STIMULUS DEPENDENCE OF DISPARITY CODING IN PRIMATE VISUAL AREA V4

Jay Hegdé1 and David C. Van Essen2
Department of Anatomy and Neurobiology
Washington University School of Medicine, St. Louis, MO 63110
REVISED –9/1/04
Submitted Online to J. Neurophysiology as a Report

Abbreviated Title: Stimulus-dependent Disparity Coding in V4
Total # characters with spaces: 30,812
1Current address: Vision Center Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines
Road, La Jolla, CA 92037.
2Correspondence should be addressed to: Dr. David Van Essen, Department of Anatomy and Neurobiology, Box
8108, Washington University School of Medicine, St. Louis, MO 63110. Tel.: (314)362-7043; Fax: (314)747-
3426; email: vanessen@brainvis.wustl.edu.

ABSTRACT
Disparity tuning in visual cortex has been demonstrated using a variety of stimulus types that contain stereoscopic depth cues. It is not known whether different stimuli yield similar disparity tuning curves. We studied whether cells in visual area V4 of the macaque show similar disparity tuning profiles when the same set of disparity values were tested using bars or dynamic random dot stereograms, which are among the most commonly used stimuli for this purpose. In a majority of V4 cells (61%) the shape of the disparity tuning profile differed significantly for the two stimulus types. The two sets of stimuli yielded statistically indistinguishable disparity tuning profiles for only a small minority (6%) of V4 cells. These results indicate that disparity tuning in V4 is stimulus-dependent. Given the fact that bar stimuli contain 2-D shape cues and the random dot stereograms do not, our results also indicate that V4 cells represent 2-D shape and binocular disparity in an interdependent fashion, revealing an unexpected complexity in the analysis of depth and 3-D shape.

INTRODUCTION
Stereoscopic depth perception makes use of small disparities in the retinal location of local features in binocularly viewed images. The neurophysiological basis for stereopsis is presumed to arise from disparity-tuned neurons, which are encountered in V1 and many extrastriate visual areas (Poggio, 1995; Gonzalez and Perez, 1998; Cumming and DeAngelis, 2000; Parker and Cumming, 2001). Disparity tuning has been studied using a variety of stimulus types, most commonly with bars (with disparity cues restricted to the perimeter of the stimulus) or random dot stimuli (with disparity cues distributed across the entire random dot pattern). A diversity of tuning profiles, including near, far, tuned excitatory, and tuned inhibitory, have been reported using both stimulus types. If a given cell responds to both bars and random dot patterns, it is natural to anticipate that the shapes of the corresponding tuning profiles would be similar. Some previous studies have tested this issue on small numbers of cells and have reported tuning profiles to be qualitatively similar for the various types of stimuli used (Poggio et al., 1985; Poggio, 1990, 1995; Gonzalez and Perez, 1998; see Discussion). However, this issue has not been systematically examined to our knowledge.

In the present study we compared the disparity tuning of neurons in area V4 of alert macaque monkeys using two types of fronto-parallel stimuli: bars defined by luminance and disks defined by random dot stereograms, with both stimulus sets covering the same set of disparity values. We found that for many V4 neurons the shapes of the disparity tuning profiles differ markedly and in unpredictable ways. Our testing paradigm was not designed to explore the neural mechanism underlying these differences, or their functional significance. Nonetheless the findings are relevant to general issues regarding the complexity of neural coding within sensory systems.

The terms ‘cue dependence’ and ‘cue independence’ (or form cue invariance) have been used (Albright, 1992; Olavarria et al., 1992; Stoner and Albright, 1992, 1993; also see Dobkins and Albright, 1993) to indicate whether or not a cell’s selectivity for a given attribute or parameter (e.g., direction of motion) is independent of the cue used to define that parameter (e.g., motion of a luminance-defined bar vs. a texture-defined bar). Here, we use similar terminology in discussing the tuning of V4 neurons to
stereo-stereoscopic stimuli. Cue independence is computationally attractive as a strategy that allows efficient encoding and decoding of multiple stimulus attributes. However, the consistency with which this strategy is adopted in sensory systems is an empirical matter. Cue-independent tuning for direction is common in MT neurons tested with moving texture vs. bars (Albright, 1992; Olavarria et al., 1992; Stoner and Albright, 1992, 1993; also see Dobkins and Albright, 1993); however, some cells are more broadly tuned for the texture-defined motion than for luminance-defined motion (Olavarria et al., 1992). Many neurons in inferotemporal cortex show similar selectivity for shapes regardless of whether the shape cue is luminance, texture, motion, or disparity (Sáry et al., 1993; Tanaka et al., 2001). The present study provides a test of cue independence in the disparity domain using stimuli in which binocular disparity is cued by edges (step changes in mean luminance) or by patterns of random dots.

**METHODS**

The responses of single V4 units to stereoscopic stimuli were recorded in awake, fixating monkeys. The details of the surgical and the experimental procedures have been described previously (Hegdé and Van Essen, 2000, 2003). Briefly, two adult male macaques (*Macaca mulatta*) were used in this study. Each animal was implanted with a scleral search coil and an acrylic cranial patch using sterile surgical procedures. After the animal was fully trained in the fixation task, a small craniotomy 5 mm in diameter was made over the recording site, and a recording chamber was mounted over the craniotomy. Neurophysiological recording was carried out using epoxy-coated tungsten electrodes (A-M Systems, Carlsborg, WA) with initial impedances of 3-5 megohms (at 1 kHz) inserted transdurally into the cortex. All animal-related procedures used in this study were reviewed and approved in advance by the Washington University Animal Studies Committee.

**Visual Stimulation and Recording**

Stimuli were presented dichoptically on a Sony GDM-17E11 17” (1280x1024 pixels) non-interlaced CRT display (refresh rate, 72 Hz) fitted with NuVision (Beaverton, OR) 17SX polarized liquid crystal shutters and viewed through passive polarized eyeglasses from a distance of 58 cm.

Units were identified as belonging to area V4 using their receptive field properties and topographical criteria (Van Essen and Zeki, 1978; Gattas et al., 1988). Single V4 units were isolated based on both the shape and the amplitude of the waveform using a dual window discriminator (Bak Electronics, Germantown, MD). We recorded from every V4 unit we were able to isolate; no additional screening criteria were imposed for the initial analysis. Following the isolation of a unit, its classical receptive field (CRF) was plotted, and the cell’s stimulus preferences were determined, using mouse-driven bar, grating, and/or dynamic random dot stimuli (dRDS) on the computer’s monitor. For all the cells in the second monkey, the CRF was also plotted using custom-written CRF mapping software, which used a small dRDS patch as the mapping stimulus. In general, CRFs as determined by the two techniques largely agreed with each other. When they did not, usually because the given cell was unresponsive to the dRDS patches used in the automated mapping, we adopted the CRFs as determined by the manual mapping, briefly remapping the CRF manually to ensure reliability.

The stimulus set included 21 bars and 21 dRDS disk stimuli, both ranging in disparity from -1.0° (crossed, or near) to +1.0° (uncrossed, or far) in 0.1° increments. An additional five control stimuli were used, consisting of a binocularly uncorrelated dRDS, and a bar and an RDS presented in either eye alone. All cells were tested using the same set of 47 stimuli, customized for the cell under study as follows. The bar stimuli had the same length, width, color and orientation as the cell’s preferred bar. For the dRDS stimuli, the color of the central disk and the color and size of the annulus were customized so as to best drive the cell. The radial size of the annulus ranged from 2° to 5°. The colors of the central disk and the annulus differed from each other for 78 of 128 cells (61%) recorded from the two animals. The preferred dRDS center color differed from the cell’s preferred bar color for only three of the 128 cells. The stimulus color/s were chosen from a palette of eight colors (with luminances measured through active liquid crystal shutters at the center of the screen using Tektronix J17 photometer): white (6.86 cd/m²), red (1.36 cd/m²), green (4.93 cd/m²), blue (0.63 cd/m²), cyan (5.72 cd/m²), magenta (1.98 cd/m²), yellow (6.42 cd/m²), and black (0 cd/m²). All stimuli were presented against a neutral gray background (1.40 cd/m²). The cross-talk between the monocular images was low for all colors (white,
2.1%; red, 2.9%; green, 1.8%; blue, 1.7%; cyan, 1.8%; magenta, 2.4%; yellow, 2.4%), and was comparable for dRDS and bars.

For all stimuli, the annulus had a fixed disparity of 0°, the same as the fixation spot, while the disparity of the center varied systematically from one stimulus to the next. The dots in the center and surround were identical in all other respects (except dot color, as noted above). The size and the density of the dots varied depending on size of the receptive field so as to provide the percept of a smooth surface (dot size range, 0.10° - 0.21°; dot density range, 40-60%). The dRDS was rendered dynamically using color-lookup table animation at the refresh rate of the monitor (72 Hz), so that in any given frame, a random one-third of dots were invisible (i.e., rendered in the background color in both eyes), and the remaining dots were rendered in the appropriate stimulus color. No coherent motion was apparent from one frame to the next.

Fixation and vergence were monitored in each eye using a dual scleral search coil setup (Remmel Labs, Ashland, MA). The stimuli were presented in a sequential, randomly interleaved fashion for 300-400 msec each with a variable 300-400 msec interstimulus interval while the animal fixated within a window of 0.5° radius for a liquid reward. Up to six stimuli per trial were presented in this fashion. To minimize the contributions of receptive field nonuniformities, if any, the spatial placement of the stimuli was systematically jittered for both bars and dRDS, so that a given presentation of each stimulus was centered on one of the four points located 25% of the CRF radius away symmetrically around the CRF center. CRF radii ranged from 0.45° to 9.65° (mean = 2.95°; n = 128). CRF eccentricities ranged from 1.4° to 23.9° (mean = 6.6°; n = 128).

Data Analysis

Spike data were analyzed using the S-Plus (Insightful Corp., Seattle, WA) and Matlab (The Mathworks, Inc., Natick, MA) utilities or custom-written C language software. Only data from the trials throughout which the animal maintained fixation were used in this study. For each cell, the evoked responses were calculated using a time window during which the overall firing rate of the given cell (across all repetitions of all stimuli) remained above background levels. Depending on the cell, this window spanned a 60–300 ms period following the stimulus onset. The background responses were calculated from a 100 msec time window immediately preceding the stimulus onset for all cells. The response to a given stimulus was calculated as the average net firing rate across 16 repetitions (fewer than 16 but at least 9 repetitions for 14 cells). A given cell was included in the analysis only if the evoked responses of the cell significantly exceeded the background responses for at least one stimulus (one-tailed t test, p < 0.05 after Bonferroni correction). Of the 128 cells recorded from the two animals, 119 cells (53 cells from the first animal, and 66 cells from the second animal) passed this test and were included in this study.

Indices

The cue sensitivity index (CSI) measured the difference between the disparity tuning curves obtained using bars vs. dRDS. To calculate CSI, we performed a two-way ANOVA for each cell, with disparity and stimulus type as the two factors. The F ratio for the interaction factor is a parametric measure of the degree to which the two disparity tuning curves differ as a function of disparity. This F ratio was used because it measures the extent to which the two tuning profiles have different shapes, and is insensitive to differences in the absolute firing rates. To calculate the CSI value for a given cell, we first calculated the F ratio for the interaction factor from the actual data. We then randomized the data by reassigning the spike counts from individual presentations randomly across different disparities within the same stimulus type, and recalculated the F ratio (for overviews of randomization, see Manly, 1991; Edgington, 1995). The CSI value was defined as the actual F ratio divided by the average F ratio from 10^6 rounds of randomization. This scaling by the average randomized F ratio effectively corrected for deviations of the data set from normality. The disparity-stimulus type interaction was considered statistically significant at the level of p < 0.05 if the randomized F ratio exceeded the actual F ratio in less than 5% of the rounds. The statistical significance of the disparity factor was also similarly determined.
The interaction factor $F$ ratio (and by extension, the CSI) cannot distinguish disparity-dependent differences in the tuning profile from the effects of multiplicative scaling, in which the two tuning curves are scalar multiples of each other (see Dean, 1981; McAdams and Maunsell, 1999). We used correlation coefficient $r$ to measure the scaling between the two tuning profiles for each cell and used randomization procedures to assess the statistical significance of the correlation. Assessing multiplicative scaling by comparing scaling factors of fitted curves (cf. McAdams and Maunsell, 1999) was not feasible for most V4 cells (Hegdé and Van Essen, unpublished results). Passing the multiplicative scaling test required that the two tuning curves were correlated at $p < 0.05$. Note that previous studies of cue invariance have not considered multiplicative scaling as criterion for assessing cue invariance (see Albright, 1992; Olavarria et al., 1992; Sáry et al., 1993).

A cell was classified as cue-dependent if the $p$ values were less than 0.05 for both the disparity factor and the interaction factor, and if it failed the multiplicative scaling test. A cell was classified as cue-independent if it had $p < 0.05$ for the disparity factor, and if it either had $p > 0.05$ for the interaction factor and/or if it passed the multiplicative scaling test. All other cells were classified as not significantly tuned for disparity.

**Regression**

We carried out regression analysis using either conventional least squares method (S-Plus function `lsfit`) which took into account all data points including outliers, or using the robust regression method of least trimmed squares method (S-Plus function `ltsreg`; Rousseeuw, 1984), which minimizes the influence of outliers.

**Vergence Eye Movements**

Vergence eye movements can potentially confound the neuronal responses to stereoscopic stimuli (Cumming and Parker, 1997; Masson et al., 1997). However, vergence eye movements (calculated from the same time window as the evoked response) were not a major confound in our dataset, as assessed by four different criteria. First, the average SD of the vergence angle was small (0.051 for the horizontal position and 0.056 for the vertical position), and did not vary between bars and dRDS or between the two animals (MANOVA, horizontal angles x vertical angles x stimulus types x animals, $p > 0.05$ for all factors). Second, vergence angles did not systematically vary with the disparity of the stimuli (correlation coefficient $r = -0.0002$ for bars; $r = 0.0018$ for dRDS; $p > 0.05$ in both cases). Third, the vergence angles were not correlated with the neuronal responses for the corresponding stimuli ($r = 0.0002$ for bars; $r = -0.0072$ for dRDS; $p > 0.05$ in both cases). Finally, when tested using a two-way ANOVA (disparity x stimulus type) for individual cells, vergence angles varied as a function of disparity and stimulus type (i.e., $p < 0.05$ for the disparity and stimulus factors) for 7 and 4 cells, respectively. The proportions of these cells were indistinguishable from those expected random (binomial proportions tests, $p > 0.05$ in both cases). Vergence data calculated using the time windows described above were indistinguishable from those calculated from the entire 300 ms stimulus duration ($t$ tests, $p > 0.05$, data not shown).
RESULTS
Disparity Tuning of Individual V4 Cells for Bar vs. dRDS Stimuli

Using stimuli of the type illustrated in Figure 1, V4 cells were generally responsive to bar stimuli, dRDS disks, or both. Figure 2 shows the disparity tuning curves for the two types of stimuli for six individual V4 cells, each of which responded both to bars and to dRDS disks significantly above background levels (one-tailed \(t\) tests, \(p < 0.05\) in all cases). For each of the six cells, the disparity tuning was statistically significant for both stimulus types (one-way ANOVAs, \(p < 0.05\) in all cases). The cell shown in panel A was broadly tuned for near disparities for both bars (solid line) and dRDS disks (dotted line). The two tuning curves were statistically indistinguishable from one another (two-way ANOVA, disparity x stimulus type; \(p > 0.05\) for the interaction factor).

However, for many other V4 cells the disparity tuning differed markedly for the bars vs. dRDS disks, as illustrated by cells in panels B-F. For the cell shown in panel B, the disparity tuning profile for dRDS was a broad tuned inhibitory pattern, whereas the disparity tuning profile for the bar stimuli was mainly tuned excitatory, with a peak near zero disparity plus smaller peaks at extreme far (+1.0\(^\circ\)) and near (-1.0\(^\circ\)) disparities. The difference between the two tuning profiles was highly significant (two-way ANOVA, \(p < 0.0001\) for interaction factor). The cells in panels C-F, like that in panel B, also showed highly significant disparity tuning for both stimulus types, but marked differences in the shape of the tuning curves. The cell in panel C showed a tuned inhibitory pattern for bars, but a tuned excitatory pattern for dRDS disks. The cell in panel D had a tuned inhibitory pattern for bars, and a complex (but relatively flat) response profile for dRDS disks. Cells in both panels E and F showed a tuned near pattern for both types of stimuli. However, the tuning profiles had different shapes and peaked at different disparities for the cell in panel E, whereas for the cell in panel F, the two tuning curves were largely scalar multiples of each other (see below).

Figure 1. Bar and random dot stereogram (RDS) stimuli. The left and the right eye views of either type of stimulus are schematically illustrated in this figure. The actual stimuli used in the experiments differed from the ones shown in many respects (including the fact that the dots in the RDS were dynamic); see Methods for details. The perimeter of the CRF as determined during the manual mapping is denoted by the dashed circle around the bar stimuli. In case of the RDS, the central disk was located within the CRF (not shown), while the annulus was presented in the non-classical surround. The disparity of the bars and central disk of the RDS varied systematically from one stimulus to the next, whereas the RDS annulus was always presented at zero disparity. Note that in case of bars, the disparity information is obtained by matching the two monocular images of the bar. The RDS stimuli contain no such monocularly visible shapes; in this case, disparity is cued by the optimal correspondence between the dots in the two monocular images.
Measuring Cue Sensitivity

To assess the sensitivity of individual V4 cells to disparity cues, we measured the extent to which the disparity tuning curves of bars vs. dRDS differed from one another using the cue sensitivity index CSI (see Methods). The CSI measured only the disparity-dependent variations between the shapes of the two tuning profiles and was, by design, insensitive to the differences in the cell’s responsiveness to bars vs. dRDS disks. A histogram of the CSI values for 119 V4 cells (Fig. 3) shows an average CSI value of 2.8 (arrow), indicating that the disparity tuning curves for the two types of stimuli differed from each other 2.8-fold more than expected from chance-level fluctuations. To assess the statistical significance of cue dependence, we tested each cell using a modified two-way ANOVA, with disparity and stimulus type as the two factors, and a multiplicative scaling test, which tested whether the two tuning curves for a given cell were scalar multiples of each other (see Methods and Fig. 3, top). For 72 of the 119 V4 cells (61%; black bars in Fig. 3), the disparity tuning was statistically significant (p < 0.05 for the disparity factor), but the disparity tuning differed significantly between the two sets of stimuli (p < 0.05 for the interaction factor), and this difference was not attributable to multiplicative scaling (multiplicative scaling test, p > 0.05), indicating that these cells conveyed disparity information in a cue-dependent fashion. These cells were classified as cue-dependent. The average CSI value for these 72 cells was 3.7 (not shown). Nine cells (8%; white and hatched bars) were classified as disparity cue insensitive. For
seven of these nine cells (6%; white bars), the disparity tuning was statistically significant, but the two tuning curves were indistinguishable between the two sets of stimuli ($p > 0.05$ for the interaction factor). For the other two cells (hatched bars), the two tuning curves were statistically distinguishable ($p < 0.05$ for the interaction factor), but were multiplicatively scaled versions of each other (multiplicative scaling test, $p < 0.05$). The remaining 38 cells (31%; gray bars) were not significantly disparity tuned.

As a further test for scaling effects, we rescaled the two tuning curves for each cell to have identical means, and repeated the above analysis. The results (not shown) were qualitatively similar to those shown in Figure 3. The average CSI value across all cells was 2.52. A total of 71 cells (60%) were classified as cue sensitive, and the mean CSI value for these cells was 3.31. Nine cells (8%) were classified as cue insensitive, including three cells which were multiplicatively scaled. The remaining 39 cells (33%) were classified as not significantly tuned for disparity. Thus, cue sensitivity of the cells was not an attributable to differences in the average responsiveness of cell to bars vs. dRDS. To determine whether and to what extent the cue dependency of disparity tuning was attributable to the disparity-dependent differences in the monocular positions of bar stimuli within the CRF (or monocular ‘artifacts’, see Cumming and DeAngelis, 2000), we compared the disparity tuning curves across the four jitter positions using a three-way ANOVA (jitter x disparity x stimulus type; not shown). RDS stimuli minimize this ‘artifact’, since the disparity information in these stimuli is dispersed throughout the
stimulus, and not confined to the edges, as in the case of the bar stimuli. The jitter factor for bar stimuli was statistically significant \((p < 0.05)\) for only a small minority of V4 cells \((6/119, 5\%)\), and 4 of the 72 cue-dependent cells \((6\%)\). For the dRDS stimuli, by comparison, the jitter factor was statistically significant for 9 cells \((8\%)\) overall, and for 6 \((8\%)\) of the 72 cue-dependent cells. Thus, the cue dependency of disparity tuning is unlikely to have arisen from a widespread sensitivity to the positioning of the stimulus within the CRF.

Another plausible scenario is that for some cells the bar stimuli may have stimulated the non-classical surround with the non-optimal color. This scenario was possible for 12 cells \((10\%)\), after taking all the relevant factors into account, including the color, length, width and orientation of the preferred bar, optimal color of the surround, CRF size, stimulus jitter, stimulus disparity, and size of the fixation window. However, the CSI values of these 12 cells were indistinguishable from those of the other cells (Wilcoxon rank-sum test, \(p > 0.05\)), indicating that this scenario was unlikely to have played a substantial role in the cue dependency of disparity tuning. The cross-talk between the two monocular views (see Methods) was unlikely to have been an important factor either, since the magnitude of the cross-talk was generally low and did not affect the CSI values (one-way ANOVA [unbalanced replicates design], \(p > 0.05\)).

To compare the overall effectiveness of bar vs. dRDS stimuli, Figure 4A plots the peak bar response versus the peak dRDS response, and Figure 4B plots the mean bar response versus the mean dRDS response. By either measure the two stimulus classes were comparably effective for the population as a whole, but there was only a modest, albeit statistically significant, correlation in the bar vs. dRDS responses (correlation coefficient \(r = 0.38\) for peak responses, 0.35 for mean responses; \(p < 0.05\) in both cases). Using robust regression, which minimizes the influence of the outliers (see Methods), the corresponding \(r\) values were 0.54 for peak responses and 0.50 for mean responses \((p > 0.05\) in both cases). As is evident from several of the exemplar cells in Figure 2, the preferred disparities for bars vs. dRDS differed for many individual V4 cells. The mean absolute difference in peak disparity value for bars vs. dRDS was \(0.74^\circ\) (range, \(0^\circ - 2^\circ\); first quartile, \(0.27^\circ\); median, \(0.57^\circ\); third quartile, \(1.16^\circ\)). A scatterplot of preferred disparities for bar and dRDS disks reveals no significant correlation (Fig. 4C).

Importantly, the three main subclasses of cells (denoted by different symbols in Fig. 4) were statistically indistinguishable from each other in terms of their mean responses, peak responses, or preferred disparities (one-way ANOVAs [unbalanced replicates design], \(p > 0.05\) in all cases; data not shown), indicating that the cue sensitivity, or lack thereof, is unlikely to have arisen from distinct subpopulations of V4 cells.
Figure 4. Comparisons of responses and preferred disparities of V4 cells for bars vs. dRDS disks. A, Peak responses for bar vs. dRDS disks. B, Mean responses above background for bars vs. dRDS disks. Outliers >80 Hz were plotted at 80 Hz in panel A, and outliers >60 Hz were plotted at 60 Hz in panel B. The solid lines denote the best-fitting regression lines, calculated with outliers at their actual values. The dashed lines denote robust regression lines. C, Preferred disparities for bars vs. dRDS disks. To reduce the uncertainty in the estimated preferred disparity arising from to trial-to-trial fluctuations in responsiveness and from multiple peaks in some of the response profiles, the preferred disparities were calculated from tuning curves smoothed using a Gaussian filter ($\sigma = 0.1^\circ$). See Methods for details.
DISCUSSION

V4 neurons are known to be selective for disparity using bar stimuli (Hinkle and Connor, 2001; Watanabe et al., 2002). Disparity selectivity using random dot stimuli has not to our knowledge been tested, nor disparity tuning for bar versus random dot stimuli compared, in V4. Some previous studies have compared the responses of neurons to solid figure vs. random dot stimuli in areas V1, V2, V3-V3A (Poggio et al. 1985; Poggio, 1990, 1995; also see Gonzalez and Perez, 1998; Prince et al., 2002), and area CIP (Taira et al., 2000). While these studies reported many notable differences in the magnitude and the response range of individual neurons to the two types of stimuli, they nonetheless found that the overall tuning profiles were qualitatively similar for the two types of stimuli (also see below). Hence, our finding of pronounced stimulus dependence in the disparity domain constitutes a novel finding for V4 in particular and for visual cortex in general. Moreover, the nature and magnitude of the cue dependence reported here differs markedly from that previously reported for motion processing in MT (Olavarria et al., 1992).

It remains to be seen whether and to what extent neurons in other visual areas code disparity in a cue-dependent fashion. As noted above, previous studies have found cue-invariant disparity tuning in areas V1, V2, V3-V3A, and CIP. The precise reasons for the discrepancy between these results and ours are difficult to ascertain, but may be related to the stimuli and the analytical methods used, or genuine differences among areas. It is also possible that disparity selectivity in other areas besides V4 will turn out to be cue-dependent when tested with other stimuli (e.g., lines or gratings). By the same token, it is possible that cue dependence of disparity coding is more common in V4 than our results suggest, since at least some of the cells classified as cue invariant or as not selective for disparity in the present study may indeed have been cue-dependent for other stimuli. Even if disparity selectivity in V4 turns out to be cue invariant for all other stimuli, our results represent the minimum estimate of cue dependence, and the maximum estimate of cue invariance, of disparity coding in V4. Moreover, the fact remains that the two most common stimulus types used for determining disparity selectivity, i.e., bars and dRDS, often yield widely different disparity tuning curves in area V4.

In interpreting these results, it is important to recognize that the nature and spatial distribution of disparity cues differed markedly for the bars and dRDS used in this study (also see Poggio, 1990, 1995). For the bar stimuli, disparity cues were restricted to the perimeter of a rectangular region lying within the CRF (but jittered in exact location from trial to trial). In contrast, the dRDS disparity cues included one disparity value for the random dots throughout the disk-shaped center, a different (albeit constant) disparity value for the annular surround, and a disparity discontinuity along their common boundary. Also, dRDS stimuli by their nature contain numerous false matches involving various dot combinations. Finally, many cells in V2 respond to disparity discontinuities, not just to uniform disparities within the CRF (von der Heydt et al., 2000). Hence, the signals reaching V4 from disparity-sensitive cells in V1 and V2 are likely to be very different for bar vs. dRDS disk stimuli.

These considerations suggest several mechanistic factors that might contribute to differences in disparity tuning for bars vs. dRDS. If, for example, the disparity tuning within the CRF of a V4 cell is spatially non-homogeneous, bar stimuli might predominantly activate a subregion of one type (e.g., a tuned excitatory strip), whereas the dRDS disk might activate a combination of subregions having different disparity preferences. Differential activation of subregions might persist even with the trial-to-trial jittering of stimulus positions (see Methods) if the subregions encompass a substantial fraction of the CRF. If V4 cells are sensitive to depth discontinuities, as occurs in some V2 neurons, the tuning for mean disparity within the CRF might be modulated by the depth discontinuities of the disk/annulus border. Another possible factor is an influence of tuning for relative disparity, rather than strictly absolute disparity, as has been reported for some V2 cells (Thomas et al., 2002). Relative disparity tuning might be expected to sharpen the tuning profiles for dRDS (which include a reference disparity in the annular region) compared to bar stimuli (where the fixation point is the only reference disparity), yet this is not the primary difference between the tuning profiles encountered in our study (see Fig. 2).

These issues can be further investigated by tests that were not included in our original paradigm, as it was designed to explore a different set of issues relating to 3-D shape processing (Hegdé and Van Essen, in preparation). For example, it would be of interest to quantitatively compare tuning profiles using luminance versus random dot stimuli that represent similar geometric shapes (e.g., bars for both or disks for both; see, e.g., Poggio, 1990, 1995; Taira et al., 2000; Prince et al., 2002).
Irrespective of the mechanistic issues, it is of interest to know what role, if any, cue-dependent disparity tuning plays in 3-D form perception. Conceivably, the stimulus dependence reported here might be epiphenomenal, if the visual system could somehow make use of selectivities for shape, color, and other attributes while ignoring the disparity-related tuning of cells in V4. However, it seems implausible that this explanation would apply generically to all cue-dependent cells in V4. Alternatively, disparity tuned cells in V4 may contribute to stereoscopic depth perception by a process that entails a more complex decoding strategy for inferring depth and shape. This might involve the establishment of cue-independent tuning in individual neurons at higher stages (e.g., in inferotemporal cortex) or by cue independence that is manifest at a population level rather than in individual tuning profiles.


ACKNOWLEDGMENTS

We thank Dr. Greg DeAngelis for advice and help throughout the project, Drs. Ed Connor and David Hinkle and Ms. Donna Hanlon for help with UNIX hardware and software, and Aki Anzai for comments on the manuscript, and Susan Danker for assistance in manuscript preparation. Supported by NIH grant EY02091.