Stereoscopic visual displays: Principles, viewing devices, alignment procedures*

GEORGE SPERLING†

Bell Telephone Laboratories, Murray Hill, New Jersey 07974

and

New York University, New York, New York 10003

The principle of binocular (stereoscopic) depth perception is that the visual system interprets the slight differences between the views seen by the two eyes as depth cues. In computer-generated displays, two slightly different images are produced on the left and right halves of the display surface and viewed by a prism, mirror, or binoculars system that delivers the appropriate image to each eye. The prism system is the simplest, the mirror system gives the best optical quality, and the binoculars system is useful for producing large apparent images from small display surfaces. All three systems can be adapted for group viewing and all require careful alignment (null adjustment of accommodative distance and vergence distance). Objective and subjective methods of alignment are described.

THE PRINCIPLE OF STEREOSCOPIC DEPTH PERCEPTION

In normal vision, the two eyes receive slightly different views of the world, and the differences between the views are interpreted by the visual system as visual depth. The name given to the depth perception that arises from differences between the images viewed by the eyes is *stereoscopic depth perception*; the two corresponding images are called a *stereo pair (of images)* or a *stereogram*, when they are produced artificially. The principle of stereoscopic depth perception is illustrated in Fig. 1.

In normal vision there are many other important cues to depth, such as those provided by the change in view produced by a head movement, the interposition of objects, perspective transformations, sizes of known objects, distance haze, vergence of the eyes, etc. Of these, only the depth cue produced by head movements is not readily simulated by a visual display system which presents an appropriate stereogram to the two eyes. In fact, most stereoscopic viewing devices require the head to remain still; even when they do not, head movements are relatively easy to measure. As this head-movement information readily can be incorporated into a stereoscopic display system, the simulation of normal binocular vision by a computer display system introduces virtually no limitations that are not already inherent in monocularly viewed displays, i.e., the

limitations of image quality and of information content.¹

VIEWING DEVICES

I shall be concerned here not with the psychological problem of how to compose a stereogram in order to provide the desired illusions of depth,² but rather with the simple technical problem of how to use a cathode ray oscilloscope (CRO) display system economically and efficiently to provide two different views. Basically, there are only two methods of producing the two images of the stereogram: (1) to use two different CROs, one for each image, and (2) to use two different subareas of a single CRO display surface (e.g., the left and right halves), and to view each subarea separately with each eye.

The method of two CROs is more expensive and has inherent in it the problem of adjusting the two CROs to be electronically equivalent. The two-CRO solution is useful mainly when it is necessary to separate the stereogram members (not readily possible on a single display surface because the two display areas are constrained to be adjacent to each other) and when independent hardware control of such parameters as x- and y-axis magnification is desired. For example, to control independently the stimulus area in a large annular neighborhood of each computer-generated display, it is convenient to separate the stereogram members. In principle, however, the problems of viewing two CROs are the same as those in viewing two subareas of one CRO, and the examples will refer only to the latter case.

The three main methods of viewing subareas of a display surface separately with each eye are illustrated in Fig. 2; they are (1) a prism viewing system, (2) a mirror viewing system, and

(3) binoculars. All three systems are improved by providing a viewing hood to position the O's head in front of the device to reduce extraneous stimulation. The viewing hood and prisms (or mirrors) usually are attached to one end of a sturdy, lightweight frame. The other end of the frame should be quickly and conveniently attachable to the display device itself in order to provide rapid changeovers between normal and stereoscopic viewing of the visual display. The binoculars system also

Fig. 1. The principle of stereoscopic depth perception. (a) Top view of an O viewing two surfaces (S_1 , S_2). The two eyes see the nearer surface against different parts of the background, as indicated. The displacement (between the two views) of the foreground relative to the background is the stereoscopic depth cue. (b) The stereogram, i.e., the views seen by the left and right eye. The stereo pair of images may, of course, be produced either on a visual display surface or by viewing a natural object. (c) Perspective drawing of the illusion of depth that results from viewing the stereogram in (b). Note that the projection of S_2 against S_1 in (b) exaggerates its size; when S_2 is seen in stereoscopic depth (c), the apparent size of S_2 approximates its actual size (a) rather than its projected size (b).

*Originally presented at the North Atlantic Treaty Organization, Advanced Studies Institute on: On-Line Computing for Behavioral Science, Department of Psychology, Sheffield, England, July 19, 1970.

†Requests for reprints should be sent to: G. Sperling, Room 2D518, Bell Telephone Laboratories, Murray Hill, New Jersey 07974.

Fig. 2. Three stereoscopic viewing systems: (a) prisms, (b) mirrors, (c) binoculars. L and R indicate the nominal centers of the left and right images, respectively, of the stereogram: M = mask, CRO = cathode ray oscilloscope (the visual display surface). Dashed lines indicate lines of sight extended to their intersection. In (a) and (b) their intersection occurs at the accommodative distance of the display indicating correct alignment of the systems. In (c), the accommodative distance is assumed to be infinity, and the lines of sight are parallel, again indicating correct alignment.

can be made detachable, but as the binoculars can be located far from the display, they do not interfere with normal viewing and can be left permanently in place.

Prism System (Brewster, 1856)

The advantage of the prism viewing system is that it is the easiest to implement. A prism is needed only in one eye, and this may even be held by hand. In fact, with practice, Os can learn to deviate their eyes (either diverged or crossed) so that the prism is superfluous. However, unnatural deviation of the eyes is undesirable for reasons to be discussed later.

The disadvantages of the prism method are: the prism produces a distortion of straight lines ("prism distortion"); the power of a prism varies depending on the eye's direction of gaze and on their vergence, and this effect interacts with any lenses the O may be using (Fry, 1937; Ogle, 1951); the prism reduces the optical quality of the image being viewed (particularly by introducing chromatic aberration, which is disturbing against dark backgrounds); and the centerline-separation shield, which extends forward from the nose to the center of the display, is an awkward restriction of the image being viewed.

The optical quality of the image can be improved by using two low-power prisms, one in front of each eye, instead of one strong prism, and by using achromatic prisms. (Achromatic prisms are composed of two different

kinds of glass, combined so as to minimize the chromatic aberration inherent in the use of an elementary prism.) In a useful variation of the prism method, the single prism (in either or in both eyes) is replaced by a pair of equal prisms. The prisms in such a pair can be rotated against each other to produce—in effect—a single prism of variable power. This is a convenience, but it adds to optical distortion.

The prism method for group viewing. With some slight modifications, the prism method can be adapted for group viewing. The prism (or prisms), preferably adjustable, is fitted on spectacle frames that are worn either alone or immediately in front of the viewer's usual spectacles (when these are used). Because there is now no centerline-separation shield, each eye will see both images of the stereogram, i.e., the entire display. This extraneous stimulation may create serious viewing problems. The simplest way to restrict the viewing of each eye to its intended member of the stereo pair is by means of a corresponding pair of crossed Polaroids, one pair of which is fixed to the display surface and the other to the spectacle lenses (Fig. 3); the remaining extraneous stimulation is removed by darkening the room.

The more persons who wish to view a display, the farther away they must stand. (The adjustable prism permits the O to vary his viewing distance.) A distant viewer not only sees a

"smaller" stereogram than a near viewer, but he perceives differently shaped objects because perceived depth relations do not vary linearly with viewing distance. Whereas the shape of a two-dimensional object remains invariant with changes in viewing distance, the three-dimensional shape of an object defined by a stereogram changes as the viewing distance changes. In producing stereograms, the absolute retinal size at which they will be viewed must be taken into account, and the viewer must also consider this.

Mirror Method (Wheatstone, 1838)

The advantages of the mirror method are: it is virtually free from optical distortion when good quality front-surface mirrors are used; it is extremely easy to adjust by rotating the mirrors slightly; and it allows convenient placement of masks to limit the field of view, thereby giving good control of the overlap area (the shared display area between the two subareas which serves as a buffer zone for each subarea). The disadvantage of the mirror system relative to the prism system is that it is more complicated to construct. Like prisms, mirrors can be worn as spectacles to facilitate group viewing.

Fig. 3. Prism-plus-Polaroid system for group viewing. Spectacle frame (F) contains Polaroid filters and an adjustable prism (AP). V (vertical) and H (horizontal) indicate the direction of polarization of the Polaroid filters.

The mirror system would seem to have an advantage over the prism system in enabling each half of a curved display surface to be viewed perpendicularly, thus reducing the foreshortening that results from the nonperpendicular view in the prism system. In any viewing system, the first-order horizontal foreshortening error can be compensated by adjustment of the x-axis magnification. However, perpendicular viewing eliminates the first-order vertical foreshortening error, as well as the horizontal error. When the display surface is a true circular arc, the second-order errors (due to noncorresponding screen curvature in the two images) also are eliminated by perpendicular viewing, but these errors are unlikely, in any case, to be significant in comparison to other image defects.

Binoculars Method

The usual implementation of mirror and prism systems limits precisely controlled viewing to one O per CRO. When there are several Os, these systems require the multiplication of CROs and the associated viewing device, one system for each O. One way to overcome this multiplication is to use binoculars to view the displays. Because they magnify the images, binoculars can be placed farther from the displays than the direct viewing devices; consequently, many Os can view the same display simultaneously. If desired, they can be viewing it at different magnifications.

The disadvantages of the binoculars method are: the viewing stand with binoculars on it must be rigidly fixed relative to the display; the system is difficult to adjust (particularly centering the images) and difficult to maintain in adjustment; viewing is difficult for the O because his head must be exactly positioned vis a vis the binoculars, and the optical quality is likely to be the poorest of the methods. On the other hand, it is far cheaper to obtain a large effective display area by viewing a small CRO with binoculars than by buying a large CRO. When very small displays are required, they are conveniently produced by inverting the binoculars. The optical quality of binoculars is much less critical when they are used to minify than when they are used to magnify.

ALIGNMENT

Accommodation and Vergence

Accommodation is the technical term for the amount of focusing of the lens of the eye. Vergence refers to the angle between the two eyes that is required to place the image of a single object onto corresponding points of the two retinas.

It is not necessary here to define additive scales of accommodation and vergence (see Emsley, 1948; Sperling, 1970). It is sufficient simply to refer to the actual distance from the O to an object for which accommodation or vergence is appropriate. The stimulus dimension that is crucial to accommodation is called the *accommodative distance* of an object; the eye's response is the *accommodated distance*. Similarly, the critical stimulus dimension for vergence is the *vergence distance* of an object; the response is the *verged distance*. For example, in normal vision an object at a distance of 1 m from an O causes the lenses of his eyes to adjust accommodation until they accurately focus light rays emanating from a distance of 1 m, and it causes his eyes to converge so that the lines of sight of the two eyes intersect at a distance of 1 m.

In the world of real objects, the stimulus to accommodation and the stimulus to vergence are perfectly correlated, and the visual system takes this into account. For example, when a stimulus only to accommodation and not to vergence is presented (i.e., one eye is kept covered), then the eyes nonetheless verge upon it (i.e., even the covered eye points at it). Similarly, a stimulus to vergence causes the eyes to accommodate appropriately. Fortunately, there is some flexibility in the visual system so that the eyes can be accommodated for one distance and verged upon a different distance, but only a limited amount of *accommodation-vergence disparity* can be tolerated. When this amount is exceeded, then either a failure of accommodation occurs (producing a blurred image) or a failure of vergence occurs (producing a "double" image), or both occur.

A computer can readily generate stereograms that define objects at numerous different vergence distances (and thereby give complex illusions of stereoscopic depth), but no convenient way has yet been discovered to concomitantly covary the stimuli to accommodation. For example, by varying their vergence distance, different objects defined by a stereogram may subjectively appear to be at different depths, ranging from immediately in front of the nose to infinity, even though accommodative distance remains fixed (at the distance from the O to the display). Generally, the inability to vary independently the accommodative distance of segments of a computer-generated display is not an important limitation. It is sufficient to vary the overall accommodative distance, and this is easily done by placing spectacle lenses in front of the O's eyes.

Matching Accommodative Distance and Vergence Distance (Alignment)

An important design consideration in stereoscopic displays is insuring that vergence is appropriate for the accommodative distance of the display. For example, let the accommodative distance be 1 m. Then, when the lines of sight of each eye are pointed at the nominal centers of their respective visual display, these lines of sight should intersect at 1 m. It is the task of the engineer to insure that the lateral separation of the left and right stereo images on the display surface and the optics of the viewing device are adjusted to accomplish this.

The alignment procedure may be divided into four steps which are best carried out in the following sequence: (1) determine the accommodative distance of the display; (2) alter the accommodative distance if desired; (3) measure the vergence distance of the display; (4) adjust the vergence distance to equal the accommodative distance. For many purposes, a simple subjective method that bypasses the first three steps will suffice, but it is useful also to know objective methods.

(1) *Measuring accommodative distance.* The easiest way to determine accommodative distance in the prism and mirror systems is to measure the optical path with a ruler. In binoculars systems, a camera with a ground-glass viewer probably is the simplest way to determine the accommodative distance of the display. The camera lens is placed at eye position and focused to produce a sharp image. The determination depends, of course, on the previous calibration of the camera lens, and it is the more accurate the greater the focal length.

(2) *Altering accommodative distance.* To alter accommodative distance, an ordinary spectacle lens placed before the eye is all that is required. The power D in *diopters* of the required lens is given by the formula

$$D = \frac{1}{A_2} - \frac{1}{A_1} \quad (1)$$

where A_1 is the initial accommodative distance (in meters) and A_2 is the desired final distance. Positive values of D indicate convex (positive) lenses and negative values indicate concave (negative) lenses.

(3) *Measuring vergence distance.* The best method requires constructing a device for producing two narrow collimated beams of light (e.g., two penlights with small apertures at the front). The collimators should be mounted on a board with the front ends 6 cm apart and the rear ends adjustable. The main requirement is that the dual beams should produce clearly defined spots at a distance

$$V = \frac{Ap}{m - S} \text{ and } S = m - \frac{Ap}{V} \quad (2)$$

Equation 2 relates the lateral separation of any two corresponding points in the stereogram to their vergence distance provided A , p , and m are known. Determination of the accommodative distance, A , was outlined above. The interpupillary distance, p , may be taken as .06 m. (The average population value is about 6.2 cm.) The dual parallel light beams project on the screen with the lateral separation of m (Fig. 4a).

It may be noted here that perceived stereoscopic depth depends far more on relative values of lateral separation, S , for two different objects in the stereogram than it does on the absolute value of S . Also, exact alignment (and perceived depth) will vary slightly between Os because of their different interpupillary distances.

(4) *Adjusting vergence distance.* In prism and mirror systems, vergence distance is easily adjusted by varying the power of the prism and the angle of rotation of the mirror. Binoculars systems also require a prism adjustment. The prism is placed on the large side of the binoculars (i.e., away from the eye); because of the larger distances, very weak prisms will suffice. Fine adjustments of vergence distance usually are made by varying the lateral separation, S , of the stereogram directly under program control. The final value of vergence distance should normally be set equal to the accommodative distance, assuming an interpupillary distance of 6 cm.

Subjective Method of Aligning Vergence Distance (Haploscopic Method)

Fig. 4. Calibration of vergence distance. (a) Two parallel beams of light are produced by the collimators (c). The light beams (f) pass through the stereoscopic viewing device (d) to project onto the display surface with a lateral separation of m cm. The apertures of the collimators are mounted on a board (b) and are separated by the interpupillary distance (p), nominally 6 cm. (b) Diagram to illustrate that when the lateral separation is diminished by p (twice $p/2$), the vergence distance changes from infinity (parallel lines) to a vergence distance of A (the accommodative distance). (c) The display surface, showing a stereogram (L, R). The Ms are assumed to be separated by a lateral separation of m , corresponding to a vergence distance of infinity—e.g., as in (a). The lateral separation between the As is $m - p$, corresponding to a vergence distance of A (the accommodative distance), e.g., as in (b). The Cs are laterally separated by $m - 2p$, corresponding to a vergence distance of $A/2$. In a stereoscopic view, the Cs would appear to be in front of the screen, the As in the plane of the screen, and the Ms behind it. The stereogram is correctly aligned when vergence distance equals accommodative distance; for prism viewing this occurs when the nominal centers of the L and R subfields have the same lateral separation as the As in (c).

equal to accommodative distance of the display surface.

The dual beams are placed at the same vantage point as the eyes and pointed at the display; there they produce two spots on the display surface (Fig. 4a). When the angle between the beams has previously been adjusted to the desired vergence distance (the point at which the beams intersect), the distance between the two spots (*lateral separation*) will define this vergence distance on the display surface.

As a practical matter, it is useful to set the collimated beams parallel because their projections on the display surface then define the lateral separation of objects at infinite vergence distance. This is the maximum lateral separation of "real" objects. When the lateral separation of two images on the display surface is decreased by the interpupillary distance (typically 6 cm) then the vergence distance equals the accommodative distance.

The relation between lateral separation on the display surface (S) and vergence distance (V) can be expressed in general terms.³ Let the lateral separation on the display surface between points at infinite

accommodative distance be m , the interpupillary distance be p , and the unaltered accommodative distance of the display be A . Then the vergence distance of any object on the screen is related to the lateral separation of the two images by

Fig. 5. Haploscopic method of adjusting vergence. (a) The test patterns. View of the display surface showing the left (L) and right (R) members of the stereogram. Each line is centered in its field. (b) Stereoscopic appearance when the viewing system is properly adjusted. (c) Stereoscopic appearance of a maladjusted system.

In this method, a stereogram such as that illustrated in Fig. 5 is presented. The engineer adjusts the prisms, the mirrors, or the stereo pair's lateral separation on the display surface until the O reports that the vertical line (which is seen only by the right eye) falls in the middle of the horizontal line (which is seen only by the left eye).

Haploscopic adjustment is virtually incumbent in the binoculars system. It is also useful in group viewing, as it allows Os to make accurate alignments, and it is particularly helpful in diagnosing failures to perceive stereoscopic depth.

Other Considerations

In all three systems, the viewer should be encouraged to wear spectacles if he normally uses them for objects at the accommodative distance of the display. Only in his normal viewing mode will his accommodation and his vergence be in their normal

relation, i.e., appropriate to real objects.

When viewing continuously visible stereograms in which the two images are very similar, there is great tolerance for alignment error, and successful vergence occurs even with large accommodation-vergence disparities. Some experiments, however, require presentation of basically different stimuli to each eye. To insure that the eyes are correctly verged in such cases, the engineer should do two things: (1) He should provide a "visual frame" that is seen by both eyes, within which the diverse stimulus materials appear; and (2) he should adjust the vergence distance individually for each O (subjective method) under conditions as similar as possible to those of the experiment.

REFERENCES

- BREWSTER, D. *The stereoscope, its history, theory and construction*. London: Murray, 1856.
 EMSLEY, H. H. *Visual optics*. (4th ed.) London: Hatton Press, 1948.

FRY, G. A. An experimental analysis of the accommodation-convergence relation. *American Journal of Optometry*, 1937, 14, 402-414.

OGLE, K. N. Distortion of the image by prisms. *Journal of the Optical Society of America*, 1951, 41, 1023-1028.

SPERLING, G. Binocular vision: A physical and a neural theory. *American Journal of Psychology*, 1970, 83, 461-534.

WHEATSTONE, C. Contributions to the physiology of vision. Part the first. On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, 1838, 8, 371-394.

NOTES

1. The divergence of light from an object—the stimulus to accommodation (focusing) of the eye—is exceedingly awkward to simulate in any kind of display; see section on Accommodation and Vergence.

2. See the section entitled, (3) Measuring vergence distance.

3. The treatment is simplified by assuming small angles, i.e., that the unaltered accommodative distance (the "viewing distance") is large relative to the interpupillary distance and relative to the picture width of each member of the stereogram. In practice, angles almost always are small.