

## INDEPENDENT STEREOSCOPIC CHANNELS FOR DIFFERENT EXTENTS OF SPATIAL POOLING

ROBERT SCHUMER and LEO GANZ

Department of Psychology, Stanford University, Stanford, California 94305, U.S.A.

(Received 18 September 1978; in revised form 12 April 1979)

**Abstract**—We report two experiments with stimuli composed of dynamic noise stereograms in which binocular disparity was modulated sinusoidally with changes in vertical spatial position. The experiments demonstrated that the human visual system contains independent stereoscopic mechanisms selectively tuned for different spatial frequencies of disparity modulation. The first experiment showed that observers can just detect a compound disparity grating when at least one of two presented sinusoidal components is close to its own independent threshold amplitude. The second experiment demonstrated selective threshold elevation following prolonged viewing of a disparity grating (selective adaptation). The results suggest that the stereoscopic mechanisms have rather broad sensitivity to the spatial frequency of disparity modulation: 2–3 octaves full-bandwidth at half-amplitude. We conclude, first, that there exist multiple stereoscopic mechanisms each characterized by a different extent of spatial pooling. Secondly, the bandpass characteristic we have observed implies that these mechanisms must receive lateral inhibition from disparity detectors tuned to adjacent positions in space.

### INTRODUCTION

Direct psychophysical evidence for the existence of disparity-selective mechanisms in the human visual system was presented by Blakemore and Julesz (1971). They reported a repulsion aftereffect in the perceived depth of a random-dot stereogram (Julesz, 1971) following prolonged viewing of a cyclopean form in depth. But Blakemore and Julesz were unable to draw conclusions concerning the nature of the stereoscopic mechanisms involved. Logically, there exist a number of alternative underlying mechanisms which could account for such results.

As has been previously pointed out (e.g. Ganz, 1966; Blakemore and Sutton, 1969; Anstis, 1975; Nelson, 1977), the presence of figural aftereffects suggests the existence of a domain of tuned channels along which stimuli are represented. In this view, an aftereffect would be due to a resulting imbalance in the overall response distribution following fatigue or inhibition of a subset of channels through adaptation or simultaneous induction. What is left unspecified by the Blakemore and Julesz finding, however, is the specific nature of the underlying stereoscopic domain. In this paper we explore the idea that different stereoscopic mechanisms exist which are each sensitive to only a selective range of the sinusoidal components present in a form-in-depth.

Study of the visual analysis of form in the luminance domain has revealed the existence of mechanisms selectively sensitive to different spatial frequencies of luminance modulation (e.g. Blakemore and Campbell, 1969; Graham and Nachmias, 1971). Julesz and Miller (1975) have studied the relevance of these channels to stereopsis. The experiments we describe are *not* concerned with these channels, but rather with the characteristics of visual sensitivity to stereoscopic form-in-depth when luminance-domain form is held *constant*.

A relevant study was performed by Tyler (1975a) who reported tilt and size aftereffects following prolonged viewing of a random-dot stereogram portraying a disparity grating. This stimulus consists of stationary horizontally-oriented bars which are perceived to ripple sinusoidally in depth from the top to bottom of the stereogram. Disparity gratings are produced through sinusoidal variation of disparity, along the vertical axis, in random-dot stereograms which contain no monocular cues to either form or depth. After adaptation to a disparity grating in which the bars were tilted off true horizontal (that is, rotated about an axis passing through the line-of-sight), a horizontally oriented test disparity grating appeared tilted in the opposite direction to that of the adapting grating. In addition, Tyler observed a perceived shift in the spatial frequency of a disparity grating after adapting to a disparity grating of slightly different spatial frequency. By analogy with interpretations given similar results using luminance gratings (Campbell and Maffei, 1971; Blakemore and Sutton, 1969; Blakemore *et al.*, 1970; Anstis, 1975), Tyler inferred the existence of channels selective for the orientation and spatial frequency of depth patterns. Evidence from our first experiment supports the conclusion that there exist independent mechanisms selectively tuned to the spatial frequency of disparity modulation. Using the threshold elevation technique of Pantle and Sekuler (1968) and Blakemore and Campbell (1969), we have also found that these mechanisms are more broadly tuned than the analogous channels in the luminance domain. In discussing these results, we will describe several implications for the pooling processes in global stereoscopic processing.

### GENERAL METHODS

A block diagram of our apparatus is shown in Fig. 1. The dynamic stereoscopic display was generated on the

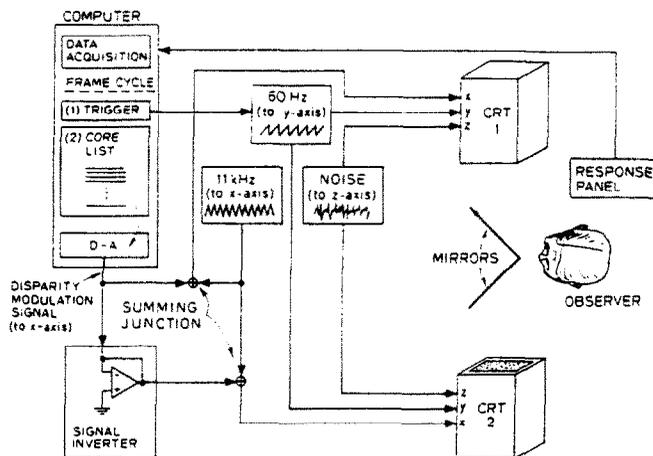


Fig. 1. Block diagram for experimental apparatus; for description, see text.

faces of two identical Tektronix 602 oscilloscopes (P31 phosphor), which were positioned at the left and right arms of a long enclosure and viewed through a pair of front surface mirrors set approximately at a right angle to one another. The two mirrors were mounted on spring loaded bolts and were adjusted so that the two displays were both easily fused and accommodated by observers. The oscilloscope screens were viewed from a distance of 80 cm and subtended  $5.7^\circ$  vertically and  $7.4^\circ$  horizontally. Stereoscopic diagonal cross-hairs were added to aid fixation.

The X-axis position of the beam of both scopes was swept by the same triangular wave from a Wavetek 180 function generator at a frequency of 11 kHz. The Y-axis position of both scopes was also identically driven by a 60 Hz sawtooth, each cycle of which was triggered by a pulse from a PDP-8 computer. Finally, the signal from a General Radio 1390-A noise generator was sent to the Z-axis input of both oscilloscopes. (The noise was low-pass and of uniform spectrum level to 5 MHz, which under the conditions of display resulted in a uniform spatial frequency spectrum, in the vertical orientation, to 31 cycles per degree of visual angle.) Hence, for a zero disparity condition, the display on each scope was identical. When viewed stereoscopically the observer saw a flat surface in the fronto-parallel plane composed of random and continuously changing noise.

The mean amplitude of the signal from the noise generator determined the mean luminance of the screens, which was the same for each screen, namely  $5.5 \text{ cd/m}^2$ . The experimental room was otherwise dark. Conditions were such that there was strong apparent movement of the specks of light on the screen, but no perception of 60 Hz flicker.

The generation of patterns in stereoscopic depth was accomplished by having the PDP-8 access repeatedly a list of 500 numbers stored in core memory. Each cycle of 500 numbers was initiated at the same time as the trigger pulse which controlled the onset of the Y-axis sawtooth. Each number was sent in sequence to a digital-to-analog converter whose voltage output was then electronically summed with the X-axis raster voltage. The D-A voltage was summed directly with the raster going to the X-axis of the right eye's scope, but was electronically inverted before summing with the X-axis raster of the left eye's scope.

The read-out of the numeric array was synchronized and phase-locked with the 60 Hz frame rate of the oscilloscopes. Hence, the numeric profile of 500 numbers was converted to a horizontal positional displacement of the noise fields, of equal magnitude but in opposite directions for the two scopes. By this means, a horizontal binocular

disparity was introduced between corresponding noise elements on an entire horizontal line, but each individual line could have any desired disparity. Each such line subtended  $56''$  arc of vertical width. The left- and right-most edges of each noise field were occluded by the edge of the oscilloscope screen under conditions of the largest displacement used. In this fashion, there were no monocular cues to the form of the surface to be portrayed in depth.

Our experiments employed stimuli in which disparity was modulated sinusoidally from the top to the bottom of the screens. The stereoscopic view was of a stationary sinusoidal grating in depth, the long axis of the bars extending from left to right. Under control of the computer were the spatial frequency of the bars (number of cycles per degree of visual angle), the amplitude of disparity modulation, and the mean disparity around which modulation occurred. In the present experiments, the mean disparity value was always held constant at zero. Our gratings are thus similar to stimuli used by Tyler (1974, 1975a), except that his stimuli were static while the noise we use is dynamic in time. A photograph of the display (integrating 4 frames) appears in Fig. 2. Viewed stereoscopically, the display shows a sinusoid in depth.

The results which we report here all are based on psychophysical measurements of the minimum amplitude of disparity modulation required for the observer to detect the presence of a briefly flashed disparity grating. Thresholds were measured by the method of limits with steps of  $2.5''$  of arc. For a single threshold determination, measurements were taken on four ascending and four descending runs, and a threshold estimate was computed by taking the mean of the eight settings. For simple gratings, four different stimuli were measured concurrently, with the computer randomly intermixing stimuli as well as ascending or descending runs. Occasionally, thresholds for only two different stimuli were concurrently determined.

At all times when a test stimulus was not being presented, a zero disparity conditioning field was present. When ready, the observer initiated a trial by pressing a microswitch. After a random delay of up to 100 msec, a disparity grating was briefly presented and was then immediately replaced by the conditioning field. The stimulus duration was 50 msec for the highly practiced of our two observers, and 100 msec for the other, less experienced observer. (For the latter, 100 msec was the briefest duration that could reliably be used without much additional training; both durations used were too brief to allow vergence or saccadic eye movements correlated with the onset of the stimulus.)

Basic data on the disparity thresholds of simple gratings

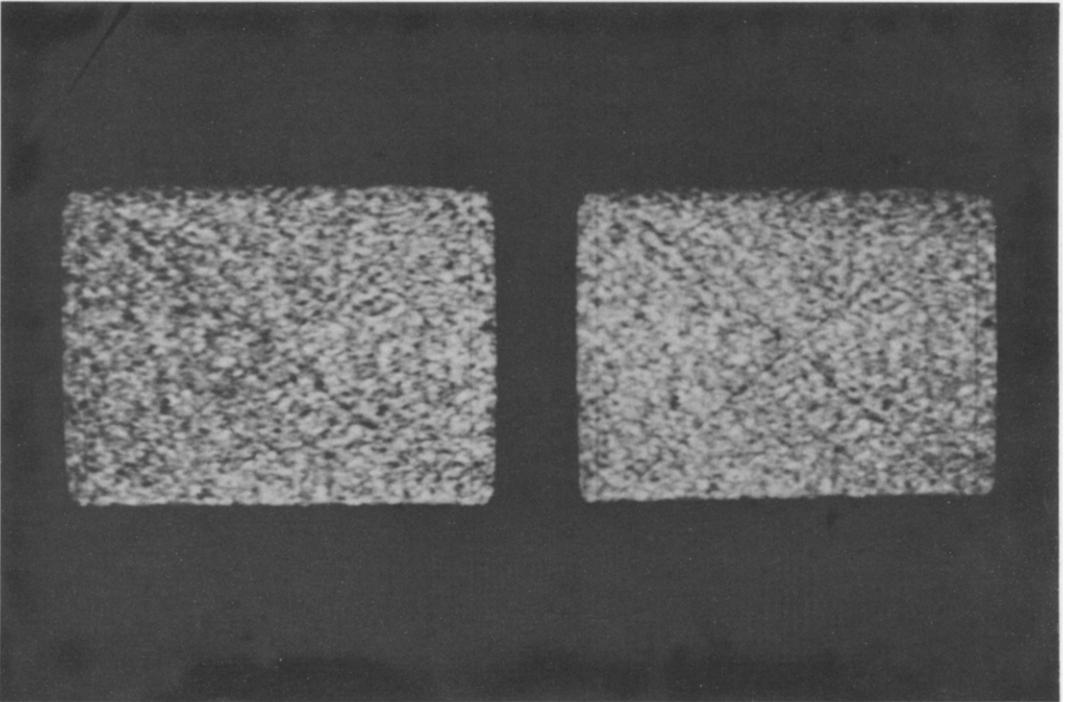


Fig. 2. Photograph of display. Shutter speed was chosen to integrate four successive frames; this gives an appearance similar to when display is viewed by eye.



were gathered by repeating the block of eight settings from five to eight times for each spatial frequency investigated. For each spatial frequency, our estimate of threshold was the mean of the results of these repetitions. Thus, each threshold estimate represents 40 to 64 settings.

Observers were dark-adapted for 2–4 min to the luminance level of the oscilloscope screens before each session. Head movements were minimized by use of a chin rest and a forehead support. Normal pupils were used. One observer, RS, (one of the authors), was highly practiced, while the other observer, LA, was inexperienced at the beginning of training and was naive throughout with respect to the experimental goals. Both wore normal refractive correction and had good stereovision as judged from accurate perception of the complex stereogram figures in Julesz (1971).

## EXPERIMENT 1

### Introduction

In order to give direct test to the idea that there exist multiple mechanisms tuned to various spatial frequencies of disparity modulation, we employed the methodology developed by Graham and Nachmias (1971) in their study of the detection of patterns varying in the spatial frequency of luminance modulation. Since the theory underlying the models which are being contrasted here has been described in detail elsewhere (see, for example, Graham and Nachmias, 1971; Sachs *et al.*, 1971; Graham, 1978; Kulikowski and King-Smith, 1973; Graham *et al.*, 1978), we will provide here only a general summary of the reasoning to be pursued.

Consider that the sensitivity of the visual system to cyclopean stimuli (Julesz, 1971) varying in disparity across space may be modeled by an appropriate neural network or set of networks. The characteristics of such networks would probably include the extent of excitation and inhibition between units sensitive to different disparities and spatial positions. Such networks may be stimulated by periodic stereoscopic patterns which, for convenience, can vary in disparity only in one spatial direction. One may characterize the rate of repetition of such a pattern as its spatial frequency, or number of cycles of disparity modulation per degree of visual angle. Each neural network would respond in a characteristic way to each spatial frequency of disparity modulation. With regard to sensory thresholds, we may assume that an observer's threshold is reduced just when the networks sensitive to the test pattern reach a criterion response. Further, we may assume for convenience that near threshold a network's response is roughly linear with small changes in the intensity of the stimulus.

Now, if there exists only a single mechanism sensitive to cyclopean disparity modulation, then, in this single-channel model, there will be a single network which underlies sensitivity to all disparity gratings of any spatial frequency. Assume that such a network is at threshold when input with a grating of spatial frequency  $i$  c/deg and having an amplitude which can be denoted  $A_i^0$  ( $^0$  for *threshold*). Then, simultaneously inputting another grating of spatial frequency  $j$  c/deg at subthreshold disparity amplitude  $A_j^{-2}$  will, in general, increase the response of the network and consequently the visibility of the compound stimulus. To set the compound stimulus  $i + j$  back to threshold,

one could reduce the amplitude  $A_i$  until the network response is back to criterion. Whether the network is linear or not should be considered as a separate question from the fact that the response of a single network will, in general, be increased by any additional signal. The magnitude of the increase may, however, be phase-dependent.

But now consider the possibility that there are several, independent, mechanisms, each sensitive to a different range of spatial frequencies of cyclopean disparity modulation. In this multiple-channel model, if the network sensitive to pattern  $i$  is responding at threshold, and if this network has characteristics making it insensitive to pattern  $j$ , then we would expect to be able to add quite substantial (but subthreshold) amounts of  $j$  to  $i$  without affecting the detectability of the compound stimulus. This will still be determined by the response to  $i$ .

In order to test these two predictions, we measured the dependence of the threshold for a compound stimulus upon the relative presence of two component disparity gratings of different spatial frequencies. The single-mechanism model predicts a lowering of the threshold for detection of a compound stimulus when increasing (but still subthreshold) amounts of either component are added. In our measurements we also sampled two phase-combinations of the components in order to reveal any effects phase might have had. On the other hand, the multiple-mechanism model predicts that substantial additions of one component should not affect the threshold for the compound (if the spatial frequencies of the components are sufficiently different), since only the most detectable of two independently detected components will determine the threshold of the compound. In addition, for multiple and independent mechanisms, the phase of the components should not matter.

### Methods

Initial measurements were made of the threshold for disparity gratings over a range of spatial frequencies from 0.35 to 3.5 c/deg. The amplitude sensitivity function for two observers is shown in Fig. 3. Notice that the peak of the sensitivity function is between 0.5 and 0.8 c/deg, which is considerably lower than the peak sensitivity to pulse presentations of luminance gratings at comparable durations (Nachmias, 1967). This conforms to an earlier finding of Tyler's (1974). Note further that while the high spatial frequency segment of the two observer's curves overlap closely, there is a marked divergence in the lower spatial frequency portion of the curves. It is possible that this result could have relevance to the commonly observed, but little understood, phenomenon of the wide range of individual difference in the speed and ease with which observers perceive forms in random-dot stereograms.

Next, experiments were run to determine sensitivity to simple and compound gratings. Sessions were arranged in which eight concurrent thresholds were measured. Two of these conditions involved the remeasurement of the threshold for each of two simple gratings. These gratings always had frequencies in the ratio of 1:3 and may be denoted by  $f$  and  $3f$ . Conditions 3 through 5 consisted of the determination of the threshold for compound stimuli that were composed of both  $f$  and  $3f$  as components. These components were combined in peaks-subtract mode, or as we shall refer to it,  $0^\circ$  phase. In all three conditions, both components were always present in the compound in some fixed proportion, relative to their own respective independent thresholds as they had been previously estimated. This can

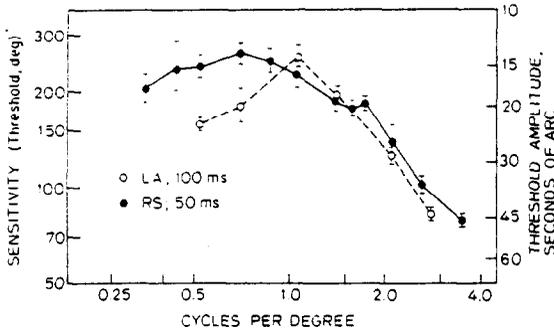


Fig. 3. Threshold disparity modulation amplitude as a function of spatial frequency for the two observers. Vertical bars show  $\pm 1$  SD around each point. Same data is plotted on two ordinates: the reciprocal of the threshold amplitude in degrees (left), and the threshold amplitude in seconds of arc (right).

be expressed as follows. Let  $A_f$  and  $A_{3f}$  refer to the disparity modulation amplitudes at threshold of  $f$  and  $3f$ , respectively, and let  $A_f^0$  and  $A_{3f}^0$  refer to the amplitudes of the  $f$  and  $3f$  components in the compound grating at any given amplitude. Then, for example, if the fixed proportion is unity:

$$\frac{A_f}{A_f^0} = \frac{A_{3f}}{A_{3f}^0}$$

More generally, this fixed proportion can be expressed by the ratio shown in equation (1), and that ratio can take on values other than 1.

$$\frac{\frac{A_f}{A_f^0}}{\frac{A_{3f}}{A_{3f}^0}} = R \quad (1)$$

This ratio can be denoted  $R$ ; any compound stimulus can then be denoted  $R_{i,j}$ . The first subscript indicates the value of  $R$ , while the second subscript refers to the angular phase of the two components in the compound. The condition just described is thus  $R_{1,0}$ . In two other conditions, the  $f$  component was always present to either twice or one-half the degree of the  $3f$  component, each relative to its own threshold. Thus, these conditions are described as  $R_{2,0}$  and  $R_{0.5,0}$ . Finally, conditions 6 through 8 involved measurement of thresholds for compounds in which the proportionate amplitudes of  $f$  and  $3f$  were as in the previous three conditions, but the components were combined in peaks-add, or  $180^\circ$  phase. These conditions were denoted  $R_{1,180}$ ,  $R_{2,180}$ , and  $R_{0.5,180}$ . For observer RS we studied three different values of  $f$  and  $3f$ : 0.5 and 1.5 c/deg, 0.7 and 2.1 c/deg, and 0.87 and 2.62 c/deg. All eight conditions of the first pair were repeated three times while the 8 conditions of the other two pairs were repeated twice each. For observer LA we only examined 0.7 and 2.1 c/deg and repeated the eight conditions three times. In addition, for observer RS, we measured thresholds for a few other values of  $R$  in two smaller sessions in each of which only four conditions were concurrently run. In each case, in addition to remeasurement of thresholds for  $f$  and  $3f$ , we measured thresholds for a different pair of compound stimuli. In one case  $f$  and  $3f$  were 0.7 and 2.1 c/deg and we measured thresholds for  $R_{1.5,180}$  and  $R_{0.67,180}$ . In the other case  $f$  and  $3f$  were 0.87 and 2.62 c/deg and we examined  $R_{1.2,0}$  and  $R_{0.8,0}$ . Both sets of four conditions were repeated three times.

In setting a compound stimulus to threshold, the computer always modified the stimulus from trial to trial so that the component with the physically larger amplitude

was modified in  $2.5^\circ$  arc steps, and the lesser amplitude component was adjusted accordingly to keep  $R$  constant. All eight (four) conditions were run concurrently, with the computer randomizing order of presentation, as before.

### Results

In order to summarize the data, we have pooled the data over replications of conditions. In Fig 4 we show plotted the averaged amplitude of the  $3f$  component present in each compound at threshold, versus the averaged amplitude of the  $f$  component present in the same compound stimulus. We have expressed these amplitudes as fractions of the concurrently determined values of  $A_f^0$  and  $A_{3f}^0$ , the independent thresholds of each component of the compound stimulus. As described in the legend to Fig. 4, different symbols represent different spatial frequency values of  $f$  and  $3f$ , while solid and open symbols represent  $0^\circ$  (peaks-subtract) and  $180^\circ$  (peaks-add) phases, respectively. The points marked by diamonds on the ordinate and abscissa are, by definition, the thresholds of each of the components alone, as estimated by the concurrently made measurements.

Examine first the points clustering near the coordinate (1.0, 0.5). These points represent the results of the  $R_{2,0}$  conditions. (This notation will be used to refer to a particular amplitude ratio collapsed across the two phase conditions.) As can be seen, when the  $R_{2,0}$  compound stimulus is at its threshold, the  $f$  component is still present just at its own threshold amplitude, while the  $3f$  component is present at 50% of its independent threshold amplitude. Similarly, the points near the (0.5, 1.0) coordinate are data for the

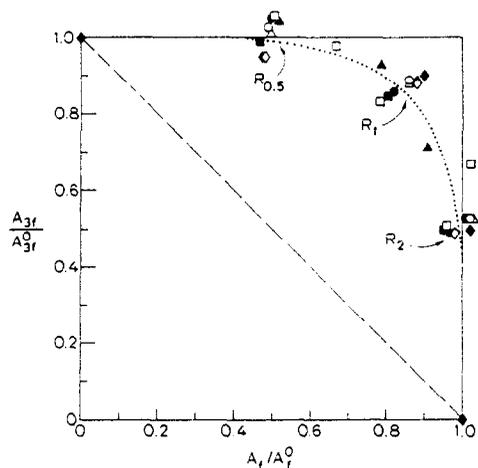


Fig. 4. Results of experiment 1. Horizontal axis shows amplitude of  $f$  component, relative to its independent threshold, when the compound of which it is a part is at its threshold. Vertical axis shows same for  $3f$  component. Observer RS (stimulus duration = 50 msec):  $\bullet$ ,  $\circ$  =  $0.52 + 1.56$  c/deg;  $\blacksquare$ ,  $\square$  =  $0.7 + 2.1$  c/deg;  $\blacktriangle$ ,  $\triangle$  =  $0.87 + 2.62$  c/deg. Observer LA (duration = 100 msec):  $\cdot$ ,  $\diamond$  =  $0.7 + 2.1$  c/deg. Solid symbols =  $0^\circ$  (peaks-subtract) phase, open symbols =  $180^\circ$  (peaks-add) phase. Diamonds on axes are single component thresholds. Diagonal line is prediction of one version of single-channel model for  $180^\circ$  phase. Vertical and horizontal lines are prediction of multiple channel model with narrow tuning and no probability summation. Dotted line is fit of probability summation model; see text for description.

$R_{0.5}$  condition, and it can be seen that when the  $R_{0.5}$  compound stimulus is at its threshold, the  $3f$  component is still present at its independent threshold amplitude, while the  $f$  component is present at 50% of its threshold amplitude. Finally, the  $R_1$  compound is at threshold when each component is at approximately 85% of its own independent threshold, and the other  $R$  conditions occupy various intermediate values between these extremes. It also is evident that the phase of the two components does not systematically affect the detectability of any of the compound stimuli.

It is interesting that the spatial phase of the two components did not affect the observer's sensitivity to the compound grating, even though the manipulation of phase resulted in substantial change in the peak-to-peak disparity. For example, if a component at  $f$  c/deg has disparity amplitude  $A_f^2$  when at its own threshold, then adding a  $3f$  component at 50% of its threshold amplitude,  $A_{3f}^2/2$ , results in a compound grating with peak amplitude equal to  $(0.38 A_f^2 + A_{3f}^2/2)$  in peaks-subtract phase and  $(A_f^2 + A_{3f}^2/2)$  in peaks-add phase. Concretely, for observer RS when  $f = 0.87$  c/deg,  $A_{3f}^2$  was found to be equal to  $2.4 \times A_f^2$ . Hence, for the example just given, the maximum amplitude of the  $R_2$  compound at threshold (expressed relative to  $A_f^2$ ) was 1.58  $A_f^2$  in peaks-subtract phase and 2.2  $A_f^2$  in peaks-add phase. Yet both these compounds were found to be equally detectable, and moreover, were no more detectable than  $f$  alone at amplitude  $A_f^2$ . (These facts about phase are, of course, only interesting if a peak-to-peak detection mechanism is assumed. In contrast, a detection device which integrated disparity over space would respond identically to both phases.)

The results confirm that for the  $R_2$  and  $R_{0.5}$  compound depth gratings, threshold is reached only when at least one of the components is at its own independent threshold, and phase does not matter. Remarkably, the threshold for a cyclopean disparity grating of  $f$  c/deg is unaffected by the presence of another disparity grating of  $3f$  or  $(f/3)$  c/deg, at 50% of its own threshold modulation amplitude. As was discussed in the introduction, this result is the prediction of the multiple-channel model and is inconsistent with the single-channel model. We conclude that the detec-

tion of disparity-modulated cyclopean gratings is mediated by channels selectively tuned to the spatial frequency of modulation, analogously to what occurs in the luminance domain.

When  $0.5 < R < 2$ , however, the two components can be at somewhat lesser an amplitude than their independent threshold amplitudes. This may reflect, in some part, probability summation between mechanisms detecting the two components, and in some part, lack of complete independence between these mechanisms.<sup>1</sup> Without knowledge of either the degree to which noise in the mechanisms is correlated, of the "steepness" of each mechanism's response function, or of the spatial frequency tuning width of the mechanisms, it is impossible to evaluate the contribution of each factor.

However, we have obtained a good fit to the data based on the expression  $(A_f/A_f^2)^p + (A_{3f}/A_{3f}^2)^p = 1$ , with  $p = 4.26$  (shown as dotted curve in Fig. 4). In this expression (derived in the appendix to Graham *et al.*, 1978), the exponent  $p$  reflects the "steepness" of each mechanism's psychometric function (probability-of-seeing curve) according to the Weibull function-type parameterization suggested by Quick (1974); lower values of  $p$  would express increased opportunity for probability summation. Preliminary direct measurements of the psychometric function for one observer suggested that a value of  $p$  in the range 3–5 is probably correct, though a precise estimate, using this method, will evidently be difficult to obtain.

## EXPERIMENT 2

### Introduction

In this experiment we sought to measure the spatial frequency sensitivity bandwidth of the mechanisms whose existence was indicated by the results of the first experiment. We used the threshold elevation technique of Pantle and Sekuler (1968) and Blake-more and Campbell (1969) to demonstrate selective adaptation of individual mechanisms. We measured the threshold amplitudes of disparity gratings following prolonged inspection of an adapting grating of a fixed spatial frequency and compared these with thresholds for the same stimuli without adapting. It is of interest to see (a) if any change in threshold occurs (in particular, an elevation), and (b) if so, how the change in threshold varies as the spatial frequency of the test grating is altered relative to that of the adapting grating. While the precise mechanism of threshold elevation is not well understood, it is generally believed that any threshold elevation is due to the consequences of prolonged activity of the mechanism(s) active during adaptation. Test patterns detected by adapted mechanisms will show threshold elevation while thresholds of those test patterns which are not normally detected by the adapted mechanisms will be unaffected by adaptation. By using a variety of test stimuli, an estimate of the sensitivity profile of the underlying mechanism can be determined.

In conducting this experiment, we also were interested in seeing if we could show the existence of band-pass channels using a different paradigm than the compound grating technique used in the first experiment.

<sup>1</sup> The data reported here have greater curvature than comparable points reported by Graham and Nachmias (1971), but Graham *et al.* (1978) report data nearly identical to ours. The difference in these authors' results, it seems, lies largely in the use by Graham *et al.* of a procedure in which single trials of different concurrently measured stimuli are randomly intermixed, while Graham and Nachmias blocked trials with a particular stimulus. Graham *et al.* suggest that intermixing stimuli causes a "frequency uncertainty effect" leading to a relative decrease in the detectability of simple gratings but not of compound gratings as compared to estimates from blocked runs. This naturally leads to greater curvature in "independence plots" like Fig. 1. With regard to the blocking of stimuli, our procedure was somewhat intermediate between the two, since all conditions were run concurrently, but only on ascending runs was there "frequency uncertainty". We feel that our data does not, consequently, exaggerate the degree of independence that actually exists.

### Methods

The apparatus was the same as in the previous experiment. Observers adapted to suprathreshold amplitude-modulated cyclopean disparity gratings by slowly moving their eyes up and down the central part of the stereoscopic screen, perpendicular to the bars in depth, while fixating the modulating surface. This procedure was designed to eliminate local disparity aftereffects (Blakemore and Julesz, 1971) by insuring that each visual cortex location to which our stimuli projected was exposed to every disparity which was used. That the procedure was effective is indicated by the absence of any perceived negative aftereffect in depth, an effect which can be readily observed after staring at a fixed point on a disparity grating for only a few seconds.

Observers adapted to a disparity grating whose steady-state amplitude was three times the previously determined threshold amplitude for brief presentations of that spatial frequency. The end of the initial adaptation period, which lasted 5 min, was indicated by two rapid blinks of the screen, which lasted 200 msec, and which were followed by the zero disparity conditioning field. Immediately there began a threshold determination for a test grating using the method of limits described in the General Methods section. Two modifications were made in the procedure in order to insure that any adaptation that occurred remained robust: after each three trials 15 sec of adaptation was interposed. Also, each time the ascending or descending criterion was reached, 45 sec of adaptation was interposed.

A session involved measurement of the threshold for one test spatial frequency and consisted of four ascending and four descending runs. Again, threshold was the average of these eight settings. Observer RS adapted to two spatial frequencies, 0.52 and 1.57 c/deg, and observer LA adapted to 1.4 c/deg. Thresholds following adaptation were measured for the adapting spatial frequency and a representative number of other spatial frequencies. As before, stimulus duration was 50 msec for RS and 100 msec for LA.

### Results

We summarized the data by computing a threshold elevation index (TEI) for each test spatial frequency we examined. This was done by taking the ratio of the threshold determined in the adaptation procedure to the average of 5 to 8 measures of the unadapted threshold, for each spatial frequency. For convenience, we then subtracted one from this ratio. This formula results in a TEI of zero when no threshold

elevation occurs and a TEI of 1.0 when threshold has been exactly doubled.

In order to be sure that any change in threshold which might have occurred would not be due simply to a change in procedure (viz. the somewhat tedious interposition of lengthy adaptation sequences between and within threshold runs), we performed several estimations of threshold using the adaptation procedure but with the zero-disparity conditioning field as the adapting stimulus. Thresholds determined in this fashion were indistinguishable from thresholds measured without the adaptation sequences, justifying the comparisons we made in computing the TEI.

Figure 5 exhibits the resulting indices for both observers. The ordinate shows the TEI as a function of the spatial frequency of the test grating, which is plotted on the abscissa. Different symbols show different adaptation frequencies, and arrows near the abscissa show the spatial frequency of the respective adapting gratings. In all cases the TEI is greatest near the adapting frequency and falls off, on both sides, as the spatial frequency of the test grating becomes increasingly dissimilar to that of the adapting grating.

That the inverted U-shape of the TEI curves is not due to experimental error is indicated by the high reliability of our threshold estimates for simple disparity gratings before adaptation (see Fig. 3). For both observers, these estimates had an average standard deviation of about 7% of the threshold amplitude at each spatial frequency. Consequently, a threshold estimate lying 2 SD units above the before-adaptation estimate gives an average TEI of only 0.15 for both LA and RS. Our effects lie well outside this confidence interval and are thus extremely unlikely the result of chance.

In order to appreciate more fully the elevation fall-off shown in Fig. 5, we replot in Fig. 6 the data of Fig. 5 normalized for spatial frequency, with values on the abscissa now expressed as octaves from the adapting spatial frequency. The TEI curves can be seen to have a full-bandwidth at half-amplitude in the range of 2-3 octaves. This value is somewhat larger than the bandwidth of mechanisms tuned to the spatial frequency of luminance modulation, estimated to be about one octave (Blakemore and Campbell, 1969;

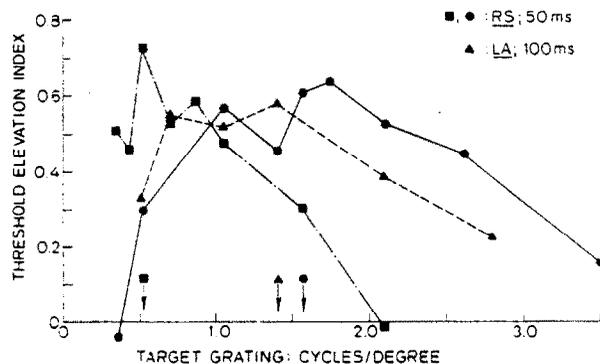


Fig. 5. Results of experiment 2. Abscissa: spatial frequency of disparity modulation of targets whose threshold was measured following adaptation to a disparity-grating. Each symbol represents a different adapting pattern; the spatial frequency of each adapting disparity-grating is shown by an arrow near the abscissa. Ordinate: threshold elevation index (disparity modulation amplitude threshold after adapting  $\div$  threshold for same target before adapting, minus 1.0). Peak elevation occurs near adapting frequency and falls off as target frequency increasingly differs from adapting frequency.

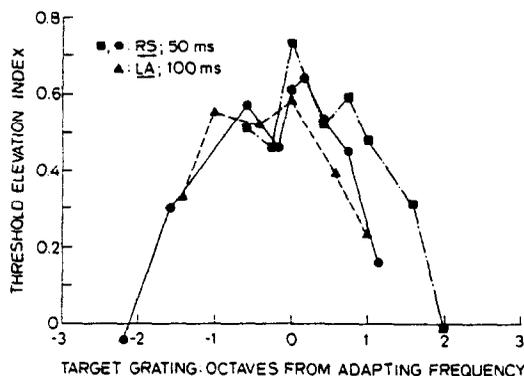


Fig. 6. Same data as shown in Fig. 5, replotted on new abscissa which shows target spatial frequency expressed in octaves from adapting frequency:  $(\log(\text{target frequency} \div \text{adapting frequency}) \div \log 2)$ . This "normalizes" each curve on a logarithmic axis. Ordinate is the same as in Fig. 5.

Stromeyer and Julesz, 1972; Graham, 1972) or less (De Valois, 1977). If we can assume that the mechanisms revealed by adaptation are the same as those studied in the first experiment, then it is evident that a mechanism with peak sensitivity at  $f$  c/deg also has substantial sensitivity at  $3f$  c/deg. This could, in part, explain why in Experiment 1 the points for  $R = 1$  fall somewhat closer to the origin than do comparable points in the study by Graham and Nachmias (1971).

We would however like to suggest caution in the interpretation of our tuning estimate since this figure may be influenced considerably by the characteristics of our display. In particular, we have not explored the effects on tuning width of changing the spatio-temporal characteristics of the dynamic noise or of systematically manipulating the stimulus duration. As a final caution, we cannot prove that the mechanisms estimated by simple vs compound grating detection are indeed the same as those estimated by the adaptation technique. Nevertheless, our evidence does point to the existence of multiple, broadly tuned "cyclopean" spatial frequency channels.

Finally, it is worth noting that there seems to be a high degree of spatial frequency selectivity even at 0.52 c/deg, which corresponds to a fairly slow rate of change of disparity. Some hint that this might be observed was given by Tyler (1975b) who observed that stereoscopic resolution of sinusoidal variations in classical line stimuli is limited to much lower frequencies than is monocular resolution of sinusoidal variation.

#### DISCUSSION

We summarize our conclusions by the following two points: First, the summation experiment (experiment 1) permits us to conclusively reject a single-channel pooling model of global stereopsis. This model may be rejected for both the case where the detection mechanism is a peak disparity detector or a total disparity detector.

Second, the adaptation experiment (experiment 2) both confirms this conclusion, through the demonstration of selective adaptation, and further suggests that the pooling mechanisms possess a rather broad

sensitivity to the spatial frequency of disparity modulation. This interpretation is valid under the following assumptions. Since we used a range of test stimuli to probe the effects of adaptation at one spatial frequency, in order to be able to infer the spectral sensitivity of a single mechanism we assume that all the mechanisms involved in our experimental conditions have approximately the same tuning characteristic. In this case the tuning characteristic of a single mechanism is simply the threshold elevation curve of Fig. 6 reflected about its main vertical axis.

A second assumption that must be made is that threshold elevation is proportional to the effective strength of the adapting stimulus. If threshold elevation increases more slowly than linearly with increases in the strength of the adapting stimulus, then the curves in Fig. 6 are over-estimates of bandwidth. Alternatively, if threshold elevation grows more rapidly than linearly, then these are under-estimates of tuning breadth.

#### Contrast between different kinds of channels

It might be useful to clarify the difference between the present results and those supporting the idea of "disparity channels". Of the latter sort are those showing that the spatial frequency-selective threshold elevation effect for luminance gratings is greatest when the test and adapting pattern are viewed at the same disparity (Blakemore and Hague, 1972; Felton *et al.*, 1972). Those studies have shown that luminance domain channels are tuned not only for spatial frequency and orientation but also for disparity. Also in this class is the negative aftereffect in depth with random-dot stereoscopic figures reported by Blakemore and Julesz (1971) and the perceived-depth-shift aftereffect reported by Mitchell and Baker (1973).

Additionally, our results should be distinguished from those showing that in global stereopsis—that resulting from fusion of random-element half-images—the global mechanisms have similar selectivities in each eye for the spatial frequency of luminance modulation (Kaufman and Pitblado, 1965; Julesz, 1971 pp. 90–102; Julesz and Miller, 1975; Frisby and Mayhew, 1977). Julesz and Miller, for example, showed that stereopsis for two-dimensional spatial frequency bandpass filtered random-dot images was unaffected by two-dimensional filtered noise with a band 2 octaves distant from that of the stereogram, while stereopsis was entirely disrupted when the noise band overlapped that of the stereogram.

For the results we have reported, every stimulus condition involved precisely identical spatial and temporal frequencies of luminance modulation. The manipulation always involved a change in the spatial pattern of binocular correlation between the noise in each eye's view. The characteristics of the noise itself did not change. Our only experimental manipulation was the distribution in space of disparity information.

In order to understand more clearly our domain of stimulation, consider Fig. 7a. Here we have depicted a schematic for the cortical structure which likely underlies early stereoscopic processing. The intersections of the matrix represent a subset of binocular disparity selective neurons all tuned to locations in a single horizontal plane of visual space. The neurons are labeled (in arbitrary anatomical coordinates) with

reference to their sources of innervation on the two retinae. Such a matrix has been called a binocular neural projection field by Boring (1933). (For clarity, we portray the field as a discrete matrix, but each detector should be considered to have a small sensitivity spread in all directions; also, the density of detectors is assumed to be very great.) These detectors represent cortical neurons of a type observed by many investigators (for summary, see Julesz, 1978). All detectors lying on a line having a given coordinate in a particular eye are detectors with the same "line-of-sight" in that eye. As another example, all detectors having identical binocular coordinates (e.g.  $jj$ ) are sensitive to binocular stimulation at anatomically corresponding points. In the lower part of the projection field of Fig. 7a, we have portrayed an illustrative sinusoidal curve, which depicts the neural projection of a sinusoidal disparity input. It was the wavelength of this sinusoidal input that was varied across conditions.

A low-spatial frequency disparity grating consisted of disparity cues between the right and left eye's images which changed *slowly* and sinusoidally as a function of space; a high-spatial frequency disparity grating's disparity cues changed *more quickly* as a function of space. We interpret our finding of multiple channels for this particular parameter of stimulation to mean that in the human visual system, there are mechanisms for detecting broad surfaces in depth, where depth changes are small with changes in spatial position, and that there are separate and independent mechanisms for detecting abrupt changes in depth. This suggests that the visual system has channels where broad spatial pooling among mechanisms tuned to the same disparity occurs and other channels where only narrow pooling occurs. In other words, at a place in the visual system at or beyond the locus

where binocular convergence occurs, the visual scene is dissected into separate channels selectively tuned to different slopes of depth changes, i.e. selectively sensitive to different values of the first derivative of disparity as a function of space.

### Global stereopsis

A central problem in global stereopsis concerns the accomplishment of a single, dense, depth surface and the suppression of numerous subsidiary surfaces which might arise due to the essential ambiguity of local stereoscopic cues; a related issue involves the resulting global stability of the stereopsis system (see, for example, Julesz and Johnson, 1968; Fender and Julesz, 1967; Julesz and Chang, 1976). Several current "cooperative" models of stereopsis (e.g. Julesz, 1971, 1978; Sperling, 1970; Dev, 1975; Nelson, 1975, 1977; Marr and Poggio, 1976; Marr *et al.*, 1978) derive a solution to this problem by invoking some degree of global and non-linear spatial pooling (i.e. of cooperativity), of disparity selective "analytic" units tuned to approximately the same disparity, and some degree of recurrent inhibition among units differing in disparity tuning.

An analogy may be drawn between such a cooperative mechanism for stereopsis and certain visual processes of the retina. One function of the retina, carried out by the photoreceptors, is to register the spatial distribution of light energy falling on the eye. Similarly, one function of the stereopsis system must be to represent the spatial pattern of disparity which results from binocular viewing. Cooperative processing is required because even this most basic task is complex, as revealed by random-dot stimulation. As the consequence of cooperative processing, noise is eliminated and stable and correct disparity detection takes place.

In the case of the luminance domain, mechanisms subsequent to the receptors must exist to carry out processing in the interests of higher perceptual functions; much has already been learned about their nature. Little attention, however, has been paid to the mechanisms which, analogously, must continue the processing of disparity information. And, just as spatial frequency selectivity in the luminance domain implies an organization beyond that of photoreceptor detection, we believe the existence of the mechanisms we have investigated implies an organization beyond that of disparity detection. Thus, mechanisms selective for the spatial frequency of disparity modulation could represent an early stage in stereo-form processing. What might be the structure of these mechanisms?

Let us examine first the sort of model most frequently proposed for global stereopsis. One version of such a model (Nelson, 1975) results in an iterated local neural spread function in which each disparity detecting neuron exerts inhibitory or facilitatory force upon its neighbors in the disparity or space domain, respectively. This is basically a single-channel model in which there is fairly broad excitatory pooling (and virtually no inhibition) along a spatial domain combined with broad inhibitory pooling along a disparity domain. One of the weaknesses of such a single-channel model is that while it can explain why broad stereo surfaces can be generated with only a very low

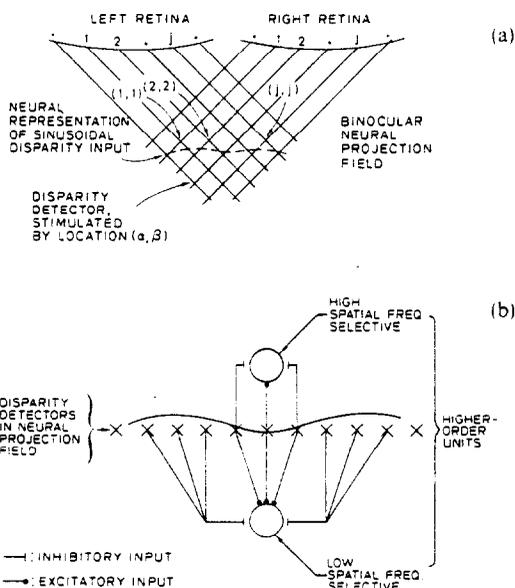


Fig. 7. (a): Schematic portrayal of cortical structure for early stages of stereoscopic processing. For description, see text. (b): Schematic for higher-level stereoscopic processing exhibiting, in particular, selectivity for the spatial frequency of disparity modulation; described more fully in text.

density of random dots (see for example, Julesz, 1971, Fig. 4.5–5), it cannot explain why stereogram-generated objects in the cyclopean field are perceived with sharp corners and edges. How is the high spatial frequency component mediated? Also, such a model fails to account for the existence of low stereo frequency band-limitation, which we observe in both the overall amplitude sensitivity function (Fig. 3) and in individual adaptation curves (Fig. 5).

Our results indicate that such a model of cooperativity among disparity analytic units must be modified or at least supplemented in the following regards. First, the system has multiple independent channels, each characterized by different spatial extents of pooling. The low spatial frequency channels are characterized by more extensive pooling, even among small-disparity mechanisms; correspondingly, the high spatial frequency channels are characterized by relatively local pooling of disparity information. This allows different mechanisms to mediate the perception of large surfaces which are fairly homogeneous in depth and an independent set of mechanisms to mediate the perception of highly curved surfaces in depth. Stereo-edges would then be perceived via the latter type of channel.

Secondly, the channels, as measured in Experiment 2, are bandwidth-limited, both on the high and low spatial frequency side. This indicates the presence of inhibition along the spatial domain: the pooling of units which are sensitive to the same disparity has a spatial characteristic of an excitatory-center antagonistic-flank type. Such inhibitory flanks would lead to low-frequency insensitivity as the result of a decrease in response when a flat or nearly-flat segment of a stereo-surface exceeds a characteristic extent. Anstis *et al.* (1978) reached a similar conclusion based on their observation of a stereoscopic-depth analogue of the Craik–O'Brien–Cornsweet illusion.

One possible organization which would be successful in modeling our results is hierarchical; a schematic is provided in Fig. 7b. It is possible that higher-order integrating units pool the responses of spatially neighboring detector-level units of similar disparity tuning such that there is antagonism between flanking regions. Different higher-order units might have central excitatory input from spatial regions of varying size, but each higher-order unit would have inhibitory inputs from spatially adjacent regions of the same disparity as the center region. In Fig. 7b, two higher-order units are illustrated, each with a fixed configuration of input from detector elements of the neural projection field (Fig. 7a). One unit is shown with a small extent of spatial pooling, the other with a larger pooling extent. We assume that many different extents can and do exist. The optimal stereoscopic stimulus for such a unit would be a surface that had fusible pattern at the critical disparity but only over the region of space just covered by the center of the unit's "cyclopean receptive field"; in the flanks of this "receptive field" an absence of fusible pattern at the critical disparity would be the optimal feature. The sine-wave shown in Fig. 7b is such an optimal stimulus for the low frequency selective unit. If the central region of pattern were too large or too small for a given higher-order unit, then the inhibitory region would be stimulated or the central region would be

under-stimulated, respectively, and the response of that unit would be sub-optimal. For the high frequency selective unit of Fig. 7b, the former obtains and the sine-wave shown would cause little or no activity.

Populations with different "cyclopean receptive field" center size would thus respond best to different spatial frequencies of disparity modulation. However, the spatial frequency tuning of all higher-order units would be broad. Since Tyler (1975a) found tilt-after-effects using disparity gratings, the "receptive fields" we describe might well be elongated and orientation-selective. Finally, we note that it is possible to devise a simple arrangement of the tuned excitatory and tuned inhibitory depth-sensitive cells discovered by Poggio and Fischer (1977) which could easily achieve the organization we describe.

*Acknowledgements*—We wish to thank Lucy Anderson for her help as an experimental subject, Cecilia Bahlman for help in preparation of the manuscript, Ben Dawson for generous technical assistance, and Drs Bela Julesz, Gerald Long, Christopher Tyler, and Brian Wandell, as well as an anonymous referee for this journal, for helpful comments on an earlier draft.

This work was supported, in part, by NIH Integrative Biology Training Grant GM 07181 to Robert Schumer and Program Project NICHD HD 09814 to Leo Ganz.

#### REFERENCES

- Anstis S. M., Howard I. P. and Rogers B. (1978) A Craik–O'Brien–Cornsweet illusion for visual depth. *Vision Res.* **18**, 213–217.
- Anstis S. M. (1975) What does visual perception tell us about visual coding? In *Handbook of Psychobiology* (edited by Gazzaniga M. S. and Blakemore C.). Academic Press, New York.
- Blakemore C. and Campbell F. (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol.* **203**, 237–260.
- Blakemore C. and Hague B. (1972) Evidence for disparity detecting neurones in the human visual system. *J. Physiol.* **225**, 437–455.
- Blakemore C. and Julesz B. (1971) Stereoscopic depth aftereffect produced without monocular cues. *Science* **171**, 286–288.
- Blakemore C., Nachmias J. and Sutton P. (1970) The perceived spatial-frequency shift: Evidence for frequency-selective neurons in the human brain. *J. Physiol.* **210**, 727–750.
- Blakemore C. and Sutton P. (1969) Size adaptation: A new aftereffect. *Science* **166**, 245–247.
- Boring E. G. (1933) *The Physical Dimensions of Consciousness*, p. 118. The Century Co. New York.
- Campbell F. W. and Maffei L. (1971) The tilt aftereffect: a fresh look. *Vision Res.* **11**, 833–840.
- Dev P. (1975) Perception of depth surfaces in random-dot stereograms: A neural model. *Int. J. Man-Machine Studies* **7**, 511–528.
- De Valois K. K. (1977) Spatial frequency adaptation can enhance contrast sensitivity. *Vision Res.* **17**, 1057–1065.
- Felton T. B., Richards W. and Smith R. A. (1972) Disparity processing of spatial frequencies in man. *J. Physiol.* **225**, 349–362.
- Fender D. H. and Julesz B. (1967) Extension of Panum's fusional area in binocularly stabilized vision. *J. opt. Soc. Am.* **57**, 819–830.

- Frisby J. P. and Mayhew J. E. W. (1977) Global processes in stereopsis: Some comments on Ramachandran and Nelson (1976). *Perception* **6**, 195–206.
- Ganz L. (1966) Mechanism of the figural aftereffects. *Psychol. Rev.* **73**, 128–150.
- Graham N. (1972) Spatial frequency channels in the human visual system: effects of luminance and pattern drift rate. *Vision Res.* **12**, 53–68.
- Graham N. (1979) Spatial-frequency channels in human vision: detecting edges without edge detectors. In *Visual Coding and Adaptability* (edited by Harris C. S.), Erlbaum, Hillsdale, N.J.
- Graham N. and Nachmias J. (1971) Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channel models. *Vision Res.* **11**, 251–259.
- Graham N., Robson J. G. and Nachmias J. (1978) Grating summation in fovea and periphery. *Vision Res.* **18**, 815–825.
- Julesz B. (1971) *Foundations of Cyclopean Perception*. University of Chicago Press, Chicago.
- Julesz B. (1978) Global stereopsis: Cooperative phenomena in stereoscopic depth perception. In *Handbook of Sensory Physiology* (edited by Held R., Leibowitz H. and Teuber H.-L.), Vol. VIII. Springer, Berlin.
- Julesz B. and Chang J.-J. (1976) Interaction between pools of binocular disparity detectors tuned to different disparities. *Biol. Cybernetics* **22**, 107–119.
- Julesz B. and Johnson S. C. (1968) Stereograms portraying ambiguously perceivable surfaces. *Proc. natn. Acad. Sci.* **61**, 437–441.
- Julesz B. and Miller J. (1975) Independent spatial-frequency-tuned channels in binocular fusion and rivalry. *Perception* **4**, 125–143.
- Kaufman L. and Pitblado C. B. (1965) Further observations on the nature of effective binocular disparities. *Am. J. Psychol.* **78**, 379–391.
- Kulikowski J. J. and King-Smith P. E. (1973) Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. *Vision Res.* **13**, 1455–1478.
- Marr D., Palm G. and Poggio T. (1978) Analysis of a cooperative stereo algorithm. *Biol. Cybernetics* **28**, 223–239.
- Marr D. and Poggio T. (1976) Cooperative computation of stereo disparity. *Science* **194**, 283–287.
- Mitchell D. E. and Baker A. G. (1973) Stereoscopic aftereffects: evidence for disparity specific neurones in the human visual system. *Vision Res.* **13**, 2273–2288.
- Nachmias J. (1967) Effect of exposure duration on visual contrast sensitivity with square-wave gratings. *J. opt. Soc. Am.* **57**, 421–427.
- 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: Aftereffects, illusions and horopter shifts in depth perception. *J. theor. Biol.* **66**, 203–266.
- Pantle A. and Sekuler R. (1968) Size-detecting mechanisms in human vision. *Science* **162**, 1146–1148.
- Poggio G. F. and Fischer B. (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving Rhesus monkey. *J. Neurophysiol.* **40**, 1392–1405.
- Quick R. F. (1974) A vector-magnitude model of contrast detection. *Kybernetik* **16**, 65–67.
- Sachs M., Nachmias J. and Robson J. G. (1971) Spatial-frequency channels in human vision. *J. opt. Soc. Am.* **61**, 1176–1186.
- Sperling G. (1970) Binocular vision: A physical and a neural theory. *Am. J. Psychol.* **83**, 461–534.
- Stromeyer C. F. and Julesz B. (1972) Spatial-frequency masking in vision: Critical bands and spread of masking. *J. opt. Soc. Am.* **62**, 1221–1232.
- Tyler C. W. (1974) Depth perception in disparity gratings. *Nature* **251**, 140–142.
- Tyler C. W. (1975a) Stereoscopic tilt and size aftereffects. *Perception* **4**, 187–192.
- Tyler C. W. (1975b) Spatial organization of binocular disparity sensitivity. *Vision Res.* **15**, 583–590.