PERGAMON

Vision Research 40 (2000) 3245–3255

Vision Research

www.elsevier.com/locate/visres

Relative contributions of sustained and transient pathways to human stereoprocessing

Leonid L. Kontsevich, Christopher W. Tyler *

Smith–Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115, USA

Received 18 January 1995; received in revised form 9 February 2000

Abstract

It has been proposed [Hubel & Livingstone (1987) *Journal of Neuroscience*, *7*, 3378–3415] that stereopsis is mediated solely by magnocellular pathway in primates. This hypothesis was evaluated for humans in psychophysical experiments with dynamic random-noise stimuli, based on the sustained/transient relationship of behavior mediated by the two divisions of the LGN [Merigan & Maunsell (1993) *Annual Review of Neuroscience*, *16*, 369–402]. The stereoscopic limits show that stereoscopic system is more sensitive to sustained random-dot stimuli than to transient ones. Quantitative modeling of the result implied a weak role for magnocellular input, suggests that human stereopsis is more strongly influenced by parvocellular input through the LGN.

Keywords: Stereopsis; Disparity; Sustained; Transient; Parvo; Magno

1. Introduction

The clear anatomical and functional distinction between visual parvocellular (P) and magnocellular (M) pathways from the retina through the primate lateral geniculate nucleus (LGN) to the cortex (for review see Kaplan, 1991; Schiller, Logothetis & Charles, 1990; Casagrande & Norton, 1991) suggests that there could be perceptual tasks specific for each of these pathways. An important problem for psychophysics and neurophysiology, therefore, is to identify these tasks. On the basis of physiological data of the binocularity of receptive fields in monkey cortex, such an attempt was made for stereoscopic processing by Hubel and Livingstone (1987) and Livingstone and Hubel (1987), who advanced the hypothesis that primate stereopsis is mediated solely by the M pathway.

This hypothesis, although based on reasonable physiological evidence, is counterintuitive. Only about 10% of LGN neurons are magnocellular in primate, while stereopsis is a demanding computational task requiring an extensive neural substrate (Tyler, 1977). Moreover, M cells generally have large receptive fields and are selectively sensitive to transient stimuli, in contrast to P cells, which have substantially higher spatial resolution and typically have a sustained or lowpass temporal sensitivity (see Derrington & Lennie, 1984). M-cell properties therefore seem suited for processing coarse and dynamic aspects of stereopsis and for stereomovement processing. Other aspects of stereoscopic depth perception are the ability to appreciate depth in a static task based on fine spatial resolution, which seems a better match for the properties of the P cells. It is therefore hard to envisage how P cells would play no role in stereoscopic processing. It seems possible that Hubel and Livingstone (1987) missed the P-system contribution because they did not evaluate the selectivity of these cells with disparities of less than 15 arcmin, the level that would be required to probe the fine disparity processing range (Norcia, Sutter, & Tyler, 1985).

To evaluate the purely magnocellular hypothesis, we need to consider the roles of the various processing stages involved in depth perception. Neurophysiological studies are typically designed on the assumption that any cell that responds selectively to disparate stimuli is a specifically stereoscopic neuron, but this view may be too restrictive. A more elaborated view of the stereo-

* Corresponding author. Fax: +1-415-3458455.
E-mail address: cwt@ski.org (C.W. Tyler).

0042-6989/00/$ - see front matter
PII: S0042-6989(00)00159-0

scopic system was provided by Tyler (1983, 1991) and others, in which at least five roles of disparity-selective neural mechanisms may be envisaged.

1. The initial disparity registration or local matching between the two eyes.
2. Solving the correspondence problem among the large number of potential spurious matches to form a unified stereoscopic depth image.
3. Combining the depth information from multiple depth cues in to a generic depth map.
4. Scaling the generic depth map to a calibrated distance map.
5. Processing the configurational form properties of the depth image (such as the orientation, elongation and sharpness of the various shapes making up the depth image).

Neurons at each of these processing levels may be selective for binocular disparity (although not necessarily so at the higher stages). It is perhaps not surprising, therefore, that disparity selective neurons have been found in many primate visual areas such as V1, V2, V3, MT and MST. Of course, if the monkey is trained in a disparity-image/flat-plane depth discrimination, then any cell supporting this discrimination performance will exhibit disparity tuning, even in frontal cortex. So disparity selectivity is not a definitive indicator of a tight coupling to an early disparity matching process.

For example, DeAngelis and Newsome (1999) find that disparity-tuned neurons are organized into cortical columns by preferred disparity, and that preferred disparity is mapped systematically within larger, disparity-tuned patches of area MT. Combined with other recent findings that electrical stimulation of disparity clusters in MT can bias responses of depth judgments (DeAngelis, Cumming, & Newsome, 1998), the data suggest that MT plays an important role in depth perception in addition to its well known role in motion perception. However, the level of stereoscopic disparity processing in monkey MT may not be just a simple disparity matching process. It may be that MT combines sensory inputs into a multisensory depth map, with inputs from motion, disparity, shape-from-shading, perspective, and so on. In this case, MT would be viewed as the site of the generic depth image, which is a broader concept than that of a simple disparity tuning, even though the generic depth map would exhibit disparity tuning if the input were purely stereoscopic.

Conversely, there is clear evidence that the P system participates in stereopsis in primates. Schiller et al. (1990) show that PLGN lesions in monkey disrupted behavioral detection of a cyclopean stereotarget, whereas MLGN lesions had no effect on performance for these fine random-dot targets. This experiment provides a double dissociation between the M system and fine stereopsis, although coarser targets should be expected to have stimulated the M system. Similarly, DeYoe and Van Essen (1988) report that P cells in the interblob regions of area V2 are binocular, suggesting a role in stereoscopic vision although they did not describe the disparity tuning. (The interblob system may also have inputs from the M pathway, however.) Moreover, Tyler (1990) has argued on the basis of a wide range of psychophysical data that stereopsis appears to be well-supported in both the color-sensitive blob and luminance-sensitive interblob subdivisions of the P stream, with M stream participation in stereopsis limited to the range of coarse disparities.

The goal of the experiments was thus to compare visual performance in stereoscopic tasks for stimuli that reveal the signature of the M or the P pathway. The magnocellular hypothesis would predict that the stimuli tuned for M-pathway response would be more effective than those tuned to P-pathway response.

2. Experimental rationale

The goal of the experiments was to compare visual performance in stereoscopic tasks for stimuli that reveal the signature of the M or the P pathway. The M and P pathways have been assessed with many techniques at many levels of neural processing, often with apparently contradictory results. For simplicity we adhere to the original definition that the M processing is mediated by signals passing through the magnocellular layer of the LGN while the P processing is mediated by signals passing through the parvocellular layer of the LGN. The fate of, or neural mechanisms involved in, the subsequent processing of these signals is an issue for neurophysiologists to resolve. For our purpose, the question is whether there is a significant difference between the net behavioral properties of the processing associated with these two neural streams.

The basis for such a distinction is described in the review by Merigan and Maunsell (1993), who combined the results of several psychophysical studies in monkeys (when either the P or the M layers of the LGN were chemically blocked) to estimate the temporal and spatial modulation threshold functions (MTFs). The data (reproduced in Fig. 1) reveal that the P pathway has lowpass temporal MTF with cutoff frequency at 1–2 Hz while the M pathway is bandpass with the peak at 10 Hz. Within the range 2–10 Hz, the two pathways have opposite slopes in their sensitivity profiles; these slopes provide the signature for which pathway is mediating performance in the stereoscopic task. A positive slope indicates a superiority of transient stimulation while a negative one implies a superiority of sustained stimulation. If stereoscopic processing receives input only from the M stream, performance with transient stimuli should be better than with sustained; if only the

P system were involved, sustained stimulation should be superior to transient[1].

It is important to note that the two behaviorally evaluated pathways established by Merigan and Maunsell (1993) were complementary in their spatiotemporal properties. Although the M system characteristic revealed by the lesion of the parvocellular layer was temporally bandpass, it was spatially lowpass. Conversely, the temporally lowpass response of the P system revealed by the lesion of the magnocellular layer was spatially bandpass. The key aspect of this pattern concerns the sensitivity at low temporal frequencies, where the slow M system sensitivity is low for uniform fields and remains low for all spatial frequencies (as are present in the random-dot stimuli that we need for the stereo studies). Conversely, the high P system sensitivity for uniform fields at low temporal frequencies remains high for a wide range of spatial frequencies. Thus, the spatial characteristics of Merigan and Maunsell results (and the extended interpretation that they provide for these results) imply that the temporal signatures they report will remain effective across spatial frequency,

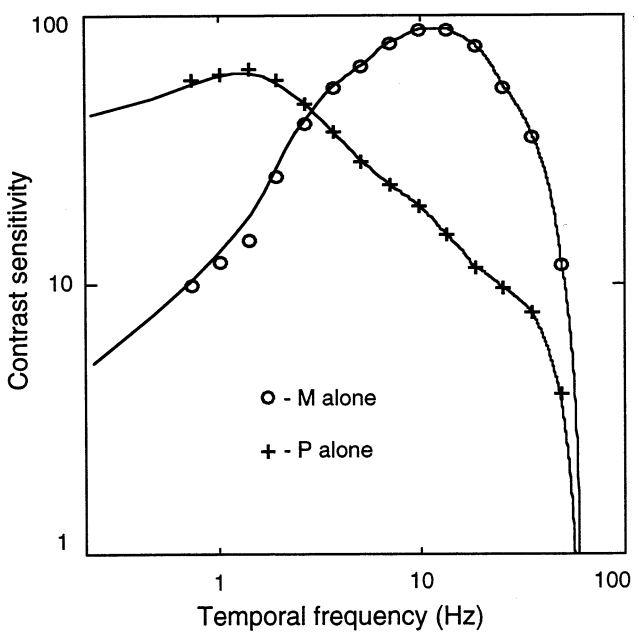

Fig. 1. The symbols represent the psychophysical temporal contrast sensitivity profiles of the P and M pathways determined from chemical lesions in the M and P layers of monkey LGN. The solid curves represent the rational-function fits we used for the impulse response reconstruction. Data adapted from Merigan and Maunsell (1993) with permission.

[1] This logic might be suspect if the stereoscopic system performed additional low-pass temporal filtering that would distort the slopes of the temporal MTFs of M and P pathways. Such a possibility is ruled out by Smallman and MacLeod (1994) who have demonstrated that the ratio between contrast sensitivities for detection and for stereopsis does not vary notably between brief 150 ms and long 2 s stimulus durations.

and justify the assumption that they are valid for the random-dot stimulus arrays used in our study.

To evaluate the relative M/P contributions to stereopsis, we designed dynamic random-dot stereograms in which the lifetime of each dot that appeared in the dynamic noise stream was defined by either the transient temporal profile of up/down pulse doublets or the sustained temporal profile of a raised cosine bell waveform (see Section 3). The quantitative ratios of the effective contrast of these controlled lifetime stimuli for each mechanism were determined by modeling the effective contrast ratio for the sustained and transient stimuli in a computer simulation.

2.1. Effective contrast in M and P pathways

To obtain numerical estimates for the signals evoked by temporally modulated noise in the P and M pathways, we made three assumptions:

1. The sensitivity of the P and M systems in psychophysical contrast threshold tasks is the same for humans and monkeys, so that the data from Merigan and Maunsell (1993) shown in Fig. 1 are representative of human response properties.
2. The shapes of the temporal MTF curves mediated by the M and P subdivisions of the LGN (see Fig. 1) are assumed to be valid for our stimuli, although originally they were measured with the stimuli of lower spatial frequency content (as discussed in the previous section).
3. The impulse response functions for the both pathways comply with the minimum phase assumption (that the reconstructed impulse response functions are the most compact in time relative to zero delay; Victor, 1989).

We approximated the temporal MTFs for the M and P streams with rational functions (the ratio of two high-order polynomials) that provided an excellent fit to the data (Fig. 1). Because these rational functions conform to the minimum phase assumption, they allow inverse Fourier transformation to generate the minimal-phase impulse response functions shown as the insets in Fig. 2B,C. To avoid ringing artifacts in the approximation, we added extra points in the high-frequency range above the cut-off frequency of the MTFs to diminish the high-frequency energy. Inclusion of points according to a slope of -6 up to 500 Hz was effective in minimizing the high-frequency ringing in the reconstructed impulse responses (Dagnelie, 1992).

To obtain the estimated response of the M and P systems to our controlled-lifetime stimuli, the reconstructed impulse response functions were convolved with the temporal luminance modulation for a summation field of a given size; examples of the sustained and transient modulations in a single pixel and their convolutions with the M and P impulse responses are shown

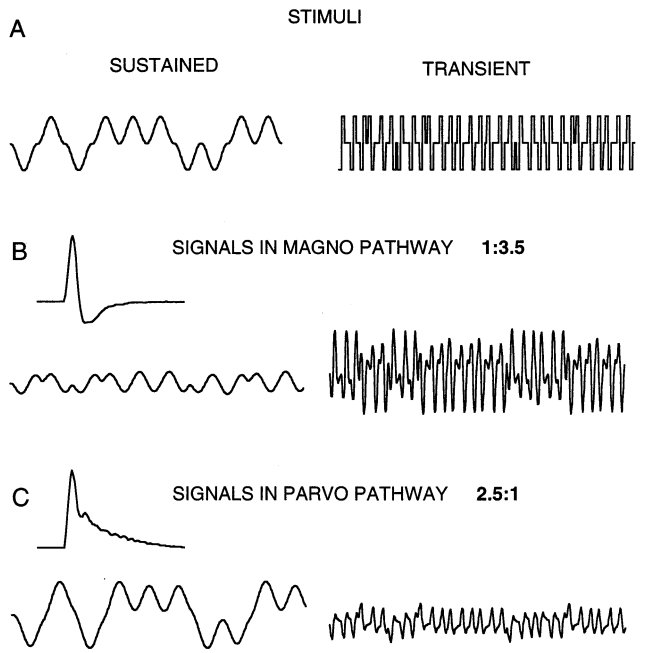

Fig. 2. (A) Three-second samples of luminance profiles for one pixel of the dynamic noise stimuli. Left-sustained profile, right-transient profile. (B) M-impulse response reconstructed from the monkey contrast sensitivity data presented in Fig. 1 and the result of its convolution with the sustained and transient signals from (A). The effective contrast ratio is about 3.5:1 in favor of M system. (C) As (B) but for the P impulse response. Now the effective contrast ratio is about 2.5:1 in favor of P system.

in Fig. 2. The temporal profile of the convolved response was integrated over time to obtain the effective contrast according to the standard power summation rule:

$$E = \left(\int |R(t)|^{\beta} \, \mathrm{d}t \right)^{1/\beta} \quad (1)$$

We considered four different power summation rules for combining the effective contrast over time: average magnitude ($\beta = 1$), square root of energy ($\beta = 2$), fourth power probability summation ($\beta = 4$) and winner-take-all ($\beta = \infty$) decision rules. We evaluated the effect of these rules for four different sizes of linear spatial summation fields: 1, 3, 10 and 100 pixels. The computed ratios between the effective contrasts for sustained and transient modulations within each pathway are presented in Fig. 3. These results show that the ratios are largely insensitive to variation of the summation area and the power summation parameter β. The P pathway is about twice as sensitive to the sustained stimulation profile than to the transient; the M pathway conversely is about three times more sensitive to the transient stimulation than to the sustained, a combined separation of a factor of six. As pointed out in the previous section, the monkey spatiotemporal data, on which the operational M/P distinction for this study is based, imply that similar relationships should hold at all spatial frequencies.

2.2. Application to stereopsis

To define the disparity sensitivity under each temporal regime, we measured the disparity detection threshold and the upper disparity limit as a function of the (vertical) width of a horizontally elongated cyclopean bar[2] (as depicted in Fig. 4). The values of these two threshold disparities across a range of widths defined the boundary of the detectability region for stereoscopic depth (Fig. 5). The region to the left of each curve defines the conditions where the stereoscopic depth of the bar stimulus is not resolved. Increasing the bar width shifts performance into the zone to the right of the curve where the target's depth is discriminable. Comparison of such curves rather than isolated points should provide greater sensitivity to any differences in processing of the sustained and transient stimuli that may exist.

A particular prediction of this analysis is that, if stereopsis gets inputs from only one of the neural streams, the shapes of the detectability regions should be identical for sustained and transient stimuli when their effective contrasts are equated. Moreover, if stereopsis is driven solely by P input, the ratio of P responses to the two stimuli (Fig. 3) predicted that the threshold curve for the sustained stimuli should be to the left of the curve for the transient stimuli, as shown in the left panel of Fig. 5. Conversely, if the input is provided by the M pathway, as predicted by Hubel and Livingstone (1987), the ratio of M system to the two stimuli (Fig. 3) predicts that the arrangement of the threshold curves should be the opposite (Fig. 5, right panel). Under the assumptions of this analysis, our experiment thus provides a test of which of the two systems dominates performance in human discrimination of depth in random-dot targets.

3. Methods

3.1. Stimuli

The dynamic stereogram displays were generated by a Macintosh IIfx computer on a monochrome Apple High-Resolution monitor with P31 phosphor. The observers fixated a central fixation spot to view the dynamic stereogram in the dark from the distance of 1.24 m (to provide a pixel size of 1 arcmin), which was maintained by a forehead support. The size of each monocular image was 5 deg in width and 8 deg in height.

[2] Our stimulus is a bar elongated in the horizontal direction. The *length* parameter specifies the horizontal dimension of the bar and *width* defines the vertical dimension.

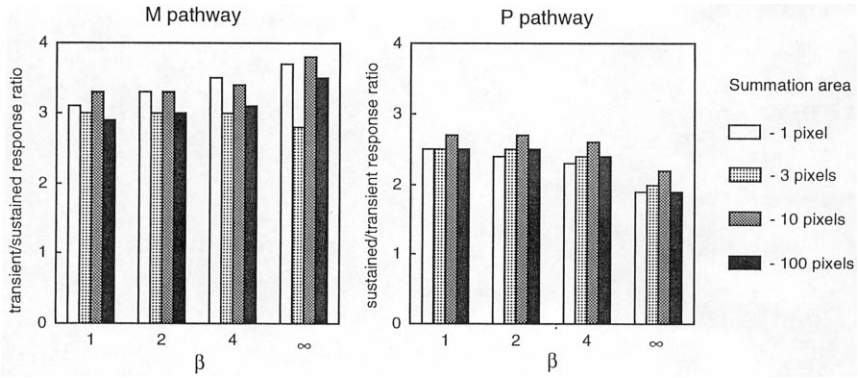

Fig. 3. The results of the simulation of the responses of M and P pathways to the transient and sustained stimuli. Left panel. Each set of bars shows the calculated transient/sustained ratios over a range of field sizes, calculated as a function of the power summation parameter β and the summation area. The ratios for the M pathway are always well above the unity level and those for the P pathway are always well below. The ratios for the both channels are largely insensitive to variation of the summation area and the power summation rule.

All pixels in the dynamic random-dot stimuli fluctuated quasi-independently in either a transient or sustained regime around the mean luminance level of 50 cd/m^2. Specifically, the screen was subdivided into two fields interleaved in a one-pixel checkerboard fashion, to allow switching between two different stereogram image pairs. Each of the 72 000 dots of a stereogram was assigned randomly to one of 60 entries of the look-up table to form a random sub-field of 600 screen-pixels per entry. The test presentation was produced solely by means of random permutation of equal numbers of black and white assignments to the entries in the look-up table. Because the number of independent entries was large, the probability of recurrence of the same subgroup of positions was negligible. Use of equal numbers of black and white assignments in this algorithm eliminated any systematic bias of the average luminance in the dynamic noise stream that would occur for purely random assignment to the look-up table.

The sustained noise modulation consisted of single-cycle raised-cosine modulation of the form $\pm[1-\cos(2\pi t/300)]$, forming a smooth 300 ms ramping of the luminance of each spot either up or down from the baseline gray to its maximum contrast excursion and back to zero in the 50% of pixels assigned to the stereogram field to be presented (Fig. 2A, left) (The other, non-presented field of 50% of the pixels remained grey throughout the trial). The 300 ms ramping sequence was initiated for 10% of the dots in every 30 ms frame, providing ten such overlapping sequences to make up the sustained noise stimulus. At the start of each of the ten sequences, the ramping polarities were randomly permuted among the six random sub-fields of 600 screen pixels controlled by each sequence. The resulting noise pattern modulated smoothly in a continuous random fashion to form the appearance of oily undulation.

The modulation sequence for transient stimulation at each pixel consisted of a two-frame doublet of a positive 30 ms pulse followed immediately by a negative 30 ms pulse, repeated three times at irregular intervals within each 300 ms epoch. The sequence of pulse doublets was again initiated randomly with either positive or negative sign for 10% of the dots on any frame (Fig. 2A, right), making ten such overlapping sequences

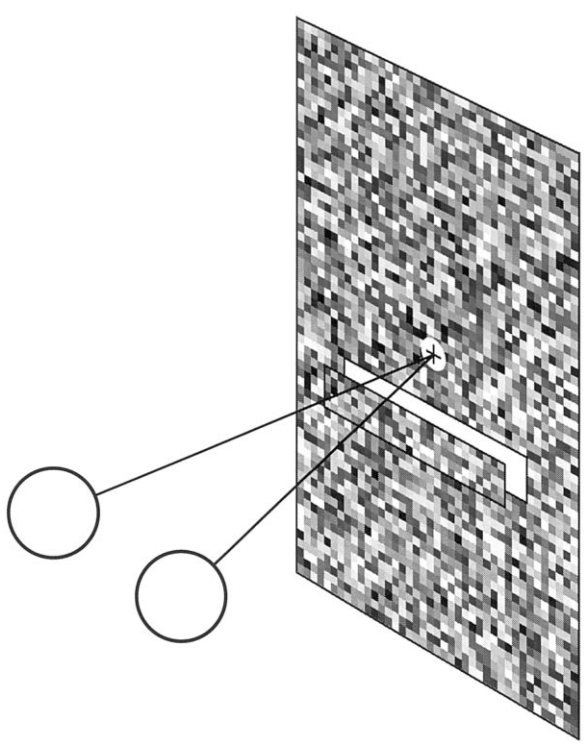

Fig. 4. The cyclopean disparity bar consisted of a horizontal rectangle of constant 200 arcmin length and variable (2–64 arcmin) vertical width, either in front of or behind the reference plane on any trial. The center of the cyclopean target was 1 deg below the fixation point. Disparities of both the plane and the target were specified by the dynamic random dot fields; the dots were modulated either slowly (sustained conditions) or rapidly (transient conditions).

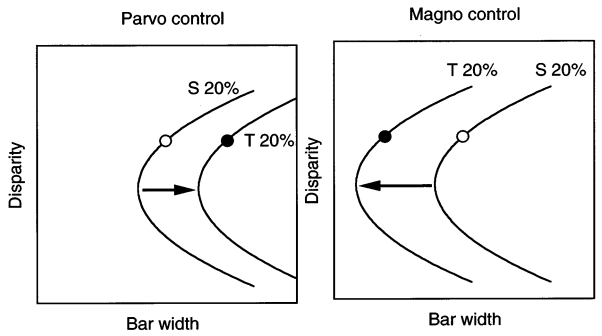

Fig. 5. Relative predictions for threshold contours in the disparity/bar width plane of the cyclopean target under the two temporal regimes. Depth is discriminable only to the right of the curve in each experimental condition. When the stereoscopic system gets input solely from the P pathway, discriminability for transient stimuli should be poorer than for sustained as shown in the left panel; if the M pathway controls stereo function, the relative position of the threshold curves should be reversed (right panel).

to form the transient noise stimulus with a staccato appearance. The pixel modulation magnitude was identical for all pixels, around the mean texture luminance of 50 cd/m^2.

The Michelson contrasts of both types of 2D noise were 20% for the transient stimuli and either 10 or 20% in different experiments for the sustained stimuli. It should be noted that these values are in the low range of effective contrasts; the contrast threshold for both stimulus types was about 3%. This can be understood in terms of the broad spectrum for noise relative to the narrow spectrum for classical grating stimuli. If expressed in terms of a 2D spatial frequency plot, the roughly circular area (strictly, volume) covered by a typical orientation-selective receptive field is of the order of 1/10th of the area under the doughnut of the overall MTF (see De Valois and De Valois, 1990, p. 271, Fig. 9.4). Thus, 3% contrast for the noise corresponds to the expected threshold of about 0.3% contrast for a grating that is specific for a particular receptive field or visual mechanism mediating performance. Correspondingly, at the 20% contrast level, the equivalent grating contrast of the noise that we used was about 2% for any receptive field mechanism.

The cyclopean test stimulus presented in the noise patterns consisted of a stereoscopic plane, near the center of which was a horizontal cyclopean bar either in front of or behind the plane (see Fig. 3). The length of the bar was 200 arcmin; its vertical width was varied from 2 to 64 arcmin. The bar was positioned at 1 deg above the fixation point because stereothresholds at smaller eccentricities were too fine to be measured with the pixel resolution of the display. This distance is somewhat smaller than 6 deg eccentricity (in the temporal direction of the fovea on the horizontal meridian) at which the monkey sensitivity curves from Fig. 1 were measured.

Between trials, the two monocular images of the dynamic noise were anticorrelated, thus providing no organized depth information. Fixation was maintained on the central fixation square four pixels on a side without difficulty. When the trial began, the test disparity configuration replaced the anticorrelated dynamic noise for 1 s. The switchover from one pattern to another was gradual, accomplished by switching the disparity of 10% of the dots on every 30 ms frame to complete the switchover in a duration of 300 ms.

3.2. Procedure

The targets were presented when the observer pressed a key to indicate that the eyes were converged stably at the fixation point, which the observers were adjured to fixate throughout the experiment. That they succeeded is indicated by the fact that the disparity thresholds always exceeded 4 arcmin, for if they had shifted fixation to the target the thresholds would have fallen to below 1 arcmin, the disparity resolution of the display. The experimental task was a single-interval two-alternative forced-choice discrimination: observers were instructed to indicate by a key-press whether the bar was presented in front or behind the fixation plane on each stimulus presentation. The disparity between the bar and the plane was constant in each block. Decision time was unlimited; the response initiated the start of the subsequent trial with no delay.

Measurements were conducted in blocks of 20 trials; in each block the probability of correct discrimination was measured for a particular combination of bar width and disparity magnitude. For each width/disparity pair, the cyclopean bar width could be set to 2, 4, 8, 16, 32 or 64 arcmin and the disparity to 1, 2, 4, 8, 16 or 32 arcmin, comprising a grid of 36 pairs. Because this concatenation defines a two-dimensional space of bar disparity and bar width, three kinds of thresholds may be considered: spatial resolution at constant disparity, disparity threshold at constant bar width and the upper disparity limit for depth discrimination at constant bar width. Note that, since the target was horizontally elongated, its horizontal length did not constrain the upper disparity limit within the range of measurement.

In the main experiment, we measured the discriminability at relevant points of the grid to obtain the 75% correct threshold in width and disparity by linear interpolation. Performance was measured for all pairs of adjacent points on the grid whose detectability spanned the 75% level in either the bar-width or the disparity direction. Interpolation was then per-

formed in each direction to measure all possible 75% points throughout the space of this grid. To estimate the complete curve, the percent correct was measured at about half of the 36 points of the grid; the remainder of points were beyond the useful range for threshold evaluation. Percent correct at each measured point in the parameter plane was estimated in one, two or three blocks of trials depending on the steepness of the percent correct surface in its neighborhood. The discriminability variance in the measured points was computed on the basis of binomial distribution; the variance of the interpolated threshold was a sum of the estimated discriminability variances in the adjacent measured points.

3.3. Observers

Two observers participated: one of the authors (LK) and a high-school student (MK) who was unaware of the purpose of the experiment. Both were emmetropic in both eyes.

4. Results

4.1. Experiment 1

We first compared the effects of the bar width on disparity sensitivity for transient and sustained stimulation at the same contrast level of 20% (Fig. 6A,B). The figures may be read as defining a sensitivity contour, to the left of which the cyclopean bar was invisible, as in Fig. 5. Thus, the observers could detect the thinnest bars at medium disparities in the range of 4–8 arcmin. For the widest bars, detectability extended from about 4 to about 32 arcmin of disparity. For both observers, spatial resolution for the sustained stimuli (○) was greater than to the transient (●), matching the pattern of behavior shown in the left panel of Fig. 5. Thus, qualitatively, the results of this experiment indicate that stereopsis is mediated predominantly by the response to sustained stimulation.

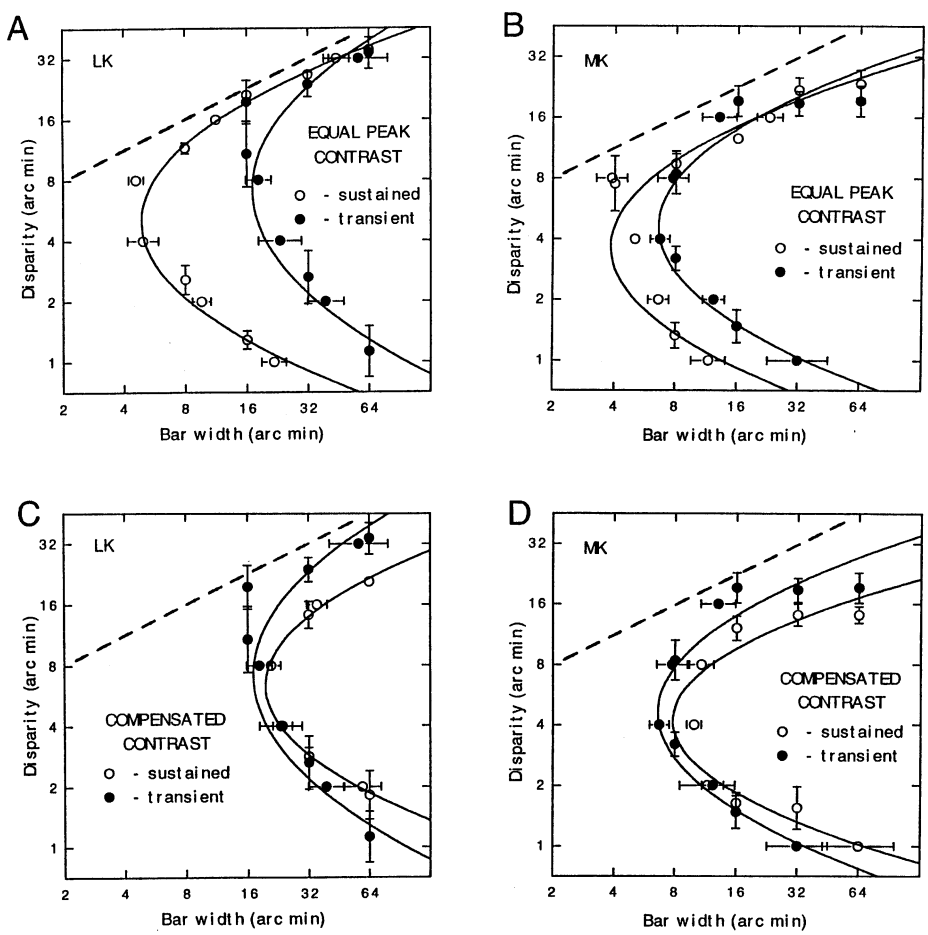

Fig. 6. Disparity discrimination thresholds as a function of bar width. Regions to the left of each curve correspond to conditions where the depth was indiscriminable. (A, B) Points in the width/disparity plane corresponding to a 75% discriminability level for two observers with sustained and transient stimulation at 20% contrast. The 75% contours are fitted by horizontal parabolas. Sustained stimuli produced better performance than transient. (C, D) Replication of the experiment with contrast of the sustained stimulus reduced to 10%. Now the discriminability of the two stimulus types is comparable, with a small advantage for transient stimuli. Dashed lines show the rate of increase in maximum disparity resolution expected according to ideal observer theory.

Table 1
Parameter values for fitted parabolas

	Spatial resolution (arcmin)	Peak disparity (arcmin)	Disparity range (octaves)
MK			
Transient (20%)	6.55	4.60	2.83
Sustained (20%)	3.81	3.65	2.77
Sustained (10%)	7.75	4.20	2.34
LK			
Transient (20%)	17.02	7.13	3.57
Sustained (20%)	4.87	4.99	2.99
Sustained (10%)	19.49	6.34	2.70

4.2. Experiment 2

In the second experiment, we reduced the contrast of the sustained modulation to a value of 10% in an attempt to match the threshold curves for the sustained and transient stimuli, which are separated horizontally by roughly a factor of two at equal physical contrast. With reduced contrast (Fig. 6C,D), the sensitivity to the sustained stimulus (○) became slightly poorer than for the transient (●).

5. Discussion

To encapsulate the shape of the 75%-confidence curves, we fit the data points by horizontal parabolas. The parabolas were specified by three parameters: the two coordinates of the parabola's tip and its width (defined at one octave to the right of the tip). These parameters provide estimates of spatial resolution, peak disparity and lower-to-upper disparity range for each kind of temporal modulation (Table 1). All parameters had a 95% confidence interval of about $\pm 15\%$ (0.2 oct).

The spatial resolution parameter varied across conditions by a factor of 2–4 (within observers), much higher than the 30–50% variation of the other two parameters. This variation means that changes in either the temporal or contrast parameters of the stimulation lead mainly to the horizontal translation of the threshold curves (change in spatial resolution) without substantial shape change. The question to be addressed is whether all three functions are mediated by a single mechanism consistent with M pathway sensitivity. If this were the case, the equivalent contrast sensitivity for the transient stimulus is predicted to be 3.5 times greater than for sustained stimulation. The data show that the spatial resolution for transient stimulation is, in fact, two to four times *less* than for sustained, in the opposite direction from the M-pathway hypothesis. How can this result be related to the effective contrast prediction?

5.1. Sustained/transient effective contrast ratio derived from the data

Note that the curve for transient stimulation lies between the 20 and 10% curves for sustained stimulation for both observers, implying that the effective contrast for transient stimulation is bracketed by the range selected for the sustained stimuli. This observation allows a quantitative estimate of the equivalent effective contrast on the basis of the spatial resolution values by interpolating between the values obtained for sustained stimulation at the two contrasts (For this narrow (10–20%) range of contrasts, it is safe to assume that linear interpolation should provide a veridical estimate). The result of the interpolation procedure was that the effective contrast of the 20% transient stimulus was the same as for a sustained stimulus of 13.0% for MK and 11.7% for LK. Thus, the transient stimulation is nearly two times *weaker* for stereoscopic detection than the sustained stimulus at the same effective contrast. This ratio looks discouraging for the M-stream hypothesis advanced by Hubel and Livingstone (1987) but is close to the ratio of 2.5 predicted for purely parvocellular input to the stereoscopic mechanism.

5.2. Is stereoprocessing limited to only the P pathway?

We conclude that human stereoscopic detection is dominated by sustained disparity processing. To the extent that our assumptions relating to the M and P streams via the monkey LGN are valid, this result implies that human sensitivity to random-dot stimulation is mediated predominantly through the P pathway.

This conclusion based on overall sensitivity does not refute the possibility that M pathway plays some role in depth perception under particular conditions. For example, if a unitary P pathway mediated detection throughout the experiment, the measured functions would be expected to maintain an invariant shape (constant width and peak disparity) under all three conditions. However, the peak disparity and disparity

range parameters from Table 1 vary consistently between the subjects: the values of both parameters are largest for the transient stimuli, although in terms of spatial resolution the transient stimulus was bracketed by 10 and 20% sustained stimuli. This deviation from invariance may be interpreted as an indication of some M input to detection at larger disparities, although only at a level where the effective contrast of the stimulus is five to ten times lower than for the sustained system. The fact that the M input is operative in the large disparity range is consistent with the hypothesis (Tyler, 1990) that the M system is responsible for coarse, off-horopteral disparity processing (as opposed to fine, global stereopsis).

It is worth noting that there is an apparent contradiction between the imputed role of the P system in stereopsis and the fact that stereopsis is typically weakened or abolished by presentation of the stereoscopic stimuli under equiluminant chromatic conditions. This paradox is addressed by DeYoe and Van Essen (1988) and Tyler (1990) by postulating that fine stereopsis is mediated by the *interblob* portion of the P stream, which is also supposed to mediate fine detail vision, while color vision is mediated predominantly in the *blob* portion of the P stream. Thus, it is no contradiction to expect that the colour-blind interblob portion of the P-stream would exhibit a degradation in performance for fine stereopsis when the stimulus contains only colour cues.

5.3. Disparity scaling or the disparity gradient limit

The two-dimensional variation of disparity and cyclopean bar width in this experiment allow evaluation of how the upper disparity limit is controlled by the width of the cyclopean bar. If there were a fixed upper disparity limit, the upper thresholds in all panels of Fig. 6 would asymptote to a constant disparity level regardless of the cyclopean bar width. On the other hand, the upper disparity limit has been established to be proportional to width of disparity excursion by Tyler (1975) for line stimuli. This limit corresponds to a similar depth limit defined by Tyler (1973, 1974, 1975) and Tyler, Schor, and Coletta (1992) for sinusoidal or square-wave disparity modulations of either line or cyclopean random-dot stimuli A similar proportionality of the binocular fusion limit was determined by Tyler (1973), Burt and Julesz (1980) and Schor and Tyler (1981). This proportionality hypothesis has been variously described as a disparity scaling with stimulus size and a disparity gradient limit. Although the proportionality principle has thus attained a range of generality, it has never been tested for cyclopean stimuli based on dynamic noise.

The principle of proportionality would apply if, for example, the disparity were processed by circular summation mechanisms whose diameter matched the width of the cyclopean bar, the noise levels in each such field being the same (i.e. limited by an invariant late noise source). The proportionality hypothesis predicts that upper disparity limits would rise with a slope of 1 (which would correspond to a 45° oblique in each panel of Fig. 6). Clearly, the current data do not match this proportionality hypothesis for any of the conditions tested. This failure suggests that stereoscopic processing in dynamic noise differs from that in static noise or for line stimuli, from which the hypothesis was derived.

On the other hand, there is a study of cyclopean depth discrimination in dynamic noise that suggests a different spatial integration principle. Tyler and Julesz (1980) found that the upper disparity limit increased as a the *square root* of the width of a dynamic-noise cyclopean bar, in agreement with the Ideal Observer prediction that the signal-to-noise ratio in such stimuli is given by a direct increase in signal and a square root increase in noise with target area (i.e. early, local sources of noise). In order to achieve such performance, however, the visual system would need to contain summation mechanisms of all possible tested widths and to focus attention on the particular spatial integration mechanism matching the extent the cyclopean bar stimulus presented on each trial (This slope analysis does not consider the issue of absolute sensitivity, and hence does not require that the summation should occur over the full length of the stimulus. The Ideal Observer slope would be predicted if the summation occurred over a constant length, even though absolute sensitivity was lower than ideal for the stimulus as a whole).

Application of the Ideal Observer hypothesis to the present stimuli with varying cyclopean bar width would generate a prediction corresponding to the dashed line in each panel of Fig. 6 (a slope of 1/2). These predictions clearly provide an adequate fit to the data in the upper disparity range, despite a few apparent deviations that nevertheless fall within the range of ± 2 SEM. We conclude that the upper disparity limit for cyclopean targets in dynamic noise is consistent with the Ideal Observer prediction of an increase with the square root of stimulus area, as proposed by Tyler and Julesz (1980). This slope has an arbitrary sensitivity, so we are not evaluating the absolute sensitivity of the system, merely whether it remains at constant efficiency as the disparity area increases. The parabolic functions fitted to the data provide an approximation to this principle over the range of the measured data (beyond which retinal inhomogeneities may degrade the effectiveness of the

spatial integration). Harris and Parker (1994a,b) have also shown a good approximation to ideal observer performance in a different paradigm for low levels of added noise.

We should point out that the upper disparity data are also consistent with other hypotheses of spatial summation, such as a single summation field with elongated tails of the appropriate shape to generate a square-root summation function. It seems slightly implausible that the integral of the function just happens to match a square-root summation behavior, but the possibility cannot be ruled out. The multiple-mechanism account of the summation slope could be tested by intermixing the various stimulus widths within a single block of trials, but that kind of experiment is beyond the scope of the look-up table animation scheme used for the present study.

6. Conclusions

We measured stereoscopic threshold as a function of the width of a cyclopean test bar in sustained and transient types of base noise. Under the assumption that the M pathway has a predominantly transient response sensitivity (Merigan and Maunsell, 1993), the results fail to support the hypothesis that stereoscopic processing occurs solely in the M pathway. To the contrary, they imply that stereoscopic stimuli (of the cyclopean type employed in this study) are processed predominantly by a sustained response mechanism, with a small contribution from a transient mechanism. Based on behavioral deficiencies following LGN lesions, this result may be interpreted as implying a predominance the P pathway for human cyclopean disparity processing.

There also is evidence of a lesser contribution from a transient (M-pathway) response at larger disparities, in accord with the analysis of Tyler (1990). We further conclude that, for dynamic noise stimuli, the upper limit of disparity processing operates according to the principles of an Ideal Observer in terms of its spatial integration properties (although its absolute sensitivity was not evaluated).

Acknowledgements

Supported by NIH grant EY 7890 to the second author. Our thanks to Richard Miller, Ph.D. for programming the cyclopean dynamic noise stimulus.

References

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science*, *208*, 615–617.

Casagrande, V. A., & Norton, T. T. (1991). Lateral geniculate nucleus: a review of its physiology and function. In J. Cronly-Dillon, *Vision and visual dysfunction*, *vol. 4*, A. G. Leventhal, *The neural basis of visual function* (pp. 41–84). Boca Raton, Florida: CRC Press.

Dagnelie, G. (1992). Temporal impulse responses from flicker sensitivities: practical considerations. *Journal of the Optical Society of America*, *A9*, 659–672.

DeAngelis, G. C.., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *Journal of Neuroscience*, *19*, 1398–1415.

DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, *394*, 677–680.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 219–240.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial vision*. New York: Oxford University Press.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience*, *11*, 219–226.

Harris, J. M., & Parker, A. J. (1994a). Constraints on human stereo dot matching. *Vision Research*, *34*, 2761–2772.

Harris, J. M., & Parker, A. J. (1994b). Objective evaluation of human and computational stereoscopic visual systems. *Vision Research*, *34*, 2773–2785.

Hubel, D. H., & Livingstone, M. (1987). Segregation of form, color and stereopsis in primate area 18. *Journal of Neuroscience*, *7*, 3378–3415.

Kaplan, E. (1991). The receptive field structure of retinal ganglion cells in cat and monkey. In J. Cronly-Dillon, *Vision and visual dysfunction*, *vol. 4*, A. G. Leventhal, *The neural basis of visual function* (pp. 10–40). Boca Raton, Florida: CRC Press.

Livingstone, M., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for perception of form, color, movement and depth. *Journal of Neuroscience*, *7*, 3416–3468.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, *16*, 369–402.

Norcia, A. M., Sutter, E. E, & Tyler, C. W. (1985). Electrophysiological evidence for the existence of coarse and fine disparity mechanisms in human. *Vision Research*, *25*, 1603–1611.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1990). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience*, *5*, 321–346.

Schor, C. M., & Tyler, C. W. (1981). Spatio-temporal properties of Panum's fusional area. *Vision Research*, *21*, 683–692.

Smallman, H. S., & MacLeod, D. I. A. (1994). Size-disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America*, *A11*, 2169–2183.

Tyler, C. W. (1973). Stereoscopic vision: cortical limitations and a disparity scaling effect. *Science*, *181*, 276–278.

Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature*, *251*, 140–142.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research*, *15*, 583–590.

Tyler, C. W. (1977). Spatial limitations in stereopsis. *Proceedings of the Society for PhotoOptical Instrumentation Engineers*, *120*, 36–42.

Tyler, C. W. (1983). Sensory processing of binocular disparity. In: C. Schor, & K. J. Ciuffreda, *Basic and clinical aspects of binocular vergence eye movements* (pp. 199–295). London: Butterworths.

Tyler, C. W. (1990). A stereoscopic view of visual processing streams. *Vision Research*, *30*, 1877–1895.

Tyler, C. W. (1991). Cyclopean vision. In: D. Regan, *Vision and visual disorders*, *vol. 9*, *Binocular vision* (pp. 38–74). New York: Macmillan.

Tyler, C. W., & Julesz, B. (1980). On the depth of the cyclopean retina. *Experimental Brain Research*, *40*, 196–202.

Tyler, C. W., Schor, C. M., & Coletta, N. J. (1992). Spatiotemporal limitations on Vernier and stereoscopic alignment acuity. Stereoscopic Displays and Applications III. *SPIE*, *1669*, 112–121.

Victor, J. D. (1989). Temporal impulse responses from flicker sensitivities: causality, linearity and amplitude data do not determine phase. *Journal of the Optical Society of America*, *A6*, 1302–1303.