

# Stereoscopic depth processing in the visual cortex: a coarse-to-fine mechanism

Michael D. Menz<sup>1</sup> and Ralph D. Freeman

Group in Vision Science, School of Optometry, and the Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720-2020, USA

<sup>1</sup> Present address: Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, California 94115, USA

Published online 9 December 2002; doi:10.1038/nn986

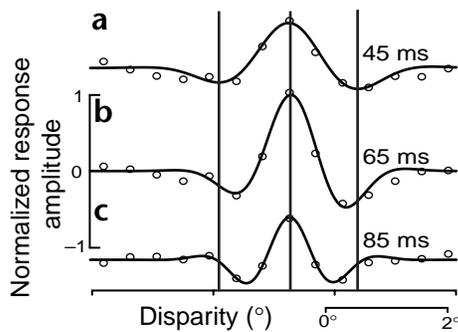
**For binocular animals viewing a three-dimensional scene, the left and right eyes receive slightly different information, and the brain uses this 'binocular disparity' to interpret stereoscopic depth. An important theoretical conjecture in this mechanism is that coarse processing precedes and constrains finely detailed processing. We present three types of neurophysiological data from the cat's visual cortex that are consistent with a temporal coarse-to-fine tuning of disparity information. First, the disparity tuning of cortical cells generally sharpened during the time course of response. Second, cells responsive to large and small spatial scale had relatively shorter and longer temporal latencies, respectively. Third, cross-correlation analysis between simultaneously recorded pairs of cortical cells showed that connections between disparity-tuned neurons were generally stronger for coarse-to-fine processing than for fine-to-coarse processing. These results are consistent with theoretical and behavioral studies and suggest that rapid, coarse percepts are refined over time in stereoscopic depth perception.**

The resolution of stereoscopic depth discrimination is extremely fine<sup>1</sup>. To achieve this, the binocular correspondence problem—which results from the slightly different inputs to each eye while viewing a scene—must first be solved. That is, the matching features in both eyes must be combined to reconstruct an accurate three-dimensional view of the scene. Spatial scale must be taken into account, as neurons respond to a limited range of disparities and spatial frequencies. Neurons that are tuned to low spatial frequencies code a relatively wide range of disparities, but with poor resolution. Conversely, cells tuned to high spatial frequencies have fine resolution, but are limited to processing a narrow range of disparities<sup>2–4</sup>. A leading theory concerning stereoscopic processing is that coarse disparity information constrains that of fine detail<sup>5</sup>. Given that coarse spatial scale covers the entire range of disparities to which the system responds, it is advantageous to first produce a global solution to the correspondence problem. This initial estimate of disparities across the visual field can then be refined by neurons that have better resolution over a narrower range. Different underlying mechanisms have been suggested for this coarse-to-fine process<sup>5–15</sup>. Although there is also an apparent role for fine-to-coarse processing<sup>13</sup>, it is clear that a coarse spatial scale constrains the correspondence solution at finer values<sup>6,8,13,14</sup>. An implicit extension of this idea is that a coarse-to-fine sequence applies to all stereoscopic neural processing, from establishing binocular correspondence to perceiving differences in depth.

Here, we report three types of neurophysiological data that are consistent with a stereoscopic mechanism in which coarse-to-fine processing is a major component. Some of our data also indicate

that there is a fine-to-coarse process involved. For the first two types of data, we analyzed response characteristics of complex cells in the cat's striate cortex. We limited the analysis to complex cells because they are considered primary disparity detectors<sup>16,17</sup>. For the third category of data, we assessed cross-correlation results from both simple and complex cells to maximize sample size. For the first analysis, we examined temporal characteristics of disparity-tuning functions for each cell to determine if disparity frequency (resolution) and range (size) change over time. Disparity frequency refers to the modulation of each cell's response for different binocular disparities. Disparity range specifies the extent of disparities that produces a response from the neuron. Together, frequency and range characterize the degree of fine or coarse tuning of a cell's disparity response. High frequencies and small range correspond to relatively fine tuning. For the second type of evidence, we determined latencies of the strongest responses and compared them with preferred disparity frequencies for each cell. The third type of evidence is derived from simultaneous recording of responses of pairs of disparity-tuned cortical cells. We used cross-correlation analysis to examine timing patterns between cells and relative strengths of interaction.

Our findings are as follows. First, disparity tuning of most cortical cells sharpened during responses. Second, population analysis showed that cells with low disparity frequency preferences tended to have relatively short response latencies. Third, cross-correlation analysis showed that more cell pairs showed fine-to-coarse cortical connections than coarse-to-fine, but the coarse-to-fine connections were generally stronger. Taken together, these results are consistent with a stereoscopic coarse-to-fine process that has a fine-to-coarse component.

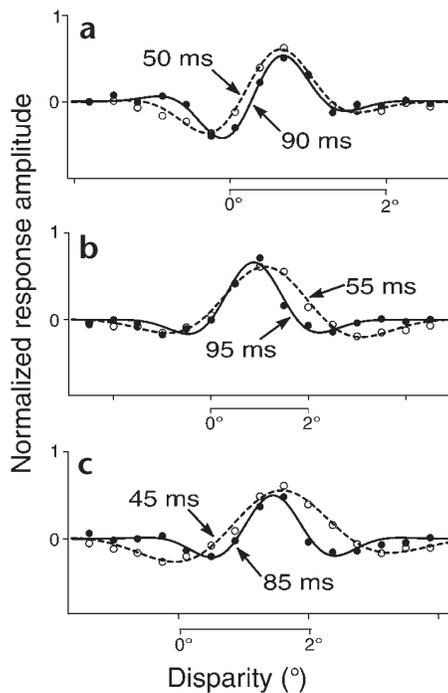


**Fig. 1.** An example of the temporal coarse-to-fine disparity tuning of complex cells. Disparity tuning data (open circles) are shown on a normalized amplitude scale where the disparity tuning at optimal time delay (b) has a maximum amplitude of 1.0. Relative disparity values are given on the abscissa. Responses curves 20 ms before (a) and 20 ms after (c) the optimal response (b) are shown. The data are fit with a Gabor function (solid lines), and the parameters of interest are the disparity frequency (the frequency of response modulation along the disparity axis) and disparity range or size (the range of disparities to which the cell responds). (a) The local maxima and minima at 20 ms before optimal delay are indicated by three vertical lines (two troughs, one peak). Note that the optimal disparity (center vertical line) did not change with correlation delay. However, the adjacent minima moved closer to the optimal disparity with time. This cell shows a 25% decrease in disparity range and a 71% increase in disparity frequency when comparing responses of (c) with those of (a).

**RESULTS**

We calculated the relative magnitude and shape of disparity tuning curves for time slices before, during and after optimal response (data for one cell shown in Fig. 1). The progressive changes in the disparity tuning curves from 20 ms before to 20 ms after the optimal time delay show that the disparity tuning curves sharpened over this time. There was a 71% increase in disparity frequency (resolution) and a 25% decrease in disparity range (size). The optimal disparity (the disparity value where the largest peak occurred) remained at a constant value for all three time slices. By integrating information across different disparity scales while maintaining the same optimal disparity, neurons become more selective for one disparity<sup>7</sup>. It is a coarse-to-fine mechanism because of the temporal order of processing.

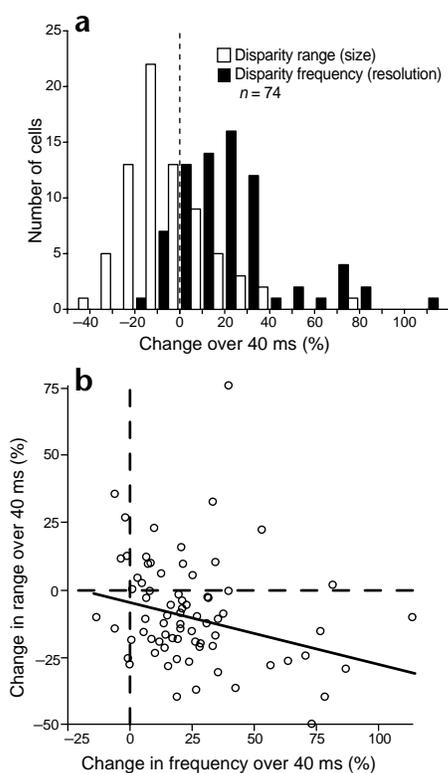
We examined the disparity tuning dynamics for several neurons and confirmed that tuning became sharper with increasing correlation delay (data from three neurons shown in Fig. 2). Responses were normalized to facilitate comparisons of waveform shape and responsivity relative to optimal time delay. Two response curves are shown for each cell; they represent the responses 20 ms before (dashed lines) and 20 ms after (solid lines) the optimal ones. These values were chosen because the temporal resolution of the stimulus is 40 ms, and it is desirable to use values that are symmetric around optimal time to compare time slices with nearly equal spike rates. In this case, differences in waveform shape cannot be attributed to variations in spike rates. Time slices of an additional 40 ms before and after these values did not produce a sufficient number of spikes to obtain a reliable disparity tuning curve. For each cell (Fig. 2a–c), the temporal slices before and after optimal have nearly identical peaks. Optimal disparity remained constant for different correlation delays. In each case, however, tuning was narrower and disparity range was smaller for the time slice following the optimal value. Disparity frequency (resolution) increased and range (size) decreased



**Fig. 2.** Three examples of complex cells show that disparity tuning becomes sharper with increasing correlation delay. (a–c) Disparity tuning data (open circles, 20 ms before optimal time delay; filled circles, 20 ms after optimal) are shown on a normalized amplitude scale where the disparity tuning at optimal delay has a maximum amplitude of 1.0. The full disparity range tested is not shown, so as to emphasize the range of disparities to which the cells respond. As the direction of eye gaze is unknown, the disparity units are relative to an arbitrary zero (the closest disparity tested). The data are fit with a Gabor function (dashed lines, 20 ms before optimal delay; solid lines, 20 ms after optimal), and the parameters of interest are the disparity frequency (frequency of response modulation along the disparity axis) and disparity range (range of disparities within which the cell responded). (a) A typical example, in which the optimal disparity (position of positive peak) did not change, but the disparity frequency was slightly greater (27%) and disparity range slightly smaller (10%) at the later correlation delay. (b) A second cell that shows a quantitatively larger effect: a 43% increase in frequency and a 36% decrease in range. (c) A third cell with a still greater change: 79% increase in frequency and a 40% decrease in range.

with increasing temporal correlation delay. The examples shown (Fig. 2a–c) represent a range of changes that we have observed from relatively small (a) to large (c). One functional implication of these changes is that the subregions, characteristic of complex cells, become smaller and closer together with time. This is shown clearly in (c). These characteristics—constant optimal disparity, increasing frequency and decreasing range—indicate a coarse-to-fine disparity mechanism.

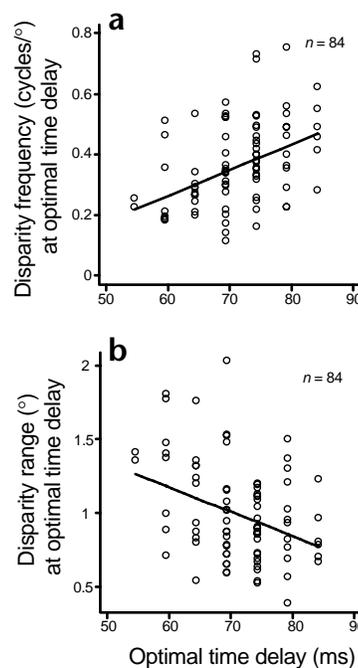
A summary of the dynamics of disparity tuning is shown for a population of 74 complex cells (Fig. 3). We plotted the change in disparity frequency (resolution) and range (size) over a 40 ms time scale (Fig. 3a). The two distributions were weighted on opposite sides of zero: the mean increase in disparity frequency was 25%, and the mean decrease in disparity range was 8.4%. Both values were significantly different from zero ( $P < 0.001$  and  $P < 0.001$ , standard normal, central limit theorem). An alternative method of analysis is to examine whether the changes in a parameter (frequency or range) are significant on a cell-by-cell



**Fig. 3.** A population summary of changes in disparity frequency and disparity range. **(a)** Distributions of disparity frequency (filled bars) and disparity size (unfilled bars) dynamics are presented for 74 complex cells. The position of the dashed vertical line (at zero) shows that the means of these distributions are shifted away from zero. Change in a parameter is considered as a percent difference between 20 ms before optimal time delay and 20 ms after. A positive value indicates an increase with greater time delay, and a negative number signifies a decrease. The average increase in frequency is 25%, and the average decrease in range is 8.4%. These values are significantly different from zero ( $P < 0.001$ ,  $P < 0.001$ ). **(b)** A relationship is shown between the changes in range and in frequency over a 40 ms time period. Dashed vertical and horizontal lines indicate no change. The robust regression (solid line) has a slope of  $-0.22$  and is significantly different from zero ( $P = 0.0097$ ). The Pearson correlation coefficient ( $-0.24$ ) and regression indicate a weak correlation such that cells with a large increase in frequency tend to have a large decrease in size.

basis, given the confidence intervals for these fitted parameters from the Levenberg-Marquardt algorithm<sup>18</sup>. Of 74 cells, 59 showed a significant increase in disparity frequency, 15 were not significantly different, and none showed a significant decrease in frequency. For disparity range, 41 cells showed no significant difference, 28 showed a significant decrease, and 5 had a significant increase. Whereas the frequency parameter was well constrained by the fit, the range parameter was relatively less constrained. The change in disparity range was not as strong an effect as the change in disparity frequency. Finally, we measured the change in optimal disparity over the same 40-ms time span during which the disparity frequency and disparity range measurements were made. Out of 74 cells, 64 had a change in optimal disparity of  $<0.1^\circ$ , seven showed a change between  $0.1^\circ$  and  $0.2^\circ$ , and one cell had a change of  $0.2$ – $0.3^\circ$ . The change in optimal disparity was minimal.

Taken together, these data are consistent with a neural mechanism in which cells make an initial, relatively coarse disparity

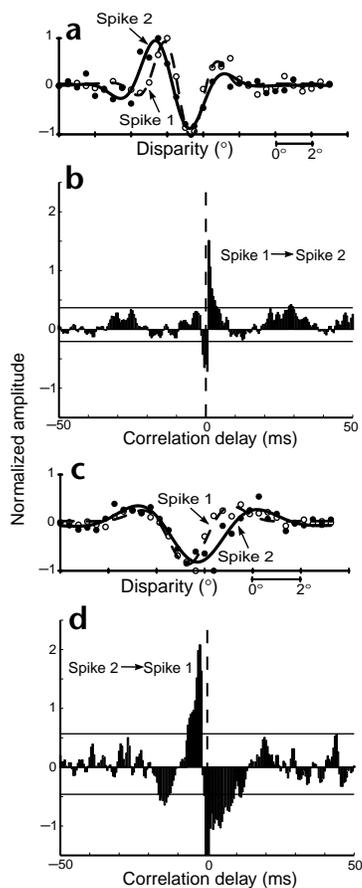


**Fig. 4.** Neurons with higher disparity frequency and smaller disparity range tend to have greater optimal correlation delays. **(a)** Relationship between disparity frequency tuning and optimal time delay is given for 84 complex cells. The optimal disparity frequency measured at the optimal time delay is plotted against the optimal time delay for these cells. Higher-frequency cells were relatively more time delayed compared to lower-frequency cells. The robust regression (solid line) has a slope ( $0.0085$  cycle/ $^\circ$ /ms) that is significantly different from zero ( $P = 0.0001$ ). **(b)** The disparity range parameter (the standard deviation of the Gaussian envelope from the Gabor fit) measured at the optimal time delay is plotted against optimal time delay. These are the same cells as shown in **(a)**. Cells with relatively smaller ranges tended to have greater delays. The robust regression (solid line) has a slope ( $-0.017$   $^\circ$ /ms), which is significantly different from zero ( $P = 0.001$ ).

judgment, which is refined during the response as the range is narrowed. If a single mechanism is responsible for both the increase in frequency and decrease in range, then these two changes should be correlated. Cells with larger increases in frequency tended to have larger decreases in range (Fig. 3b). However, the correlation was relatively weak (correlation coefficient,  $-0.24$ ). The weakness of the correlation suggests the presence of more than one mechanism.

Coarse-to-fine dynamic disparity tuning could be the result of pooling from cells in which relatively lower-frequency cells have shorter latencies. To test this hypothesis, we examined the disparity frequency and disparity range at optimal time delay (Fig. 4a). Cells with higher disparity frequency were clearly more time-delayed. The slope (solid line;  $0.0085$  cycles/ $^\circ$ /ms) was significantly different from zero ( $P = 0.0001$ ). A similar result was found for disparity range (Fig. 4b). This slope (solid line;  $-0.017$   $^\circ$ /ms) was also significantly different from zero ( $P = 0.001$ ). Neurons with smaller ranges tended to have greater time delays. These results are consistent with a process in which disparity tuning becomes finer over time.

The third type of analysis we report here is that of simultaneously recorded pairs of disparity-tuned neurons. For this analysis, we included both simple and complex cells to increase our



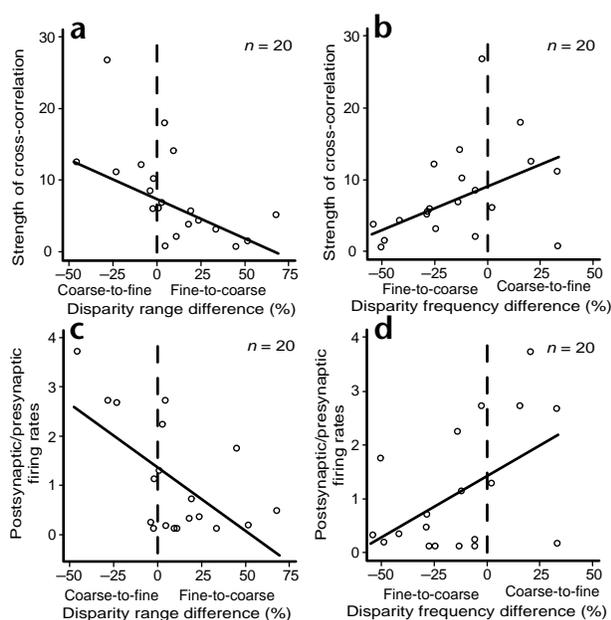
**Fig. 5.** Examples of fine-to-coarse and coarse-to-fine connections between disparity-tuned complex cells. **(a)** Spike 1 disparity tuning data (open circles), Gabor fit (dashed line); Spike 2 disparity tuning data (closed circles), Gabor fit (solid line). **(b)** Shuffle-subtracted, cross-correlogram of the neurons in **(a)**, normalized by the number of pre-synaptic spikes ('effectiveness')<sup>30</sup>. The two horizontal lines are the mean  $\pm$  3 s.d. of the response in the range 50–100 ms. Bins that exceeded this were considered significant. Spike 1, which has a disparity frequency of 0.31 cycle/° and size of 1.48° is clearly pre-synaptic to spike 2, which has a frequency of 0.23 cycle/° and a range of 1.8°. This is a fine-to-coarse connection and the effective connectivity (area of the peak) is relatively weak. **(c)** The disparity tuning function of a different pair of cells is shown here. **(d)** The cross-correlogram of the pair in **(c)** shows a coarse-to-fine connection. Pre-synaptic spike 2 has a frequency of 0.17 cycle/° and a range of 2.1°, whereas the post-synaptic spike 1 Gabor function fit yields a frequency of 0.21 cycle/° and a range of 1.3°. This coarse-to-fine connection displays a much stronger effective connectivity compared to the fine-to-coarse connection in **(b)** (12.5 versus 5.6).

sample size. We analyzed data for 20 pairs of disparity-tuned simple and complex cells: 17 pairs were recorded from the same electrode, and 3 pairs were from different electrodes. Three combinations of monosynaptic and polysynaptic connections (see Methods for classification criteria) were included: simple-to-simple, complex-to-simple and complex-to-complex. The simple-to-complex connection was not used because it is a special case in which we believe that simple cells are the subunits of complex cells<sup>3,4</sup>; it should therefore be considered separately. Representative cross-correlation data are shown for two pairs of complex cells (Fig. 5). For the first pair, the shuffle-subtracted cross-correlogram indicates that cell 1 drove cell 2, and the dis-

parity frequency and range data indicate that this cell combination represents a fine-to-coarse connection (Fig. 5a and b). The second cell pair showed a stronger cross-correlation peak. In this case, cell 2 drove cell 1 (both complex cells). Disparity frequency and range data indicate that this cell combination represents a coarse-to-fine connection (Fig. 5c and d).

For the pairs of neurons from which we derived cross-correlogram data, the type of connectivity between the cells was assessed according to either disparity range (size) or disparity frequency (resolution). If the disparity range was larger for the pre-synaptic cell compared to the post-synaptic cell, we considered the connection to be coarse-to-fine. Conversely, a smaller pre-synaptic cell range signified a fine-to-coarse connection. To derive a quantitative measurement, we considered the difference in range expressed as a percentage of the average range (Methods). Similarly, if the disparity frequency was lower for the pre-synaptic cell, then the connection was considered coarse-to-fine. Cell connectivity was asymmetrical (Fig. 6). Whether evaluated by disparity range or disparity frequency, there were more fine-to-coarse than coarse-to-fine connections. This difference was statistically significant only for frequency ( $P = 0.021$ , binomial sign test; Fig. 6b and d). This type of asymmetry indicates that the quantities of connections are not equal. Another type of asymmetry involves the strength of connections and firing rates. In Fig. 6a and b, we show the relationship between the strength (that is, area) of the cross-correlation peak and the type of connectivity. Coarse-to-fine connections, whether defined by range or frequency, tend to be stronger (robust regressions,  $P = 0.001$  and  $P = 0.005$ , respectively). The firing rates of our pairs of cells were generally not equal. Relative firing rates (firing rate of the post-synaptic cell divided by that of the pre-synaptic cell) instead of correlation strength are also shown (Fig. 6c and d). Firing rate is a much simpler, more intuitive measure of neural response than shuffle-subtracted, normalized, cross-correlogram peak area. Coarse-to-fine connections, whether defined by range or frequency, tend to have higher post-synaptic firing rates compared to those of the pre-synaptic cells (robust regressions,  $P = 0.0098$  and  $P = 0.039$ , respectively). If the fine-to-coarse connections have weak synaptic strengths, it follows that their post-synaptic firing rates are low compared to those at pre-synaptic sites. Similarly, it is logically consistent that strong coarse-to-fine connections result in more robust post-synaptic firing rates compared to those at pre-synaptic levels. Two clear points emerge from these data. First, there were more cell pairs with fine-to-coarse connections. Second, the coarse-to-fine connections (to the left of the dashed lines for range plots Fig. 6a and c and to the right for frequency plots Fig. 6b and d) were generally stronger with post-synaptic cells that had higher firing rates.

Combining information across spatial scale is most useful if the neurons are tuned to the same optimal disparity. Therefore, it is important to note whether disparity-tuned neurons that are connected are tuned to the same optimal disparity, even though their disparity frequencies and disparity ranges may be different. We found that many connected pairs had nearly opposite disparity tuning curves. The 'similarity index' (SI) of tuning curves can be quantified by taking the Pearson correlation coefficient of the pair (where +1 and -1 indicate identical and exactly opposite tuning, respectively). Our 20 pairs of cells were evenly split, such that half had an SI > 0, and the other half had an SI < 0. Connections between neurons are likely to have multiple purposes. The resolution of horizontal disparity across spatial scale is one possibility. It is not clear what functional role is performed by cell pairs with nearly opposite tuning



**Fig. 6.** Asymmetries in coarse-to-fine and fine-to-coarse connections. **(a, b)** Strength of neural cross-correlation (area of the peak) is given between monosynaptic and polysynaptic pairs of disparity-tuned cells (all combinations of simple/complex except simple-to-complex). Coarse-to-fine and fine-to-coarse connections are defined by either the difference in disparity range **(a)** or disparity frequency **(b)** between the pairs, expressed as a percentage of the average value for the pair. The dashed vertical lines indicate that pre and post-synaptic cells are not different with respect to the given parameter (frequency or range). **(a)** This plot defines the relationships of the cells in terms of differences in disparity range. To the left of the vertical dashed line, the pre-synaptic cell has a larger range (coarse-to-fine); to the right, the pre-synaptic cell has a smaller range (fine-to-coarse). The most common type of connection is fine-to-coarse. The robust regression (solid line) is significantly different from zero ( $P = 0.001$ ). The Pearson correlation coefficient is  $-0.62$ . Coarse-to-fine connections tend to be stronger than fine-to-coarse connections. **(b)** This plot defines the relationships of the cells in terms of differences in disparity frequency. To the left of the vertical dashed line, the pre-synaptic cell has a smaller frequency (fine-to-coarse); to the right, the pre-synaptic cell has a larger frequency (coarse-to-fine). There are far more fine-to-coarse connections. The robust regression (solid line) is significantly different from zero ( $P = 0.005$ ). The Pearson correlation coefficient is  $0.40$ . Consistent with the results in **(a)**, coarse-to-fine connections tend to be stronger. **(c, d)** The relationship between relative firing rate (firing rate of the post-synaptic cell divided by that of the pre-synaptic cell) and coarse/fine connectivity, as defined above. The stronger coarse-to-fine connections defined either by range differences **(c)** or by frequency differences **(d)** tend to have a relatively higher postsynaptic firing rates (robust regression,  $P = 0.0098$ ,  $P = 0.039$ , respectively).

( $SI < 0$ ). However, among the ten pairs with an  $SI > 0$ , the difference in optimal disparity was less than 10% for eight of these pairs (difference in optimal disparity expressed as a percentage of the average disparity period of the pair, where disparity period is the inverse of disparity frequency and has units of degrees). In the example shown above (Fig. 5a), the optimal disparities defined by the positive peaks differed by 13% (this example is one of only two with differences greater than 10%), but the inhibitory troughs were almost perfectly aligned. In the 'tuned inhibitory' type of cells (Fig. 5b), the location of the troughs differed by 8%. Thus, for a subclass of connected cells, the dis-

parity tuning peaks are approximately aligned, whereas the other parameters (notably disparity frequency and disparity range) may differ considerably.

## DISCUSSION

The three types of neurophysiological evidence presented here are consistent with coarse-to-fine sequential processing as a major component of the neural mechanism of stereopsis. The improvement in resolution with greater correlation delay is similar to that for other parameters such as orientation<sup>19</sup> and spatial frequency<sup>20,21</sup>. A previous investigation also shows direction tuning changes of medial temporal (MT) neurons during a similar temporal window (60 ms)<sup>22</sup>. There is some behavioral evidence that supports coarse-to-fine temporal processing of information. For example, low-pass and high-pass spatially filtered versions of images have been studied<sup>23</sup>. The spatially filtered versions were presented in different temporal sequences: low-to-high and high-to-low frequencies. The subject's task was to discriminate a full-bandwidth image from the spatially filtered versions. The low-to-high temporal sequence was more often mistaken for the full bandwidth presentation than the high-to-low sequence. These perceptual and neurophysiological results are consistent with our findings in support of a general, temporal coarse-to-fine processing mechanism in the visual cortex.

A model of neural connectivity that is compatible with all of our results consists of a post-synaptic neuron that pools over many neurons of slightly different spatial scales (some larger and others smaller). All of the neurons are tuned to approximately the same optimal disparity. The pre-synaptic neurons of relatively lower disparity frequency and larger disparity size have shorter latencies that would dominate the early response of the post-synaptic neuron. Pre-synaptic neurons with relatively larger disparity range (size) form stronger connections compared to the more numerous pre-synaptic neurons with smaller disparity range. Considered together, the post-synaptic neuron performs a weighted averaging across spatial scale, with greater and earlier weight given to coarse-to-fine connections.

The concept of weighted averaging across spatial scale with the same optimal disparity is also consistent with theoretical and behavioral studies. For example, it has been proposed that averaging disparity signals across spatial scales helps to determine the correct disparity, reducing the ambiguity inherently present in any one spatial scale<sup>7</sup>. Psychophysical studies that show both fine-to-coarse and coarse-to-fine interactions are also compatible with weighted disparity averaging, which incorporates both types of connections. Perceptual weighted disparity averaging across spatial scales is a well-known phenomenon when the stimulus consists of multiple disparities in the same location<sup>24</sup>.

Our results are consistent with other neurophysiological studies in which the dynamics of neural tuning have been examined<sup>19,20,21</sup>. They are also compatible with psychophysical studies that support a temporal coarse-to-fine processing system, and with theoretical models that advocate the advantages of combining information across spatial scale. Together, these neurophysiological results and the behavioral studies in stereopsis suggest that coarse spatial scales constrain finer scales, but there is also a role for fine-to-coarse processing.

## METHODS

**Isoflurane.** (2–4%) was used to anesthetize animals during surgery. After surgery, anesthesia was switched to thiopental sodium (continuous infusion as required of around 1.4 mg per kg per h). Paralysis was maintained with pancuronium (continuous infusion of 0.2 mg per kg per h). Arti-

ficial respiration was maintained with 30% O<sub>2</sub> and 70% N<sub>2</sub>O at 25 breaths/min, and the stroke volume was adjusted according to the cat's weight. All procedures complied with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of California Animal Care Committee. Visual stimuli were generated by a computer with two high-resolution graphics boards that ran custom software. Images were displayed on a pair of video monitors that the cat views dichoptically by means of beam splitters (mean luminance, 23 cd/m<sup>2</sup>). Action potentials were discriminated by custom-made software and time-stamped with 40 μs resolution.

**Receptive field mapping.** The results from drifting grating runs were used to determine optimal stimulus parameters for receptive field mapping with dichoptic, one-dimensional, binary m-sequence noise. Sixteen adjacent bars were presented to each eye at optimal orientation. The width of the bars was approximately one-fourth the period of the optimal spatial frequency. The length of the bar was equal to sixteen times the width. This square pattern was centered over the receptive field. Each of the sixteen bars was either bright or dark, and the mean luminance of the bars constituted the background. One stimulus pattern lasted for three frames (39 ms).

Each spike train was cross-correlated with the stimulus sequence at each location of binocular combination of stimuli (which varies in depth and position along a fronto-parallel plane) to obtain a two-dimensional binocular interaction map for a particular time delay. Excitatory response to bars of the same contrast or of opposite contrast was represented as a positive or a negative number, respectively. The two-dimensional binocular interaction maps were reduced to one-dimensional disparity tuning data by integration along lines of equal disparity. This was repeated for all time delays of interest (0–200 ms) in increments of 5 ms. Time delay was measured relative to the middle point in the three frame pattern. The optimal time delay was defined as that which produced the greatest root mean squared (RMS) signal strength. The time slices that preceded and followed optimal by 20 ms were chosen for further analysis. These time slices were chosen because the temporal resolution of the stimulus is 40 ms, and symmetrical slices on either side and close to optimal yield high signal-to-noise values and best represent receptive field dynamics. A detailed description of these procedures is provided elsewhere<sup>2–4,25</sup>.

**Curve fitting.** The disparity tuning curves were fitted with a Gabor function:

$$S(d) = \exp(-(d - d_0)^2/2\sigma_s^2) \cos(2\pi f_{ds}(d - d_0) + \phi),$$

where  $d$  is disparity,  $d_0$  is the center position,  $\sigma_s$  is the disparity range (size) parameter,  $f_{ds}$  is the disparity frequency, and  $\phi$  is the phase. If there were insufficient spikes (the Gabor fit was not good), that neuron was not included in the data set. The criteria for a good fit were that (i)  $R^2 > 0.8$ , where  $R^2$  is the coefficient of determination, and (ii) the disparity frequency and disparity range parameters were well-constrained based on the confidence interval provided by the Levenberg-Marquardt algorithm<sup>18</sup>. These criteria are more suitable than the signal/noise ratio or the total quantity of spikes. The percentage changes in disparity resolution (frequency) and range (size) with correlation were calculated as

$$[100 \times (P_{+20ms} - P_{-20ms})/P_{-20ms}]$$

where  $P$  is either a frequency or a size parameter used to measure dynamic alterations in the coarseness of disparity tuning. The maximum amplitude of the signal at optimal time delay was used to normalize all amplitudes from the other time delays (Figs. 1 and 2). Summary data in scatter plots were fit with robust regressions, which minimize the influence of outliers<sup>26</sup>.

**Neural cross-correlation.** Two or more cells were recorded from either the same electrode or adjacent electrodes for which the difference in cortical depth did not exceed 500 μm. Data for this study includes two or more binocular disparity tuned cells with similar orientation and spatial frequency tuning, and adjacent or overlapping receptive field loca-

tions. Cell pairs included in the analysis had clear structure in their neural cross-correlograms and were either mono- or polysynaptic<sup>27</sup>.

Disparity tuning curves were fit with a Gabor function at optimal delay only. Differences in the range or frequency parameters between pre- and post-synaptic cells were normalized by the average value as follows:

$$100 \times (S_{post} - S_{pre}) / ((S_{post} + S_{pre})/2),$$

where  $S$  is either the size or frequency parameter of the pre (pre-synaptic) or post (post-synaptic) neuron. This is a measure of the magnitude of the coarse-to-fine and fine-to-coarse connections.

All raw neural correlograms were shuffle-subtracted to eliminate stimulus-based correlations<sup>28</sup>. Shuffle-subtraction was obtained by first taking the cross-correlogram between pairs of cells from all the repetitions of a stimulus. Cross-correlation between repetitions, which can only be stimulus-based, were then subtracted out. The result was a neurally based cross-correlogram. In this study, we need to distinguish common input connections from those of a mono- or polysynaptic nature. In mono- and polysynaptic connections, there was a relatively narrow peak shifted away from zero so that one cell (e.g. spike 1) can be said to consistently fire before another (e.g. spike 2)<sup>27</sup>. From this, we infer that spike 1 is presynaptic to spike 2. If the peak was broad and straddled zero, then the pair of cells was assumed to be receiving a common input<sup>27</sup>. These categories are somewhat arbitrary, as there is a continuum of interactions. It is also possible for one cell to be pre-synaptic to another while they both receive common input. Three quantitative criteria were used for eliminating a common input type of cross-correlogram from the analysis. First, the correlogram asymmetry index [(AI = (R - L)/(R + L), where R and L are the areas of the bins to the right and to the left of zero (by ± 5 ms, respectively)], is a measure of how much the correlogram peak was shifted from zero<sup>29</sup>. Second, the latency of the cross-correlogram peak is the bin of maximum amplitude. Third, the width of the peak is defined as the width at half-maximum amplitude. Common input connections were defined as having a latency of ≤ 5 ms and an AI of < 0.2. Some common input connections can have narrow widths, so this criterion was not included. As noted above, common input connections were excluded from this study. To be included, cells must be classified as either monosynaptic or polysynaptic. Monosynaptic connections were defined as having a peak latency < 3.5 ms, a width < 5 ms, and an asymmetry index > 0.2. Polysynaptic connections were defined as non-monosynaptic connections having a peak latency < 10 ms and an AI > 0.2. These quantitative criteria are consistent with the known properties and characteristics of monosynaptic and polysynaptic connections<sup>27</sup>. The strength of cross-correlation is the area of the peak between half-heights of maximum amplitude (width) normalized by the number of pre-synaptic spikes ('effectiveness')<sup>30</sup> and multiplied by 100 to give percentage.

## Acknowledgments

This work was supported by research and CORE grants (EY01175 and EY03716) from the National Eye Institute. We gratefully acknowledge the use of data collected by I. Ohzawa and A. Anzai.

## Competing interests statement

The authors declare that they have no competing financial interests.

RECEIVED 1 AUGUST; ACCEPTED 18 NOVEMBER 2002

1. Wheatstone, C. Contributions to the physiology of vision. Part the first: on some remarkable and hitherto unobserved phenomena of binocular vision. *Phil. Trans. R. Soc. Lond.* **128**, 371–394 (1838).
2. Anzai, A., Ohzawa, I. & Freeman, R.D. Neural mechanisms for processing binocular information I. Simple cells. *J. Neurophysiol.* **82**, 891–908 (1999).
3. Anzai, A., Ohzawa, I. & Freeman, R.D. Neural mechanisms for processing binocular information II. Complex cells. *J. Neurophysiol.* **82**, 909–924 (1999).
4. Ohzawa, I., DeAngelis, G.C. & Freeman, R.D. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. Neurophysiol.* **77**, 2879–2909 (1997).
5. Marr, D. & Poggio, T. A., computational theory of human stereo vision. *Proc. R. Soc. Lond. B Biol. Sci.* **204**, 301–328 (1979).

6. Wilson, H.R., Blake, R. & Halpern, D.L. Coarse spatial scales constrain the range of binocular fusion on fine scales. *J. Opt. Soc. Am. A* **8**, 229–236 (1991).
7. Fleet, D.J., Wagner, H. & Heeger, D.J. Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Res.* **36**, 1839–1857 (1996).
8. Rohaly, A.M. & Wilson, H.R. Nature of coarse-to-fine constraints on binocular fusion. *J. Opt. Soc. Am. A* **10**, 2433–2441 (1993).
9. Anderson, C.H. & Van Essen, D.C. Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proc. Natl. Acad. Sci. USA* **84**, 6297–6301 (1987).
10. Nishihara, H.K. Practical real-time imaging stereo matcher. *Opt. Eng.* **23**, 536–545 (1984).
11. Nomura, M. A model for neural representation of binocular disparity in striate cortex: distributed representation and veto mechanism. *Biol. Cybern.* **69**, 165–171 (1993).
12. Quam, L. H. Hierarchical warp stereo. in *Readings in Computer Vision* (eds. Fischler, M.A. & Firschein, O.) 80–86 (Kauffman, Los Altos, California, 1987).
13. Smallman, H. S. Fine-to-coarse scale disambiguation in stereopsis. *Vision Res.* **35**, 1047–1060 (1995).
14. Smallman, H.S. & MacLeod, D.I. Spatial scale interactions in stereo sensitivity and the neural representation of binocular disparity. *Perception* **26**, 977–994 (1997).
15. Sanger, T.D. Stereo disparity computation using gabor filters. *Biol. Cybern.* **59**, 405–418 (1988).
16. Qian, N. & Zhu, Y. Physiological computation of binocular disparity. *Vision Res.* **37**, 1811–1827 (1997).
17. Ohzawa, I., DeAngelis, G.C. & Freeman, R.D. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* **249**, 1037–1041 (1990).
18. Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. *Numerical Recipes in C* (Cambridge Univ. Press, Cambridge, UK, 1992).
19. Ringach, D.L., Hawken, M.J. & Shapley, R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature* **387**, 281–284 (1997).
20. Bredfeldt, C.E. & Ringach, D.L. Dynamics of spatial frequency tuning in macaque V1. *J. Neurosci.* **22**, 1976–1984 (2002).
21. Mazer, J.A., Vinje, W.E., McDermott, J., Schiller, P.H. & Gallant, J.L. Spatial frequency and orientation tuning dynamics in area V1. *Proc. Natl. Acad. Sci. USA* **99**, 1645–1650 (2002).
22. Pack, C.C. & Born, R.T. Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature* **409**, 1040–1042 (2001).
23. Parker, D.M., Lishman, J.R. & Hughes, J. Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception* **26**, 1169–1180 (1997).
24. Rohaly, A.M. & Wilson, H.R. Disparity averaging across spatial scales. *Vision Res.* **34**, 1315–1325 (1994).
25. DeAngelis, G.C., Ohzawa, I. & Freeman, R.D. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiol.* **69**, 1091–1117 (1993).
26. Li, G. Robust regression. in *Exploring Data Tables, Trends and Shapes* (eds. Hoaglin, D.C., Mosteller, F. & Tukey, J.W.) 281–340 (Wiley, New York, 1985).
27. Moore, G.P., Segundo, J.P., Perkel, D.H. & Levitan, H. Statistical signs of synaptic interaction in neurons. *Biophys. J.* **10**, 876–900 (1970).
28. Perkel, D.H., Gerstein, G.L. & Moore, G.P. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J.* **7**, 419–440 (1967).
29. Alonso, J.M. & Martinez, L.M. Functional connectivity between simple cells and complex cells in cat striate cortex. *Nat. Neurosci.* **1**, 395–403 (1998).
30. Aertsen, A.M. & Gerstein, G.L. Evaluation of neuronal connectivity: sensitivity of cross-correlation. *Brain Res.* **340**, 341–354 (1985).