Ann. Rev. Neurosci. 1984. 7:379–412
Copyright © 1984 by Annual Reviews Inc. All rights reserved

THE ANALYSIS OF STEREOPSIS

Gian F. Poggio

The Philip Bard Laboratories of Neurophysiology, Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Tomaso Poggio

Department of Psychology and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

INTRODUCTION

Our eyes capture different 2-D images of the objects around us and our brain uses these images to recover a description of the 3-D structure of the environment. Stereopsis is the process responsible for this reconstruction of the depth dimension in our visual world. Since Wheatstone invented the stereoscope in 1838, the processes underlying primate stereo vision have been intensively studied, first with psychophysical techniques, and more recently in terms of the underlying physiological mechanisms in the visual cortex. During the last few years, the computational aspects of the problem of stereo vision have received increased attention. As with so many other visual tasks that humans perform easily and effortlessly, the development of automatic systems of stereoscopic vision, which would yield immediate and important applications, has proven surprisingly difficult.

It is now clear that the problem of stereopsis is not only a problem in the area of psychophysics and physiology, but also a complex problem in information processing. Because of its knowledge-free, low-level character, the solution of the stereopsis problem may thus uncover some fundamental principles that apply equally well to artificial and natural vision systems.

This somewhat new perspective is reflected in the organization of this paper. We first review the computational problems posed by stereopsis and describe briefly the main theories and models that have been proposed in the context of human and computer vision. The structure of these algorithms and their

0147-006X/84/0301-0379$02.00

performance and limitations suggest a number of critical questions about human stereopsis. For each of these areas of investigation, we review relevant psychophysical data and try to put in focus the main open questions. The final two parts of the paper review the physiology of stereopsis and attempt to relate it to both the psychophysical and computational levels. This review is by no means exhaustive; we have concentrated on a few recent developments and aspects of the analysis of stereopsis that we consider of particular importance in providing new information.

INFORMATION PROCESSING IN STEREOPSIS

What is stereopsis? Because our two eyes are located in different positions in the head, their views of a 3-D scene are slightly disparate. One can easily experience directly this binocular disparity by looking at objects not too distant and noting their different relative positions when closing each eye in turn. The disparity of each "point" depends on its distance from the fixation point of the two eyes. Our brain is capable of measuring this disparity and using it to produce the sensation of depth that is the subjective estimate of relative distance. This is stereopsis and its sole basis is the horizontal disparity between the two retinal images. There are, of course, several cues to depth, like perspective, texture gradients, and shading, that are used in our everyday vision. Binocular disparity is one of the most important and accurate of them. We are concerned here only with stereopsis, considered as the information processing module in our brain that is responsible for measuring and using binocular disparity.

The Computational Problems of Stereopsis

From the experience of everyday life, it is not clear how to separate stereoscopic processing from the monocular analysis of each image. Line stereograms (like the ones of a chemical structure) already show that stereopsis only needs disparity information: shading, perspective and texture gradients are not critical. An early step in stereo processing is to compute disparity between the two images. With knowledge of disparity, an estimate of distance can then be recovered from the geometry of the situation. Thus, the processing involved in stereopsis includes matching corresponding points of the images in the two eyes, measuring their disparity, and from this information recovering the 3-D structure of the objects seen by the viewer.

THE CORRESPONDENCE PROBLEM If one could identify with certainty a location in the two images, the first two steps could be avoided and the problem would be easy. In practice, one cannot mark spots in the scene and the difficult part of the computation is solving the correspondence problem. Before Julesz's

work, it was thought that this problem was in fact avoided by first recognizing objects and their components, and then performing an unequivocal correspondence between the same recognized detail in the left and in the right image. In this way, there is never any real problem in deciding what should match what. This comes at the cost of placing the correspondence process underlying stereopsis quite late in the processing of visual information, certainly after object recognition. Julesz (1960) demonstrated that it is not so by inventing the random-dot stereogram, in which there is no information whatever about visible surfaces except for disparity.

Thus, experiments with random-dot stereograms proved that binocular combination need not happen after object recognition. They strongly suggest, in fact, that the correspondence problem preliminary to stereopsis is solved early on, independently from higher level processing. This fits well with our present knowledge of visual physiology. Binocularity appears very early in the visual pathway, before any complex recognition process has taken place.

Julesz's observation is critical because it allows one (*a*) to consider the processing module involved in decoding stereopsis independent from other processes (at least to a first approximation), and (*b*) to formulate the computational goal of human stereopsis as the extraction of disparity information from a pair of images, without the need of monocular cues. The main problem that human stereovision has to solve is what has been called the correspondence problem—how to find corresponding points in the two images without recognizing objects or their parts. Random-dot stereograms (Figure 1, top) would seem to confront our brain with an immense number of possible matches between the two images. After all, dots in one image are all the same, of the same size and contrast: Any given dot in one image could in theory be matched with any one of a large number of dots in the other image. And yet our brain solves the false target problem and comes up with the right answer. How does the brain know what corresponds to what?

There is an additional problem preliminary to the matching operation at the level of a computational theory of stereopsis. We have to specify what has to be put in correspondence between the two images. To be sure, the basic information available to a visual system is an array of measurements of light intensity at different image points. In our eye, these measurements are taken by the photoreceptors. Each point in an image is the projection along a line of sight of a point on a physical surface (unless the camera is looking at the sky!). If a point is visible by both eyes or cameras, then its projections are corresponding points in the image and should be matched. Attempts to use directly the intensity values at each pixel as the elements to be matched have had little success. It is not difficult to see why a scheme solely based on intensity is doomed to failure. Typically, corresponding points in the two images do not have exactly the same shade of gray, because of differences in the vantage points of the observer's

eyes and noise intrinsic to every imaging and sensing system like our eye or a camera. Furthermore, numerous points in a surface might fortuitously be the same shade of gray, especially in the neighborhood of the correct match. Finally, it is easy to show that images of a stereopair that have rather different contrasts can be easily fused by our brain (Julesz 1971). On the other hand, we have already seen that a high level representation of recognized objects is not used by the stereomodule. More primitive measurements taken on the intensity arrays are used by the correspondence process and a critical problem in stereopsis is the nature of this representation. In brief, for understanding the stereo computation we have to answer two questions: what to match and how to match.

The matching primitives Gray levels themselves are the most elementary form of "features" that could be used for the correspondence problem. They are, however, as we mentioned above, unreliable. Moreover, the intensity values measured by the photoreceptors are not directly transmitted to the cerebral cortex, where binocular combination occurs. An alternative approach is to match higher level features, typically oriented edges, extracted from each image. In 1967, Barlow, Blakemore & Pettigrew suggested just this solution to the correspondence problem, by pointing out that "the number of identical trigger features lying in the same appropriate region of each eye can safely be assumed to belong to the same object." Most recent computer algorithms and models of human stereovision that rely on some form of edges as the basic primitives perform far better than gray level correlation techniques on most images. This, of course, seems to be in encouraging agreement with the physiological evidence about the "edge detection properties" (without for now specifying what is meant exactly) of binocular neurons in the visual cortex.

Yet, one should not overlook the possibility that the primitive measurements used for stereo matching may be far more dense and specific than just location and parameters of edges. Many different measurements at each point of the image, for instance various types of derivatives, may provide a rich and robust description of each image, suitable for matching. One can easily see how this may work: Almost every point in each image of a stereopair may be characterized in terms of several measures of local variations of light intensity; corresponding locations in the other image can be found by identifying the point characterized by the most similar set of measurements (see Marr & Poggio 1976). An algorithm of this type for solving the correspondence problem has been recently developed and tested with encouraging results (Kass 1983). It is based on the binocular conjunction of a large collection of independent linear measurements, specifically partial derivatives of images smoothed via a few oriented filters of different size.

The other possibility is that a simpler set of measurements on the image intensity allows satisfactory matches to be assigned. Several recent theories of stereopsis are indeed mainly based on one set of primitives—called zero-crossings—which are specific relatives of edges. The hypothesis that a zero-crossing map could be one of the main products of the first phase of visual processing was suggested by a combination of psychophysical and physiological data on human vision (Marr & Poggio 1979). The first step in the scheme was to filter the image through a low-pass filter in order to cut out the high spatial frequency components of the signal. The next step was to take the second derivative of this filtered intensity array and to detect the zero-crossings in it. These zero-crossings then correspond to the inflection points of intensity changes in the filtered array, i.e. the points at which intensity is changing most rapidly. Originally, these steps were conceived as the operation of convolving the image with a particular function, the difference of two Gaussians of opposite sign (DOG function), which effectively performs both operations simultaneously. These filters are not tuned to orientation, but may be quite similar to a center-surround type of receptive field, as suggested by psychophysical observations (Mayhew & Frisby 1978a, 1979a) and computational arguments (Marr & Hildreth 1980). Spatial filters of this type are band-pass: They respond optimally to a certain range of spatial frequencies in the image. In summary, for a given resolution, the process of finding intensity changes consists of convolving the image with a center-surround filter whose spatial dimensions reflect the scale over which the changes have to be detected, and then of locating the zero-crossings in the filtered image. This process is performed in parallel at several different resolutions, that is, with filters of several different dimensions. A theorem by Logan (1977), which does not strictly apply to images filtered through the DOG filters, suggests nevertheless that the relatively sparse number of "discrete" symbols provided by zero-crossing are very rich in information about the filtered image and represent, therefore, one of the candidates for an optimal encoding scheme used for later processes. Marr & Poggio (1979) have suggested that they are the most important (but not the only!) primitives to be matched between the two images.

The constraints of matching Even for quite specific matching primitives, the false matches problem cannot be completely avoided. Ambiguous matches have to be solved, taking into account constraints that the world and the geometry of the imaging process dictate. Some of these constraints are very simple and straightforward and some are somewhat more subtle. Existing algorithms differ in the emphasis given to these constraints and in their implementation, that is, in their matching rules. The most obvious constraints to be taken into account are the following:

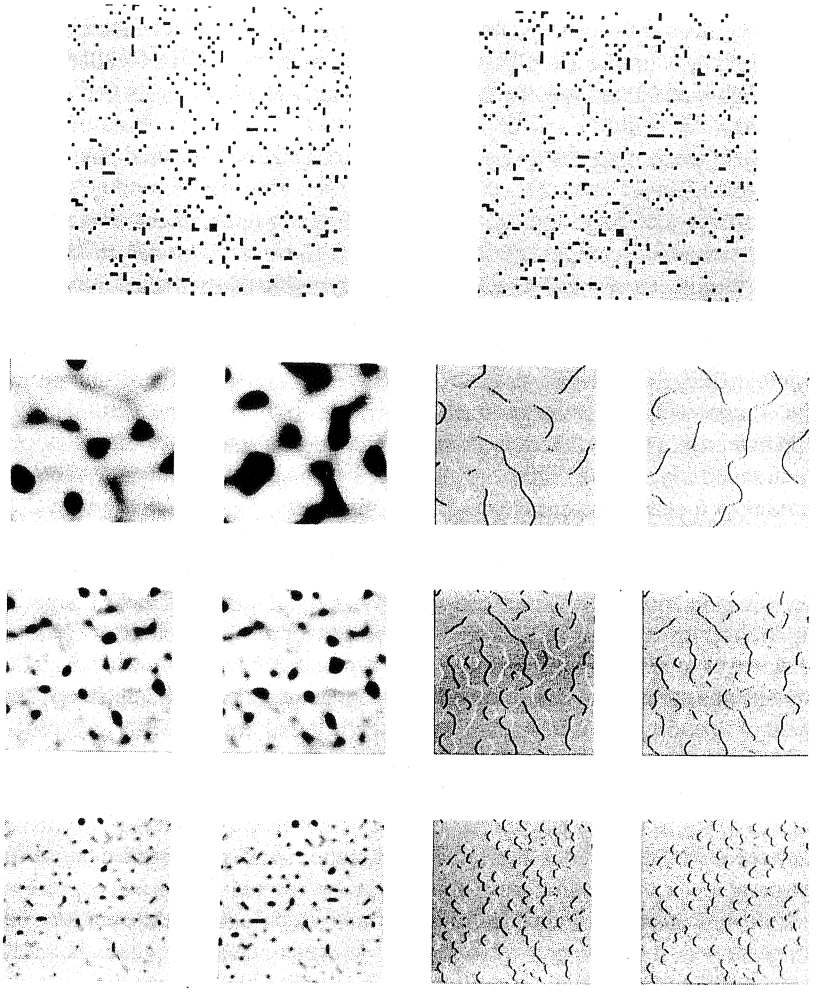


Figure 1 A random-dot stereogram (density 10%) is shown at the top. Its convolutions with center-surround masks of three different sizes are indicated in the *two left columns*. The convolution values are positive and negative: they are represented here in terms of shades of gray, positive values black and negative ones white. The coarse structure of the image is captured by the large filters; increasingly finer details are seen by the smaller masks. If the size of the dots in the stereogram is taken to be $4' \times 4'$ (at the corresponding distance the stereogram subtends slightly more than 4° of visual angle) the diameter of the center region of the three masks is $35'$, $17'$, and $9'$ respectively, from top to bottom.

The *two right columns* show the zero-crossings obtained from the convolutions in the left two columns. If positive values in the convolution were associated with activity of ON-center retinal ganglion cells and negative values with the activity of OFF-center cells, then the zero-crossings would correspond to the ON-OFF transitions between the activity of the two types of cells.

(These images are courtesy of E. Grimson of the Artificial Intelligence Laboratory at M.I.T.)

1. The uniqueness constraint: A given point on a physical surface has only one 3-D location at any given time. This translates into the rule that an item in one eye should usually be matched with only one item in the other eye (Marr & Poggio 1976).
2. The continuity and ordering constraint: Variations in the distance of surfaces from the viewer are generally smooth, with discontinuous changes being encountered at object boundaries (Marr & Poggio 1976). Thus, except at object boundaries, the disparity gradient should not be too high and, in particular, ordering of edges in the two images should be preserved (Baker 1982, Mayhew 1983).
3. Trigonometric constraint: This determines the orientation of the epipolar lines, that is, the pairs of straight lines in the two images that match point by point. When the observer fixates a distant point (ideally, at infinity), the epipolar lines are horizontal and parallel: In general, they depend on the direction of gaze. If the epipolar lines are known, the matching problem is essentially one-dimensional: The search for matches is limited to corresponding rasters in the images in the two eyes. If the epipolar lines are not known, the matching problem, now truly two-dimensional, becomes dramatically more difficult, because of the increase in number of false matches.

Other constraints have also been proposed by various authors to help solve the correspondence problem. Two of the most interesting ones are:

4. Connectivity in 3-D space of matched edges (Baker & Binford 1981).
5. Relative orientation of corresponding edges in the two images can be slightly different, depending on the viewing geometry. For the typical geometry and distance characteristic of human vision, the expected half-width of orientation difference is about nine degrees (Arnold & Binford 1980).

In conclusion, most of the recent computational approaches to the matching problem emphasize the use of suitable constraints to eliminate false matches among simple primitives. For "unspecific" primitives, the question of how to match is, of course, critical. An alternative approach is to rely on very specific primitives (as a large set of independent measurements at each point of the image): In this way, the false matches problem may be avoided in most situations, without any need for global constraints. Further on we discuss again this dichotomy of approaches to the matching problem from the perspective of the physiology of stereoscopic vision.

THE PROBLEM OF "STRUCTURE FROM STEREOPSIS" The human visual system can recover the three-dimensional shape of objects and estimate their absolute distance from binocular disparities. Horizontal disparities, which are

the carrier of stereoscopic depth perception, are themselves uninterpretable without information about eye position. Knowledge of the interocular separation and of the convergence angle of the eyes is needed to recover the quantitative 3-D structure of a scene from horizontal disparity alone. In artificial systems, the distance between the two "eyes" and their angle of convergence are known precisely (camera model). In biological systems, some absolute depth information has generally been assumed to be extraretinal in origin, most likely derived from the convergence angle of the eyes (Foley 1980). Although this information may indeed be useful, its effective availability and its precision are unknown.

In principle, however, a binocular system that takes into account vertical as well as horizontal disparities can solve the interpretation problem without recourse to extraretinal sources of information. This has been well known to photogrammetrists for several years, but its relevance to human stereopsis was pointed out only recently (Longuet-Higgins 1981). The main result is that if the correspondence problem has been solved for seven points, the 3-D structure can be recovered uniquely (unless the seven points are transversed by two planes with one plane containing the origin or by a cone containing the origin; see Tsai & Huang 1981). Furthermore, there exists a remarkably simple approximate method of deriving the 3-D parameters of a planar surface from the horizontal and vertical disparities of a small number of corresponding points (Mayhew & Longuet-Higgins 1982). The question of whether the human visual system does measure vertical disparities and uses them for estimating distance (instead of simply correcting for them with eye movement) is still open. On the other hand, since vertical disparities are quite small even at short fixation distances and eccentric locations, their accurate measurement may turn out to be unattainable; in addition, the main psychophysical support for the idea—the Ogle-induced effect—has been criticized (Westheimer 1978, Gillam et al 1983) as unreliable and quantitatively in disagreement with the prediction of the Mayhew & Longuet-Higgins (1982) hypothesis. The question of whether vertical disparities are used to recover the 3-D structure is important also for its potential implications for the correspondence problem. If vertical disparities are to be measured precisely, matching may become a genuine two-dimensional problem, instead of a one-dimensional process along scan lines.

Theories of Stereoscopic Matching

The following outline of a few of the current computational approaches to stereopsis takes into account only algorithms that are more directly related to human stereoscopic vision. There is no theory encompassing all aspects of stereopsis; the algorithms described here mainly focus on the specific problem of stereomatching.

1. First, we mention a model proposed by Sperling (1970) based on correlation between two gray-level images. The model was never implemented, but

performance of gray-level correlation schemes is well known to be quite limited. Sperling's work does, however, make an interesting point of the connection between stereopsis and vergence movements. His neurophysiological theory of stereopsis includes several interesting notions, such as inhibition among cells tuned to different disparities and detectors for fine and coarse stereopsis.

2. Another model that has not yet been cast in the form of a precise algorithm is Julesz's dipole model (1971). Each position on each retina is associated with a magnetic dipole whose polarity is determined (in the case of random-dot stereograms) by the retinal intensity value. Spring coupling between the tips of adjacent dipoles implements the continuity rule and, of course, the ordering constraint. The orientation of a dipole represents a disparity value, and the fact that each dipole can have only one orientation at a time provides an implementation of the uniqueness rule. Taken literally, this model would correspond to a scheme in which disparity at each position is signaled by the rate of firing of a single neuron. The model exemplifies a matching scheme based on nonspecific primitives (intensity values). False matches are therefore a real problem, and to cope with them, "global" constraints are enforced (via the magnets and the springs).

3. We consider now a cooperative algorithm that was devised by Marr & Poggio (1976) (the approach was first tried by Dev 1975; see also Nelson 1975, 1977, Marr et al 1978), and which is successful at solving random-dot stereograms. The algorithm, which embodies some of the features of Julesz's dipole model, requires a 3-D network of nodes or "neurons," each of which lies at the intersection of a line of sight from each image. The algorithm implements the uniqueness constraint in that the nodes lying along a given line of sight strictly inhibit each other. To implement the continuity constraint, each node excites its immediate neighbors at the same disparity. In the unnatural case of a random-dot stereogram, each pixel of the stereogram can be made to correspond to one line of sight and will have one of two distinct intensities (black or white). The algorithm does not specify the type of matching primitives to be used, whether values in images filtered through center-surround receptive fields, or the more complex features envisioned by the "primal sketch" (Marr 1976), or a large set of measurements at each point in the image. The algorithm is only concerned with a simple cooperative implementation of the basic matching rules.

The range of effectiveness of this algorithm can be extended to natural images by transforming the images to obtain better primitives than the intensity values themselves. By convolving the image with a center-surround filter and representing the sign of the resulting values, it is possible to convert natural scenes into patterns that bear a striking resemblance to Julesz's binary random-dot stereograms (the sign of this convolution array is equivalent to the zero-crossing map). The cooperative algorithm can now operate on this binary image exactly as on random-dot stereograms and extract the correct disparities.

4. Marr & Poggio (1979) later proposed a rather different algorithm specifically motivated by human psychophysics. It exploits the intrinsic possibility offered by sparse matching primitives like zero crossings at different spatial resolutions. At low spatial resolutions, zero-crossings of a given sign are quite rare and never too close. Consequently, false matches, that is, matches between noncorresponding zero-crossings, are essentially absent over a relatively large disparity range. From a mathematical analysis of the probability of occurrence of zero-crossings in bandpass images, it turns out that if the disparity range that is considered is on the order of the size of the receptive field used to filter the image (more precisely, the diameter of the field's center), false matches are virtually absent. These observations led to a different algorithm for solving the correspondence problem. In its simplest version, the algorithm matches zero-crossings of the same sign in image pairs filtered with receptive fields of three or more different sizes. These ideas lead to several algorithms with somewhat different properties: for instance, matching could take place between the binary function of the filtered images over a certain area (H. K. Nishihara, in preparation).

The main aspect of this theoretical framework is its avoidance of the false matches problem by trading off resolution with disparity range in a coarse-to-fine strategy. [Moravec (1980) was the first to use a matching scheme based on a search from the lowest to the highest resolution]. A computer implementation of this algorithm by Grimson (1981) performs satisfactorily on random-dot stereograms as well as on natural images, showing some of the properties of human depth perception. Psychophysical data against some of the detailed predictions made by Marr & Poggio (1979) have been extensively discussed by Mayhew & Frisby (1981). For instance, vergence movements are not driven only by matches in the largest channels; simpler measurements of disparity are used as well (e.g. based on monocular contours). In addition, the role of eye movements and the independence of the matching at the various resolutions are apparently inconsistent with some of the results reported by Mayhew & Frisby (1979b) and by Julesz & Schumer (1981), for instance, concerning the size of Panum's area for filtered stereograms in the absence of eye movements. Though the theoretical framework may still be correct in its main lines, it is already clear that the scheme of Marr & Poggio will have to be modified in several respects.

One of the main weaknesses of the algorithm is its sensitivity to vertical disparities. Vertical disparities can be well tolerated up to a fraction, about 1/4 of the filter size. Interestingly, human performance deteriorates quite rapidly with vertical disparities in the absence of eye movements, almost as much as predicted by the model (Nielsen & Poggio 1983).

5. An important extension of the Marr-Poggio (1979) theory has been proposed by Mayhew & Frisby (1981). The main modifications they suggest

are (*a*) concerning the matching primitives: in addition to zero-crossings, Mayhew & Frisby advocate the use of peaks in the convolution values, both because of psychophysical evidence and for computational reasons; (*b*) concerning the matching rules: instead of binocular matching performed independently on each spatial frequency channel, Mayhew & Frisby propose a matching rule which takes into account simultaneous correspondences in all channels. In addition, continuity of zero-crossing segments to be matched is also explicitly used and thus the matching primitives are oriented, continuous zero-crossing segments (as in the neural implementation suggested by Marr & Poggio 1979). Computational experiments support the usefulness of these modifications for solving correspondence problems in random-dot stereograms. Extended tests on natural images are not yet available and would be highly desirable.

6. Several other algorithms of stereopsis have been developed without a direct concern with human stereopsis (e.g. Arnold 1982, Gennery 1980). Among them, the most interesting one has been developed by Baker (1980, 1982) and Baker & Binford (1981). Its matching primitives are "edges" (similar to zero-crossings) at various spatial resolutions. In addition to position, contrast orientation and intensity to either side of the edges are used as matching primitives, and provide a weighting factor for the correspondence process. The matching process takes place on epipolar lines (assumed known and parallel) and exploits the ordering constraint (corresponding edges in the two images have the same order). It starts with low resolution edges to bring the two images into rough correspondence. A successive refinement in resolution brings finer details into the analysis. The matching performs the optimal correspondence on each epipolar line. A global constraint is then applied, checking for edge connectivity: a connected sequence of edges in one image should be seen as a connected sequence of edges in the other. Finally, intensity values are matched between matched edges.

7. To recapitulate: The main matching rules proposed in the cooperative algorithm and used by the Marr-Poggio theory (1979) have been refined and extended by Mayhew & Frisby (1981) and by Baker & Binford (1981). Cross channel correspondence, figural continuity, and edge connectivity are the main additions to continuity and uniqueness of matching. Computer simulations suggest that these rules may be advantageous but experiments must be conducted to find out whether they are used by our own visual system.

All these theories agree that "edges" are the main matching primitives. This is consistent with the neurophysiological observations that binocular cortical cells respond to oriented edges. The thrust of the Marr-Poggio theory (1979) is that simple primitives of the zero-crossing type, together with a coarse to fine strategy, already go a long way in avoiding the false matches problem, without disambiguation processes. More complex and specific primitives, as pointed

out by Barlow et al (1967), would clearly show an even simpler false matches problem and allow a larger Panum's fusional area. A step further, one may consider whether a larger set of independent measurements on each image point, possibly including zero-crossings, directional derivatives, etc, may be used by the visual system to avoid almost completely the problem of incorrect matches.

All theories discussed above consider the problem of full stereo matching, for the extraction of a precise disparity map. Simpler and more primitive schemes could be used, and are probably used by the visual system, when the only information that is needed is a rough estimate of convergent or divergent depth (see section on Vergence).

PSYCHOPHYSICS OF STEREOPSIS

Primitives and Constraints of Matching

So far, psychophysics does not tell much about the nature of the primitives used in stereoscopic matching. We know they are not intensity values themselves or complex micropatterns of intensity values (Julesz 1971). Available evidence is consistent with the idea that zero-crossings, as described above, are among the matching primitives. In particular: (*a*) stereograms in which the two images have different contrast (of the same sign) can be easily fused (Julesz 1971), (*b*) if the contrast of one half-pair is reversed, fusion is impossible (Julesz 1960), and (*c*) fusion survives blurring of both or just one of the two images. The experiments of Mayhew & Frisby (1981), on the other hand, suggest that zero-crossings alone cannot explain the perception of stereograms composed of a triangle ramp grating paired with a phase-varying ramp grating. They have proposed that peaks in the channel convolutions are also used by the process of stereo matching. The idea, presented above, that several independent measurements, possibly including or being in part equivalent to zero-crossings and peaks, are used in the matching process has still to be tested.

The constraints that the matching process obeys are only indirectly explored with psychophysical techniques. The continuity and ordering constraint is consistent with the "nail illusion" (Kroll & van der Grind 1980) and with the disparity gradient limit found by Burt & Julesz (1980a,b). Convincing evidence is, however, lacking. On the whole, the question of the nature of the primitives and constraints of matching waits for new psychophysical experiments.

Vertical Disparity, Registration, and Eye Movements

Vertical disparity arises naturally because of eye misalignments and because of the stereo geometry. In the latter case, vertical disparity contains valuable information about the parameters of gaze. Can the visual system measure the corresponding small vertical disparity and use it to compute the "camera"

parameters? As a separate question, how does the matching process cope with perspective deformation, and in particular, with vertical disparity?

One may expect that eye movements would perform the kind of registration process that must precede full, high-resolution stereomatching. The psychophysical evidence is consistent with this view. In particular, Duwaer (1981) and Duwaer & van der Brink (1981, 1982) have argued that the sensorimotor system eliminates vertical disparities with high accuracy: Vergence eye movements can be initiated by vertical disparities that are much too small to induce diplopia. Residual vertical disparities that can be fused are in the range of 8′–15′. The stimuli Duwaer used did not pose, however, any false match problem. The interesting question concerns the maximum range of vertical disparities that allows fusion without eye movements in the presence of potential false matches. Recently, Nielsen & Poggio (1983) found that for briefly flashed random-dot stereograms about 4′–7′ of vertical disparity represent the fusion limit. Because of the small value, this finding may be consistent with a simple one-dimensional matching scheme having a built-in tolerance to small vertical disparities. These data have a clear implication: A registration process based on eye movements must precede the stereo matching process to remove vertical disparities larger than about 4′–7′. Some questions immediately arise about this registration stage; for instance, which measurements on the image are used to control the registration process? These measurements must be simpler and coarser than what is involved in the full matching process. Otherwise, in order to register and control the eyes the visual system would have to solve a 2-D correspondence problem.

Vergence

Given the obvious importance of eye movements in stereopsis, it is surprising that so little is known about the role of vergence. In most computational schemes for stereo, the false matches problem is simplified by considering only those potential left/right fusions with disparities under a certain limiting size (defined by Panum's area). This disparity limit requires that when disparities greater than Panum's area are presented, a suitable vergence change must be initiated to bring corresponding features within the allowed range for fusion. As we mentioned, it is critical to know the nature of the measurements made on the image that guide these vergence movements. Furthermore, one would like to know how vergence is controlled and maintained.

It is long known that vergence movements are independent of conjunctive eye movements, are smooth rather than saccadic, are accurate within 2′ of disparity, and are continuously controlled by a disparity signal (Jones & Kerr 1972, Rashbash & Westheimer 1961(a,b), Riggs & Niehl 1960, Westheimer & Mitchell 1969). Marr & Poggio (1979) conjectured that matches between large receptive fields (that can cope with disparities of the same order of the size of

the receptive field) initiate vergence movements bringing into correspondence higher spatial frequency channels mediating smaller disparities. Recently, it has become clear that this is not the only mechanism that can trigger correct vergence movements. Kidd, Frisby & Mayhew (1979) have shown that texture contours can initiate correct vergence eye movements even for disparities larger than the ones predicted by Marr & Poggio. Later, Mowforth, Mayhew & Frisby (1981) claimed that high-frequency filtered stereograms could initiate vergence movements well outside the range of the largest receptive field presumably active. The conclusion is inescapable that not only matching of zero-crossings but some other more primitive measurements of disparity control at least the direction of vergence movements. Possibly, prominent monocular cues (like strong edges in some, or all, of the channels) could drive eye movements (Mowforth, Mayhew & Frisby 1981; see also Marr & Poggio 1980). For isolated and diplopic cues, there is no false matches problem and no reason that vergence movement and rough sensation of depth could not be provided even for large disparities.

For dense textures, there are two basic possibilities: (*a*) Either a similar mechanism works on prominent monocular features of the images (for instance, a few sparse strong zero-crossings, possibly coincident in all channels), or (*b*) all convergent matches are weighted against all divergent matches (false and correct) over relatively large areas. This scheme may correctly find the overall sign of disparity, but fails to identify the shape of a stereo figure. Moreover, the disparity limit in this case may depend on the size of the figure (see Marr & Poggio 1980). Interestingly, Tyler & Julesz (1980) have reported that detection ability depends on the square root of the area.

Spatial Frequency Channels in Stereopsis

Several data suggest that monocular detection of targets relies upon a set of independent channels of different coarseness—probably corresponding to receptive fields of different sizes. That similar independent channels may be used in stereopsis was demonstrated by Julesz & Miller (1975), who found that masking noise disrupts stereopsis only if it has a spatial spectrum overlapping that of the stereogram. These findings, as well as the work of Frisby & Mayhew (1977), provide the basis for the assumption that binocular matches are made on independently filtered images. The natural consequence of the assumption is that coarse channels encode large disparities whereas fine channels can match only small disparities (without eye movements). Psychophysical observations by Felton et al (1972), Kulikowski (1978), and Levinson & Blake (1979) support this view and are consistent with the physiological observation that neurons with small receptive fields often have a narrower range of disparities than neurons with large receptive fields (see below).

Mayhew & Frisby (1978b), however, found that the contrast threshold for

stereopsis does not depend on the independent contrast of different spatial frequency components. The contrast threshold for stereopsis is higher than for monocular detection of narrow-band filtered, random-dot noise at all spatial frequencies by 0.3 to 0.4 log units (Frisby & Mayhew 1978). This may reflect processes at the level of 2 1/2-D sketch of Marr & Poggio (1979), the dynamic buffer where successful matches are stored. More likely, however, it supports the idea that different channels are combined in the matching process (the cross-channel correspondence rules of Mayhew & Frisby 1980). More evidence in this direction has been provided by the same authors (with the "missing fundamental" experiment and with spatial frequency filtered stereograms portraying corrugated surfaces). It seems, therefore, that the strictly independent channel model of Marr & Poggio has to be modified, allowing for some interactions between the channels at the level of disambiguating ambiguous matches. An alternative possibility is the coupling of the channels, not simply via eye movement, but also via a hypothetical cortical shift mechanism.

An important property of the binocular spatial frequency channels is that according to Mayhew & Frisby (1978a, 1979a) they are not orientation selective, a conclusion consistent with the notion that the matching primitives are extracted from the images filtered through center-surround receptive fields. This is not in disagreement with the orientation tuning of disparity selective cells in the cortex since the matching primitives themselves, such as the zero-crossings, can be locally oriented. Oriented receptive fields, however, cannot be completely ruled out on the basis of these experiments alone.

Panum's Area

There is a vast literature, and a correspondingly large amount of data, on Panum's fusional area, and almost as much disagreement on its properties and even its precise definition.

Panum's area represents the total range of disparities that can be fused about the fixation distance without the help of eye movements. The values reported for line stereograms are typically around 10′ for the maximum amount of convergent or divergent disparity without diplopia; the extent of Panum's fusional area is, therefore, twice this (Mitchell 1966). Comparable values have been reported for vertical disparities (Schor & Tyler 1980).

While these data refer to tachistoscopic presentations, Fender & Julesz (1967) measured the limits of fusion under stabilized image conditions, and found that fusion occurred between line targets at a maximum disparity of 40′. The corresponding value for random-dot stereograms was 8′ for horizontal disparities and 10′ for vertical disparities (for the onset of fusion). Fusion, once established, could be maintained over a much larger range. Diner (1978) replicated the experiments of Fender & Julesz but failed to obtain foveal

diplopic thresholds beyond the classical limit of 20′ for binocularly stabilized bars.

All these data show a relatively small Panum's area with a size comparable to the maximum size of the receptive fields operating in foveal vision (around 20′ in human according to Wilson & Bergen 1979). However, recent data obtained with dynamic random-dot stereograms imply that Panum's area can be larger. In particular, Schumer & Julesz (1980; see also Julesz & Schumer 1981) have reported that plus or minus 50′ disparities can be successfully fused. Mayhew & Frisby (1979b) observed that convergent disparities could be fused in a flash up to 18′ of disparity: At this value, the latency for discrimination (longer than 180 msec) indicated the need of vergence movements. With filtered stereograms, disparities larger than twice the values predicted by the Marr-Poggio (1979) independent channel algorithm can be successfully fused (Mayhew & Frisby 1979b, Schumer & Julesz 1982). A critical prediction of this theory is that the maximum fusable disparity should scale with the spatial frequency of the stimulus, since the lower spatial frequencies will be detected only by the larger channels. Julesz & Schumer (1981) concluded that there is probably an association between large disparities and coarse channels, but that the complementary association between small disparities and fine channels lacks experimental support.

The presence of Panum's area seems to be the simplest explanation for the persistence of single binocular vision, even though there are constant changes in retinal image disparity caused by ocular drifts and saccades (2′–4′; St. Cyr & Fender 1969). In this view, the disparities for which Panum's area compensates are dynamic. Schor & Tyler (1981) studied the temporal properties of Panum's area and found that for slow changes in disparity (0.2 Hz), the area increasing horizontally by as much as a factor of ten, but remaining basically unchanged vertically, compared with fast disparity variations (5 Hz). They were, however, unable to obtain Panum's areas much greater than 20′ and thus to approach the large extended fusional ranges reported by Fender & Julesz (1967) under stabilized image conditions.

Tyler (1975) and Schor & Tyler (1981) found that Panum's area was reduced with increasing spatial frequency variations of disparity (both for horizontal and vertical directions). Interestingly, Panum's area dependency on temporal modulation of disparity holds only for low spatial frequency.

In order to interpret correctly measurements of Panum's area obtained with line stereograms and with random-dot stereograms, it should be kept in mind that the mechanisms of rivalry and fusion may not coincide. Thus, different measured ranges for diplopia and for fusion may reflect two different underlying processes (Duwaer 1981). Furthermore, successful depth discrimination, especially in a forced-choice paradigm, may not need a solution of the corre-

spondence problem for all matches; a high enough proportion of correct, possibly sparse, matches could be sufficient.

Stereoacuity

The impressive quality of stereoacuity would seem to impose quite restrictive requirements on the mechanisms subserving stereopsis. The threshold for depth discrimination is only a few seconds of arc, which is much less than the size and the separation of foveal photoreceptors. In addition, stereoscopic acuity does not suffer when targets move laterally or in depth by five or more minutes of arc during a short 200-msec exposure (Westheimer & McKee 1978, 1980a). Stereoacuity has also its limitations, however: Synchrony of presentations of left and right image and absence of a standing disparity (of even one minute of arc) are necessary requirements for small discrimination thresholds (Westheimer 1979). In addition, blur of the images decreases stereoacuity more than ordinary visual acuity: For optimal stereoacuity, the full representation of the spatial frequency range admitted by the eye's optics is needed (Westheimer & McKee 1980b).

All these results refer to tachistoscopic presentations of line stereograms. One may ask how stereoacuity is affected by the potential presence of false targets as in a random-dot stereogram, since Westheimer reports that "crowding" of features decreases stereoacuity even with line stereograms. The available data are contradictory on this point. Harwerth & Rawlings (1977) report relatively high acuity thresholds for random-dot stereograms presented in a flash. Uttal et al (1975) find comparable values at 75% correct answer threshold. Lower thresholds have been measured for dynamic random-dot stereograms by Schumer & Ganz (1979). Recently, K. R. K. Nielsen (in preparation) has also measured thresholds as low as 30″ for dense random-dot stereograms presented in a short flash of 150 msec. Thus, the potential problem of false targets does not seem to affect dramatically stereoacuity. Interestingly, stereoacuity thresholds depend on exposure duration and are higher for discriminating the form of the (cyclopean) figure than for simply discriminating depth (Harwerth & Rawlings 1977).

Westheimer (1979) has suggested that a differencing mechanism measuring relative disparities is at the basis of stereoacuity. For two points in a stereo target, depth discrimination would be accomplished in terms of the difference of the individual disparities, a value that is independent of ocular stability. More recent experiments by G. Westheimer and G. Mitchison (personal communication) suggest that in a stereogram of several lines, the perceived disparity of a given line is determined by an average of its disparities' differences with respect to the flanking lines. As a consequence, if the lines lie exactly on a tilted plane, any given line (not at the borders) is perceived at the

same depth as the left and right flanking lines. It is clearly important to check whether this computation of relative disparity takes place only for small disparities (in the range of stereoacuity) or also for large ones. The underlying computational reasons are still unclear, although evaluation of relative disparities—actually changes in disparities—would give the stereosystem a useful tolerance to eye movements. The precision of stereoacuity requires that the monocular matching primitives are localized with the same level of relative accuracy in the image delivered by each eye. It is possible that the interpolation process underlying monocular hyperacuity—as it is revealed by vernier detection tasks—is used as an input to stereoacuity. In this view, the patterns of activity transmitted by the left and the right retinae would first be interpolated independently, localizing features such as zero-crossing with precision of a few seconds, much less than the spacing and the diameter of the foveal cones. These features would then be matched between the two eyes. Absolute disparities could not be measured reliably because of involuntary eye movements, but differences in disparity between neighboring features could be evaluated with an accuracy of a few seconds of arc.

On the other hand, it has been recently shown that vernier cues need not be input to the stereo matching process (Nishihara & Poggio 1982), as proposed earlier by Julesz & Spivack (1967). The successful fusion of random line stereograms with vernier breaks between the lines as small as 15″ can be explained in terms of a matching process operating on the coarse structure of these stereograms as extracted by relatively large receptive fields, without recourse to interpolation and detection of the vernier breaks.

Disparity Gradient

Panum's area is an upper limit to the disparities that can be fused without eye movements. Recently, evidence has been accumulating that in addition to disparity magnitude, the rate of change in disparity (the disparity gradient) also sets a limit to stereo fusion. There is in most situations a physical upper limit to the maximum disparity gradient. When one continuous surface changes in depth, one of the two eyes may not see all of the surface: this starts to happen when one line of sight grazes the surface. In situations of this type (which correspond to the Panum limiting case) human vision has difficulties and stereopsis is usually diplopic.

Tyler (1973, 1974, 1975a,b) was the first to observe that there is a limit in the rate of change of disparity across the retina both for line stereograms and for random-dot stereograms. Burt & Julesz (1980a,b) measured the disparity gradient limit of pairs of points at different orientations defined as their total disparity divided by their "binocular" separation. This definition implies that the disparity gradient is $d' = 2$ for the Panum case. They find a value for d'

around 1, roughly independent of orientation for the limit between fusion and diplopia.

The critical question, of course, is the nature of this constraint. Is it a constraint used in matching or is it a property of later processes, for instance, of the way depth is represented? As argued by Baker (1982) and by Mayhew (1983), stereo projection almost always preserves the order of successive pairs of matches in each eye and thus it seems natural to implement this constraint during the matching process. The ordering constraint, however, would correspond to d' less than or equal to 2 and is therefore only indirectly implemented by the human visual system.

There may be additional, nonexclusive reasons for the disparity gradient limit. A large disparity gradient locally compresses and rotates one of the two images relative to the other: thus, the same line segments will differ in orientation and length in the two images and these differences may by themselves make the matching impossible. The results of filtering the two images for extracting the matching primitives (for instance, by ganglion cells' receptive fields) may give different results for high disparity gradients, depending on the size of the operator and of course, on the pattern itself. In other words, the matching features extracted from the monocular patterns may be affected by the disparity gradient. Braddick (1979) has found that the geometrical relations of different parts of the pattern presented to each eye may influence binocular single vision.

Stereopsis by Binocular Delay

Morgan (1975, 1976), Ross & Hogben (1975), and Tyler (1977) studied a Pulfrich-like effect in stereopsis that has some fascinating implications about the spatiotemporal interpolation performed by the visual system. The new observation was that a Pulfrich effect—dimming the view by one eye—still occurs for an object in apparent motion produced by flashing it stroboscopically at a sequence of positions. Burr & Ross (1979) were then able to show that binocular delay of a spot of light moving stroboscopically against a background of dynamic noise yields a vivid stereoscopic depth. Temporal delay of one image seems equivalent to binocular disparity and has an equivalent threshold of a few seconds of arc. The simplest explanation is provided by another (monocular) experiment by Burr (1979): If two line segments are displayed sequentially at a series of stations, an illusory displacement occurs if the line segments are aligned in space but are displayed with a slight delay in one sequence relative to the other. Not the actual portions of the line segments, but the positions estimated by a temporal interpolation process are displaced relative to each other. From this experiment, it seems that the spatial pattern of activity between the flashes is actually reconstituted. This monocular reconsti-

tuted pattern of activity could be used by stereopsis and could satisfactorily explain the results of Burr & Ross. An apparent implication of this view is that at least some binocular cells should have access to interpolated monocular activity, reconstructed between the positions at which the stimulus was presented stroboscopically.

Motion and Color in Stereopsis

It is known that contours of discontinuities in a motion field can be matched stereoscopically. Thus, primitives more complex than simple edges or zero-crossings can be used in stereomatching. A more general question is whether motion can interact with stereopsis at the level of matching (helping to disambiguate false matches) or at the level of structure from stereopsis.

On the whole, the interaction of motion with stereopsis still requires an analysis at both the computational and the experimental levels. Similarly, not much is known about color information and its use in the process of stereoscopic matching and stereofusion (but see Julesz 1978).

Development of Stereopsis

Recently, a number of investigators in different laboratories have been able to measure the onset of binocular vision and the development of stereoacuity in infants. These data and their comparison with developmental studies of binocular cortical mechanisms in cats and monkeys promise to provide some important information about the neural organization of stereopsis. Coarse stereopsis appears around the fourth month of age; at this time stereoacuity (fine stereopsis) begins to increase very rapidly, approaching adult levels around the fifth month after birth (Fox et al 1979, Held et al 1980, Petrig et al 1981). There is some evidence that stereoacuity for crossed disparity develops earlier than, but approximately at the same rate as, the uncrossed disparities (Birch et al 1982). This finding is consistent with the notion that at least two stereoscopic mechanisms exist in the adult (Richards 1970, 1971). The rapid development of stereoacuity contrasts sharply with the development of monocular visual acuity for gratings, which increases rapidly during the first six months of age and then slowly over the first few years (Teller 1981). Held (1983) has conjectured that segregation of the ocular dominance columns in layer IVc of the primary visual cortex is responsible for the onset and the development of stereoacuity.

NEUROPHYSIOLOGY OF STEREOPSIS

While psychology and psychophysics provide knowledge of its functional properties and computational capacities of stereoscopic vision, neurophysiology must attempt to unravel the extent and nature of the "cyclopean brain," the processing neural network upon which depends the vivid and rich three-

dimensional visual world we may enjoy. The term "cyclopean," after the one-eyed giants, foes of Ulysses, aptly suggests the egocentric one-image visual world that we usually perceive, even though the hapless, monocular Cyclops never experienced stereoscopic depth.

In preceding sections of this article, we have described the psychophysical evidence showing that binocular positional disparity provides the most important and accurate information, indeed the only necessary one, for stereopsis. While it may be argued that "stereopsis constitutes only a fraction of the information upon which depth perception is based" (DeValois & DeValois 1980), this fraction is undeniably of some importance, for it alone makes the monocular and binocular experiences of depth so dramatically and fundamentally different.

The analysis of the neural mechanisms of stereopsis must begin with an attempt to discover where and how binocular disparity is derived and measured by the cells in the brain, and how it is signaled in neural code. In primates, and generally in higher mammals, the region of brain where the inputs from the two eyes converge in the region of cerebral cortex is known as "area striata," area 17 of Brodman, the primary visual cortex, or V1. In the macaque monkey, nearly all visual cortical neurons are binocular. Binocular neurons have two receptive fields, one for each eye, and their activity reflects the dynamic interaction of excitatory and inhibitory influences from each eye; thus, the responses of these neurons to binocularly viewed patterns may vary markedly depending on the receptive field characteristics and on the spatio-temporal configuration of the stimulus in the two eyes.

Neuron Sensitivity to Positional Disparity: "Local" Stereopsis

Barlow, Blakemore & Pettigrew (1967) and Nikara, Bishop & Pettigrew (1968) first presented evidence that neuronal stereoscopic processing can be identified in the primary visual cortex of the cat based on the phenomenon of receptive field disparity. These studies, conducted with isolated line/bar stimuli, showed that the receptive fields of binocular cortical neurons subserving central vision could be in exact spatial correspondence in the two eyes, or could have different relative positions, some field pairs having convergent disparities, others, divergent disparities. These disparity sensitive cells might well play an important role in stereoscopic depth perception because, with normal convergent fixation of binocular vision, they would be selectively activated by objects at different relative depths.

The properties of disparity sensitive neurons in the visual cortex of the cat, and more generally of the neuronal characteristics of binocular interaction, were analyzed in some detail by Bishop and his collaborators (Pettigrew et al 1968, Joshua & Bishop 1970, Bishop & Henry 1971, Bishop et al 1971). These studies, together with more recent ones by von der Heydt et al (1978), Fischer

& Krueger (1979), and Ferster (1981), have confirmed and extended the original observations. Disparity sensitive neurons have been found also in the visual cortex (V1 and V2) of the sheep (Clarke et al 1976), and in the Wulst of the owl (Pettigrew & Konishi 1976).

The stereoscopic properties of cortical neurons of the macaque monkey, an animal whose discriminatory capacities for depth are very similar to those of humans (Bough 1970, Sarmiento 1975), were first studied by Hubel & Wiesel (1970) in anesthetized and paralyzed preparations. These authors were unable to identify disparity sensitive neurons in area 17, but did observe numerous neurons in prestriate area 18, or V2, which were termed "binocular depth cells" for they were uniquely sensitive to simultaneous stimulation of the two eyes, some neurons having receptive fields in exact binocular correspondence, others disparate fields. On the basis of these observations and of similar ones they made in the cat, Hubel & Wiesel (1970, 1973) suggested that the elaboration of stereoscopic mechanisms occurs outside the primary visual cortex.

Studies in alert and visually attentive macaque monkeys under conditions of normal binocular vision, unequivocally demonstrated that a large number (60–70%) of neurons in striate cortex, and an even higher proportion in prestriate cortex, are sensitive to horizontal disparity, and that different types of depth neurons exist, many with the same binocular properties as those described by Hubel & Wiesel (1970) in area 18 (Poggio & Fischer 1977, Poggio & Talbot 1981, Poggio 1984). An account of the findings of those studies is given in that which follows.

Response profiles of single cells in area 17 and 18 were obtained for binocular stimulation with isolated line/bar patterns of optimal size and orientation for the cell under study, and presented in real (Poggio & Fischer 1977), or simulated (dichoptic) depth (Poggio & Talbot 1981). Under these stimulus conditions, in which false matches are not present, two main types of stereoscopic neurons may be recognized: 1. A group of neurons are disparity selective over a limited and often narrow range; excitatory binocular facilitation (*Tuned excitatory neurons*) and, less frequently, suppressive interaction (*Tuned inhibitory neurons*) may be observed. 2. A second type of stereoscopic neurons includes cells with a reciprocal selectivity for crossed and uncrossed disparities: One set of neurons (*Far neurons*) gives excitatory responses to objects farther than the points of fixation and inhibitory responses to nearer objects. Other neurons (*Near neurons*) have the opposite behavior: excitation for nearer objects and inhibition for farther ones. About one third of the neurons in A17, and fewer still in A18, are not disparity selective and give similar responses to stimuli of different disparities within the disparity range of binocular interaction (*Flat neurons*).

TUNED NEURONS Approximately one half of the disparity selective neurons are Tuned excitatory (Figure 2, left). In the area of striate cortex subserving

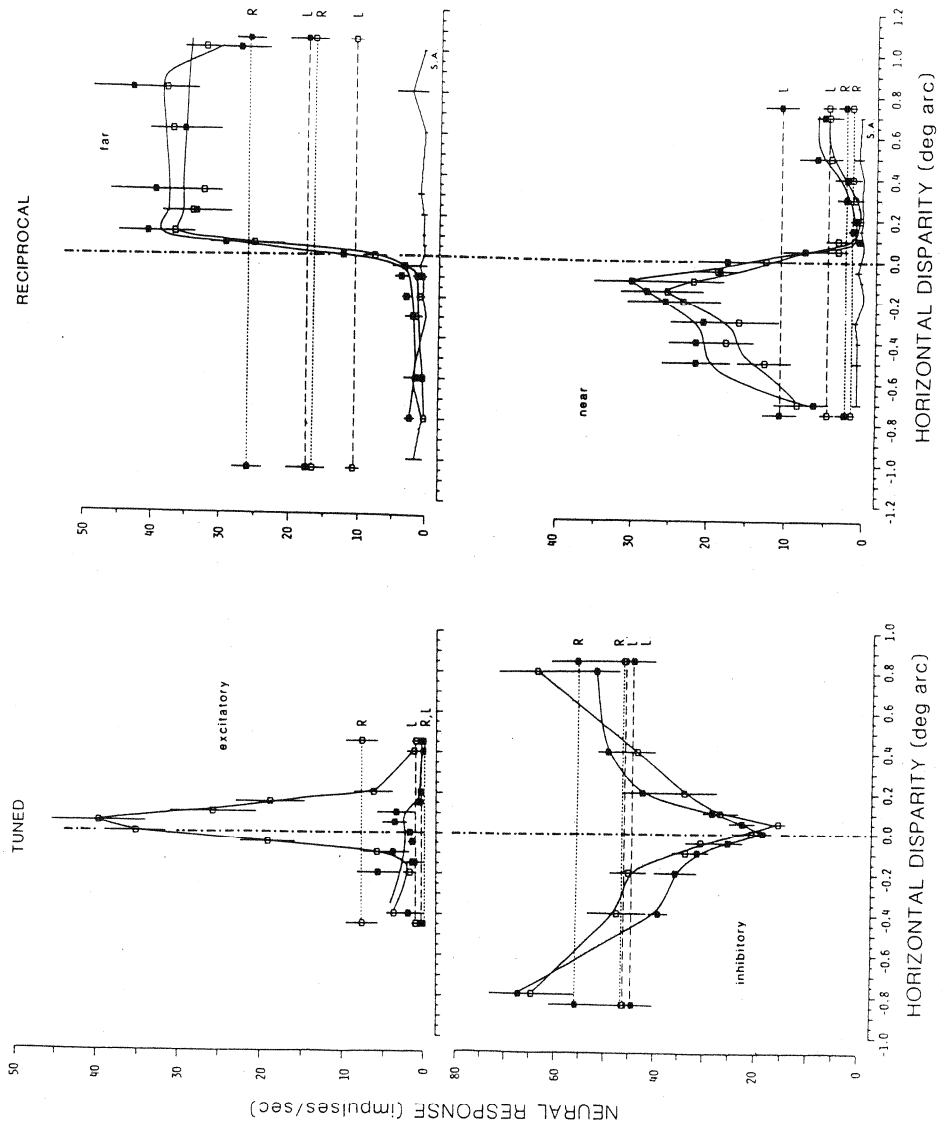
central vision, peak disparity sensitivity of these neurons is very seldom found outside 12′ of crossed or uncrossed disparity, with most neurons maximally excited by disparities within ± 6′. The width of the disparity sensitivity curve for a group of these neurons (taken at the response level of maximal response ÷ $\sqrt{2}$; Schiller et al 1976) was 10′ ±4 ′(SD) of disparity. (Poggio, 1980, 1984). Frequently, tuned excitatory neurons are directionally selective and many are strictly unidirectional. Most characteristics of depth tuned cells are neurons, found both in striate and prestriate cortex, whose properties appear to be the same as those of the "binocular depth cell" of Hubel & Wiesel (1970). These cells do not respond at all, or only minimally, to monocular stimulation over a very narrow range of disparities at or about zero disparity. No monocularly silent neurons were ever found with other types of stereoscopic properties.

Tuned inhibitory neurons have disparity response profiles opposite to those of the tuned excitatory neurons, in that their binocular responses are suppressed within the same narrow range of disparities near the horopter over which tuned excitatory neurons are binocularly facilitated. For foveal neurons, maximal suppression occurs within ± 6′ of disparity and response facilitation at larger crossed and uncrossed disparities is often observed.

RECIPROCAL NEURONS Neurons with the second main type of stereoscopic properties, the Near and the Far neurons (Figure 2, right), give excitatory responses over a range of disparities of one sign and inhibitory responses over a similar range of opposite sign. For many of these neurons, disparity sensitivity for bar stimuli extends over a range of one degree or more on either side of the horopter; other reciprocal neurons have a narrower, S-shaped depth response profile. Both groups, however, have common properties characterized by a steep response gradient from maximal excitation to maximal inhibition with the mid-point of response activity at, or very close to, zero disparity.

The distribution of stereoscopic types among binocularly "simple" and "complex" cells appears to be remarkably similar, and no functional relationship was observed between the simple/complex classification and the tuned/reciprocal types of depth sensitivity for bar patterns (Poggio 1984). Similar conclusions have been reached by Ferster (1981) for the visual cortex of the cat. Stereoscopic neurons typically signal depth with little response uncertainty; the response to the preferred excitatory or inhibitory disparity is consistently stronger than the response to any other disparity, irrespective of any intrinsic response variability. Thus, under normal binocular vision, the mechanisms of binocular interaction are remarkably secure and capable of compensating dynamically for the small movements of fixation (Motter & Poggio 1981, 1982, Poggio 1984).

The neurophysiological evidence, both from cats and from monkeys, indicates that the characteristics of the binocular response to isolated line patterns are determined by the spatial organization of the component regions of the


RECIPROCAL
TUNED
far
near
excitatory
inhibitory
R
L
R,L
S.A
NEURAL RESPONSE (impulses/sec)
HORIZONTAL DISPARITY (deg arc)
HORIZONTAL DISPARITY (deg arc)

neuron's receptive fields in left and right eyes, and the dynamic processing of the signals reaching the cortical cell. Maximal responses occur when synergistic regions of the two monocular receptive fields are concurrently engaged, and minimal responses when antagonistic regions are stimulated (Barlow et al 1967, Bishop et al 1971, Ferster 1981). Both simple and complex cells measure retinal image disparity by receptive field disparity and may operate in mechanisms of "local" stereopsis appropriate for detecting unambiguous disparities in sparse and simple contour patterns.

The psychophysics of stereoacuity is not easily interpreted in terms of the known properties of binocular cortical neurons. The threshold of stereoacuity is more than one order of magnitude smaller than the width of tuning of disparity-sensitive cells. Perhaps sets of stereoscopic neurons provide the desired precision by means of interpolation process similar and subsequent to the one postulated to occur in layer IVc-beta for monocular hyperacuity (Barlow 1979, Crick et al 1981).

The existence of different types of depth neurons, Tuned neurons and Near and Far neurons, gives support to the suggestion of Richards (1971) that normal stereopsis is based on the activity of three populations of neurons preferentially activated by crossed, near zero, and uncrossed disparity. This suggestion was based on studies of stereoanomalies (Richards 1970, 1971) showing that some individuals are unable to localize stimuli presented with large (>0.5°) uncrossed disparities (Far), while other individuals are unable to utilize large crossed disparities (Near). Jones (1977) demonstrated that these forms of stereoblindness may be present in subjects who have normal fine stereopsis. The proposition may be advanced that these persons lack either functionally normal Far or Near neurons but possess a normal Tuned stereosystem (Fischer & Poggio 1979).

Neuron Responses to Random-Dot Stereograms: "Global" Stereopsis

The experimental results described above provide evidence that cortical visual neurons signal and measure horizontal binocular disparity in two distinct ways

Figure 2 Positional disparity sensitivity profiles for line stereograms of *Tuned* (excitatory and inhibitory) and *Reciprocal* (far and near) foveal cortical neurons of the macaque. The graphs were constructed by plotting mean response magnitude vs. horizontal binocular disparity of stimuli. For each neuron, the response profiles for the two diametrically opposite directions of stimulus motion perpendicular to stimulus orientation are shown (*open* and *filled square symbols*). The *vertical lines* at each point indicate ±1 SE of mean. The magnitude of the response to monocular stimulation of each eye for each of the two directions are shown by the *horizontal lines* across the graphs, identified at each end by the appropriate directional *square symbols*. *Broken line* = left eye (L); *dotted line* = right eye (R).

(tuned and reciprocal), but do not give information on how the brain handles the problem of false matches that may occur during normal binocular vision.

An attempt to understand this aspect of the analysis of stereopsis was made by studying the response of cortical neurons to dynamic random-dot stereograms. Using this powerful form of "cyclopean" stimulation[1], Julesz (1960) conclusively demonstrated that correct stereoscopic matches occur quite early in the processing of visual information, and that basic stereoscopic operation probably precedes form recognition. Thus, it was reasonable to speculate that the neuronal mechanisms for the resolution of ambiguities and for determining the correct correspondence associated with depth should be present at early stages of binocular interaction. In 20% of a sample of 230 neurons whose depth sensitivity was examined both with line/bar stereograms and with (dynamic) cyclopean random-dot stereograms, the latter evoked evident responses, both from disparity selective neurons as well as from neurons responding equally well to all disparities. In most (but not all) instances, the depth response profiles for line patterns and random-dot patterns were qualitatively the same (Poggio 1980, 1984).

The proportion of cells responding to cyclopean random-dot stereograms is not significantly different among the tuned (excitatory and inhibitory) or reciprocal (near and far) neurons. Certain functional properties of the cell, such as eye dominance and directionality, are qualitatively the same with line patterns as with random-dot patterns. For some cells, however, other properties, stimulus orientation and size selectivities in particular, are different for the two types of stimuli in that the response to cyclopean random-dot patterns may show little or no selectivity for the orientation of the monocularly invisible figure and/or its size.

At variance with what obtains with disparity unambiguous line/bar stimuli, an evident correlation exists between sensitivity to random-dot stimulation and simple/complex receptive field organization in that essentially all cells that responded to cyclopean stereograms are complex cells (Poggio 1984). This finding is particularly significant for it assigns to the complex cell the unique capacity of solving the correspondence problem by responding to the "correct" binocular matches over the receptive fields in the two eyes.

CONCLUSION: TOWARD AN INTEGRATED VIEW

The foregoing account of the results of computational, psychophysical, and physiological analyses of stereopsis shows that in recent years, these three

[1]The term *cyclopean* identifies that form of visual stimulation with patterns such as random-dot stereo pairs that are visible only when the brain combines the inputs from the two eyes, patterns that do not exist "physically" under monocular vision (Julesz 1971).

approaches have converged on similar problems. It is now possible to identify a common ground, and for the information gained with one approach to contribute to an explanation of the findings of another. The processing involved in stereopsis includes matching corresponding points of the images in the two eyes, measuring their disparity, and recovering the 3-D structure of the objects seen by the viewer. In what follows, we attempt to discuss some aspects of this processing that may be considered from an integrated perspective.

The correspondence problem has dominated much of the recent computational and psychophysical investigations. As unequivocally shown by the psychophysics with random-dot stereograms, during everyday, normal binocular vision, the brain solves the correspondence problem of stereopsis with amazing ease, precision, and immunity to noise. Potentially numerous false matches do not disturb the performance of our visual system. Existing algorithms are still far from solving the correspondence problem as effectively as our visual system, but suggest ways of how it may be solved at all. Recently, cells have been found in the visual cortex of the macaque that signal the correct disparity of dynamic random-dot stereograms in which there are many possible false matches. These neurons reflect the results of the correspondence process; they are at least one of the neural correlates of "global" stereopsis (Julesz 1971). The computational approach strongly suggests that these neurons must play a critical role in the perception of depth, not only from random-dot stereograms, but most importantly, in natural images. Of singular importance for the understanding of the organization of the visual system, these neurons are found at the earliest stages of binocular interaction in cortical area 17. This discovery suggest new insights about the brain mechanisms of stereoscopic matching.

In terms of the correspondence problem, we ask first what are the matching primitives for the stereoscopic neurons or, in other words, what are the significant "measurements" of the visual images on which the neuronal matching process operates. Second, we ask whether these neurons exploit global matching constraints of the type discussed above such as continuity and uniqueness.

The results of neurophysiological experiments show that neurons in monkey striate and prestriate cortex are sensitive to binocular disparity. When presented with disparity unambiguous stimuli, such as isolated line contours of optimal size and orientation, the majority of cortical neurons, both simple and complex, display disparity selectivity. The evidence suggests that under these conditions cortical neurons measure binocular disparity by receptive field disparity. For each cell a response profile can be constructed along the disparity domain over a range centered on the functionally optimal spatial superposition of the two monocular fields in space, and extending over a range of disparities that depends on the extent of field superposition and on receptive field size (Poggio

& Talbot 1981). In the cat, and presumably in the macaque as well, receptive field width and the width of the disparity tuning curve correlate strongly (Pettigrew et al 1968, Ferster 1981), as expected by some of the models of stereopsis described above. These stereoscopic cells then may be thought to use matching primitives that correspond to isolated oriented edges.

When, on the other hand, the cortical binocular neuron is presented with disparate figures embedded in or as part of a textured pattern, such as occurring in random-dot stereograms (and commonly in natural images as well), only about one out of five of the binocular cells in primary visual cortex, A17, is capable of detecting disparity. A major functional difference is that the simple cells do not and the complex cells do respond to texture patterns without sharp, oriented monocular edges (Hammond & MacKay 1977, Burr et al 1981, Poggio 1984). It seems unlikely that the complex cells require for activation isolated "edges" or zero-crossings of a certain orientation in their receptive field. Indeed, neurons have been found early in the cortical stereoprocessor (in area 17) that are orientation insensitive for cyclopean random-dot stereograms. Under these stimulus conditions, simple positional disparity between the receptive fields in the two eyes does not appear to be a sufficient mechanism for disparity detection of the correct matches. At these cortical neurons, binocular matching appears to be spatially very precise, for these cells are capable of signaling the correct disparity of a few dots whose positions over the cell's receptive fields change 100 times a second (Poggio 1984). A possible arrangement for these neurons could be based on sets of discrete and numerous receptive sites, or subfields, in the receptive field in one eye, "gated" with a similar set of appropriate positional disparity in the other eye. False matches would be avoided by having cells with subfield disparity of the order of the associated precortical (or IVc) monocular receptive field's size, and not larger. This scheme is computationally similar to the correlation of some function of the filtered image over the area defined by the cortical receptive field. Another conjecture on the neural organization for recognizing correct correspondence would regard the stereoscopic neuron matching a large set of local monocular measurements on the image such as, for example, derivatives of image intensities. In this case, one would expect afferents to the cell to carry in parallel the different types of measurement taken at each location in the visual field. The disparity range may easily be larger than in the neural network hypothetized previously (see section on the Correspondence Problem).

In both cases, a key question is how the activity of the binocularly interacting subfields, or measurements, is "integrated" over the area of the cortical receptive field. Candidate operations range from an approximately linear summation to an almost logical combination of the binocular matches. Data from responses of cortical cells to random-dot patterns with different degrees of binocular correlation could provide information of the characteristics of the integration.

Difficulties with matching schemes involving linear summation over relatively large areas (like the receptive fields of complex cells) are well known (Baker 1982). Disparity discontinuities—like the border of a figure in a random-dot stereogram—are neither precisely nor easily detected in this way. Thus, one wonders whether the complex cells' sensitivity to random-dot stereograms may also represent the substrate for the psychophysical percept of sharp boundaries or whether other stereo mechanisms are required.

Little can be said about the neurophysiological correlates of "global" matching constraints. Constraints like continuity and uniqueness are enforced by the mechanisms suggested above. Moreover, a direct implementation of the constraints would also be possible, for example in terms of inhibitory interactions among cells tuned to different disparities, and with overlapping receptive field locations, but there is no physiological evidence either in favor or against it.

A set of complex cells, therefore, seems to represent at least an initial solution to the correspondence problem. But what, then, is the role of disparity sensitive simple cells? The receptive field of simple cells suggests that oriented edges or bars (possibly oriented zero-crossings or peaks) are their matching primitives. Simple cells, however, do not respond to random-dot stereograms. Is it possible that their disparity range and their input organization would make them particularly susceptible to false matches? And that this is the reason they are shut off in situations in which there is a large number of potential false matches, as in random-dot stereograms? One may think that in the presence of textured patterns that pose a severe correspondence problem, simple cells are suppressed monocularly (Burr et al 1981) and binocularly, possibly by complex cells. What would then be their function? Psychophysical as well as computational findings suggest that monocular contours play an important role in driving eye movements for registering the two images. Isolated, oriented, sharp zero-crossing (or edges) are in addition useful matching primitives, as shown by Mayhew & Frisby (1981), both computationally and psychophysically. Simple cells may represent and match monocularly strong isolated borders with a given orientation and their output may be used in the matching process and in the control of vergence movements (see also Marr & Poggio 1979, figure 6).

Other problems also have been brought into sharper focus by the interaction between different approaches, for instance, the role of vertical disparity in stereopsis and the control of eye movements. From the computational perspective, vertical disparity and more generally geometric distortions pose some of the hardest problems for stereoscopic matching. It would be interesting to know the extent to which the complex cells responding to random-dot stereograms can tolerate vertical disparity, and to compare it with psychophysical data and with the performance of different algorithms. Since registration of the two eyes seems so important, it would be natural to expect cells tuned to vertical disparities, especially Near/Far with simple receptive field organization (be-

cause monocular contours should play a prominent role in registration), as well as Tuned cells with peak at zero disparity, to have an important role in controlling eye movements.

In addition to the solution of the correspondence problem, the computation of stereoscopic information entails the recovery of the 3-D structure of the visual scene from binocular disparity and knowledge of the fixation distance (for example, from angle of vergence and gaze) or assessment of vertical disparity (Mayhew & Longuet-Higgins 1982). Psychophysical observations have shown that brains do perform this computation, though approximately, but nothing is known of the physiological mechanisms underlying it. Neurons or, more likely, neuronal networks sensitive to viewer-centered coordinates and distances (as opposed to retino-centered coordinates) are expected to be involved. This aspect of stereoscopic analysis is likely to be beyond the initial stages of visual processing and is possibly organized within a nontopographic representation of the visual field.

Many other questions are readily suggested by any attempt to connect the computational point of view with the physiological and psychophysical data. Among the most obvious ones, a general problem is whether cells represent disparity directly or rather the relative disparity with respect to another point in the image. The question is motivated by the apparent stability of the cortical neuron's response in the presence of small involuntary fixation disparities (Poggio 1984) as well as by the psychophysical findings of Westheimer & Mitchison on stereoacuity (see section on Stereoacuity). An associated question concerns the possible existence of a cortical shift mechanism similar to vergence movements over a small range of disparities (less than ten degrees); in computer algorithms, these two possibilities are indistinguishable.

In conclusion: The interaction of the different approaches to stereopsis promises to be very fruitful for understanding both the mechanisms and the information processing aspects of binocular vision. The first effect of an integrated approach is a number of new questions and a recasting of older ones in a different and more meaningful form. Work on algorithms for stereopsis is helping to clarify what the central issues are in terms of information processing; physiology and psychophysics will tell us how and where these problems are solved. Stereopsis is an early but very difficult state of vision; it provides, we believe, an unique opportunity for a combined understanding of information processing and of neural mechanisms. We expect that future research on stereopsis will show that the attempt of understanding human vision and of developing computer vision can fruitfully interact.

ACKNOWLEDGMENTS

We wish to thank the Neuroscience Research Institute of the Neuroscience Research Program for the kind hospitality and for the facilities provided to us during the writing of this review.

We are grateful to E. Grimson, M. Kass, and especially to K. Nishihara and K. Nielsen for many helpful discussions.

This work was supported by NIH grant EYO2966 (G. F. P.), by the Advanced Research Project Agency of the Department of Defense under Office of Naval Research contract N00014–75–C–0643 to the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology (T. P.), and by NIH grant EYO4206 (T. P.).

Literature Cited

Arnold, R. D. 1982. *Automated stereo perception*. PhD thesis. Stanford Univ.

Arnold, R. D., Binford, T. O. 1980. Geometric constraints in stereo vision. *SPIE J.* 238:281–92

Baker, H. H. 1980. Edge-based stereo correlation. *Proc. ARPA Image Understanding Workshop*, Univ. Md., pp. 168–75

Baker, H. H. 1982. *Depth from edge- and intensity-based stereo*. PhD thesis. Univ. Ill., Urbana, IL

Baker, H. H. Binford, T. O. 1981. Depth from edge- and intensity-based stereo. *Proc. 7th Intern. Joint Conf. on Artif. Intell.* Vancouver, British Columbia, pp. 631–36

Barlow, H. 1979. Reconstructing the visual image in space and time. *Nature* 279:189–90

Barlow, H., Blakemore, C., Pettigrew, J. D. 1967. The neural mechanism of binocular depth discrimination. *J. Physiol.* 193:327–42

Birch, E. E., Gwiazda, J., Held, R. 1982. Stereoacuity development for crossed and uncrossed disparities in human infants. *Vision Res.* 22:507–13

Bishop, P. O., Henry, G. H., 1971. Spatial vision. *Ann. Rev. Psychol.* 22:119–61

Bishop, P. Q., Henry, G. H., Smith, C. J. 1971. Binocular interaction fields of single units in the cat striate cortex. *J. Physiol.* 216:39–68

Bough, E. W. 1970. Stereoscopic vision in the macaque monkey: A behavoural demonstration. *Nature* 225:42–44

Braddick, O. J. 1979. Binocular single vision and perceptual processing. *Proc. R. Soc. London Ser. B* 201:503–12

Burr, D. C. 1979. Acuity for apparent Vernier offset. *Vision Res.* 19:835–37

Burr, D. C., Ross, J. 1979. How does binocular delay give information about depth? *Vision Res.* 19:523–32

Burr, D. C., Morrone, C., Maffei, L. 1981. Intra-cortical inhibition prevents simple cells from responding to textured visual patterns. *Exp. Brain Res.* 43:455–58

Burt, P., Julesz, B. 1980a. A disparity gradient limit for binocular fusion. *Science* 208:615–17

Burt, P., Julesz, B. 1980b. Modifications of the classical notion of Panum's fusional area. *Perception* 9:671–82

Clarke, P. G. H., Donaldson, I. M. L., Whitteridge, D. 1976. Binocular visual mechanisms in cortical areas I and II of the sheep. *J. Physiol.* 256:509–26

Crick, F., Marr, D., Poggio, T. 1981. An information-processing approach to understanding the visual cortex. In *The Organization of the Cerebral Cortex*, ed. F. O. Schmitt, F. G. Worden, G. Edelman, S. G. Dennis, pp. 505–33. Cambridge, Mass: MIT Press

Dev, P. 1975. Perception of depth surfaces in random-dot stereograms: A neural model. *Int. J. Man-Machine Studies* 7:511–28

DeValois, R. L., DeValois, K. K. 1980. Spatial vision, *Ann. Rev. Psychol.* 31:309–41

Diner, D. 1978. *Hysteresis in human binocular fusion: A second look*. PhD thesis. Calif. Inst. Technol., Pasadena

Duwaer, A. L. 1981. *Binocular Single Vision: Psychophysical Studies on Underlying Sensory and Motor Processes*. Rotterdam: Delfta Univ. Press. 173 pp.

Duwaer, A. L., van den Brink, G. 1981. Diplopia thresholds and the initiation of vergence eye movements. *Vision Res.* 21:1727–37

Duwaer, A. L., van den Brink, G. 1982. Detection of vertical disparities. *Vision Res.* 22:467–78

Felton, B., Richards, W., Smith, A. Jr. 1972. Disparity processing of spatial frequencies in man. *J. Physiol.* 225:319–62

Fender, D., Julesz, B. 1967. Extension of Panum's fusional area in binocularly stabilized vision. *J. Opt. Soc. Am.* 57:819–30

Ferster, D. 1981. A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *J. Physiol.* 311:623–55

Fischer, B., Krueger, J. 1979. Disparity tuning and binocularity of single neurons in the cat visual cortex. *Exp. Brain Res.* 35:1–8

Fischer, B., Poggio, G. F. 1979. Depth sensitivity of binocular cortical neurons of behaving monkeys. *Proc. R. Soc. London Ser. B* 204:409–14

Foley, J. M. 1980. Binocular distance perception. *Psych. Rev.* 87(5):411–34
Fox, R., Aslin, R. N., Shea, S. L., Dumais, S. T. 1979. Stereopsis in human infants. *Science* 207:323–24
Frisby, J. P., Mayhew, J. E. W. 1977. Global processes in stereopsis: Some comments on Ramachandran and Nelson (1976). *Perception* 6:195–206
Frisby, J. P., Mayhew, J. E. W. 1978. Contrast sensitivity function for stereopsis. *Perception* 7:423–29
Gennery, D. B. 1980. Modeling the environment of an exploring vehicle by means of stereo vision. *Stanford Artif. Intell. Lab. Memo* 339
Gillam, B., Lawergren, B. 1983. The induced effect, vertical disparity, and stereoscopic theory. *Percept. Psychophys.* 34:121–30
Grimson, W. E. L. 1981. *From Images to Surfaces: A Computational Study of the Human Early Visual System*. Cambridge, Mass: MIT Press. 274 pp.
Hammond, P., MacKay, D. M. 1977. Differential responsiveness of simple and complex cells in cat striate cortex to visual texture. *Exp. Brain Res.* 30:275–96
Harwerth, R. S., Rawlings, S. C. 1977. Viewing time and stereoscopic threshold with random-dot stereograms. *Am. J. Optom. Phys. Opt.* 54(7):452–57
Held, R. 1983. Binocular Vision . . . behavioral and neuronal development. In *Neonate Cognition: Beyond the Blooming, Buzzing Confusion*, ed. J. Mehler, R. R. Fox. Hillsdale, N.J.: Erlbaum. In Press
Held, R., Birch, E., Gwiazda, J. 1980. Stereoacuity of human infants. *Proc. Nat. Acad. Sci. USA* 77:5572–74
Hubel, D. H., Wiesel, T. N. 1970. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225:41–42
Hubel, D. H., Wiesel, T. N. 1973. A reexamination of stereoscopic mechanisms in the cat. *J. Physiol.* 232:29–30P
Jones, R. 1977. Anomalies of disparity in the human visual system. *J. Phys.* 264:621–40
Jones, R., Kerr, K. E. 1972. Vergence eye movements to pairs of disparity stimuli with shape selection cues. *Vision Res.* 12:1425–30
Joshua, D. E., Bishop, P. O. 1970. Binocular single vision and depth discrimination. Receptive field disparities for central and peripheral vision and binocular interaction on peripheral units in cat striate cortex. *Exp. Brain Res.* 10:389–416
Julesz, B. 1960. Binocular depth perception of computer-generated patterns. *Bell System Tech. J.* 39:1125–62
Julesz, B. 1971. *Foundations of Cyclopean Perception*. Chicago: Univ. Chicago Press. 406 pp.
Julesz, B. 1978. Global stereopsis: Cooperative phenomena in stereoscopic depth perception. In *Hand. Sensory Physiol.* 8:215–52
Julesz, B., Miller, J. 1975. Independent spatial-frequency-tuned channels in binocular fusion and rivalry. *Perception* 4:125–43
Julesz, B. Schumer, R. A. 1981. Early visual perception. *Ann. Rev. Psychol.* 32:575–627
Julesz, B., Spivack, G. J. 1967. Stereopsis based on vernier acuity cues alone. *Science* 157:563–65
Kass, M. 1983. A computational framework for the visual correspondence problem. *Proc. ARPA Image Understanding Workshop, Washington DC*
Kidd, A. L., Frisby, J. P., Mayhew, J. E. W. 1979. Texture contours can facilitate stereopsis by initiating vergence eye movements. *Nature* 280:829–32
Kroll, J. D., van der Grind, A. 1980. The double-nail illusion: Experiments on binocular vision with nails, needles and pins. *Perception* 9:651–69
Kulikowski, J. J. 1978. Limit of single vision in stereopsis depends on contour sharpness. *Nature* 275:126–27
Levinson, E., Blake, R. 1979. Stereopsis by harmonic analysis. *Vision Res.* 19:73–78
Logan, B. F. Jr. 1977. Information in the zero crossings of bandpass signals. *Bell System Tech. J.* 66:487–510
Longuet-Higgins, H. C. 1981. A computer algorithm for reconstructing a scene from two projections. *Nature* 293:4133–35
Marr, D. 1976. Early processing of visual information. *Philos. Trans. R. Soc. London Ser. B* 275:483–524
Marr, D., Hildreth, E. 1980. Theory of edge detection. *Proc. R. Soc. London Ser. B* 207:187–217
Marr, D., Palm, G., Poggio, T. 1978. Analysis of a cooperative stereo algorithm. *Biol. Cyber.* 28:223–39
Marr, D., Poggio, T. 1976. Cooperative computation of stereo disparity. *Science* 194:283–87
Marr, D., Poggio, T. 1979. A computational theory of human stereo vision. *Proc. R. Soc. London Ser. B* 204:301–28
Marr, D., Poggio, T. 1980. Some comments on a recent theory of stereopsis. *MIT Artif. Intell. Memo* 558
Mayhew, J. E. W. 1983. Stereopsis. In *Physiological and Biological Processing of Images*, ed. O. J. Braddick, A. C. Sleigh. Berlin: Springer-Verlag
Mayhew, J. E. W., Frisby, J. P. 1978a. Stereopsis masking in humans is not orientationally tuned. *Perception* 7:431–36
Mayhew, J. E. W., Frisby, J. P. 1978b. Contrast summation effects and stereopsis. *Perception* 7:537–50
Mayhew, J. E. W., Frisby, J. P. 1979a. Sur-

faces with steep variations in depth pose difficulties for orientationally tuned disparity filters. *Perception* 8:691–98
Mayhew, J. E. W., Frisby, J. P. 1979b. Convergent disparity discriminations in narrow-band-filtered random-dot stereograms. *Vision Res.* 19:63–71
Mayhew, J. E. W., Frisby, J. P. 1980. The computation of binocular edges. *Perception* 9:69–86
Mayhew, J. E. W., Frisby, J. P. 1981. Psychophysical and computational studies towards a theory of human stereopsis. *Artif. Intell.* 16:349–85
Mayhew, J. E. W., Longuet-Higgins, C. 1982. A computational model of binocular depth perception. *Nature* 297:376–78
Mitchell, D. E. 1966. Retinal disparity and diplopia. *Vision Res.* 6:441–51
Moravec, H. P. 1980. Obstacle avoidance and navigation in the real world by seeing a robot rover. *Stanford Artif. Intell. Lab. Memo* 340
Morgan, M. J. 1975. Stereoillusion based on visual persistence. *Nature* 256:639–40
Morgan, M. J. 1976. Pulfrich effect and the filling-in of apparent motion. *Perception* 5:187–95
Motter, B. C., Poggio, G. F. 1981. The three-dimensional locus of fixation of the rhesus monkey. *Invest. Ophthalmol. Vis. Sci. Suppl.* 20:27
Motter, B. C., Poggio, G. F. 1982. Spatial invariance of receptive field location in the presence of eye movements of fixation for neurons in monkey striate cortex. *Soc. Neurosci. Abstr.* 8:707
Mowforth, P., Mayhew, J. E. W., Frisby, J. P. 1981. Vergence eye movements made in response to spatial-frequency-filtered random-dot stereograms. *Perception* 10:299–304
Nelson, J. I. 1975. Globality and stereoscopic fusion in binocular vision. *J. Theor. Biol.* 49:1–88
Nelson, J. I. 1977. The plasticity of correspondence: After-effects, illusion and horopter shifts in depth perception. *J. Theor. Biol.* 66:203–66
Nielson, K. R. K., Poggio, T. 1983. Vertical image registration in human stereopsis. *MIT Artif. Intell. Memo* 743
Nikara, T., Bishop, P. O., Pettigrew, J. D. 1968. Analysis of retinal correspondence by studying receptive fields of binocular signal units in cat striate cortex. *Exp. Brain Res.* 6:353–72
Nishihara, H. K., Poggio, T. 1982. Hidden cues in random line stereograms. *Nature* 300:347–49
Petrig, B., Julesz, B., Kropfl, W., Baumgartner, G., Anliker, M. 1981. Development of stereopsis and cortical binocularity in human infants: Electrophysiological evidence. *Science* 213:1402–5
Pettigrew, J. D., Konishi, M. 1976. Neurons selective for orientation and binocular disparity in the visual Wulst of the barn owl *(Tyto alba). Science* 193:675–78
Pettigrew, J. D., Nikara, T., Bishop, P. O. 1968. Binocular interaction on single units in cat striate cortex: Simultaneous stimulation by single moving slits with receptive fields in correspondence. *Exp. Brain Res.* 6:391–410
Poggio, G. F. 1980. Neurons sensitive to dynamic random-dot stereograms in areas 17 and 18 of the rhesus monkey cortex. *Soc. Neurosci. Abstr.* 6:672
Poggio, G. F. 1984. Processing of stereoscopic information in primate visual cortex. In *Dynamic Aspects of Neocortical Function*, ed. G. Edelman, W. M. Cowan, W. E. Gall. New York: Wiley. In press
Poggio, G. F., Fisher, B. 1977. Binocular interaction and depth sensitivity of striate and prestriate cortical neurons of the behaving rhesus monkey. *J. Neurophysiol.* 40:1392–405
Poggio, G. F., Talbot, W. H. 1981. Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *J. Physiol.* 315:469–92
Rashbass, C., Westheimer, G. 1961a. Disjunctive eye movements. *J. Physiol.* 159:339–60
Rashbass, C., Westheimer, G. 1961b. Independence of conjunctive and disjunctive eye movements. *J. Physiol.* 159:361–64
Richards, W. 1970. Stereopsis and stereoblindness. *Exp. Brain Res.* 10:380–88
Richards, W. 1971. Anomalous stereoscopic depth perception. *J. Opt. Soc. Am.* 61:410–14
Riggs, L. A., Niehl, E. W. 1960. Eye movements recorded during convergence and divergence. *J. Opt. Soc. Am.* 50:913–20
Ross, J., Hogben, J. H. 1975. Short-term memory in stereopsis. *Vision Res.* 14:1195–1201
Sarmiento, R. F. 1975. The stereoacuity of macaque monkey. *Vision Res.* 15:493–98
Schiller, P. H., Finlay, B. L., Volman, S. 1976. Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance. *J. Neurophysiol.* 39:1320–33
Schor, C. M., Tyler, C. W. 1981. Spatio-temporal properties of Panum's fusional area. *Vision Res.* 21:683–92
Schumer, R. A., Ganz, L. 1979. Independent stereoscopic channels for different extents of spatial pooling. *Vision Res.* 19:1303–44
Schumer, R. A., Julesz, B. 1980. Maximum disparity limit for detailed depth resolution. *Invest. Ophthalmol. Visual Sci.* 19:106–7 (Suppl.)
Schumer, R. A., Julesz, B. 1982. Disparity limits in bandpass random-grating stero-

grams. *Invest Ophthalmol. Visual Sci.* 22:272 (Suppl.)
Sperling, G. 1970. Binocular vision: A physical and a neural theory. *Am. J. Psychol.* 83:461–534
St. Cyr, G., Fender, D. H. 1969. The interplay of drifts and flicks in binocular fixation. *Vision Res.* 9:245–65
Teller, D. Y. 1981. The development of visual acuity in human and monkey infants. *Trends NeuroSci.* 4:21–24
Tsai, R. Y., Huang, T. S. 1981. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *Tech. Rep. R921, Coordinated Sci. Lab., Univ. Ill., Urbana.* 60 pp.
Tyler, C. W. 1973. Stereoscopic vision: Cortical limitations and a disparity scaling effect. *Science* 181:276–78
Tyler, C. W. 1974. Depth perception in disparity gratings. *Nature* 251:140–42
Tyler, C. W. 1975a. Spatial organizations of binocular disparity sensitivity. *Vision Res.* 15:583–90
Tyler, C. W. 1975b. Spatial limitations of human stereoscopic vision. *SPIE J.* 120:36–42
Tyler, C. W. 1977. Stereomovement from interocular delay in dynamic visual noise: A random spatial disparity hypothesis. *Am. J. Optom. Phys. Opt.* 54(6):374–86
Tyler, C. W., Julesz, B. 1980. On the depth of the cyclopean retina. *Exp. Brain Sci.* 40:192–202
Uttal, W. R., Fitzgerald, J., Eskin, T. E. 1975. Parameters of tachistoscopic stereopsis. *Vision Res.* 15:705–23
von der Heydt, R., Adorjani, C., Hanny, P., Baumgartner, G. 1978. Disparity sensitivity and receptive field incongruity of units in the cat striate cortex. *Exp. Brain Res.* 31:423–545
Westheimer, G. 1978. Vertical disparity detection: Is there an induced size effect? *Invest. Ophthalmol. Visual Sci.* 17:545–51
Westheimer, G. 1979. Cooperative neural processes involved in stereoscopic acuity. *Exp. Brain Res.* 36:585–97
Westheimer, G., McKee, S. P. 1978. Stereoscopic acuity for moving retinal images. *J. Opt. Soc. Am.* 68:450–55
Westheimer, G., McKee, S. P. 1980a. Stereogram design for testing local stereopsis. *Invest. Ophthalmol. Visual Sci.* 19:802–9
Westheimer, G., McKee, S. P. 1980b. Stereoscopic acuity with defocused and spatially filtered retinal images. *J. Opt. Soc. Am.* 70(7):772–78
Westheimer, G., Mitchell, D. E. 1969. The sensory stimulus for disjunctive eye movements. *Vision Res.* 9:749–55
Wilson, H. R., Bergen, J. R. 1979. A four mechanism model for threshold spatial vision. *Vision Res.* 19:19–32