UC Irvine

UC Irvine Previously Published Works

Title

Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse perspective.

Permalink

<https://escholarship.org/uc/item/07237885>

Journal

Perception, 8(2)

ISSN

0301-0066

Authors

Yellott, JI
Kaiwi, JL

Publication Date

1979

DOI

10.1068/p080135

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse perspective

John I Yellott, Jr, Jerry L Kaiwi

Cognitive Science Group, School of Social Sciences, University of California at Irvine, Irvine, California 92717, USA

Received 20 October 1977, in revised form 18 October 1978

Abstract. Inside-out relief masks of faces can be depth-inverted (i.e. seen in reverse perspective) during close-up binocular viewing. If a random-dot stereogram is projected onto such a mask, stereopsis can be achieved for the stereogram, and its depth planes are correctly seen while the mask itself, including the region covered by the stereogram, is simultaneously perceived as depth-inverted. This demonstration shows that binocular depth inversion cannot be explained by a complete loss of stereoscopic information (e.g. through monocular suppression), or by a process analogous to pseudoscopic viewing whereby retinal disparities are incorporated into perception, but with their signs uniformly reversed.

1 Introduction

Reversible perspective is probably most often thought of in connection with ambiguous pictures like the Necker cube (figure 1a), but it is also well-known that depth reversals can sometimes be experienced with solid objects, and with far more dramatic perceptual consequences. [Gregory (1970) describes these in detail; Helmholtz (1925) reviews references back to 1712, and Hochberg (1972) and Robinson (1972) review the modern literature, which is surprisingly thin.] In this note we use 'depth inversion' to refer specifically to this second kind of reversible perspective, in which real objects (rather than depicted ones) are perceived, from the standpoint of depth, as inside-out, so that their concave surfaces appear convex and vice versa. Figure 1b illustrates a convenient demonstration due to Mach (1866, 1914), which requires only a bent note card, monocular viewing, and a little practice. Eventually an observer at O can perceive the bent edge as depressed instead of raised, and maintain this inverted percept for extended periods of time⁽¹⁾.

With binocular viewing, however, this same demonstration does not work at all well: even after extensive practice, binocular depth inversions of Mach's card (and most other objects) are very rare and, when they do occur, invariably brief, strenuous, and 'curiously unreal' in appearance (Gregory 1970). This difference can readily be attributed to the fact that binocular parallax provides depth information which is unambiguous in a way that monocular information is not, and previous accounts of depth inversion seem to have left the matter there, implicitly assuming that binocular inversion does not need to be explained because it effectively does not occur.

However, it is not hard to demonstrate that under appropriate conditions binocular depth inversion can occur almost as easily as monocular depth inversion, and lead to

⁽¹⁾ In trying this demonstration for the first time one should be aware that depth inversion is a learned illusion which proceeds in distinct stages corresponding to the various cues to depth. Starting with a stable monocular inversion achieved with the head stationary and no shadows, one finds that as each additional depth cue becomes available the illusion is initially disrupted and then, with practice, restored to stability—as though the visual system were becoming more adept at reinterpreting that cue in a manner consistent with the inversion. Binocular parallax plays an anomalous role in this learning process because, as emphasized in the text, no amount of practice makes it easy to obtain a stable binocular inversion of such depth-neutral objects as Mach's card.

percepts that seem equally natural and stable, even though the observer has normal stereoscopic ability and distances are such that binocular parallax should be an effective cue. The two simplest explanations for this apparent anomaly would seem to be either that one eye's information is somehow lost in transit, so that at the final stages of visual processing only monocular (hence readily invertible) information remains, or else that binocular depth inversion occurs after stereopsis, so that left and right images are first combined into a single cyclopean percept, which is then depth-inverted as a unit—a kind of mental pseudoscopy, since the effect would be equivalent to interchanging the two eyes.

However, neither of these explanations will work, for the following reason. If the anaglyph version of one of Julesz's (1971) random-dot stereograms is projected onto a concave surface, and viewed binocularly through red-green glasses, *correct* depth can be seen in the stereogram even though the surface itself is simultaneously perceived as depth-inverted—so that the stereogram, including its floating planes (visible only through stereopsis), appears convex.

This demonstration of depth inversion despite stereopsis is the substance of the present paper. However, before giving details, we detour briefly in order to explain the motivation for the experiment more carefully. (Readers who already see the point should turn directly to the methods section.) First we review the perceptual consequences of monocular depth inversion and their usual interpretation; second we consider the implications of that interpretation for binocular depth inversion, and finally we describe techniques for obtaining binocular inversion and possible interpretations of that phenomenon—these last being the direct motivation for our experiment.

Monocular depth inversion has three major perceptual consequences. (i) Changes in apparent size, and consequently shape, occur along the lines of Emmert's law: parts appear larger or smaller according to whether inversion makes them seem farther or nearer, so that an inverted cube, for example, assumes the shape of a truncated pyramid. (ii) Depth-inverted objects always appear to rotate when the observer moves his head (Shopland and Gregory 1964)—an inverted inside-out face, for example, seems to rotate in such a way as to keep the observer under constant surveillance. (iii) When shadows are present, there are changes in apparent reflectance: if the light in figure 1b comes from the left, so that the right side is in shadow, after

Figure 1. (a) The Necker cube. (b) Mach's card. (c) Cross-sectional diagram illustrating depth inversion with a skeleton cube. Solid figure cabd represents the real cube; ac^*d^*b represents an inverted object that would project the same retinal image. (d) Top-view diagram showing why monocularly depth-inverted objects appear to rotate during head movements. The left circle represents an eye viewing two real points, x and y . Depth inversion entails projecting x backwards, beyond y , to position x_1^* . When the eye moves rightwards, as indicated by the arrow, depth inversion can be reconciled with motion parallax by assuming that x^* has moved from x_1^* to x_2^* . (e) Depth inversion of an inside-out mask. O denotes the observer, L a light transilluminating the mask. The inverted surface indicated by the dotted line is not intended to be geometrically accurate in the same way as in part (c).

inversion “the left side appears to become much brighter and the right side much darker. Light and shade appear as if painted upon it” (Mach 1866).

To explain monocular inversion and its consequences the obvious starting point is the intrinsic ambiguity of the monocular retinal image: since such an image can always be the projection of an infinite number of possible three-dimensional objects, including inverted ones, the visual system’s selection of a single ‘object hypothesis’ [to borrow Gregory’s (1970) terminology] is necessarily arbitrary, and so it is understandable that under deliberately ambiguous conditions a mistake is sometimes made, and the system adopts an incorrect hypothesis. When this happens the perceptual significance of sensory information is apparently automatically reinterpreted—so as to be consistent with the false object hypothesis—hence the dramatic consequences of depth inversion. Thus when the three-dimensional cube in figure 1c is inverted, so that line *cd* is taken to be more distant (say, at c^*d^*) than line *ab*, the relative sizes of the retinal images of lines *cd* and *ab* can be reconciled with the inversion by assuming that the object is actually a truncated pyramid, and this is what one perceives. Similarly, the motion-parallax information produced by head movements can be made consistent with an inverted-depth hypothesis by introducing the auxiliary assumption that the object is rotating (as shown in figure 1d), and the depth cues provided by shadows can be reinterpreted by attributing them to reflectance properties of the object, rather than its illumination. The fact that all these monocular cues are logically susceptible to reinterpretations consistent with an inverted-object hypothesis allows monocular depth inversion to occur with no loss of sensory information: all of the depth information provided by the retinal image is still present in the inverted percept, and only the perceptual significance of that information needs to be changed.

These arguments make monocular depth inversion understandable, and at the same time suggest that binocular depth inversion should be either impossible, or at best a very unnatural experience, quite unlike the monocular case—provided the observer has normal stereoscopic vision, and the object is near enough to make retinal disparity an effective cue. This conclusion follows from the fact that retinal-disparity information is not susceptible to any obvious reinterpretation consistent with an inverted-object hypothesis. To make this point explicit, imagine that the two circles in figure 1d represent the left and right eyes, simultaneously viewing points *x* and *y*. The image in either eye alone can be inverted by imagining *x* projected backwards to x_1^* or x_2^* . However, none of the inversions possible for the left eye can be reconciled with any of those possible for the right eye. Consequently, it would seem that if binocular depth inversion can occur at all, it must be at the expense of a loss or distortion of retinal-disparity information.

In fact, everyone agrees that binocular viewing normally does make depth inversion much more difficult: typically, stable inversions achieved monocularly are disrupted by opening the other eye, and with such depth-neutral objects as Mach’s card and Gregory’s skeleton cube this cannot be overcome even with extensive practice. However, this is not always true: other objects which are not depth-neutral can be quite readily depth-inverted despite close-up binocular viewing, the prime example being inside-out reliefs of human faces. (This seems to be a fairly well-known piece of visual folklore, but we have not found any explicit reference to it in the scientific literature.) We have experimented with life-size, inside-out relief masks, and find that with appropriate lighting these objects provide stable depth inversions despite binocular viewing at distances on the order of 1 m. An ordinary Halloween mask made of rigid translucent plastic works well for demonstrations. If the mask is mounted in a sheet of cardboard and viewed transilluminated (as shown in figure 1e) in a dark room, monocular depth inversion occurs very readily and, with a little

practice, binocular inversion is almost as easy. At first, opening the other eye tends to disrupt a stable monocular inversion, just as head movements tend initially to disrupt inversions achieved with the head stationary. However, under the conditions described, where the lighting casts no shadows to suggest true depth, observers typically require only a few minutes of practice (starting perhaps at a viewing distance of 2 m and gradually working closer) to achieve binocular inversions that are not disrupted by head movements. The appearance of a binocularly depth-inverted inside-out mask is quite stable and natural, unlike that of a skeleton cube or Mach's card.

Now from one standpoint there is nothing remarkable in the fact that inside-out faces can be binocularly depth-inverted quite easily whereas cubes and bent cards cannot: this simply reflects the role of past experience in the object-hypothesis-selection process. It has been known since Wheatstone's original observations with the pseudoscope that stereopsis can fail when the depth implied by binocular parallax would result in an unlikely apparition, such as a human face turned inside-out, or for that matter, a landscape. Accordingly, binocular depth inversion could simply be regarded as a failure of stereopsis, akin to those commonly experienced with difficult stereograms.

However, there still remains the question of what happens to the retinal-disparity information that should, if properly incorporated, prevent binocular depth inversion? Put another way, how can two monocular views be combined into a single visual experience when that experience cannot be simultaneously reconciled with both views? Two possibilities suggest themselves immediately. The first, and simplest, is that during binocular inversion one eye's information is lost, or suppressed, as in binocular rivalry, so that the ultimate perception is based entirely on monocular, hence consistently depth-invertible, information. (This undoubtedly does happen with observers who have defective binocular vision. However, the question is whether it is always necessary.) The second is that binocular depth inversion entails a kind

Figure 2. Inside-out mask with superimposed random-dot stereogram. Mask dimensions: height, 29 cm; width, 16 cm; maximum depth in region covered by stereogram, 8 cm.

of mental pseudoscopy, in which the depth given by binocular parallax is altogether reversed by a process that simply changes the signs of all the disparities—thus in effect interchanging the signals from the left and right eyes. This second possibility would provide a simple means of actually incorporating retinal-disparity information into the inverted percept, and thereby allowing it to play a role analogous to the monocular depth cues, which as we have seen are not suppressed, but simply reinterpreted.

To test these two possibilities, we have experimented with random-dot stereograms projected onto the surface of inside-out face masks, as illustrated in figure 2. We reasoned that if binocular depth inversion entails a complete loss of one eye's information, then it should not be possible to see depth at all in such a stereogram when the surface on which it is superimposed is simultaneously perceived as depth-inverted. On the other hand if binocular inversion involves uniformly reversing the signs of the retinal disparities (mental pseudoscopy), then one might expect to see depth in the stereogram during depth inversion, but the apparent direction of this depth should be uniformly opposite to that implied by the actual physical disparities.

As it turns out, neither of these predictions is correct. The actual result is tantalizingly clear-cut, but requires a more subtle explanation. The discussion section considers some possibilities suggested by the work of Sperling (1970) and Kaufman (1964, 1974).

2 Method

The stereogram used here was the anaglyph version of Julesz's (1971) figure 8.1-2, which depicts three concentric square planes of stereoscopic depth, stacked one on top of the other. This was viewed through red and green glasses oriented so that the smallest and most central plane should appear stereoscopically nearest the observer. (To simplify the description of results we ignore the intermediate depth plane in the figures and elsewhere, and concentrate simply on the small central square and the background plane. This has no special significance: the intermediate plane behaved exactly like the other two.)

This stereogram was projected onto the surface of the inside-out mask shown in figure 2, in the position indicated, and viewed in an otherwise dark room from a distance of roughly 1.2 m. (This particular mask was used because its derby hat provided a smoothly concave surface large enough to contain the stereogram.) There were four observers (the authors, plus two colleagues), all familiar with random-dot stereograms and depth-inversion phenomena, and all having normal stereoscopic ability. Their task was to report whether depth could be seen in the stereogram while the mask was simultaneously perceived as inverted, and to describe the appearance of the depth planes, both with the head stationary and during head movements.

We should add that none of the details of this particular arrangement seem especially critical: we have performed the same experiment with naive observers, and with different stereograms, and the results are always the same as long as the observer can see depth in random-dot stereograms to begin with, and is willing to spend a few minutes practicing depth inversion. It is also possible to do away with the inside-out mask, and perform the experiment directly with a printed anaglyph cut from Julesz's book. In this case the stereogram itself can be curved into a gently concave surface, and then depth-inverted—though not without difficulty. This leads to the same result, but takes more patience, and not everyone can succeed in a reasonable time. Interestingly, in our experience the observers who are best at seeing depth in difficult stereograms are also best at binocular depth inversion.

3 Results

Perceptually the results of this experiment are quite clear-cut and automatic, as though the visual system already had a rule for dealing with such peculiar stimuli. Every observer reported that depth could be seen in the random-dot stereogram while the mask was simultaneously perceived as depth-inverted, and the direction of this depth was always that implied by the physical disparities—never the reverse. Figure 3a illustrates the general appearance of things during inversion. The fact that depth inversion and stereopsis are occurring simultaneously for the same surface is shown by two critical signs. First (as shown in the left panel of figure 3a) the surface of the stereoscopically perceived central square appears convex, sharing the illusory convexity of the inverted mask. Second, when the observer moves his head sideways (as shown in the right panel), maintaining depth inversion, the entire mask, including the stereogram and its in-depth central square, all rotate together as one rigid body, with the central square moving as though attached to the background by an invisible pedestal. Both effects are distinctly different from those experienced when the surface beneath the stereogram is not perceived as depth-inverted. In that case (as illustrated in figure 3b), the central square of the stereogram appears concave (left panel), and during head movements (right panel) the background remains stationary, and only the central square moves. This effect is well-known (Julesz 1971) and can readily be accounted for by projection theories of stereopsis, as shown in the figure.

It is important to note that figure 3a has a different logical status than figure 1c (which represents the perceptual geometry of monocular depth inversion) and figure 3b (which is a typical projection diagram designed to account for ordinary stereopsis). The latter two diagrams both depict consistent relationships between physical geometry and visual experience. In the case of monocular inversion it is natural to assume that the inverted 'object' lies along the same projection lines as the real one, so that both would project the same retinal image. In stereoscopic projection diagrams like figure 3b it is natural to assume that stereoscopically visible objects lie at the intersections of lines projected out from their left and right retinal images. However, in the case of binocular depth inversion it is not easy to see how

Figure 3. (a) Apparent rotation of a perceptually inverted mask and stereogram during rightwards head movement. Solid line ab is the actual mask surface with the stereogram background at cd and the central square of nasally disparate dots at ef . Dashed lines a^*b^* , c^*d^* , and e^*f^* represent the inverted mask surface, stereogram background, and stereoscopically perceived central square, respectively. (b) Appearance of random-dot stereogram when inside-out mask is not perceptually inverted. Solid line ab indicates mask surface, line cd the background (zero disparity) portion of the stereogram, $e-r$, $e-l$ and $f-r$, $f-l$ are representative anaglyph dots at the edges of the central square, dashed line e^*f^* is the stereoscopic appearance of the square. During head movement (right panel) only this square appears to move.

one can draw a diagram consistent with both physical and perceptual geometry. In fact, this is precisely the puzzle that binocular inversion poses for projection theories of stereopsis. Clearly we cannot simply ignore either retinal image, because every observer sees the depth planes of the stereogram. Nor can we draw a purely pseudoscopic diagram, in which objects are located at the positions they would have in a standard projection diagram if the left and right images were interchanged, because in that case the central square of the stereogram would have to appear behind the background, rather than in front as actually happens. Thus figure 3a is only intended to convey the unambiguous ordinal properties of the visual experience produced by this arrangement, rather than account for it in geometrical terms.

4 Discussion

Since the two easiest explanations for binocular depth inversion appear to be ruled out by this demonstration, something more complex is required. One possibility is that our result might be explained along the same lines as Kaufman's (1964, 1974) demonstration that under binocular-rivalry conditions, stereopsis can occur despite an apparently complete monocular suppression. In that demonstration, stereopsis fails for the high spatial frequencies of a stereogram, which are seen as alternating between one monocular view and the other, but succeeds for the low frequencies, which lend a sense of stereoscopic depth to coarse regions of the cyclopean picture. Sperling's (1970) model of stereopsis accounts for this by postulating two separate disparity-calculating mechanisms, one tuned to low-frequency channels, which computes a global sense of depth, the other tuned to high-frequency channels, so as to determine the disparities of fine details. Stereopsis can fail on either channel if the differences between left-eye and right-eye inputs are too great to be reconcilable, and in that case the output of the mechanism corresponds simply to one monocular input or the other—in patches, presumably, as happens perceptually in binocular rivalry.

To apply this idea to the present experiment one could imagine that the high-frequency mechanism is responsible for the perception of depth in the random-dot stereogram, while the low-frequency mechanism is responsible for the inverted appearance of the mask as a whole. Our case then could be interpreted as the reverse of Kaufman's, in that here stereopsis succeeds for the high-frequency mechanism, but fails for the low-frequency mechanism, which consequently yields only monocular output ambiguous with respect to depth, posing no barrier to global depth inversion. (This possibility and the demonstration that refutes it were suggested to us by D I A MacLeod.)

The difficulty with this explanation is that it seems to imply that the global appearance of a binocularly inverted mask should correspond to its monocularly inverted appearance in one eye or the other. However, it is easy to demonstrate that this does not happen during binocular inversion. For example, if the observer is eyeball-to-eyeball with an inside-out mask, then during a monocular inversion the nose of the inverted face seems to point towards whichever eye is seeing. During binocular inversion, however, it points straight ahead. Moreover there is no suggestion of the kind of alternations between monocular views characteristic of binocular rivalry.

However, it might be possible to explain our results in terms of something like Sperling's (1970) model if we assume that the outputs of the various spatially tuned mechanisms can each be independently inverted in a pseudoscopic fashion, at a point before they are recombined into a final cyclopean image. We have no clear idea exactly how such a process might work (in particular, what form recombination would take), but nothing in our results seems immediately to rule out an interpretation along these lines. One difficulty is that it is not clear whether purely stereoscopic

depth can be inverted at all. For example, we know of no demonstration that depth inversion can be achieved with random-dot stereograms, and we have never succeeded at this ourselves. [Hochberg (1964, cited in Julesz 1971), has shown that reversible perspective can be seen in a cyclopean Necker cube, but this is a different demonstration from the one we have in mind because the cyclopean lines depicting this cube lie in the same stereoscopic plane, and so these are reversals of pictorial depth.] However, this may be only a perceptual learning problem: all the stereograms we have encountered are rather neutral with respect to depth, in the same sense as Mach's card. We are not aware of any published accounts of experiments with random-dot stereograms depicting inside-out faces, but this would certainly seem the best way to try for a purely stereoscopic inversion.

A referee has suggested that our result might be alternatively interpreted in terms of 'fusion without stereopsis', the high-frequency information from the stereogram being fused and perceived stereoscopically in the normal way, while the low-frequency information from the mask would be fused but its depth implications simply not registered. Presumably in this case the low-frequency 'points' would be stereoscopically located at an indeterminate distance along lines projecting from a cyclopean nodal point midway between the observer's eyes. The critical difference between this interpretation and the one suggested in the text is that ours implies a deterministic relationship between the apparent magnitude of inverted depth and the physical disparities, while the referee's does not. If it should prove impossible to invert random-dot stereograms of inside-out faces, his interpretation would seem to be more sensible than ours.

Acknowledgements. We thank T Batey, T Boda, M Braunstein, and M Vernoy for assistance with the experiment, and A Ahumada, W Estes, B Julesz, D MacLeod, and L Narens for helpful comments on an earlier draft of the paper.

References

- Gregory R L, 1970 *The Intelligent Eye* (New York: McGraw-Hill)
- Helmholtz H von, 1925 *Treatise on Physiological Optics*, volume 3, translated from the 3rd German edition, Ed. J P C Southall (Optical Society of America; New York: Dover, republished 1962)
- Hochberg J, 1972 "Perception: space and movement" in *Woodworth and Schlosberg's Experimental Psychology* volume 1, Eds J W Kling, L A Riggs (New York: Holt, Rinehart and Winston)
- Julesz B, 1971 *Foundations of Cyclopean Perception* (Chicago: University of Chicago Press)
- Kaufman L, 1964 "On the nature of binocular disparity" *American Journal of Psychology* 77 398-401
- Kaufman L, 1974 *Sight and Mind* (New York: Oxford University Press)
- Mach E, 1866 "On the physiological effect of spatially distributed light stimuli, III" translated and reprinted by F Ratliff (1965) in *Mach Bands: Quantitative Studies on Neural Networks in the Retina* (San Francisco: Holden-Day)
- Mach E, 1914 *The Analysis of Sensations and the Relation of the Physical to the Psychical* (New York: Dover, republished 1959)
- Robinson J O, 1972 *The Psychology of Visual Illusion* (London: Hutchinson)
- Shopland G, Gregory R L, 1964 "The effect of touch on a visually ambiguous three-dimensional figure" *Quarterly Journal of Experimental Psychology* 16 66-70
- Sperling G, 1970 "Binocular vision: a physical and a neural theory" *American Journal of Psychology* 83 461-534