In this case, we found only a single exception, albeit a very significant one. Observers apparently notice the inconsistency of their visual experience, as the response patterns clearly deviate from those for the conventional case of the circular disk with linear gradient. However, they fail to reach a consistent, stable visual awareness of a definitely curved surface.

All this implies that one should take the established literature consensus with a grain of salt. One reason may be that interobserver variability is not generally appreciated (for an exception, see Liu and Todd 2004) and possibly leads to suppression of reports. All this also indicates that theories based on machine vision algorithms of ‘shape from shading’ are not likely to be applicable as models of human monocular stereopsis. It would be easy enough to frame such theories that would beat our observers in the case of our stimulus set. But, of course, the stimuli are extreme abstractions of what occurs in natural images.

The data suggest to us that the human observer tends to assume the boundary to be an occluding one, and the background to be irrelevant. This would apply to many cases in the real world, most cases where a small, convex object is seen against a relatively distant background. This would explain the problematic nature of saddle shapes (boundary a flag edge), and the precarious nature of square outlines (only part of the outline can be an occluding boundary or dihedral edge, the remainder flag edge). (See figure 16.) It also explains the low frequency of concave responses, since these require either dihedral edges or flag edges. From a more general perspective, these assumptions boil down to genericity (general viewpoint, general position of unrelated parts) assumptions.

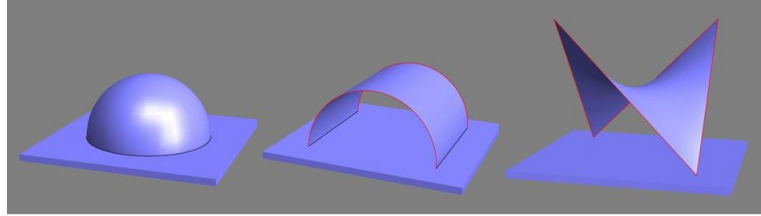


Figure 16. Left: the spherical cap (like the spherical cup) fits a planar support by way of a dihedral edge (the black curve). Center: the convex cylinder may only share two of its generators with the supporting plane (the dihedral edges drawn in black). The remainder of the boundary has to ‘lift out of the plane’ and become flag edge. Right: the square saddle patch does not ‘fit’ the supporting plane at all. All its edges have to be interpreted as flag edges.

In retrospect, one may trace the curious lacunæ in Alberti’s ‘exhaustive’ list of surface shapes to this. The analog in geometry may be the fact that objects of limited size with smooth skins cannot be bounded with overall hyperbolic surfaces (as proven by Hilbert 1901), whereas they can with overall elliptic ones (as with an egg).

That the hyperbolic areas are typically ignored in human understanding of form is also evident from common academic practice. For instance, the sculptor/author Rogers (1969, page 5152) remarks:

Sculpture students modeling from the living model used to be told to take care of the positive forms and leave the negative ones to take care of themselves. They would arrive at the hollow of an armpit or a navel or the channel of the spine by building up the convex shapes that surround them. By working in this way they would come to see more clearly that these hollows are not concave at all but are formed by the grouping of convex forms which are blended together by the unifying effect of the skin.

This is illustrated in [figure 17](#). The awareness of the channel between the two convexities emerges as a secondary effect due to the primary perception of the two convexities. Thus, the fact “that we can see hyperbolic regions” (as was pointed out by reviewers of the present paper) is in no way at odds with the finding that human observers are largely ‘saddle blind’. The saddle regions appear default, as a kind of glue that patches the convexities together.

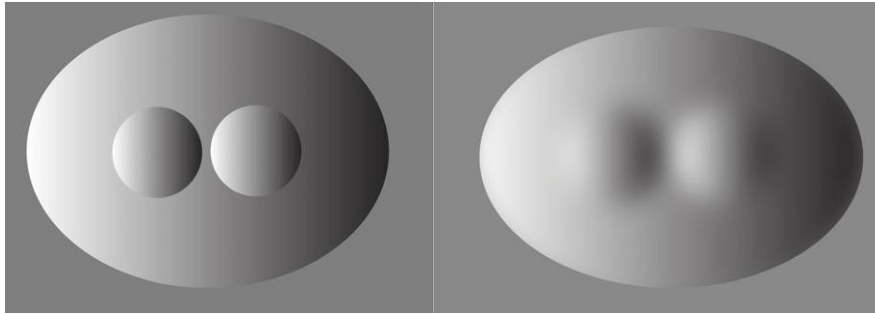


Figure 17. At left, an ovoid with two cups grafted on it; at right, an ovoid with two smoothly curved protuberances. These hills have the same curvature as the cups at their summits, but they are joined to the overall ovoid by smooth fillets. The surface in between the hills then defaults to a smooth saddle or pass. If one sees the hills, one at least implicitly ‘sees the saddle’, but it is by no means necessarily the case that the visual system detects the saddle as an individual surface element. In many cultures the form at left—which is locally all convex—would be considered an apt sculptural rendering of the form at right. This makes sense because the elliptical regions (ovoid, hills) look like ‘things’, whereas the fillet between the hills looks like nothing specific.

In summary, although perhaps surprising from the perspective of shape from shading theory, our findings might be related to the (statistical) fact that many objects are small, bounded by elliptic, convex surfaces, and seen against backgrounds to which they bear no relation.

Acknowledgements. This work was supported by the Methusalem program by the Flemish Government (METH/08/02), awarded to JW. We would like to acknowledge technical support by Frank Amand and administrative support by Stephanie Poot. JK was a visiting fellow at the Flemish Academic Centre for Science and the Arts (VLAC) at the time of the research.

References

- Adams W J, 2007 “A common light-prior for visual search, shape, and reflectance judgments” *Journal of Vision* **7** 1–7 ◀
- Adams W J, 2008 “Frames of reference for the light-from-above prior in visual search and shape judgements” *Cognition* **107** 137–150 doi:10.1016/j.cognition.2007.08.006 ◀
- Adams W J, Graf E W, Ernst M O, 2004 “Experience can change the ‘light-from-above’ prior” *Nature Neuroscience* **7** 1057–1058 doi:10.1038/nn1312 ◀
- Alberti L B, 1435 *De pictura* Latin and Italian at http://www.liberliber.it/biblioteca/a/alberti/de_pictura/html/index.htm, English at <http://www.noteaccess.com/Texts/Alberti/> ◀
- Baxandall M, 1972 *Painting and experience in fifteenth century Italy: a primer on the social history of pictorial style* (New York: Oxford University Press) ◀
- Bouguer P, 1729 *L'Essai d'optique* (Paris: C Jombert) ◀
- Brewster D, 1832 *Letters on natural magic* (London: John Murray) ◀
- Cate A D, Behrmann M, 2010 “Perceiving parts and shapes from concave surfaces” *Attention Perception & Psychophysics* **72** 153–167 doi:10.3758/72.1.153 ◀
- Forsyth D A, Zisserman A, 1991 “Reflections on shading” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13** 671–679 doi:10.1109/34.85657 ◀
- Forsyth D A, Ponce J, 2002 *Computer vision: A modern approach* (Upper Saddle River, NJ: Prentice-Hall) ◀
- Freeman W T, 1994 “The generic viewpoint assumption in a framework for visual perception” *Nature* **368** 542–545 doi:10.1038/368542a0 ◀

-
- Gauß C F, 1827 “Disquisitiones generales circa superficies curvas” *Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores* **6** 99–146 ◀
- Hayakawa H, Nishida S, Wada Y, Kawato M, 1994 “A computational model for shape estimation by integration of shading and edge information” *Neural Networks* **7** 1193–1209 doi:10.1016/0893-6080(94)90002-7 ◀
- Hershberger W A, 1970 “Attached-shadow orientation perceived as depth by chicken reared in an environment illuminated from below” *Journal of Comparative and Physiological Psychology* **73** 407–411 doi:10.1037/h0030223 ◀
- Hess E H, 1950 “Development of the chick’s responses to light and shade cues of depth” *Journal of Comparative and Physiological Psychology* **43** 112–122 doi:10.1037/h0059235 ◀
- Hilbert D, 1901 “Über Flächen von constanter Gausscher Krümmung” *Transactions of the American Mathematical Society* **2** 87–99 ◀
- Hill H, Bruce V, 1994 “A comparison between the hollow-face and ‘hollow-potato’ illusions” *Perception* **23** 1335–1337 doi:10.1068/p231335 ◀
- Horn B K P, Brooks M J, 1989 “*Shape from shading*” (Cambridge, MA: MIT Press) ◀
- Humphrey G K, Symons L A, Herbert A M, Goodale M A, 1996 “A neurological dissociation between shape from shading and shape from edges” *Behavioural Brain Research* **76** 117–125 doi:10.1016/0166-4328(95)00190-5 ◀
- Kleffner D A, Ramachandran V S, 1992 “On the perception of shape from shading” *Perception & Psychophysics* **52** 18–36 ◀
- Koenderink J J, van Doorn A J, 1983 “Geometrical modes as a general method to treat diffuse interreflections in radiometry” *Journal of the Optical Society of America* **73** 843–850 doi:10.1364/JOSA.73.000843 ◀
- Koenderink J J, van Doorn A J, Kappers A M L, 1994 “On so-called paradoxical monocular stereoscopy” *Perception* **23** 583–594 doi:10.1068/p230583 ◀
- Koenderink J J, van Doorn A J, Christou C, Lappin J S, 1996 “Perturbation study of shading in pictures” *Perception* **25** 1009–1026 doi:10.1068/p251009 ◀
- Lambert J H, 1760 *Photometria, sive, De mensura et gradibus luminis, colorum et umbrae* (Augsburg, Germany: V E Klett) ◀
- Langer M S, Bühlhoff H H, 2001 “A prior for global convexity in local shape-from-shading” *Perception* **30** 403–410 doi:10.1068/p3178 ◀
- Lipschutz M, 1969 *Differential geometry* (New York: McGraw-Hill) ◀
- Liu B X, Todd J T, 2004 “Perceptual biases in the interpretation of 3D shape-from-shading” *Vision Research* **44** 2135–2145 doi:10.1016/j.visres.2004.03.024 ◀
- Luckiesh M, 1916 *Light and shade and their applications* (New York: D Van Nostrand Company) ◀
- Mamassian P, Goutcher R, 2001 “Prior knowledge on the illumination position” *Cognition* **81** 1–9 doi:10.1016/S0010-0277(01)00116-0 ◀
- Metzger W, 1975 *Gesetze des Sehens* (Frankfurt a M, Germany: Verlag Waldemar Kramer) ◀
- Minnaert M, 1993 *Light and color in the outdoors* (New York: Springer-Verlag) ◀
- Möbius A, 1827 *Der baryzentrische Kalkül* (Leipzig, Germany: Gesammelte Werke) ◀
- Palmer S E, 1999 *Vision science: Photons to phenomenology* (Cambridge, MA: Bradford Books/MIT Press) ◀
- Ramachandran V S, 1988a “Perceiving shape from shading” *Scientific American* **259** 76–83 doi:10.1038/scientificamerican0888-76 ◀
- Ramachandran V S, 1988b “Perception of shape from shading” *Nature* **331** 163–166 doi:10.1038/331163a0 ◀
- Riedel R, 1984 *Biology of knowledge: The evolutionary basis of reason* (Chichester, U.K.: John Wiley & Sons) ◀
- Rittenhouse D, 1786 “Explanation of an optical deception” *Transactions of the American Philosophical Society* **2** 37–42 doi:10.2307/1005164 ◀
- Rogers L R, 1969 *Sculpture. The appreciation of the arts/2* (London: Oxford University Press) ◀
- Süffert F, 1932 “Phänomene visueller Anpassung” *Zeitschrift für Morphologie und Ökologie der Tiere* **26** 147–316 doi:10.1007/BF00446393 ◀
- Sun J, Perona P, 1998 “Where is the sun?” *Nature Neuroscience* **1** 183–184 doi:10.1038/630 ◀
- Thayer G H, 1909 *Concealing coloration in the animal kingdom: An exposition on the laws of disguise through color and pattern* (New York: The Macmillan Co.) ◀
- Turhan M, 1935 “Über räumliche Wirkungen von Helligkeitsgefällen” *Psychologische Forschung* **21** 1–49 ◀
- van Diggelen J, 1951 “A photometric investigation of the slopes and heights of the ranges of hills in the maria of the moon” *Bulletin of the Astronomical Institutes of the Netherlands* **11** 283–289 ◀

Zhang R, Tsai P S, Creyer J E, Shah M, 1999 "Shape from shading: a survey" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21** 690–706 [doi:10.1109/34.784284](https://doi.org/10.1109/34.784284) ◀