

# Hierarchical Processing of Horizontal Disparity Information in the Visual Forebrain of Behaving Owls

Andreas Nieder and Hermann Wagner

*Lehrstuhl für Zoologie/Tierphysiologie, Institut für Biologie II, Rheinisch-Westfälische Technische Hochschule Aachen, D-52074 Aachen, Germany*

According to their restricted receptive fields and input-filter characteristics, disparity-sensitive neurons at early processing levels of the visual system perform rather ambiguous computations; they respond vigorously to disparity in false-matched images and show multiple response peaks in their disparity-tuning profiles. On the other hand, the perception of depth from binocular disparity is reliable, thus raising the question as to where and how in the brain additional processing is accomplished leading toward behaviorally relevant disparity detection. To address this issue, tuning data during stimulation with correlated and anticorrelated random-dot stereograms (a-RDS) were obtained from 52 disparity-sensitive visual Wulst neurons in three behaving owls. From the disparity-tuning curves, several quantitative measures were derived that allowed to determine the response ambiguity of a cell. A systematic decline of

response ambiguities with increasing response latencies was observed. An increase in response latencies of neurons was correlated with a decrease of the strength of responses to a-RDS. Declining responses to a-RDS are expected for global detectors, because an owl was not able to discriminate depth in psychophysical tests with a-RDS. In addition, suppression of response side peaks was increased and disparity tuning was enhanced with growing response latencies. These results suggest a functional hierarchy of disparity processing in the owl's forebrain, leading from spatial filters to more global disparity detectors that may be able to solve the correspondence problem. Nonlinear threshold operations and inhibition are proposed as candidate mechanisms to resolve coding ambiguities.

*Key words: binocular disparity; stereovision; coding ambiguity; hierarchical processing; visual forebrain; radiotelemetry; owl*

Horizontal binocular disparity is one of the dominant cues to derive a three-dimensional representation from two-dimensional images projected onto the retinas of both eyes. A major problem the visual system faces when extracting depth is the so-called "correspondence problem:" which point in the left eye corresponds to which point in the right eye? By using random-dot stereograms (RDS) (Fig. 1*a*), Julesz (1960) demonstrated that our visual system is able to solve the correspondence problem before monocular form recognition.

Neurons responding to horizontal disparity have been known for over three decades (Barlow et al., 1967). Poggio and coworkers (Poggio et al., 1985; Poggio, 1995) were the first to show that neurons in the visual cortex of behaving monkeys also signal disparity in global RDS. Such neurons were implicitly thought to possess the capacity to eliminate false matches and solve the correspondence problem (Poggio and Poggio, 1984). Recent physiological studies, however, provided data consistent with the view that disparity-sensitive neurons at early visual levels perform more local filtering rather than global image matching (Cumming and Parker, 2000). Cumming and Parker (1997)

clearly demonstrated that many neurons in V1 of the fixating monkey cannot discard false matches. Using anticorrelated RDS (Fig. 1*b*) that cannot be matched in the two eyes and, thus, do not support depth perception, these authors demonstrated that most neurons in V1 signaled disparity in false-matched images and inverted their tuning profile, as expected for local disparity detectors (Qian, 1994; Ohzawa, 1998). The resulting discrepancy was that V1 neurons signal disparity in a stimulus that contains no visible depth information. Thus, it was concluded that V1 neurons cannot be a direct correlate for depth perception (Cumming and Parker, 1997).

Another major response ambiguity of local disparity detectors refers to their spatial-filter characteristics. Local disparity detectors found in V1 of mammals and the visual Wulst of owls are well explained by a combination of monocular receptive fields that can be modeled as Gabor functions (Marceljā, 1980; Field and Tolhurst, 1986; Jones and Palmer, 1987; Ohzawa et al., 1990; Nieder and Wagner, 2000) (Fig. 1*c*). Because of their spatial-frequency filter characteristics, local detectors respond quasiperiodically as a function of disparity. Tuning curves typically exhibit several response peaks, even after integration across spatial frequencies (Wagner and Frost, 1993, 1994; Fleet et al., 1996; Ohzawa et al., 1997) (Fig. 1*c*) and, thus, may signal images at quite different depth planes.

It remains an open question as to where in the brain a postulated global processing stage might be realized by neurons that signal disparity unambiguously. Like other complex visual tasks (Van Essen and DeYoe, 1995), stereopsis is thought to arise from a hierarchy of increasingly sophisticated representations ranging from spatial filtering to perceptually relevant, global disparity detection (Marr and Poggio, 1979; Tyler, 1994; Neri et al., 1999).

Received Jan. 10, 2001; revised April 4, 2001; accepted April 5, 2001.

This work was supported by Deutsche Forschungsgemeinschaft Grant WA606/6 (to H.W.). We are indebted to Drs. Doug P. Hanes, Kirk G. Thompson, and Jeffrey D. Schall (Vanderbilt University, Nashville, TN) for generously providing the algorithm and source code for Poisson spike train analysis. We are especially grateful to Jörg Lippert for his valuable discussion of data. Jörg Lippert and Kathleen C. Anderson provided helpful comments on earlier drafts of this manuscript.

Correspondence should be addressed to A. Nieder at his present address: Center for Learning and Memory, Department of Brain and Cognitive Sciences, E25-236, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. E-mail: nieder@mit.edu.

Copyright © 2001 Society for Neuroscience 0270-6474/01/214514-09\$15.00/0

In the current study, single-unit data are presented that suggest a functional hierarchy toward global disparity detection in the visual Wulst of behaving barn owls. Mechanisms that may account for the resolution of coding ambiguities are evaluated.

## MATERIALS AND METHODS

**Psychophysics.** The method for behavioral investigation has been described previously (Nieder and Wagner, 1999). Briefly, a barn owl was trained using a two-alternative choice paradigm to discriminate depth stimuli displayed on a computer monitor by pecking on one of two keys. Presentation of stereoscopic stimuli is described below.

In the baseline task, square-sized ( $7 \times 7^\circ$ ), binocularly correlated RDS (c-RDS) with eighteen different disparity values (nine positive and nine negative disparities) were presented on a random-dot background ( $0^\circ$  disparity) that covered the entire monitor. Nine different disparity values for the near (crossed) and far (uncrossed) stimulus configuration were chosen to ensure that the owl generalized depth information into the categories “near” versus “far” rather than discriminating defined disparity values. For each stimulus presentation, the dot pattern of the static RDS (with identical stimulus features as used for physiology) was newly randomized to avoid local discrimination cues. Baseline stimuli were displayed in pseudorandom order with a probability of  $p = 0.5$  for both the near (negative disparity) and far (positive disparity) configurations. Errors were followed by correction trials. For the well trained bird, the rate of reward was reduced to 85% to habituate the bird to the occasional absence of a reward after correct response to probe stimuli in transfer tests.

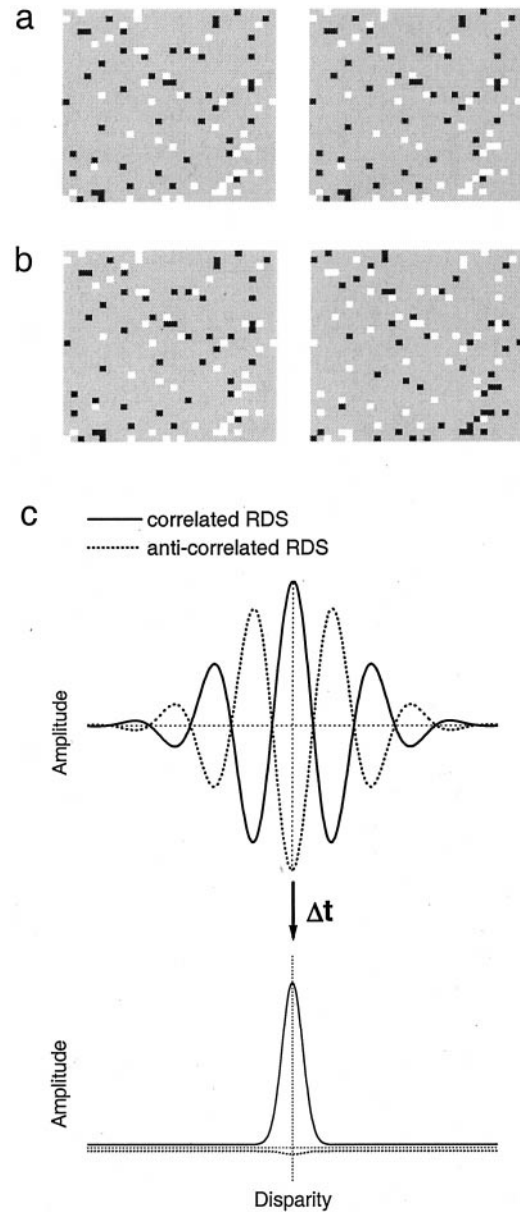
In transfer tests, anticorrelated RDS (a-RDS) with a disparity of  $-0.3$  and  $+0.3^\circ$  were presented with a probability of  $p = 0.1$ . a-RDS were identical to c-RDS except for the opposite contrast of the dot pattern in one eye relative to the other. Independent of the owl’s response, a-RDS were randomly rewarded at  $p = 0.5$ . No correction trials were applied for a-RDS stimuli. Performance was evaluated using a binomial test based on 50 observations for each stimulus.

**Electrophysiological recordings.** The method for single-unit recordings in behaving barn owls has been described previously (Nieder and Wagner, 1999, 2000). Owls were perched in front of a computer monitor and were trained to perform a visual fixation task. Gaze orientation was detected automatically by means of an infrared photoelectric device in combination with a reflective foil attached to the birds head. A trial was interrupted whenever the birds made head movements larger than  $\pm 1.5^\circ$ . During the training, owls learned to avoid head movements while fixating, which was monitored by observing the gaze and eyes under infrared illumination. Eye movements were not measured because they are virtually absent in owls (Steinbach and Money, 1973; Pettigrew and Konishi, 1976). In addition, tuning curves were analyzed to confirm that data were not contaminated by vergence (Nieder and Wagner, 2000).

Microdrives supplied with one or two tungsten microelectrodes (10 M $\Omega$ ; Frederick Haer Co., Bowdoinham, ME) were chronically implanted under general anesthesia to record from the hyperstriatum accessorium of the visual Wulst (Pettigrew, 1979). A custom-built miniature frequency modulation stereo radio transmitter (Nieder, 2000) attached to the skull transmitted neuronal activity. After filtering and amplification, the waveforms of the signals were digitized at a sampling rate of 32 kHz and stored to disk using a personal computer-based recording system (Discovery; DataWave Technologies, Minneapolis, MN). Preliminary cluster cutting was performed on-line, and definitive single-unit isolation was repeated off-line. Care and treatment of the owls were in accordance with the guidelines for animal experimentation as approved by the Regierungspräsidentium Köln (Germany).

**Visual stimulation.** Visual stimulation was performed by means of a Silicon Graphics (Mountain View, CA) workstation. After receptive fields had been determined, graphics were switched to stereo mode with a spatial resolution of  $1280 \times 496$  pixels and a refresh rate of 120 Hz (60 frames per second for each eye). Stereoscopic presentation was accomplished using a liquid crystal polarizer (model SGS17S; NuVision, Beaverton, OR) placed in front of the display. The polarizer allowed alternate transmission of images to the left and right eye with opposite light polarization in synchrony with the refresh rate of the monitor. Interocular cross talk was 11%. Owls wore glasses filtering polarized light to allow the passage of the image of the right eye to the right eye but blocking it for the left eye and vice versa.

Static RDS covering the entire screen of the monitor (except the fixation target) were flashed for 500 msec on a gray background. All



**Figure 1.** Random-dot stimuli and schematic responses of disparity detectors. *a*, c-RDS, which allows depth perception. *b*, In a-RDS, a black dot is shown to the left eye, and the right eye sees a white dot at the corresponding position, and vice versa. This stimulus is rivalrous and does not allow depth perception. *c*, Schematic disparity response profiles simulated with a Gabor function. Tuning curves of local disparity detectors (*top panel*) display two major response ambiguities: multiple response peaks according to the spatial filter characteristics of input neurons and responses to “false” disparities by inverting the response profile during stimulation with a-RDS. Global disparity detectors (*bottom panel*), however, should only show a response peak at the preferred disparity and ignore disparity in false-matched images. Because global detectors should emerge from local detectors, additional neuronal processing ( $\Delta t$ ) is expected, which increases response latency of global detectors.

receptive fields were entirely filled by the RDS. The RDS consisted of 5% white, 5% black, and 90% gray rectangular dots (size of  $0.15^\circ$ ). By shifting one of the two RDS images horizontally, positive or negative disparities could be induced. The fixation target was always set to zero disparity as a reference. After each stimulus presentation, a new dot pattern was shown. The sequence of disparities was pseudorandomized and repeated 5–10 times. Presentation of c-RDS was alternated with presentation of a-RDS for each disparity (Fig. 1*a,b*).

**Data analysis.** Response latency of disparity-sensitive neurons was determined using a Poisson spike train analysis (Hanes et al., 1995; Thompson et al., 1996). This algorithm has been used previously to determine the occurrence of bursts in spike trains (Legéndy and Salzman, 1985), as well as visual response latencies across the macaque visual system (Schmolesky et al., 1998). The Poisson spike train analysis defines times of neuronal modulation in single spike trains and not deviations from mean rates. It compares the number of spikes that occurred within a given time interval with the number of spikes that would be predicted to occur in an interval of the same length if spike timing would follow a Poisson distribution. The algorithm detects intervals with significant changes in neuronal activity. Intervals of 200 msec after stimulus onset were taken into account ( $p < 0.05$ ). Latencies of single spike trains were determined for the three disparities that elicited the largest (mean) response (i.e., for the three preferred disparities). The median of the derived single-trial latencies was used as the response latency of the cell. Spontaneous activity was derived in 100 msec intervals before physical stimulus onset (black screen) and averaged across all trials.

Whether neurons were disparity-selective was determined by calculating a nonparametric ANOVA (Kruskal–Wallis  $H$  test;  $p < 0.05$ ). To derive quantitative measures of the tuning curve, a Gabor function  $f(d)$  (Gabor, 1946) was fitted to the mean firing rates as a function of disparity:

$$f(d) = A * e^{-0.5 \left( \frac{d-xc}{\sigma} \right)^2} * \cos(2\pi(\omega(d-xc) + \phi)) + B, \quad (1)$$

where  $A$  and  $B$  are the amplitude of the envelope and the firing rate offset (baseline),  $xc$  and  $\sigma$  are the position offset and the SD of the Gaussian, and  $\omega$  and  $\phi$  are the frequency and phase of the cosine. To characterize quantitatively the occurrence of side peaks during stimulation with c-RDS, a side peak-suppression index (SSI) was calculated:

$$SSI = \omega^{-1}/\sigma, \quad (2)$$

where SSI would be 0 for a pure sine wave; values of 10 and larger indicate single-peaked curves with essentially no periodic modulation.

The disparity tuning index (DTI) of a tuning curve was determined by the following:

$$DTI = \frac{(R_{max} - R_{min})}{(R_{min} + R_{max})}, \quad (3)$$

where  $R_{max}$  and  $R_{min}$  are the maximum and minimum mean spike rate (Cumming and Parker, 1999).

Correlated and anticorrelated response profiles were fitted simultaneously ( $\chi^2$  minimization after Levenberg–Marquardt). Mean spike rate data points were weighted with SEs when computing  $\chi^2$ . The fitting procedure shared all parameters except the amplitude  $A$  and the phase  $\phi$ . For the few cases in which parameters could not be constrained by the fitting algorithm, the estimates of the SEs from the variance–covariance matrix after each iteration were balanced. Tuning curves for c-RDS and a-RDS were compared by deriving the envelope amplitude ratio (EAR) ( $A_a/A_c$ ) and the phase difference ( $\phi_c - \phi_a$ ) for each neuron (Cumming and Parker, 1997; Ohzawa et al., 1997). The output of the neurons was compared with a local filtering model of disparity detection (Ohzawa et al., 1990; Qian, 1994) that predicts a total inversion of the response profile during stimulation with anticorrelated images (EAR of 1;  $\Delta\phi$  of  $180^\circ$ ) (Ohzawa et al., 1990, 1997; Qian, 1994; Fleet et al., 1996; Cumming and Parker, 1997) (Fig. 1c).

The amount of inhibition occurring during stimulation with c-RDS was calculated:

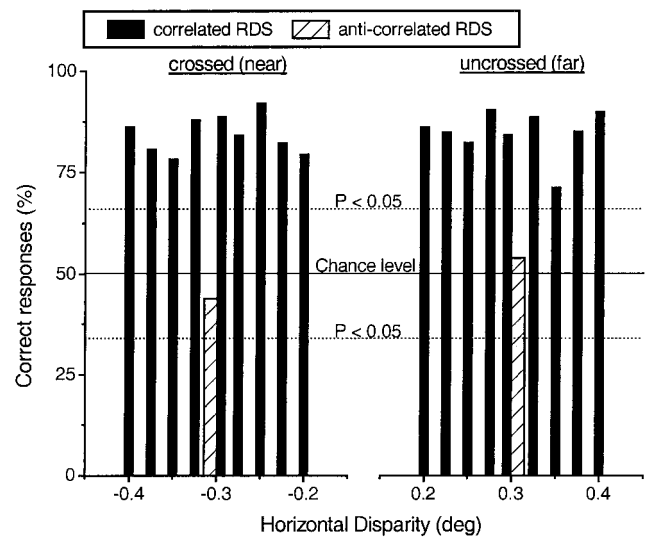
$$\text{Inhibition} = S - R_{min}, \quad (4)$$

where  $S$  is the spontaneous activity (spikes per second) derived in 100 msec intervals before RDS stimulation (i.e., black screen). For all correlation analyses, Spearman's rank correlation coefficient  $r_s$  was computed (all  $p$  values were two-tailed) to account for nonparametric distributions and nonlinear relationships.

## RESULTS

### Psychophysical data

Anticorrelated random-dot stereograms do not support stereoscopic vision in humans (Cogan et al., 1993) and monkey (Cum-



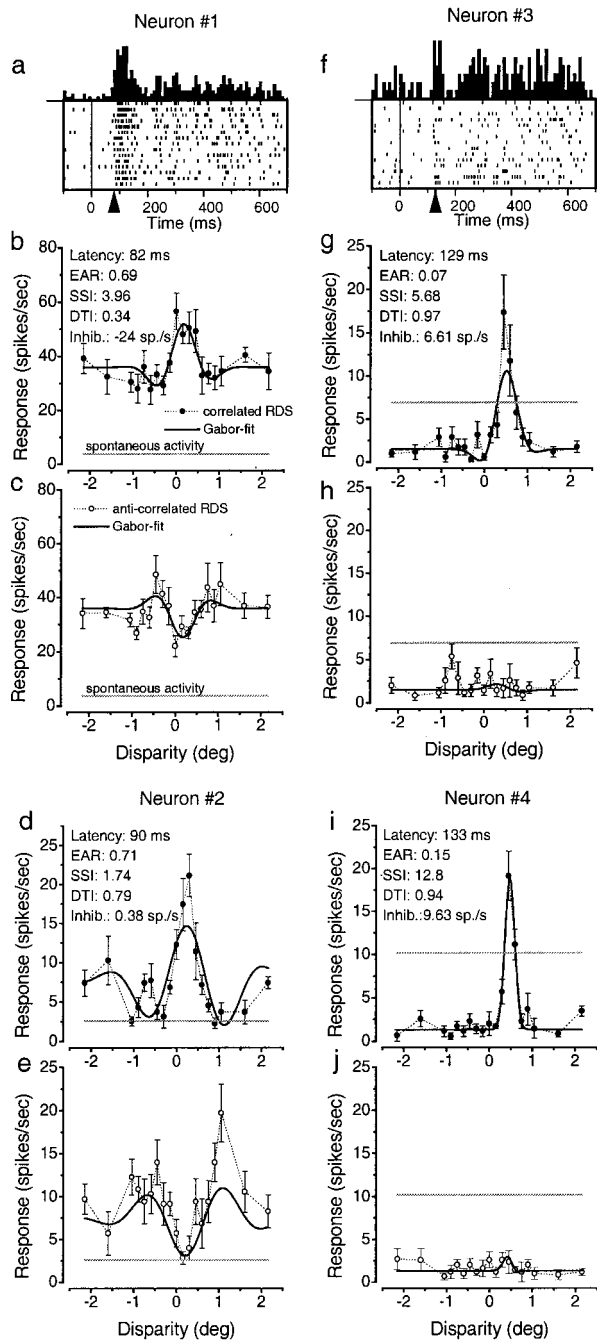
**Figure 2.** Behavioral discrimination of random-dot stereograms by an owl. The owl significantly discriminated between near (crossed) and far (uncrossed) configurations of correlated RDS in the baseline task (black columns). During transfer tests, the bird's performance to anticorrelated stereograms with positive and negative disparities was at chance (hatched columns), indicating that it was not able to discriminate depth in false-matched images. Each column represents performance for 50 stimulus presentations. Chance level (50%) and confidence interval above chance ( $p < 0.05$ ; binomial test; two-tailed;  $n = 50$ ) are marked.

ming and Parker, 1997), which has important consequences when investigating the neural basis of conscious depth perception. To find out whether this phenomenon is also found in owls, birds with an independently evolved binocular system (Pettigrew, 1986), one barn owl was tested with a two-alternative choice discrimination paradigm. The owl had to signal depth configurations in random-dot stereograms by pecking one of two keys (the ability of owls to extract depth in c-RDS has been demonstrated by van der Willigen et al., 1998). Both the stimulus and the background consisted of dot patterns with identical visual features as used for physiology (5% white, 5% black, and 90% gray dots), which permits comparison of forebrain recordings with the owl's depth perception.

Once the owl performed the baseline discrimination with c-RDS reliably, transfer tests with a-RDS began. Anticorrelated stereograms of negative ( $-0.3^\circ$ ) or positive ( $+0.3^\circ$ ) disparity, respectively, were occasionally inserted among ongoing baseline trials displaying correlated RDS. Although the owl significantly discriminated correlated stereograms for all tested disparity values, responses to anticorrelated stereograms were not different from chance performance (Fig. 2). In other words, the bird was not able to transfer the depth percept in correlated RDS to anticorrelated RDS. We conclude, first, that the barn owl was able to discriminate depth in random-dot stereograms with an overall dot density of 10% and, second, that the bird was not able to see stereoscopic depth in anticorrelated random-dot stereograms of the same dot density.

### Neural data

Quantitative disparity-tuning data during stimulation with static RDS were obtained from 52 disparity-sensitive single units in three awake owls that were trained to perform a visual fixation task (Nieder and Wagner, 2000). For 41 units, response latency was determined using a Poisson spike train analysis (Hanes et al.,



**Figure 3.** Relationship between responses to c-RDS and a-RDS. Responses of four neurons are shown. For each neuron, tuning curves to c-RDS (solid symbols, *b, d, g, i*) and a-RDS (open symbols, *c, e, h, j*) are shown. Gabor functions (solid lines) were fitted simultaneously to both tuning curves of each cell. The dot-raster display and corresponding peristimulus time histogram (bin width, 10 msec) in *a* and *f* illustrate the temporal response pattern of *Neuron #1* and *Neuron #3*, respectively, during stimulus presentation (physical stimulus onset at  $t = 0$  msec). Arrowheads at the bottom of the dot-raster displays indicate the response latency of the cells as derived by the Poisson spike train analysis. Response latencies for *Neuron #3* and *Neuron #4*, which responded only weakly to disparity in a-RDS, were longer compared with response latencies of *Neuron #1* and *Neuron #2*, which showed a pronounced, inverted tuning curve to contrast-inverted images. Values of derived parameters are given in the top left corner of c-RDS tuning curves. *sp./s*, Spikes per second.

1995; Thompson et al., 1996). In the remaining 11 neurons, discharge was either too uniformly distributed to detect the occurrence of spike bursts for determining response latency or mainly inhibitory.

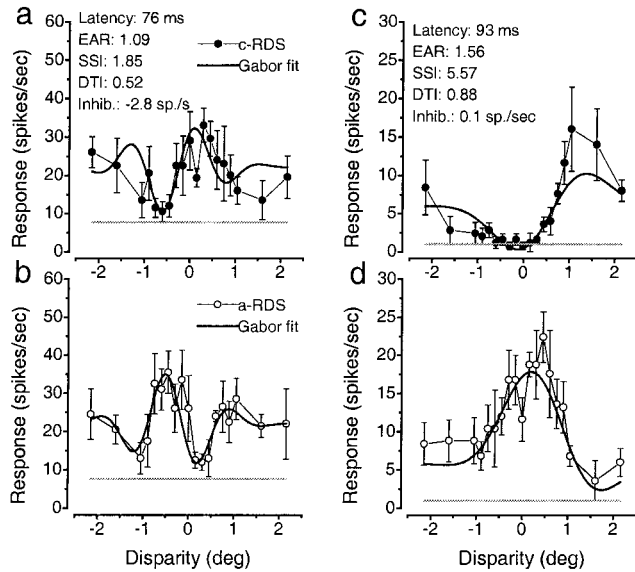
From the disparity-tuning data, the envelope amplitude ratio, side peak suppression index, disparity tuning index, and baseline of Gabor fit (see Materials and Methods) were derived as indicators for ambiguous or unequivocal disparity detection. These parameters and the response latency of all tested cells were not different in the three owls (all  $p > 0.05$ ; Kruskal–Wallis one-way ANOVA; two-tailed) and, thus, pooled for additional analysis.

### Cellular responses to correlated and anticorrelated random-dot stereograms

Single-unit responses to both c-RDS (Fig. 1*a*) and a-RDS (Fig. 1*b*) were recorded. To quantify the effect of contrast inversion, tuning curves of single neurons to correlated and anticorrelated RDS were fitted simultaneously with a Gabor function (Gabor, 1946), and the ratio of the envelope amplitude of the fits of individual neurons was derived. The sample contained two double-peaked tuning profiles (Nieder and Wagner, 2000) that were fitted like any other profile with a Gabor function for the sake of objective quantification. For local disparity detectors, the envelope amplitude in the two stimulation conditions should be equal and, thus, the EAR should be near 1. The modulation phase of the tuning curves in both stimulus conditions should exhibit a difference of half of a cycle (Cumming and Parker, 1997; Ohzawa et al., 1997). Figure 3 displays tuning profiles of four neurons to c-RDS and a-RDS. As predicted by the local filtering model, *Neuron #1* (Fig. 3*a–c*) and *Neuron #2* (Fig. 3*d,e*) responded with an almost complete inversion of their disparity-tuning profile during stimulation with contrast-inverted RDS. However, *Neuron #3* (Fig. 3*f–h*) and *Neuron #4* (Fig. 3*i,j*), although sharply tuned to c-RDS, were not significantly activated by any disparity in a-RDS; accordingly, they showed a more or less flat tuning curve to contrast-inverted stereograms with an EAR near 0. Figure 4 displays the responses of two example neurons with near odd-symmetric disparity tuning profiles.

The distribution of phase differences and EARs for all 52 tested cells is shown in Figure 5*a*. In accordance with the prediction of the local filtering model, the mean phase difference between the correlated and anticorrelated stimulus condition was close to 0.5 cycles (mean  $\pm$  SD,  $0.47 \pm 0.18$ ;  $n = 52$ ). Thirty-eight percent of the neurons (20 of 52) responded during stimulation with a-RDS with an inversion of the disparity-response profile of at least half of the envelope amplitude compared with c-RDS (EAR of  $\geq 0.5$ ) (Fig. 3, *Neuron #1*, *Neuron #2*). The remaining units, however, responded only weakly to a-RDS, and 33% of the sample (17 of 52) exhibited EARs of  $\leq 0.2$  (Fig. 3, *Neuron #3*, *Neuron #4*). For all neurons tested, a continuum of EARs ranging from  $\sim 1$  to almost 0 was observed (Fig. 5*a*), with a mean  $\pm$  SD EAR of  $0.46 \pm 0.36$ . Similar EAR values have also been reported for cells in the primary visual cortex of monkeys and cats (see Discussion).

We tested the hypothesis that the different EARs might represent a processing hierarchy. Therefore, the correlation between the EAR and response latency of neurons to RDS was measured. Indeed, cells with longer response latencies showed smaller EARs (Fig. 5*b*) (Spearman's rank correlation coefficient;  $r_s = -0.49$ ;  $p = 0.001$ ;  $n = 41$ ). This reduction of responses to false-matched images did not depend on the type of tuning profile (Fig. 5*c*). No correlation was found between EAR and the phase derived from



**Figure 4.** Two example neurons showing more or less odd-symmetric disparity-tuning curves. The phase of the Gabor function fitted to the first neuron (response profile to c-RDS and a-RDS in *a* and *b*, respectively) was 0.33 cycles. The second neuron in *c* and *d* had a phase of 0.28 cycles. Again, the values for the most important measures are given in the top left corner. The gray horizontal line indicates spontaneous activity.

the Gabor fit ( $r_s = 0.15$ ;  $p = 0.29$ ). It should be mentioned, however, that our sample consisted predominantly of tuning profiles with phases at  $\sim 0^\circ$  that are generally more abundant in the owl's Wulst (Nieder and Wagner, 2000). Both EAR (Fig. 5*d*) and response latency (Fig. 5*e*) were not correlated with the maximum discharge rate of the neurons (EAR-maximum discharge:  $r_s = 0.21$ ,  $p = 0.13$ ; latency-maximum discharge:  $r_s = -0.11$ ,  $p = 0.48$ ).

### Suppression of tuning-curve side peaks with response latency

Apart from responses to a-RDS, the occurrence of side peaks in the tuning curves derived with c-RDS represents another major ambiguity in local disparity detection. The amount of side peak suppression was used as a second indicator to determine whether a neuron had an improved coding capacity. Tuning profiles in our sample varied from periodic curves exhibiting several prominent side peaks (Fig. 6*a*) to single-peaked curves (Fig. 6*b*). Neurons showing a high SSI (see Materials and Methods) had, on average, longer response latencies ( $r_s = 0.50$ ;  $p = 0.001$ ;  $n = 41$ ) (Fig. 6*c*). Furthermore, a significant correlation was found between SSI and EAR ( $r_s = -0.37$ ;  $p = 0.006$ ;  $n = 52$ ); in other words, side peaks and responses to false-matched images became suppressed in parallel (Fig. 6*d*). This general rule is reflected in the responses shown in Figure 3; *Neuron #1* and *Neuron #2* showed both a strongly modulated tuning curve and a profile inversion to a-RDS, whereas *Neuron #3* and *Neuron #4* exhibited a single-peaked tuning curve without responses to a-RDS. As shown in Figure 6*e*, SSI did not depend on the maximum discharge rate of the cells ( $r_s = -0.04$ ;  $p = 0.77$ ).

In addition, the DTI (a measure of the signal-to-noise ratio of the coding capacity of a cell) significantly increased with response latency ( $r_s = 0.39$ ;  $p = 0.011$ ;  $n = 41$ ) (Fig. 7*a*), and neurons with a high DTI had, on average, lower EAR values (Fig. 7*b*) ( $r_s = -0.61$ ;  $p < 0.001$ ;  $n = 52$ ). The DTI, however, was significantly

correlated with the maximum spike rate ( $r_s = -0.40$ ;  $p = 0.001$ ) (Fig. 7*c*).

### Decline of baseline activity with response latency

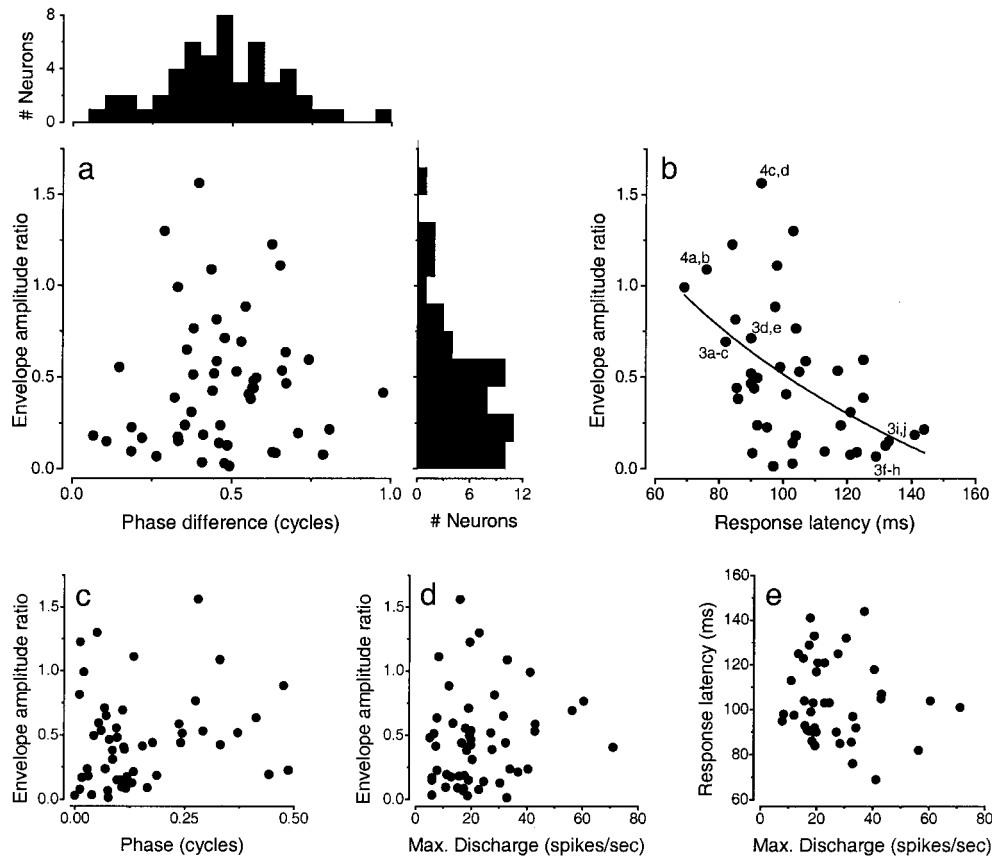
A recent computational study (Lippert et al., 2000) suggested that nonlinear threshold operations might account for the generation of unambiguous disparity detectors. Interestingly, the tuning profiles of many neurons with a single response peak (Fig. 6*b*) and neurons that did not respond to false-matched patterns (Fig. 3, *Neuron #3*, *Neuron #4*) exhibited baseline discharges to nonpreferred disparities close to zero. To test whether this effect, which could provide evidence for a nonlinear threshold mechanism (see Discussion), was present throughout the entire population of tested cells, baseline activity *B* was derived from the Gabor fits (i.e., the discharge offset of the fit; see Materials and Methods) and correlated with response latency (Fig. 8*a*). Indeed, neurons with longer response latency showed, on average, lower baseline activity ( $r_s = -0.39$ ;  $p = 0.011$ ;  $n = 41$ ). Significant correlations were also observed between baseline and EAR ( $r_s = 0.59$ ;  $p < 0.001$ ), baseline and SSI ( $r_s = 0.38$ ;  $p = 0.006$ ), and baseline and DTI ( $r_s = -0.80$ ;  $p < 0.001$ ) of all 52 tested cells (Fig. 8*b–d*). There was a weaker but still significant correlation between baseline and EAR for a subsample of cells that had fitted phases of  $>0.1$  cycles ( $r_s = 0.39$ ;  $p = 0.04$ ;  $n = 28$ ) (Fig. 8*b*, indicated by different symbols). Figure 8*e* shows that spontaneous activity and baseline activity were not correlated ( $r_s = 0.19$ ;  $p = 0.25$ ). Together, unambiguously responding disparity detectors showed lower offset activity in disparity tuning curves.

### Inhibitory influences

Inhibition plays a major role in generating response selectivity in a variety of sensory systems. In the visual system, orientation selectivity of cells in the primary visual cortex is greatly enhanced by cortical inhibition (for review, see Ferster and Miller, 2000). We tested whether inhibition could be an additional mechanism responsible for a reduction of response ambiguities in disparity-sensitive neurons with longer latencies. The amount of inhibition was determined by subtracting the minimum response rate at any given disparity in c-RDS from spontaneous activity. Thus, positive values indicate inhibition. Inhibition tended to increase with response latency of the neurons ( $r_s = 0.30$ ;  $p = 0.06$ ;  $n = 41$ ) (Fig. 9*a*). Most interesting, cells with strong inhibition exhibited, on average, significantly weaker responses to a-RDS ( $r_s = -0.55$ ;  $p < 0.001$ ) (Fig. 9*b*) but larger tuning indices ( $r_s = 0.81$ ;  $p < 0.001$ ) (Fig. 9*d*). Side peak suppression, in contrast, was not significantly correlated with the amount of inhibition ( $r_s = 0.19$ ;  $p < 0.23$ ) (Fig. 9*c*).

### DISCUSSION

In this study, we first demonstrated a continuous distribution of three fundamental tuning parameters (SSI, EAR, and DTI) in disparity-sensitive neurons. Correlation analyses showed a systematic relationship between all three parameters and response latency. Neurons with higher latencies showed significantly more characteristics of postulated behaviorally relevant disparity detectors. Because neurons exhibiting response characteristics typical for local disparity detectors have the shortest response latencies, the most parsimonious explanation of our data is that the output of local disparity detectors is further processed to generate more unambiguous detectors at later stages of computation. Nonlinear threshold operations together with inhibition are discussed as putative mechanisms to eliminate coding ambiguities.



**Figure 5.** Comparison of quantitative tuning parameters with c-RDS and a-RDS. *a*, Distribution of phase differences and envelope amplitude ratios. A local filtering model would predict phase differences of 0.5 cycles per degree and an EAR of 1. On average, cells in the owl's visual Wulst responded much less to a-RDS, as indicated by a shift of EARs toward values below 1. *b*, Envelope amplitude ratio was negatively correlated with the response latency of the neurons (see Results for statistics). The best fit to the data (logarithmic regression) is shown. The data points of example neurons of Figures 3 and 4 are indicated in the scatterplot. *c*, No correlation was observed between EAR and the phase of the Gabor function fitted to the tuning curves (derived by c-RDS stimulation). EAR and response latency was not correlated with the maximum discharge rate of the neurons (*d*, *e*).

### Scopes and limits of the binocular disparity energy model

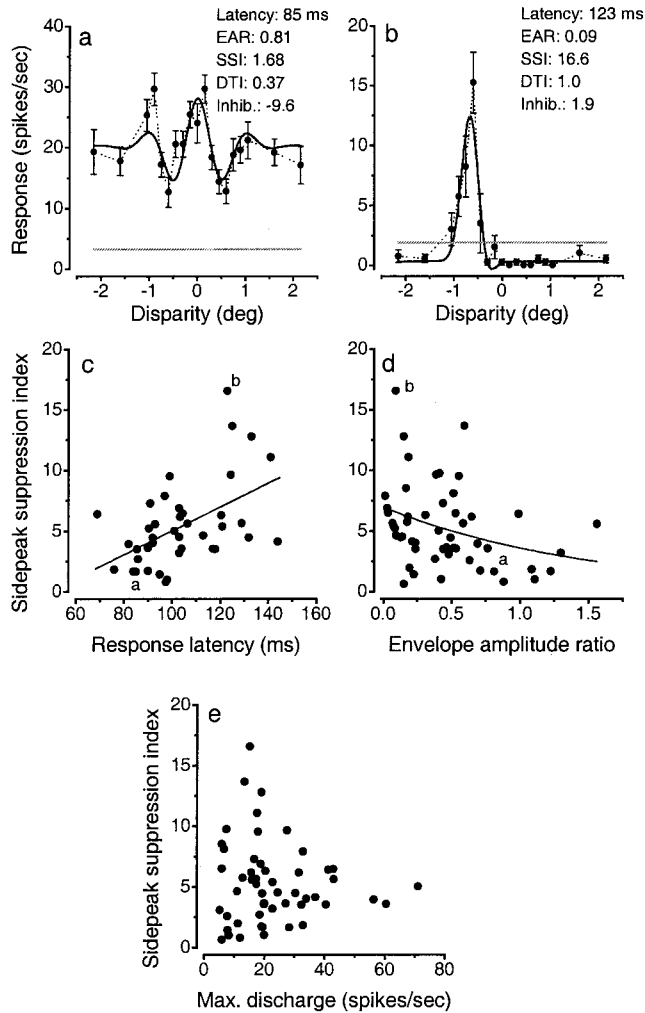
The binocular disparity energy model in its original form assumes that a disparity-sensitive (complex-like) output neuron sums the squared responses of four half-wave rectified, linear simple cells (Ohzawa et al., 1990, 1997; Ohzawa 1998). The squaring nonlinearity of the model can explain translation invariance of real disparity-sensitive neurons. Simple squaring, however, fails to account for the effect of reduced responses to false-matched images, because both negative and positive responses in the tuning profile are conveyed without attenuation. Thus, the energy model predicts a complete inversion of the disparity-tuning profile for opposite-contrast patterns like a-RDS. Many neurons in V1 of cats and primates, as well as visual Wulst neurons in the owl, however, show substantially reduced activity for contrast-inverted stimuli. In the owl, a continuum of EARs ranging from  $\sim 1$  to almost 0 was observed (Fig. 5*a*), with a mean  $\pm$  SD EAR of  $0.46 \pm 0.36$ . Similar yet slightly higher EAR values have been reported for cells in the primary visual cortex of monkey applying RDS (mean EAR of 0.52) (Cumming and Parker, 1997) and cat using bar stimuli (mean EAR of 0.79) (Ohzawa, 1998) (see also Livingstone and Tsao, 1999). As pointed out by Ohzawa (1998), these results are clear deviations from what is expected based on the local filtering model, but "it is also a deviation in the desired direction, in the sense that, ideally, there should be no responses

to reversed-contrast stimuli if these neurons support conscious perception of depth" (Ohzawa, 1998).

In the present study, disparity-sensitive neurons from the behaving owl's visual forebrain were investigated for systematic deviations from the local filtering model. We observed a gradual transition from neurons showing ambiguities typical for energy neuron detectors (EAR close to 1, low SSI) to unequivocally responding neurons that discarded false matches. This transition was highly correlated with an increase in response latency. Latency differences imply hierarchical computation because time is required to transfer information from one stage of processing to the next (Schmolesky et al., 1998). It cannot be excluded that the differences in response latency might reflect thalamic input from functionally different neurons (comparable with M or P neurons in mammalian lateral geniculate nucleus). In such a case, however, we should have observed distinct populations of neurons with different latencies.

### Candidate mechanisms to resolve coding ambiguities: nonlinear threshold operation

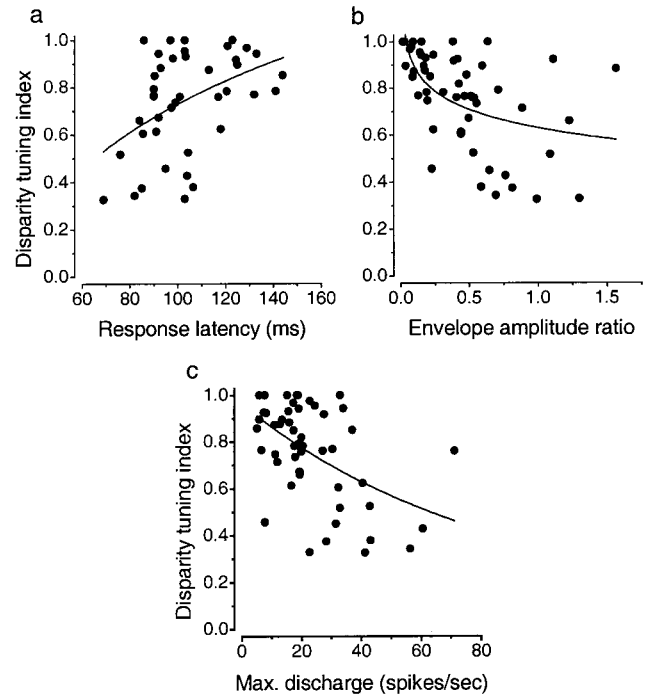
Coding ambiguities arise in several sensory systems and are mainly caused by the narrow filter characteristics of peripheral sensory neurons. In stereovision, Fleet et al. (1996) argued that side peaks could be eliminated if the output of disparity detectors that show the same preferred disparity but different spatial fre-



**Figure 6.** Emergence of side peak suppression. *a, b*, Disparity-tuning curves of two single neurons during stimulation with c-RDS. Tuning curves were fitted with a Gabor function (*solid line*). *a*, Neuron showing extensive modulation and large side peaks. *b*, Response profile of a cell exhibiting one single response peak without any sideband modulation. *c*, SSI and response latency for all tested neurons were significantly correlated. The *solid line* shows the best, linear fit to the data. *d*, Significant correlation between SSI and EAR (*line* represents best, exponential regression). *e*, SSI was not correlated with the maximum spike rate of the cells.

quencies and/or stimulus orientations were linearly pooled. Evidence for spatial frequency integration has been provided for neurons in the anesthetized owl (Wagner and Frost, 1993, 1994). Although across-channel integration would be very effective in reducing side peaks, such pooling alone is insufficient to generate global detectors because it cannot explain suppression of responses to opposite-contrast stimuli. Additional mechanisms must be postulated to explain the elimination of responses to false-matched images.

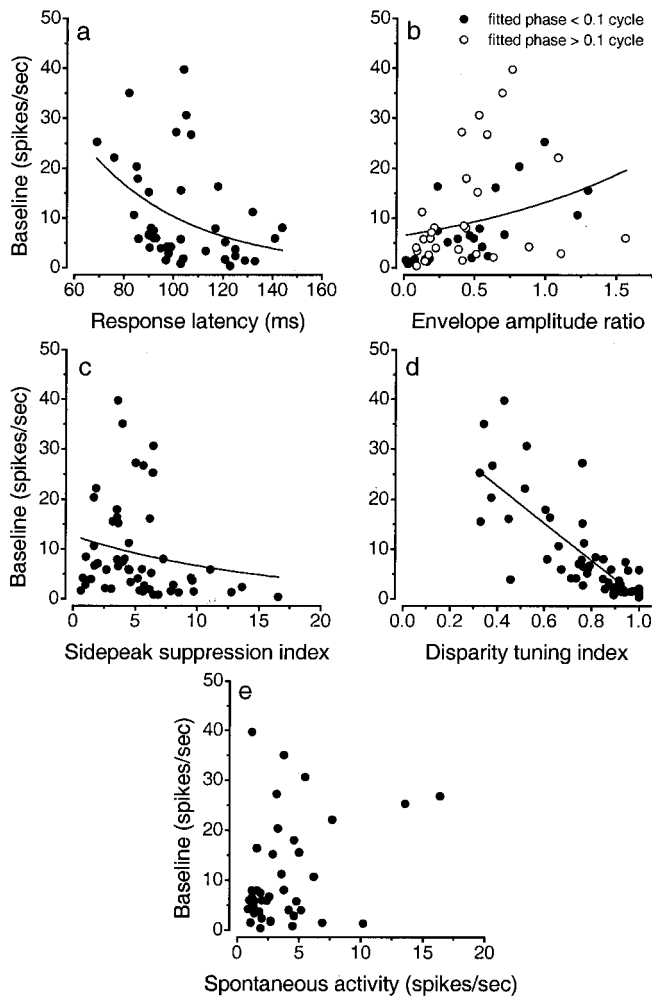
A simple yet very effective mechanism to eliminate responses to false matches is an implementation of higher discharge thresholds for higher-order neurons that get input from local detectors. This would enable the visual system to “clip” side peaks as well as response dips caused by profile inversion in opposite-contrast stimuli. Such a mechanism could also decrease the response offset of the tuning curves. Threshold mechanisms have been shown to play a significant role in shaping the responses of simple cells in



**Figure 7.** Correlation of disparity-tuning index with response latency (*a*) and envelope amplitude ratio (*b*). *Lines* show best fit to the data (logarithmic regression for both *a* and *b*). *c*, DTI was also significantly correlated with the maximum discharge rate of the neurons.

the visual system (Ferster and Miller, 2000). Recent intracellular recordings showed that orientation tuning and direction selectivity of cells measured from their action potentials was considerably sharper compared with orientation tuning and direction selectivity measured directly from the membrane potentials, a phenomenon termed “iceberg effect” (Carandini and Ferster, 2000). In the owl’s Wulst, the significant decline of baseline activity (i.e., tuning-curve offset) of cells with response latency in parallel to the decrease of ambiguities (as defined by EAR, SSI, and DTI) suggests threshold operations. The most unequivocally responding cells had, on average, the longest latencies and very low discharge rates for nonpreferred disparities. Evidently, the disparity-response profile of such low-firing neurons cannot invert, because activity cannot become negative.

Based on recent results obtained with a hierarchical feedforward network (Lippert et al., 2000), we suggest that nonlinear threshold operations during stereo information processing might, at least in part, account for the generation of global disparity detector characteristics in higher-order neurons. Lippert et al. (2000) designed a three-layered network (input, hidden, and output units) that consisted of physiologically motivated monocular Gabor input filters and created output responses mirroring disparity-selective neurons. In contrast to the responses of most V1 neurons (Cumming and Parker, 1997, 2000; Ohzawa et al., 1997; Livingstone and Tsao, 1999) and several visual Wulst cells (current study), however, output neurons of the network were trained to very low baseline activity and, as a result, did not respond to a-RDS (Lippert et al., 2000, their Fig. 6). The authors attributed this effect to the nonlinear threshold functions implemented in their model, a major difference compared with current local filtering models (Qian, 1994; Ohzawa, 1998). Even more interesting, although output units suppressed responses to false-

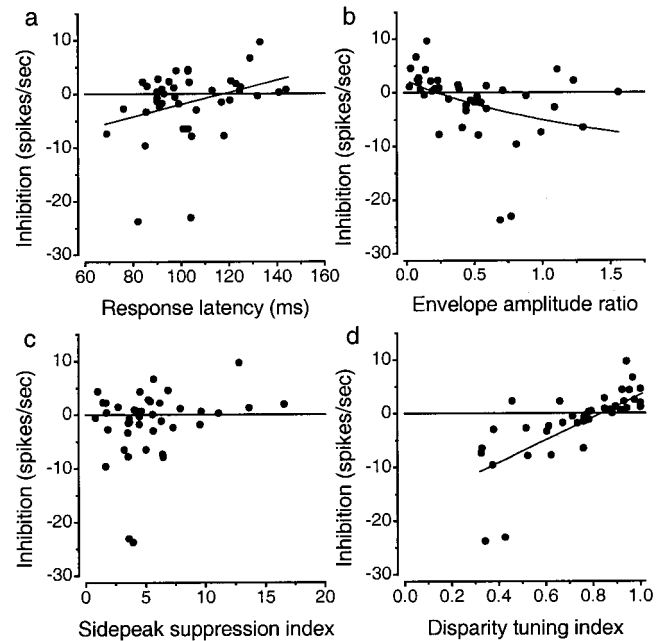


**Figure 8.** Relationship between baseline activity and other physiological parameters. Baseline activity was significantly correlated with response latency (*a*), EAR (*b*), SSI (*c*), and DTI (*d*). Lines are the best fits to the data (all regressions were exponential, except for a linear regression in *d*). *e*, Baseline activity and spontaneous activity were not correlated.

matched images completely, preceding hidden units still showed substantial responses to disparity in a-RDS by profile inversion, as well as extensive modulation of the tuning curve and a corresponding higher baseline activity (J. Lippert, personal communication). This shows that hierarchical processing of disparity information applying nonlinear threshold function can contribute to the elimination of the major response ambiguities inherent to local detectors along processing stages.

### Candidate mechanisms to resolve coding ambiguities: inhibition

A threshold operation, however, is very likely not the only mechanism leading toward higher-order detectors. In particular, the reduction of responses to false-matched images for neurons with tuning curve phases of  $90^\circ$  (odd-symmetric) to  $180^\circ$  (even-symmetric “tuned inhibitory” neurons) cannot be explained by simple threshold mechanisms. Furthermore, simple binocular summation and threshold operation are not able to explain phase differences other than 0.5 cycles between c-RDS and a-RDS tuning profiles. Although the mean phase difference was close to 0.5 cycles, Figure 5*a* illustrates that the deviations were considerable. A similar illustration has been reported previously for



**Figure 9.** Correlation of inhibition with physiological parameters. *a*, Inhibition tended to increase with response latency (positive values indicate inhibition). Cells with low EAR (*b*) and high DTI (*d*) showed, on average, significantly more inhibition. No correlation was found between inhibition and SSI (*c*).

monkey V1 neurons (Cumming and Parker, 1997). Part of this effect, however, might be attributed to the fact that phase determination becomes unreliable for profiles that are more or less flat during a-RDS stimulation.

Even for orientation tuning of simple cells, a threshold is not sufficient to explain all observed effects (Sompolinsky and Shapley, 1997). Recent models that are able to explain contrast invariance in simple cells incorporate stimulus-induced synaptic inhibition in addition to pure feedforward mechanisms (Crook et al., 1998; Ferster and Miller, 2000). Our results from the owl’s visual forebrain suggest that inhibitory influences also contribute to generate more selective and unambiguous disparity detection. Disparity-sensitive neurons with the longest response latencies tended to show more pronounced inhibition, and cells that suppressed responses to a-RDS showed significantly more inhibition compared with neurons that signaled disparity in opposite-contrast patterns. Side peak suppression, on the other hand, seemed not to be influenced by inhibition.

The fact that approximately half of the disparity-sensitive neurons exhibited discharge rates less than spontaneous activity suggests that disparity detection cannot be explained by mere binocular summation. On average, the spontaneous activity was  $3.8 \pm 3.3$  spikes per second. None of the derived measures (response latency, EAR, SSI, DTI, inhibition, or baseline) was correlated with spontaneous activity (all  $p > 0.10$ ). In contrast to dynamic RDS, which might elicit primarily onset (phasic) responses throughout stimulus presentation, static RDS used in the current study evoked predominantly sustained (tonic) responses several milliseconds after stimulus onset; this might have favored the occurrence of suppression or inhibition (because inhibition needs some time to become active).

Based on our results, we suggest a hierarchical framework that can primarily explain the physiological data: local detectors are implemented according to a local filtering model. The thresh-



olded output of several local disparity-sensitive neurons (that may exhibit different selectivity to spatial frequency and/or orientation) converges successively onto higher-order neurons. Inhibitory influences additionally contribute to suppression of ambiguous responses, thus gradually leading toward disparity detectors that may ultimately represent a direct correlate of depth perception. The broad distribution of response latencies and different degrees of coding ambiguities in disparity-sensitive neurons argues against discrete classes of detectors (local versus global) that have been suggested based on recent psychophysical investigations in humans (Neri et al., 1999).

## REFERENCES

- Barlow HB, Blakemore C, Pettigrew JD (1967) The neural mechanisms of binocular depth discrimination. *J Physiol (Lond)* 193:327–342.
- Carandini M, Ferster D (2000) Membrane potential and firing rate in cat primary visual cortex. *J Neurosci* 20:470–484.
- Cogan AI, Lomakin AJ, Rossi A (1993) Depth in anticorrelated stereograms. *Vision Res* 33:1959–1975.
- Crook JM, Kisvárdy ZF, Eysel UT (1998) Evidence for a contribution of lateral inhibition to orientation tuning and direction selectivity in cat visual cortex: reversible inactivation of functionally characterized sites combined with neuroanatomical tracing techniques. *Eur J Neurosci* 10:2056–2075.
- Cumming BG, Parker AJ (1997) Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389:280–283.
- Cumming BG, Parker AJ (1999) Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *J Neurosci* 19:5602–5618.
- Cumming BG, Parker AJ (2000) Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. *J Neurosci* 20:4758–4767.
- Ferster D, Miller KD (2000) Neural mechanisms of orientation selectivity in the visual cortex. *Annu Rev Neurosci* 23:441–471.
- Field DJ, Tolhurst DJ (1986) The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc R Soc Lond B Biol Sci* 228:379–400.
- Fleet DJ, Wagner H, Heeger DJ (1996) Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Res* 36:1839–1857.
- Gabor D (1946) Theory of communication. *J Inst Elec Eng* 93:429–457.
- Hanes DP, Thompson KG, Schall JD (1995) Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. *Exp Brain Res* 103:85–96.
- Jones JP, Palmer LA (1987) An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol* 58:1233–1258.
- Julesz B (1960) Binocular depth perception of computer-generated patterns. *Bell Sys Tech J* 39:1125–1162.
- Legéndy CR, Salzman M (1985) Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. *J Neurophysiol* 53:926–939.
- Lippert J, Fleet DJ, Wagner H (2000) Disparity tuning as simulated by a neural net. *Biol Cybern* 83:61–72.
- Livingstone MS, Tsao DY (1999) Receptive fields of disparity-selective neurons in macaque striate cortex. *Nat Neurosci* 2:825–832.
- Marceljà S (1980) Mathematical description of the responses of simple cortical cells. *J Opt Soc Am* 70:1297–1300.
- Marr D, Poggio T (1979) A computational theory of human stereo vision. *Proc R Soc Lond B Biol Sci* 204:301–328.
- Neri P, Parker AJ, Blakemore C (1999) Probing the human stereoscopic system with reverse correlation. *Nature* 401:695–698.
- Nieder A (2000) Miniature stereo radio transmitter for simultaneous recording of multiple single-neuron signals from behaving owls. *J Neurosci Methods* 101:157–164.
- Nieder A, Wagner H (1999) Perception and neuronal coding of subjective contours in the owl. *Nat Neurosci* 2:660–663.
- Nieder A, Wagner H (2000) Horizontal-disparity tuning of neurons in the visual forebrain of the behaving barn owl. *J Neurophysiol* 83:2967–2979.
- Ohzawa I (1998) Mechanisms of stereoscopic vision: the disparity energy model. *Curr Opin Neurobiol* 8:509–515.
- Ohzawa I, DeAngelis GC, Freeman RD (1990) Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* 249:1037–1040.
- Ohzawa I, DeAngelis GC, Freeman RD (1997) Encoding of binocular disparity by complex cells in the cat's visual cortex. *J Neurophysiol* 77:2879–2909.
- Pettigrew JD (1979) Binocular visual processing in the owl's telencephalon. *Proc R Soc Lond B Biol Sci* 204:435–454.
- Pettigrew JD (1986) The evolution of binocular vision. In: *Visual neuroscience* (Pettigrew JD, Sanderson KJ, Levick WR, eds), pp 208–222. Cambridge, UK: Cambridge UP.
- Pettigrew JD, Konishi M (1976) Neurons selective for orientation and binocular disparity in the visual Wulst of the barn owl (*Tyto alba*). *Science* 193:675–678.
- Poggio GF (1995) Mechanisms of stereopsis in monkey visual cortex. *Cereb Cortex* 3:193–204.
- Poggio GF, Poggio T (1984) The analysis of stereopsis. *Annu Rev Neurosci* 7:379–412.
- Poggio GF, Motter BC, Squatrito S, Trotter Y (1985) Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Res* 25:397–406.
- Qian N (1994) Computing stereo disparity and motion with known binocular cell properties. *Neural Comp* 6:390–404.
- Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG (1998) Signal timing across the macaque visual system. *J Neurophysiol* 79:3272–3278.
- Sompolinsky H, Shapley R (1997) New perspectives on the mechanisms for orientation selectivity. *Curr Opin Neurobiol* 7:514–522.
- Steinbach MJ, Money KE (1973) Eye movements of the owl. *Vision Res* 13:889–891.
- Thompson KG, Hanes DP, Bichot NP, Schall JD (1996) Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J Neurophysiol* 76:4040–4055.
- Tyler CW (1994) Cyclopean riches: cooperativity, neurotrophs, hysteresis, stereoattention, hyperglobality, and hypercyclopean processes in random-dot stereopsis. In: *Early vision and beyond* (Papathomas TV, Chubb C, Gorea A, Kowler E, eds), pp 5–15. Cambridge, MA:MIT.
- van der Willigen RF, Frost BJ, Wagner H (1998) Stereoscopic depth perception in the owl. *NeuroReport* 9:1233–1237.
- Van Essen DC, DeYoe EA (1995) Concurrent processing in the primate visual cortex. In: *The cognitive neurosciences* (Gazzaniga MS, ed), pp 383–400. Cambridge, MA: MIT.
- Wagner H, Frost B (1993) Disparity sensitive cells in the owl have a characteristic disparity. *Nature* 364:796–798.
- Wagner H, Frost B (1994) Binocular responses of neurons in the barn owl's visual Wulst. *J Comp Physiol [A]* 174:661–670.