Review

Three-dimensional displays and stereo vision

Gerald Westheimer*

Division of Neurobiology, University of California, Berkeley, CA 94720-3200, USA

Procedures for three-dimensional image reconstruction that are based on the optical and neural apparatus of human stereoscopic vision have to be designed to work in conjunction with it. The principal methods of implementing stereo displays are described. Properties of the human visual system are outlined as they relate to depth discrimination capabilities and achieving optimal performance in stereo tasks. The concept of depth rendition is introduced to define the change in the parameters of three-dimensional configurations for cases in which the physical disposition of the stereo camera with respect to the viewed object differs from that of the observer's eyes.

Keywords: human stereoscopic vision; binocular disparity; three-dimensional cameras; depth rendition

1. INTRODUCTION: THE THIRD DIMENSION OF VISUAL SPACE

There are three spatial dimensions in the world of real objects but only two in the standard modes of its capture and depiction—photographic film, TV camera, the printed page, movie screen or video monitor. Condensing visual information from three dimensions into two can be reasonably successful. Painters have for centuries given clues to the three-dimensional disposition of objects on paper or canvas [1]: the further a given figure the smaller it appears; closer objects partially obscure those behind; light and shadow give hints about fore and aft position. Active participation on the part of the observer or camera can help: when one target is far and another close, focusing on one will blur the other; activity of the eye's accommodation can add to a target's sense of nearness, and so can change in relative position when the head is moved. These clues to depth—often called monocular because they are available even when one eye is closed—require, however, some prior knowledge or good guesses: that people and trees and houses are all approximately the same sizes, that roads and railroad tracks remain constant in width, that the sun shines from above. The ubiquity of visual representations in modern life using only two dimensions bears testimony to the effectiveness of these stratagems. But being based on supposition and not physical certainty, they could, and do, fail in unusual circumstances or novel observation.

As technology comes to grips with genuine three-dimensional displays, it recapitulates the more secure process for estimating the third visual dimension that evolution perfected by placing the two eyes forward in the head instead of laterally, enabling a process of triangulation based on two vantage points separated by a few inches horizontally. A price was paid: a panoramic representation of the world was sacrificed to a restricted one, albeit with overlapping visual fields of the two eyes. This now means that the same object is imaged on two separate two-dimensional surfaces and a correspondence between them has to be established. In an innate anatomical arrangement, neural paths from points on the two retinas with the same spatial signature converge on single cells and, when stimulated together, the observer will see single points in fixed locations. In particular, the foveal centres are corresponding points. In addition, humans have the capacity to move their two eyes not only in parallel but also relative to each other, to converge the foveal lines of sight to various positions along the *z*-axis, allowing objects in different planes to be brought into register.

Hence, one way of surveying the layout of objects in their three-dimensionality is to keep track of the eye movements, laterally, vertically and convergent, needed to image them bifoveally. The readout of the oculomotor stance is, however, not very sensitive and is not used as the main source of depth information. To the contrary, there is little if any change in the apparent disposition of objects in the world when the eyes are moved or their convergence changed by a prism in front of one eye. Instead, in a subtle refinement that is the essence of what we now call stereoscopic vision, the small differences between the right and left eyes' images are analysed for relative object distances.

The two-dimensional sketch in figure 1 might represent a book seen opened either towards or away from the viewer. Bare of monocular cues, the configuration remains ambiguous. But viewing it simultaneously from two vantage points can resolve the ambiguity.

For the human stereoscopic mechanism, the so-called bipolar coordinate system applies in which the three-dimensional location of a picture element is specified by the angle it makes at the two eyes, the binocular parallax for its distance coordinate, as well as two angular coordinates: azimuth, for how far lateral with respect to the
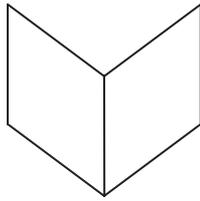
*gwestheimer@berkeley.edu

Figure 1. This sketch, devoid of monocular cues to depth, can represent a book or a folder open towards or away from the viewer. The ambiguity is resolved by getting a simultaneous look from two separate vantage points.

straight-ahead, and elevation, for how much the plane containing the two eyes has to be tilted around the line joining the two eyes. Translation from x,y,z Cartesian coordinates is easily accomplished. Figure 2 shows how the difference in the z-coordinates between the front and back levels of the configuration in figure 1 is converted into differences in relative placement in the images received by the right and left retinas. When edge B is nearer than edges A and C, the angle made by B at the two eyes is larger than the corresponding angles for A and C. It is this angular difference, called binocular disparity, which constitutes the stimulus for the stereoscopic detection of depth.

2. OPTIMAL CONDITIONS FOR STEREOPSIS

Much has been learned since Wheatstone invented the stereoscope in 1838 and was thus enabled to examine the feature combination in the two retinas that is involved in stereopsis. Here, some of the conditions that optimize the detection of depth differences between adjoining feature elements will be surveyed. To visualize what is involved, place two knitting needles at arm's length and try to make their tips touch. Doing this with just one and then with both eyes open makes a convincing case for the value of stereopsis. For a good observer in a practised task, the disparity threshold can be as small as a few seconds of arc, which translates into a depth difference of less than a tenth of a millimetre, or the height of the profile on a coin, at arm's length.

Human depth discrimination is best when there are several objects in different three-dimensional locations. The task then requires the examination of the right- and left-eyed images, the pairing of corresponding elements and the determination of their relative disparity. The virtue of the operation in the domain of disparities is that values, being differences, are independent of the state of convergence of the eyes.

The most immediate questions relate to the temporal and spatial feature disposition. In the time domain, the arrival of input to the two retinas should be synchronized to better than 50 ms. Alternating presentation to the two eyes at a rate of 15 Hz already leads to a performance decrement [2]. Similarly, when the depths of two neighbouring features are compared, their binocular onset should be synchronized to better than 50 ms [3]. Further, the untrained observer does poorly when targets are exposed for less than a few hundred milliseconds [4], although crude stereoscopic depth can be detected with a sub-millisecond lightning flash, if bright enough [5].

Stereoacuity is better the more sharply focused the target. Low contrast [6] as well as blunt or distributed patterns like Gabor or Gaussian patches or gratings do not favour performance [7], which also suffers when vision is not equal in the two eyes.

For foveal vision, features would ideally be shown free of encumbrance, separated by no less than 3–4 arcminutes and no more than about one-half a degree of arc [8]. These values become correspondingly larger for peripheral vision and for features much further or nearer than the fixation plane [9].

Perceptual learning, leading to an improvement in performance as familiarity is built up, is a significant feature of the stereo system [10]. In all these aspects, depth discrimination is more fragile than ordinary visual acuity, because the stereo system is more demanding of neural processing.

3. TESTING STEREOPSIS

In a clinical setting, the examiner can give more attention to a patient than in quick checks for the presence of stereopsis in broader surveys for suitability to perform certain tasks. But in all instances stereopsis tests should use patterns that are not too dim, cluttered, brief or unfamiliar. This is best accomplished with simple sharply delineated targets, an unobstructed view, at bright room light luminance levels ($30-100$ cd m^{-2}) and unlimited observation time. Values depend on the pattern, and transfer from one kind of stimulus to another is not the rule [11]. The classical Howard–Dolman two-rod [12] procedure has not been improved (figure 3). For children, one test that works well is the fly test, a tablet showing a fly in a transilluminated polarized view, where the wing is seen well above the plane of the tablet, and the attempt to pull it reveals stereopsis.

The random-dot stereogram (figure 4) has attained popularity because the contours of the test patch are invisible in the absence of stereopsis [14]. Two otherwise identical panels contain many small tokens, and a subset is being given disparity by a horizontal shift through the width of one or more tokens. Monocularly the panels contain no internal contours, but in binocular view, the shifted patch's disparity makes it appear in front or behind the plane of the full panel. The procedure has both virtues and drawbacks. For the observer, there is neither depth nor shape recognition until the disparity has been sorted out in the visual cortex, whereas in tests with clear singular pattern elements, the presence and the identity of the feature are not an issue, only its apparent depth. When administered in standard eye testing or visual surveys, the simpler patterns give more reliable results and are to be preferred.

The reason is to be found in the neural unravelling needed for the disparity to arise in a random-dot stereogram. In figure 4 [13], two rows are shown, one just above the disparity patch, one just within it. A dot in one eye should be matched with just one in the other, leaving in the end only two levels of disparity. Some trial and error is needed, choosing among ambiguous matches until a coherent solution emerges. The neural mechanism by which this is achieved has yet to be elucidated, but it takes time; this is the reason why on first showing, longer exposure durations are needed than for
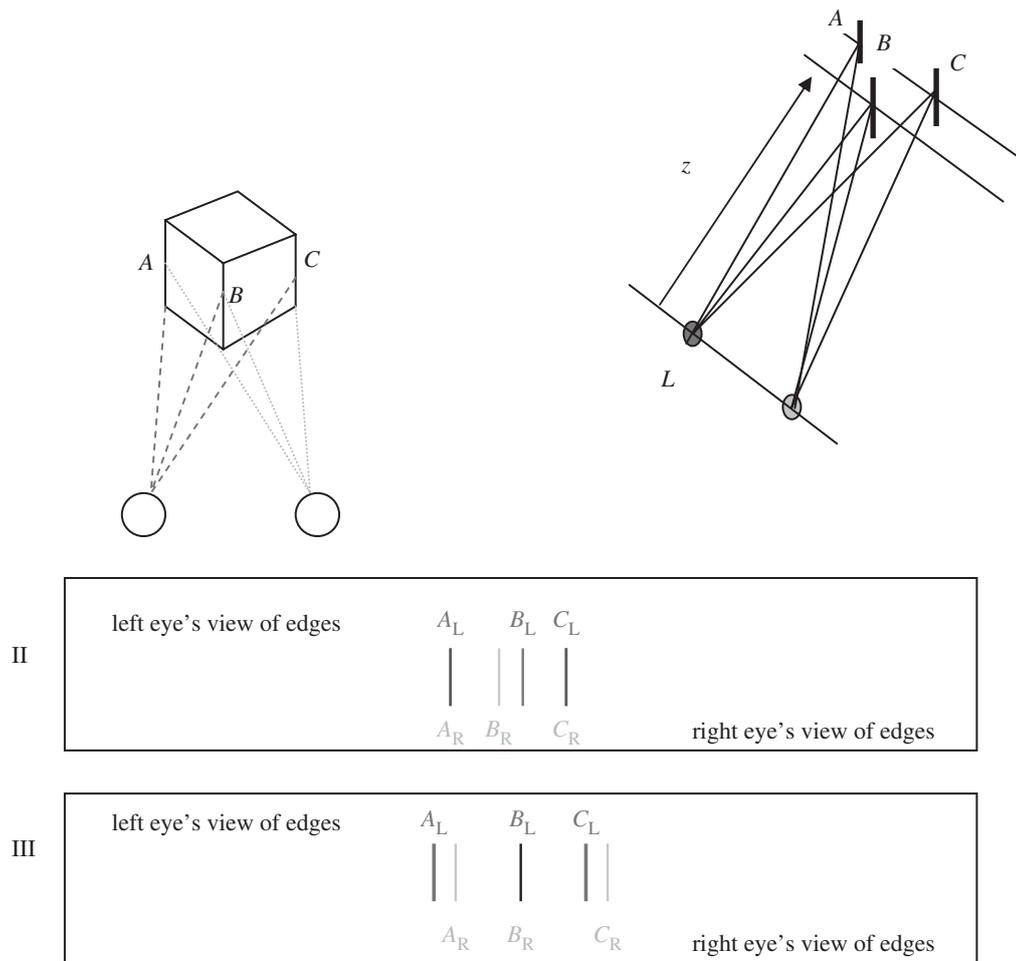
Figure 2. Geometrical layout when a cube is viewed by two eyes. Edge B is nearer than A and C. The angle subtended by AB at the left eye is larger than by BC, and the reverse is true of the right eye. When the eyes are converged on plane AC, targets A and C are on corresponding points and B has disparity, as shown in panel II. When the convergence of the eyes is adjusted for the plane of B, the two retinal images have configuration as shown in panel III.
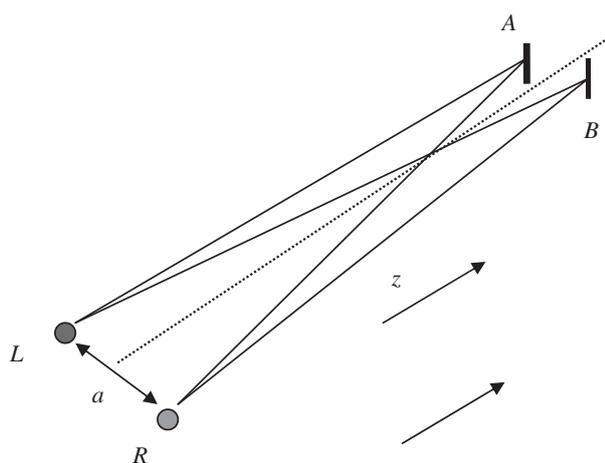

Figure 3. Principle of Howard–Dolman test for stereoacuity. The vertical bar B is fixed in the midline at a distance z of several metres. The vertical bar A above it is adjusted by the observer until it appears just nearer. Alternatively, many settings of equal depth are made and the variance is an estimate of the precision of detectability of depth differences. For a distance of 6 m, a good observer's just discriminable difference in depth is a disparity; i.e. $\delta\gamma = LAR - LBR$, of a few arcseconds, which using the equation $\delta z = a\delta\gamma/z^2$ computes to a depth separation δz of about 1 cm.

simpler stereo ones. However, once a given random-dot stereogram has been solved, it becomes easier on subsequent trials; perceptual learning, known to play a significant role in stereoscopic viewing, is an integral part of testing with random-dot stereograms.

For an uncomplicated, easily administered evaluation of a subject's stereo performance, a panel that uses the time-tested features of the Snellen acuity chart (figure 5) by giving one character in each line successively smaller disparities is ideal. Like all stereograms, it requires presentation of a different image to the right and left eyes and involves one of the methods to which we now turn.

4. METHODOLOGY OF THREE-DIMENSIONAL DISPLAYS

The indispensable condition for a three-dimensional stereo display is to give each eye its own separate view of the world. Here, we will leave aside elaborate virtual displays that, instead of presenting two flat images stereoscopically, use holography or multiple-stage optical imaging to produce suitably differentiated electro-magnetic disturbances reaching the two eyes of the observer.

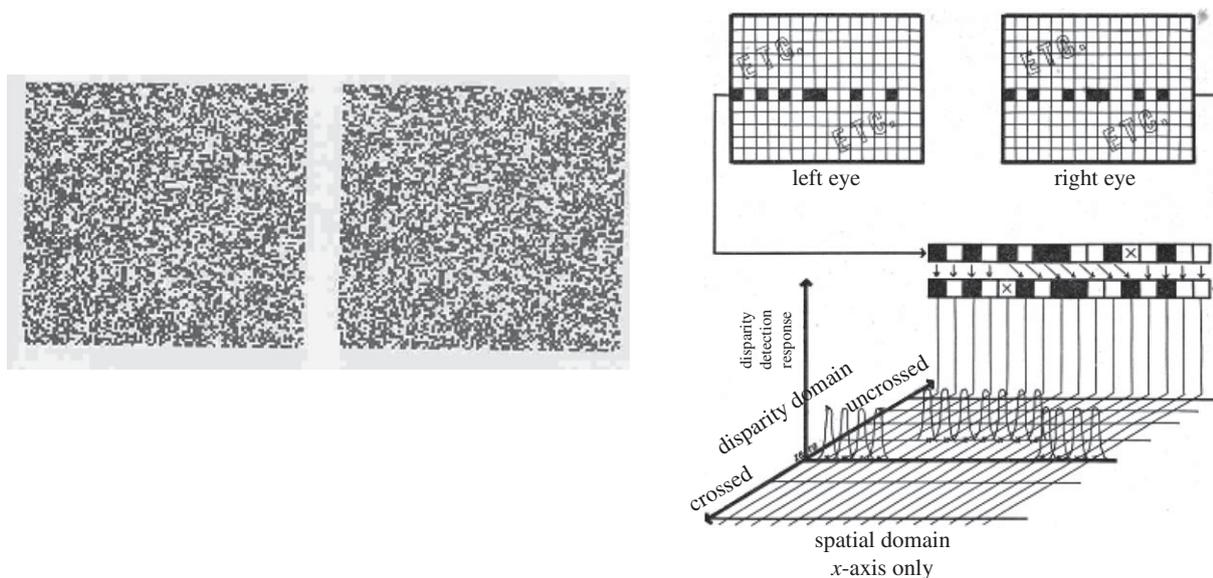Re-representation through photography requires the capture of a two-dimensional image from each of two

Figure 4. Principle of random-dot strereograms. The random-dot pattern is the same in the two eyes, except for a patch in the middle, which is displaced laterally in one eye by the width of one dot with respect to the other. On the right is shown diagrammatically the disposition of the dots in a single row, showing how the visual system is required to establish matches, favouring some over others in the interest of global coherence that then makes the central patch appears behind the remainder of the panel [13].

Figure 5. Three lines of a letter chart for assessing stereo performance, in which the task is the identification of the one letter in each line that differs in depth from the others. Successive lines would have progressively smaller disparities allowing measurement of threshold of stereoacuity. This test, based on an approach used in eye charts used in the clinic, uses clear, sharp, well-articulated targets, known to the observer, and thus optimizes conditions for good stereo performance.
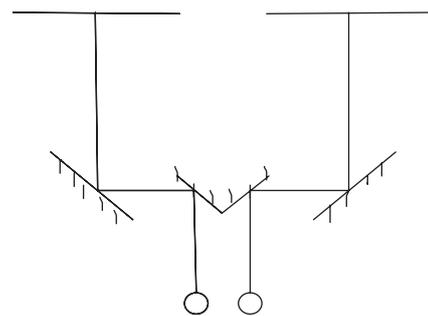

Figure 6. Mirror stereoscope in which the observer sees the left and right images, displayed on side-by-side panels, superimposed by a system of mirrors.

horizontally separated vantage points followed by their display to the two eyes individually. The most direct if technologically most demanding way is to provide each eye with its own optical path and miniature screen on a head-mounted device, though the exclusion of the real outside visual world would introduce problems. Widely used are dual projectors equipped with orthogonal polarizers, directed at an aluminium screen reflection from which retains polarization. Viewing through orthogonal polarizing lenses, cheap and disposable, ensures that each eye receives only its own image.

Alternatively, one eye's picture can be photographed and viewed through a blue filter and the other eye's a red filter, with non-overlapping wavelength transmission bands. The technique, to which the name anaglyph is attached, of course interferes with the normal chromatic properties of the images.

Less convenient but no less effective are mirror systems, in which side-by-side panels, containing the right and left eyes' images, are brought into register by the use of mirrors (figure 6). Practised observers can bring about superimposition voluntarily by fusion without any appliances. Acquiring this skill has become a distinct

advantage for readers of the scientific literature showing such displays.

One procedure in which the right and left eyes' images share the same viewing plane while still retaining segregation, though somewhat awkward in practice, is to populate alternate narrow vertical strips with the image components associated with the right, respectively, left eyes' view in these locations, and direct light from these to the appropriate eye by, for example, alternating vertical strips of right and left deviating prisms or by some means of transillumination by deviating light beams (figure 7). As with mirror systems, correct registration has to be meticulously assured: the deviated beams must not overlap and be separated by the interocular distance in the observer's facial plane.

At the time of writing, the most promising approach is to make the two eyes' images share the whole viewing surface, with a temporal alternation at a high frames rate for the two eyes, preferably 120 Hz or more, with some means of synchronizing each eye's exposure to only the frames meant for it. This can be achieved by viewing through goggles with suitable triggered binocularly alternating occlusion (figure 8). In a related technique, light
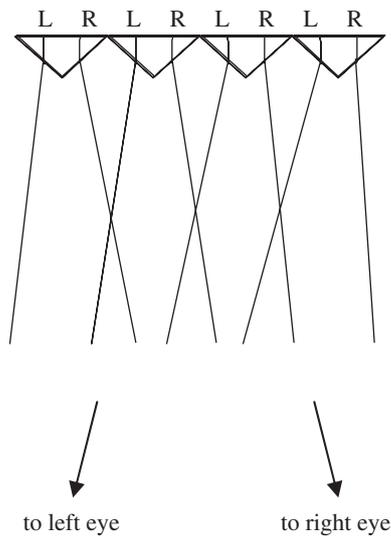
Figure 7. Schema for partitioning a monitor screen into alternate narrow vertical strips containing image segments directed to the left and right eyes, respectively, by suitably deviating their optical paths. Screen and eye positioning are critical.
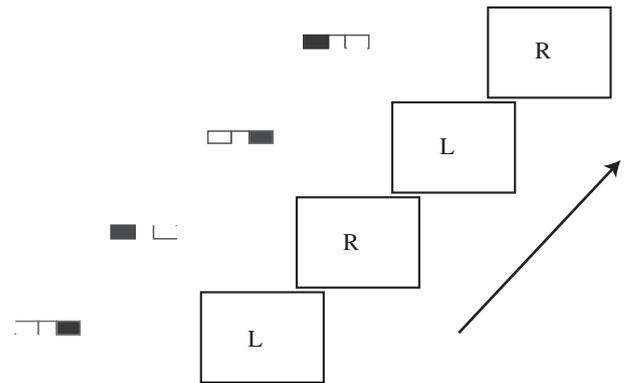

Figure 8. Alternating frames on the monitor display picture for left and right eyes' viewing, synchronized with appropriate alternating occlusion of liquid crystal display goggles. Frame rate needs to be fast enough to prevent flicker.

from the display is passed through a screen with suitably synchronized changing angles of circular polarization. Here, the goggles need only be passively circular polarizing. The advantage of these procedures is that the full capabilities in the spatial and colour domains of the display are preserved with no performance decrement or flicker if the temporal alternation rate is high enough.

All methods of re-representation described collapse a three-dimensional view onto two fixed two-dimensional surfaces, albeit with relative image placements within them appropriate to location in 3-space. But in negotiating the real world, the human visual system has the capacity to adjust to different fore and aft planes both by converging the foveal lines of sight to superimpose individual objects, however far they are, and to change focus as needed. Ordinarily these two functions are yoked; when focusing on the target plane, the eyes will converge on it also. A mismatch can, however, occur. The two eyes may not be in register because of the physical arrangement on the screens or deliberately exaggerated three-dimensional effects. Depending on the observer's accommodation–convergence relationship [15] and the tightness of the linkage, complaints of ocular discomfort may be encountered. This should be handled in the first instance by ensuring that the instrumentation and mode of viewing leave the observer without focus and convergence errors. But even in situations that are, optometrically speaking, unexceptional, prolonged stereo viewing is not infrequently found uncomfortable.

5. DEPTH RENDITION IN THREE-DIMENSIONAL DISPLAYS
When depth in a scene had been captured by stereo-photography and then re-created by one of the methods just described, attention should be paid to the fidelity of the rendition. Brewster [16] early on drew attention to the problem and Helmholtz [17] made reference to it.

However, their accounts lack transparency and fail to clearly articulate the useful distinction, first developed by von Rohr [18], between a display that gives a view identical to the observers' of the original target (tautomorphic), one that preserves relative if not actual depth relationships (homeomorphic), and one that gives different depth (heteromorphic).

In its essence the problem is most conveniently analysed by examining the angular relationships in which an observer with interocular separation a_o views face-on a small cubical element of side length Δl at a distance z_o (in terms of the Cartesian coordinates z for the distance from the observer, and x and y for horizontal, respectively, vertical distances in the observation plane). For the observer, the view is characterized by two variables, both expressed in angular measure: $\Delta l/z_o$, the width of the element, and $\Delta l/a_o z_o^2$, the binocular disparity between front and back surfaces. (If z is large compared with Δl, radians can be substituted for the angles' tangents.) Both are converted by the eyes' optics into retinal positions and from there to neural impulses to the visual cortex [19]. The analysis is laid out in figure 9 in which one edge of the cube has been aligned for convenience with the line of sight of one eye.

Transferring from Cartesian coordinates (where for a cube $\Delta z/\Delta x = 1$) to the angles of disparity and of the width subtended at the eye, the condition for a display to be that of a cube is that the disparity/width ratio be given by

$$r_o = \frac{\Delta l/a_o z_o^2}{\Delta l/z_o} = \frac{a_o}{z_o}.$$

As seen on the screen, the rendition is that of a cube if the ratio of the distances AC/AD (which for large z and small Δl is equal to the ratio of the difference of the angles LAR and LCR to the angle BLA) is equal to a_o/z_o. It expresses a veridical depiction of the cubical element within the polar coordinates operative in binocular vision and remains invariant with magnification.

Suppose now a photographic record is obtained of this target placed at a distance z_c from the principal points of a twin camera with lens separation a_c. In such a record, the rendition will have value $r_c = a_c/z_c$, which remains unchanged with magnification.
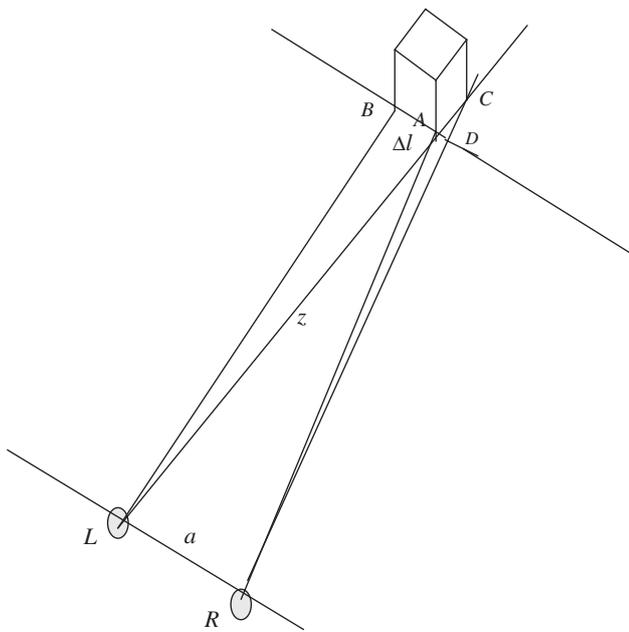
Figure 9. A cube of side length Δl is viewed head-on by an observer with interocular separation a from a distance z, large compared with the other distances. Each side of the front face of the cube has angular subtense $\Delta l/z$ and the cube's depth has disparity $\sim a\Delta l/z^2$. *Depth rendition* is defined as the ratio of the binocular disparity associated with the cube's depth AC to the angular size of its width AB. It is numerically equal to the ratio AB/AD. Angular size $= \Delta l/z$; disparity $= LAR - LCR = \Delta l \ \ a/z^2$; depth rendition $=$ disparity/size $= a/z$.

For an observer with interocular separation a_o, screen images of this record will provide the correct stimulus configuration of a cube only at a specific viewing distance z_o for which $a_o/z_o = a_c/z_c$. At shorter viewing distances, the screen representation of the disparity (which changes as the inverse square of the distance) will be proportionally larger than that of the cube's face width (which changes inversely with distance). As a consequence, the view will be of a structure more compressed in depth than a cube. Conversely, a screen viewed further from the observer than the designated veridical position will give stimulus dimensions of a structure with an expanded depth. For example, for veridicality of depth rendition, a scene photographed at 250 cm with camera interlens separation 25 cm must be viewed by an observer with interocular separation 6.25 cm at 62.5 cm, regardless of any overall magnification with which it is presented. If the reproduction is viewed instead at 50 cm, the observer will experience the cube's angular width reduced by 20 per cent but the disparity by 36 per cent and the disparity/width ratio will no longer be that appropriate to a cubical target.

It must be understood that the discussion has been in optically defined terms, and that there is a fundamental distinction in psychophysics between a stimulus variable and its perceptual correlate. Angular size and binocular disparity are stimulus variables; while related to the observer's sensory task, they exist and are measured in physical object space and hence are subject to simple laws of scaling and superposition. For their counterparts in the realm of perception, however, this is usually not the case. The relationship between the physically defined disparity and its perceptual counterpart, the apparent depth,

is complex, nonlinear and dependent on context [20]. The distinction drawn by von Rohr between homeomorphic and tautomorphic re-representation in a stereoscopic display comes into play because scaling of $r =$ disparity/angular width in the realm of a configuration's visual stimulus does not necessarily lead to a parallel scaling of $\rho =$ seen depth/apparent width in that of its percept. If m is a multiplying factor, $m \times r$ does not, in general, entail a corresponding $m \times \rho$. Veridicality in the perceived depth of a scene is to be expected only when $a_c = a_o$ and $z_c = z_o$. The condition $a_c/z_c = a_o/z_o$ does not suffice.

There is a specific example. Relative depths in a scene at moderate distances seen with binoculars without augmentation of the base distance, i.e. merely with significant magnification, are substantially less than in normal view [21]. Yet, the retinal images delivered by such a device are merely scaled up, with the disparity/width ratios unchanged, i.e. they are homeomorphic. Several factors are at play in the relationship between the stimulus depth rendition factor and an observer's report whether the reproduced scene appears to be veridical, foreshortened or extended in depth. With increasing disparity values, the seen depth becomes a smaller multiple of that at threshold disparity: depth increments do not scale linearity with disparity.

6. PSYCHO-PHYSIOLOGY OF STEREOSCOPY

Small identical targets imaged on the centre of the fovea in the two eyes are seen in the same spatial location; they 'correspond'. With both eyes remaining fixed, the locus of all corresponding points in a horizontal plane containing the two eyes and the fixation point is a line called 'longitudinal horopter' [22]. When a target is shown in places on the two retinas that do not correspond, i.e. some distance in front or behind the horopter, double vision, or diplopia, ensues, although this may not be immediately reported by the observer. There is, however, a little leeway when moving a target out of exact correspondence before diplopia occurs. This is most clearly analysed by separating the view of the two eyes, finding a pair of corresponding points, say the centres of the two foveas, placing a small target in that location in one eye and with rigidly maintained convergence, mapping the range of spatial locations in the other eye over which this target is still seen single, the so-called Panum's fusional area. It has an extent of several arcmins, depending somewhat on the targets being used, and is subject to training [23]. An important experimental demonstration follows: targets, one in each eye, on corresponding points are seen as one and not two, but targets on slightly non-corresponding points—within Panum's areas and hence not yet seen double—will, by definition, have disparity and give rise to stereoscopic depth (figure 10). Thus, the phenomenon of disparity and its use for the purpose of stereoscopic depth perception is to be decoupled from that of fusion, the report of singleness of binocularly presented targets. (Dysfunctions of binocularity, e.g. strabismus, are beyond the scope of this review.).

The decoupling of fusion and stereoscopic depth is also evident in reports of qualitative depth, i.e. the ability to distinguish between the sensation 'nearer' or 'farther', with targets that are far apart from corresponding points on the two retinas and are very prominently seen as double [24].
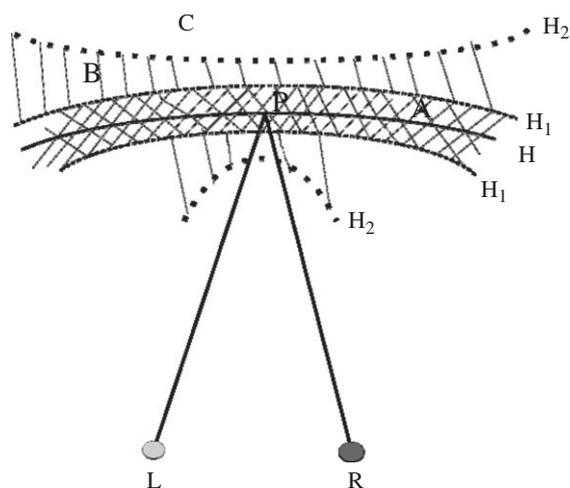
Figure 10. Zones in binocular vision in which points are seen single or double, and with or without depth. Horizontal plane containing right and left eyes and fixation point P on which the lines of sight of the two eyes are converged. Curve H traced out all corresponding points in this situation, i.e. points that are single and have no disparity. Binocular stimuli in the hatched area between the two curves H_1, for example, point A, are still seen single although their depth, viz. their distance from curve H, can be detected. They lie within Panum's fusional areas. Points in the zone between curves H_1 and H_2, for example, B, are seen in diplopia yet have some qualitative depth, i.e. some vague sense of 'nearer' or 'farther' than P, associated with them. Points in the region beyond H_2, for example, C, are seen double and in general lack depth location.

The properties of the primitive neural apparatus subserving stereopsis in the brain are usually charted with geometrical configurations made of simple components. Woven in and superimposed are more complex factors where the individual's previous experience, expectation and attention matter. Their consideration falls into the disciplines of perception and cognition.

Reduction in stereoacuity when targets are too close together, called crowding [25], is one of many subtle interaction effects in the domain of disparity where the seen depth of targets relative to each other is affected in unexpected ways. For example, good depth articulation of isolated features can be lost when they are connected, or become part of a uniform plane surface [26]. Whereas small depth differences in closely adjacent features can be lost by 'pooling' of their depth values, they may be enhanced by further separation in the manner of the well-known centre/surround antagonism of other visual attributes [27].

7. THE PERCEPTION OF DEPTH

The stereoscopic apparatus, intended to tease out the difference in the retinal images of the two eyes that result from their view of the three-dimensional object world from dual vantage points some distance apart in the head, is only one of the components used in the overall appreciation of depth. Its capability of discriminating parallax differences of a few seconds of arc under the best circumstances is a remarkable achievement, and the occasional drawback, such as unwanted diplopia, can be forgiven, as also its fragility under adverse time and

contrast conditions, and its notable dependence on learning and experience. These factors are so well integrated in the operation of the visual system as a whole that they are essentially transparent to all but the most astute observer. But this is precisely what must make us conscious of the framework needed for its seamless functioning. Eye movements, in particular, take place at intervals up to several times a second. This means that the placement of the target images on the retina changes often and, since spatially the retina is in a fixed relationship to the cortex, the analysing circuits there have to be robust to various kinds of stimulus displacements, both in parallel in the two eyes, and relative to each other. This is a tremendous task and the understanding of its neural substrate is only in its beginnings.

Although disparity is processed near the entrance of visual information into the brain [28], often called 'early visual processing [29]', at all times stereoscopy is only one component in our judgement of the third dimension. It is overlaid and interdigitated with highly sophisticated analyses of visual forms and their comparison with stored memory signals leading to deductions about the current disposition of objects in the visual world. Contours individually created on the two retinas need pairing, and this can lead to errors if there is incompleteness or ambiguity among members available for matching [30]. Intervening objects can obscure the view to one eye and block out whole sections of the retinal image. These conflicts are resolved internally usually without rising to consciousness, in the interest of a consistent visual world. A connection is occasionally made with Bayesian Inference and indeed what is now understood under this term (as distinct from what Bayes himself proposed [31]) is, in outline, what actually takes place: an image is created, analysed and the probability is determined that it is one of an array of originating objects, known to exist in the world with a certain prior probability distribution. This enables the computation of a 'posterior probability' that under the particular circumstances, a given object is in fact out there [32]. The procedure, effective in machine vision [33], is only a hazy outline of the human perceptual process, because the probability distributions cannot as yet be expressed with the precision that would make Bayesian Inference a meaningful enterprise. But the idea does capture the essence that the full act of depth perception comprises a variety of components, some of current visual input, some of attention and expectation, and many related to previous memory storage and learning. Though the stereoscopic apparatus is only one of them, it has the advantage of being explicable—and therefore specifiable—in terms of the geometry of optical imagery. Knowledge of its properties, and the situations in which its contributions are unique as well as those in which it can give rise to conflicts, helps implementation on modern electro-optical and computer devices and interfacing the observer with them.

REFERENCES

1 Kemp, M. 1990 *The science of art*. New Haven, CT: Yale.
2 Wist, E. & Gogel, W. C. 1966 The effect of inter-ocular delay and repetition interval on depth perception. *Vis. Res.* **6**, 325–334. (doi:10.1016/0042-6989(66)90066-6)

3 Westheimer, G. 1979 Cooperative neural processes involved in stereoscopic acuity. *Exp. Brain Res.* **36**, 585–597. (doi:10.1007/BF00238525)

4 Ogle, K. N. & Weil, M. P. 1958 Stereoscopic vision and the duration of the stimulus. *A.M.A. Arch. Ophthalmol.* **59**, 4–17.

5 Dove, H. W. 1860 Ueber Stereoskopie. *Poggendorffs Annalen der Physik u. Chemie* **60**, 494–498. (doi:10.1002/andp.18601860710)

6 Westheimer, G. & Pettet, M. W. 1990 Contrast and duration of exposure differentially affect vernier and stereoscopic acuity. *Proc. R. Soc. Lond. B* **241**, 42–46. (doi:10.1098/rspb.1990.0063)

7 Westheimer, G. & McKee, S. P. 1980 Stereoscopic acuity with defocused and spatially filtered retinal iamges. *J. Opt. Soc. Am.* **70**, 772–778. (doi:10.1364/JOSA.70.000772)

8 Westheimer, G. & McKee, S. P. 1980 Stereogram design for testing local stereopsis. *Invest. Ophthalmol. Vis. Sci.* **19**, 802–809.

9 Blakemore, C. 1970 The range and scope of binocular depth discrimination in man. *J. Physiol.* **211**, 599–622.

10 Fendick, M. & Westheimer, G. 1983 Effects of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vis. Res.* **23**, 145–150. (doi:10.1016/0042-6989(83)90137-2)

11 Coutant, B. E. & Westheimer, G. 1993 Population distribution of stereoscopic ability. *Ophthalmic Physiol. Opt.* **13**, 3–7. (doi:10.1111/j.1475-1313.1993.tb00419.x)

12 Howard, H. J. 1919 A test for the judgment of distance. *Am. J. Ophthalmol.* **2**, 656–675.

13 Nelson, J. I. 1975 Globality and stereoscopic fusion in bincular vision. *J. Theor. Biol.* **49**, 1–88. (doi:10.1016/S0022-5193(75)80020-8)

14 Julesz, B. 1960 Binocular depth perception of computer-generated patterns. *Bell Syst. Tech. J.* **39**, 1125–1162.

15 Carpenter, R. H. S. 1988 *Movements of the eyes.* London, UK: Pion.

16 Brewster, D. 1856 *The stereoscope. Its history, theory and construction.* London, UK: John Murray.

17 Helmholtz, H. 1867 *Handbuch der Physiologischen Optik. Abt. 3.* Hamburg, Germany: Voss.

18 von Rohr, M. 1907 *Die binokularen Instrumente.* Berlin, Germany: Julius Springer.

19 Hubel, D. H. & Wiesel, T. N. 1977 Ferrier Lecture. Functional architecture of macaque visual cortex. *Proc. R. Soc. Lond. B* **198**, 1–59. (doi:10.1098/rspb.1977.0085)

20 Ritter, M. 1977 Effect of disparity and viewing distance on perceived depth. *Percept. Psychophys.* **22**, 400–407. (doi:10.3758/BF03199707)

21 Wallach, H. & Zuckerman, C. 1963 The constancy of stereoscopic depth. *Am. J. Psychol.* **76**, 404–412. (doi:10.2307/1419781)

22 Ogle, K. N. 1950 *Researches in binocular vision.* Philadelphia, PA: Saunders.

23 Mitchell, D. E. 1966 Retinal disparity and diplopia. *Vis. Res.* **6**, 441–541. (doi:10.1016/0042-6989(66)90052-6)

24 Ogle, K. N. 1953 Precision and validity of stereoscopic depth perception from double images. *J. Opt. Soc. Am.* **53**, 906–913. (doi:10.1364/JOSA.43.000906)

25 Westheimer, G. & Truong, T. T. 1988 Target crowding in foveal and peripheral stereoacuity. *Am. J. Optom. Physiol. Opt.* **65**, 395–399.

26 Mitchison, G. J. & Westheimer, G. 1984 The perception of depth in simple figures. *Vis. Res.* **24**, 1063–1073. (doi:10.1016/0042-6989(84)90084-1)

27 Westheimer, G. 1986 Spatial interaction in the domain of disparity signals in human stereoscopic vision. *J. Physiol.* **370**, 619–629.

28 Westheimer, G. In press. Path dissociation of visual signals entering the cortex: disparity, contour orientation and position. *Vis. Res.* (doi:10.1016/j.visres.2011.02.020)

29 Regan, D. 2000 *Human perception of objects: early processing of spatial form.* Sunderland, MA: Sinauer.

30 Marr, D. & Poggio, T. 1979 A computational theory of stereo vision. *Proc. R. Soc. Lond. B* **204**, 301–328. (doi:10.1098/rspb.1979.0029)

31 Bayes, T. 1763 An essay towards solving a problem in the doctrine of chances. *Phil. Trans.* **53**, 370–418. (doi:10.1098/rstl.1763.0053)

32 MacKay, D. J. C. 2003 *Information theory, inference and learning algorithms.* Cambridge, UK: Cambridge University Press.

33 Kersten, D., Mamassian, P. & Yuille, A. 2004 Object perception as Bayesian inference. *Annu. Rev. Psychol.* **55**, 271–304. (doi:10.1146/annurev.psych.55.090902.142005)